DESY-26-023

A quantitative study of two-loop splitting in double parton distributions

Markus Diehl and Peter Plößl

Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany

Double parton distributions at small distances between the two partons are dominated by a mechanism in which the two observed partons originate from the splitting of a single parton. This contribution can be computed in terms of single-parton distributions and perturbative splitting kernels. We demonstrate that two-loop corrections to these kernels can have a substantial quantitative impact and considerably improve the stability of predictions for double parton scattering. We also consider the impact of heavy quark masses in the two-loop splitting kernels in an approximate manner.

Contents

---

1 Introduction

An intriguing mechanism in proton-proton collisions at high energy is double parton scattering (DPS), where two partons in each proton enter a hard-scattering process that produces a part of the observed final state. This mechanism is generically suppressed compared with single-parton scattering (SPS), where the observed particles in the final state originate from a single hard scatter, but there are various final states and kinematical regions in which DPS can be important or even dominant. A prominent example is the production of like-sign W pairs (W^+W^+ or W^-W^-), which was proposed for studying DPS in [1, 2] and experimentally observed in Run II of the LHC [3, 4].

The theoretical description of DPS in QCD rests on pioneering work in the 1980s [5, 6], followed by substantial new developments in the 2010s [7–18]. The inclusion of double (and more generally multiple) parton interactions is also an integral part of general-purpose event generators for proton-proton collisions [19–23]. A comprehensive overview of the field up to 2017 is given in the monograph [24], and for recent theoretical or phenomenological studies we refer to [25–31]. Experimentally, double parton scattering has been observed both at the Tevatron and the LHC, see [3,4,32–35] and references therein. In addition to like-sign W pairs, observed final states in these studies are J/Ψ pairs [34–37], $W + J/\Psi$ [38], four jets [39, 40], $W+$ jets [41, 42], and several others.

The parton-level initial state in DPS is described by double parton distributions (DPDs), which quantify the joint distribution of two partons with momentum fractions x_1 and x_2 and a given transverse distance y from each other. These are genuinely non-perturbative quantities

and remain poorly known compared with single-parton distributions (PDFs), which have by now entered the realm of precision physics. However, at small distance y between the two partons, the dominant contribution to a DPD comes from graphs in which the two observed partons originate from the perturbative splitting of a single parton. For ease of language we refer to this as "parton splitting in DPDs" or "DPD splitting" henceforth.[1] This mechanism has been studied long ago in the context of scale evolution [7, 43–46] and later from more general points of view [8–14, 16, 17]. Depending on the parton combination and kinematics (parton momentum fractions and hard scales), the importance of this splitting contribution to the DPS cross section can be anything between negligible and dominant. In cases where it is dominant, one has the remarkable situation that DPS can be largely predicted in terms of ordinary PDFs and perturbatively calculable quantities. At leading order (LO) in α_s, the kernels for parton splitting in DPDs have a simple connection to the one-loop splitting functions in the DGLAP equation. This connection is lost at higher orders, and the DPD splitting kernels at next-to-leading order (NLO) have been computed for massless unpolarised partons in [47,48]. The impact of heavy quark masses in the splitting process was investigated in [49].

In physical cross sections, DPDs are evaluated at the characteristic scales μ_1 and μ_2 of the two hard-scattering processes, whereas the splitting process just described should be evaluated at its natural scale $\mu_{\text{init}} \sim 1/y$. Numerical studies in [49] have shown that the dependence of the cross section on the specific choice of μ_{init} can be substantial if the splitting is evaluated at LO, reaching an order of magnitude or more in certain kinematics. In such situations one barely has a useful theoretical prediction. The first purpose of the present study is to see how this changes when the splitting is evaluated at NLO.

Parton splitting in DPDs is also of primary importance because it gives rise to double counting between DPS and certain higher-loop graphs for SPS [8, 50]. The scheme developed in [17] avoids this double counting by imposing a lower cutoff y_{cut} on y in DPS and by writing the overall cross section as a sum of contributions from SPS, DPS, and a subtraction term. The second purpose of the present work is to revisit the construction of that subtraction term and to study the sensitivity of predictions on the choice of y_{cut} when DPD splitting is evaluated at LO or NLO.

The paper is organised as follows. In section 2 we recall the basics of DPS and of DPD splitting in the formalism of [8, 17, 18] that underlies our study. In section 3 we recall our scheme for avoiding double counting between DPS and SPS, and we present a new prescription for the subtraction term in that scheme. The quantitative impact of NLO corrections to the DPD splitting kernels is studied in section 4 for the case where quark masses are neglected and in section 5 for the case where they are included in an approximate way. We summarise our results in section 6.

2 Double parton scattering basics

In this section we recall the theory description of double parton scattering developed in [8, 18, 51], with an emphasis on the perturbative splitting mechanism for DPDs [17, 47, 48].

Consider the cross section for a process $p + p \to A_1 + A_2 + X$, where A_i denotes the system of observed particles produced in the hard subprocess number i ($i = 1, 2$), and X denotes the unobserved part of the final state. If the hard subprocesses are computed at LO, then the

[1] Note that "parton splitting" in this sense does *not* include additional splitting processes that are described by DGLAP evolution.

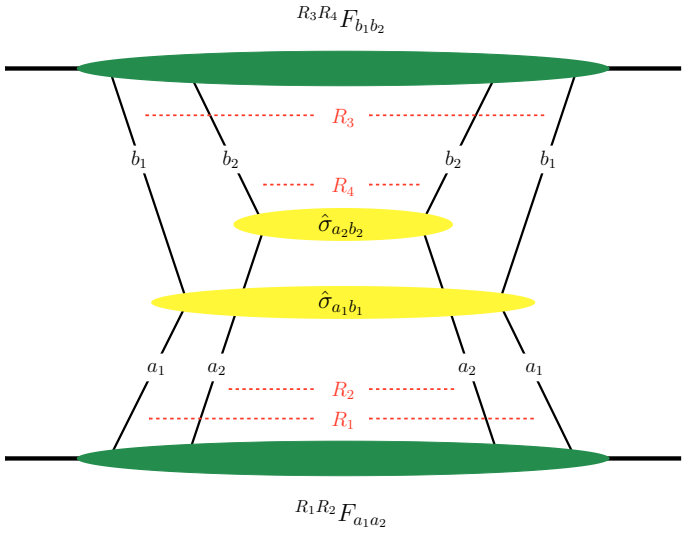


Figure 1: Graph corresponding to the factorisation formula for double parton scattering, given by (2.1) and (2.2). Pairs of parton lines have their colour indices coupled to irreducible representations R_i of SU(3). The final-state cut would be a vertical line through the centre of the graph and is omitted for clarity. The horizontal lines at the very left and the very right represent the incoming protons.

DPS cross section is given by

$$\frac{d\sigma_{\mathrm{DPS}}}{dM_1^2\,dY_1\,dM_2^2\,dY_2} = \frac{1}{s^2}\,\frac{1}{1+\delta_{A_1A_2}} \sum_{R_1R_2R_3R_4} \sum_{a_1a_2\,b_1b_2} {}^{\overline{R}_1\overline{R}_3}\hat{\sigma}_{a_1b_1}(M_1;\mu_1)\; {}^{\overline{R}_2\overline{R}_4}\hat{\sigma}_{a_2b_2}(M_2;\mu_2) \times {}^{R_1R_2,R_3R_4}\mathcal{L}_{a_1a_2,b_1b_2}(M_1,Y_1,M_2,Y_2;\mu_1,\mu_2,\nu,s)\,. \tag{2.1}$$

Kinematic quantities appearing here are the c.m. energy $\sqrt{s}$ of the proton-proton collision, along with the invariant mass M_i and the rapidity Y_i of the observed system A_i. Depending on the process, the parton-level cross sections $\hat{\sigma}$ and the physical cross section may depend on further quantities, such as the transverse momentum in jet production.
R_1 denotes the colour representation of the pair of quark or gluon lines associated with the parton label a_1, as indicated in figure 1, with corresponding associations of $R_2 \leftrightarrow a_2$, $R_3 \leftrightarrow b_1$, and $R_4 \leftrightarrow b_2$. The relevant representations are $R = 1$ and $R = 8$ for quarks and antiquarks, and $R = 1, A, S, 10, \overline{10}, 27$ for gluons, where A and S respectively stand for the antisymmetric and symmetric colour octet. By $\overline{R}_i$ we denote the conjugate of the representation R_i, which is equal to R_i for all cases except 10 and $\overline{10}$. More details about colour non-singlet DPDs can be found in [18, 52].
The parton-level cross sections $\hat{\sigma}$ in (2.1) are multiplied by double parton luminosities $\mathcal{L}$, which are products of two DPDs integrated over the transverse distance y between the two partons extracted from each proton:

$$\begin{aligned} &{}^{R_1R_2,R_3R_4}\mathcal{L}_{a_1a_2,b_1b_2}(M_1,Y_1,M_2,Y_2;\mu_1,\mu_2,\nu,s) \\ &\quad = 2\pi \int_{y_{\mathrm{cut}}}^{\infty} dy\, y\; {}^{R_1R_2}F_{a_1a_2}(x_1,x_2,y;\mu_1,\mu_2,\zeta)\; {}^{R_3R_4}F_{b_1b_2}(\bar{x}_1,\bar{x}_2,y;\mu_1,\mu_2,\bar{\zeta})\,. \end{aligned} \tag{2.2}$$

In full generality, the labels a_1, a_2 and b_1, b_2 also specify the polarisation of the partons. For transversely polarised quarks and for linearly polarised gluons, both DPDs and parton-level

cross sections carry transverse Lorentz indices, and the DPDs depend on the direction of the transverse distance vector $\vec{y}$ [8]. Our present study is restricted to unpolarised partons, since NLO splitting effects have not yet been computed for the polarised cases. We therefore have DPDs depending only on $y = |\vec{y}|$, and the trivial integration over the azimuth of $\vec{y}$ has been carried out in (2.2).

The DPDs in (2.2) are evaluated for longitudinal momentum fractions

$$x_i = \frac{M_i}{\sqrt{s}}\, e^{Y_i}\,, \qquad \bar{x}_i = \frac{M_i}{\sqrt{s}}\, e^{-Y_i}\,, \qquad i = 1, 2, \tag{2.3}$$

which are fixed by the process kinematics. If the parton-level cross sections are evaluated beyond LO, then the factorisation formula (2.1) must be modified to contain convolution integrals over the parton momentum fractions, in the same way as is done for single parton scattering.

The DPDs and parton-level cross sections depend on factorisation scales μ_1 and μ_2, and the dependence of the DPDs on these scales is given by DGLAP equations whose full form can be found in section 2.2 of [53]. For colour singlet distributions ($R_1 = R_2 = 1$) these are straightforward generalisations of the DGLAP equations for PDFs and contain the same DGLAP splitting functions. The evolution of colour non-singlet DPDs up to NLO accuracy has been studied in [51, 53].

As will be reviewed in section 3, an important property of factorisation for DPS is that the integral over y has a lower cutoff y_{cut}, which gives rise to a further momentum scale ν:

$$y_{\text{cut}} = b_0/\nu\,, \tag{2.4}$$

where

$$b_0 = 2e^{-\gamma_E} \approx 1.12 \tag{2.5}$$

and γ_E is the Euler-Mascheroni constant. As explained in section 6 of [17], an appropriate choice of ν is

$$\nu = \min(\mu_1, \mu_2)\,, \tag{2.6}$$

which will be adopted as a default throughout this work. The rapidity parameters ζ and $\bar{\zeta}$ on the r.h.s. of (2.2) will be explained in section 2.1.

We write the DPDs in terms of splitting and intrinsic contributions as

$$F = F^{\text{spl}} + F^{\text{intr}}, \tag{2.7}$$

where each contribution depends on the same labels and variables as F in (2.2). This separation is model independent at small y, where the two contributions can be defined via an operator product expansion [17]. When evaluated at $\mu_1 \sim \mu_2 \sim \sqrt{\zeta} \sim 1/y$, they have a power behaviour $F^{\text{spl}} \sim y^{-2}$ and $F^{\text{intr}} \sim y^0$, and they can be evaluated by the convolution of perturbative kernels with either PDFs (for F^{spl}) or twist-four distributions (for F^{intr}). For y in the non-perturbative region, we retain the separation (2.7) in the sense of a model ansatz. The separation (2.7) induces a separation of double parton luminosities into four parts:

$$\mathcal{L}_{1\text{v}1} \leftrightarrow F^{\text{spl}}\, F^{\text{spl}}\,, \quad \mathcal{L}_{1\text{v}2} \leftrightarrow F^{\text{spl}}\, F^{\text{intr}}\,, \quad \mathcal{L}_{2\text{v}1} \leftrightarrow F^{\text{intr}}\, F^{\text{spl}}\,, \quad \mathcal{L}_{2\text{v}2} \leftrightarrow F^{\text{intr}}\, F^{\text{intr}}\,, \tag{2.8}$$

where "1" indicates that there is a single parton at the origin of the two partons that enter the DPS process [10].

The perturbative splitting formula for DPDs has the form

$$
\begin{aligned}
{}^{R_1R_2}F^{\text{spl}}_{a_1a_2}(x_1,x_2,y;\mu,\mu,\zeta) &= \frac{1}{\pi\, y^2} \sum_{a_0} \int_{x_1+x_2}^{1} \frac{dz}{z^2}\; {}^{R_1R_2}V_{a_1a_2,a_0}\left(\frac{x_1}{z},\frac{x_2}{z},\alpha_s(\mu),L_y,L_\zeta\right) f_{a_0}(z;\mu) \\
&\stackrel{\text{def}}{=} \frac{1}{\pi\, y^2} \sum_{a_0} {}^{R_1R_2}V_{a_1a_2,a_0} \underset{12}{\otimes} f_{a_0}
\end{aligned}
\tag{2.9}
$$

with logarithms

$$
L_y = \ln \frac{y^2\mu^2}{b_0^2}\,, \qquad\qquad L_\zeta = \ln \frac{\mu^2}{x_1x_2\,\zeta}\,. \tag{2.10}
$$

The splitting kernels ${}^{R_1R_2}V_{a_1a_2,a_0}$ have a perturbative expansion starting at order α_s. The masses of heavy quarks (c, b, t) can be taken into account in the splitting, as we will discuss in section 5.

2.1 Rapidity evolution

As explained in [18], colour non-singlet DPDs depend not only on the factorisation scales μ_1 and μ_2 for the two partons, bit also on a rapidity parameter. In the double parton luminosity (2.2), the rapidity parameters ζ and $\bar\zeta$ of the two DPDs must satisfy the kinematic relation

$$
\left(\zeta\bar\zeta\right)^{1/2} = s\,, \tag{2.11}
$$

which is realised by the explicit choice

$$
\zeta = \frac{M_1M_2}{x_1x_2}\,, \qquad\qquad \bar\zeta = \frac{M_1M_2}{\bar x_1\bar x_2}\,. \tag{2.12}
$$

The reason for choosing rapidity parameters that depend on the parton momentum fractions is explained in section 3.2 of [53].
The rapidity dependence of DPDs is described by a Collins-Soper equation

$$
\frac{d}{d\ln\zeta}\, \ln {}^{R_1R_2}F_{a_1a_2}(x_1,x_2,y;\mu_1,\mu_2,\zeta) = \frac{1}{2}\, {}^{R_1}J(y;\mu_1,\mu_2)\,, \tag{2.13}
$$

in close analogy to the case of transverse-momentum dependent parton distributions (TMDs) [54]. The Collins-Soper kernel ${}^{R}J$ satisfies a renormalisation group equation in μ_1

$$
\frac{d}{d\ln\mu_1}\, {}^{R}J(y;\mu_1,\mu_2) = -\, {}^{R}\gamma_J(\mu_1) \tag{2.14}
$$

and its analogue for the scale μ_2. The kernel ${}^{R}J$ is zero for the colour singlet ($R = 1$) and nonzero for all other colour representations. For the colour octet channels, one has the exact relation [18, 55]

$$
{}^{8}J(y;\mu,\mu) = {}^{S}J(y;\mu,\mu) = {}^{A}J(y;\mu,\mu) = K_g(y,\mu)\,, \tag{2.15}
$$

where K_g is the Collins-Soper kernel for the rapidity evolution of single-gluon TMDs.
The system (2.13) and (2.14) of differential equations is readily solved analytically. The DGLAP equations for the μ_1 and μ_2 dependence of DPDs depend on ζ and γ_J as well. Solving the full system of evolution equations and using the condition (2.11), one finds that the explicit dependence of the DPDs on ζ or $\bar\zeta$ drops out in the double parton luminosity (2.2).

P	β/α_s	γ_J	J
α_s	α_s	α_s^2	α_s
α_s^2	α_s^2	α_s^3	α_s^2

Table 1: Combination of perturbative orders used in the present work for scale and rapidity evolution. The entries for P refer to the DGLAP kernels for the evolution of the DPDs and of the PDFs in the DPD splitting formula, those for β/α_s to the running of α_s, and those for J to the fixed-order perturbative input ${}^8J^{\text{pt}}$ of the Collins-Soper kernel in (2.16).

A detailed discussion is given in section 4.2 of [53]. The particular choice (2.12) is hence not relevant for the value of the double parton luminosities, but it simplifies the discussion of rapidity evolution effects in the DPDs.
In the present work we will only be concerned with colour singlet and colour octet channels. We use (2.15) and assume Casimir scaling between the Collins-Soper kernels K_g and K_q for gluon and quark TMDs, which holds at small distances up to order α_s^3 [56, 57]. This allows us to construct a model for ${}^8J$ using input from phenomenological determinations of K_q. We follow the model construction laid out in section 3.1 of [53], which has the form

$$\begin{aligned} &{}^8J(y;\mu_1,\mu_2) \\ &\quad = {}^8J^{\text{pt}}\big(y^*(y);\mu^*(y),\mu^*(y)\big) + {}^8\Delta J(y) - \int_{\mu^*(y)}^{\mu_1} \frac{d\mu}{\mu}\, {}^8\gamma_J(\mu) - \int_{\mu^*(y)}^{\mu_2} \frac{d\mu}{\mu}\, {}^8\gamma_J(\mu) \end{aligned} \tag{2.16}$$

with functions $\mu^*(y)$ and $y^*(y)$ specified in (2.17) and (2.19) below. The short-distance part ${}^8J^{\text{pt}}$ is evaluated in fixed-order perturbation theory, and the long-distance part ${}^8\Delta J$ is fixed such that the full kernel ${}^8J$ is equal to C_A/C_F times the kernel K_q in the SV19 parametrisation [58].
The combinations of perturbative orders used for ${}^8J^{\text{pt}}$, for the anomalous dimension ${}^8\gamma_J$, and for the QCD β function are discussed in section 4.3 of [53]. In the present work we use the two combinations given in table 1. We note that the β function is taken at one order less than what is required for NLL or NNLL accuracy in the nomenclature of resummation. The reason for this is that we take the PDFs in the DPD splitting formula from global fits, which use the same perturbative order for the DGLAP splitting functions and for β/α_s. For further discussion, we refer to section 2.4 in [59] and to section 4 in [60].

2.2 Initial scales for DPD evolution

In DPS cross sections, DPDs are needed at the scales μ_1 and μ_2 of the two parton-level subprocesses. We obtain them by DGLAP evolution from a suitable starting scale μ_{init}, which is always taken equal for the two partons.
To evaluate the initial condition for the splitting part F^{spl} using a fixed-order formula, we take a y dependent initial scale $\mu_{\text{init}}(y)$ so as to avoid large logarithms of $y\mu_{\text{init}}$ that would spoil the convergence of the perturbative series. Our default choice is $\mu_{\text{init}}(y) = \mu^*(y)$, where

$$\mu^*(y) = b_0/y^*(y) \tag{2.17}$$

with a function $y^*(y)$ that satisfies

$$y^*(y) \to \begin{cases} y & \text{for } y \ll 1/\Lambda\,, \\ y_{\text{max}} & \text{for } y \gg 1/\Lambda\,, \end{cases} \tag{2.18}$$

where Λ represents a typical non-perturbative scale. As discussed in section 3.2 of [53], it is natural to use $\mu_{\text{init}}(y)$ as an initial scale also for the intrinsic part F^{intr}. In our numerical studies we take

$$y^*(y) = \frac{y}{\sqrt[4]{1 + y^4/y_{\text{max}}^4}}\,, \tag{2.19}$$

which ensures that μ^* is always greater than $\mu_{\text{min}} = b_0/y_{\text{max}}$. We set

$$\mu_{\text{min}} = 2\,\text{GeV} \tag{2.20}$$

as we did in our evolution study [53], mainly in order to avoid the gluon PDF in the DPD splitting formula becoming negative at low x and low scales.
For later use, we note that the choice (2.19) implies

$$\begin{aligned} \mu^*(y_{\text{cut}}) &= \nu\ \sqrt[4]{1 + \mu_{\text{min}}^4/\nu^4} \\ &\approx \nu \qquad\qquad\qquad\qquad \text{for } \nu \gg \mu_{\text{min}}\,. \end{aligned} \tag{2.21}$$

The corrections to the approximation in the second line are of order $(\mu_{\text{min}}/\nu)^4$ and thus tiny in the range $\nu \geq 10\,\text{GeV}$ we will consider later in this work.
Again in the spirit of avoiding large logarithms in the fixed-order splitting formula, we take the initial scale of rapidity evolution as

$$\zeta_{\text{init}}(y) = \frac{\mu_{\text{init}}^2(y)}{x_1\,x_2}\,, \qquad\qquad \bar{\zeta}_{\text{init}}(y) = \frac{\mu_{\text{init}}^2(y)}{\bar{x}_1\,\bar{x}_2} \tag{2.22}$$

in the relevant DPD. To assess the dependence of our results on the choice of $\mu_{\text{init}}(y)$, we will vary that scale within the limits

$$\min\bigl\{\mu_{\text{min}},\, \mu^*(y)/2\bigr\} \leq \mu_{\text{init}}(y) \leq 2\mu^*(y) \tag{2.23}$$

around our default choice $\mu_{\text{init}}(y) = \mu^*(y)$. This corresponds to the conventional variation by a factor of 2 up and down when $\mu^*(y) > 2\mu_{\text{min}}$, while avoiding scales below μ_{min} at large y. We regard any initial scale within these limits as a reasonable theoretical choice.
We will shortly show that splitting DPDs obtained with initial scales in the region (2.23) differ by an amount that is of higher perturbative order than the order used in the splitting formula. Variation of scales that are internal to perturbative calculations is often used to estimate the uncertainty due to missing higher orders, but the limitations of such an approach have long been known, and a detailed recent discussion may be found in section 2.3 of [61]. Nevertheless, the impact of varying μ_{init} within the limits (2.23) carries important information. If this results in a large variation of double parton luminosities, higher-order corrections are manifestly large. Conversely, a moderate or small impact of the scale variation on double parton luminosities is a necessary although not a sufficient condition for having a reliable theory result.
Let us assume that the DPD splitting kernels are evaluated up to order α_s^k, with $k = 1$ at LO, $k = 2$ at NLO, etc. We introduce the notation

$$F^{\text{spl}(k)}(x_1, x_2, y; \mu_1, \mu_2, \zeta | \mu_0, \zeta_0) \tag{2.24}$$

for a DPD that is initialised from the k-th order splitting formula at scale μ_0 and rapidity parameter ζ_0 and then evolved to the final scales μ_1, μ_2 and the final rapidity parameter ζ. For brevity, we omit parton and colour labels in (2.24) and the following equations.
We must of course require that y is in the perturbative regime, where the splitting formula for DPDs is applicable and where the Collins-Soper kernel ${}^R\!J$ can be expanded in α_s. We

also require that $\mu_0 \sim \sqrt{\zeta_0} \sim 1/y$, so that no large logarithms in the ratio of these scales can spoil counting powers of α_s in the following perturbative estimates.[2]
In terms of an evolution operator U, which is the Green's function of the combined system of DGLAP and Collins-Soper equations, we can write

$$\begin{aligned} F^{\mathrm{spl}(k)}(x_1,x_2,y;\mu_1,\mu_2,\zeta|\mu_0,\zeta_0) = \int_{x_1}^1 \frac{dz_1}{z_1} \int_{x_2}^1 \frac{dz_2}{z_2}\, U(x_1,x_2,\mu_1,\mu_2,\zeta|z_1,z_2,\mu_0,\mu_0,\zeta_0) \\ \times F^{\mathrm{spl}(k)}(z_1,z_2,y;\mu_0,\mu_0,\zeta_0|\mu_0,\zeta_0)\,, \end{aligned} \tag{2.25}$$

where the DPD on the r.h.s. is given by the DPD splitting formula (2.9) at order α_s^k.
It is sufficient for the following results that the evolution of DPDs and of the PDFs in the splitting formula is evaluated with DGLAP and Collins-Soper kernels of order α_s^{k-1}. This is consistent because the explicit logarithms of μ or ζ in the k-th order splitting formula are accompanied by evolution kernels of at least one order less, given that the splitting DPD itself starts at order α_s [47, 48]. In practice one will typically take the evolution kernels at order α_s^k, and γ_J even at order α_s^{k+1} (see also table 1).
The splitting DPD evaluated up to order k satisfies the Collins-Soper equation up to terms of order α_s^{k+1} or higher, i.e.

$$\begin{aligned} &\frac{d}{d\ln\zeta_0} F^{\mathrm{spl}(k)}(x_1,x_2,y;\mu_0,\mu_0,\zeta_0\,|\,\mu_0\zeta_0) \\ &\quad = \frac{1}{2}\, J^{(k-1)}(y;\mu_0,\mu_0)\, F^{\mathrm{spl}(k)}(x_1,x_2,y;\mu_0,\mu_0,\zeta_0\,|\,\mu_0\zeta_0) + \mathcal{O}\bigl(\alpha_s^{k+1}(\mu_0)\bigr) \end{aligned} \tag{2.26}$$

where $J^{(i)}$ is the Collins-Soper kernel evaluated up to order α_s^i. A corresponding statement holds for the DGLAP equation. As a consequence, the dependence on μ_0 and ζ_0 does not exactly cancel between U and F on the r.h.s. of (2.25), and one has

$$\begin{aligned} \frac{d}{d\ln\mu_0} F^{\mathrm{spl}(k)}(x_1,x_2,y;\mu_1,\mu_2,\zeta|\mu_0,\zeta_0) &= \mathcal{O}\bigl(\alpha_s^{k+1}(\mu_0)\bigr) \\ \frac{d}{d\ln\zeta_0} F^{\mathrm{spl}(k)}(x_1,x_2,y;\mu_1,\mu_2,\zeta|\mu_0,\zeta_0) &= \mathcal{O}\bigl(\alpha_s^{k+1}(\mu_0)\bigr) \end{aligned} \tag{2.27}$$

We emphasise that the orders on the r.h.s. are dictated by the truncation of the DPD splitting kernel and not by the orders at which evolution is performed, as long as the latter includes the order α_s^{k-1}.
Integrating the differential equations (2.27) between two scales μ_a and μ_b, we obtain

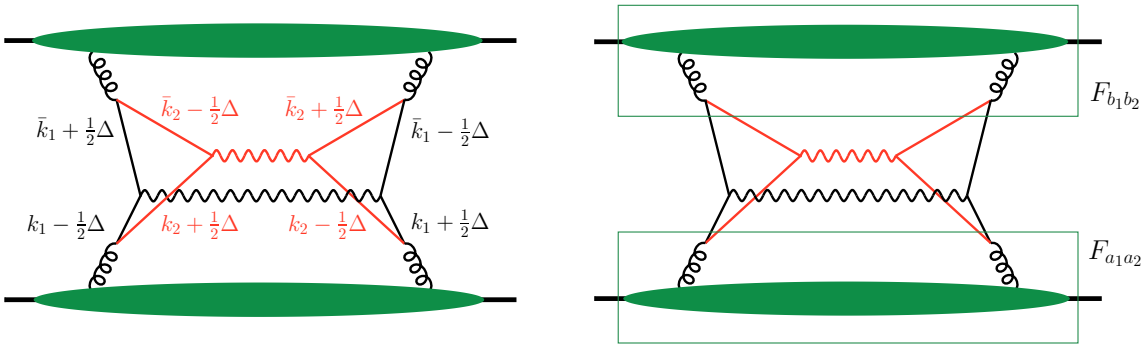
$$\begin{aligned} &F^{\mathrm{spl}(k)}\left(x_1,x_2,y;\mu_1,\mu_2,\zeta \,\middle|\, \mu_a, \frac{\mu_a^2}{x_1x_2}\right) - F^{\mathrm{spl}(k)}\left(x_1,x_2,y;\mu_1,\mu_2,\zeta \,\middle|\, \mu_b, \frac{\mu_b^2}{x_1x_2}\right) \\ &\quad = \mathcal{O}\bigl(\alpha_s^{k+1}(\mu_a)\bigr) \qquad\qquad \text{for } \mu_a \sim \mu_b \sim 1/y\,. \end{aligned} \tag{2.28}$$

The dependence of an evolved splitting DPD on the initial scale is thus one order higher than the perturbative accuracy at which the DPD is evaluated, as we claimed above.

