THE EUROPEAN
PHYSICAL JOURNAL C

Evolution and interpolation of double parton distributions using Chebyshev grids

Markus Diehl[1,a] , **Riccardo Nagar**[2], **Peter Plößl**[1], **Frank J. Tackmann**[1]

[1] Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany
[2] Università degli Studi di Milano-Bicocca and INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, 20126 Milan, Italy

Abstract Double parton distributions are the nonperturbative ingredients needed for computing double parton scattering processes in hadron–hadron collisions. They describe a variety of correlations between two partons in a hadron and depend on a large number of variables, including two independent renormalization scales. This makes it challenging to compute their scale evolution with satisfactory numerical accuracy while keeping computational costs at a manageable level. We show that this problem can be solved using interpolation on Chebyshev grids, extending the methods we previously developed for ordinary single-parton distributions. Using an implementation of these methods in the C++ library CHILIPDF, we study for the first time the evolution of double parton distributions beyond leading order in perturbation theory.

Contents

[a] e-mail: markus.diehl@desy.de (corresponding author)

1 Introduction

Double parton scattering (DPS) is a mechanism in hadron–hadron collisions in which two partons from each hadron initiate two separate hard-scattering processes. The Tevatron and the LHC have provided evidence for this mechanism for a wide range of final states with heavy flavors, jets, photons, or electroweak gauge bosons [1–6]. While often suppressed compared with single-parton scattering, DPS can be significant and even dominate in specific final states or kinematic regions. A prominent example is like-sign W production [5,7–13], which (via its leptonic decay) also provides a background to searches for physics beyond the Standard Model [14,15]. Significant progress in the theory description of DPS has been made in the last decade [16–24], and there is a sustained interest in phenomenological aspects (see for instance the recent study in [25] and the references therein). A detailed account of theoretical and experimental aspects of DPS is given in Ref. [26].

Springer

Double parton distributions (DPDs) are the nonperturbative quantities needed to compute DPS cross sections. They quantify a variety of correlations between two partons in the proton [19,20,27] and thus reveal aspects of hadron structure that are not accessible in ordinary parton distributions (PDFs). Our knowledge of DPDs is still quite limited, although there is considerable activity in devising theory guided ansätze [28–31] and in computing DPDs using lattice QCD [32] or quark models [33–40].

Reflecting the multitude of information they contain, DPDs depend on a large number of variables, namely the momentum fractions x_1 and x_2 of the two partons, the transverse distance y between them, and the renormalization scales μ_1 and μ_2 associated with each parton (corresponding to the factorization scales in the two hard-scattering processes that the partons initiate). Their scale dependence is well understood (see Eq. (1) below), but the numerical delivery of evolved DPDs is a computational challenge. We believe that this represents one of the bottlenecks for using realistic forms of DPDs in predictions, either at the analytical level or in the form of Monte Carlo event generators [41–45]. In the present work, we present a method that addresses this challenge, allowing for high numerical accuracy at a moderate computational cost.

We recall that the evolution equations for PDFs can be solved numerically either by discretization in the momentum fraction x or in Mellin space, and that public codes are available for both options, see for instance Refs. [46–50] and [51–53]. For DPDs, even the simplest realistic initial conditions for evolution have a correlated dependence on x_1 and x_2 (see Eqs. (7) and (8) below), which precludes the analytic computation of complex Mellin moments in these two variables. As far as we can see, this makes a discretization in x_1 and x_2 inevitable.

The main reason why the handling of evolved distributions is much more demanding for DPDs than for PDFs is the sheer amount of computer memory needed to store the former. As an example, consider interpolation grids for x_1 and x_2 with 64 points (for PDFs, this is at the lower end of typical grid sizes in the LHAPDF interface [54]), and grids with 48 points for y, μ_1, and μ_2. This corresponds to $64^2 \times 48^3 \approx 4.5 \times 10^8$ real numbers and thus (with 8 bytes for a double precision floating-point number) to 3.4 GiB per parton flavor combination. A full unpolarized DPD with 5 active quark flavors has $11^2 = 121$ parton flavor combinations. Getting evolved DPDs from pre-computed grids would thus require an enormous amount of memory (with an added penalty of accessing this memory for interpolation in the five variables). The path we have chosen instead is to store a DPD only for a pair of initial scales, and to evolve "on the fly" to the desired

values of μ_1 and μ_2. In the above example, this reduces the memory imprint to $64^2 \times 48$ points and thus to 1.5 MiB per flavor combination.

We will see in Sect. 3.4 that the computational effort for evolving a DPD with p_x grid points in x_1 and x_2 scales like p_x^3, whilst the scaling with the number p_y of points in y is at most linear. It is therefore of great importance to limit the size of interpolation grids for the momentum fractions, without sacrificing numerical accuracy. In our previous work [55] we found that a highly accurate interpolation and evolution of PDFs in x space is possible using Chebyshev grids. In the present paper, we show how to adapt this approach to the case of DPDs. We find that with $p_x \sim 54$ and $p_y \sim 40$ to 56 points, one can achieve high accuracy for the interpolation and integration of DPDs with $x_1, x_2 \geq 10^{-5}$ and $y \geq (7\,\mathrm{TeV})^{-1}$.

We have implemented our approach in the C++ library CHILIPDF,[1] which is under development and which will be made public in the future. Our implementation demonstrates that the methods we will describe do work in practice. It also allows us to explore a number of physics aspects. So far, DPD evolution has been studied only with DGLAP kernels at leading order (LO). We extend this to next-to-leading and next-to-next-to-leading order (NLO and NNLO). Moreover, our implementation includes the change in the number of active quark flavors at specified matching scales, and access to distributions with different scales (μ_1 and μ_2) and flavor numbers (n_1 and n_2) for the two partons. Last but not least, we can evolve DPDs for polarized partons, which may for instance have measurable impact on observables in like-sign W production according to the studies in Refs. [12,13].

This paper is organized as follows. In Sect. 2.1 we recall some basics about DPDs and set up the corresponding notation. Using our implementation for a physics study, we compare in Sect. 2.2 the evolution of DPDs at LO, NLO, and NNLO. In Sect. 3, we summarize the method we developed in Ref. [55] for interpolating and evolving PDFs and describe its extension to DPDs. We study the joint interpolation in the two momentum fractions x_1 and x_2 in Sect. 4, and the evolution in the two scales μ_1 and μ_2 in Sect. 5. Section 6 is devoted to the interpolation in the distance y between the partons, from some minimum value to infinity. In Sect. 7 we cross check our implementation against the DPD evolution code DOVE, which was originally introduced in Ref. [28] and further developed in subsequent work. We conclude in Sect. 8. In Appendix A we prove an important statement about the independence of DPD evolution and flavor matching on the path taken in the (μ_1, μ_2) plane.

---

[1] <u>C</u>hebyshev <u>I</u>nterpolation <u>L</u>ibrary for <u>PDF</u>s.

2 Basics and main results

2.1 Theory framework

To begin with, let us recall some properties of DPDs that are relevant to this work. A DPD $F_{a_1 a_2}^{n_1, n_2}(x_1, x_2, y; \mu_1, \mu_2)$ depends on the longitudinal momentum fractions x_1 and x_2 of the two partons, on their distance y in the transverse plane, and on the factorization scales μ_1 and μ_2 associated with each parton. The labels a_1 and a_2 specify the flavor and polarization of each parton. DPDs involving transverse quark or linear gluon polarization have open Lorenz indices and depend on the difference $\vec{y}$ of the transverse parton positions rather than on the length y of this vector. As specified in section 2 of Ref. [56], they can be decomposed into scalar distributions that depend on y. The evolution and matching equations discussed below hold separately for these scalar distributions. Throughout this work, we consider only DPDs in the color-singlet channel, i.e. where the color is summed over separately for each of the two extracted partons.

Notice that we allow different numbers n_1 and n_2 of active quark flavors for the two partons, which is useful when the hard-scattering processes initiated by a_1 and a_2 take place at very different scales. Technically, n_1 and n_2 are specified by the renormalisation prescription for the twist-two operator pertaining to parton a_1 and a_2, respectively.

The scale dependence of a DPD is given by the evolution equations

$$
\frac{d}{d \ln \mu_1^2} F_{a_1 a_2}^{n_1, n_2}(x_1, x_2, y; \mu_1, \mu_2)
$$
$$
= \sum_{b_1} \left[P_{a_1 b_1}^{n_1}(\mu_1) \underset{1}{\otimes} F_{b_1 a_2}^{n_1, n_2}(x_2, y; \mu_1, \mu_2) \right](x_1),
$$
$$
\frac{d}{d \ln \mu_2^2} F_{a_1 a_2}^{n_1, n_2}(x_1, x_2, y; \mu_1, \mu_2)
$$
$$
= \sum_{b_2} \left[P_{a_2 b_2}^{n_2}(\mu_2) \underset{2}{\otimes} F_{a_1 b_2}^{n_1, n_2}(x_1, y; \mu_1, \mu_2) \right](x_2) \tag{1}
$$

with separate Mellin convolutions

$$
\left[P(\mu_1) \underset{1}{\otimes} F(x_2, y; \mu_1, \mu_2) \right](x_1)
$$
$$
= \int_{x_1}^1 \frac{dz}{z} \, P(z; \mu_1) \, F\left(\frac{x_1}{z}, x_2, y; \mu_1, \mu_2 \right),
$$
$$
\left[P(\mu_2) \underset{2}{\otimes} F(x_1, y; \mu_1, \mu_2) \right](x_2)
$$
$$
= \int_{x_2}^1 \frac{dz}{z} \, P(z; \mu_2) \, F\left(x_1, \frac{x_2}{z}, y; \mu_1, \mu_2 \right) \tag{2}
$$

for the two momentum fractions. Although we define each convolution integral in Eq. (2) with the upper integration limit equal to 1, the effective integration range is reduced by the

support property

$$
F_{a_1 a_2}^{n_1, n_2}(x_1, x_2, y; \mu_1, \mu_2) = 0 \quad \text{for } x_1 + x_2 > 1 \tag{3}
$$

of DPDs, which is conserved by the evolution equations (1). The evolution kernels $P_{ab}^n(z; \mu)$ in Eq. (1) are identical to the familiar DGLAP kernels for the evolution of PDFs with n active quark flavors. Note also that the distance y plays no active role in the scale evolution, so that each "slice" of a DPD at fixed y evolves by itself.

We remark in passing that the homogeneous evolution equations (1) hold for DPDs depending on the interparton distance y. An inhomogeneous term appears if one integrates the distributions over $\vec{y}$ or performs a Fourier transform from $\vec{y}$ to its conjugate transverse momentum [19,24]. This term is closely related with the splitting contribution (7) discussed below and has been extensively studied in the earlier literature [28,57–60]. Throughout this work, we will exclusively work with DPDs that depend on y and evolve as given in Eq. (1).

The transition from $n - 1$ to n active flavors in a DPD is described by matching equations

$$
F_{a_1 a_2}^{n_1, n_2}(x_1, x_2, y; \mu_1, \mu_2) = \sum_{b_1} \Big[A_{a_1 b_1}^{n_1}(m_{n_1}; \mu_1)
$$
$$
\underset{1}{\otimes} F_{b_1 a_2}^{n_1-1, n_2}(x_2, y; \mu_1, \mu_2) \Big](x_1),
$$
$$
F_{a_1 a_2}^{n_1, n_2}(x_1, x_2, y; \mu_1, \mu_2) = \sum_{b_2} \Big[A_{a_2 b_2}^{n_2}(m_{n_2}; \mu_2)
$$
$$
\underset{2}{\otimes} F_{a_1 b_2}^{n_1, n_2-1}(x_1, y; \mu_1, \mu_2) \Big](x_2), \tag{4}
$$

where m_n is the mass of the quark with flavor number n and the matching kernels $A_{ab}^n(z, m_n; \mu)$ are the same as the ones for PDFs [61]:

$$
f_a^n(x; \mu) = \sum_b \left[A_{ab}^n(m_n; \mu) \otimes f_b^{n-1}(\mu) \right](x). \tag{5}
$$

Equation (1) is a coupled set of differential equations in μ_1 and μ_2. In Appendix A we show that with an initial condition at some point (μ_{01}, μ_{02}) one obtains a unique DPD regardless of the evolution path taken in the (μ_1, μ_2) plane. We will also show that the result of successive flavor matching does not depend on the order in which the flavor thresholds for the two partons are crossed. Note that the path independence of evolution and flavor matching in the (μ_1, μ_2) plane is exact even if the perturbative expansions of the DGLAP and matching kernels are truncated. This is in contrast to the independence on the scale at which the flavor matching is performed, which only holds to the perturbative order of the truncation.

The following evolution and matching kernels are currently implemented in CHILIPDF and used in the present work:

- DGLAP evolution for unpolarized and longitudinally polarized partons up to NNLO (order α_s^3). For the NNLO kernels we use the approximate parameterized forms given in Refs. [62–65]. The kernels in these references were confirmed by the calculations in Refs. [66,67].
- DGLAP evolution up to NLO for transversely polarized quarks and linearly polarized gluons, with the kernels given in Refs. [68,69].
- Flavor matching for unpolarized partons up to NNLO, i.e. order α_s^2. The corresponding kernels were computed independently in Refs. [61,70–72], and we verified that they agree with each other.
- Flavor matching for polarized partons up to NLO. At order α_s, the matching kernels are directly proportional to the DGLAP kernels for the same polarization.

We now take a closer look at the dependence of DPDs on the kinematic variables x_1, x_2, and y. The y dependence is not directly observable, because DPDs enter cross sections (differential or integrated) via so-called double parton luminosities

$$
\begin{aligned}
&\mathcal{L}_{a_1 a_2, b_1 b_2}^{n_1, n_2}(x_1, x_2, \bar{x}_1, \bar{x}_2; \mu_1, \mu_2, y_{\min}) \\
&= \int \mathrm{d}^2\vec{y}\ \theta(y - y_{\min})\, F_{a_1 a_2}^{n_1, n_2}(x_1, x_2, y; \mu_1, \mu_2) \\
&\quad \times F_{b_1 b_2}^{n_1, n_2}(\bar{x}_1, \bar{x}_2, y; \mu_1, \mu_2),
\end{aligned}
\tag{6}
$$

where y is integrated over with a lower cutoff $y_{\min}$ of order $1/\min(\mu_1, \mu_2)$. The origin of this cutoff is explained in Ref. [24], where it is also shown how the dependence on this cutoff is removed when the cross sections for single and double parton scattering are combined.

The form in Eq. (6) holds for unpolarized and longitudinally polarized partons; for transverse quark or linear gluon distributions additional tensors depending on $\vec{y}$ appear under the integral.

At sufficiently small distances y, the dominant contribution to a DPD arises from the perturbative splitting of a single parton into the two observed partons a_1 and a_2. At leading order in the coupling, this mechanism gives [19,24]

$$
\begin{aligned}
F_{a_1 a_2}^{\mathrm{spl}}(x_1, x_2, y; \mu_y, \mu_y) &= \frac{1}{\pi y^2}\, \frac{\alpha_s(\mu_y)}{2\pi} \\
&\times V_{a_1 a_2, a_0}^{(1)}\!\left(\frac{x_1}{x_1 + x_2}\right) \frac{f_{a_0}(x_1 + x_2; \mu_y)}{x_1 + x_2},
\end{aligned}
\tag{7}
$$

where the scale μ_y should be taken of order $1/y$ so as to avoid large logarithms $\ln(y\mu)$ in higher-order corrections.

In addition, there is an "intrinsic" two-parton contribution, which lacks the $1/y^2$ enhancement of the splitting part at small y but grows more strongly at small x_1 and x_2 for gluons and sea quarks.

At distances y of nonperturbative size, our knowledge of DPDs is rather limited, and in practice one needs to make a model ansatz. Where needed, we will take recourse to the model used in Refs. [24,30], where a DPD is written as a sum $F^{\mathrm{int}} + F^{\mathrm{spl}}$ of intrinsic and splitting parts at all y. At a starting scale μ_0, the intrinsic part of the DPD is assumed to be proportional to the product of two PDFs:

$$
\begin{aligned}
F_{a_1 a_2}^{\mathrm{int}}(x_1, x_2, y; \mu_0, \mu_0) &= \frac{1}{4\pi h_{a_1 a_2}}\, \exp\!\left[-\frac{y^2}{4 h_{a_1 a_2}}\right] \\
&\times \rho(x_1, x_2)\, f_{a_1}(x_1; \mu_0)\, f_{a_2}(x_2; \mu_0),
\end{aligned}
\tag{8}
$$

with a phase space factor

$$
\rho(x_1, x_2) = \frac{(1 - x_1 - x_2)^r}{(1 - x_1)^r (1 - x_2)^r}
\tag{9}
$$

to ensure that the DPD smoothly goes to zero at the kinematic boundary $x_1 + x_2 \to 1$. Other models in the literature take only a power $(1 - x_1 - x_2)^r$; for the motivation of the form (9) we point to section 3.2 in Ref. [28].

The splitting term in the model of Refs. [24,30] is given by the perturbative splitting expression (7) multiplied with the same exponential factor $\exp\!\left[-y^2/(4 h_{a_1 a_2})\right]$ as in the intrinsic part. This retains the correct small-y limit whilst providing a more realistic decrease at large y.

2.2 Quantitative impact of higher orders in DPD evolution

With the methods presented in this work, we are able to evolve DPDs at different perturbative orders, for all relevant polarization combinations and with different scales and active flavor numbers of the two partons. As a demonstration, we now compare DPDs evolved at LO, NLO, and NNLO from a common starting condition.

As starting condition we take the perturbative splitting form (7) with $n_1 = n_2 = 3$ active flavors. The splitting formula is evaluated at $\mu_y = 2\,\mathrm{GeV}$ and $y \approx 0.561\,\mathrm{GeV}^{-1}$, which corresponds to the scale choice $\mu_y = b_0/y$ with $b_0 = 2 e^{-\gamma_E} \approx 1.12$, where γ_E is the Euler–Mascheroni constant. This choice is suggested by the form of the NLO corrections to the splitting formula (7) and was found to keep these corrections at a moderate size [73,74]. For the PDFs in the DPD splitting formula, we take the default NNLO set of the MSHT20 parameterization [75].[2] The parameterization

---

[2] We use the central PDF from the `MSHT20nnlo_as118` set obtained via LHAPDF [54]. The CHILIPDF library includes an interface for importing data from LHAPDF.
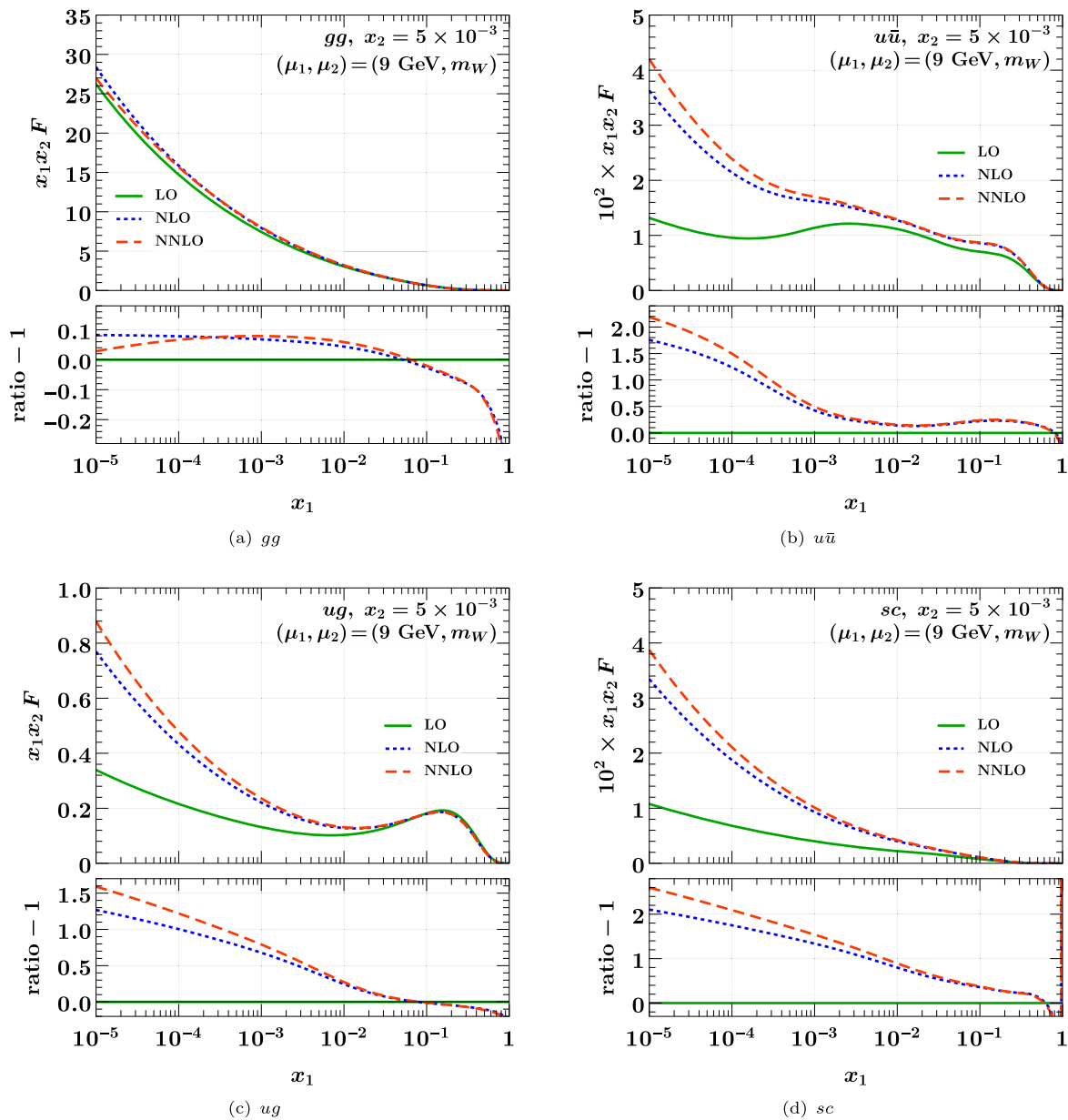
Fig. 1 Unpolarized DPDs at $x_2 = 5 \times 10^{-3}$, evolved from $(\mu_1, \mu_2) = (2\,\text{GeV}, 2\,\text{GeV})$ to $(9\,\text{GeV}, m_W)$ at different perturbative orders. The starting conditions and the procedure of flavor matching are specified in the text. The ratio shown in the small panels is taken with respect to the LO result

assumes quark masses $m_c = 1.4\,\text{GeV}$, $m_b = 4.75\,\text{GeV}$ and a five-flavor coupling $\alpha_s(m_Z) = 0.118$.