3 Double counting subtraction

Since the two partons in a DPD can originate from the splitting of a single parton, there is an overlap between DPS with two splitting DPDs and higher-loop corrections to single parton

[2] This statement holds up to possible large logarithms $\ln x_1$ and $\ln x_2$, which are of dynamical origin and beyond the scope of our analysis.



(a) SPS (b) DPS

Figure 2: (a): Box graph for the production of two electroweak gauge bosons via SPS. All loop momenta are routed such that they flow towards the gauge boson vertex. (b): The same graph, interpreted as a contribution to DPS with two splitting DPDs. The final state cut would be a vertical line through the centre of the graph and is omitted for clarity.

scattering [8]. In different regions of the loop momenta, the same graph can be regarded as either DPS or SPS, as shown in figure 2. A scheme to delineate the two mechanisms was developed in [17]. In this scheme, double counting between SPS and DPS is avoided by a subtraction term. It is the purpose of this section to revisit the construction of this term.
In the scheme of [17], the full cross section for a process to which DPS can contribute reads

$$\sigma = \sigma_{\text{DPS}}(\nu) - \sigma_{\text{sub}}(\nu) + \sigma_{\text{SPS}} \,. \tag{3.1}$$

To understand this construction, we write all three terms in (3.1) as integrals over the transverse distance between the two partons:

$$\begin{aligned}
\sigma_{\text{DPS}}(\nu) &= \int d^2y \; \theta(y - y_{\text{cut}}) \, I_{\text{DPS}}(y) \,, \\
\sigma_{\text{sub}}(\nu) &= \int d^2y \; \theta(y - y_{\text{cut}}) \, I_{\text{sub}}(y) \,, \\
\sigma_{\text{SPS}} &= \sigma_{\text{SPS},0} + \int d^2y \; I_{\text{SPS}}(y) \,,
\end{aligned} \tag{3.2}$$

where we use y_{cut} and $\nu = b_0/y_{\text{cut}}$ interchangeably. Throughout this section, it is understood that $\nu \sim \min(\mu_1, \mu_2)$.
The expression of I_{DPS} in terms of DPDs and parton-level cross sections is readily obtained from (2.1) and (2.2). In analogy to (2.8) it can be split into four contributions that involve the splitting and intrinsic parts of the two DPDs in all combinations: $I_{\text{DPS}} = I_{\text{1v1}} + I_{\text{1v2}} + I_{\text{2v1}} + I_{\text{2v2}}$. At small y this decomposition is model independent.
The full SPS cross section σ_{SPS} can be expressed by a standard factorisation formula involving PDFs and parton-level cross sections. The part $\sigma_{\text{SPS},0}$ is due to graphs that do not have any overlap with DPS, whereas I_{SPS} is obtained from the loop graphs that do have a DPS region. In these graphs, a particular loop momentum Δ can be identified such that its transverse component is Fourier conjugate to the transverse distance $\vec{y}$. An example is given in figure 2a.
The integrand of the double counting subtraction term σ_{sub} must satisfy the conditions

$$I_{\text{sub}}(y) \to \begin{cases} I_{\text{1v1}}(y) & \text{for } y \sim y_{\text{cut}} \,, \\ I_{\text{SPS}}(y) & \text{for } y \gg y_{\text{cut}} \,, \end{cases} \tag{3.3}$$

in order to ensure that

1. for $y \sim y_{\text{cut}}$ the subtraction term cancels the 1v1 contribution to DPS in the sum (3.1), such that one is left with the SPS term, which correctly describes the process in that region of y,

2. the dependence on ν cancels between the DPS and subtraction terms,

3. for $y \gg y_{\text{cut}}$ the subtraction term cancels the part of SPS that has overlap with DPS. One is then left with the DPS term and with the part of SPS that has no overlap with DPS.

In all three points, the cancellation is understood to be accurate up to remainder terms that are beyond the accuracy of the calculation.

Our discussion so far is oversimplified in the sense that the double counting subtraction in (3.1) only addresses the 1v1 part of the DPS cross section. The 1v2 and 2v1 parts of σ_{DPS} also have an overlap, namely with contributions that involve a twist-four distribution in one of the colliding protons and a PDF in the other proton. In principle one should include these contributions, along with a further double counting subtraction term, but this would be quite impractical, given the paucity of knowledge about twist-four distributions and the fact that the corresponding hard-scattering cross sections have not been computed. As explained in section 4.1 of [17], omitting these twist-four contributions while keeping the 1v2 and 2v1 parts of σ_{DPS}, one retains the leading logarithmic part of graphs in which the hard scattering is initiated by two partons in one proton and one parton in the other proton. The ν dependence of the 1v2 and 2v1 contributions to σ_{DPS} goes like $\ln \nu$ and is much weaker than the quadratic ν dependence of the 1v1 contribution.

Returning to the construction of I_{sub}, we recall that the integrand I_{DPS} is constructed from DPDs

$$F\left(x_1, x_2, y; \mu_1, \mu_2, \frac{M_1 M_2}{x_1 x_2} \,\middle|\, \mu_{\text{init}}(y), \frac{\mu^2_{\text{init}}(y)}{x_1 x_2}\right) \tag{3.4}$$

with $F = F^{\text{spl}} + F^{\text{intr}}$. Here we use again the notation introduced in (2.24) that specifies both the initial and final scales of evolution. In the colour singlet sector, the construction of I_{sub} is very simple if the two hard scales are equal, $\mu_1 = \mu_2 = \mu$ and if one takes $\mu_h = \mu$ as factorisation scale in the SPS cross section. In that case, I_{sub} is obtained from I_{DPS} by replacing the DPDs in (3.4) by F^{spl} evaluated with the splitting formula at the scale μ.

For unequal scales $\mu_1 \neq \mu_2$, section 6.3 of the original work [17] outlined a construction in which the splitting DPDs in the subtraction term have y dependent initial *and* final scales μ_1, μ_2, and μ_0. In the present work we present an alternative that we find both simpler and more flexible. We also include the case of colour non-singlet DPDs, which was not considered in [17]. In our new scheme, I_{sub} is constructed from two components $I_{\text{sub,DPS}}$ and $I_{\text{sub,SPS}}$ in the following manner.

- $I_{\text{sub,DPS}}(y)$ is constructed with DPDs

$$F^{\text{spl}}\left(x_1, x_2, y; \mu_1, \mu_2, \frac{M_1 M_2}{x_1 x_2} \,\middle|\, \mu_{\text{sub,init}}, \frac{\mu^2_{\text{sub,init}}}{x_1 x_2}\right) \tag{3.5}$$

  instead of those in (3.4), where

$$\mu_{\text{sub,init}} = \mu_{\text{init}}(y_{\text{cut}})\,. \tag{3.6}$$

  By construction, $I_{\text{sub,DPS}}(y)$ and $I_{\text{1v1}}(y)$ are equal at $y = y_{\text{cut}}$ and gradually deviate from each other as y increases. This holds for the default choices of μ_{init} and $\nu = b_0/y_{\text{cut}}$ as well as for the cases where these scales are varied.

Note that for our central choice of initial scale one has

$$\mu_{\text{sub,init}} = \mu^*(y_{\text{cut}}) \approx \nu\,, \tag{3.7}$$

where the approximation is very accurate for the range of ν in our numerical studies, as mentioned after equation (2.21).

If we also take our default choice (2.6) for ν, we get

$$\mu_{\text{sub,init}} \approx \min(\mu_1, \mu_2)\,. \tag{3.8}$$

For $\mu_1 < \mu_2$ the DPDs in $I_{\text{sub,DPS}}(y)$ are then initialised with both scales equal to μ_1, after which the second scale is evolved to μ_2. There is evolution in ζ as well. An analogous statement holds for $\mu_1 > \mu_2$. If $\mu_1 = \mu_2$, then the DPDs are essentially not evolved at all.

- $I_{\text{sub,SPS}}(y)$ is constructed from DPDs

$$F^{\text{spl}}\left(x_1, x_2, y; \mu_h, \mu_h, \frac{M_1 M_2}{x_1 x_2} \,\middle|\, \mu_h, \frac{M_1 M_2}{x_1 x_2}\right) \tag{3.9}$$

that are not evolved in either the scale or the rapidity parameter. Here the factorisation scale μ_h is the same as the one taken in the SPS factorisation formula. We take

$$\mu_h = \frac{1}{2}(\mu_1 + \mu_2) \tag{3.10}$$

in our numerical studies; alternative choices such as $\mu_h = \sqrt{\mu_1\,\mu_2}$ would work as well.

- The full subtraction term interpolates between these two functions as

$$I_{\text{sub}}(y) = \big[1 - \rho(y)\big]\, I_{\text{sub,DPS}}(y) + \rho(y)\, I_{\text{sub,SPS}}(y) \tag{3.11}$$

where ρ is a function that smoothly interpolates between 0 and 1 in the interval $y \in [r_1\, y_\rho, r_2\, y_\rho]$ with $r_1 < r_2$ and $y_\rho \sim y_{\text{cut}}$. We chose

$$\rho(y) = \tilde{\rho}\left(\frac{2y - (r_2 + r_1)\, y_\rho}{(r_2 - r_1)\, y_\rho}\right) \tag{3.12}$$

with

$$\tilde{\rho}(x) = \begin{cases} 0 & \text{for } x \le -1, \\ \frac{1}{2} + \frac{1}{2}\sin\Big(\frac{\pi}{2}\sin\big(\frac{\pi}{2}\,x\big)\Big) & \text{for } -1 < x < 1, \\ 1 & \text{for } x \ge 1. \end{cases} \tag{3.13}$$

$\tilde{\rho}(x)$ has the feature that its first and second derivatives are continuous and zero at $|x| = 1$. Alternatively, one may replace $\sin\frac{\pi}{2}z$ with $\frac{3}{2}z\big(1 - \frac{1}{3}z^2\big)$ for one or both sine functions in (3.13).

For the results derived in section 3.1 it is required (*i*) that $r_2\, y_\rho$ is in the perturbative region and (*ii*) that the size of r_2 is limited by the condition $\alpha_s(\nu)\, \ln r_2 \ll 1$.

With this new scheme it is rather easy to assess the dependence of results on the specific form of the interpolation function (3.13), because a change of that function does not require re-computation of DPDs. This is not the case for the construction in [17], where the interpolation

between the SPS and DPS regimes is achieved at the level of the initial and final scales of the splitting DPDs.
In the case $\mu_1 = \mu_2$ our default choice of ν gives $\mu_{\text{sub,init}} \approx \mu_1 = \mu_h$ according to (3.8) and (3.10). In the colour singlet sector, one then has $I_{\text{sub}} \approx I_{\text{sub,DPS}} \approx I_{\text{sub,SPS}}$ to high accuracy and recovers our simple original construction for equal scales. The same holds for colour non-singlet channels if one additionally takes $\mu_i = M_i$. In these cases, the choice of the interpolation function $\rho(y)$ has no impact on the overall double counting subtraction. It does, however, determine the individual terms in the sum (3.14) below.
As will be discussed in the following sections, the DPD splitting formula should be evaluated for a number n_f of active quark flavours that depends on y, possibly with the inclusion of heavy quark masses. The logic of the subtraction mechanism then requires that

- $I_{\text{sub,DPS}}$ should be evaluated with the flavour settings used for DPDs at $y = y_{\text{cut}}$,
- the value of n_f and the treatment of heavy quark masses in $I_{\text{sub,SPS}}$ should be the same as in the computation of the SPS cross section.

Since both the DPS cross section and the double counting subtraction involve products of DPDs, their relative size and their combination can be discussed at the level of double parton luminosities, which are independent of specific hard-scattering cross sections. The counterpart of the double parton luminosity $\mathcal{L}$ that goes into σ_{DPS} is then the double parton luminosity $\mathcal{L}_{\text{sub}}$ from which σ_{sub} is constructed. Separating the two terms in the interpolation prescription (3.11), we can further write

$$\mathcal{L}_{\text{sub}}(\nu) = \mathcal{L}_{\text{sub,DPS}}(\nu) + \mathcal{L}_{\text{sub,SPS}}(\nu)\,, \tag{3.14}$$

where the first term corresponds to the y-integral of $(1-\rho)\, I_{\text{sub,DPS}}$ and the second one to the y-integral of $\rho\, I_{\text{sub,SPS}}$. Both terms involve the product of two DPDs without hard-scattering cross sections.

3.1 Dependence on the cutoff scale

Let us now show to which formal accuracy the ν dependence cancels in the overall cross section (3.1) with our new construction. As in section 2.2, we assume that splitting DPDs are computed up to order α_s^k and that evolution is performed at order α_s^{k-1} or higher. As a compact notation let us write

$$\begin{aligned} P(y;\mu) = 2\pi y\, F^{\text{spl}(k)}&\left(x_1, x_2, y; \mu_1, \mu_2, \frac{M_1 M_2}{x_1 x_2} \,\middle|\, \mu, \frac{\mu^2}{x_1 x_2}\right) \\ &\times F^{\text{spl}(k)}\left(\bar{x}_1, \bar{x}_2, y; \mu_1, \mu_2, \frac{M_1 M_2}{\bar{x}_1 \bar{x}_2} \,\middle|\, \mu, \frac{\mu^2}{\bar{x}_1 \bar{x}_2}\right), \end{aligned} \tag{3.15}$$

where for brevity we suppress parton and colour labels, as well as the dependence of P on all arguments that are left constant in the following derivation. We then have

$$\begin{aligned} \mathcal{L}_{\text{1v1}}(\nu) &= \int_{b_0/\nu}^{\infty} dy\, P\big(y; \mu_{\text{init}}(y)\big)\,, \\ \mathcal{L}_{\text{sub,DPS}}(\nu) &= \int_{b_0/\nu}^{\infty} dy\, \big[1-\rho(y)\big]\, P\big(y; \mu_{\text{init}}(b_0/\nu)\big)\,. \end{aligned} \tag{3.16}$$

We now study the difference of subtracted 1v1 luminosities at two cutoff scales $\nu_1 \sim \nu_2$. We take the *same* interpolating function $\rho(y)$ and require that the parameter y_ρ in (3.12) satisfies

$y_\rho \sim 1/\nu_1 \sim 1/\nu_2$ and $r_1\, y_\rho \geq \max(b_0/\nu_1, b_0/\nu_2)$. The second condition ensures that $\rho(y) = 0$ for y between b_0/ν_1 and b_0/ν_2. We then have

$$\begin{aligned}
&\big[\mathcal{L}_{1\text{v}1}(\nu_1) - \mathcal{L}_{\text{sub}}(\nu_1)\big] - \big[\mathcal{L}_{1\text{v}1}(\nu_2) - \mathcal{L}_{\text{sub}}(\nu_2)\big] \\
&\quad = \int_{b_0/\nu_1}^{\infty} dy \, \Big\{ P(y, \mu_{\text{init}}(y)) - \big[1 - \rho(y)\big]\, P(y; \mu_{\text{init}}(b_0/\nu_1)) \Big\} \\
&\qquad - \int_{b_0/\nu_2}^{\infty} dy \, \Big\{ P(y, \mu_{\text{init}}(y)) - \big[1 - \rho(y)\big]\, P(y; \mu_{\text{init}}(b_0/\nu_2)) \Big\} \\
&\quad = \int_{b_0/\nu_1}^{b_0/\nu_2} dy \, \Big\{ P(y, \mu_{\text{init}}(y)) - P(y; \mu_{\text{init}}(b_0/\nu_1)) \Big\} \\
&\qquad + \int_{b_0/\nu_2}^{\infty} dy \, \big[1 - \rho(y)\big] \, \Big\{ P(y; \mu_{\text{init}}(b_0/\nu_2)) - P(y; \mu_{\text{init}}(b_0/\nu_1)) \Big\} \,. \qquad (3.17)
\end{aligned}$$

Note that the contribution from $\mathcal{L}_{\text{sub,SPS}}$ has dropped out because it is independent of the cutoff parameter ν. Because of the factor $1 - \rho(y)$, the last integral in (3.17) is limited to $y \leq r_2\, y_\rho$, and the integrand is strongly damped in the vicinity of its upper limit. Under the conditions stated after equation (3.13), we can apply the result (2.28) to both integrals on the r.h.s. of (3.17). This gives

$$\big[\mathcal{L}_{1\text{v}1}(\nu_1) - \mathcal{L}_{\text{sub}}(\nu_1)\big] - \big[\mathcal{L}_{1\text{v}1}(\nu_2) - \mathcal{L}_{\text{sub}}(\nu_2)\big] = \mathcal{O}\big(\alpha_s^{k+2}(\nu_1)\big) \,, \qquad (3.18)$$

where the power of α_s results from the fact that a splitting DPD starts at order α_s. The product of two splitting DPDs computed at order k contains significant terms from order α_s^2 up to order α_s^{k+1}, so that the difference in (3.18) is beyond the accuracy of the computation, as it should be.

3.2 Summary of the subtraction formalism

At this point we can review the overall logic of our subtraction formalism.

- For $y \sim y_{\text{cut}} \sim 1/\min(\mu_1, \mu_2)$, the DPS formalism is not valid, and the correct description of the process is given by SPS alone. In the region $y \leq r_1 y_\rho$, the subtraction term we constructed removes the 1v1 part of DPS up to contributions that are at least one order of α_s higher than the accuracy of the double parton luminosities. These contributions may be regarded as an artefact of our procedure, and their size can be estimated by varying y_{cut} in $\sigma_{1\text{v}1} - \sigma_{\text{sub}}$ around a suitably chosen central value.

- For $y \gg y_{\text{cut}}$, the DPS formalism is valid. In the region where y is still perturbative, the 1v1 contribution is fully calculable in terms of PDFs and DPD splitting kernels, and it correctly resums large logarithms of $y\mu_1$ and $y\mu_2$ (as well as large Sudakov double logarithms in colour non-singlet channels) that are missed when SPS is evaluated at fixed order. In the region $y \geq r_2\, y_\rho$, our subtraction term removes the contribution of SPS graphs that have DPS topology, up to power corrections in y_{cut}/y that are due to applying the DPS approximation to the graphs.

- In the region $r_1\, y_\rho \leq y \leq r_2\, y_\rho$, our subtraction term smoothly interpolates between the two cases just described. As is typical of the interpolation between two regimes with different approximations, there is no simple parametric estimate of the error incurred by this procedure. This error can be estimated by varying the parameters of the interpolating function (i.e. r_1 and r_2 in our case) within appropriate limits.

We recall from the discussion above equation (3.4) that our master formula (3.1) misses explicit contributions with a twist-four distribution in one proton and a PDF in the other, but that the 1v2 and 2v1 parts of DPS include the part of these contributions that is enhanced by $\ln(\nu/\Lambda)$. The variation of the 1v2 and 2v1 parts of DPS with y_{cut} around an appropriate central value may hence be regarded as an estimate of the uncertainty of this specific leading logarithmic approximation. The variation of the 2v2 part of DPS with y_{cut} is negligible, because this contribution is dominated by a broad range of y values in the non-perturbative region.
Overall, the variation of $\sigma_{\text{DPS}} - \sigma_{\text{sub}}$ or $\mathcal{L}_{\text{DPS}} - \mathcal{L}_{\text{sub}}$ with y_{cut} may thus be regarded as a measure for the theoretical uncertainty of our scheme for handling the overlap between DPS and other contributions to a given physical process.

4 Splitting with massless quarks

In this section, we study DPD splitting without heavy quark masses. We use the general perturbative expansion (2.9) with splitting kernels up to NLO:

$$\begin{aligned} {}^{R_1R_2}V_{a_1a_2,a_0}(z_1,z_2,\alpha_s(\mu),L_y,L_\zeta) = {}& \frac{\alpha_s(\mu)}{2\pi}\, {}^{R_1R_2}V^{(1)}_{a_1a_2,a_0}(z_1,z_2) \\ & + \left(\frac{\alpha_s(\mu)}{2\pi}\right)^2 {}^{R_1R_2}V^{(2)}_{a_1a_2,a_0}(z_1,z_2,L_y,L_\zeta) + \mathcal{O}(\alpha_s^3)\,. \end{aligned} \tag{4.1}$$

The lowest-order splitting kernels are

$${}^{R_1R_2}V^{(1)}_{a_1a_2,a_0}(z_1,z_2) = c_{a_1a_2,a_0}(R_1R_2)\,\delta(1-z_1-z_2)\,P^{(0)}_{a_1a_0}(z_1)\,, \tag{4.2}$$

where $P^{(0)}_{a_1a_0}(z)$ is the leading-order DGLAP splitting function for PDFs without its distributional part at $z=1$. The colour factors $c_{a_1a_2,a_0}(R_1R_2)$ can for instance be found in section 4 of [48]. The NLO splitting kernels have the form

$$\begin{aligned} {}^{R_1R_2}V^{(2)}_{a_1a_2,a_0}(z_1,z_2,L_y,L_\zeta) = {}& {}^{R_1R_2}V^{[2,0]}_{a_1a_2,a_0}(z_1,z_2) + L_y\, {}^{R_1R_2}V^{[2,1]}_{a_1a_2,a_0}(z_1,z_2) \\ & + \frac{{}^{R_1}\gamma^{(0)}_J}{2}\left(L_yL_\zeta - \frac{L_y^2}{2} - \frac{\pi^2}{12}\right) {}^{R_1R_2}V^{(1)}_{a_1a_2,a_0}(z_1,z_2)\,, \end{aligned} \tag{4.3}$$

where

$${}^1\gamma^{(0)}_J = 0\,, \qquad {}^8\gamma^{(0)}_J = {}^S\gamma^{(0)}_J = {}^A\gamma^{(0)}_J = 2C_A \tag{4.4}$$

are the LO coefficients of the anomalous dimension for rapidity evolution.
The number n_f of active quark flavours for both the DPDs and the PDFs in the splitting formula (2.9) is taken according to the characteristic scale

$$\mu_y = b_0/y \tag{4.5}$$

of the splitting process. We include a heavy quark Q in the active flavours if $\mu_y > \gamma m_Q$ with a parameter $\gamma \sim 1$. Specifically, we take

$$n_f = \begin{cases} 3 & \text{for } \mu_y < \gamma m_c\,, \\ 4 & \text{for } \gamma m_c < \mu_y < \gamma m_b\,, \\ 5 & \text{for } \gamma m_b < \mu_y < \gamma m_t\,, \\ 6 & \text{for } \gamma m_t < \mu_y\,. \end{cases} \tag{4.6}$$

In the present section we set $\gamma = 1$; different values will be explored in section 5.3. We follow here the procedure of our earlier work [49] and take the scale μ_y as criterion for determining n_f. One could alternatively take the scale $\mu^*(y)$ from (2.17), in which case the lowest number of active flavours for DPDs would be $n_f = 4$ because $\mu^*(y) > \mu_{\min} = 2\,\text{GeV}$.
Starting from initial conditions with n_f from (4.6), DPDs with additional active flavours are obtained by matching, adding the heavy quark Q at $\mu_1 = m_Q$ for the first parton and at $\mu_2 = m_Q$ for the second one. The matching proceeds in the same way as for PDFs and is described in section 2.2 of [49] for DPDs in the colour singlet channel. The one-loop flavour matching kernels for colour octet DPDs are related to their colour singlet analogues as

$$ {}^{AA}A_{gg}^{Q(1)} = {}^{SS}A_{gg}^{Q(1)} = {}^{11}A_{gg}^{Q(1)} \,, \qquad {}^{AS}A_{gg}^{Q(1)} = {}^{SA}A_{gg}^{Q(1)} = 0 \tag{4.7} $$

for gluon-gluon transitions, where the relevant graphs are purely virtual, and as

$$ {}^{8R}A_{Qg}^{Q(1)} = c_{qg}(8R)\; {}^{11}A_{Qg}^{Q(1)} \qquad \text{with } R = A, S \tag{4.8} $$

for the transition from a gluon to a heavy quark, where the relevant graphs are purely real and involve the same colour factors

$$ c_{qg}(8A) = \sqrt{\frac{N_c^2}{2(N_c^2-1)}} \,, \qquad c_{qg}(8S) = \sqrt{\frac{N_c^2-4}{2(N_c^2-1)}} \tag{4.9} $$

as the leading-order DGLAP kernels ${}^{8R}P_{qg}^{(0)}$ [8, 51].