We evolve these initial conditions to $\mu_1 = 9\,\text{GeV}$ with $n_1 = 3$ for the first parton and to $\mu_2 = m_W$ with $n_2 = 5$ for the second one. This is a setting relevant for the production of a J/Ψ and a W, a channel for which double parton scattering has in fact been observed experimentally [4,76]. Flavor matching for the second parton is performed at the charm and bottom quark masses, with kernels of the same order (LO, NLO, or NNLO) as the evolution kernels. An exception is

longitudinally polarized evolution at NNLO, where we use NLO flavor matching.

We show the effect of the higher-order DGLAP evolution for a selection of unpolarized DPDs in Fig. 1. The distributions are given in the top panel of each plot, and their ratios with respect to the leading-order distribution in the lower panel. At small momentum fractions, higher-order evolution has an effect of order 10% on F_{gg}, whereas for the other flavor combinations shown in the figure, the effect is of order 1. In all cases, we find that the change from LO to NLO is
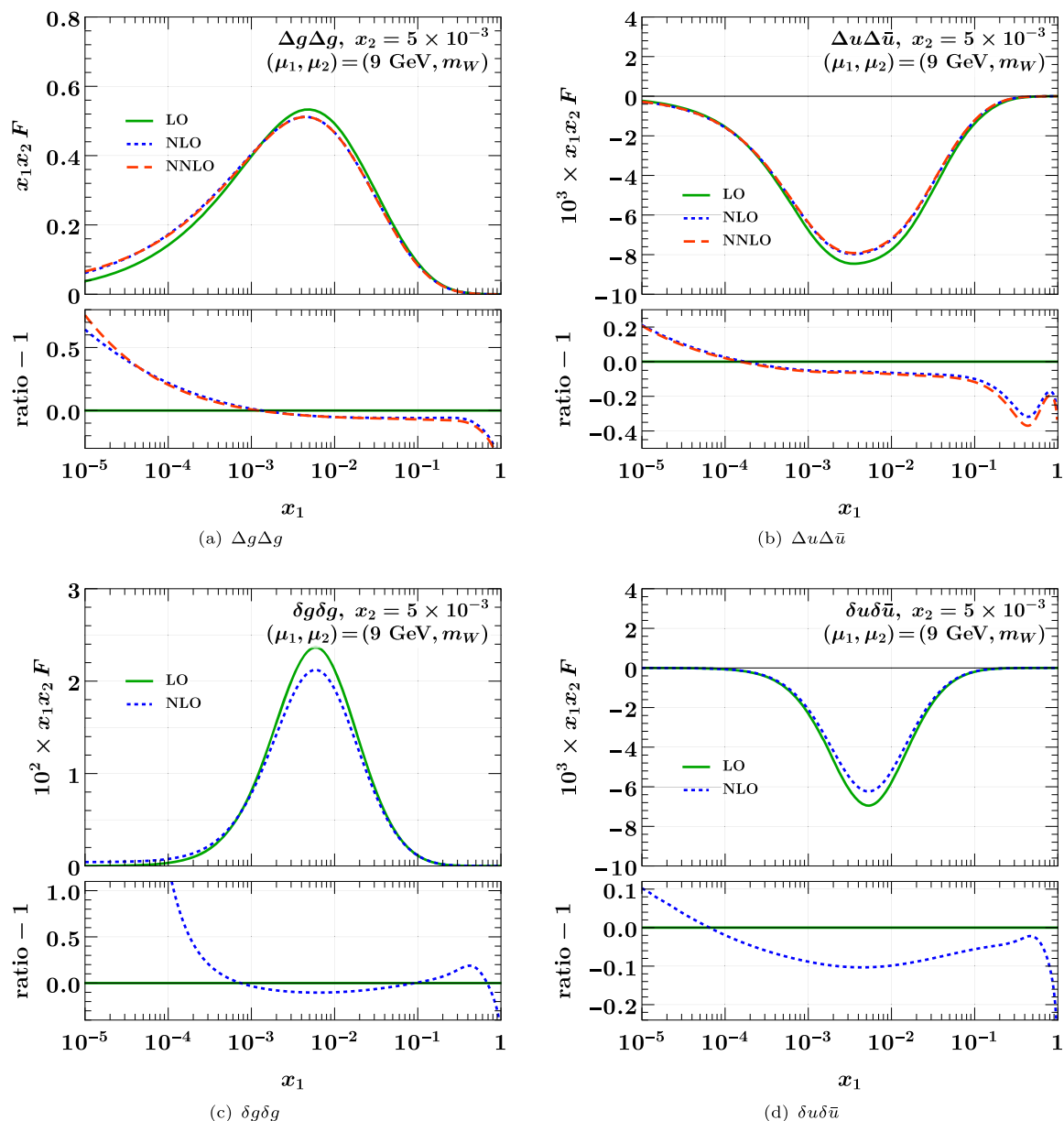
Fig. 2 As Fig. 1 but for polarized DPDs. Panels **a** and **b** are for longitudinal, panel **c** for linear, and panel **d** for transverse polarization

much larger than the rather moderate change from NLO to NNLO.

A selection of polarized DPDs is shown in Fig. 2. For linear gluon and transverse quark polarization, we have only implemented evolution at LO and NLO. We find that the NLO corrections to evolution introduce effects of order 10% for $\Delta u \Delta \bar{u}$ and $\delta u \delta \bar{u}$ in different regions of x_1, where as for $\Delta g \Delta g$ they reach 70% at small x_1. For $\delta g \delta g$, one obtains a huge relative difference between NLO and LO evolution in the limit $x_1 \ll x_2$. This is because for $z \to 0$ the splitting function $P_{\delta g \delta g}(z)$ vanishes like z at LO but grows like $1/z$ at NLO [69]. At both orders, the evolved distribution at $x_1 \ll x_2$ remains however small compared with its peak around $x_1 \sim x_2$.

As is the case for PDFs, DPDs are not observables, and distributions evolved at different orders are combined with parton-level cross sections computed at different orders. Nevertheless, the strong evolution effects we observed in several channels and kinematic regions suggest that NLO corrections in double parton scattering can be of substantial size.

In Fig. 3, we illustrate the scale dependence of DPDs, using evolution at NNLO and flavor matching at the highest available order (NNLO for unpolarized partons and NLO for polarized ones). We show the distributions at the starting point $(\mu_1, \mu_2) = (2\,\text{GeV}, 2\,\text{GeV})$, and after evolution to different combinations of the scales $\mu_i = 9\,\text{GeV}$ or $\mu_i = m_W$ for the two partons, with associated flavor numbers $n_i = 3$ in the first and $n_i = 5$ in the second case. For
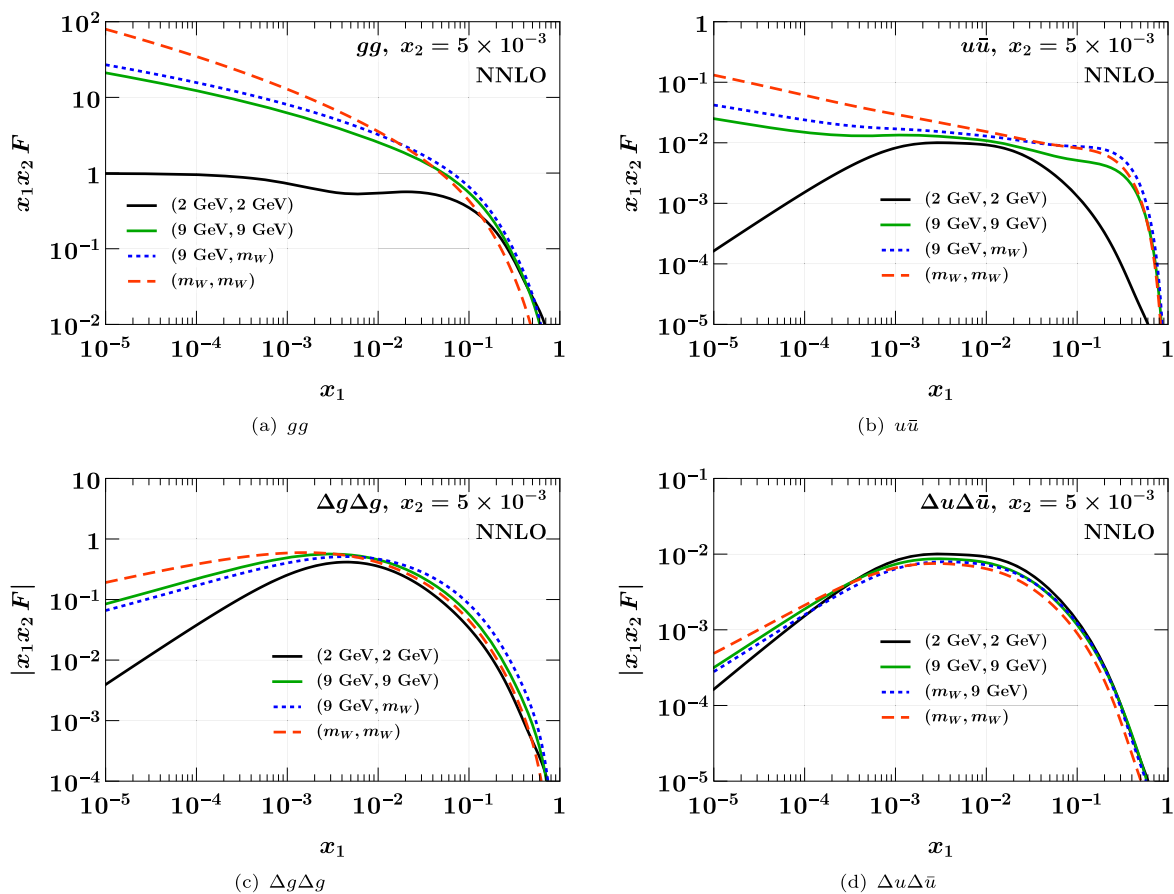
(a) gg

(b) $u\bar{u}$

(c) $\Delta g \Delta g$

(d) $\Delta u \Delta \bar{u}$

Fig. 3 Comparison of DPDs at different scales. The starting conditions at $(\mu_1, \mu_2) = (2\,\text{GeV}, 2\,\text{GeV})$ and $(n_1, n_2) = (3, 3)$ are the same as those used for the previous two figures. At higher scales, the number of active flavors is $n_i = 3$ for $\mu_i = 9\,\text{GeV}$ and $n_i = 5$ for $\mu_i = m_W$. Evolution and flavor matching are carried out at the highest available order

unpolarized partons, we find that the effect of evolution from $\mu_1 = \mu_2 = 2\,\text{GeV}$ to $9\,\text{GeV}$ is huge at small x, turning a shape that is flat or increasing with x at the low scale into a shape decreasing with x at the higher scale. Subsequent evolution to m_W for one or both partons has a significant, although less dramatic, effect. For polarized partons, evolution is generally weaker, and we find DPDs increasing with x at small x even when both partons are taken at the high scale m_W.

3 Interpolation and evolution of DPDs in ChiliPDF

In this section we sketch the mathematical framework underlying CHILIPDF, summarizing the relevant methods presented in Ref. [55] and extending them from PDFs to DPDs. In particular, we give explicit formulae for interpolation and integration, and for the discretization of the DGLAP evolu-

tion equations. A detailed account of Chebyshev interpolation and its applications can be found in Ref. [77].[3]

3.1 Chebyshev interpolation

We start by reviewing Chebyshev interpolation of a function $f(t)$ on the interval $t \in [-1, 1]$. The function is discretized on a grid consisting of the so-called *Chebyshev points*, which for given N read

$$t_i = \cos\left(\frac{i\pi}{N}\right) \quad \text{with } i = 0, \ldots, N. \tag{10}$$

These are the points where the Nth Chebyshev polynomial of the first kind, $T_N(t)$, assumes its maxima $+1$ and minima -1. They form a descending series from $t_0 = 1$ to $t_N = -1$ and satisfy the symmetry property $t_{N-i} = -t_i$. We call the set of Chebyshev points a *Chebyshev grid*.

---

[3] The first chapters of this book are available on https://people.maths.ox.ac.uk/trefethen/ATAP.

A sufficiently smooth function $f(t)$ on $t \in [-1, 1]$ can be approximated by the *Chebyshev interpolant* $p_N(t)$. This is a series of Chebyshev polynomials $T_j(t)$ with $j = 0, \ldots, N$ whose expansion coefficients are such that one has $p_N(t_i) = f(t_i)$ at the Chebyshev points. For details see Eqs. (2.6) to (2.8) in Ref. [55].

Barycentric formula. A simple and efficient way to compute the Chebyshev interpolant without having to evaluate Chebyshev polynomials is given by the *barycentric formula*

$$p_N(t) = \sum_{i=0}^{N} f(t_i) \, b_i(t) \tag{11}$$

with the barycentric basis functions

$$b_i(t) = \beta_i \frac{(-1)^i}{t - t_i} \bigg/ \sum_{j=0}^{N} \beta_j \frac{(-1)^j}{t - t_j}, \tag{12}$$

where $\beta_0 = \beta_N = 1/2$ and $\beta_i = 1$ otherwise. Even though it is not directly evident from Eq. (12), the $b_i(t)$ are polynomials of order N. The number of operations for evaluating the barycentric formula scales linearly with N. The barycentric formula is found to be numerically stable in the interpolation interval (however, this is *not* the case for extrapolating the function $f(t)$ outside this interval, see chapter 5 of Ref. [77]).

Accuracy estimate. A computationally inexpensive estimate for the accuracy of Chebyshev interpolation is obtained by interpolating $f(t)$ on the Chebyshev grid *without* the end points t_0 and t_N. The resulting interpolant $q_{N-2}(t)$ is a polynomial of order $N - 2$ and can be computed with an adapted form of the barycentric formula, see Eqs. (2.23) to (2.27) in Ref. [55]. Comparing $p_N(t)$ with $q_{N-2}(t)$ yields an estimate for the interpolation accuracy, i.e. for the difference between $p_N(t)$ and $f(t)$. This estimate was found to work well for PDF interpolation [55], and we will investigate in later chapters how reliable it is in the case of DPDs.

Integration. Using the expansion of f in terms of Chebyshev polynomials $T_i(t)$ and the analytic form for integrals of these polynomials, one readily obtains the integration rule

$$\int_{-1}^{1} dt \, f(t) \approx \sum_{i=0}^{N} w_i \, f(t_i) \tag{13}$$

with weights

$$w_i = \frac{4\beta_i}{N} \sum_{\substack{j=0 \\ \text{even}}}^{N} \beta_j \frac{\cos(j\theta_i)}{1 - j^2}. \tag{14}$$

This integration rule is known as *Clenshaw–Curtis quadrature*. For a detailed discussion of its accuracy (and comparison with Gauss quadrature), we refer to chapter 19 of Ref. [77] and to section 2 of Ref. [55].

An estimate for the integration accuracy can be obtained by using a quadrature rule on the Chebyshev grid without its end points, in full analogy to the procedure described above for interpolation. The result is known as *Fejér's second rule*. Details are given in Eqs. (2.31) to (2.33) of Ref. [55].

3.2 Interpolation strategy

In this subsection, we consider the scales μ_1 and μ_2 to be fixed and omit them as arguments for brevity. We discuss the discretization of DPDs in the three variables x_1, x_2, and y, which we generically denote by z. Depending on the typical dependence of a DPD on z, different variable transformations $z \to u$ are used to map an interval $[z_{\min}, z_{\max}]$ onto a finite interval $[u_{\min}, u_{\max}]$. For the momentum fractions x_1 and x_2 we always use the transformation

$$u(x_i) = \ln x_i \quad \text{for } i = 1, 2, \tag{15}$$

which corresponds to the choice for interpolating PDFs in Ref. [55]. For both x_1 and x_2, the interpolation interval has a fixed upper limit $x_{\max} = 1$, whereas the lower limits may be equal or different (but must be nonzero). For the y dependence (where $y_{\max}$ may be infinite or finite) we use a number of different transformations, which are presented in Sect. 6.1.

For reasons given below, the z intervals are usually split into a few subintervals. On each subinterval, we perform a linear transformation from u to a variable $t \in [-1, 1]$, which is used for Chebyshev interpolation as described in Sect. 3.1.

Consider a particular subgrid with $N + 1$ points $u_0, \ldots, u_N$, which is mapped by a linear transform onto the Chebyshev grid $t_0, \ldots, t_N$ given in Eq. (10). The corresponding grid points in the physical variable are then given by the inverse variable transformation $z_i = z(u_i)$. Similarly to the PDF case, it is advantageous to interpolate DPDs scaled by their momentum fractions,

$$\widetilde{F}(x_1, x_2, y) = x_1 x_2 F(x_1, x_2, y), \tag{16}$$

because this leads to less steep functions in x_1 and x_2. The interpolation in one variable is given by the barycentric formula

$$\widetilde{F}(z, \ldots) \approx \sum_{i=0}^{N} \widetilde{F}_i(\ldots) \, b_i(u(z)) \quad \text{for } z_0 \leq z \leq z_N, \tag{17}$$

where the ellipsis denotes the other variables and

$$\widetilde{F}_i(\ldots) = \widetilde{F}(z_i, \ldots),$$

$$b_i(u) = \beta_i \frac{(-1)^i}{u - u_i} \Big/ \sum_{j=0}^{N} \beta_j \frac{(-1)^j}{u - u_j}. \tag{18}$$

The interpolation of the full DPD is obtained by discretization in both momentum fractions and the interparton distance and reads

$$\widetilde{F}(x_1, x_2, y) \approx \sum_{i=0}^{N_{x,1}} \sum_{j=0}^{N_{x,2}} \sum_{k=0}^{N_y} \widetilde{F}_{ijk}$$
$$\times b_i\big(\ln(x_1)\big)\, b_j\big(\ln(x_2)\big)\, b_k\big(u(y)\big), \tag{19}$$

where $N_{x,1}$, $N_{x,2}$, and N_y specify the polynomial order of interpolation in each variable, and

$$\widetilde{F}_{ijk} = \widetilde{F}(x_{1,i}, x_{2,j}, y_k) \tag{20}$$

denotes the values of $\widetilde{F}$ on the interpolation grid. Equation (18) reflects that the form of the barycentric basis functions (12) remains unchanged under a linear transform of the interpolation variable. Analogous formulae can be used to interpolate functions derived from DPDs, such as the Mellin convolutions of DPDs with an integral kernel (see Sect. 3.3).

Note that we interpolate $\widetilde{F}(x_1, x_2, y)$ on a full rectangle in the (x_1, x_2) plane, given by $x_1 \in [x_{1,\min}, 1]$ and $x_2 \in [x_{2,\min}, 1]$. This includes the region $x_1 + x_2 > 1$ where the DPD is zero. In this way, we can use the same interpolation grid in x_1 for all values of x_2 and vice versa, which in turn permits a simple implementation of DGLAP evolution for the two partons. The implication of this choice on the interpolation accuracy is discussed in Sect. 4.

Chebyshev interpolation is "global" in the sense that the interpolant at a given point z depends on the function values at *all* points z_i in the interpolation interval. For grids with many points, this can lead to an accumulation of rounding errors, especially in regions of z where the function is much smaller than elsewhere in the interval. To limit this effect, it is advantageous to split the full z domain into several subintervals and to use the barycentric formula (19) separately on each subinterval. In Ref. [55] this was found to substantially improve the interpolation accuracy for PDFs, provided that the number of points in each subinterval does not become too small. We find that the same is true for the interpolation of DPDs in x_1 and x_2, as well as in y.

To specify such a composite grid with k individual subgrids, we use the notation

$$[z_0, z_1, \ldots, z_k]_{(p_1, p_2, \ldots, p_k)}, \tag{21}$$

where z_i are the subinterval boundaries and $p_i = N_i + 1$ is the number of Chebyshev points in subgrid number i. We refer to this as a $(p_1, p_2, \ldots, p_k)$-point grid. Note that neighboring subgrids share their end points, such that the total number of grid points is $p_{\text{total}} = \sum_{i=1}^{k} p_i - (k-1)$.

Splitting the interpolation interval in y has the additional benefit that one can use different variable transformations $u(y)$ for regions with different characteristic y dependence, such as the power law behavior from perturbative splitting and an exponential decrease at large distances. Details will be discussed in Sect. 6.1.

3.3 Mellin convolution

Consider now the convolution of a DPD with an integral kernel, which may be a DGLAP evolution kernel, a flavor matching kernel, or a hard-scattering coefficient. The general expression to be computed reads

$$\big(K \otimes \widetilde{F}\big)(x) = \int_x^1 \frac{\mathrm{d}z}{z}\, K(z)\, \widetilde{F}\Big(\frac{x}{z}\Big), \tag{22}$$

where x stands for x_1 or x_2 and the dependence of the rescaled DPD on all other variables is omitted for now. The kernel $K(z)$ is appropriately rescaled, given that $g_1 = g_2 \otimes g_3$ implies $h_1 = h_2 \otimes h_3$ with $h_k(z) = z\, g_k(z)$.

To discretize Eq. (22), we consider for simplicity a single Chebyshev grid with p_x points for $x \in [x_{\min}, 1]$. The result of the convolution $(K \otimes \widetilde{F})(x)$ is a function defined in the same x interval as $\widetilde{F}(x)$ and can hence be interpolated on the same grid. It is thus sufficient to evaluate Eq. (22) at the grid points x_i,

$$\big(K \otimes \widetilde{F}\big)(x_i) = \int_{x_i}^1 \frac{\mathrm{d}z}{z}\, K(z)\, \widetilde{F}\Big(\frac{x_i}{z}\Big)$$
$$\approx \int_{x_i}^1 \frac{\mathrm{d}z}{z}\, K(z) \sum_{i'=0}^{p_x - 1} \widetilde{F}(x_{i'})\, b_{i'}\Big(\ln \frac{x_i}{z}\Big), \tag{23}$$

where in the second step we have approximated $\widetilde{F}(x_i/z)$ by its interpolated form. Here and in the following, we use the number of grid points p_x rather than the polynomial order $N_x = p_x - 1$ to specify summation ranges. At the grid points, the convolution is thus obtained by a matrix multiplication

$$\big(K \otimes \widetilde{F}\big)(x_i) \approx \sum_{i'=0}^{p_x - 1} K_{ii'}\, \widetilde{F}(x_{i'}), \tag{24}$$

where the *kernel matrix*

$$K_{ii'} = \int_{x_i}^1 \frac{\mathrm{d}z}{z}\, K(z)\, b_{i'}\Big(\ln \frac{x_i}{z}\Big) \tag{25}$$

can be computed using standard numerical integration techniques (we use an adaptive Gauss–Kronrod routine). If $K(z)$

contains plus- or δ distributions, this requires rewriting Eq. (25) as specified in Eqs. (4.5) to (4.10) of Ref. [55].

It is straightforward to generalize the preceding discussion to the case where the interval $[x_{min}, 1]$ is split into several subintervals with corresponding subgrids. The relation (24) remains valid, with an appropriately generalized definition of the kernel matrix and with p_x being the *total* number of grid points in x.

The kernel matrix method can be readily used for flavor matching of DPDs, discretizing the Mellin convolutions with the matching kernels in Eq. (4). It is also a crucial part in DPD evolution, which is presented next.

3.4 DGLAP evolution

We now discuss the evolution of DPDs, starting with evolution in μ_1 at a fixed value of μ_2. The evolution equation for the rescaled DPD reads

$$\frac{d}{d \ln \mu_1^2} \widetilde{F}(x_1, x_2, y; \mu_1, \mu_2)$$
$$= \left[\widetilde{P}(\mu_1) \underset{1}{\otimes} \widetilde{F}(x_2, y; \mu_1, \mu_2) \right](x_1), \quad (26)$$

where $\widetilde{P}(z; \mu_1) = z P(z; \mu_1)$ is the rescaled splitting function and we omit parton labels and the associated sums for the time being. After discretization on one or several subgrids with a total number of $p_{x,1}, p_{x,2}, p_y$ points in x_1, x_2, y, the integro-differential equation (26) turns into a linear system of ordinary differential equations in μ_1,

$$\frac{d}{d \ln \mu_1^2} \widetilde{F}_{ijk}(\mu_1, \mu_2) = \sum_{i'=0}^{p_{x,1}-1} \widetilde{P}_{ii'}(\mu_1) \widetilde{F}_{i'jk}(\mu_1, \mu_2), \quad (27)$$

where $\widetilde{F}_{ijk}(\mu_1, \mu_2)$ is given in Eq. (20) and $\widetilde{P}_{ii'}$ is the matrix for the kernel $\widetilde{P}(z)$ as defined in Eq. (25) or its generalization to several subgrids. At NNLO accuracy, we have

$$\widetilde{P}_{ii'}(\mu) = \sum_{k=0}^{2} \left[\frac{\alpha_s(\mu)}{4\pi} \right]^{k+1} \widetilde{P}_{ii'}^{(k)}, \quad (28)$$

and after a change of variables from the scales μ_i to the evolution times

$$t_i = -\ln(\alpha_s(\mu_i)) \quad \text{for } i = 1, 2, \quad (29)$$

we obtain

$$\frac{d}{dt_1} \widetilde{F}_{ijk}(t_1, t_2) = \sum_{i'=0}^{p_{x,1}-1} K_{ii'}(t_1) \widetilde{F}_{i'jk}(t_1, t_2) \quad (30)$$

with a new kernel matrix given by

$$K_{ii'}(t) = \left[\beta_0 + \frac{e^{-t}}{4\pi} \beta_1 + \frac{e^{-2t}}{(4\pi)^2} \beta_2 \right]^{-1}$$
$$\times \left[\widetilde{P}_{ii'}^{(0)} + \frac{e^{-t}}{4\pi} \widetilde{P}_{ii'}^{(1)} + \frac{e^{-2t}}{(4\pi)^2} \widetilde{P}_{ii'}^{(2)} \right], \quad (31)$$

where β_k are the expansion coefficients of the QCD β function (see Eq. (5.4) in [55] for our normalization conventions).

To solve Eq. (30) we use a Runge–Kutta algorithm with a fixed maximal step size in t_1. Notice that the t_1 dependence of $K_{ii'}(t_1)$ starts only at NLO, which leads to a more uniform accuracy of the algorithm compared with the direct solution of Eq. (27), where $\widetilde{P}_{ii'}(\mu_1)$ changes with $\ln \mu_1$ already at LO. As documented in appendix A of Ref. [55], we find that the use of a higher-order Runge–Kutta algorithm – in particular the 8th order algorithm of Dormand and Prince [78] – greatly improves the accuracy of evolution for a given number N_{RK} of intermediate time values at which the r.h.s. of the differential equation is evaluated for evolution from one value of t_1 to another.

The Runge–Kutta algorithm is also used to compute the transformation (29), solving the differential equation for the running of α_s^{-1} as a function of $\ln \mu$.

Flavor labels and sums have to be added in Eq. (30) as in the original version (1). The mixing between parton flavors under evolution is the same as for PDFs, and correspondingly we solve the evolution equations in the same basis that is used for PDFs in Ref. [55]. For each of the two partons, we thus form the linear combinations

$$\Sigma^{\pm} = \sum_i q_i^{\pm}, \quad u^{\pm} - d^{\pm}, \quad d^{\pm} - s^{\pm},$$
$$s^{\pm} - c^{\pm}, \quad c^{\pm} - b^{\pm}, \quad b^{\pm} - t^{\pm}, \quad (32)$$

where $q^{\pm} = q \pm \bar{q}$. The combination Σ^- and all flavor differences then evolve by themselves, and mixing is reduced to the combination Σ^+ and the gluon.

The system (30) needs to be solved separately for each combination of the indices j, k that correspond to the "inactive" variables x_2 and y during evolution in t_1, and for each of the $2n_2 + 1$ active flavors of the second parton. This gives a large number of $p_{x,2} \, p_y \, (2n_2 + 1)$ combinations – typically several thousand. It is much more efficient to apply the Runge–Kutta algorithm not to each of these combinations, but to a set of $p_{x,1}$ basis vectors e_i^j that generates all possible initial conditions on the grid for x_1. With the choice of basis $e_i^j = \delta_{ij}$, evolution from t_{01} to t_1 requires solving the system

$$\frac{d}{dt_1} U_{ij}(t_1, t_{01}) = \sum_{i'=0}^{p_{x,1}-1} K_{ii'}(t_1) U_{i'j}(t_1, t_{01}), \quad (33)$$

with the boundary condition $U_{ij}(t_{01}, t_{01}) = \delta_{ij}$. The solution of Eq. (30) is then obtained by a matrix multiplication

$$\widetilde{F}_{ijk}(t_1, t_2) = \sum_{i'=0}^{p_{x,1}-1} U_{ii'}(t_1, t_{01}) \, \widetilde{F}_{i'jk}(t_{01}, t_2), \qquad (34)$$

where the evolution matrix $U_{ii'}(t_1, t_{01})$ acts as the discretized version of the Green function for the original evolution equation (26). With flavor labels reinstated, Eq. (33) is solved separately for the evolution kernels of the different channels (i.e. for $q_i^+ - q_j^+$, $q_i^- - q_j^-$, Σ^-, and the coupled system of Σ^+ and the gluon).

This approach is also attractive in the PDF case when a large set of different PDFs is evolved simultaneously, and it is for instance used in the PDF evolution codes APFEL [49] and EKO [53], with $U_{ii'}(t_1, t_{01})$ being called an evolution operator and an evolution kernel operator, respectively. In CHILIPDF, the evolution of a DPD in one scale is implemented in the same way as the evolution of a set of PDFs.

The number of operations for evaluating Eq. (34) for all indices i, j, k and flavors for the second parton scales like $(p_{x,1})^2 \, p_{x,2} \, p_y \, (2n_2 + 1)$, and the number of operations for solving the evolution equation (33) scales like $(p_{x,1})^3 \, N_{RK}$ with N_{RK} introduced below Eq. (31). The number of grid points for the interpolation in the momentum fractions is therefore a major factor determining the computational speed of DPD evolution.

The preceding discussion is trivially extended to evolution in the scale μ_2 at fixed μ_1. To evolve from (μ_{01}, μ_{02}) to (μ_1, μ_2), we first evolve from (μ_{01}, μ_{02}) to (μ_1, μ_{02}) and from there to (μ_1, μ_2). We verified that evolving to (μ_{01}, μ_2) in the first step leads to the same final result with very high numerical accuracy (see also Sect. 5.1).

Let us briefly sketch how the initial conditions for DPD evolution are set up in CHILIPDF. In a first step, the values of $F_{a_1 a_2}(x_1, x_2, y; \mu(y), \mu(y))$ for a selected number n_0 of active flavors are computed on the grid points in x_1, x_2 and y, either from pre-defined functions for the splitting form (7), or from functions implementing a product ansatz like the one in (8), or from functions entirely provided by the user. As is adequate for the physics, the input scale $\mu(y)$ may depend on y as specified by the user. For each grid point y_k, the discretized DPD is then evolved to a pair of scales (μ_0, μ_0) that is independent of y and stored as an initial condition. After this, all evolution calls can use the same set of evolution matrices for all grid points y_k, which significantly reduces computing time. One may also add two DPDs (after evolution to a common scale pair), which is for instance needed for calculating $F^{int} + F^{spl}$.

In the procedure just described, the flavor number n_0 is equal for both partons. DPDs with larger values of n_1 or n_2 are then obtained by flavor matching, with initial conditions

at the pair of scales (μ_1, μ_2) where the matching is carried out. More detail is given in Sect. 5.1.

4 Interpolation of DPDs in both momentum fractions

As shown in the previous section, our approach to interpolating DPDs in x_1 and x_2 is a straightforward two-dimensional generalization of the method presented in Ref. [55] for the interpolation of PDFs in x. In the present section, we study the performance of this approach. Let us point out some particularities of the DPD case:

1. To limit computation time, interpolation grids should not have too many points, while still giving numerically reliable results. For PDFs one can typically afford denser grids, and the demands on interpolation accuracy are often higher than for DPDs.
2. The DPD is a non-analytic function of x_1 at the point $x_1 = 1 - x_2$ because it vanishes identically above that value. Since we interpolate on the same grid in x_1 for all values of x_2, this singular point is generally in the middle of an interpolation interval. Polynomial interpolation is not very accurate in the vicinity of such a point, but we will see that this effect is in general relatively mild.
3. The perturbative splitting mechanism gives a dependence on x_1 that is qualitatively different from the x dependence of a PDF. In some channels, the splitting contribution (7) goes to zero for $x_1 \ll x_2$ and thus results in a *strong* rise of the DPD as a function of x_1.

Of course, the last two points also apply to interpolation in x_2 at a given x_1.

In the following, we will study the interpolation accuracy for DPDs that have a known analytic expression. To this end, we simplify the forms (8) and (7) to

$$\begin{aligned} &\widehat{F}_{a_1 a_2}^{int(r)}(x_1, x_2) \\ &= \frac{(1 - x_1 - x_2)^r}{(1 - x_1)^r (1 - x_2)^r} \, f_{a_1}(x_1) \, f_{a_2}(x_2) \end{aligned} \qquad (35)$$

and

$$\begin{aligned} &\widehat{F}_{a_1 a_2}^{spl}(x_1, x_2) \\ &= V_{a_1 a_2, a_0}^{(1)}\left(\frac{x_1}{x_1 + x_2}\right) \frac{f_{a_0}(x_1 + x_2)}{x_1 + x_2}, \end{aligned} \qquad (36)$$

respectively. For the PDFs, we take the simple parameterizations of the Les Houches benchmark PDFs given in section 1.32 of Ref. [79] and section 4.4 of Ref. [80].

Corresponding to the typical use case, we take the same interpolation grids for x_1 and x_2 (a situation in which different

grids are useful is encountered in Sect. 7). We consider three composite grids of the form

$$[10^{-5}, 5 \times 10^{-3}, 0.5, 1]_{(p, p, p)}$$

$$\text{with } p = \begin{cases} 12 & (\text{coarse } x \text{ grid}) \\ 16 & (\text{medium } x \text{ grid}) \\ 24 & (\text{fine } x \text{ grid}) \end{cases} \tag{37}$$

where the shorthand notation from Eq. (21) has been used. We find that for the given x range, three subgrids with the above interval limits generally give the best accuracy for a given total number of points. As a measure for the quality of interpolation, we use the relative accuracy defined as

relative accuracy
$$= \left| \text{interpolated result/exact result} - 1 \right|. \tag{38}$$

In the following, this interpolation accuracy will be shown as a function of x_1 for fixed values of x_2. The latter are *not* grid points, so that interpolation is performed for both momentum fractions.

4.1 Interpolation of the product form

We start with the product form (35) for the three values $r = 0, 1, 2$ of the power in the phase space factor. The case $r = 0$ corresponds to a "naive" product form with a discontinuity at $x_1 + x_2 = 1$ that is not very plausible on physics grounds. In fact, evolution to higher scales quickly changes such a discontinuity to a steep but continuous decrease. We nevertheless include the case $r = 0$ here, in order to see how our method performs in this limiting case. The following plots show the $u\bar{u}$ distribution; we verified that for other parton combinations the interpolation accuracy is not significantly better or worse.

In the left panels of Fig. 4 we see that all three grids yield an excellent interpolation accuracy at small momentum fractions. Not surprisingly, there is little dependence on the power r in this region, given that the phase space factor is very close to 1 for $x_1, x_2 \ll 1$. In the right panels of the figure, we see that a degradation of the relative accuracy sets in as x_1 approaches its kinematic limit $1 - x_2 = 0.7$, and that the degradation is stronger for smaller r. Nevertheless, the accuracy is better than 1% as long as the distance $1 - (x_1 + x_2)$ from the kinematic threshold is above 0.15 for $r = 1, 2$ and above 0.2 for $r = 0$.

The plots in Fig. 5 show the x_1 region around the point where the DPD has a non-analytic behavior, which is at $x_1 = 0.7$ in this example. In these plots, we include the region $x_1 > 1 - x_2$, where the interpolated DPD oscillates around its true value zero. Note that this region is included when one computes a Mellin convolution for the DPD with the kernel

matrix method described in Sect. 3.3. However, the effect of oscillations around the true value will partially cancel out in convolution integrals.

We see that for $r = 0$, one obtains a rather poor approximation of the DPD in the range $x_1 \in [0.6, 0.8]$ with the coarse or medium grid. In situations where this region is deemed important, one should take a dense grid and check the reliability of the results by comparing with an even denser grid. Of course, this is only needed in the relevant subinterval (i.e. for $x_1 \in [0.5, 1]$ in our example) and hence entails only a moderate increase in the total number of grid points. For $r = 1$, where the DPD is continuous but has a discontinuous first derivative at $x_1 = 0.7$, we see that the fine grid performs very well, and for $r = 2$ already the medium grid yields a good approximation of the DPD.

4.2 Interpolation of the splitting form

We now turn to the interpolation of the splitting form. We see in Eq. (36) that it decreases almost as fast as a PDF for $x_1 + x_2 \to 1$, so that according to the results of the previous subsection one does not expect particular problems with interpolation in that region. This is confirmed by our numerical studies.

The shape of the splitting form is strongly influenced by the end-point behavior of the splitting kernel $V^{(1)}_{a_1 a_2, a_0}$, and in the following we show results for two qualitatively different cases:

$$V^{(1)}_{u\bar{u},g}(z) = T_F \big[z^2 + (1-z)^2 \big],$$
$$V^{(1)}_{\delta g \delta g, g}(z) = 2C_A \, z(1-z). \tag{39}$$

Whereas the first kernel is finite for $z \to 0$ and $z \to 1$, the second one goes to zero in both limits. This implies a strong decrease of $\widehat{F}^{\text{spl}}_{\delta g \delta g}(x_1, x_2)$ for both $x_1 \ll x_2$ and $x_1 \gg x_2$. Other splitting kernels diverge at the end points – namely when an unpolarized gluon produced by the splitting becomes soft – and we verified that the interpolation accuracy for these channels is not worse than the one for $u\bar{u}$.

We see in the top row of Fig. 6 that already the medium grid yields a very good interpolation as long as the momentum fractions are not too close to the kinematic limit. At small x_1 and x_2, the accuracy is somewhat less good than for the product form in Fig. 4, which can be understood because the splitting form has more structure in that region. This is also seen for the interpolation of $\widehat{F}^{\text{spl}}_{\delta g \delta g}$ at small momentum fractions in panel (c), where the accuracy of the medium grid is at the permille level. In panel (d) we see that the linear decrease of the distribution for $x_1 \ll x_2$ requires the fine grid for a satisfactory interpolation at the smallest x_1, where the DPD is many orders of magnitude smaller than its maximum value. On the other hand, the medium grid still performs well
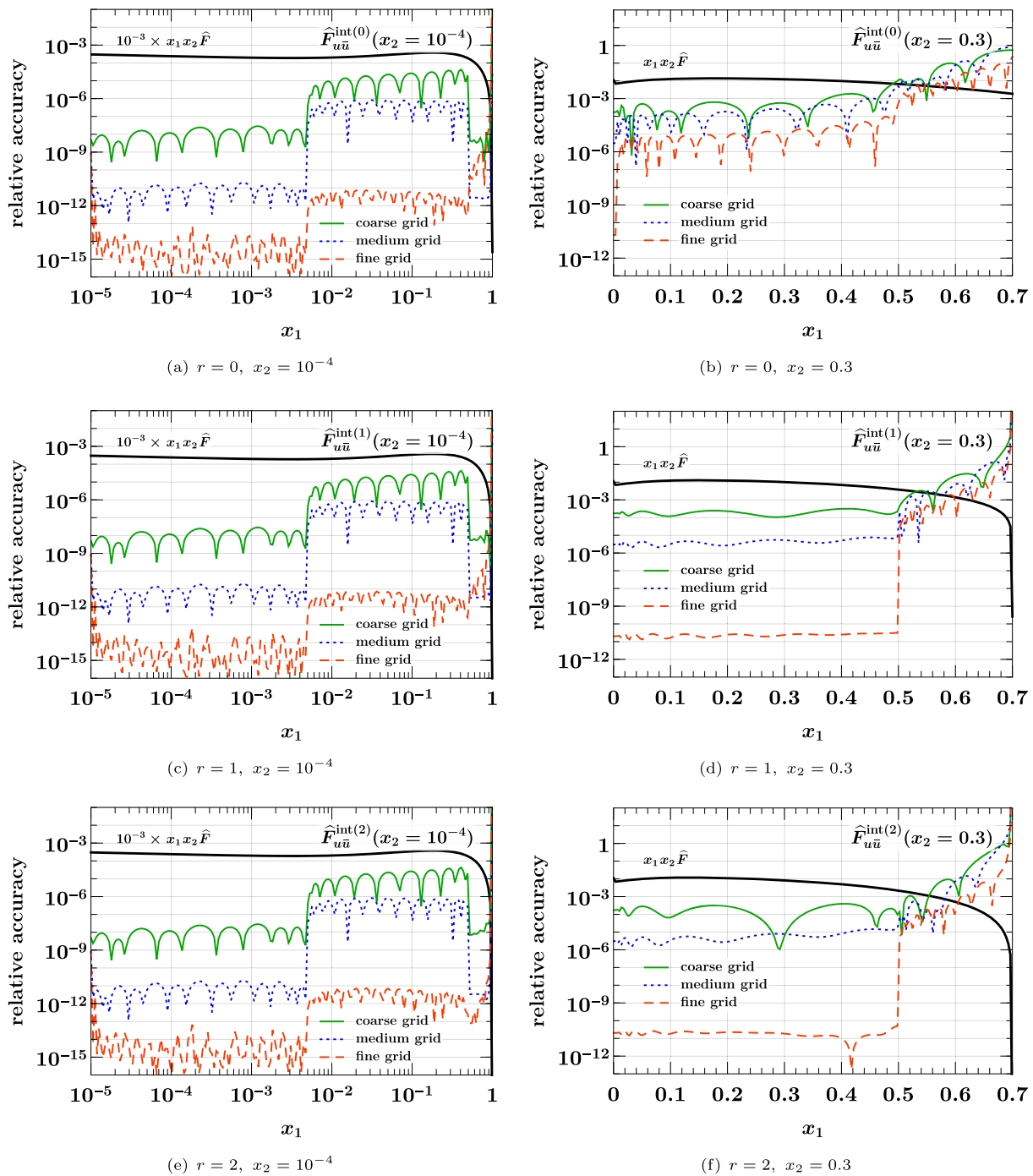
(a) $r = 0$, $x_2 = 10^{-4}$

(b) $r = 0$, $x_2 = 0.3$

(c) $r = 1$, $x_2 = 10^{-4}$

(d) $r = 1$, $x_2 = 0.3$

(e) $r = 2$, $x_2 = 10^{-4}$

(f) $r = 2$, $x_2 = 0.3$

Fig. 4 Relative accuracy (38) for interpolation in x_1 and x_2, evaluated for the product form $\widehat{F}_{u\bar{u}}^{\mathrm{int}(r)}(x_1, x_2)$ of Eq. (35) with $r = 0, 1, 2$. The interpolation grids are specified in Eq. (37). Here and in similar plots, the exact result that is being interpolated is shown in black

for $x_1 > 10^{-4}$. We note that the coarse grid does not give a good accuracy in any of the cases considered here, and it is entirely unreliable in the case of panel (d). We conclude that at least 16 points per subgrid should be taken, unless one considers only unpolarized DPDs and has only very moderate accuracy requirements.