4.1 Parton kinematics and scales

In the following numerical studies we consider DPDs and DPD luminosities for x_i and $\bar{x}_i$ evaluated from (2.3) with

$$ \sqrt{s} = 14\,\text{TeV} \,, \qquad Y_1 = -Y_2 = Y \,, \tag{4.10} $$

where $Y \in [-4, 4]$. For the invariant masses of the produced systems, we take one of the two combinations

$$ M_1 = M_2 = 80\,\text{GeV} \tag{4.11} $$

or

$$ M_1 = 80\,\text{GeV} \,, \qquad M_2 = 10\,\text{GeV} \,. \tag{4.12} $$

An invariant mass of 80 GeV is for instance relevant for the production of a W or a dijet. An invariant mass of 10 GeV could be a typical scale for producing a pair of mini-jets, or for a J/Ψ with some transverse momentum. The corresponding parton momentum fractions are

$$ x_1 = \bar{x}_1 \approx \begin{cases} 7.1 \times 10^{-4} & \text{for } M_1 = 10\,\text{GeV} \\ 5.7 \times 10^{-3} & \text{for } M_1 = 80\,\text{GeV} \end{cases} \qquad \text{if } Y_1 = 0 \tag{4.13} $$

and

$$ x_1 \approx \begin{cases} 3.9 \times 10^{-2} \\ 3.1 \times 10^{-1} \end{cases} \qquad \bar{x}_1 \approx \begin{cases} 1.3 \times 10^{-5} & \text{for } M_1 = 10\,\text{GeV} \\ 1.0 \times 10^{-4} & \text{for } M_1 = 80\,\text{GeV} \end{cases} \qquad \text{if } Y_1 = 4 \,, \tag{4.14} $$

with corresponding values for x_2 and $\bar{x}_2$ obtained by symmetry considerations.
The factorisation scales in the double parton luminosities are always taken as

$$ \mu_i = M_i \,, \qquad i = 1, 2 \tag{4.15} $$

and the rapidity parameters as specified in (2.12). We will not vary the scales μ_i, which would be more instructive when DPDs are multiplied by parton-level cross sections evaluated at the same scales.

4.2 Splitting DPDs

The DPD splitting formula (2.9) is valid at small y, whereas DPD luminosities involve an integral over y up to infinity. Clearly, DPDs at large y are outside the reach of perturbative computations, and we need to model them. Following our strategy in earlier work [17, 62] we use the decomposition $F = F^{\text{spl}} + F^{\text{intr}}$ at all values of y, modelling both terms when y is large. For the splitting part, we do so by multiplying the perturbative form $F^{\text{spl, pt}}$ in (2.9) with a Gaussian factor:

$$\begin{aligned} &{}^{R_1 R_2}F^{\text{spl}}_{a_1 a_2}(x_1, x_2, y; \mu_{\text{init}}, \mu_{\text{init}}, \zeta_{\text{init}}) \\ &\quad = \exp\left[\frac{-y^2}{4 h_{a_1 a_2}}\right] {}^{R_1 R_2}F^{\text{spl, pt}}_{a_1 a_2}(x_1, x_2, y; \mu_{\text{init}}, \mu_{\text{init}}, \zeta_{\text{init}}) \,, \end{aligned} \tag{4.16}$$

where ζ_{init} is given in (2.22) and the parameters are taken from section 9.2.1 of [17]:

$$h_{gg} = 4.66\,\text{GeV}^{-2}\,, \qquad h_{gq} = h_{qg} = 5.86\,\text{GeV}^{-2}\,, \qquad h_{qq} = 7.06\,\text{GeV}^{-2}\,, \tag{4.17}$$

with equal values for quarks and antiquarks.
For the PDFs in the splitting formula we use the default LO or NLO sets of the MSHT20 fit [63], depending on the order at which we evolve the DPDs. The strong coupling used in these sets is

$$\alpha_s^{(n_f=5)}(m_Z) = 0.13 \qquad \text{at LO} \tag{4.18}$$

and

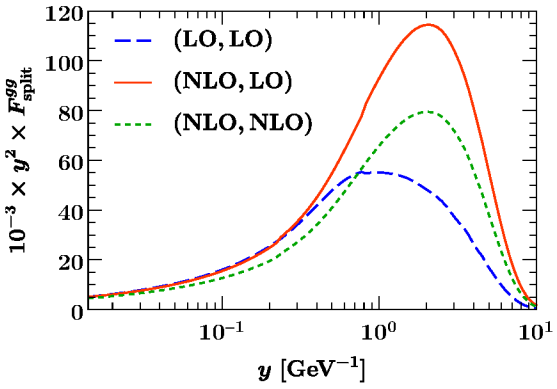
$$\alpha_s^{(n_f=5)}(m_Z) = 0.118 \qquad \text{at NLO.} \tag{4.19}$$

The PDFs are initialised at $\mu_0 = 1\,\text{GeV}$ from the LHAPDF interface [64], where the respective sets have the keys `MSHT20lo_as130` and `MSHT20nlo_as118`. The quark masses are

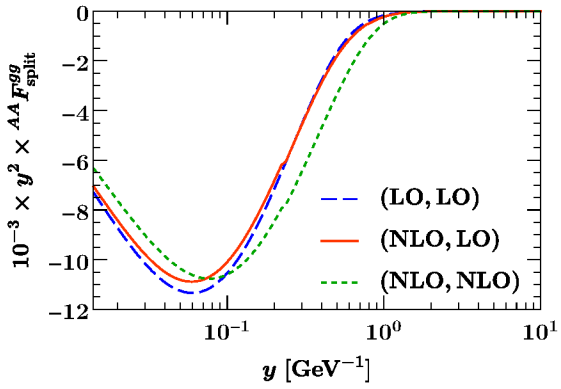
$$m_c = 1.4\,\text{GeV}\,, \qquad m_b = 4.75\,\text{GeV} \tag{4.20}$$

for both sets.

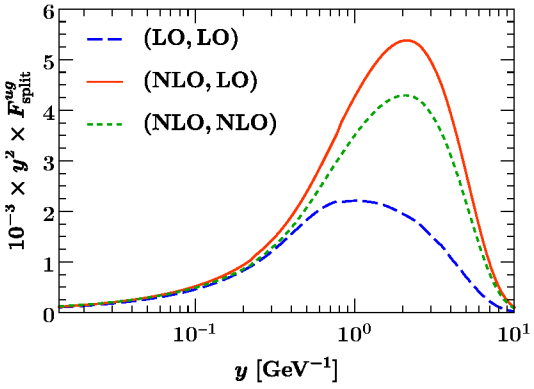
Comparison of different perturbative orders. In figures 3 and 4 we show DPDs for the kinematic setting with $M_1 = M_2 = 80\,\text{GeV}$ and $Y = 0$, which gives $x_1 = x_2 \approx 5.7 \times 10^{-3}$ according to (4.13). Figure 5 is for DPDs at $M_1 = M_2 = 80\,\text{GeV}$ and $Y = 4$, which gives $x_1 \approx 0.31$ and $x_2 \approx 10^{-4}$ according to (4.14). All DPDs are evolved from $\mu_{\text{init}} = \mu^*(y)$ to $\mu_i = M_i$ and from ζ_{init} in (2.22) to ζ in (2.12).
The lowest y value in the plots is $y_{\text{min}} = b_0/\mu_1$. At this point one has $\mu^*(y) \approx \mu_1$, so that there is essentially no evolution of the DPD after its initialisation. As y becomes bigger than y_{min}, the amount of evolution to higher scales increases. The DPDs in the figures are multiplied with y^2, so as to compensate the $1/y^2$ behaviour of splitting DPDs at the scale where they are initialised.
We observe that in the colour singlet channel, the scaled DPDs grow significantly with y as a consequence of evolution, which was already observed in [17]. A notable exception are $q\bar{q}$ distributions: we find strong evolution effects in the y dependence for asymmetric momentum fractions (figure 5a) but only mild ones in our setting with $x_1 = x_2$ (figure 4c). With the global factor $1/y^2$ included, $q\bar{q}$ colour singlet DPDs in this setting are hence more strongly concentrated at $y \sim y_{\text{cut}}$ than all other parton combinations. This has important consequences for double parton luminosities at central rapidities, as we shall see.



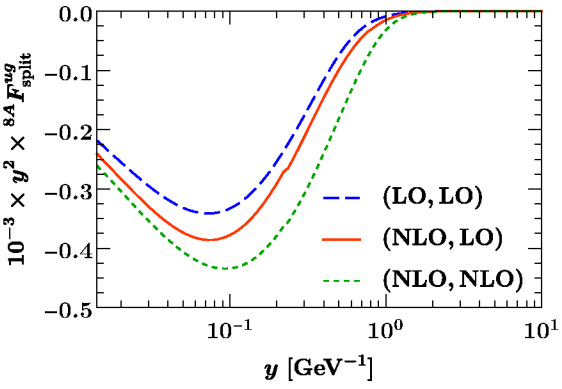
(a) gg, colour singlet



(b) gg, AA

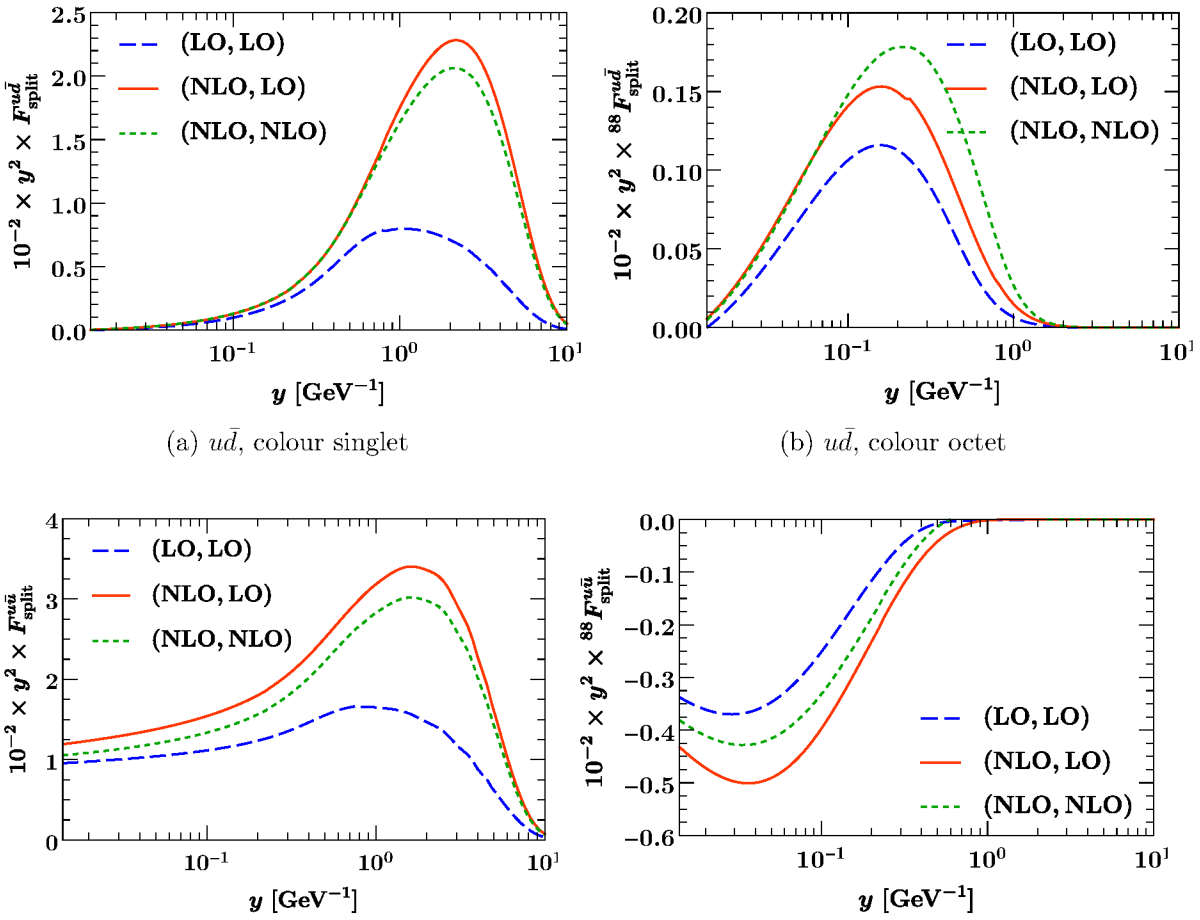


(c) ug, colour singlet



(d) ug, $8A$

Figure 3: Splitting DPDs for gg and ug at $x_1 = x_2 \approx 5.7 \times 10^{-3}$, evolved to scales $\mu_1 = \mu_2 = 80\,\text{GeV}$. The first specification of "LO" or "NLO" in the plots refers to the DPD splitting kernels, and the second one refers to the order of DGLAP evolution of the DPDs and of the PDFs in the splitting formula.



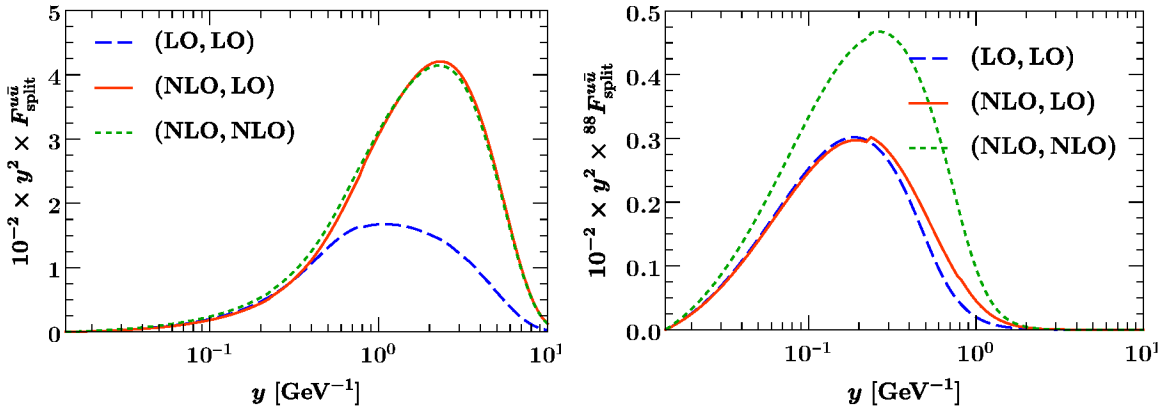
(a) $u\bar{d}$, colour singlet

(b) $u\bar{d}$, colour octet

(c) $u\bar{u}$, colour singlet, $x_1 = x_2$

(d) $u\bar{u}$, colour octet, $x_1 = x_2$

Figure 4: As figure 3, but for $u\bar{d}$ and $u\bar{u}$.



(a) $u\bar{u}$, colour singlet, $x_1 \gg x_2$

(b) $u\bar{u}$, colour octet, $x_1 \gg x_2$

Figure 5: As the bottom row of figure 4, but for asymmetric momentum fractions $x_1 \approx 0.31$ and $x_2 \approx 10^{-4}$.

Scaled colour octet DPDs behave very differently: at small y they are similar in magnitude to their colour singlet counterparts, whilst at large y they are strongly suppressed due to Collins-Soper evolution. Our study in [53] showed that this conclusion does not depend on the specific model for the non-perturbative part of the Collins-Soper kernel. For gluons in the decuplet or the 27 representation (not shown here), the suppression at large y is even stronger.

We now turn to the role of higher perturbative orders. The difference between the solid red and dashed blue curves in the figures shows the impact of using DPD splitting kernels at NLO instead of LO, with LO PDFs in the splitting formula in both cases:

- We find large differences at intermediate or large y for colour singlet DPDs, with the NLO result being bigger than the one at LO for all parton combinations considered.
- For colour octet distributions at y below $1\,\text{GeV}^{-1}$, the impact of NLO corrections to the DPD splitting is clearly visible. At larger y these distributions are negligibly small.

The difference between the solid red and dotted green lines in the figures shows the impact of switching from LO to NLO DGLAP evolution for DPDs and PDFs (along with changing the PDF set). In most cases, this impact is significant but more moderate than the change from LO to NLO in DPD splitting.

In the remainder of this paper we will focus on the impact of the NLO kernels for DPD splitting. We will therefore always evolve DPDs and PDFs at LO, even when the DPD splitting is computed at NLO. This may appear unusual, but we recall from section 2.2 that the combination of NLO splitting and LO evolution kernels is theoretically consistent. We also note that evolving DPDs at LO is adequate if one uses the DPS formula (2.1) with LO hard-scattering cross sections $\hat{\sigma}$. If the latter are evaluated at NLO, one gets convolution integrals rather than a simple multiplication of $\hat{\sigma}$ with double parton luminosities.

4.3 Double parton luminosities and subtraction terms

To investigate full double parton luminosities, we need not only the 1v1 but also the 1v2, 2v1, and 2v2 contributions and therefore require a model for the intrinsic part of the DPDs. In the colour singlet case we adopt our previous choice from [53], which is

$$
\begin{aligned}
&{}^{11}F^{\text{intr}}_{a_1 a_2}(x_1, x_2, y; \mu_{\text{init}}, \mu_{\text{init}}) \\
&\quad = n_{a_1 a_2}\, \frac{1}{4\pi h_{a_1 a_2}} \exp\left[\frac{-y^2}{4 h_{a_1 a_2}}\right] \frac{(1-x_1-x_2)^2}{(1-x_1)^2\,(1-x_2)^2}\, f_{a_1}(x_1; \mu_{\text{init}})\, f_{a_2}(x_2; \mu_{\text{init}})
\end{aligned}
\tag{4.21}
$$

with

$$
n_{a_1 a_2} = \begin{cases} 0 & \text{if } a_1 = a_2 = d - \bar{d}, \\ 1/2 & \text{if } a_1 = a_2 = u - \bar{u}, \\ 1 & \text{otherwise.} \end{cases}
\tag{4.22}
$$

Here the parton labels a_1 and a_2 have been transformed to the valence-sea basis, i.e. to the linear combinations $q - \bar{q}$ and $\bar{q}$ for each quark flavour. The parameters $h_{a_1 a_2}$ are the same as in (4.17).

In [53] we also explored models for F_{intr} in colour non-singlet channels. We found that the resulting double parton luminosities for the 1v2, 2v1, and 2v2 combinations are always considerably smaller than those for the pure splitting combination 1v1. This is easily understood: rapidity evolution strongly suppresses large y in the luminosity integral, and in the remaining region of small y the splitting part of DPDs is much larger than their intrinsic part. We

will hence not include the latter in the present study, also because modelling intrinsic colour non-singlet DPDs is poorly constrained and thus affected by huge uncertainties.
In the colour singlet channel there is no Collins-Soper suppression at large y, and depending on parton combinations and kinematic settings, both the intrinsic and the splitting parts of DPDs can be important in the overall double parton luminosities. We will see examples for this in the following.

Impact of the initial scale. Let us take a first look at double parton luminosities and their dependence on the initial scale μ_{init} at which DPDs are computed from the splitting formula (4.16) or from the model (4.21) for the intrinsic part of DPDs.
In the remainder of this paper we use the following shorthand notation for colour representations in double parton luminosities:

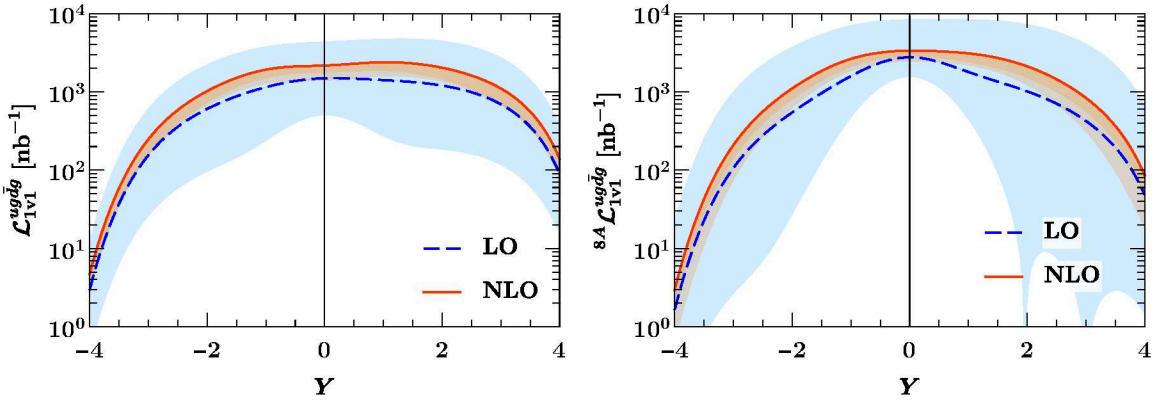
$$ {}^{11,11}\mathcal{L} \to \mathcal{L}\,, \qquad {}^{88,88}\mathcal{L} \to {}^{88}\mathcal{L}\,, \qquad {}^{8A,8A}\mathcal{L} \to {}^{8A}\mathcal{L}\,, \qquad {}^{AA,AA}\mathcal{L} \to {}^{AA}\mathcal{L}\,. \tag{4.23} $$

Note in particular that we simply omit colour labels to designate the colour singlet.
In figures 6 and 7 we show luminosities for the parton combinations $ug, \bar{d}g$ and $u\bar{d}, \bar{d}u$ at $M_1 = M_2 = 80\,\text{GeV}$, which are respectively relevant for $W+$ dijet and for W^+W^+ production. Note that splitting DPDs for the $u\bar{d}, \bar{d}u$ channel are generated only by NLO corrections to the splitting formula and by evolution from the initial to the final scale.
Comparing the 1v1, 1v2 + 2v1, and 2v2 contributions in the colour singlet (panels a, c, and d in both figures) we see that splitting and intrinsic contributions are comparable in size for both parton combinations. The change from LO to NLO splitting kernels leads to a noticeable upward shift of the central curves (except of course for the 2v2 contribution, where splitting DPDs do not contribute).
Comparing panels a and b in both figures 6 and 7 we see that the 1v1 contribution in the octet channels is of similar magnitude as for the singlet, and that the impact of including NLO splitting is similar as well.
The bands in the two figures show the result of varying the initial scale μ_{init} by a factor 2 up and down (with the restriction specified in (2.23)). For the intrinsic DPDs and hence for the 2v2 contribution this has almost no effect. This is expected for our ansatz (4.21): it is known that for small momentum fractions x_i there is little difference between the scale evolution of a single DPD and of the product of two PDFs [65]. For the splitting DPDs the situation is very different. With LO splitting kernels, the dependence of double parton luminosities on μ_{init} is huge, covering an order of magnitude or more in several cases. With NLO kernels this dependence is significantly reduced, although still visible in our logarithmic plots. We also note that at NLO the maximum of the luminosities is typically acquired near our central scale choice $\mu_{\text{init}} = \mu^*(y)$. We do not have an explanation for this quite remarkable finding.
In order to see details on a linear rather than logarithmic scale, we switch to plotting normalised double parton luminosities

$$ \begin{aligned} &{}^{R_1R_2,R_3R_4}\mathcal{R}_{a_1a_2,b_1b_2}(M_1, Y_1, M_2, Y_2; \mu_1, \mu_2, \nu, s) \\ &\quad = \frac{{}^{R_1R_2,R_3R_4}\mathcal{L}_{a_1a_2,b_1b_2}(M_1, Y_1, M_2, Y_2; \mu_1, \mu_2, \nu, s)}{f_{a_1}(x_1;\mu_1)\, f_{a_2}(x_2;\mu_2)\, f_{b_1}(\bar{x}_1;\mu_1)\, f_{b_2}(\bar{x}_2;\mu_2)} \\ &\quad = 2\pi \int_{y_{\text{cut}}}^{\infty} dy\, y\, \frac{{}^{R_1R_2}F_{a_1a_2}(x_1, x_2, y; \mu_1, \mu_2, \zeta)}{f_{a_1}(x_1;\mu_1)\, f_{a_2}(x_2;\mu_2)}\, \frac{{}^{R_3R_4}F_{b_1b_2}(\bar{x}_1, \bar{x}_2, y; \mu_1, \mu_2, \bar{\zeta})}{f_{b_1}(\bar{x}_1;\mu_1)\, f_{b_2}(\bar{x}_2;\mu_2)}\,, \end{aligned} \tag{4.24} $$

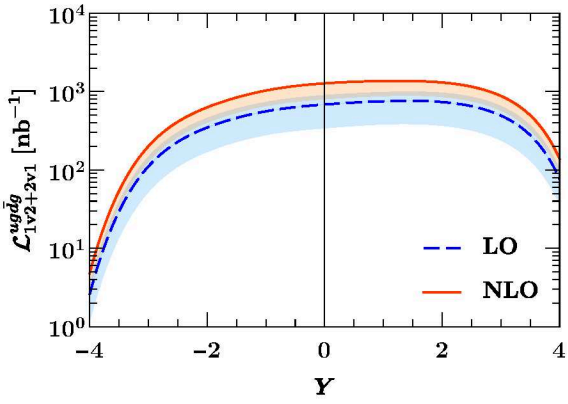
where the momentum fractions of the PDFs and DPDs are given in (2.3). This normalisation actually carries interesting physics information. If one makes the simplistic ansatz

$$ {}^{11}F_{a_1a_2}(x_1, x_2, y; \mu_1, \mu_2) \to f_{a_1}(x_1;\mu_1) f_{a_2}(x_2;\mu_2)\, G(y) \tag{4.25} $$

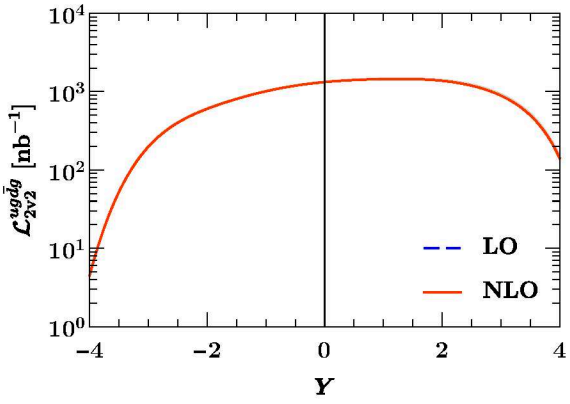


(a) $ug, \bar{d}g$, 1v1

(b) $ug, \bar{d}g$, $8A, 8A$, 1v1

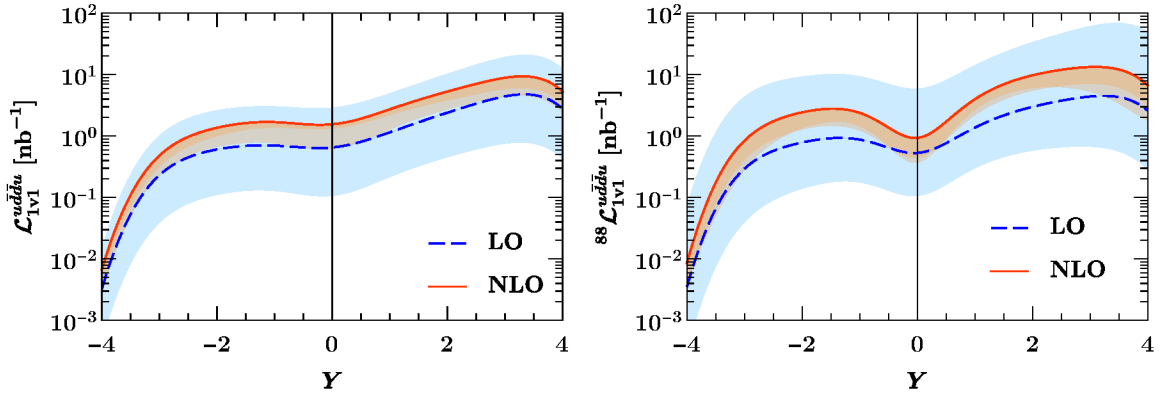


(c) $ug, \bar{d}g$, 1v2+2v1

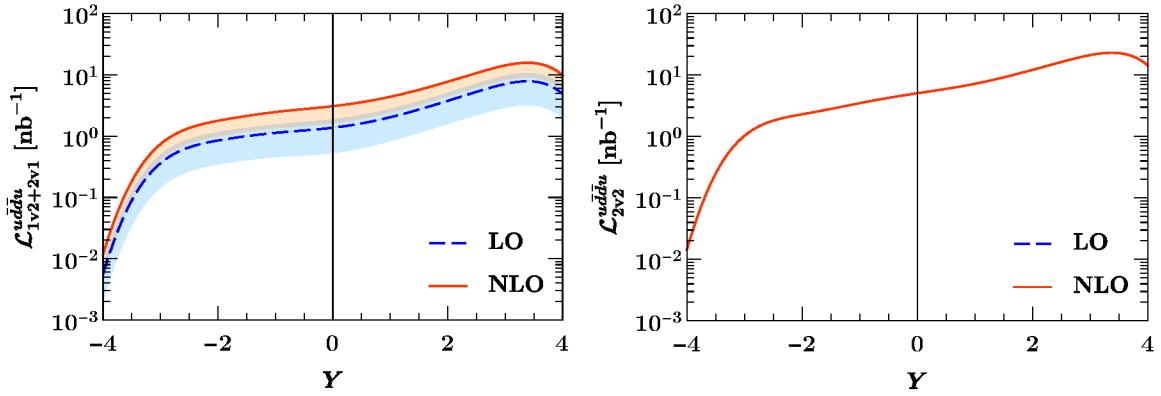


(d) $ug, \bar{d}g$, 2v2

Figure 6: Double parton luminosities for the product of DPDs $F_{ug}\, F_{\bar{d}g}$ with parton kinematics and scales set for $M_1 = M_2 = 80\,\text{GeV}$ as specified in section 4.1. The simplified notation (4.23) is used for colour representations. Bands correspond to a variation of the initial scale μ_{init} of the DPDs by a factor 2 as specified in (2.23). Here and in all following plots, the labels "LO" and "NLO" indicate the highest order of the DPD splitting kernels included in the calculation; with evolution being carried out at LO.



(a) $u\bar{d}, \bar{d}u$, 1v1 (b) $u\bar{d}, \bar{d}u$, colour octet, 1v1



(c) $u\bar{d}, \bar{d}u$, 1v2+2v1 (d) $u\bar{d}, \bar{d}u$, 2v2

Figure 7: The same as figure 6, but for the product of DPDs $F_{u\bar{d}}\, F_{\bar{d}u}$.

with a function $G(y)$ that varies over a typical hadronic scale, and if one neglects all colour non-singlet and all polarised contributions, then one gets the so-called pocket formula for the DPS cross section

$$\frac{d\sigma_{\text{DPS}}}{dM_1^2\, dY_1\, dM_2^2\, dY_2} \to \frac{1}{1+\delta_{A_1A_2}} \frac{1}{\sigma_{\text{eff}}} \frac{d\sigma_{A_1}}{dM_1^2\, dY_1} \frac{d\sigma_{A_2}}{dM_2^2\, dY_2} \tag{4.26}$$

with

$$\sigma_{\text{eff}}^{-1} = 2\pi \int_0^\infty dy\, y\, G^2(y)\,, \tag{4.27}$$

where σ_{A_i} is the cross section for producing the system A_i by single parton scattering. In (4.27) we have taken 0 instead of y_{cut} as lower integration boundary, which is a good approximation for small y_{cut}. Under the same assumptions, the ratio in (4.24) becomes

$${}^{11,11}\mathcal{R}_{a_1a_2,b_1b_2} \to \sigma_{\text{eff}}^{-1}\,. \tag{4.28}$$

The amount of variation of $\mathcal{R}$ with kinematic variables or parton channels hence indicates the extent to which the simple ansatz (4.25) works or fails. In turn, the size of $\mathcal{R}$ may be compared with the range $\sigma_{\text{eff}}^{-1} \sim 50 \ldots 200\,\text{b}^{-1}$ of values that have been extracted from measurements using the above pocket formula (see e.g. [66, 67] for compilations of σ_{eff}).

Impact of the subtraction term. In the following plots we compare the 1v1 part of the double parton luminosity with the full double counting subtraction term $\mathcal{L}_{\text{sub}}$ and with the partial contribution $\mathcal{L}_{\text{sub,SPS}}$ in (3.14). We recall that the latter is concentrated in the region $y \gg y_{\text{cut}}$, where it is designed to compensate contributions from SPS graphs that have an overlap with DPS.

In figure 8 we show this comparison for equal scales $M_1 = M_2 = 80\,\text{GeV}$ and for parton combinations that allow splitting already at LO . The total subtraction term is important in all cases; for the parton combination $u\bar{u}, \bar{d}d$ at central rapidities it is nearly equal to $\mathcal{L}_{\text{1v1}}$. We always find $|\mathcal{L}_{\text{sub,SPS}}| \ll |\mathcal{L}_{\text{sub}}|$, i.e. $\mathcal{L}_{\text{sub}} \approx \mathcal{L}_{\text{sub,DPS}}$, which reflects that the product of two splitting DPDs is most important at small y. In some cases $\mathcal{L}_{\text{sub,SPS}}$ is comparable in size to the difference $\mathcal{L}_{\text{1v1}} - \mathcal{L}_{\text{sub}}$ and in this sense has an impact on the overall cross section. This holds for $u\bar{u}, \bar{d}d$ and also for the four-gluon case at NLO. In the other cases shown in figure 8, $\mathcal{L}_{\text{sub,SPS}}$ is of minor importance.

For parton combinations where splitting only starts at NLO, we find that the subtraction term evaluated at NLO is generally of minor size, although not necessarily negligible as seen for $u\bar{d}, \bar{d}u$ in figure 9. At LO (not shown here) the corresponding subtraction terms are very close to zero, since the splitting DPDs vanish at the initial scale and there is essentially no evolution in this case.

We now turn to the case of unequal scales with $M_1 = 80\,\text{GeV}$ and $M_2 = 10\,\text{GeV}$. The difference between the integrands for the DPS and SPS subtraction now becomes important, and we need to specify the interpolating function in (3.11). As a default we take $y_\rho = y_{\text{cut}}$ and

$$r_1 = 2\,, \qquad\qquad r_2 = 4\,, \tag{4.29}$$

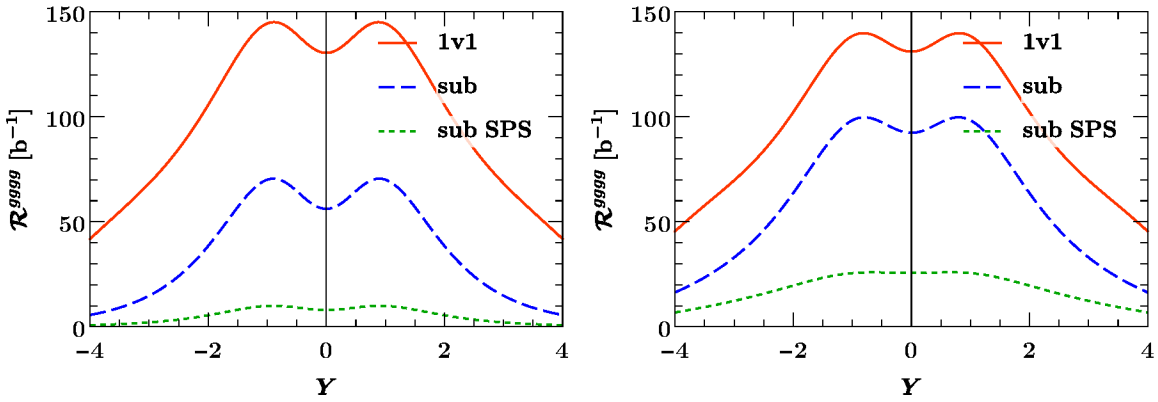
which corresponds to a rather quick changeover from $I_{\text{sub,DPS}}(y)$ to $I_{\text{sub,SPS}}(y)$ as a function of y. We make this choice in order to see how important $\mathcal{L}_{\text{sub,SPS}}$ can plausibly be, since with larger values of r_2 that term will decrease even further. We also computed double parton luminosities with the choices $r_1 = 2, r_2 = 6$ and $r_1 = 3, r_2 = 5$ and find that $\mathcal{L}_{\text{sub}}$ and $\mathcal{L}_{\text{1v1}} - \mathcal{L}_{\text{sub}}$ barely change for the parton combinations considered in our study.

Comparing figure 10 with the first row in figure 8, we see that in the four-gluon channel the subtraction term with one small and one large scale is less important than with two large scales. We find a similar situation for the combination $ug, \bar{d}g$ (not shown here).

Size of the subtracted 1v1 term. According to the master formula (3.1), the full cross section of a process involves the subtracted luminosity $\mathcal{L}_{\text{1v1}} - \mathcal{L}_{\text{sub}}$ plus the partial luminosities for 1v2, 2v1, and 2v2. In figures 11 to 13 we compare these three terms for different kinematics and parton combinations. Figure 11 shows that the relative importance of the different contributions can change to some extent when going from LO to NLO splitting.

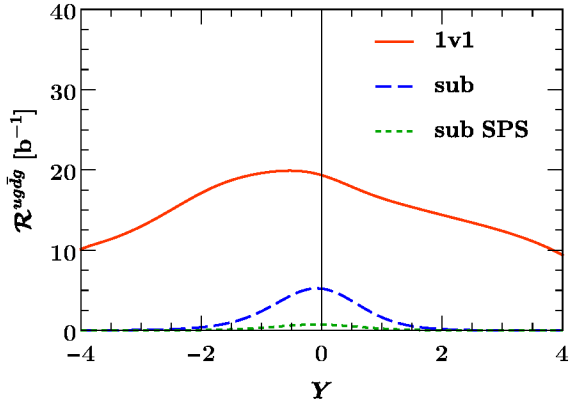
In almost all considered channels we find that the contributions with or without splitting DPDs are of comparable magnitude and thus all play a role in the final result.

The only exception is the parton combination $u\bar{u}, \bar{d}d$ in figure 12a, where for central rapidities $|Y| \lesssim 2$ the 1v1 contribution is completely dominant, even after double counting subtraction. This is remarkable, because the subtracted 1v1 term in this case is already the result of a substantial cancellation between $\mathcal{L}_{\text{1v1}}$ and $\mathcal{L}_{\text{sub}}$, as we saw in the bottom row of figure 8. The same is observed for other channels with a $q\bar{q}$ pair in each DPD. An important consequence is that at central rapidities these luminosities have very little uncertainty due to modelling the intrinsic part of DPDs, which is not the case for the other channels in figures 11 to 13. Since the $q\bar{q}$ splitting DPDs are concentrated at small y, there is also less dependence on how the perturbative splitting form is extended to the non-perturbative y region.

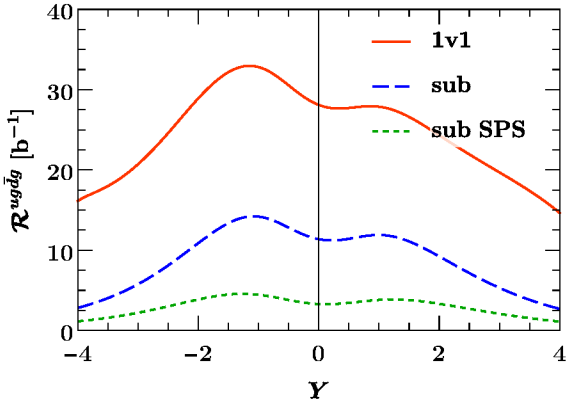


(a) gg, gg, LO

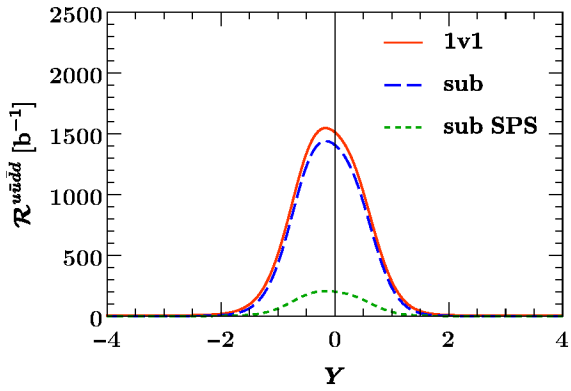
(b) gg, gg, NLO



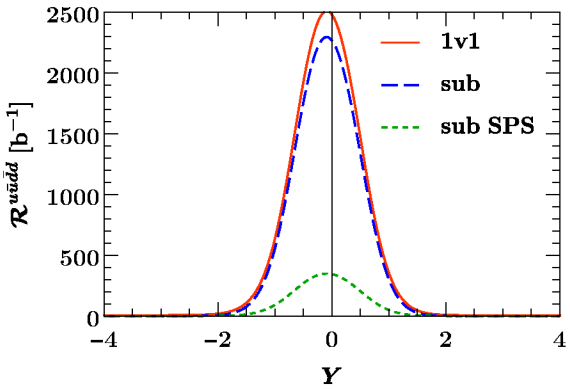
(c) $ug, \bar{d}g$, LO



(d) $ug, \bar{d}g$, NLO

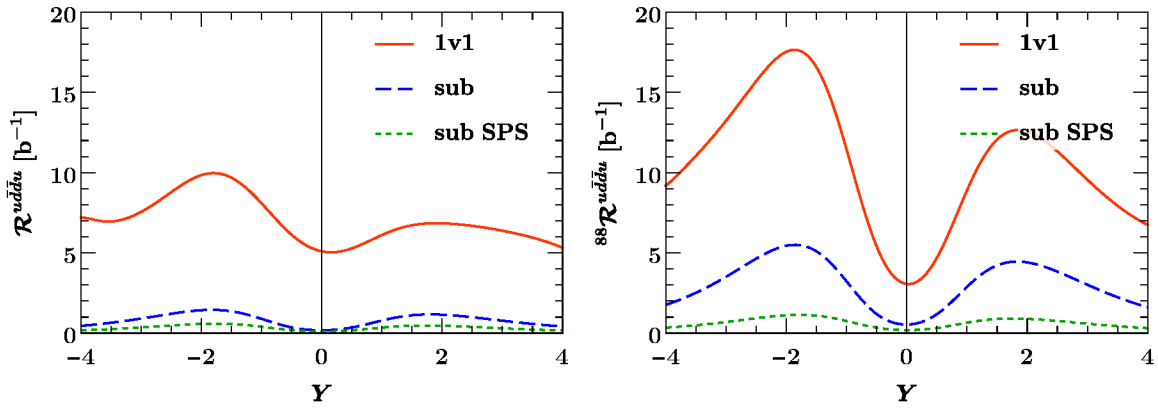


(e) $u\bar{u}, \bar{d}d$, LO



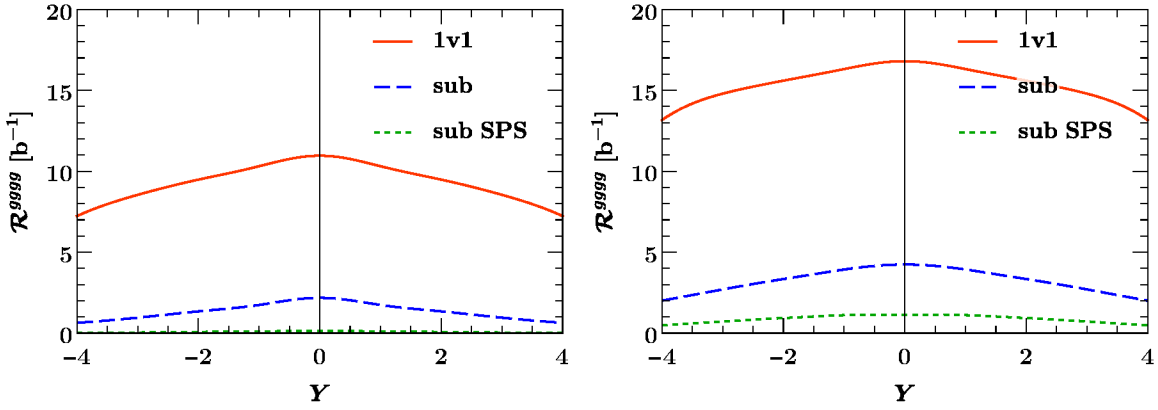
(f) $u\bar{u}, \bar{d}d$, NLO

Figure 8: Normalised double parton luminosities for selected parton combination, with parton kinematics and scales set for $M_1 = M_2 = 80\,\text{GeV}$ as specified in section 4.1. Shown are the 1v1 component $\mathcal{L}_{1\text{v}1}$ of the luminosity for DPS, the luminosity $\mathcal{L}_{\text{sub,SPS}}$ for the double counting subtraction term, and its large-y part $\mathcal{L}_{\text{sub,SPS}}$ in the decomposition (3.14). All luminosities are normalised as specified in (4.24).



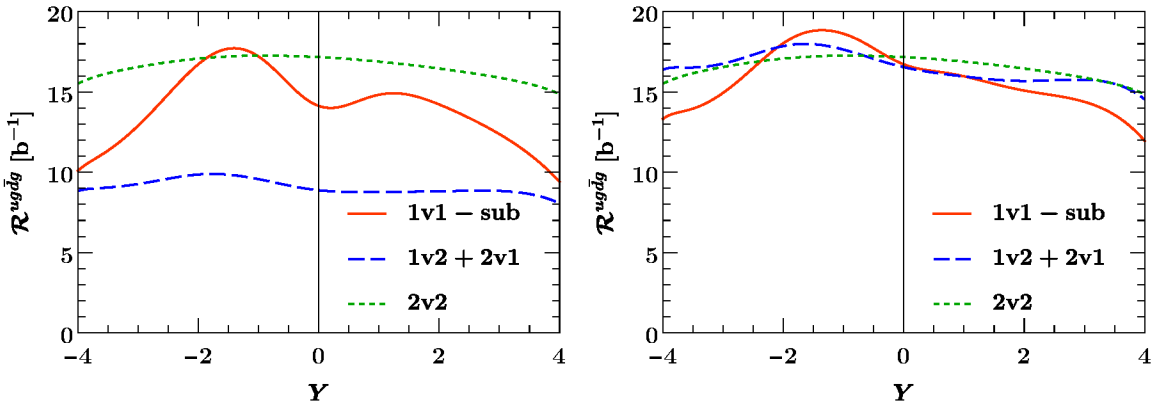
(a) $u\bar{d}, \bar{d}u$, NLO, colour singlet (b) $u\bar{d}, \bar{d}u$, NLO, colour octet

Figure 9: As figure 8, but for a parton combination where the DPDs splitting formula starts at NLO.



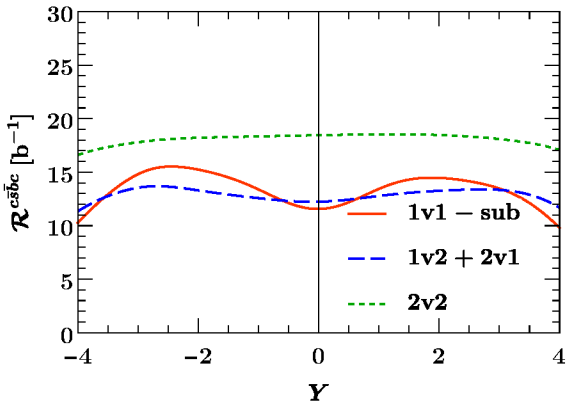
(a) gg, gg, LO (b) gg, gg, NLO

Figure 10: As the first row of figure 8, but for kinematics with $M_1 = 80\,\text{GeV}$ and $M_2 = 10\,\text{GeV}$.

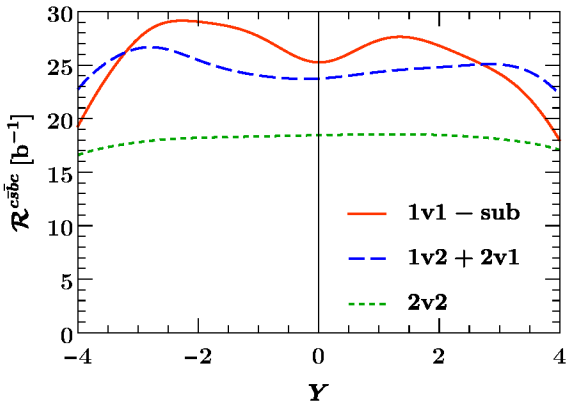


(a) $ug, \bar{d}g$, LO

(b) $ug, \bar{d}g$, NLO

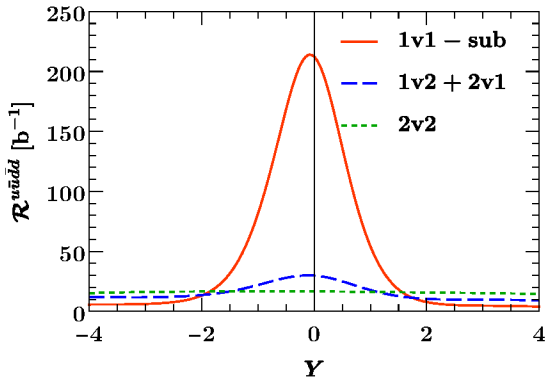


(c) $c\bar{s}, \bar{b}c$, LO

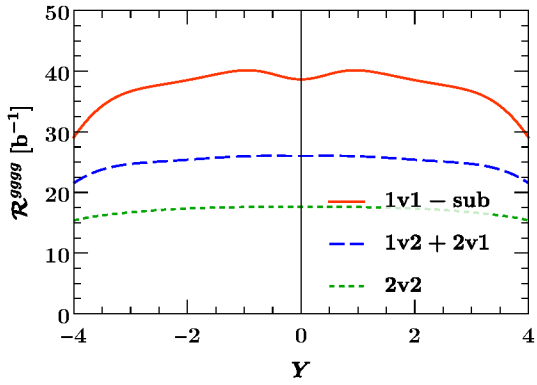


(d) $c\bar{s}, \bar{b}c$, NLO

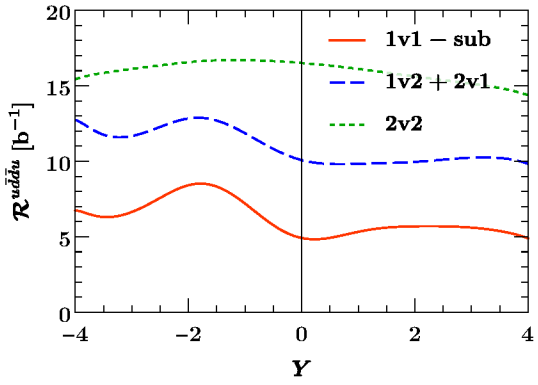
Figure 11: Different contributions to normalised double parton luminosities in kinematics with $M_1 = M_2 = 80\,\text{GeV}$. The sum of the three contributions gives $\mathcal{R}_{\text{DPS}} - \mathcal{R}_{\text{sub}}$.



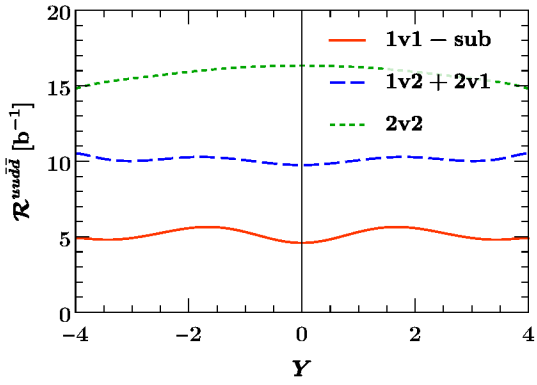
(a) $u\bar{u}, \bar{d}d$, NLO



(b) gg, gg, NLO

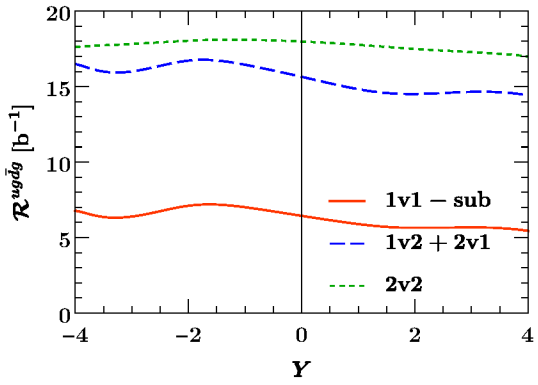


(c) $u\bar{d}, \bar{d}u$, NLO

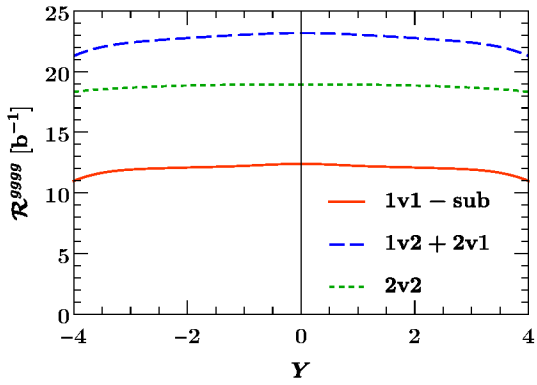


(d) $uu, \bar{d}\bar{d}$, NLO

Figure 12: As the right panels in figure 11, but for other parton combinations.



(a) $ug, \bar{d}g$, NLO



(b) gg, gg, NLO

Figure 13: Different contributions to normalised double parton luminosities in kinematics with $M_1 = 80\,\text{GeV}$ and $M_1 = 10\,\text{GeV}$. The corresponding plots for equal scales are shown in figures 11b and 12b.