4.3 Estimating the interpolating accuracy

Let us finally study the reliability of the accuracy estimate for interpolation described in Sect. 3.1, given by

(a) $r = 0$, $x_2 = 0.3$

(b) $r = 0$, $x_2 = 0.3$

(c) $r = 1$, $x_2 = 0.3$

(d) $r = 1$, $x_2 = 0.3$

(e) $r = 2$, $x_2 = 0.3$

(f) $r = 2$, $x_2 = 0.3$
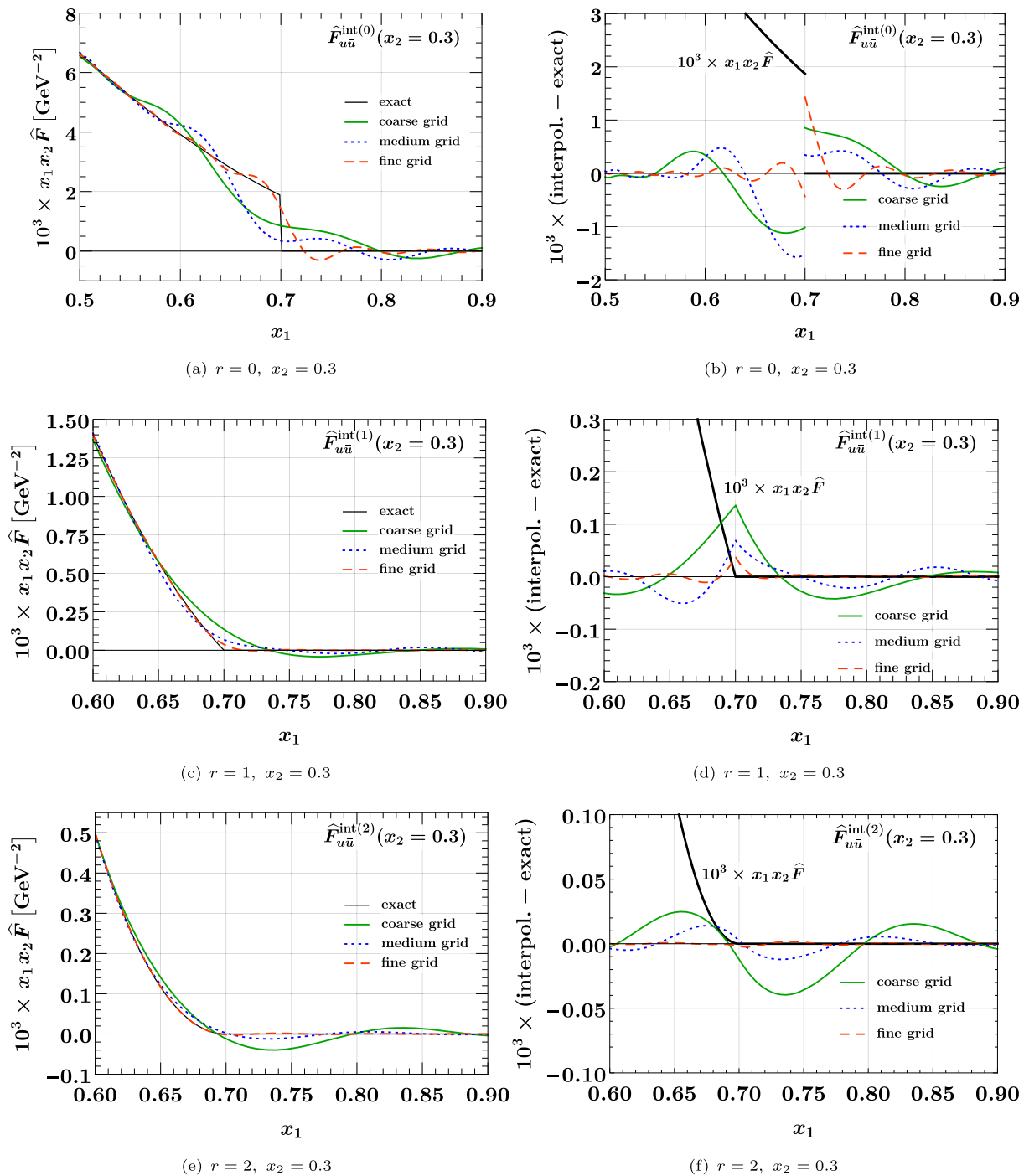
Fig. 5 Interpolation of DPDs at large momentum fractions. The DPDs considered are the same as in the right panels of Fig. 4. The plots on the left compare the exact and interpolated forms of the DPDs. The plots on the right show the difference between the interpolated and the exact values, along with the exact form. Notice that the plots in the top row have a different x_1 range than the remaining ones

relative accuracy estimate

$$= \left| \frac{\text{value interpolated without end points}}{\text{value interpolated with end points}} - 1 \right|. \quad (40)$$

In Fig. 7, we compare this estimate with the true relative accuracy (38) for the medium x grid in Eq. (37). We see that the estimate is typically not more than two orders of magnitude away from the true accuracy (except around points where the error estimate or the true error has a zero crossing).

(a) $u\bar{u}$, $x_2 = 10^{-4}$

(b) $u\bar{u}$, $x_2 = 0.3$

(c) $\delta g \delta g$, $x_2 = 10^{-4}$

(d) $\delta g \delta g$, $x_2 = 0.3$
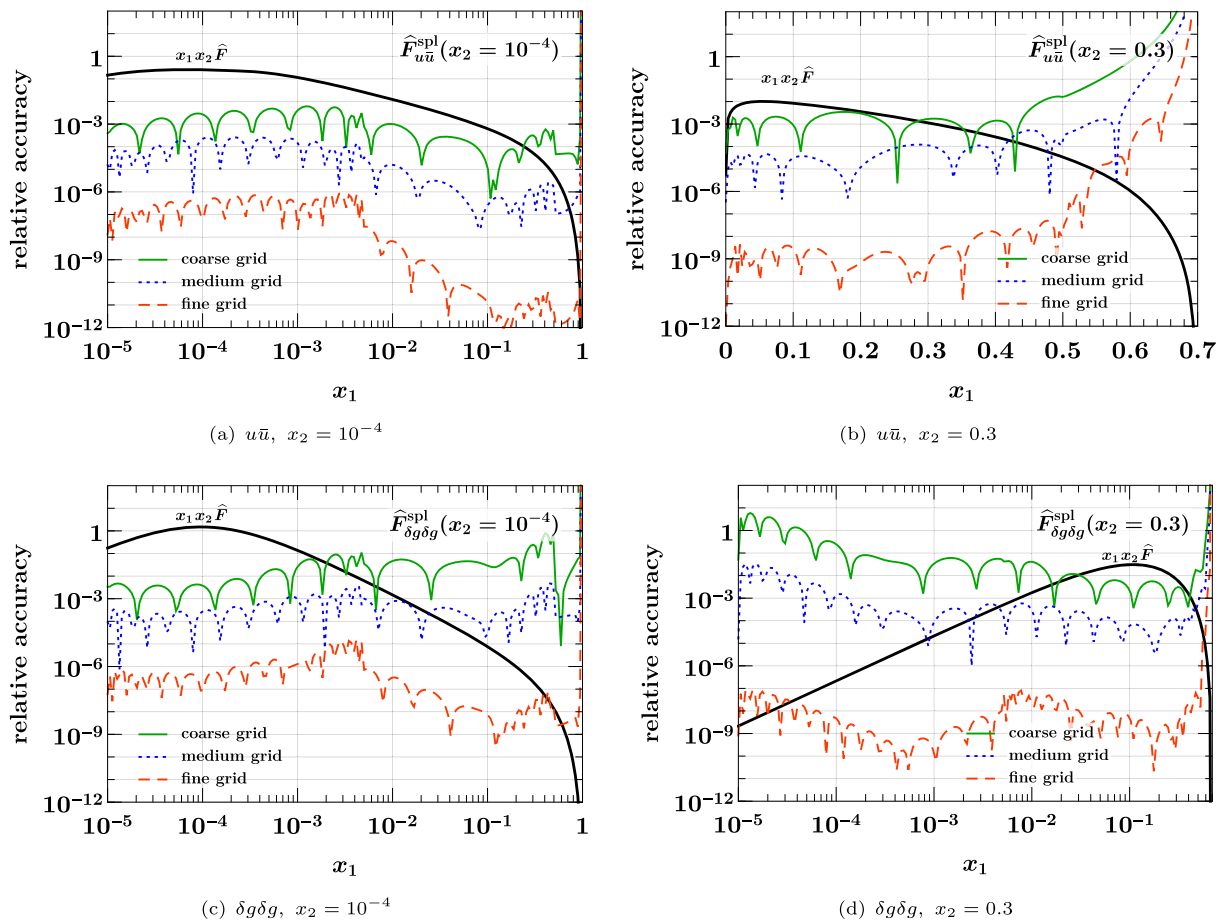
Fig. 6 Relative accuracy (38) for interpolation in x_1 and x_2, evaluated for the splitting form $\widehat{F}^{\text{spl}}_{a_1 a_2}(x_1, x_2)$ of Eq. (36). The interpolation grids are specified in Eq. (37). The plots in the top row are for an unpolarized $u\bar{u}$ pair, and those in the bottom row for a pair of linearly polarized gluons. Notice that in panel (**d**) a logarithmic scale for x_1 is used to plot the DPD at $x_2 = 0.3$.

Somewhat surprisingly, the estimate almost coincides with the true accuracy in some cases, as seen in Fig. 7a. Closer inspection reveals that the estimate tends to be above the true accuracy close to subinterval end points; this is expected because in that region interpolation without the end points is particularly bad by construction. We find a qualitatively similar picture for the accuracy estimate in other kinematic settings, and for the other two grids in Eq. (37). Although not perfectly accurate in all regions, the method based on Eq. (40) provides a good and easy-to-compute indicator for the quality of interpolation.

5 Evolution and flavor matching of DPDs

In this section, we demonstrate the accuracy of DGLAP evolution and flavor matching for DPDs with the methods presented in Sects. 3.3 and 3.4. As a default setting for DPD evolution, we take a maximal step size of $h = 0.22$ for the

evolution time (29) in the Runge–Kutta method. We use the 8th order algorithm of Dormand and Prince [78], where each Runge–Kutta step is divided into 13 substeps for which the r.h.s. of the evolution equation is evaluated.

As initial conditions we take the splitting form (7) evaluated with the Les Houches benchmark PDFs of Refs. [79,80]. Following the settings in these references, we take

$$\mu_0 = \sqrt{2}\,\text{GeV} \tag{41}$$

as initial scale with $n = 3$ active flavors for both partons. The evolution accuracy is independent of y, and we fix this variable to the value $y = b_0/\mu_0 \approx 0.794\,\text{GeV}^{-1}$. Unless specified otherwise, flavor matching is carried out at the canonical scales $\mu_q = m_q$, with the quark masses $m_c = \sqrt{2}\,\text{GeV}$, $m_b = 4.5\,\text{GeV}$, and $m_t = 175\,\text{GeV}$. The strong coupling at the initial scale is taken to be $\alpha_s(\mu_0) = 0.35$. Throughout this section, we consider unpolarized partons and use evolution and flavor matching at NNLO. We specify the scales and
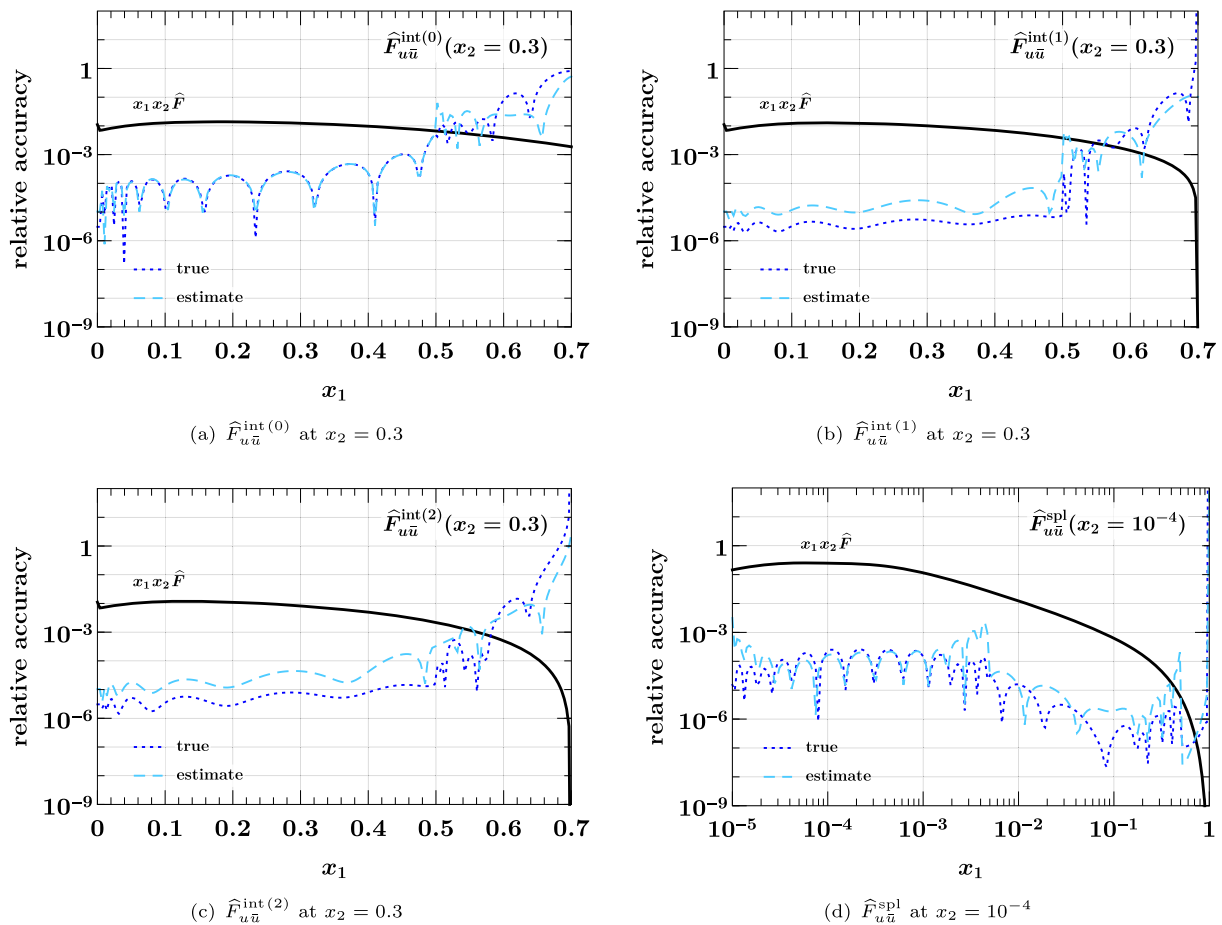
Fig. 7 True interpolation accuracy in x_1 and x_2 compared with its estimate (40). All values refer to the medium x grid in Eq. (37). The DPDs used are specified in Eqs. (35) and (36)

the number of active flavors for the two partons in the form (μ_1, μ_2) and (n_1, n_2).

Guided by the studies in Sect. 4, we interpolate both x_1 and x_2 on the grid

$$[10^{-5}, 5 \times 10^{-3}, 0.5, 1]_{(16,16,24)}. \qquad (42)$$

This corresponds to the medium grid in Eq. (37), except for the third subinterval, where we afford a finer grid for improved interpolation accuracy at large momentum fractions.

5.1 Path independence of evolution and flavor matching

For a given heavy quark flavor q, we apply the matching equations (4) at a fixed scale μ_q, so that in the (μ_1, μ_2) plane we match at the line (μ_q, μ_2) for the first parton and at the line (μ_1, μ_q) for the second one. We show in Appendix A that, under these conditions, the result of evolution and flavor matching between two points in the (μ_1, μ_2) plane is inde-

pendent of the evolution path, including the order in which the different flavor matching steps are carried out.

We verify this numerically by comparing the results of evolving and matching from (1 GeV, 1 GeV) with (3, 3) flavors to (500 GeV, 1000 GeV) with (6, 6) flavors along the two paths shown in Fig. 8. The path in black is the default taken in CHILIPDF when matching from 3 to 6 flavors for both partons and then evolving to the final pair of scales. The staircase-like path in red can be selected by subsequent evolution calls for each line segment. Each point marked by a dot in the figure corresponds to the initial condition of the DPD with given flavor numbers (n_1, n_2) in CHILIPDF. For any scale pair (μ_1, μ_2), the DPD with these flavor numbers is computed by evolution from this initial condition.

The results of evolving and matching along the two specified paths are in excellent agreement, as illustrated in Fig. 9 for the $t\bar{t}$ distribution, for which the differences are largest among all flavor combinations. The agreement deteriorates close to the kinematic limit $x_1 = 1 - x_2$ but remains small compared with the interpolation accuracy in that region. The
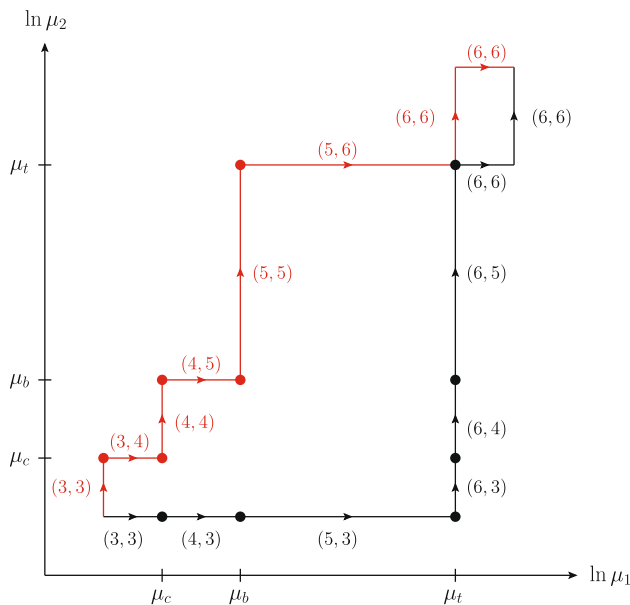
Fig. 8 Two paths for evolution and flavor matching in the (μ_1, μ_2) plane. Next to each line segment, we specify the numbers (n_1, n_2) of active flavors for which the DPD is evolved. Flavor matching is performed at the points marked by dots. The path in black is the CHILIPDF default for matching from $(3, 3)$ to $(6, 6)$ and for subsequent evolution up to the final scale pair

choice for a particular evolution path in our implementation is therefore of no consequence for the accuracy.

5.2 Evolution and flavor matching accuracy

We now study evolution and flavor matching to higher scales, using the initial conditions specified at the beginning of this section. To assess the accuracy, we use three settings:

1. our default setting with the grid $[10^{-5}, 5 \times 10^{-3}, 0.5, 1]_{(16,16,24)}$ and a maximal step size $h = 0.22$ for the Runge–Kutta algorithm,
2. a finer grid $[10^{-5}, 5 \times 10^{-3}, 0.5, 1]_{(40,40,40)}$ with the same maximal step size $h = 0.22$,
3. the same finer grid with a smaller maximal step size $h = 0.0088$.

The difference between settings 1 and 2 is taken as an estimate of the error from discretization in x_1 and x_2 with our default grid, and the difference between settings 2 and 3 is taken as an estimate of the error due to the Runge–Kutta algorithm.

These errors are shown in Fig. 10 for evolution and matching to (100 GeV, 10 TeV) with (5, 6) active flavors. The plots are for selected parton combinations; we verified that the errors for the other combinations are of similar size or even smaller. We see that the Runge–Kutta errors are always negligible compared with the errors from discretization. The latter are of similar size as (or even smaller than) the interpolation errors for the initial conditions, which are shown for $u\bar{u}$ in Fig. 6. Evolution to high scales does therefore *not* degrade the overall accuracy of DPDs with our method. Only close to the kinematic limit does the relative error from discretization become large. With the chosen grid, it remains below 1% for $1 - (x_1 + x_2)$ above 0.2.

In Fig. 11, we show the errors for evolution and matching to $(1.01 \mu_c, 1.01 \mu_b)$ with (4, 5) flavors, i.e. to scales just slightly above the matching scales for the final flavor numbers. The accuracy remains excellent, even for DPDs with charm as first or bottom as second parton, which are still tiny at the chosen scales.
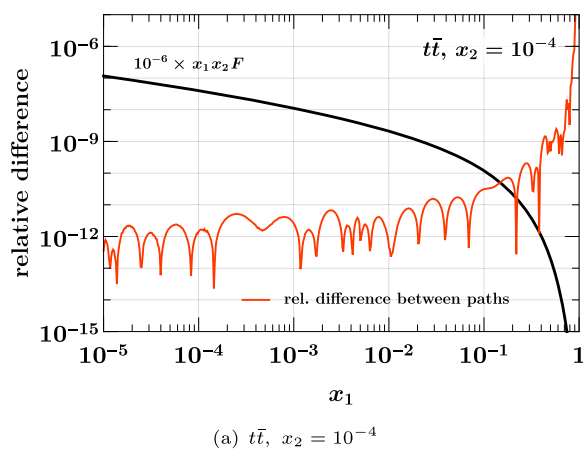
(a) $t\bar{t}, \ x_2 = 10^{-4}$
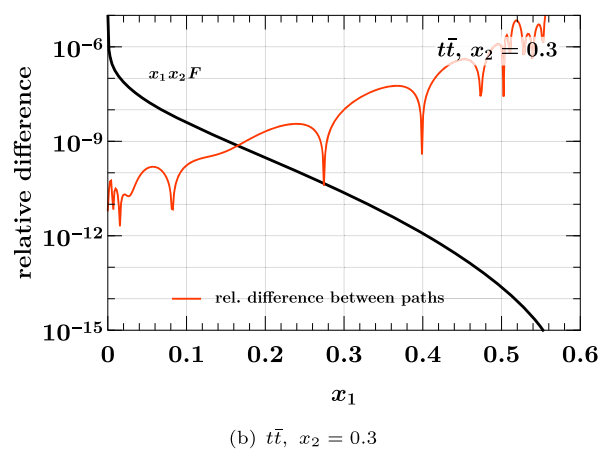
(b) $t\bar{t}, \ x_2 = 0.3$

Fig. 9 Relative difference between the $t\bar{t}$ distribution obtained by evolving and matching along one of the two paths in Fig. 8. Both paths lead from (1 GeV, 1 GeV) with (3, 3) flavors to (500 GeV, 1000 GeV) with (6, 6) flavors. The distributions at (1 GeV, 1 GeV) are obtained by evolving backwards from the starting conditions at (μ_0, μ_0) specified at the beginning of the section

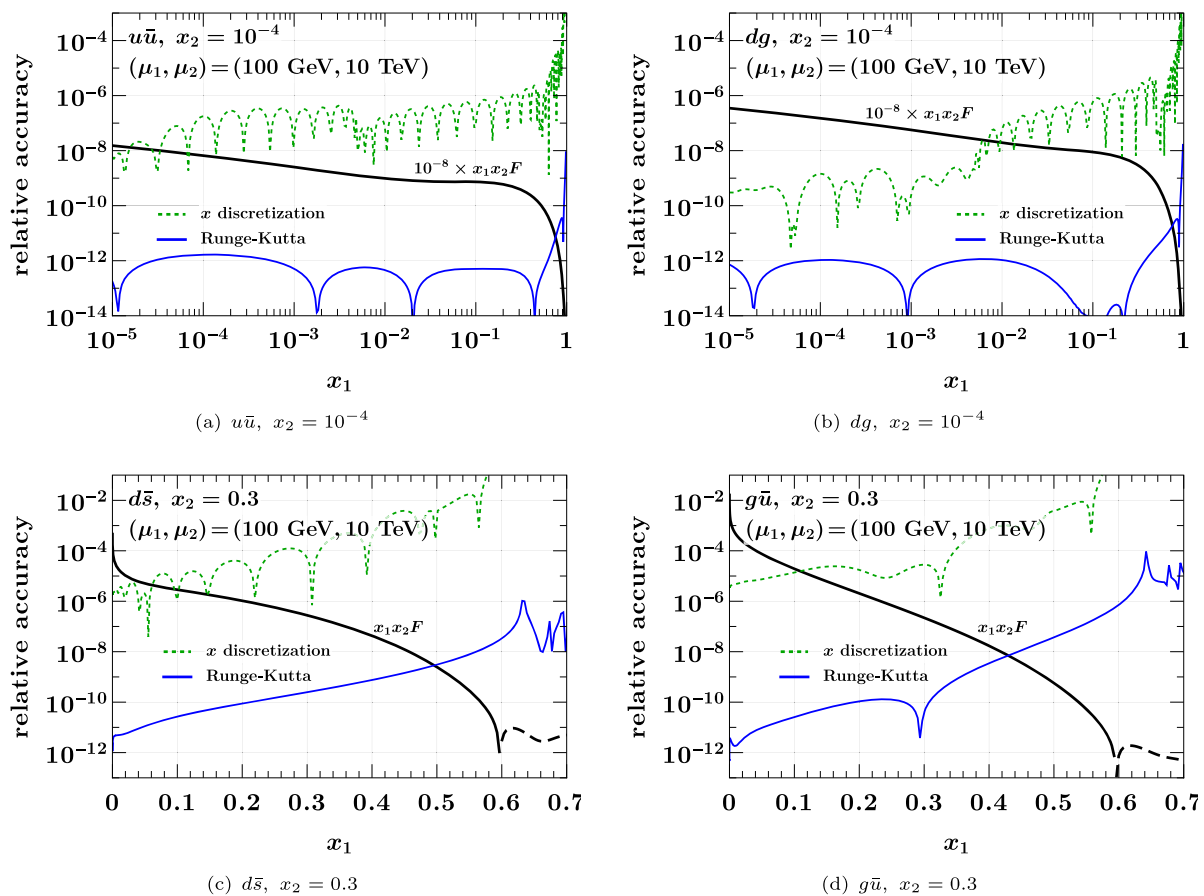**Fig. 10** Relative Runge–Kutta and discretization errors for DPD evolution and flavor matching from scales (μ_0, μ_0) with $(3, 3)$ flavors to $(100\,\text{GeV}, 10\,\text{TeV})$ with $(5, 6)$ flavors. The initial conditions are specified at the beginning of the section

5.3 Backward evolution accuracy

A stringent way to test the accuracy of an evolution algorithm is backward evolution from large to small scales. Whilst small inaccuracies of distributions at a given scale are diminished when evolving to higher scales, they are amplified by backward evolution. In the following, we adapt the backward evolution test performed for PDFs in Ref. [55] to the case of DPDs.

We start with the three-flavor input DPD specified at the beginning of this section, match to 4 flavors at the input scales (μ_0, μ_0), and then evolve to $(\mu_{\text{low}}, \mu_{\text{low}})$ with $\mu_{\text{low}} = m_b/2 = 2.25\,\text{GeV}$. At this point, we match from 4 to 5 flavors for both partons. Since the matching scale differs from m_b, all DPDs are nonzero at that point, including the ones with bottom quarks or antiquarks. We take these distributions as starting conditions for evolution with 5 flavors, evolving

1. from $(\mu_{\text{low}}, \mu_{\text{low}})$ to $(\mu_{\text{high}}, \mu_{\text{high}})$ with $\mu_{\text{high}} = 1\,\text{TeV}$ in step 1,
2. from $(\mu_{\text{high}}, \mu_{\text{high}})$ back to $(\mu_{\text{low}}, \mu_{\text{low}})$ in step 2.

We use our default value of $h = 0.22$ for the maximal Runge–Kutta step size.

The relative difference between the input to the first step and the output of the second step is sensitive to the accuracy of forward *and* backward evolution. In Fig. 12, this quantity is shown for selected parton combinations; the accuracy is not worse for other combinations. We find that backward evolution is very precise, with the relative errors in our exercise being below 10^{-6} over most of the phase space. This also holds for distributions containing bottom quarks or antiquarks, which have zero crossings at the scale μ_{low} where the ratio shown in the figure is evaluated.

We note that backward evolution is inherently more demanding for two scales than for a single scale. Indeed, step 2 of our study corresponds to backward evolution first from $(\mu_{\text{high}}, \mu_{\text{high}})$ to $(\mu_{\text{low}}, \mu_{\text{high}})$ and then from $(\mu_{\text{low}}, \mu_{\text{high}})$ to $(\mu_{\text{low}}, \mu_{\text{low}})$. Inaccuracies accumulated in the first substep are amplified in the second one. We repeated the above exercise for evolution in only one scale, evolving from $(\mu_{\text{low}}, \mu_{\text{low}})$ up to $(\mu_{\text{high}}, \mu_{\text{low}})$ and then down to $(\mu_{\text{low}}, \mu_{\text{low}})$. We find that indeed the accuracy is even higher in that case, by one or two orders of magnitude in some kinematic regions.
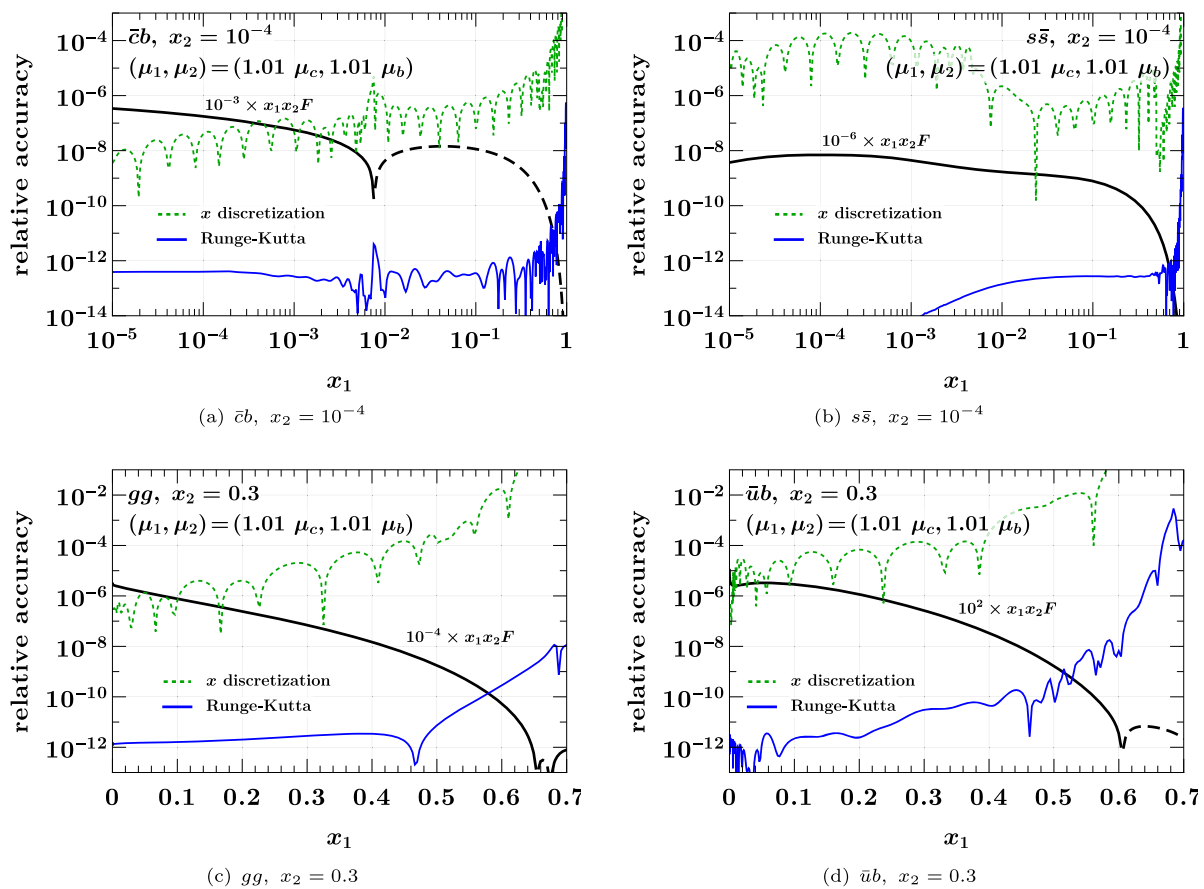
Fig. 11 As Fig. 10, but for different flavor combinations evolved and matched to scales $(1.01\,\mu_c, 1.01\,\mu_b)$ with $(4, 5)$ flavors. A dashed black line indicates that the interpolated function is negative
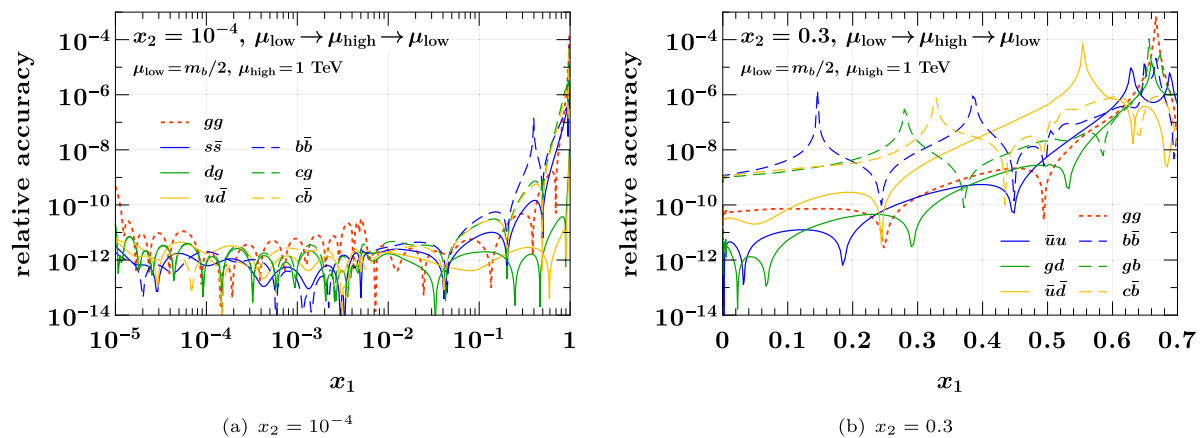

Fig. 12 Relative accuracy of evolving DPDs from $\mu_{\text{low}} = m_b/2$ to $\mu_{\text{high}} = 1$ TeV and back to μ_{low} for both partons. Details are given in the text

6 Interpolation of DPDs in the distance between partons

In the discussion so far, the transverse distance y between partons did not play an active role, given that evolution and flavor matching of DPDs are performed for each discretized value of y individually, without any cross talk between different values of y. Once a DPD is computed for selected scales

(μ_1, μ_2) and flavor numbers (n_1, n_2), two different tasks are concerned with the y dependence:

1. interpolation of the DPD for specific values of y, in order to study how evolution and matching have modified the y dependence of the initial conditions. In the study described in Ref. [81], visible effects of evolution

Table 1 Variable transformations suitable for interpolation in y. The parameters α and m must be positive. All transformation satisfy $u \leq 0$, $y(0) = \infty$, and $dy/du > 0$ for $u \in [-1, 0]$. We abbreviate $L(u) = \ln\big(1/|u|\big)$

No.	$-u(y)$	$y(u)$	$y(-1)$	dy/du
1	$y^{-\alpha}$	$|u|^{-1/\alpha}$	1	$\frac{1}{\alpha}\,|u|^{-(1+\alpha)/\alpha}$
2	$\exp\!\left[-\frac{1}{4}\left(m^2 y^2 + my\right)\right]$	$\frac{1}{2m}\left(\sqrt{16L(u)+1}-1\right)$	0	$\frac{4}{m}\,|u|^{-1}\left(16L(u)+1\right)^{-1/2}$
3	$\exp\!\left[-\frac{1}{4}\,my\right]$	$\frac{4}{m}\,L(u)$	0	$\frac{4}{m}\,|u|^{-1}$
4	$\exp\!\left[1-\sqrt{1+my/2}\right]$	$\frac{2}{m}\left(L^2(u)+2L(u)\right)$	0	$\frac{4}{m}\,|u|^{-1}\left(L(u)+1\right)$

on the large-y dependence of DPDs were reported, and in Ref. [24] it was found that evolution can substantially flatten the y^{-2} behavior generated by the perturbative splitting mechanism at small distances. Numerically reliable interpolation of DPDs over a wide range of y is hence important for understanding the dynamics of evolution.

2. integration over y from a minimum value $y_{\min}$ up to infinity. First and foremost, this is needed for computing the double parton luminosities (6) that enter in cross sections of double parton scattering processes. It is also needed for evaluating sum rules (see Sect. 7), which DPDs must satisfy and which provide valuable constraints for their modeling. Note that one integrates the product of two DPDs when computing double parton luminosities and single DPDs when evaluating the sum rules.

Both tasks can be carried out using Chebyshev grids, as described in Sect. 3.1. To this end, one first needs to find variable transformations that map the characteristic y dependence of DPDs to a dependence in the variable u that is well suited for interpolation by high-order polynomials. In the next two subsections, we show how this can be achieved. After that, we will demonstrate the resulting accuracy of interpolation and integration.

6.1 Characteristic y dependence and associated variable transformations

It is natural to distinguish the region of small y, where perturbative mechanisms are at work from the domain of large y.

As mentioned earlier and discussed at length in Ref. [24], DPDs at small y can be represented as the superposition of a part due to perturbative splitting, with a power behavior like y^{-2}, and an "intrinsic" part that behaves like y^0 and is related to twist-four distributions. Both power laws are modified by logarithms from loop corrections. The relative weight of the two parts strongly depends on the parton combination and on the momentum fractions x_1 and x_2 in the DPD. As already mentioned, the y^{-2} behavior can be substantially flattened by evolution to high scales. In the small-y region,

the transformation $u(y)$ should therefore be suited to handle the interpolation of a power behavior y^{-q} with $0 \leq q \leq 2$.

A complication arises from the fact that for certain parton combinations, such as uu or $u\bar{d}$, the leading-order splitting contribution (7) is zero. Quark-gluon mixing results in nonzero values for these combinations when evolving from the scale $\mu_y \sim 1/y$ at which the splitting formula is evaluated to the final scale μ. Hence, the splitting part of the DPD at scale μ has a zero around $y \sim 1/\mu$. The position of this zero is shifted when the intrinsic part is added.

At large distances, one expects a gradual decrease of DPDs governed by some characteristic nonperturbative mass scale. This is often modeled by a Gaussian, $F \sim \exp(-M^2 y^2)$, but one may also argue in favor of a Yukawa-like decrease, $F \sim \exp(-My)$. The characteristic mass scale M may depend on the parton combination and on x_1 and x_2 [41,81], and it can slowly change under evolution.

We now discuss variable transformations that we found to work well for the types of y dependence just discussed. An overview of the transformations and their properties is given in Tables 1 and 2. By convention, all transformations are such that $u(y)$ has a positive derivative in the domain where it is used. In the case of interpolation, the function $f(y)$ in Table 2 represents a DPD at fixed momentum fractions or scales. For sum rule integrals $f(y)$ is $2\pi y$ times a DPD (see Eq. (55)), and for double parton luminosities $f(y)$ is $2\pi y$ times the product of two DPDs (see Eq. (52)). The overall normalization of $f(y)$ is irrelevant for the interpolation or integration accuracy and therefore set to an arbitrary value here.