Comparing the plots for unequal scales $M_1 = 80\,\text{GeV}$ and $M_2 = 10\,\text{GeV}$ in figure 13 with their counterparts for $M_1 = M_2 = 80\,\text{GeV}$ in figures 11b and 12b, we observe that the subtracted 1v1 part is less important for unequal scales.

4.4 Subtracted double parton luminosities and their scale dependence

Let us finally study the dependence of the subtracted luminosities $\mathcal{L}_{\text{DPS}} - \mathcal{L}_{\text{sub}}$ on the choice of the scales $\mu_{\text{init}}(y)$ and ν. This dependence propagates into the overall cross section, albeit with possible cancellations in the sum over different parton and colour channels. Apart from such cancellations, the uncertainty due to choosing a scale directly translates into an uncertainty of cross section predictions.

The panels on the left of the following figures show bands resulting from varying $\mu_{\text{init}}(y)$ around the central choice $\mu^*(y)$ by a factor 2, with a lower bound μ_{min} as specified in (2.23). The value of ν is fixed to its default choice $\min(\mu_1, \mu_2)$ in this case.

The panels on the right of the same figures show bands resulting from varying the cutoff parameter ν by a factor 2 up and down while keeping the interpolation function $\rho(y)$ fixed. The initial scale for DPDs in $\mathcal{L}_{\text{DPS}}$ is kept at its default $\mu_{\text{init}}(y) = \mu^*(y)$ in this case, but note that the initial scale of DPDs in the subtraction term $\mathcal{L}_{\text{sub,DPS}}$ is always $\mu_{\text{sub,init}} = \mu_{\text{init}}(y_{\text{cut}})$ and hence changes with y_{cut}. If one prefers to keep $\mu_{\text{sub,init}}$ fixed, one can vary y_{cut} and μ_{init} simultaneously, or one can choose $\mu_{\text{sub,init}} = \min(\mu_1, \mu_2)$ independent of y_{cut}. We will not pursue these possibilities here.

If follows from equation (2.28) that the μ_{init} dependence of subtracted double parton luminosities decreases parametrically with the order at which the splitting DPDs are computed. We have shown in section 3.1 that the dependence of subtracted luminosities on ν is beyond the perturbative order of the calculation and hence also decreases parametrically when that order is raised. Note that the same is *not* true (and not meant to be true) for the ν dependence of $\mathcal{L}_{\text{DPS}}$ without the double counting subtraction. In the following we investigate to which extent the parametric decrease in powers of α_s shows up at the quantitative level.

Given the practical importance of subtracted luminosities and their μ_{init} and ν dependence, the following plots will cover a larger number of channels than the examples shown so far. Figures 14 to 17 are for equal scales $M_1 = M_2 = 80\,\text{GeV}$, and figures 18 to 19 are for unequal scales $M_1 = 80\,\text{GeV}$, $M_2 = 10\,\text{GeV}$.

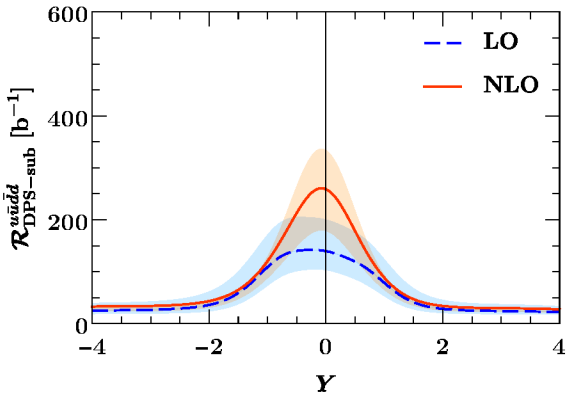
Subsequent pairs of figures show colour singlet and colour octet channels for the same parton combinations. We typically find that the central values for octet channels are much smaller than for the colour singlet, although not necessarily negligible. Only in the four-gluon case at equal scales do we find colour octet contributions of size similar to the singlet, as seen in the top panels of figures 16 and 17.

Not shown here are the colour combinations with symmetric instead of antisymmetric octets. We generally find that the SS, SS luminosity for four gluons is very similar in magnitude to AA, AA. The same holds for $8S, 8S$ compared with $8A, 8A$ for the $ug, \bar{d}g$ combination. The mixed-symmetric AS, AS luminosity for four gluons is tiny compared to AA, AA. At the starting scale $\mu^*(y)$, the smallness of AS splitting DPDs was already noticed in [48].

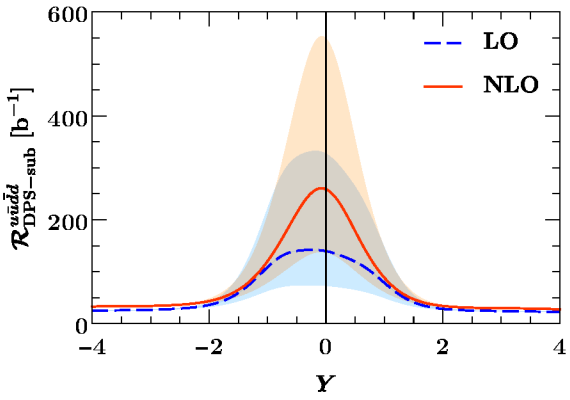
For the central values of subtracted luminosities, we find in most cases that the NLO corrections to DPD splitting are appreciable at the level of several 10%. In a few channels they reach a factor around 2, notably for $u\bar{u}, \bar{d}d$ at central rapidities in the colour singlet and for gg, gg in the colour octet case.

For scale variations we see a clear difference between $u\bar{u}, \bar{d}d$ and the other parton combinations shown here. In the channels other than $u\bar{u}, \bar{d}d$, we find the following:

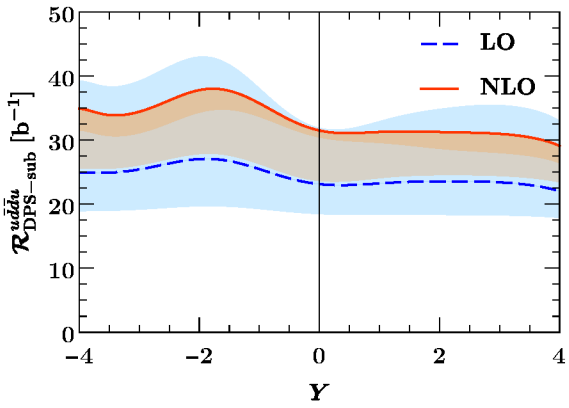
- The variation with μ_{init} significantly decreases from LO to NLO. At NLO, the scale variation is highly asymmetric, with the central curve being close to the edge of the



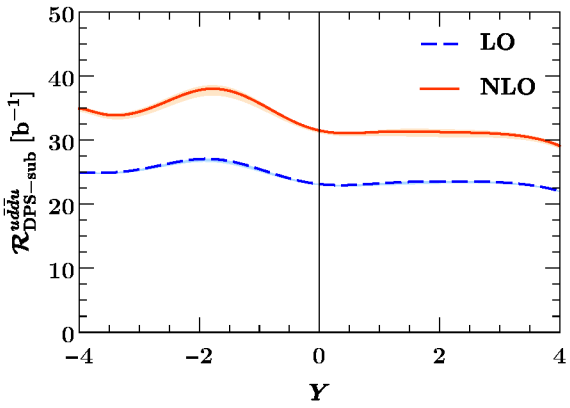
(a) μ_{init} variation, $u\bar{u}, \bar{d}d$



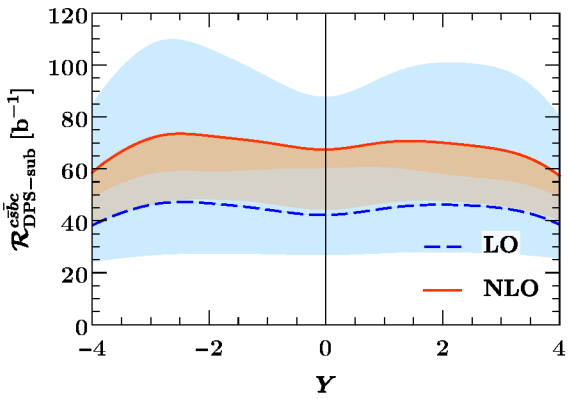
(b) ν variation, $u\bar{u}, \bar{d}d$



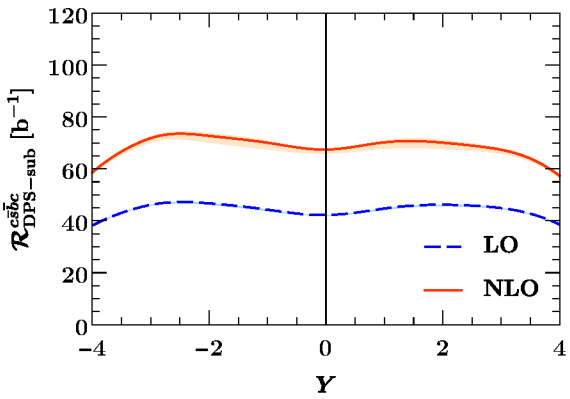
(c) μ_{init} variation, $u\bar{d}, \bar{d}u$



(d) ν variation, $u\bar{d}, \bar{d}u$

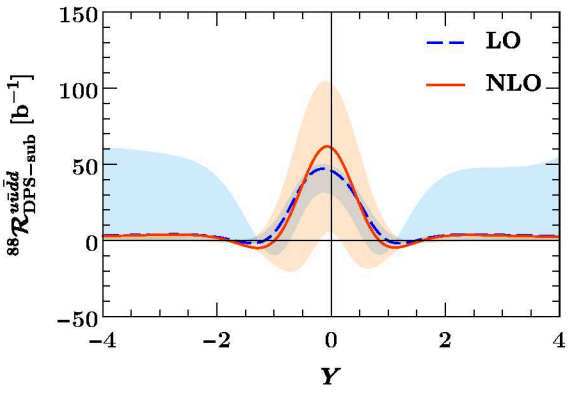


(e) μ_{init} variation, $c\bar{s}, \bar{b}c$

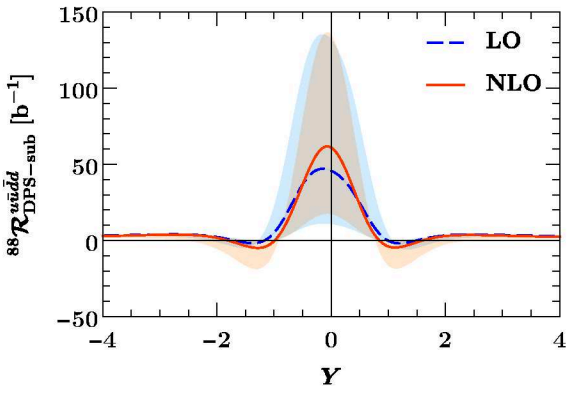


(f) ν variation, $c\bar{s}, \bar{b}c$

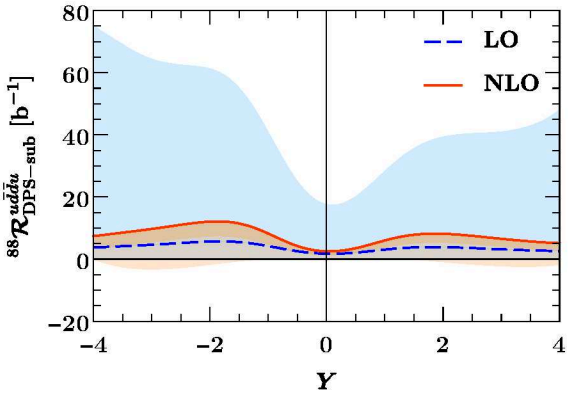
Figure 14: Subtracted double parton luminosities and their scale dependence in the colour singlet for $M_1 = M_2 = 80\,\mathrm{GeV}$. Further parton combinations are shown in figure 16.



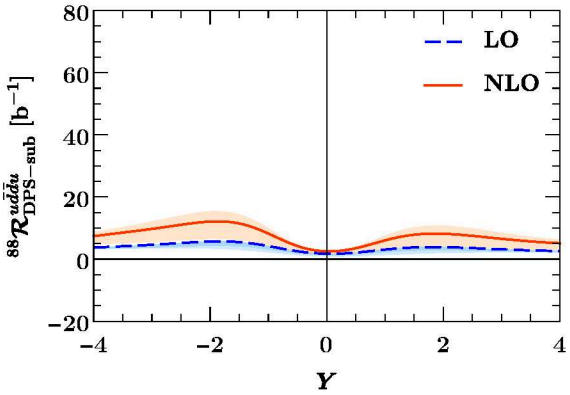
(a) μ_{init} variation, $u\bar{u}, \bar{d}d$, colour octet



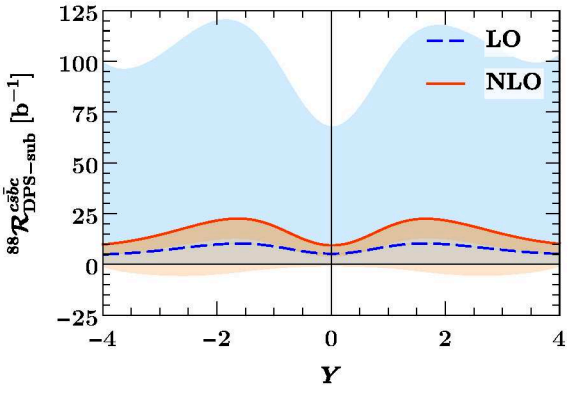
(b) ν variation, $u\bar{u}, \bar{d}d$, colour octet



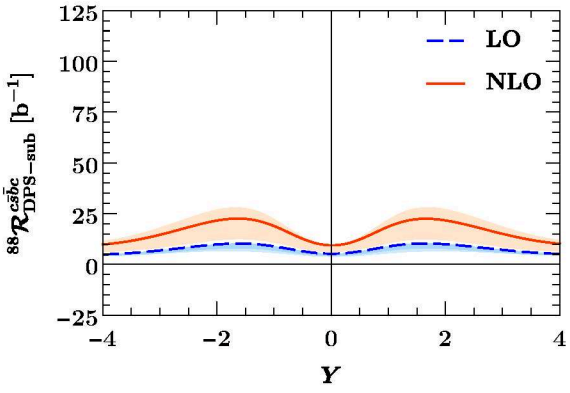
(c) μ_{init} variation, $u\bar{d}, \bar{d}u$, colour octet



(d) ν variation, $u\bar{d}, \bar{d}u$, colour octet

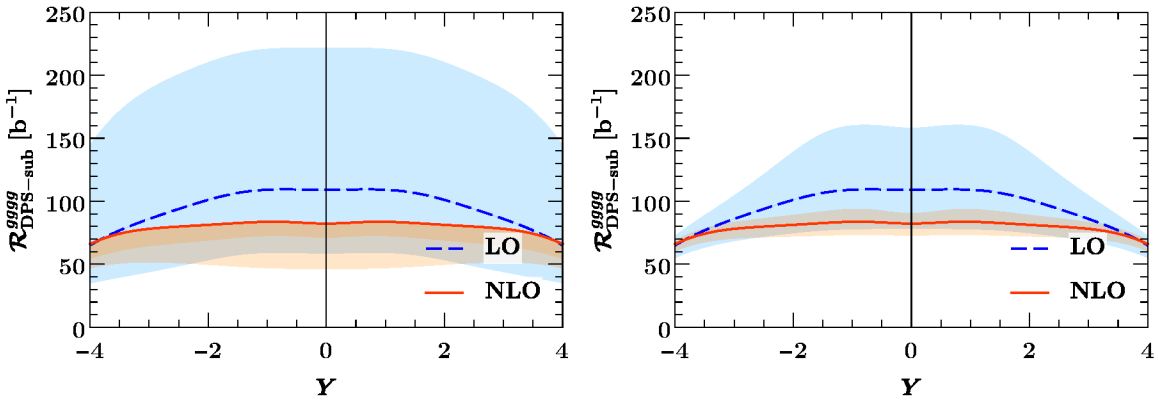


(e) μ_{init} variation, $c\bar{s}, \bar{b}c$, colour octet



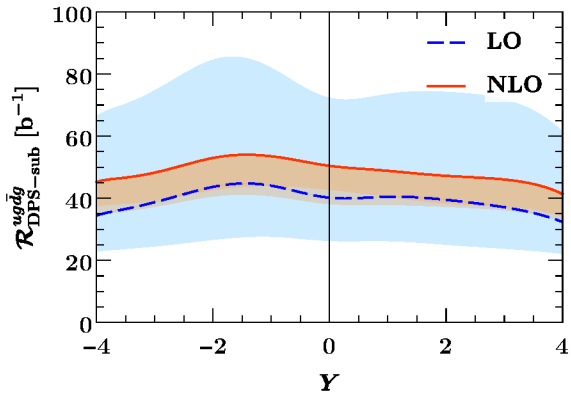
(f) ν variation, $c\bar{s}, \bar{b}c$, colour octet

Figure 15: As figure 14, but for colour octet channels. Further parton combinations are shown in figure 17.

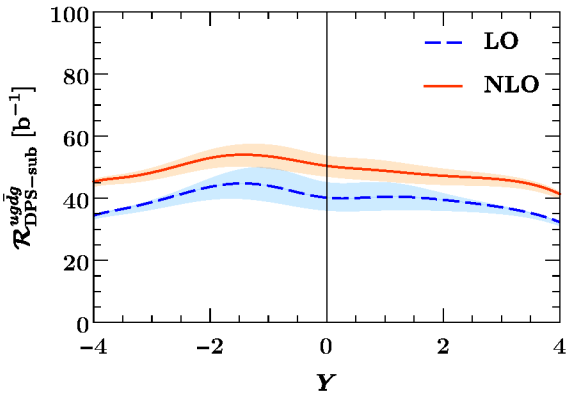


(a) μ_{init} variation, gg, gg

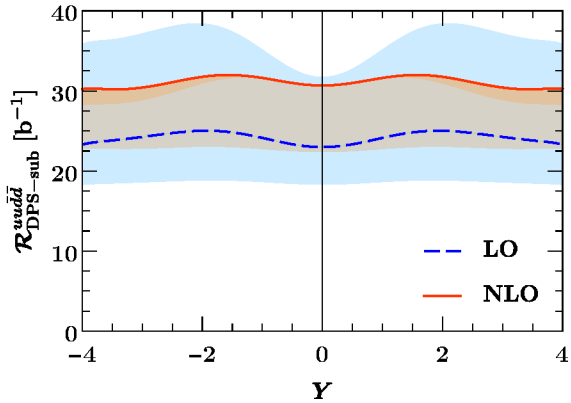
(b) ν variation, gg, gg



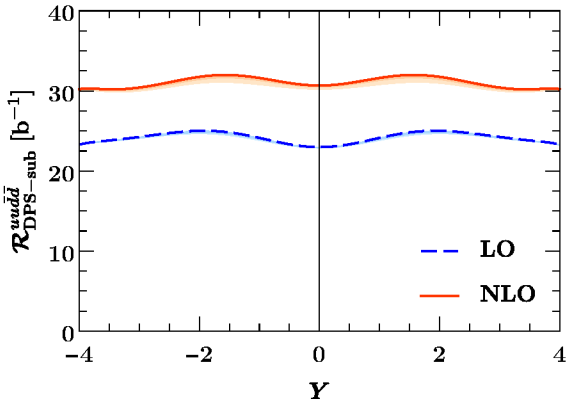
(c) μ_{init} variation, $ug, \bar{d}g$



(d) ν variation, $ug, \bar{d}g$

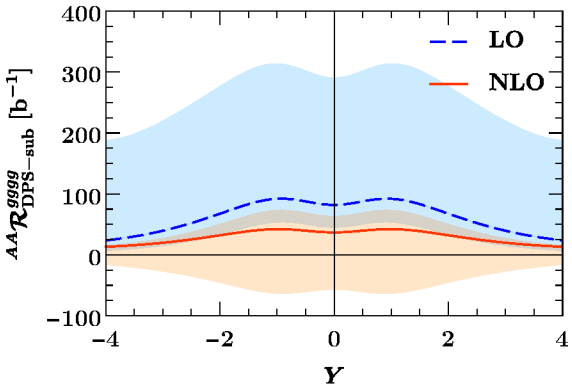


(e) μ_{init} variation, $uu, \bar{d}\bar{d}$

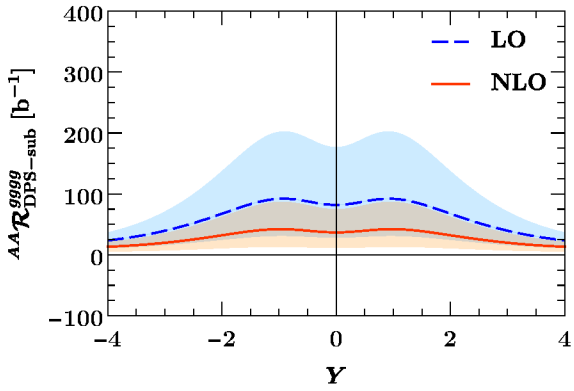


(f) ν variation, $uu, \bar{d}\bar{d}$

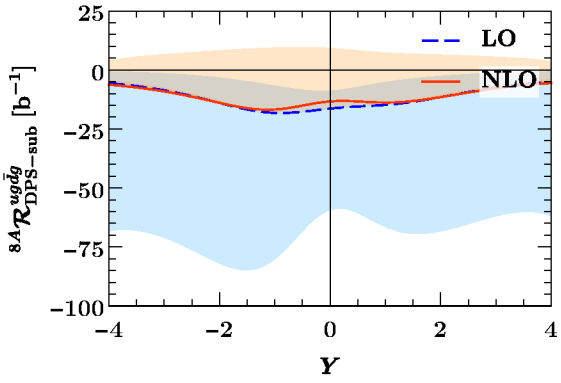
Figure 16: Further subtracted double parton luminosities and their scale dependence in the colour singlet for $M_1 = M_2 = 80\,\mathrm{GeV}$.



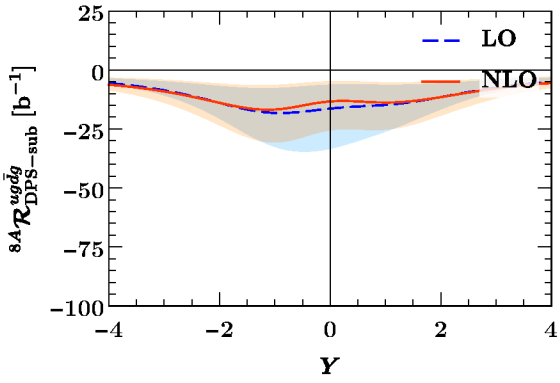
(a) μ_{init} variation, gg, gg, AA, AA



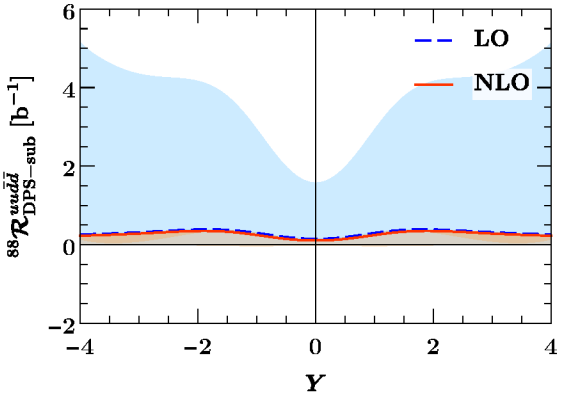
(b) ν variation, gg, gg, AA, AA



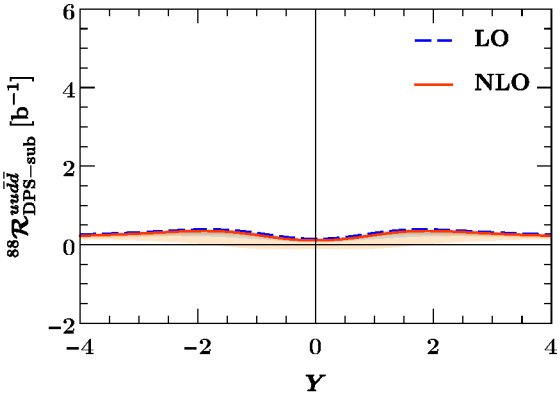
(c) μ_{init} variation, $ug, \bar{d}g$, $8A, 8A$



(d) ν variation, $ug, \bar{d}g$, $8A, 8A$

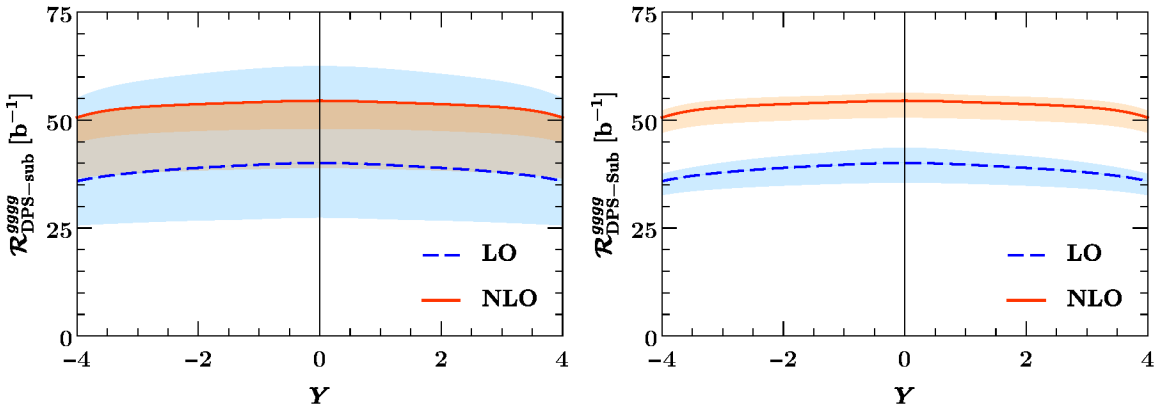


(e) μ_{init} variation, $uu, \bar{d}\bar{d}$, colour octet



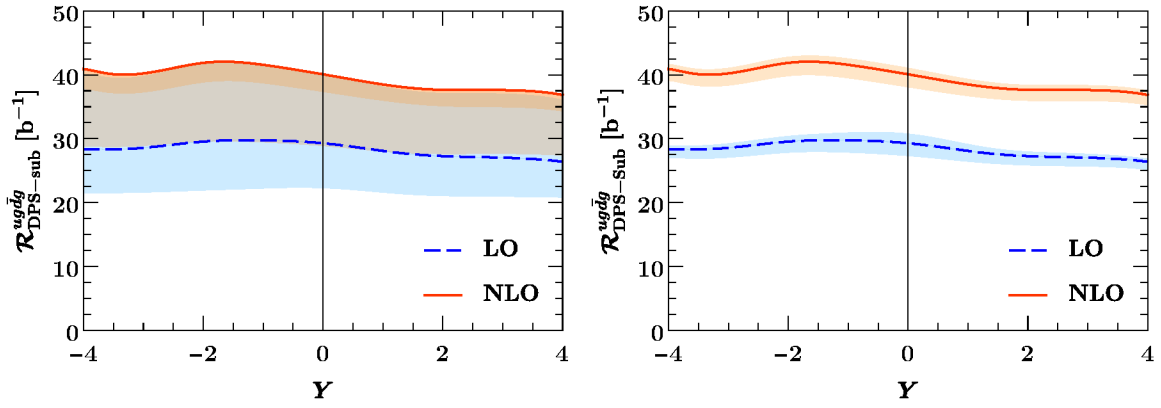
(f) ν variation, $uu, \bar{d}\bar{d}$, colour octet

Figure 17: As figure 16, but for colour octet channels.