The last three transformations in the table are designed for functions with a decrease like $\exp(-M^2 y^2)$ or $\exp(-My)$ at large y, possibly modified by a power law in y. They map the physical interval $y \in [0, \infty]$ onto $u \in [-1, 0]$ and depend on a mass parameter m, which is normalized such that $dy/du = 4/m$ at $y = 0$. The integral

Table 2 Behavior of the variable transformations in Table 1 for selected functions $f(y)$. In the third column we used $\exp[-cL(u)] = |u|^c$. The condition in the last column ensures that $\mathrm{d}y/\mathrm{d}u\, f(y(u)) \to 0$ for $u \to 0$. It is understood that the parameters M and m are always positive

No.	$f(y)$	$f(y(u))$	$\int^\infty \mathrm{d}y\, f(y)$ requires
1	$y^{-\beta}$	$|u|^{\beta/\alpha}$	$\alpha < \beta - 1$
2	$\exp(-M^2 y^2)$	$|u|^{(2M/m)^2} \exp\left[\frac{M^2}{m} y(u)\right]$	$m < 2M$
3	$\exp(-M^2 y^2)$	$|u|^{L(u)\,(4M/m)^2}$	No condition
3	$\exp(-My)$	$|u|^{4M/m}$	$m < 4M$
4	$\exp(-My)$	$|u|^{L(u)\,2M/m}\, |u|^{4M/m}$	No condition

$$\int_{y_{\min}}^{\infty} \mathrm{d}y\ f(y) = \int_{u_{\min}}^{0} \mathrm{d}u\ \frac{\mathrm{d}y}{\mathrm{d}u}\ f(y(u)) \qquad (43)$$

is computed with the Clenshaw–Curtis quadrature rule, which includes the interval boundary $u = 0$ where the Jacobian $\mathrm{d}y/\mathrm{d}u$ is infinite. The upper limit on m specified in the last column of Table 2 ensures that the product $\mathrm{d}y/\mathrm{d}u\, f(y(u))$ is zero at the point $u = 0$, so that this point does not contribute to the Clenshaw–Curtis rule. We note that the value of the mass scale M in the integrand depends on whether one integrates a product of two DPDs or a single DPD.

- **Inverse power law:** Transformation no. 1 in the table is $u(y) = -y^{-\alpha}$. We will show that with $\alpha \sim 0.2$ this is well suited for the range of power laws $y^{-\beta}$ expected for DPDs at small distances. The more general form $u(y) = -(y + a)^{-\alpha}$ can be used down to $y = 0$ if $a > 0$, but we will not need this here.
- **Gaussian:** Transformation no. 2 in the table has a Gaussian falloff at large y and is designed for functions with a Gaussian behavior, $f(y) \sim \exp(-M^2 y^2)$. The transformed function $f(y(u))$ approximately behaves like a power of $|u|$ for $u \to 0$, and we find that good interpolation and integration accuracy is obtained for $m \sim M$.
- **Exponential:** Transformation no. 3 in the table is an exponential in y. It is suitable for both $f(y) \sim \exp(-My)$ and $f(y) \sim \exp(-M^2 y^2)$. In the first case, the transformed function $f(y(u))$ behaves like a power of $|u|$ for $u \to 0$, whereas in the second case it falls off faster than any power of $|u|$. In both cases, values $m \sim M$ yield good interpolation and integration accuracy.
- **Exponential with square root:** Transformation no. 4 in the table is designed for functions $f(y) \sim \exp(-My)$ with a Yukawa-type decrease. The transformed function $f(y(u))$ decreases faster than any power of $|u|$ for $u \to 0$, and good accuracy is obtained with $m \sim M$.

When handling DPDs with different flavor combinations and in different kinematic regions, the mass scale M of $f(y)$

in the preceding discussion does not have a unique value. M should instead be understood as an average, or in the case of the bounds in the last column of Table 2 as the lower limit of the relevant values.

6.2 Composite grids

Given the different behavior of DPDs at small and large y, it is natural to use interpolation grids with different transformations $u(y)$. At small y, there is an additional physics reason for using several subgrids. Not only the scale $\mu_y \sim 1/y$ but also the number n of active flavors for which the splitting DPD is naturally initialized depends on y. Since the value of a DPD in general changes with n, one should use a unique value of n in a subgrid, so as to avoid interpolating discontinuous functions. As discussed at length in Ref. [82], it is appropriate to evaluate the perturbative splitting formula (7) with n active flavors for $\gamma m_n < \mu_y < \gamma m_{n+1}$ with $\gamma \sim 1$. We therefore place subgrid boundaries in y at values around the inverse heavy quark masses.

In the following studies, we will assume a y dependence

$$F(y; \kappa) = y^{-2\kappa}\, \exp(-M^2 y^2), \qquad (44)$$

for a single DPD, with $\kappa = 0$ for the intrinsic and $\kappa = 1$ for the splitting part. We take $M = 0.21\,\mathrm{GeV}$, which roughly corresponds to the Gaussian parameter in quark-gluon DPDs of the model used in Refs. [24,30]. As for the function $f(y)$ in Table 2, the overall normalization of $F(y; \kappa)$ is irrelevant for the interpolation accuracy.

The smallest value of y needed in a given physics setting corresponds to the inverse of the largest hard scale that will be considered. In the following, we will take $y_{\min} = 1/(7\,\mathrm{TeV})$, which covers the production of very heavy states at the LHC. For the interpolation of Eq. (44) and for the computation of the associated double parton luminosities, we use the following combination of grids and transformations:

$$\big[1/(7\,\mathrm{TeV}),\ 1/m_b,\ 1/m_c\big]_{p_1,\,p_2}$$

inverse power law transf. with $\alpha = 0.2$, $\qquad (45)$

(a) $\exp(-M^2 y^2)$

(b) $y^{-2} \exp(-M^2 y^2)$

(c) $F(y; y_0)$ with $y_0 = 10^{-2}$
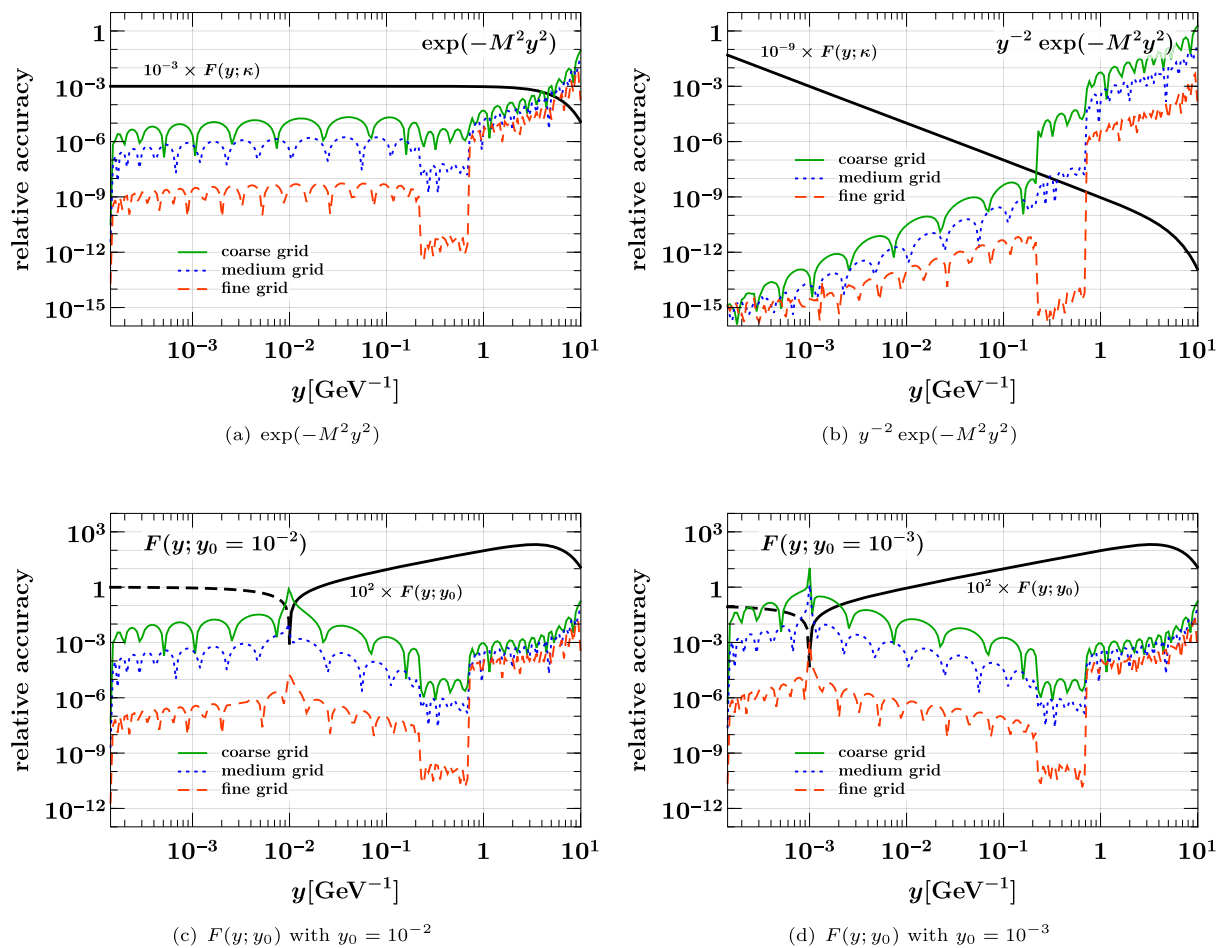

(d) $F(y; y_0)$ with $y_0 = 10^{-3}$

Fig. 13 Relative interpolation accuracy (38) on the grids specified in Eqs. (45), (46) and (47). The interpolated functions are $F(y; \kappa)$ in the upper panels and $F(y; y_0)$ in the lower ones, see Eqs. (44) and (48).

A dashed black line indicates that the interpolated function is negative. Here and in all following plots, the mass parameter in the functions is $M = 0.21\,\text{GeV}$

$$\big[1/m_c, \infty\big]_{p_3}$$

Gaussian transf. with $m = 2M$, (46)

where the "inverse power law" and the "Gaussian" transformation are respectively given in entries no. 1 and 2 of Table 1. We find that taking $m = 2M$ results in a somewhat better overall accuracy than the rule-of-thumb value $m \sim M$ given in the previous subsection. We will consider three different settings for the number of points in each subgrid:

$$(p_1, p_2, p_3) = \begin{cases} (12, 6, 12) & \text{(coarse } y \text{ grid)} \\ (16, 8, 16) & \text{(medium } y \text{ grid)} \\ (24, 12, 24) & \text{(fine } y \text{ grid).} \end{cases} \quad (47)$$

Notice the relatively small number of points in the central grid, which covers the narrow y region in which one would initialize the splitting part of the DPD with 4 active flavors. For the quark masses, we use the values specified at the

beginning of Sect. 5, which results in subgrid boundaries at $1/m_b \approx 0.22\,\text{GeV}^{-1}$ and $1/m_c \approx 0.71\,\text{GeV}^{-1}$.

6.3 Accuracy of interpolation in y

In the top panels of Fig. 13, we show the interpolation accuracy for the form (44) with $\kappa = 0$ and $\kappa = 1$. In the first and second subgrid, a very accurate interpolation is obtained even with the coarse grid setting. In the third subgrid, the accuracy degrades rather quickly for large distances, but for the medium grid it remains below 1% for y up to about $4\,\text{GeV}^{-1} \approx 0.8\,\text{fm}$. The fine grid yields an acceptable accuracy at even larger distances.

Before discussing this issue in more detail, we consider the interpolation of a DPD with a zero crossing in y, which arises for certain parton combinations as explained above. As an example for this case, we take the simple form

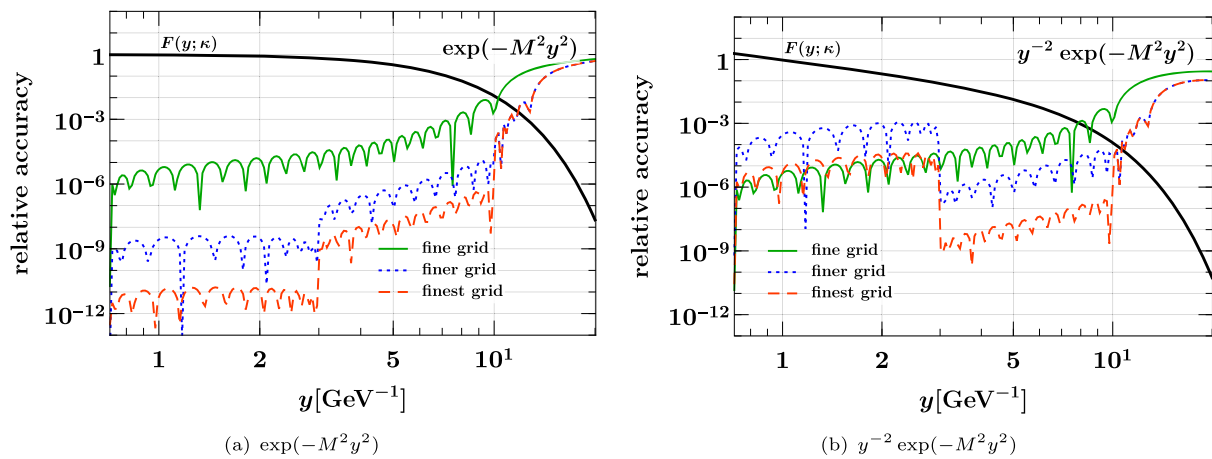$$F(y; y_0) = (y - y_0) \exp(-M^2 y^2), \quad (48)$$

Fig. 14 Relative interpolation accuracy at large y for the grid in Eq. (46) with $p_3 = 24$ points (fine grid) and for the composite grids specified by Eqs. (49), (50) and (51) (finer and finest grid). The interpolated functions are the same as in Fig. 13a and b, respectively

with different values of y_0 and the same mass parameter $M = 0.21$ GeV as before. The resulting accuracy is shown in the bottom panels of Fig. 13. We find that the coarse grid performs poorly, whereas with the medium grid the relative accuracy stays below 1% except in the immediate vicinity of the zero crossing. Very high accuracy is obtained with the fine grid.

We note that for the intermediate y grid from $1/m_b$ to $1/m_c$, the setting with $p_2 = 8$ grid points is sufficient to obtain a relative interpolation accuracy of 10^{-6} or better in all cases considered here. In each of the grids for smaller and at larger y, one may take either 16 or 24 points, depending on how stringent the accuracy requirements are. This gives 40, 48, or 56 for the total number of points.[4]

With the grids considered so far, the interpolation accuracy at very large y becomes poor mainly because there are too few grid points in that region. For many purposes, the behavior of the distributions in the extreme tail region is not of particular interest, and it contributes only little to integrals over y. Nevertheless, we wish to show that with an appropriate choice of subgrids and grid transformations, one can also accurately describe this region. To this end, we replace the large-y grid in Eq. (46) with

$$\left[1/m_c, 3\,\text{GeV}^{-1}, 10\,\text{GeV}^{-1}\right]_{p_a, p_b}$$

Gaussian transf. with $m = \sqrt{2}M$,   (49)

$$\left[10\,\text{GeV}^{-1}, \infty\right]_{p_c}$$

Gaussian transf. with $m = 2M$,   (50)

---

[4] Here we count the common end points of neighboring subgrids twice, since different subgrids may contain DPDs with different flavor numbers.

where we consider two settings for the number of points:

$$(p_a, p_b, p_c) = \begin{cases} (12, 12, 8) & \text{(finer } y \text{ grid)} \\ (16, 16, 8) & \text{(finest } y \text{ grid).} \end{cases}$$ (51)

In Fig. 14, the resulting interpolation accuracy is compared with the previous setting of Eq. (46) with $p_3 = 24$ points. We see that both the finer and the finest grid extend the domain of very accurate interpolation up to $y \approx 10\,\text{GeV}^{-1} \approx 2$ fm, with only a moderate increase in the number of grid points.

In analogy to Sect. 4.3, we can estimate the interpolation accuracy in y by comparing the results of interpolation with and without subinterval end points. We find that this method works very well, as shown by the examples in Fig. 15. The biggest discrepancies between true and estimated errors appear near subgrid boundaries, and in the central grid, where the omission of the end points decreases the polynomial order of interpolation from 11 to 9.

6.4 Accuracy of integration over y

Integration over y from a lower cutoff y_{min} to infinity is needed to evaluate the double parton luminosities (6) and DPD sum rules. If y_{min} coincides with the end point of a subinterval of the interpolation grid, one can directly use Clenshaw–Curtis quadrature to compute the integral. Otherwise, we perform a number of intermediate steps:

1. identify the subinterval $[y_a, y_b]$ that contains y_{min},
2. define a reduced subgrid on the interval $[y_{\text{min}}, y_b]$ with the same number of points and the same variable transformation as the original subgrid,
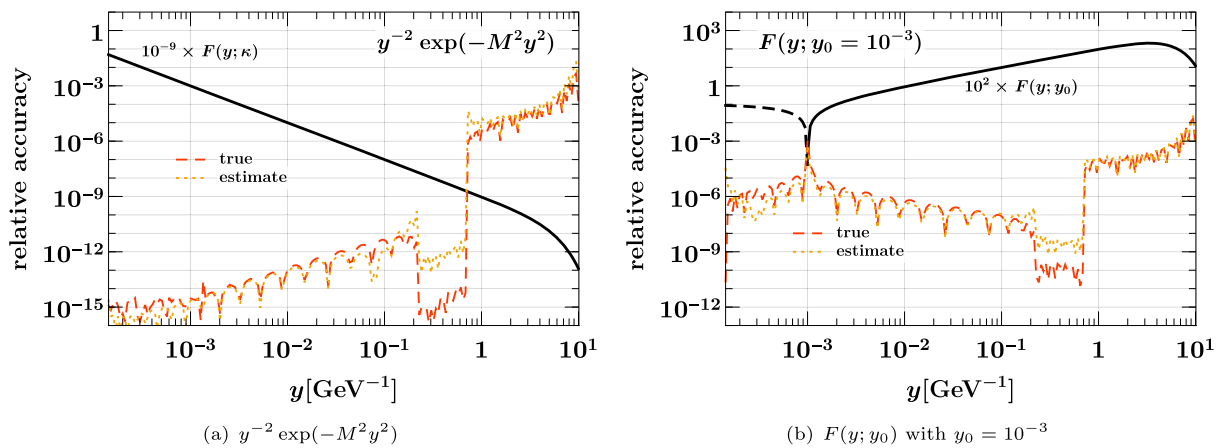3. compute the required DPDs on the reduced subgrid by interpolation on $[y_a, y_b]$.

Fig. 15 True and estimated relative interpolation accuracy in y, as defined in Eqs. (38) and 40, respectively. The curves are for the fine grid setting in Eq. (47)

The integral is then computed using Clenshaw–Curtis quadrature on the reduced subgrid and on the subgrids for $y \geq y_b$.

Let us investigate the accuracy of this procedure for the double parton luminosities

$$
\begin{aligned}
\mathcal{L}(y_{\min}, \kappa_1 + \kappa_2) &= 2\pi \int_{y_{\min}}^{\infty} \mathrm{d}y \; y \; F(y; \kappa_1) \, F(y; \kappa_2) \\
&= 2\pi \int_{y_{\min}}^{\infty} \frac{\mathrm{d}y \; y}{y^{2(\kappa_1 + \kappa_2)}} \, \exp(-2M^2 y^2)
\end{aligned} \tag{52}
$$

that correspond to the simplified shapes of DPDs in Eq. (44). We consider the cases $\kappa_1 + \kappa_2 = 0, 1, 2$, which respectively correspond to taking the intrinsic parts of both DPDs, the intrinsic part of one and the splitting part of the other DPD, or the splitting parts of both DPDs. Although only the sum of intrinsic and splitting parts enters in physical cross sections, it is useful to investigate these different combinations separately, as was for instance done in the studies in Refs. [24,82].

The integration accuracy can be computed using the exact expression of (52) in terms of an incomplete gamma function. The result is shown in the left panels of Fig. 16 for the interpolation grids introduced in Sect. 6.2. We see that for $y_{\min}$ below $1\,\mathrm{GeV}^{-1}$, good accuracy is achieved even with the coarse grid. For cutoff values above $1\,\mathrm{GeV}^{-1}$, the accuracy quickly becomes unsatisfactory unless the fine grid is used. This reflects the degrading interpolation accuracy in that region, which was shown in the upper panels of Fig. 13. In physics applications, $y_{\min}$ is of the order of an inverse hard scale, so that large values of the cutoff will only be needed in special circumstances, for instance if one is interested in the partial contribution of large y to an integral. We see that the fine grid setting is useful in such a case.

In the right panels of Fig. 16, we compare the true integration accuracy with its estimate obtained from comparing the values obtained with Clenshaw–Curtis quadrature and with Fejér's second rule (see Sect. 3.1). We see that for $y_{\min}$ below $1\,\mathrm{GeV}^{-1}$, one obtains an overestimate of the true error by one or two orders of magnitude. This is qualitatively similar to what we found for integrals of PDFs over x in section 3.4 of Ref. [55].