(a) μ_{init} variation, gg, gg

(b) ν variation, gg, gg



(c) μ_{init} variation, $ug, \bar{d}g$

(d) ν variation, $ug, \bar{d}g$

Figure 18: As figure 16, but for asymmetric scales $M_1 = 80\,\text{GeV}$ and $M_2 = 10\,\text{GeV}$.

band of variation.

Compared with the central value, the scale variation at NLO is moderate for the colour singlet, whereas in the octet channels it is still large.

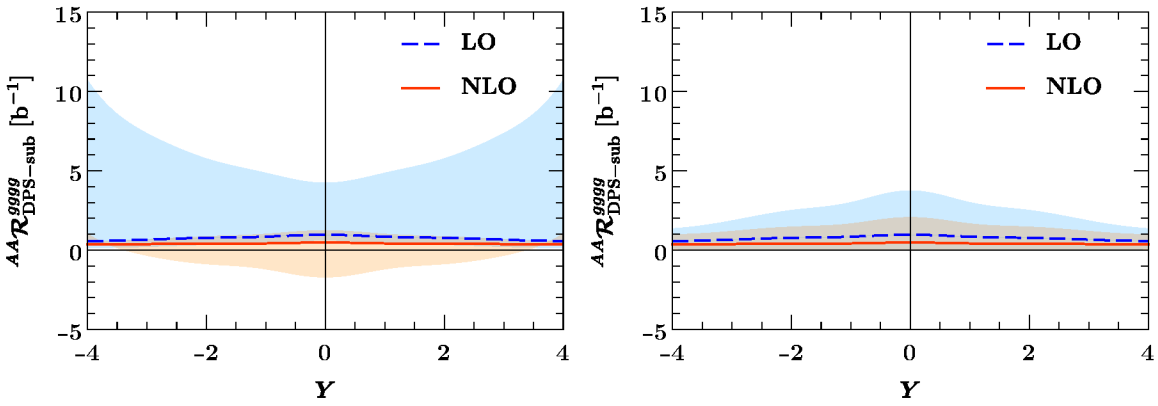
- The variation with ν is typically much less than the one with μ_{init}.

  In colour singlet channels, the ν dependence either decreases from LO to NLO, or it is very small at both orders. For colour octet channels there is no uniform picture, with the variation decreasing from LO to NLO in some channels and increasing in others.

- With the exception of the four-gluon luminosity at equal scales, the variation with μ_{init} and ν in the octet channel is moderate compared with the size of the colour singlet result.

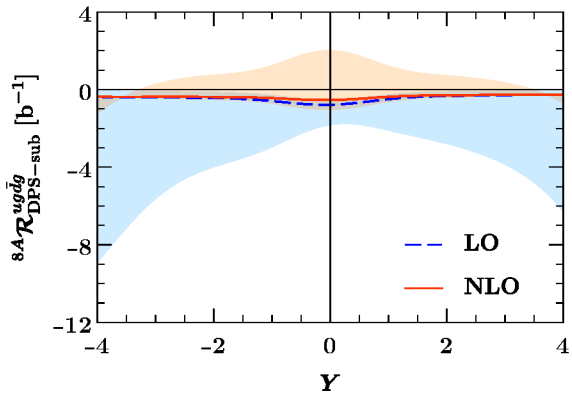
At high rapidity, the behaviour of the $u\bar{u}, \bar{d}d$ luminosity is quite similar to the one just discussed. By contrast, for central rapidities we find that

- in the colour singlet channel, the variation with μ_{init} and ν is very similar at LO and NLO. At $Y = 0$ we have a μ_{init} variation of 30 to 40% and a ν variation by about a factor of 2, with both variations being approximately symmetric around the central value.

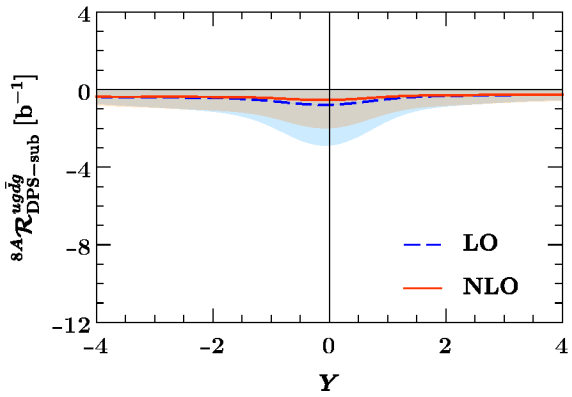


(a) μ_{init} variation, gg, gg, AA, AA

(b) ν variation, gg, gg, AA, AA



(c) μ_{init} variation, $ug, \bar{d}g$, $8A, 8A$



(d) ν variation, $ug, \bar{d}g$, $8A, 8A$

Figure 19: As figure 18, but for colour octet channels.

In the colour octet channel, the ν variation is similar at both orders, whilst the μ_{init} variation is much larger at NLO than at LO. Compared with the central value of the colour singlet, the variation of the colour octet luminosity remains moderate: at $Y = 0$ it is at level of about 25%.

We recall that for central rapidities the cancellation between the 1v1 luminosity and its subtraction term is by far the largest for $u\bar{u}, \bar{d}d$, see the bottom row in figure 8. This is in line with our earlier observation that the DPDs for $q\bar{q}$ are most strongly concentrated at $y \sim y_{\text{cut}}$, where the difference between the DPDs in the 1v1 term and its subtraction is small by design. Given this strong cancellation, a high sensitivity of the result to scale variations is perhaps not too surprising.

Whilst the relative uncertainty due to scale choices is most pronounced for $q\bar{q}$ channels at central rapidities, the uncertainty due to modelling the intrinsic DPDs is smallest in this case, as we discussed in the context of figure 12a.

5 Splitting with massive quarks

So far we have neglected the masses of the heavy quarks c, b, t in the computation of DPD splitting kernels. For a given flavour Q this is adequate if the characteristic scale $1/y$ of the splitting is much larger than the quark mass m_Q. Conversely, a heavy quark decouples from the splitting process if its mass is much larger than $1/y$. In the region where $1/y \sim m_Q$, the quark mass should be taken into account when computing the splitting process in the DPD. For colour singlet DPDs we gave a detailed analysis of this in [49]. In the present section we generalise the results from sections 3, 5.5, and 6.4 in [49] to colour non-singlet channels, which is straightforward. We then proceed to a numerical analysis of evolved splitting DPDs that include quark mass effects.

For each heavy flavour Q, we keep the quark mass m_Q finite in the region $\alpha m_Q < \mu_y < \beta m_Q$, where we write again $\mu_y = b_0/y$. In that region we replace the massless splitting formula (2.9) by

$$ {}^{R_1R_2}F^{n_F+1}_{a_1a_2}(x_1, x_2, y; \mu, \mu, \zeta) = \frac{1}{\pi\, y^2} \sum_{a_0} {}^{R_1R_2}V^{Q,n_F}_{a_1a_2,a_0} \underset{12}{\otimes} f^{n_F}_{a_0} \,, \tag{5.1} $$

where $F^{n_F+1}_{a_1a_2}$ is a DPD with $n_f + 1$ active flavours (including the heavy one) and $f^{n_F}_{a_0}$ is a PDF with n_f active flavours. We omit the superscript "spl" on F for brevity. The massive splitting kernels V^{Q,n_F} are the analogues of the massless kernels in (2.9) and depend on the additional dimensionless variable $y m_Q$. One should take $1/\alpha$ and β sufficiently large, such that

1. at scales $\mu_y < \alpha m_Q$ the heavy quark approximately decouples in the splitting process. DPDs including the flavour Q are generated by the usual matching procedure, as discussed around equation (4.7).

2. at scales $\mu_y > \beta m_Q$ the quark Q is well approximated as massless.

At the same time, $1/\alpha$ and β should not be taken too large: the massive splitting kernels contain terms with $\ln(\mu_y/m_Q)$ raised to a power that increases with the order in α_s, and when these logarithms become too large, the fixed-order truncation of the DPD splitting kernels becomes a poor approximation.

At LO the only channel in which heavy partons are produced is $g \to Q\overline{Q}$. The corresponding kernel for unpolarized quarks reads

$$ \begin{aligned} &{}^{R_1R_2}V^{Q(1)}_{Q\overline{Q},g}(z_1, z_2, y m_Q) \\ &\quad = c_{q\bar{q},g}(R_1R_2)\, \delta(1 - z_1 - z_2)\, T_F\, (y m_Q)^2 \left[(z_1^2 + z_2^2) K_1^2(y m_Q) + K_0^2(y m_Q) \right] , \end{aligned} \tag{5.2} $$

where K_1 and K_0 are modified Bessel functions of the second kind. Its polarised counterparts are given in section 3.1 of [49]. Leading-order kernels with one observed heavy parton are zero, and kernels ${}^{R_1R_2}V^{Q(1)}_{a_1a_2,a_0}$ for light partons a_1 and a_2 are equal to their massless counterpart ${}^{R_1R_2}V^{(1)}_{a_1a_2,a_0}$ in (4.2).

The NLO kernels including mass effects have not been computed so far. Their behaviour for $\mu_y \ll m_Q$ and for $\mu_y \gg m_Q$ is however known, as is their scale dependence. Using these

constraints, we can write

$$
\begin{aligned}
{}^{R_1R_2}V^{Q,n_F(2)}_{a_1a_2,a_0} &= {}^{R_1R_2}V^{n_F[2,0]}_{a_1a_2,a_0} + {}^{R_1R_2}V^{n_F[2,1]}_{a_1a_2,a_0} \ln\frac{y^2m_Q^2}{b_0^2} + {}^{R_1R_2}V^{I}_{a_1a_2,a_0}(z_1,z_2,ym_Q) \\
&\quad + g_0(ym_Q)\left({}^{R_1R_2}V^{n_F+1[2,0]}_{a_1a_2,a_0} - {}^{R_1R_2}V^{n_F[2,0]}_{a_1a_2,a_0}\right) \\
&\quad + g_1(ym_Q)\left({}^{R_1R_2}V^{n_F+1[2,1]}_{a_1a_2,a_0} - {}^{R_1R_2}V^{n_F[2,1]}_{a_1a_2,a_0}\right) \\
&\quad + \ln\frac{\mu^2}{m_Q^2}\,{}^{R_1R_2}v^{n_F,\mathrm{RGE}}_{a_1a_2,a_0}(z_1,z_2) + \frac{{}^{R_1}\gamma_J^{(0)}}{2}\left(L_yL_\zeta - \frac{L_y^2}{2} - \frac{\pi^2}{12}\right){}^{R_1R_2}V^{Q(1)}_{a_1a_2,a_0}\,,
\end{aligned}
\tag{5.3}
$$

where $V^{n_f[2,0]}(z_1,z_2)$ and $V^{n_f[2,1]}(z_1,z_2)$ are the massless splitting kernels in the decomposition (4.3), evaluated for n_f active flavours. The μ dependent part of (5.3) involves the function

$$
\begin{aligned}
{}^{R_1R_2}v^{n_F,\mathrm{RGE}}_{a_1a_2,a_0} &= \sum_{b_1,R_1'} {}^{R_1\overline{R}_1'}P^{n_F+1(0)}_{a_1b_1} \underset{1}{\otimes} {}^{R_1'R_2}V^{Q(1)}_{b_1a_2,a_0} + \sum_{b_2,R_2'} {}^{R_2\overline{R}_2'}P^{n_F+1(0)}_{a_2b_2} \underset{2}{\otimes} {}^{R_1R_2'}V^{Q(1)}_{a_1b_2,a_0} \\
&\quad - \sum_{b_0} {}^{R_1R_2}V^{Q(1)}_{a_1a_2,b_0} \underset{12}{\otimes} {}^{11}P^{n_F(0)}_{b_0a_0} + \frac{\beta_0^{n_F+1}}{2}\,{}^{R_1R_2}V^{Q(1)}_{a_1a_2,a_0}
\end{aligned}
\tag{5.4}
$$

with $\beta_0^{n_F} = 11C_A/3 - 2n_F/3$. In addition to the convolution $\otimes_{12}$ introduced in (2.9), the expression (5.4) involves convolution products

$$
P \underset{1}{\otimes} V = \int_{x_1}^1 \frac{dz}{z}\,P(z)\,V\left(\frac{x_1}{z},x_2,\dots\right)
\tag{5.5}
$$

for the first parton and its analogue $\otimes_2$ for the second parton. In the second and third lines of (5.3) we have the functions

$$
g_0(w) = w^2K_1(w)^2\,, \qquad\qquad g_1(w) = -w^2K_0(w)K_2(w)\,,
\tag{5.6}
$$

which have the limiting behaviour

$$
g_i(w) \overset{w\to0}{\longrightarrow} \left(\ln\frac{w^2}{b_0^2}\right)^i, \qquad\qquad g_i(w) \overset{w\to\infty}{\longrightarrow} 0 \qquad\qquad \text{for } i=0,1\,.
\tag{5.7}
$$

Finally, the functions $V^I_{a_1a_2,a_0}$ satisfy

$$
{}^{R_1R_2}V^{I}_{a_1a_2,a_0}(z_1,z_2,ym_Q) \overset{y\to0}{\longrightarrow} 0\,, \qquad\qquad {}^{R_1R_2}V^{I}_{a_1a_2,a_0}(z_1,z_2,ym_Q) \overset{y\to\infty}{\longrightarrow} 0
\tag{5.8}
$$

and specify the transition from small to large y. For specific parton combinations we have the additional constraints ${}^{R_1R_2}V^I_{gg,g} \propto \delta(1-z_1-z_2)$ and ${}^{R_1R_2}V^I_{q\bar{q},g} = 0$.
Without a massive two loop calculation, the functions V^I remain unknown. In our numerical studies we set them to zero for all parton and colour channels. This does not yield a correct NLO prediction in all regions of y, but it provides a smooth interpolation between the correct limiting behaviour for $y \ll 1/m_Q$ and $y \gg 1/m_Q$, and it has the correct dependence on the scales μ and ζ. In this sense we regard it as an improvement over using the massless scheme at NLO, in which the region $y \sim 1/m_Q$ is treated in a less realistic way, with often large unphysical discontinuities of the DPDs at the transition point $\mu_y = \gamma m_Q$.

Several heavy flavours The masses of charm and bottom quarks are rather close to each other, so that a region of scales μ with $m_c \ll \mu \ll m_b$ is hard to identify. It is therefore useful to consider a scheme in which u, d, s are treated as massless and both c and b as heavy. The matching of PDFs from $n_f = 3$ to $n_f = 5$ active flavours in one step has been worked out in [68] at two loops, and three-loop calculations for the relevant parton transitions are ongoing [69].
Using the two-mass scheme for DPDs in the relevant region of y, we evaluate initial conditions with mass effects in the splitting as follows:

$$\begin{aligned}
&n_f = 4, && m_c \neq 0 && && \text{for } \mu_y < \alpha m_b, \\
&n_f = 5, && m_c \neq 0, && m_b \neq 0 && \text{for } \alpha m_b < \mu_y < \beta m_c, \\
&n_f = 5, && m_c = 0, && m_b \neq 0 && \text{for } \beta m_c < \mu_y < \beta m_b
\end{aligned} \tag{5.9}$$

and

$$\begin{aligned}
&n_f = 6, && m_c = m_b = 0, && m_t \neq 0 && \text{for } \alpha m_t < \mu_y < \beta m_t, \\
&n_f = 6, && m_c = m_b = 0, && m_t = 0 && \text{for } \beta m_t < \mu_y \,.
\end{aligned} \tag{5.10}$$

Here we have assumed that $\alpha m_b < \beta m_c$ and $\alpha m_t > \beta m_b$, which is the case for our parameter choices in the following.
The DPD splitting kernels for the three regimes in (5.9) are $V^{c,3}$, V^{cb}, and $V^{b,4}$, respectively, where V^{cb} denotes the splitting kernels for 3 light flavours plus massive charm and bottom.
The analogue of the general form (5.3) for V^{cb} at NLO is given for colour singlet DPDs in equations (6.34) to (6.38) in [49]. To generalise those expressions to all colour channels, one should add labels R_1, R_2 to V^{cb} in the same way as in (5.3) and (5.4) here, and add a double logarithmic term

$$\frac{{}^{R_1}\gamma_J^{(0)}}{2} \left(L_y L_\zeta - \frac{L_y^2}{2} - \frac{\pi^2}{12} \right) {}^{R_1 R_2} V^{cb(1)}_{a_1 a_2, a_0} \tag{5.11}$$

with the massive LO kernels $V^{cb(1)}_{a_1 a_2, a_0}$ for the colour singlet being specified in and below equation (3.11) of [49].

5.1 Parton kinematics and scales

A comprehensive numerical study of mass effects in DPD splitting at LO was carried out in [49], where several shortcomings of the LO approximation were identified. In the present work we will see how the situation changes when NLO effects are included in the approximate manner described below (5.8).
We consider the same kinematics as in [49] and evaluate DPDs and double parton luminosities for

$$n_f = 5\,, \qquad M_1 = M_2 = 25\,\text{GeV}\,, \qquad \sqrt{s} = 14\,\text{TeV}\,, \tag{5.12}$$

$$n_f = 6\,, \qquad M_1 = M_2 = 1\,\text{TeV}\,, \qquad \sqrt{s} = 100\,\text{TeV}\,. \tag{5.13}$$

As in section 4 we take $Y = Y_1 = -Y_2$ in double parton luminosities.
The scale of 25 GeV, which may for instance appear in dijet production, is a compromise between being high enough to compute parton-level cross sections with $n_f = 5$ massless quarks and low enough to be sensitive to quark mass effects in the DPD splitting process.

Parton momentum fractions for the settings (5.12) and (5.13) are

$$x_1 = \bar{x}_1 \approx \begin{cases} 1.8 \times 10^{-3} & \text{if } M_1 = 25\,\text{GeV and } \sqrt{s} = 14\,\text{TeV} \\ 1.0 \times 10^{-2} & \text{if } M_1 = 1\,\text{TeV and } \sqrt{s} = 100\,\text{TeV} \end{cases} \tag{5.14}$$

for $Y_1 = 0$ and

$$x_1 \approx \begin{cases} 9.7 \times 10^{-2} \\ 5.5 \times 10^{-1} \end{cases} \qquad \bar{x}_1 \approx \begin{cases} 3.3 \times 10^{-5} & \text{if } M_1 = 25\,\text{GeV and } \sqrt{s} = 14\,\text{TeV} \\ 1.8 \times 10^{-4} & \text{if } M_1 = 1\,\text{TeV and } \sqrt{s} = 100\,\text{TeV} \end{cases} \tag{5.15}$$

for $Y_1 = 4$. The corresponding fractions x_2 and $\bar{x}_2$ are obtained by symmetry considerations. In the following studies, initial conditions of the DPDs are evaluated for the y dependent number of active flavours in the chosen flavour scheme, given by (4.6) or by (5.9) and (5.10). The DPDs are then evolved to the final scales $\mu_1 = \mu_2 = M_1$ and the rapidity parameter $\zeta = M_1 M_2/(x_1 x_2)$, with appropriate flavour number matching at scales $\mu = m_Q$.

5.2 Splitting DPDs with or without mass effects

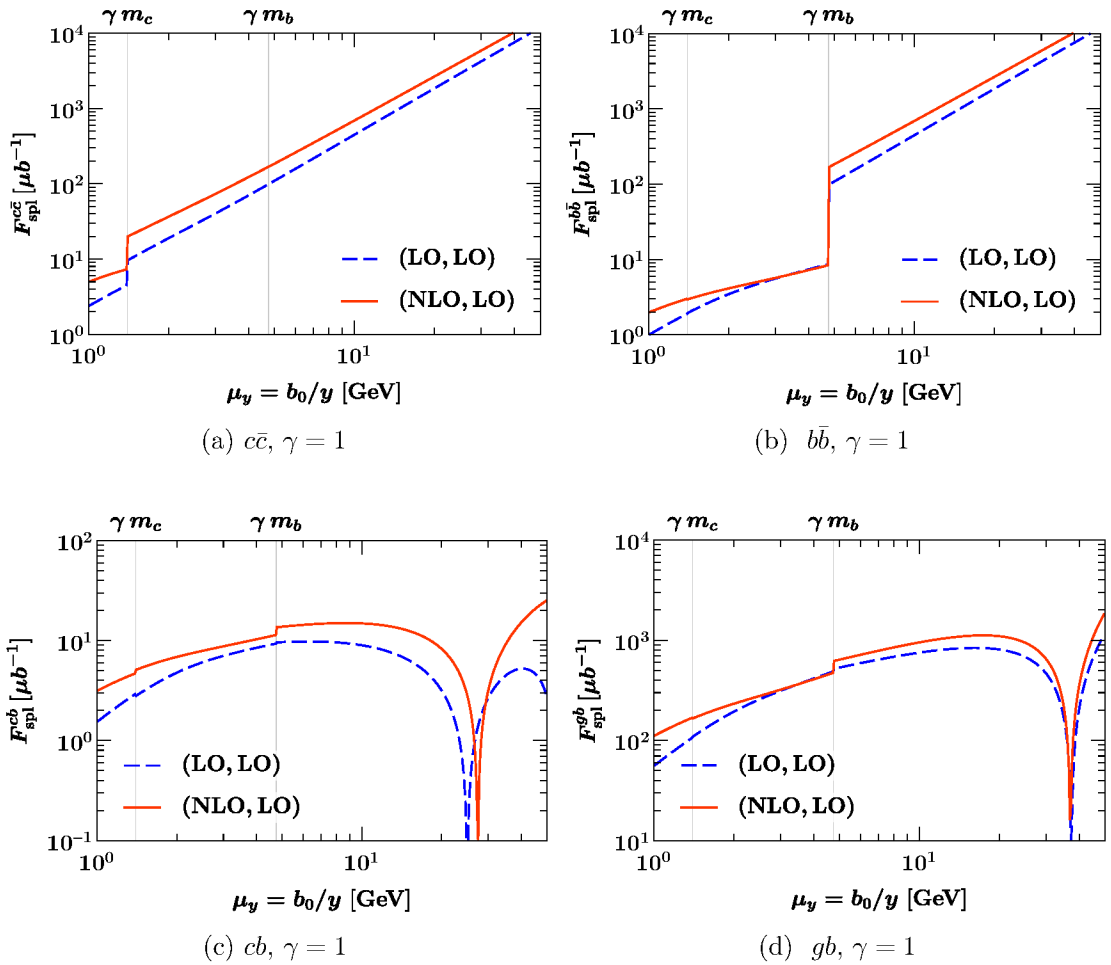
Both in the massless flavour scheme laid out in section 4 and in the massive scheme just presented, the number of active flavours in the initial conditions of DPDs is changed at specific values of y. This leads to discontinuities in the y dependence of the resulting DPDs, even after they have been evolved to final scales and matched to final n_f values. Such discontinuities can be considered as artefacts of the approximations in a given scheme, so that the size of these discontinuities can be regarded as an indicator for the quality of the scheme. Specifically, we will chose the parameters α and β in (5.9) and (5.10) so as to minimise discontinuities of evolved DPDs in y. Of course, small unphysical discontinuities are necessary but not sufficient for having realistic results. We will shortly encounter examples for this statement.

In the following we speak of the "massless scheme" in the case where splitting DPDs are initialised with massless splitting kernels for n_f given by (4.6), and of the "massive scheme" in the case where they are initialised with massless or massive kernels as specified in (5.9) and (5.10). In the following discussion and plots, we use the scale $\mu_y = b_0/y$ that determines the initial conditions of the splitting DPDs, rather than the distance y.

Not surprisingly, we find that in general the largest discontinuities occur for DPDs with one or two heavy partons. In the following plots we therefore concentrate on such cases. For DPDs with only light partons we find that discontinuities are often most pronounced in the gg channel.

In figure 20 we show colour singlet DPDs in the massless scheme. Very large discontinuities are seen for $Q\overline{Q}$ distributions at $\mu_y = m_Q$, which reflects that the direct splitting $g \to Q\overline{Q}$ does not contribute for $\mu_y < m_Q$. The same holds in the colour octet channel, as seen in figure 23a. For all channels shown here, we find that discontinuities at NLO are somewhat larger than at LO. This is an example in which smaller discontinuities do not indicate a better approximation.