7 Cross check: comparison with DOVE

As a cross check for our implementation of DPD evolution in CHILIPDF, we performed a comparison with DOVE, a code that was introduced in Ref. [28] and further developed or adapted in subsequent work in Refs. [24,30,44,45,81].[5] The comparison is done with the version that was used in Ref. [30] to compute the Gaunt-Stirling sum rules for DPDs. In the following, we show the momentum sum rule for a u quark and the number sum rule for the parton combination $g d_v$, which read

$$
\sum_{a_2} \int_0^{1-x_1} \mathrm{d}x_2 \; x_2 \; F_{u a_2}^{\overline{\mathrm{MS}}}(x_1, x_2; \mu) = (1 - x_1) \, f_u(x_1; \mu) \tag{53}
$$

and

$$
\int_0^{1-x_1} \mathrm{d}x_2 \; F_{g d_v}^{\overline{\mathrm{MS}}}(x_1, x_2; \mu) = f_g(x_1; \mu), \tag{54}
$$

where $F_{g d_v} = F_{gd} - F_{g\bar{d}}$. The expressions for general parton combinations are given in Eqs. (3.1) and (3.2) of Ref. [28]. As shown in Ref. [83], the sum rules hold for the distributions

---

[5] Results obtained with an early version of DOVE are given on the website https://gsdpdf.hepforge.org.
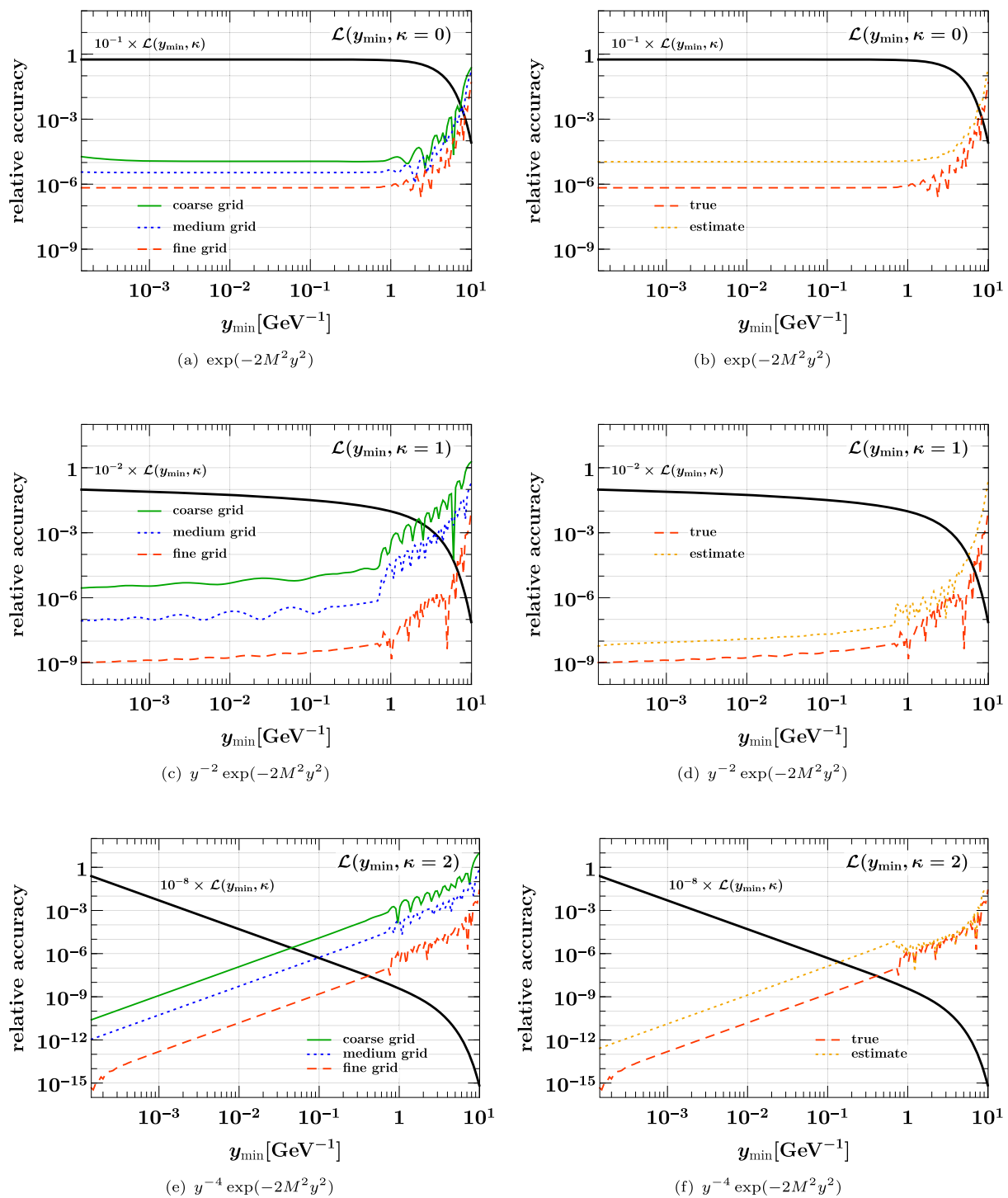
Fig. 16 Left: Relative accuracy of computing the integrals (52) with the interpolation grids specified by Eqs. (45), (46) and (47). Right: True and estimated integration accuracy for the fine grid setting of Eq. (47). The functions in the legends are the product $F(y; \kappa_1) F(y; \kappa_2)$ of DPDs in the luminosity integral

$$
\begin{aligned}
F_{a_1 a_2}^{\overline{\mathrm{MS}}}(x_1, x_2; \mu) &= 2\pi \int_{b_0/\nu}^{\infty} \mathrm{d}y \; y \; F_{a_1 a_2}(x_1, x_2, y; \mu, \mu) \\
&+ F_{a_1 a_2}^{\mathrm{match}}(x_1, x_2; \nu, \mu),
\end{aligned} \tag{55}
$$

which are integrated over y with an $\overline{\mathrm{MS}}$ renormalization prescription for the y^{-2} singularity due to perturbative splitting. The matching term $F_{a_1 a_2}^{\mathrm{match}}(x_1, x_2; \nu, \mu)$ relates this prescription with the cutoff in y used on the r.h.s. of Eq. (55). It can be

(a) u momentum sum rule

(b) gd_v number sum rule
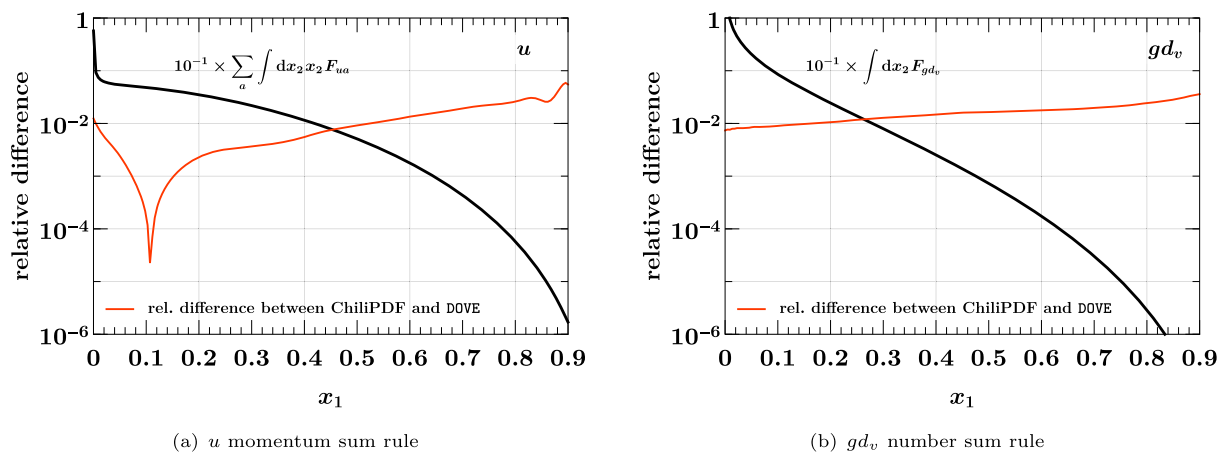
Fig. 17 Relative differences between the left-hand sides of the DPD sum rules (53) and (54), evaluated using DOVE and CHILIPDF. Details are given in the text

computed in terms of a PDF and a matching kernel in a similar fashion as the splitting part in Eq. (7), and its LO expression is given in Eq. 18 of Ref. [30]. The dependence on the matching scale ν cancels between the two terms in (55) up to power corrections in $1/\nu^2$. We recall that $b_0 = 2e^{-\gamma_E} \approx 1.12$.

The different roles played by x_1 and x_2 in the sum rules (53) and (54) motivate us to take different grids for the two momentum fractions,

$$x_1 : \quad \left[5 \times 10^{-5}, 5 \times 10^{-3}, 0.5, 0.9, 1.0\right]_{(12,12,16,8)},$$
$$x_2 : \quad \left[5 \times 10^{-5}, 5 \times 10^{-3}, 0.5, 1.0\right]_{(12,16,16)}. \quad (56)$$

The first grid has an additional small subgrid that leads to a better interpolation accuracy at large x_1, which is not needed for integration over x_2. The sum rule integrals over x_2 are computed using Clenshaw–Curtis quadrature over the full x_2 grid. This implies that the integrals are computed with a lower cutoff $x_{\min} = 5 \times 10^{-5}$. The same cutoff is used in the computation with DOVE.

We find that the integration over y in the sum rules can be computed with the rather small grids

$$\left[b_0/(170\,\text{GeV}), b_0/(1\,\text{GeV})\right]_{(12)}$$

inverse power law transf. with $\alpha = 0.2$, $\quad (57)$

$$\left[b_0/(1\,\text{GeV}), \infty\right]_{(8)}$$

Gaussian transf. with $m = M$, $\quad (58)$

where $M = 1/\sqrt{4h_{gg}} \approx 0.23\,\text{GeV}$ is the Gaussian mass parameter for gluons in the initial conditions of the model (see Eq. (8)).

We take the DPD model specified in section 3 of Ref. [30] and evolve to $\mu_1 = \mu_2 = 144.6\,\text{GeV}$ with $n = 3$ active flavors for each parton. At that scale, the momentum and number sum rules are evaluated with $\nu = 144.6\,\text{GeV}$. These μ and ν values correspond to grid points in our setup of DOVE and thus avoid additional interpolation. The detailed grid settings and integration procedure in DOVE are as specified in section 4 of Ref. [30].

In Fig. 17, we show the relative deviation between the left-hand sides of the sum rules (53) and (54) evaluated in CHILIPDF and in DOVE. The numerical differences increase with x_1 and reach several percent at $x_1 = 0.9$. For sum rules with other parton combinations, we find similar or smaller discrepancies. We also compared the sum rules evaluated at smaller values of $\mu_1 = \mu_2$ and ν and found somewhat smaller discrepancies than at high scales.

An error estimate for the accuracy of evolution with DOVE in its original version is given in figure 9 of [28]. Given that this estimate was carried out with finer grids than the ones used here, and given that relatively simple integration rules are used in the evaluation of the sum rules with DOVE, we find the level of agreement between the two codes satisfactory for the purpose of cross-validation.

8 Conclusions

We have shown that discretization on Chebyshev grids allows for a precise and economical representation of DPDs as functions of the momentum fractions x_1, x_2 and the distance y between the partons. This generalizes the methods we developed for PDFs in Ref. [55]. We use grids for x_1 that are independent of x_2 and vice versa. This gives the highest degree of simplicity and decoupling between operations that involve only one momentum fraction, which is the case for the Mellin

convolutions that appear in DGLAP evolution and in flavor matching. This method incurs interpolation errors close to the kinematic boundary $x_1 + x_2 = 1$ of the DPDs, which we find manageable. We investigated the interpolation accuracy for typical shapes of DPDs, both for the input distributions and after evolution. We find that grids with $p_x = 54$ points for $x \geq 10^{-5}$ yield a very accurate representation of DPDs, with relative errors that are much smaller than 10^{-3} in large parts of the phase space and remain below 1% for $x_1 + x_2 < 0.8$ (see Figs. 10 and 11).

For the dependence on y, Chebyshev interpolation is used after variable transformations that are adapted to the functional form and the typical mass scale for the decrease of the DPDs at large distance. We find that with $p_y \sim 48$ grid points, very good accuracy can be achieved for interpolation and integration.

To solve the DGLAP equations for DPDs, we use a high-order Runge–Kutta algorithm for computing the evolution matrices in Eqs. (33) and (34), which are the discretized version of the Green functions for the evolution equations. This provides very high numerical accuracy for both forward and backward evolution, with errors well below the ones due to discretization in x_1 and x_2. With our algorithm, it is natural to evolve DPDs separately in the scales μ_1 and μ_2 associated with the two partons. We note that the computational effort scales linearly with p_y but with the third power of p_x.

In our setup, a DPD with given flavor numbers (n_1, n_2) is specified numerically by its values at all grid points in (x_1, x_2, y) and at one pair (μ_1, μ_2) of scales. Its values at other scales are computed "on the fly" by solving the evolution equations. DPDs at other values of x_1, x_2, and y are computed using the barycentric interpolation formula, which is fast and numerically stable. As discussed in the introduction, we avoid grids in the renormalization scales, which would require a huge amount of computer memory. This is a notable difference between our approach and the widely used method to compute PDFs from grids in x and μ, for instance via the LHAPDF interface [54], and it reflects the notable difference in complexity between the distribution functions for one or for two partons inside a proton.

With our present implementation of the above methods in CHILIPDF, the evolution of a DPD with $n = 5$ active flavors (and thus 121 parton flavor combinations) from $\mu_1 = \mu_2 = 15 \, \text{GeV}$ to $150 \, \text{GeV}$ takes 1 to $2 \, \text{s}$[6] for $p_x = 54$ grid points in x_1 and x_2 and $p_y = 48$ grid points in y. With grids of this size, the largest part of computing time is spent on multiplying the evolution matrices with the initial conditions, so that there is only a weak dependence of the timing on the perturbative order of evolution and on the initial and final scales. Since the current code is based on a simple implementation of linear

algebra and matrix multiplication, we expect that significant gains can still be made with a dedicated performance tuning.

We used our code to investigate the impact of higher orders in the DGLAP kernels on the evolution of DPDs. For some parton combinations, the change from LO to NLO evolution produces $\mathcal{O}(1)$ effects at small momentum fractions, whereas we find only moderate differences when changing from NLO to NNLO kernels.

The numerical delivery of realistic DPDs has been a bottleneck in practice so far. The presented methods and their implementation provide a practical tool for computing evolved DPDs from user-given starting conditions, on a par with what has long been a standard for PDFs. This increases our ability to use the predictive power of perturbation theory and thus presents an important ingredient to making theory predictions for double parton scattering more realistic.

The evolution equations considered in this work apply to DPDs that are summed separately over the color of each parton. DPDs with color correlations between the two partons follow a different evolution pattern, with kernels that have recently been computed at NLO [84]. The methods described in the present paper can be adapted to this case, as will be reported in a future publication.

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors' comment: There is no associated data for this paper.]

---

[6] Timings were obtained on different computers with Intel® Core™ i5, Intel® Core™ i7, or AMD EPYC™ 7513 32-Core processors.

Appendix A: Path independence of evolution and flavor matching

In this appendix we show that the set (1) of DGLAP equations has a unique solution for a given initial condition at a point in the (μ_1, μ_2) plane. In other words, the solution is independent of the path in the (μ_1, μ_2) plane that is taken to evolve from the initial condition to the final scales. To establish this, it is enough to show that the second derivatives in the renormalisation scales commute, i.e. that

$$
\begin{aligned}
&\frac{\partial}{\partial L_1} \frac{\partial}{\partial L_2} F_{a_1 a_2}^{n_1, n_2}(L_1, L_2) \\
&= \frac{\partial}{\partial L_1} \left(\sum_{b_2} P_{a_2 b_2}^{n_2}(L_2) \underset{2}{\otimes} F_{a_1 b_2}^{n_1, n_2}(L_1, L_2) \right) \\
&= \sum_{b_2, b_1} P_{a_2 b_2}^{n_2}(L_2) \underset{2}{\otimes} P_{a_1 b_1}^{n_1}(L_1) \underset{1}{\otimes} F_{b_1 b_2}^{n_1, n_2}(L_1, L_2)
\end{aligned}
\tag{59}
$$

is equal to

$$
\begin{aligned}
&\frac{\partial}{\partial L_2} \frac{\partial}{\partial L_1} F_{a_1 a_2}^{n_1, n_2}(L_1, L_2) \\
&= \frac{\partial}{\partial L_2} \left(\sum_{b_1} P_{a_1 b_1}^{n_1}(L_1) \underset{1}{\otimes} F_{b_1 a_2}^{n_1, n_2}(L_1, L_2) \right) \\
&= \sum_{b_1, b_2} P_{a_1 b_1}^{n_1}(L_1) \underset{1}{\otimes} P_{a_2 b_2}^{n_2}(L_2) \underset{2}{\otimes} F_{b_1 b_2}^{n_1, n_2}(L_1, L_2).
\end{aligned}
\tag{60}
$$

Here and in the following, we change evolution variables from μ_i to

$$
L_i = 2 \ln \mu_i
\tag{61}
$$

and drop the arguments x_1, x_2, and y of the DPDs for brevity. For evolution at fixed (n_1, n_2), i.e. without flavor matching, the desired equality follows from the fact that the convolution integrals w.r.t. the first and second momentum fraction of the DPD commute with each other. This can readily be shown from their definition (2). Note that it is essential for the argument that the evolution kernel for one parton does not depend on the scale of the other parton.

We now extend this argument, showing that one obtains the same result when flavor matching along different paths in the (n_1, n_2) plane. Let us see how Eqs. (59) and (60) have to be modified to include the effects from flavor matching. We first consider the case of PDFs, where the notation is less involved. For a set of given matching scales $\tilde{\mu}_n$ for the transition from $n - 1$ to n active flavors, we define the "canonical"

number of flavors at given $L = 2 \ln \mu$ as

$$
n(L) = \begin{cases} 3 & \text{for } L < \widetilde{L}_4 \\ 4 & \text{for } \widetilde{L}_4 < L < \widetilde{L}_5 \\ 5 & \text{for } \widetilde{L}_5 < L < \widetilde{L}_6 \\ 6 & \text{for } L > \widetilde{L}_6 \end{cases}
\tag{62}
$$

where $\widetilde{L}_n = 2 \ln \tilde{\mu}_n$. We further define PDFs

$$
f_a(L) = f_a^{n(L)}(L),
\tag{63}
$$

with the canonical number of flavors at a given renormalisation scale. From these distributions, one can obtain $f_a^n(L)$ for flavor numbers different from the canonical values by evolution.[7]

At the matching scales, the distributions in Eq. (63) can have discontinuities, which complicates their description via differential equations. We therefore introduce a smoothened version of $f_a(L)$. Taking a single matching step from $n - 1$ to n flavors at $\widetilde{L} = 2 \ln \tilde{\mu}$ for simplicity, one can define a smooth approximation $\theta_\epsilon(L)$ of the step function that takes the values

$$
\theta_\epsilon(t) = \begin{cases} 0 & \text{for } t < -\epsilon \\ 1 & \text{for } t > \epsilon \end{cases}
\tag{64}
$$

and is continuous and monotonic in the intermediate region $-\epsilon < t < \epsilon$. Its derivative $\delta_\epsilon(t) = \mathrm{d}\theta_\epsilon(t)/\mathrm{d}t$ is then a smooth approximation of the Dirac distribution. Define now a PDF $f_a^\epsilon(L)$ that evolves according to

$$
\begin{aligned}
\frac{\mathrm{d} f_a^\epsilon(L)}{\mathrm{d}L} &= \sum_b P_{ab}^{n_\epsilon(L)}(L) \otimes f_b^\epsilon(L) \\
&+ \delta_\epsilon(L - \widetilde{L}) \left[\sum_b A_{ab}^n(m; \widetilde{L}) \otimes f_b^\epsilon(\widetilde{L} - \epsilon) - f_a^\epsilon(\widetilde{L} - \epsilon) \right],
\end{aligned}
\tag{65}
$$

with

$$
n_\epsilon(L) = n - 1 + \theta_\epsilon(L - \widetilde{L})
\tag{66}
$$

interpolating smoothly between $n - 1$ and n. The DGLAP kernels for the evolution of $f_a^\epsilon(L)$ are thus evaluated with flavor number $n - 1$ for $L \leq \widetilde{L} - \epsilon$, with flavor number n for $L \geq \widetilde{L} + \epsilon$, and with fractional flavor number in the

---

[7] This implies that our argument is *not* tied to the use of a specific "variable flavor number scheme".

intermediate L region.[8] One furthermore has

$$
\begin{aligned}
f_a^\epsilon(\widetilde{L} + \epsilon) &= f_a^\epsilon(\widetilde{L} - \epsilon) + \int_{\widetilde{L}-\epsilon}^{\widetilde{L}+\epsilon} dL \, \frac{df_a^\epsilon(L)}{dL} \\
&= f_a^\epsilon(\widetilde{L} - \epsilon) + \mathcal{O}(\epsilon) + \int_{\widetilde{L}-\epsilon}^{\widetilde{L}+\epsilon} dL \, \delta_\epsilon(L - \widetilde{L}) \\
&\quad \times \left[\sum_b A_{ab}^n(m; \widetilde{L}) \otimes f_b^\epsilon(\widetilde{L} - \epsilon) - f_a^\epsilon(\widetilde{L} - \epsilon) \right] \\
&= \sum_b A_{ab}^n(m; \widetilde{L}) \otimes f_b^\epsilon(\widetilde{L} - \epsilon) + \mathcal{O}(\epsilon),
\end{aligned}
\tag{67}
$$

where the $\mathcal{O}(\epsilon)$ contribution comes from the integral over the term with $P^{n_\epsilon(L)}$ in Eq. (65). In the limit $\epsilon \to 0$, this tends to the flavor matching prescription given in Eq. (5).