The situation is very different in the massive scheme, as shown in figures 21 and 22. The discontinuities for $b\bar{b}$ and $c\bar{c}$ distributions are significantly reduced compared with the massless scheme. Moreover, discontinuities tend to be smaller at NLO than at LO. This holds in particular for the discontinuity of the gb distribution at $\mu_y = \beta m_b$, which is rather pronounced at LO. The reason for this is discussed in detail in section 4.1.2 of [49]: if one uses the massive scheme at LO, then for $\mu_y < \beta m_b$ the sequential splitting processes $g \to b\bar{b} \to gb\bar{b}$ and $g \to gg \to gb\bar{b}$ require DGLAP evolution of the DPD, which is strongly suppressed for μ_y close to the final scale 25 GeV. This leads to an underestimation of the DPD at $\mu_y < \beta m_b$. To minimise this effect we took the rather small value $\beta = 2$ in [49]. At NLO, both splitting



(a) $c\bar{c}$, $\gamma = 1$ (b) $b\bar{b}$, $\gamma = 1$

(c) cb, $\gamma = 1$ (d) gb, $\gamma = 1$

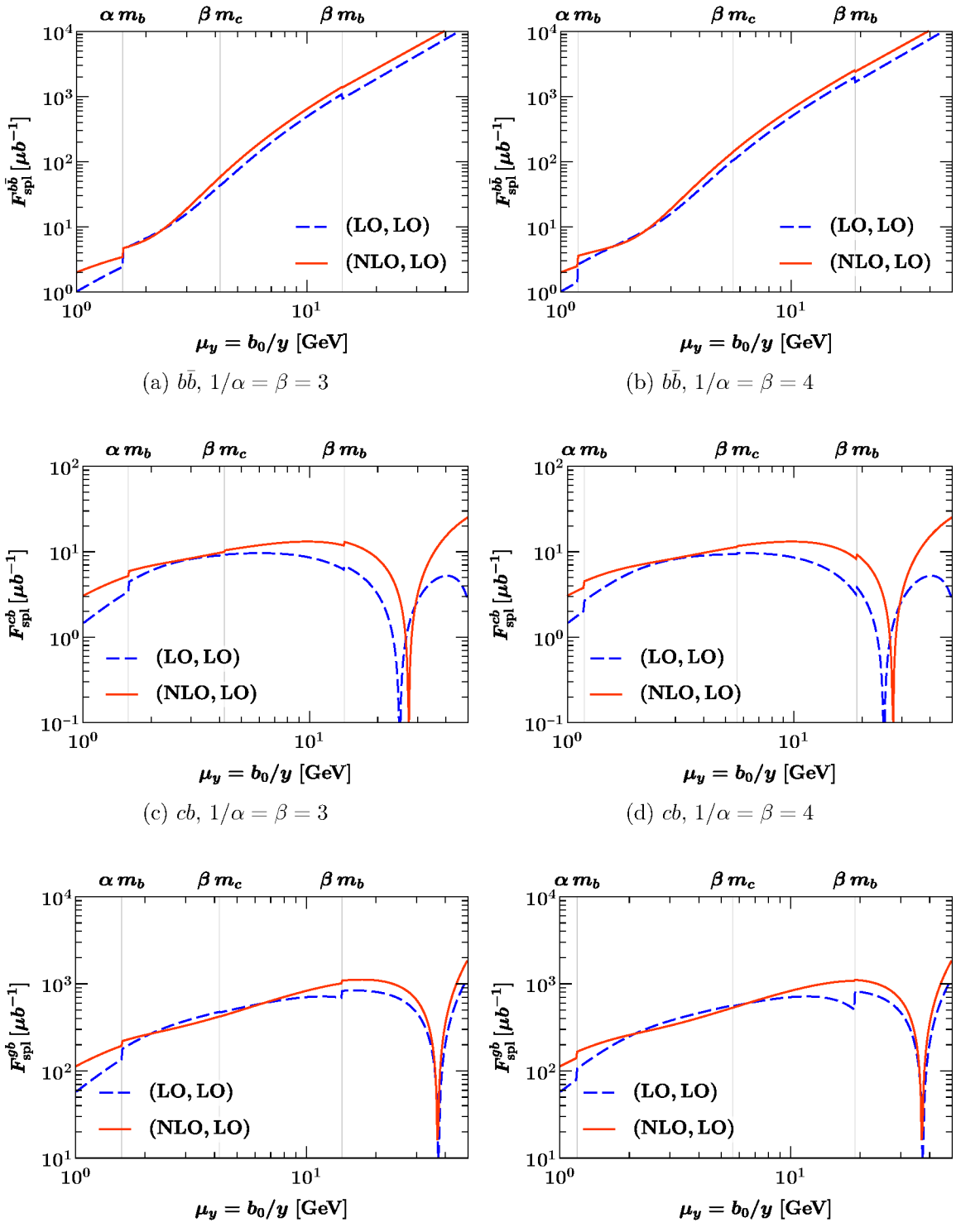
Figure 20: Splitting DPDs for $x_1 = x_2 \approx 1.8 \times 10^{-3}$. Distributions are initialised in the massless scheme (4.6) with $\gamma = 1$, followed by evolution and flavour matching to $\mu_1 = \mu_2 = 25\,\mathrm{GeV}$ and $n_f = 5$. The specification of perturbative orders is the same as in figure 3.

processes are taken into account in the DPD splitting kernels, and the problem just described is absent. Consequently, there is no problem in setting β to 3 or 4, as is seen in the bottom row of figure 21.

We remark in passing that the colour octet $b\bar{b}$ distribution has a change of sign in the massive scheme, which is absent in the massless one (see figure 23b). Since in the corresponding y region the DPD is many orders of magnitude smaller than its maximal value, we do not investigate this issue further here.

Examples for distributions involving top quarks are shown in figures 24 and 25. The differences between the massless and massive schemes, and between LO and NLO splitting are qualitatively similar to what we observed for charm and bottom quarks. We note that in the massless scheme, the discontinuity of the $t\bar{t}$ distribution at $\mu_y = \gamma m_t$ spans two orders of magnitude both at LO and at NLO.

For the channels shown here, both $1/\alpha = 3$ and $1/\alpha = \beta = 4$ yield reasonably small discontinuities in y for NLO splitting DPDs. In the following we take the smaller value 3 as our default.



(a) $b\bar{b}$, $1/\alpha = \beta = 3$

(b) $b\bar{b}$, $1/\alpha = \beta = 4$

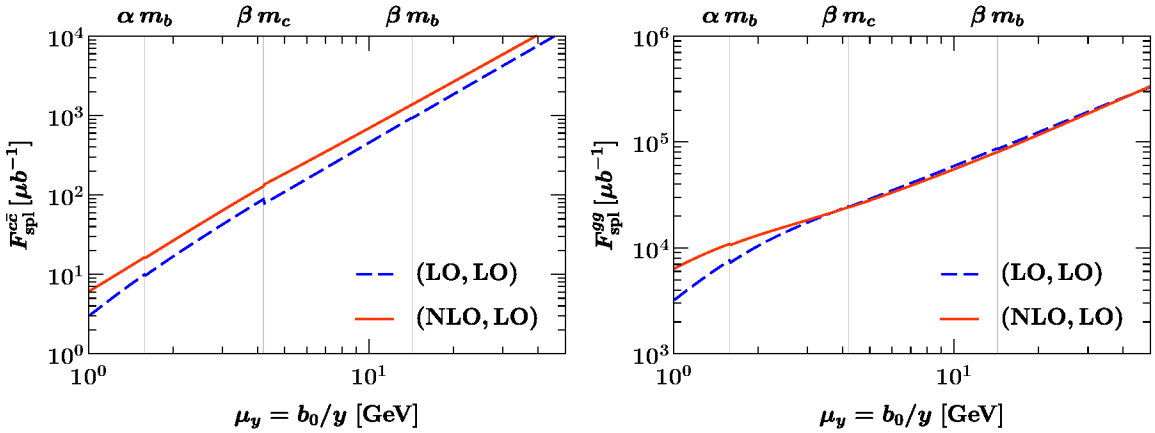
(c) cb, $1/\alpha = \beta = 3$

(d) cb, $1/\alpha = \beta = 4$

(e) gb, $1/\alpha = \beta = 3$

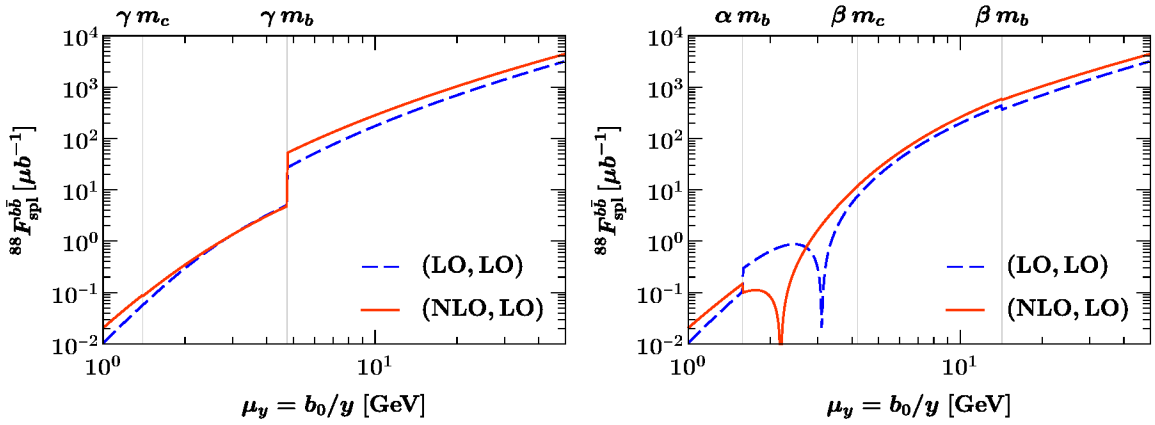
(f) gb, $1/\alpha = \beta = 4$

Figure 21: The same DPDs as in figure 20b to 20d, but initialised in the massive scheme (5.9) with two different choices of $1/\alpha = \beta$.



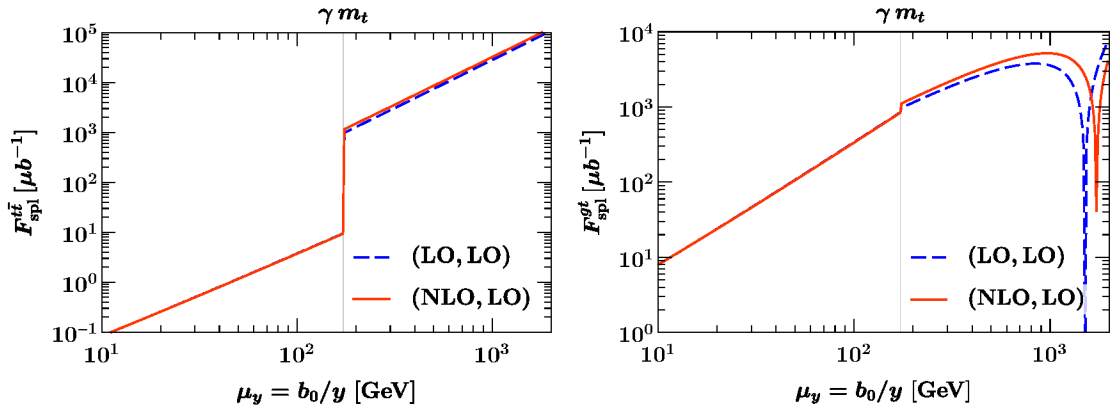
(a) $c\bar{c}$, $1/\alpha = \beta = 3$ (b) gg, $1/\alpha = \beta = 3$

Figure 22: As the left panels of figure 21, but for further parton combinations.



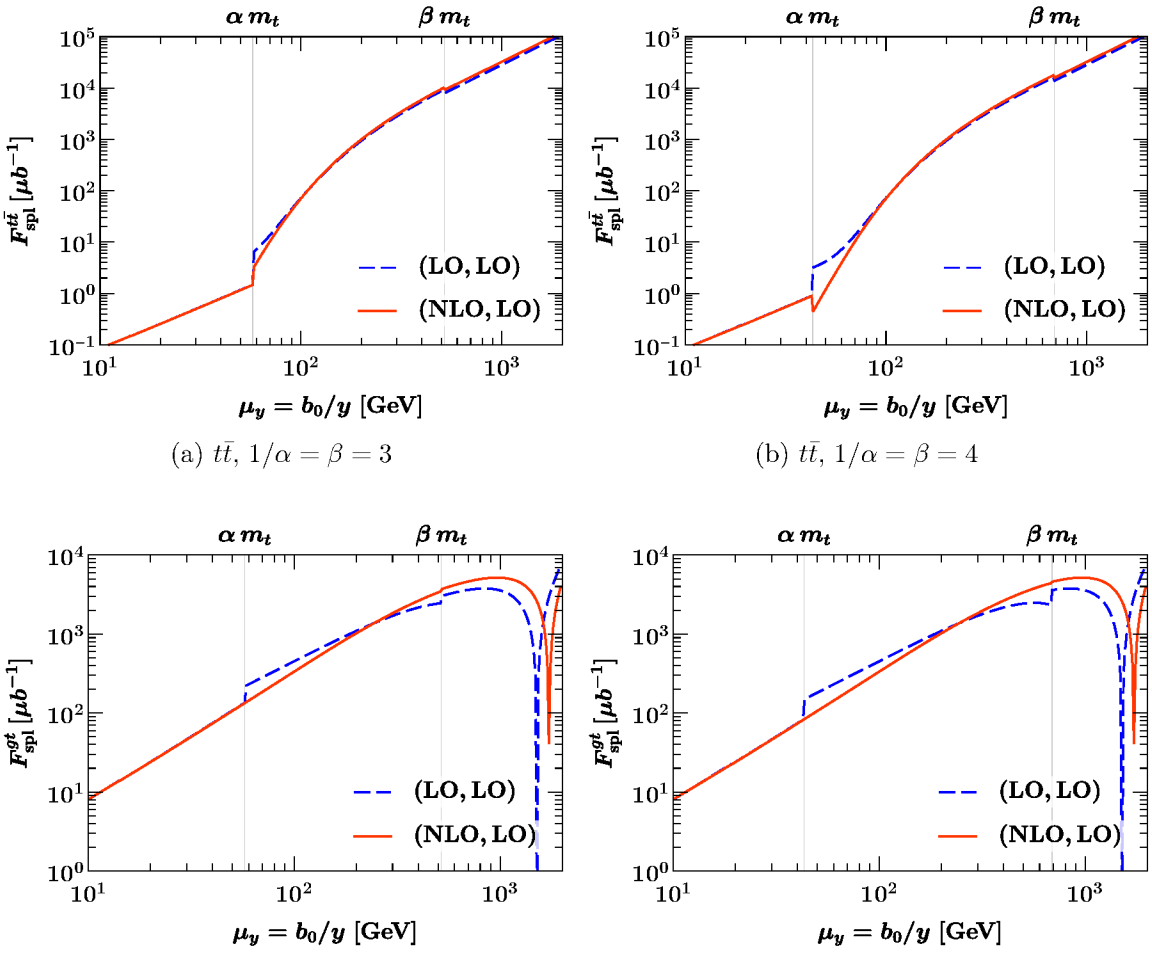
(a) $b\bar{b}$, colour octet, massless $\gamma = 1$ (b) $b\bar{b}$, colour octet, massive $1/\alpha = \beta = 3$

Figure 23: Colour octet DPDs for $b\bar{b}$ with the same settings as in figure 20 (left) or 22 (right).



(a) $t\bar{t}$, $\gamma = 1$ (b) gt, $\gamma = 1$

Figure 24: Splitting DPDs with top quarks for $x_1 = x_2 = 0.01$. Distributions are initialised in the massless scheme (4.6) with $\gamma = 1$, followed by evolution and flavour matching to $\mu_1 = \mu_2 = 1\,\mathrm{TeV}$ and $n_f = 6$.



(a) $t\bar{t}$, $1/\alpha = \beta = 3$ (b) $t\bar{t}$, $1/\alpha = \beta = 4$

(c) gt, $1/\alpha = \beta = 3$ (d) gt, $1/\alpha = \beta = 4$

Figure 25: The same DPDs as in figure 24, but initialised in the massive scheme (5.9), (5.10) with two different choices of $1/\alpha = \beta$.

5.3 Double parton luminosities: scheme parameter dependence

We find that in the massive scheme the inclusion of NLO effects has qualitatively similar effects on double parton luminosities as in the massless case, both for the central values and for the size of scale variations. We will therefore not show corresponding plots here.
Instead we now revisit the dependence of double parton luminosities on the parameters α, β or γ in the massive or the massless scheme. We take

$$\begin{aligned} &\alpha_0 = 1/3\,, && \beta_0 = 3 && \text{at NLO,} \\ &\alpha_0 = 1/4\,, && \beta_0 = 2 && \text{at LO} \end{aligned} \tag{5.16}$$

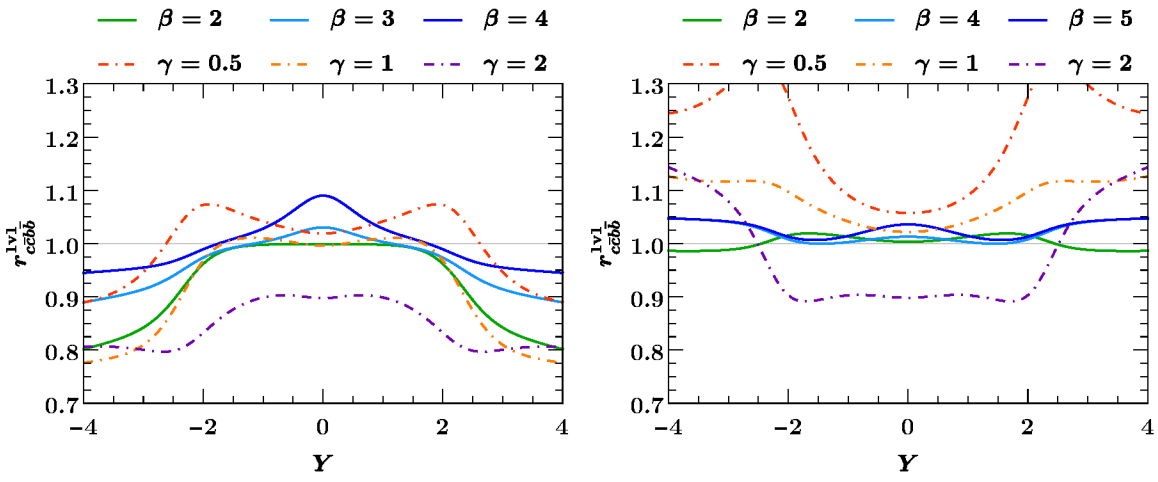
as baseline parameters in the massive scheme, where the choice for LO corresponds to our analysis in [49]. In figures 26 to 28 we show the ratios

$$r(\beta) = \frac{\mathcal{L}_{\text{massive}}(1/\alpha = \beta)}{\mathcal{L}_{\text{massive}}(\alpha = \alpha_0, \beta = \beta_0)}\,, \qquad r(\gamma) = \frac{\mathcal{L}_{\text{massless}}(\gamma)}{\mathcal{L}_{\text{massive}}(\alpha = \alpha_0, \beta = \beta_0)} \tag{5.17}$$

for selected parton combinations, either for the pure splitting contributions (1v1) or for the mixed contributions (1v2 + 2v1). The intrinsic part of the DPDs is modelled as described in section 4.3.
The variation of the ratio $r(\beta)$, shown as solid lines in the figures, quantifies how luminosities change within the massive scheme when the parameters α and β are varied. We find that in general this variation is smaller at NLO than at LO (except for the combination cb, gg in the colour octet channel). With the inclusion of NLO effects, the influence of the scheme parameters on the luminosities is at the level of 5 to 15% in the colour singlet and slightly larger in the octet case.
The ratio $r(\gamma)$, shown as dashed lines in the figures, quantifies to which extent the massless scheme is able to reproduce the (more realistic) massive scheme at the level of double parton luminosities. For c and b quarks (figures 26 and 27) an agreement at the level of 15% can be achieved with $\gamma = 1/2$ at LO. At NLO the agreement is better with $\gamma = 1$ than with $\gamma = 1/2$ and again reaches about 15% in the cases considered here.
For t quarks (figure 28) the choice $\gamma = 1$ is preferred at both LO and NLO, with deviations from the massive results reaching 20% at LO and 40% at NLO. In this case, the possibility to approximately reproduce the massive scheme with the massless one degrades rather than improves when going from LO to NLO. This does, of course, not imply that results are more accurate at LO than at NLO.

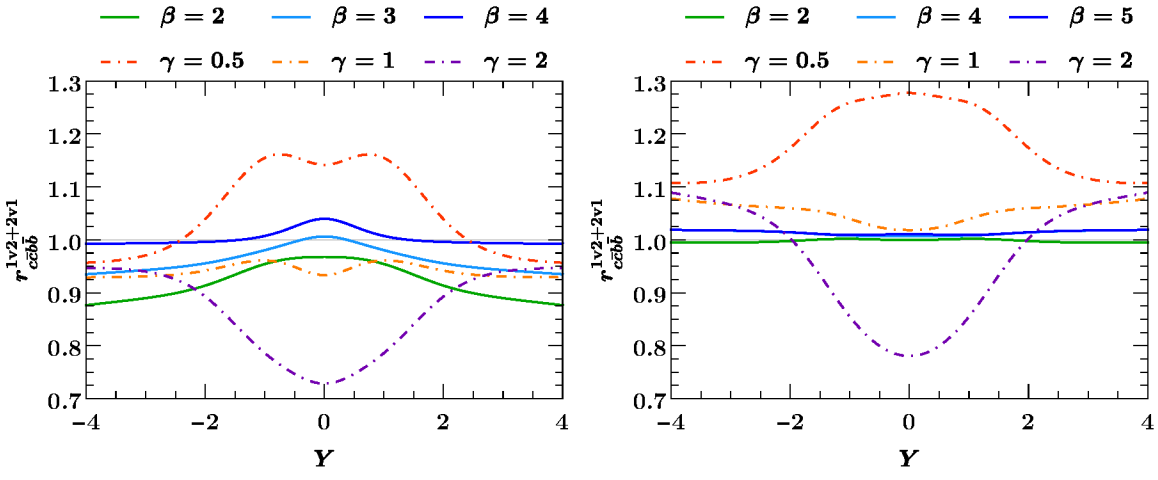
6 Summary

We have performed a detailed quantitative study of parton splitting in DPDs, i.e. of the mechanism in which the two observed partons originate from the splitting of a single one. Our study proceeds at the level of DPDs and of double parton luminosities, which are products of DPDs integrated over the transverse distance y between the partons. We considered several combinations of momentum fractions and scales that are relevant for double parton scattering at the LHC, in particular for the production of combinations of W bosons, J/Ψ, and dijets. In addition we considered kinematics in which quark mass effects in the splitting process are important for DPDs with bottom or top quarks.
The "1v1" contributions to the DPS cross section, which involve parton splitting in both DPDs, have an overlap with higher-loop corrections to single parton scattering. In the formalism of [17] this is addressed by a double counting subtraction term in the overall cross



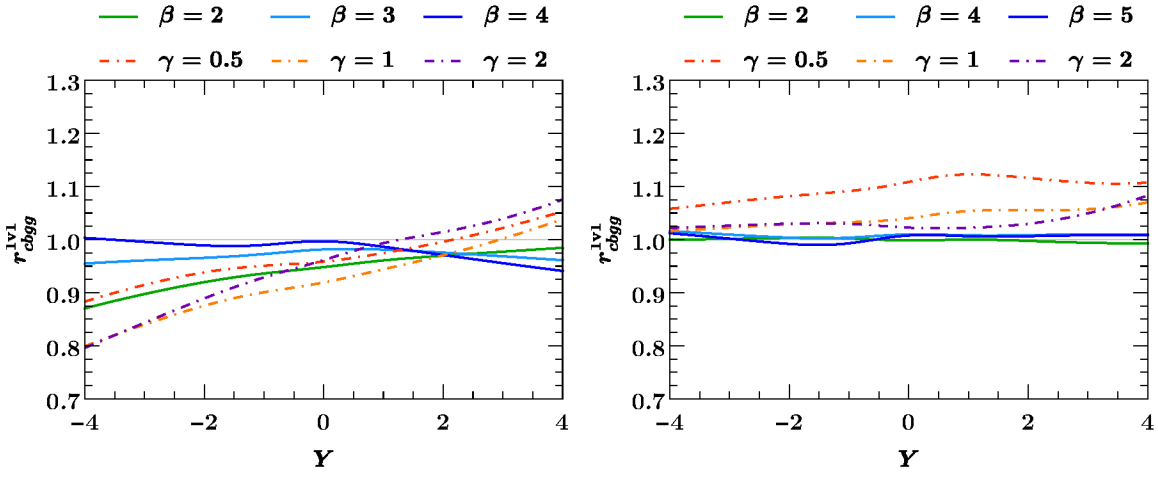
(a) $c\bar{c}, \bar{b}b$, 1v1, LO

(b) $c\bar{c}, \bar{b}b$, 1v1, NLO



(c) $c\bar{c}, \bar{b}b$, 1v2+2v1, LO

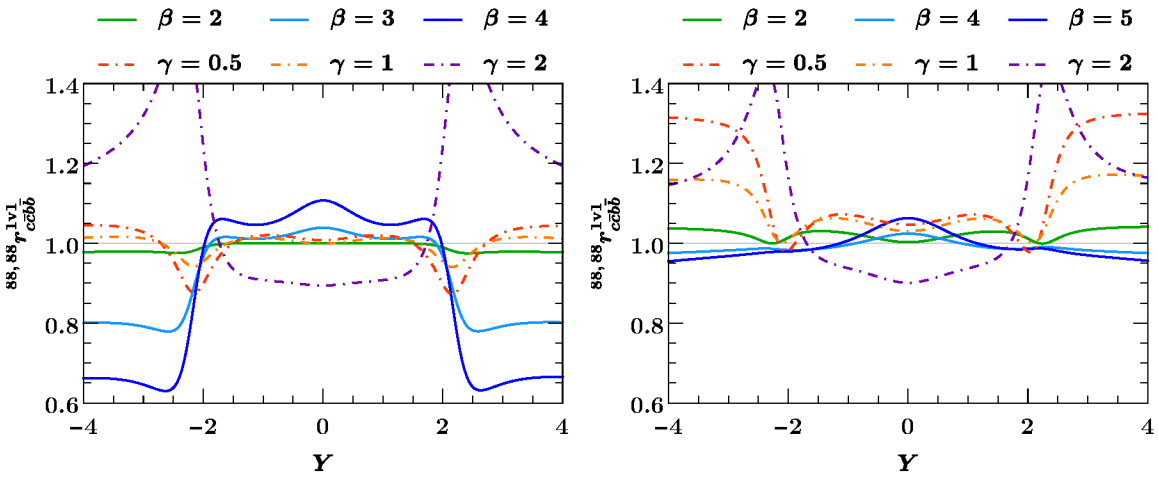
(d) $c\bar{c}, \bar{b}b$, 1v2+2v1, NLO



(e) cb, gg, 1v1, LO

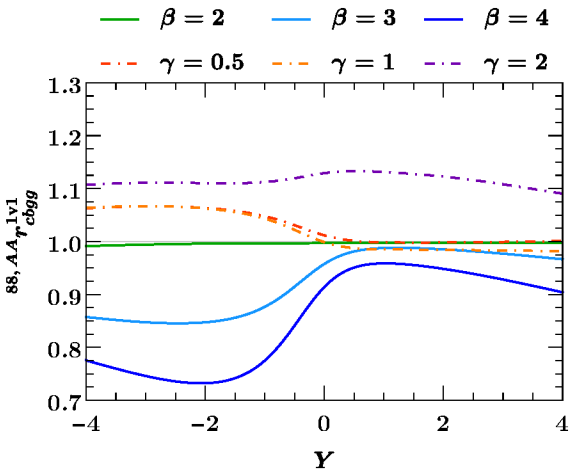
(f) cb, gg, 1v1, NLO

Figure 26: Double parton luminosity ratios (5.17) in the kinematic setting (5.12). All plots refer to the colour singlet channel. Note that at NLO the ratio $r(\beta)$ with $\beta = 3$ is equal to 1 per definition and therefore not shown. We include $\beta = 5$ in the comparison instead.