It is straightforward to extend this construction from one to several matching steps and from PDFs to DPDs. In analogy to Eq. (63), DPDs with canonical flavor numbers are defined as

$$
F_{a_1 a_2}(L_1, L_2) = F_{a_1 a_2}^{n(L_1), \, n(L_2)}(L_1, L_2),
\tag{68}
$$

and their smoothened version $F_{a_1 a_2}^\epsilon(L_1, L_2)$ evolves according to

$$
\begin{aligned}
&\frac{\partial}{\partial L_1} F_{a_1 a_2}^\epsilon(L_1, L_2) \\
&= \sum_{b_1} P_{a_1 b_1}^{n_\epsilon(L_1)}(L_1) \underset{1}{\otimes} F_{b_1 a_2}^\epsilon(L_1, L_2) \\
&\quad + \sum_n \delta_\epsilon(L_1 - \widetilde{L}_n) \sum_{b_1} \Delta A_{a_1 b_1}^n \underset{1}{\otimes} F_{b_1 a_2}^\epsilon(\widetilde{L}_n - \epsilon, L_2), \\
&\frac{\partial}{\partial L_2} F_{a_1 a_2}^\epsilon(L_1, L_2) \\
&= \sum_{b_2} P_{a_2 b_2}^{n_\epsilon(L_2)}(L_2) \underset{2}{\otimes} F_{a_1 b_2}^\epsilon(L_1, L_2) \\
&\quad + \sum_n \delta_\epsilon(L_2 - \widetilde{L}_n) \sum_{b_2} \Delta A_{a_2 b_2}^n \underset{2}{\otimes} F_{a_1 b_2}^\epsilon(L_1, \widetilde{L}_n - \epsilon),
\end{aligned}
\tag{69}
$$

where we abbreviated

$$
\Delta A_{ab}^n(z) = A_{ab}^n(z, m_n; \widetilde{L}_n) - \delta_{ab} \, \delta(1 - z).
\tag{70}
$$

As at the beginning of this section, one can then show that

$$
\begin{aligned}
&\frac{\partial}{\partial L_1} \frac{\partial}{\partial L_2} F_{a_1 a_2}^\epsilon(L_1, L_2) \\
&= \frac{\partial}{\partial L_2} \frac{\partial}{\partial L_1} F_{a_1 a_2}^\epsilon(L_1, L_2),
\end{aligned}
\tag{71}
$$

so that with an initial condition given in the (L_1, L_2) plane, Eq. (69) has a unique solution independent of the evolution path. Taking the limit $\epsilon \to 0$, the desired result is obtained.

---

8 PDFs and DGLAP kernels for inactive flavors are set to zero, so that the range of flavor sums does not explicitly depend on n.

References

1. CDF Collaboration, F. Abe et al., Double parton scattering in $\bar{p}p$ collisions at $\sqrt{s}$ = 1.8 TeV. Phys. Rev. D **56**, 3811 (1997). https://doi.org/10.1103/PhysRevD.56.3811
2. D0 Collaboration, V.M. Abazov et al., Study of double parton interactions in diphoton + dijet events in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV. Phys. Rev. D **93**, 052008 (2016). https://doi.org/10.1103/PhysRevD.93.052008. arXiv:1512.05291
3. LHCb Collaboration, R. Aaij et al., Measurement of the J/ψ pair production cross-section in pp collisions at $\sqrt{s}$ = 13 TeV. JHEP **06**, 047 (2017). https://doi.org/10.1007/JHEP06(2017)047. arXiv:1612.07451
4. ATLAS Collaboration, Measurement of J/ψ production in association with a $W^\pm$ boson with pp data at 8 TeV. JHEP **01**, 095 (2020). https://doi.org/10.1007/JHEP01(2020)095. arXiv:1909.13626
5. CMS Collaboration, Evidence for WW production from double-parton interactions in proton–proton collisions at $\sqrt{s}$ = 13 TeV. Eur. Phys. J. C **80**, 41 (2020). https://doi.org/10.1140/epjc/s10052-019-7541-6. arXiv:1909.06265
6. CMS Collaboration, Measurement of double-parton scattering in inclusive production of four jets with low transverse momentum in proton-proton collisions at $\sqrt{s}$ = 13 TeV. JHEP **01**, 177 (2022). https://doi.org/10.1007/JHEP01(2022)177. arXiv:2109.13822
7. CMS Collaboration, Observation of same-sign WW production from double parton scattering in proton–proton collisions at $\sqrt{s}$ = 13 TeV. arXiv:2206.02681
8. A. Kulesza, W. Stirling, Like sign W boson production at the LHC as a probe of double parton scattering. Phys. Lett. B **475**, 168 (2000). https://doi.org/10.1016/S0370-2693(99)01512-9. arXiv:hep-ph/9912232
9. J.R. Gaunt, C.-H. Kom, A. Kulesza, W. Stirling, Same-sign W pair production as a probe of double parton scattering at the LHC. Eur. Phys. J. C **69**, 53 (2010). https://doi.org/10.1140/epjc/s10052-010-1362-y. arXiv:1003.3953
10. K. Golec-Biernat, E. Lewandowska, Electroweak boson production in double parton scattering. Phys. Rev. D **90**, 094032 (2014). https://doi.org/10.1103/PhysRevD.90.094032. arXiv:1407.4038
11. F.A. Ceccopieri, M. Rinaldi, S. Scopetta, Parton correlations in same-sign W pair production via double parton scattering at the LHC. Phys. Rev. D **95**, 114030 (2017). https://doi.org/10.1103/PhysRevD.95.114030. arXiv:1702.05363
12. S. Cotogno, T. Kasemets, M. Myska, Spin on same-sign W-boson pair production. Phys. Rev. D **100**, 011503 (2019). https://doi.org/10.1103/PhysRevD.100.011503. arXiv:1809.09024
13. S. Cotogno, T. Kasemets, M. Myska, Confronting same-sign W-boson production with parton correlations. JHEP **10**, 214 (2020). https://doi.org/10.1007/JHEP10(2020)214. arXiv:2003.03347
14. CMS Collaboration, Search for new physics in same-sign dilepton events in proton–proton collisions at $\sqrt{s}$ = 13 TeV. Eur. Phys. J. C **76**, 439 (2016). https://doi.org/10.1140/epjc/s10052-016-4261-z. arXiv:1605.03171
15. CMS Collaboration, Search for top quark partners with charge 5/3 in the same-sign dilepton and single-lepton final states in proton–proton collisions at $\sqrt{s}$ = 13 TeV. JHEP **03**, 082 (2019). https://doi.org/10.1007/JHEP03(2019)082. arXiv:1810.03188
16. B. Blok, Y. Dokshitzer, L. Frankfurt, M. Strikman, The four jet production at LHC and Tevatron in QCD. Phys. Rev. D

83, 071501 (2011). https://doi.org/10.1103/PhysRevD.83.071501. arXiv:1009.2714

17. J.R. Gaunt, W. Stirling, Double parton scattering singularity in one-loop integrals. JHEP **06**, 048 (2011). https://doi.org/10.1007/JHEP06(2011)048. arXiv:1103.1888

18. M.G. Ryskin, A.M. Snigirev, A fresh look at double parton scattering. Phys. Rev. D **83**, 114047 (2011). https://doi.org/10.1103/PhysRevD.83.114047. arXiv:1103.3495

19. M. Diehl, D. Ostermeier, A. Schäfer, Elements of a theory for multiparton interactions in QCD. JHEP **03**, 089 (2012). https://doi.org/10.1007/JHEP03(2012)089. arXiv:1111.0910

20. A.V. Manohar, W.J. Waalewijn, A QCD analysis of double parton scattering: color correlations, interference effects and evolution. Phys. Rev. D **85**, 114009 (2012). https://doi.org/10.1103/PhysRevD.85.114009. arXiv:1202.3794

21. M.G. Ryskin, A.M. Snigirev, Double parton scattering in double logarithm approximation of perturbative QCD. Phys. Rev. D **86**, 014018 (2012). https://doi.org/10.1103/PhysRevD.86.014018. arXiv:1203.2330

22. J.R. Gaunt, Single perturbative splitting diagrams in double parton scattering. JHEP **01**, 042 (2013). https://doi.org/10.1007/JHEP01(2013)042. arXiv:1207.0480

23. B. Blok, Y. Dokshitzer, L. Frankfurt, M. Strikman, Perturbative QCD correlations in multi-parton collisions. Eur. Phys. J. C **74**, 2926 (2014). https://doi.org/10.1140/epjc/s10052-014-2926-z. arXiv:1306.3763

24. M. Diehl, J.R. Gaunt, K. Schönwald, Double hard scattering without double counting. JHEP **06**, 083 (2017). https://doi.org/10.1007/JHEP06(2017)083. arXiv:1702.06486

25. O. Fedkevych, A. Kulesza, Double parton scattering in four-jet production in proton–proton collisions at the LHC. Phys. Rev. D **104**, 054021 (2021). https://doi.org/10.1103/PhysRevD.104.054021. arXiv:2008.08347

26. P. Bartalini, J.R. Gaunt (eds.), *Multiple Parton Interactions at the LHC*, vol. 29. (World Scientific Publishing, Singapore, 2019). https://doi.org/10.1142/1064610.1142/10646

27. T. Kasemets, S. Scopetta, Parton correlations in double parton scattering. Adv. Ser. Direct High Energy Phys. **29**, 49 (2018). https://doi.org/10.1142/9789813227767_0004. arXiv:1712.02884

28. J.R. Gaunt, W.J. Stirling, Double parton distributions incorporating perturbative QCD evolution and momentum and quark number sum rules. JHEP **03**, 005 (2010). https://doi.org/10.1007/JHEP03(2010)005. arXiv:0910.4347

29. K. Golec-Biernat, E. Lewandowska, M. Serino, Z. Snyder, A.M. Stasto, Constraining the double gluon distribution by the single gluon distribution. Phys. Lett. B **750**, 559 (2015). https://doi.org/10.1016/j.physletb.2015.09.067. arXiv:1507.08583

30. M. Diehl, J.R. Gaunt, D.M. Lang, P. Plößl, A. Schäfer, Sum rule improved double parton distributions in position space. Eur. Phys. J. C **80**, 468 (2020). https://doi.org/10.1140/epjc/s10052-020-8038-z. arXiv:2001.10428

31. K. Golec-Biernat, A.M. Staśto, The momentum sum rule and factorization of double parton distributions. Phys. Rev. D **107**(5), 054020 (2023). https://doi.org/10.1103/PhysRevD.107.054020. arXiv:2212.02289

32. G.S. Bali, M. Diehl, B. Gläßle, A. Schäfer, C. Zimmermann, Double parton distributions in the nucleon from lattice QCD. JHEP **09**, 106 (2021). https://doi.org/10.1007/JHEP09(2021)106. arXiv:2106.03451

33. H.-M. Chang, A.V. Manohar, W.J. Waalewijn, Double parton correlations in the bag model. Phys. Rev. D **87**, 034009 (2013). https://doi.org/10.1103/PhysRevD.87.034009. arXiv:1211.3132

34. M. Rinaldi, S. Scopetta, V. Vento, Double parton correlations in constituent quark models. Phys. Rev. D **87**, 114021 (2013). https://doi.org/10.1103/PhysRevD.87.114021. arXiv:1302.6462

35. W. Broniowski, E. Ruiz Arriola, Valence double parton distributions of the nucleon in a simple model. Few Body Syst. **55**, 381 (2014). https://doi.org/10.1007/s00601-014-0840-4. arXiv:1310.8419

36. M. Rinaldi, S. Scopetta, M. Traini, V. Vento, Double parton correlations and constituent quark models: a Light Front approach to the valence sector. JHEP **12**, 028 (2014). https://doi.org/10.1007/JHEP12(2014)028. arXiv:1409.1500

37. W. Broniowski, E. Ruiz Arriola, K. Golec-Biernat, Generalized valon model for double parton distributions. Few Body Syst. **57**, 405 (2016). https://doi.org/10.1007/s00601-016-1087-z. arXiv:1602.00254

38. T. Kasemets, A. Mukherjee, Quark-gluon double parton distributions in the light-front dressed quark model. Phys. Rev. D **94**, 074029 (2016). https://doi.org/10.1103/PhysRevD.94.074029. arXiv:1606.05686

39. M. Rinaldi, S. Scopetta, M.C. Traini, V. Vento, Correlations in double parton distributions: perturbative and non-perturbative effects. JHEP **10**, 063 (2016). https://doi.org/10.1007/JHEP10(2016)063. arXiv:1608.02521

40. M. Rinaldi, F.A. Ceccopieri, Relativistic effects in model calculations of double parton distribution function. Phys. Rev. D **95**, 034040 (2017). https://doi.org/10.1103/PhysRevD.95.034040. arXiv:1611.04793

41. R. Corke, T. Sjostrand, Multiparton interactions with an x-dependent proton size. JHEP **05**, 009 (2011). https://doi.org/10.1007/JHEP05(2011)009. arXiv:1101.5953

42. B. Blok, P. Gunnellini, Dynamical approach to MPI in W+dijet and Z+dijet production within the PYTHIA event generator. Eur. Phys. J. C **76**, 202 (2016). https://doi.org/10.1140/epjc/s10052-016-4035-7. arXiv:1510.07436

43. B. Blok, P. Gunnellini, Dynamical approach to MPI four-jet production in Pythia. Eur. Phys. J. C **75**, 282 (2015). https://doi.org/10.1140/epjc/s10052-015-3520-8. arXiv:1503.08246

44. B. Cabouat, J.R. Gaunt, K. Ostrolenk, A Monte-Carlo simulation of double parton scattering. JHEP **11**, 061 (2019). https://doi.org/10.1007/JHEP11(2019)061. arXiv:1906.04669

45. B. Cabouat, J.R. Gaunt, Combining single and double parton scatterings in a parton shower. JHEP **10**, 012 (2020). https://doi.org/10.1007/JHEP10(2020)012. arXiv:2008.01442

46. A. Cafarella, C. Coriano, M. Guzzi, Precision studies of the NNLO DGLAP evolution at the LHC with CANDIA. Comput. Phys. Commun. **179**, 665 (2008). https://doi.org/10.1016/j.cpc.2008.06.004. arXiv:0803.0462

47. 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

48. 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

49. V. Bertone, S. Carrazza, J. Rojo, 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

50. 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

51. S. Weinzierl, Fast evolution of parton distributions. Comput. Phys. Commun. **148**, 314 (2002). https://doi.org/10.1016/S0010-4655(02)00584-2. arXiv:hep-ph/0203112

52. A. Vogt, Efficient evolution of unpolarized and polarized parton distributions with QCD-PEGASUS. Comput. Phys. Commun. **170**, 65 (2005). https://doi.org/10.1016/j.cpc.2005.03.103. arXiv:hep-ph/0408244

53. 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

54. 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

55. M. Diehl, R. Nagar, F.J. Tackmann, ChiliPDF: Chebyshev interpolation for parton distributions. Eur. Phys. J. C **82**, 257 (2022). https://doi.org/10.1140/epjc/s10052-022-10223-1. arXiv:2112.09703

56. M. Diehl, T. Kasemets, Positivity bounds on double parton distributions. JHEP **05**, 150 (2013). https://doi.org/10.1007/JHEP05(2013)150. arXiv:1303.0842

57. R. Kirschner, Generalized Lipatov–Altarelli–Parisi equations and jet calculus rules. Phys. Lett. B **84**, 266 (1979). https://doi.org/10.1016/0370-2693(79)90300-9

58. V.P. Shelest, A.M. Snigirev, G.M. Zinovev, The multiparton distribution equations in QCD. Phys. Lett. B **113**, 325 (1982). https://doi.org/10.1016/0370-2693(82)90049-1

59. A.M. Snigirev, Double parton distributions in the leading logarithm approximation of perturbative QCD. Phys. Rev. D **68**, 114012 (2003). https://doi.org/10.1103/PhysRevD.68.114012. arXiv:hep-ph/0304172

60. F.A. Ceccopieri, An update on the evolution of double parton distributions. Phys. Lett. B **697**, 482 (2011). https://doi.org/10.1016/j.physletb.2011.02.047. arXiv:1011.6586

61. 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

62. S. Moch, J.A.M. Vermaseren, A. Vogt, The three loop splitting functions in QCD: the nonsinglet case. Nucl. Phys. B **688**, 101 (2004). https://doi.org/10.1016/j.nuclphysb.2004.03.030. arXiv:hep-ph/0403192

63. A. Vogt, S. Moch, J.A.M. Vermaseren, The three-loop splitting functions in QCD: the singlet case. Nucl. Phys. B **691**, 129 (2004). https://doi.org/10.1016/j.nuclphysb.2004.04.024. arXiv:hep-ph/0404111

64. S. Moch, J.A.M. Vermaseren, A. Vogt, The three-loop splitting functions in QCD: the helicity-dependent case. Nucl. Phys. B **889**, 351 (2014). https://doi.org/10.1016/j.nuclphysb.2014.10.016. arXiv:1409.5131

65. S. Moch, J.A.M. Vermaseren, A. Vogt, On γ_5 in higher-order QCD calculations and the NNLO evolution of the polarized valence distribution. Phys. Lett. B **748**, 432 (2015). https://doi.org/10.1016/j.physletb.2015.07.027. arXiv:1506.04517

66. J. Blümlein, P. Marquard, C. Schneider, K. Schönwald, The three-loop unpolarized and polarized non-singlet anomalous dimensions from off shell operator matrix elements. Nucl. Phys. B **971**, 115542 (2021). https://doi.org/10.1016/j.nuclphysb.2021.115542. arXiv:2107.06267

67. J. Blümlein, P. Marquard, C. Schneider, K. Schönwald, The three-loop polarized singlet anomalous dimensions from off-shell operator matrix elements. JHEP **01**, 193 (2022). https://doi.org/10.1007/JHEP01(2022)193. arXiv:2111.12401

68. W. Vogelsang, Next-to-leading order evolution of transversity distributions and Soffer's inequality. Phys. Rev. D **57**, 1886 (1998). https://doi.org/10.1103/PhysRevD.57.1886. arXiv:hep-ph/9706511

69. W. Vogelsang, Q^2 evolution of spin dependent parton densities. Acta Phys. Pol. B **29**, 1189 (1998). arXiv:hep-ph/9805295

70. J. Ablinger, J. Blümlein, A. De Freitas, A. Hasselhuhn, A. von Manteuffel, M. Round et al., The transition matrix element $A_{gq}(N)$ of the variable flavor number scheme at $O(\alpha_s^3)$. Nucl. Phys. B **882**, 263 (2014). https://doi.org/10.1016/j.nuclphysb.2014.02.007. arXiv:1402.0359

71. 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

72. 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

73. M. Diehl, J.R. Gaunt, P. Plößl, A. Schäfer, Two-loop splitting in double parton distributions. SciPost Phys. **7**, 017 (2019). https://doi.org/10.21468/SciPostPhys.7.2.017. arXiv:1902.08019

74. M. Diehl, J.R. Gaunt, P. Plößl, Two-loop splitting in double parton distributions: the colour non-singlet case. JHEP **08**, 040 (2021). https://doi.org/10.1007/JHEP08(2021)040. arXiv:2105.08425

75. 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

76. ATLAS Collaboration, Measurement of the production cross section of prompt J/ψ mesons in association with a $W^{\pm}$ boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. JHEP **04**, 172 (2014). https://doi.org/10.1007/JHEP04(2014)172. arXiv:1401.2831

77. L.N. Trefethen, *Approximation Theory and Approximation Practice* (Society for Industrial and Applied Mathematics, Philadelphia, 2012)

78. P. Prince, J. Dormand, High order embedded Runge–Kutta formulae. Journal of Computational and Applied Mathematics **7**, 67 (1981). https://doi.org/10.1016/0771-050X(81)90010-3

79. W. Giele et al., The QCD/SM working group: summary report, in *Physics at TeV colliders. Proceedings, Euro Summer School, Les Houches, France, May 21–June 1, 2001.* arXiv:hep-ph/0204316

80. M. Dittmar et al., *Working Group I: Parton Distributions: Summary Report for the HERA LHC Workshop Proceedings.* arXiv:hep-ph/0511119

81. M. Diehl, T. Kasemets, S. Keane, Correlations in double parton distributions: effects of evolution. JHEP **05**, 118 (2014). https://doi.org/10.1007/JHEP05(2014)118. arXiv:1401.1233

82. M. Diehl, R. Nagar, P. Plößl, Quark mass effects in double parton distributions. arXiv:2212.07736

83. M. Diehl, P. Plößl, A. Schäfer, Proof of sum rules for double parton distributions in QCD. Eur. Phys. J. C **79**, 253 (2019). https://doi.org/10.1140/epjc/s10052-019-6777-5. arXiv:1811.00289

84. M. Diehl, F. Fabry, A. Vladimirov, Two-loop evolution kernels for colour dependent double parton distributions. JHEP **05**, 067 (2023). https://doi.org/10.1007/JHEP05(2023)067. arXiv:2212.11843