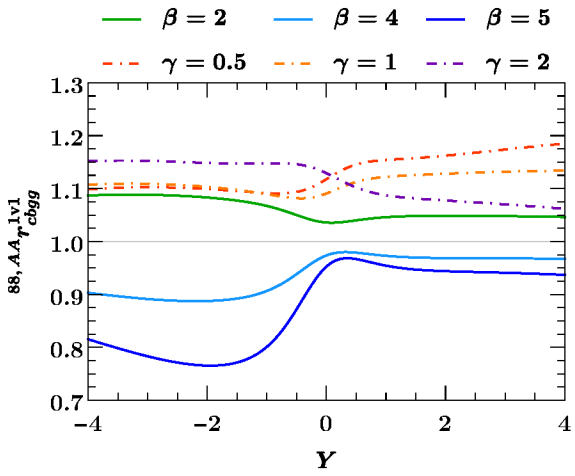


(a) $c\bar{c}, \bar{b}b$, colour octet, 1v1, LO

(b) $c\bar{c}, \bar{b}b$, colour octet, 1v1, NLO

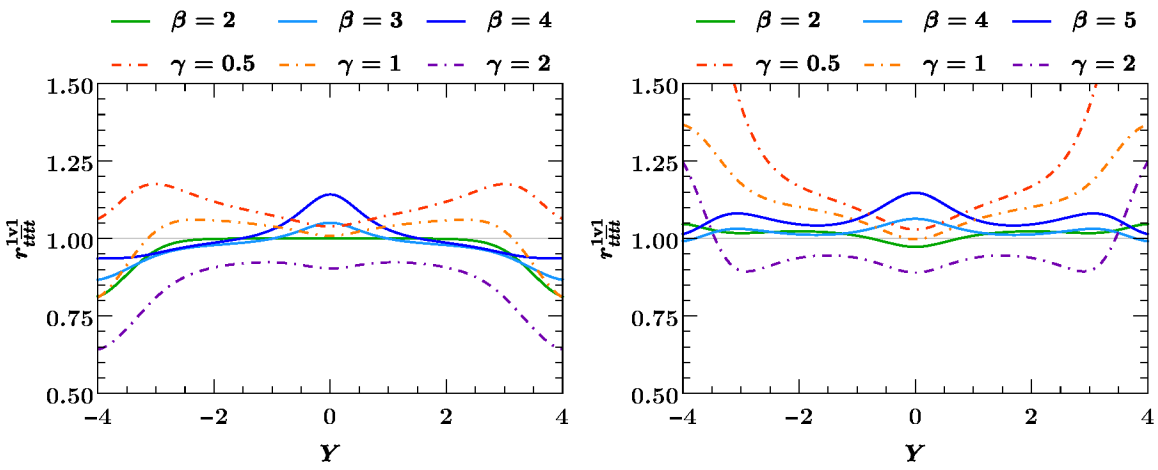


(c) cb, gg, 1v1, colour octet, LO



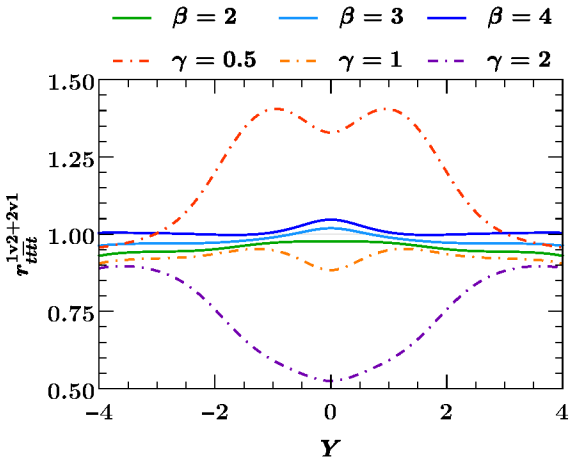
(d) cb, gg, 1v1, colour octet, NLO

Figure 27: As figure 26, but for the colour octet channel.

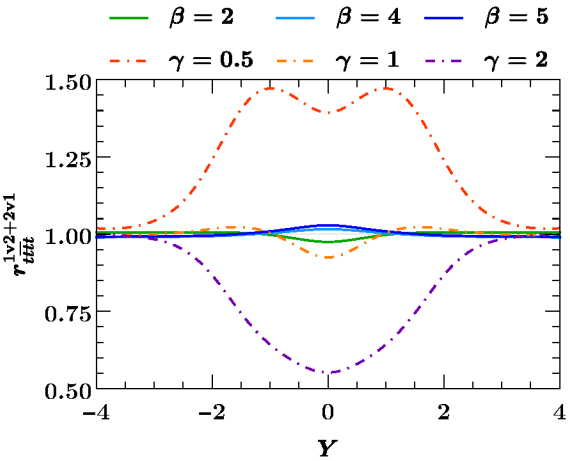


(a) $t\bar{t}, \bar{t}t$, 1v1, LO

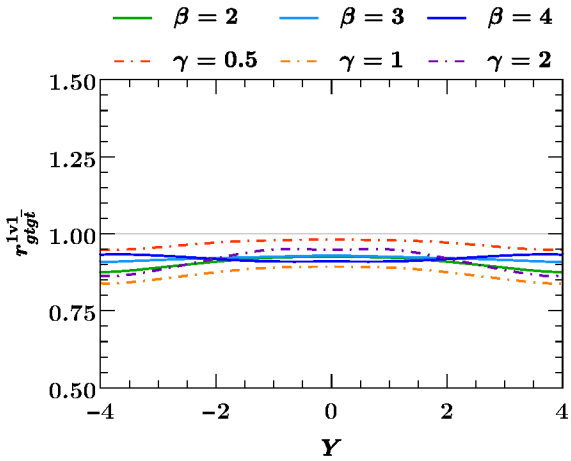
(b) $t\bar{t}, \bar{t}t$, 1v1, NLO



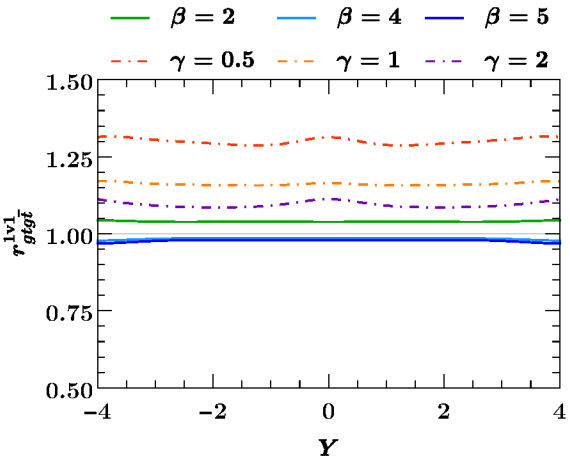
(c) $t\bar{t}, \bar{t}t$, 1v2+2v1, LO



(d) $t\bar{t}, \bar{t}t$, 1v2+2v1, NLO



(e) $gt, \bar{t}g$, 1v1, LO



(f) $gt, \bar{t}g$, 1v1, NLO

Figure 28: As figure 26, but for top quarks in the kinematic setting (5.13).

section, along with a lower cutoff y_{cut} on the integral over y. We have presented a new construction of this subtraction term, which is both flexible and simple to implement in practical calculations.

Calculations within our scheme depend on the initial scale μ_{init} of DPD evolution, at which the DPD splitting process is evaluated. We have shown that the variation of the overall cross section with μ_{init} and y_{cut} within appropriate limits can be appreciable, although it is formally beyond the accuracy of the computation. This can for instance happen when contributions with an additional power of α_s are initiated by partons with a higher density than the partons initiating the lower-order contribution (e.g. by gluons instead of quarks at small x, or by quarks instead of gluons at large x).

These are the main findings of our study for double parton luminosities that include the double counting subtraction term:

1. In many cases, the scale dependence is huge when the splitting in DPDs is evaluated at LO, whereas at NLO it is considerably reduced. The dependence on y_{cut} is typically smaller than the one on μ_{init}, and it often decreases further when going from LO to NLO.

2. The preceding statement holds in particular for parton combinations that contribute to W^+W^+ or W^-W^- pair production. We find comparable contributions from parton splitting and from "intrinsic" parton pairs in the proton (which must be modelled). This is remarkable because the parton splitting part only starts at NLO for this process.

3. A notable exception from statement 1 are parton combinations with a $q\bar{q}$ pair in each proton, which contribute in particular to W^+W^- and ZZ production. At central rapidities of the produced bosons, the μ_{init} variation is moderate and the y_{cut} variation is large, both at LO and at NLO. The fact that the double parton luminosity is strongly dominated by y close to y_{cut} indicates that SPS contributions having an overlap with DPS should be much more important than the subtracted DPS cross section $\sigma_{\text{DPS}} - \sigma_{\text{sub}}$ (see the discussion in [17]).

   We therefore expect that, although the computation of $\sigma_{\text{DPS}} - \sigma_{\text{sub}}$ is affected by large uncertainties at the presently available perturbative order, the same is not true for the overall cross section. We note that SPS contributions overlapping with DPS start at order α_s^2 (NNLO) for W^+W^- and ZZ production and are known, as well as their corrections with an additional power of α_s [70–77].

4. Colour non-singlet luminosities are most often smaller than their colour-singlet counterparts, with the exception of the four-gluon channel. They are dominated by values of y in the perturbative region, because larger values of y are strongly suppressed by rapidity evolution.

The mass m_Q of a heavy quark should be taken into account when computing splitting DPDs at $y \sim 1/m_Q$. The corresponding kernels are known at LO, whilst at NLO we currently have only "approximate" kernels that have (i) the correct scale behaviour required by the evolution equations and (ii) the correct limiting behaviour for $y \ll 1/m_Q$ and $y \gg 1/m_Q$, while interpolating smoothly between these two regimes in a model dependent way. Our main findings in this "massive scheme" are the following:

1. Unphysical discontinuities occur at y values where one switches between treating a quark flavour as massless, massive, or absent in the splitting kernels. These are greatly decreased when going from LO to approximate NLO.

2. In line with this, the dependence of double parton luminosities on the y values where the above transitions occur is reduced when going from LO to approximate NLO.

3. Double parton luminosities in the massive scheme differ from those in the scheme where only massless DPD splitting kernels are used. By adjusting the y values at which a new flavour is included in the massless splitting DPDs, one can obtain an *approximate* global agreement between the two schemes, at the level of several 10% in the cases we considered. For top quarks, this agreement happens to be worse at approximate NLO than at LO.

Overall, we find that the inclusion of NLO effects in splitting DPDs significantly improves the stability of predictions for double parton scattering. It would be good to have NLO splitting kernels also for polarised partons, and to include heavy quark masses in a more accurate way than the one described here. We leave this to future work.

Acknowledgements

It is a pleasure to thank Frank Tackmann for useful discussions. The numerical results in this work have been obtained with the ChiliPDF library [78, 79], which is under development. We gratefully acknowledge the contributions of our collaborators Florian Fabry, Oskar Grocholski, Riccardo Nagar, and Frank Tackmann to that project. The Feynman graphs in this manuscript were produced with JaxoDraw [80, 81].
This work has been supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – grant number 409651613 (Research Unit FOR 2926). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 101002090 COLORFREE).

References

[1] A. Kulesza and W. Stirling, *Like sign W boson production at the LHC as a probe of double parton scattering*, *Phys. Lett. B* **475** (2000) 168 [hep-ph/9912232].

[2] J. R. Gaunt, C.-H. Kom, A. Kulesza and W. Stirling, *Same-sign W pair production as a probe of double parton scattering at the LHC*, *Eur. Phys. J. C* **69** (2010) 53 [1003.3953].

[3] CMS collaboration, A. Tumasyan et al., *Observation of same-sign WW production from double parton scattering in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, *Phys. Rev. Lett.* **131** (2023) 091803 [2206.02681].

[4] ATLAS collaboration, G. Aad et al., *Observation of double parton scattering in same-sign W boson pair production in pp collisions at $\sqrt{s}$ =13 TeV with the ATLAS detector*, *Phys. Lett. B* **870** (2025) 139892 [2505.08313].

[5] N. Paver and D. Treleani, *Multi-Quark Scattering and Large p_T Jet Production in Hadronic Collisions*, *Nuovo Cim. A* **70** (1982) 215.

[6] M. Mekhfi, *Multiparton Processes: an Application to Double Drell-Yan*, *Phys. Rev. D* **32** (1985) 2371.

[7] J. R. Gaunt and W. J. Stirling, *Double Parton Distributions Incorporating Perturbative QCD Evolution and Momentum and Quark Number Sum Rules*, *JHEP* **03** (2010) 005 [0910.4347].

[8] M. Diehl, D. Ostermeier and A. Schäfer, *Elements of a theory for multiparton interactions in QCD*, *JHEP* **03** (2012) 089 [1111.0910].

[9] J. R. Gaunt and W. Stirling, *Double Parton Scattering Singularity in One-Loop Integrals*, *JHEP* **06** (2011) 048 [1103.1888].

[10] J. R. Gaunt, *Single Perturbative Splitting Diagrams in Double Parton Scattering*, *JHEP* **01** (2013) 042 [1207.0480].

[11] B. Blok, Y. Dokshitser, L. Frankfurt and M. Strikman, *pQCD physics of multiparton interactions*, *Eur. Phys. J.* **C72** (2012) 1963 [1106.5533].

[12] B. Blok, Y. Dokshitzer, L. Frankfurt and M. Strikman, *Perturbative QCD correlations in multi-parton collisions*, *Eur. Phys. J. C* **74** (2014) 2926 [1306.3763].

[13] M. G. Ryskin and A. M. Snigirev, *A Fresh look at double parton scattering*, *Phys. Rev.* **D83** (2011) 114047 [1103.3495].

[14] M. G. Ryskin and A. M. Snigirev, *Double parton scattering in double logarithm approximation of perturbative QCD*, *Phys. Rev.* **D86** (2012) 014018 [1203.2330].

[15] A. V. Manohar and W. J. Waalewijn, *A QCD Analysis of Double Parton Scattering: Color Correlations, Interference Effects and Evolution*, *Phys. Rev. D* **85** (2012) 114009 [1202.3794].

[16] A. V. Manohar and W. J. Waalewijn, *What is Double Parton Scattering?*, *Phys. Lett. B* **713** (2012) 196 [1202.5034].

[17] M. Diehl, J. R. Gaunt and K. Schönwald, *Double hard scattering without double counting*, *JHEP* **06** (2017) 083 [1702.06486].

[18] M. G. A. Buffing, M. Diehl and T. Kasemets, *Transverse momentum in double parton scattering: factorisation, evolution and matching*, *JHEP* **01** (2018) 044 [1708.03528v4].

[19] T. Sjostrand and P. Z. Skands, *Multiple interactions and the structure of beam remnants*, *JHEP* **03** (2004) 053 [hep-ph/0402078].

[20] J. Bellm, S. Gieseke and P. Kirchgaesser, *Improving the description of multiple interactions in Herwig*, *Eur. Phys. J. C* **80** (2020) 469 [1911.13149].

[21] Sherpa collaboration, E. Bothmann et al., *Event generation with Sherpa 3*, *JHEP* **12** (2024) 156 [2410.22148].

[22] B. Cabouat and J. R. Gaunt, *Combining single and double parton scatterings in a parton shower*, *JHEP* **10** (2020) 012 [2008.01442].

[23] O. Fedkevych, J. R. Gaunt and S. Smith, *Improving on the PYTHIA modelling of equal-scale multi-parton distribution functions*, 2510.04554.

[24] P. Bartalini and J. R. Gaunt, eds., *Multiple Parton Interactions at the LHC*, vol. 29. World Scientific Publishing, 2019, 10.1142/10646.

[25] M. Jaarsma, R. Rahn and W. J. Waalewijn, *Towards double parton distributions from first principles using Large Momentum Effective Theory*, *JHEP* **12** (2023) 014 [2305.09716].

[26] J. R. Andersen, P. F. Monni, L. Rottoli, G. P. Salam and A. Soto-Ontoso, *Exploring High-Purity Multiparton Scattering at Hadron Colliders*, *Phys. Rev. Lett.* **132** (2024) 041901 [2307.05693].

[27] BLFQ collaboration, T.-C. Peng, Z. Hu, S. Nair, S. Xu, X. Liu, C. Mondal et al., *Double parton distributions of the proton from basis light-front quantization*, *Phys. Rev. D* **112** (2025) 014018 [2410.11574].

[28] A. Dumitru and E. Kolbusz, *Quantum entanglement correlations in double quark PDFs*, *Phys. Rev. D* **111** (2025) 114033 [2501.12312].

[29] B. Blok and R. Segev, *Double parton interactions initiated by direct photons in γp and γA collisions revisited*, *Phys. Rev. D* **113** (2026) 014009 [2503.14152].

[30] J. V. C. Lovato, E. Huayra and E. G. de Oliveira, *Momentum fraction and hard scale dependence of double parton scattering*, *JHEP* **10** (2025) 163 [2506.05337].

[31] F. A. Ceccopieri, F. Fornetti, E. Pace, M. Rinaldi, G. Salmè and N. Iles, *Theoretical insights on nuclear double parton distributions*, *Eur. Phys. J. C* **85** (2025) 1265 [2507.02495].

[32] CDF collaboration, F. Abe et al., *Double parton scattering in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV*, *Phys. Rev. D* **56** (1997) 3811.

[33] 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** (2016) 052008 [1512.05291].

[34] 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** (2017) 047 [1612.07451].

[35] ALICE collaboration, S. Acharya et al., *Measurement of inclusive J/ψ pair production cross section in pp collisions at $\sqrt{s}$ =13 TeV*, *Phys. Rev. C* **108** (2023) 045203 [2303.13431].

[36] D0 collaboration, V. M. Abazov et al., *Observation and Studies of Double J/ψ Production at the Tevatron*, *Phys. Rev. D* **90** (2014) 111101 [1406.2380].

[37] ATLAS collaboration, M. Aaboud et al., *Measurement of the prompt J/ψ pair production cross-section in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **77** (2017) 76 [1612.02950].

[38] ATLAS collaboration, M. Aaboud et al., *Measurement of J/ψ production in association with a $W^{\pm}$ boson with pp data at 8 TeV*, *JHEP* **01** (2020) 095 [1909.13626].

[39] ATLAS collaboration, M. Aaboud et al., *Study of hard double-parton scattering in four-jet events in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS experiment*, *JHEP* **11** (2016) 110 [1608.01857].

[40] CMS collaboration, A. Tumasyan et al., *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** (2022) 177 [2109.13822].

[41] ATLAS collaboration, G. Aad et al., *Measurement of hard double-parton interactions in $W(\to l\nu)+$ 2 jet events at $\sqrt{s}$=7 TeV with the ATLAS detector*, *New J. Phys.* **15** (2013) 033038 [1301.6872].

[42] CMS collaboration, S. Chatrchyan et al., *Study of Double Parton Scattering Using W + 2-Jet Events in Proton-Proton Collisions at $\sqrt{s}$ = 7 TeV*, *JHEP* **03** (2014) 032 [1312.5729].

[43] R. Kirschner, *Generalized Lipatov-Altarelli-Parisi Equations and Jet Calculus Rules*, *Phys. Lett. B* **84** (1979) 266.

[44] V. P. Shelest, A. M. Snigirev and G. M. Zinovev, *The Multiparton Distribution Equations in QCD*, *Phys. Lett. B* **113** (1982) 325.

[45] A. M. Snigirev, *Double parton distributions in the leading logarithm approximation of perturbative QCD*, *Phys. Rev. D* **68** (2003) 114012 [hep-ph/0304172].

[46] F. A. Ceccopieri, *An update on the evolution of double parton distributions*, *Phys. Lett. B* **697** (2011) 482 [1011.6586].

[47] M. Diehl, J. R. Gaunt, P. Plößl and A. Schäfer, *Two-loop splitting in double parton distributions*, *SciPost Phys.* **7** (2019) 017 [1902.08019].

[48] M. Diehl, J. R. Gaunt and P. Plößl, *Two-loop splitting in double parton distributions: the colour non-singlet case*, *JHEP* **08** (2021) 040 [2105.08425].

[49] M. Diehl, R. Nagar and P. Plößl, *Quark mass effects in double parton distributions*, *JHEP* **09** (2023) 100 [2212.07736].

[50] M. Cacciari, G. P. Salam and S. Sapeta, *On the characterisation of the underlying event*, *JHEP* **04** (2010) 065 [0912.4926].

[51] M. Diehl, F. Fabry and A. Vladimirov, *Two-loop evolution kernels for colour dependent double parton distributions*, *JHEP* **05** (2023) 067 [2212.11843].

[52] M. Diehl, J. R. Gaunt, P. Pichini and P. Plößl, *Double parton distributions out of bounds in colour space*, *Eur. Phys. J. C* **81** (2021) 1033 [2109.14304].

[53] M. Diehl, F. Fabry and P. Plößl, *Evolution of colour correlated double parton distributions: a quantitative study*, *JHEP* **02** (2024) 229 [2310.16432].

[54] J. Collins, *Foundations of perturbative QCD*, vol. 32. Cambridge University Press, 11, 2013, 10.1017/CBO9780511975592.

[55] A. Vladimirov, *Soft factors for double parton scattering at NNLO*, *JHEP* **12** (2016) 038 [1608.04920].

[56] Y. Li and H. X. Zhu, *Bootstrapping Rapidity Anomalous Dimensions for Transverse-Momentum Resummation*, *Phys. Rev. Lett.* **118** (2017) 022004 [1604.01404].

[57] M. G. Echevarria, I. Scimemi and A. Vladimirov, *Unpolarized Transverse Momentum Dependent Parton Distribution and Fragmentation Functions at next-to-next-to-leading order*, *JHEP* **09** (2016) 004 [1604.07869].

[58] I. Scimemi and A. Vladimirov, *Non-perturbative structure of semi-inclusive deep-inelastic and Drell-Yan scattering at small transverse momentum*, *JHEP* **06** (2020) 137 [1912.06532].

[59] A. Bacchetta, V. Bertone, C. Bissolotti, G. Bozzi, F. Delcarro, F. Piacenza et al., *Transverse-momentum-dependent parton distributions up to N^3LL from Drell-Yan data*, *JHEP* **07** (2020) 117 [1912.07550].

[60] I. W. Stewart, F. J. Tackmann and W. J. Waalewijn, *The Quark Beam Function at NNLL*, *JHEP* **09** (2010) 005 [1002.2213].

[61] F. J. Tackmann, *Beyond scale variations: perturbative theory uncertainties from nuisance parameters*, *JHEP* **08** (2025) 098 [2411.18606].

[62] M. Diehl, J. R. Gaunt, D. M. Lang, P. Plößl and A. Schäfer, *Sum rule improved double parton distributions in position space*, *Eur. Phys. J. C* **80** (2020) 468 [2001.10428].

[63] S. Bailey, T. Cridge, L. A. Harland-Lang, A. D. Martin and R. S. Thorne, *Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs*, *Eur. Phys. J. C* **81** (2021) 341 [2012.04684].

[64] 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** (2015) 132 [1412.7420].

[65] M. Diehl, T. Kasemets and S. Keane, *Correlations in double parton distributions: effects of evolution*, *JHEP* **05** (2014) 118 [1401.1233].

[66] ATLAS collaboration, L. Adam, *Measurements of multiparton interactions at ATLAS*, *PoS* **DIS2019** (2019) 258.

[67] CMS collaboration, A. Tumasyan et al., *Observation of triple J/ψ meson production in proton-proton collisions*, *Nature Phys.* **19** (2023) 338 [2111.05370].

[68] J. Blümlein, A. De Freitas, C. Schneider and K. Schönwald, *The Variable Flavor Number Scheme at Next-to-Leading Order*, *Phys. Lett. B* **782** (2018) 362 [1804.03129].

[69] J. Ablinger, J. Blümlein, A. De Freitas, A. von Manteuffel, C. Schneider and K. Schönwald, *The two-mass contributions to the three-loop massive operator matrix elements $\widetilde{A}_{Qg}^{(3)}$ and $\Delta\widetilde{A}_{Qg}^{(3)}$*, *JHEP* **01** (2026) 111 [2510.09403].

[70] E. W. N. Glover and J. J. van der Bij, *Z boson pair production via gluon fusion*, *Nucl. Phys. B* **321** (1989) 561.

[71] J. M. Campbell, R. K. Ellis, M. Czakon and S. Kirchner, *Two loop correction to interference in $gg \to ZZ$*, *JHEP* **08** (2016) 011 [1605.01380].

[72] F. Caola, K. Melnikov, R. Röntsch and L. Tancredi, *QCD corrections to ZZ production in gluon fusion at the LHC*, *Phys. Rev. D* **92** (2015) 094028 [1509.06734].

[73] F. Caola, M. Dowling, K. Melnikov, R. Röntsch and L. Tancredi, *QCD corrections to vector boson pair production in gluon fusion including interference effects with off-shell Higgs at the LHC*, *JHEP* **07** (2016) 087 [1605.04610].

[74] M. Grazzini, S. Kallweit, M. Wiesemann and J. Y. Yook, *ZZ production at the LHC: NLO QCD corrections to the loop-induced gluon fusion channel*, *JHEP* **03** (2019) 070 [1811.09593].

[75] F. Caola, K. Melnikov, R. Röntsch and L. Tancredi, *QCD corrections to W^+W^- production through gluon fusion*, *Phys. Lett. B* **754** (2016) 275 [1511.08617].

[76] M. Grazzini, S. Kallweit, M. Wiesemann and J. Y. Yook, *W^+W^- production at the LHC: NLO QCD corrections to the loop-induced gluon fusion channel*, *Phys. Lett. B* **804** (2020) 135399 [2002.01877].

[77] B. Agarwal, S. Jones, M. Kerner and A. von Manteuffel, *Complete Next-to-Leading Order QCD Corrections to ZZ Production in Gluon Fusion*, *Phys. Rev. Lett.* **134** (2025) 031901 [2404.05684].

[78] M. Diehl, R. Nagar and F. J. Tackmann, *ChiliPDF: Chebyshev interpolation for parton distributions*, *Eur. Phys. J. C* **82** (2022) 257 [2112.09703].

[79] M. Diehl, R. Nagar, P. Plößl and F. J. Tackmann, *Evolution and interpolation of double parton distributions using Chebyshev grids*, *Eur. Phys. J. C* **83** (2023) 536 [2305.04845].

[80] D. Binosi and L. Theussl, *JaxoDraw: A Graphical user interface for drawing Feynman diagrams*, *Comput. Phys. Commun.* **161** (2004) 76 [hep-ph/0309015].

[81] D. Binosi, J. Collins, C. Kaufhold and L. Theussl, *JaxoDraw: A Graphical user interface for drawing Feynman diagrams. Version 2.0 release notes*, *Comput. Phys. Commun.* **180** (2009) 1709 [0811.4113].