The Henryk Niewodniczański Institute of Nuclear Physics
Polish Academy of Sciences

Doctoral dissertation

Defended by

Marcin Piotr Bury

Phenomenology of transverse-momentum dependent factorizations in hadronic collisions

Supervisor:

dr hab. Krzysztof Kutak

Auxiliary Supervisor:

dr hab. Sebastian Sapeta

Kraków, January 2020

Abstract

This thesis discusses QCD factorization frameworks for hadronic collisions that directly include transverse momenta of interacting partons. In the first part we show how the Transverse Momentum Dependent (TMD) distributions become process dependent and calculate TMD gluon distributions for processes involving 5 and 6 colored partons. A general result for multi-gluon process at large N_c is presented. The second part concerns the phenomenology of hadronic collisions. We investigate the importance of several physical effects that can improve theoretical description of processes with jets in the final state. We study the double parton scattering contributions to dijet cross section, constructed from the single jet production in high energy factorization (HEF). We analyze the effect of final state radiation and introduce initial state parton shower based on the unintegrated parton distributions. Finally, the results of TMD gluon distributions, obtained in the first part, were used to compare the predictions for three jet production in HEF and Improved TMD factorization.

Streszczenie

W niniejszej rozprawie przedstawione są podejścia faktoryzacyjne chromodynamiki kwantowej uwzględniające pędy poprzeczne oddziałujących partonów. W pierwszej części pokazane zostało jak rozkłady partonowe TMD uzyskują zależność od procesu oraz otrzymano rozkłady gluonowe TMD dla procesów zawierających 5 i 6 cząstek posiadających ładunek kolorowy. Przedstawiony został także ogólny wynik dla wielogluonowych procesów w granicy dużej liczby kolorów. Druga część pracy dotyczy fenomenologii zderzeń hadronów. Analizowany jest wpływ szeregu efektów fizycznych pozwalających na lepszy opis teoretyczny procesów produkcji dżetów. W szczególności, badany jest wkład podwójnego rozpraszania partonów do procesu produkcji dwóch dżetów, skonstruowanego na podstawie produkcji pojedynczego dżetu w faktoryzacji wysokoenergetycznej, oraz rozpatrywane są efekty kaskad partonowych ze stanów początkowych i końcowych. Na koniec, wyniki rozkładów gluonowych TMD, otrzymane w pierwszej części, zostały zastosowane do porównania przewidywań dla procesu produkcji trzech dżetów w formalizmie ulepszonej faktoryzacji TMD oraz faktoryzacji wysokoenergetycznej.

Acknowledgments

First and foremost I wish to express my sincere gratitude to my supervisors, Krzysztof Kutak and Sebastian Sapeta, for their guidance, patience and dedication during the last few years. I have been very fortunate to work under their supervision, and I genuinely thank them for their invaluable help and advice.

Furthermore I would like to thank Andreas van Hameren and Piotr Kotko for their support and many hours of useful discussions and explanations.

I am also indebted to Michal Deak, Hannes Jung, Hans Van Haevermaet, Pierre Van Mechelen and Mirko Serino for allowing me the privilege of participating in very fruitful collaborations during the past years.

I wish to thank my friends and colleagues from the institute for their presence and support, in particular Cosimo, Etienne, Kuba, Martin, Mirko, Pieter, Sasza and Sergii.

I wish also to thank the members of the Division of Theoretical Physics at the Institute of Nuclear Physics in Cracow for creating a friendly atmosphere for work and study.

Special thanks to my family and my friends who have been there when I have needed them the most.

Contents

Preface

The Standard Model (SM) has been the leading theory of particle physics with exceptional success in describing the strong and electroweak interactions of fundamental fermions and vector bosons. Nevertheless, despite the great success, the Standard Model cannot be the final theory - even after the discovery of the Higgs boson - as there are phenomena, like baryon asymmetry or neutrino masses, which cannot be explained within the model. For this reason a considerable effort has been dedicated to the search for new physics beyond the SM.

An unprecedented opportunity to explore new physics has emerged with the start-up of the Large Hadron Collider (LHC) at CERN. The strong interactions are the dominant mechanism of particle production at the LHC and Quantum Chromodynamics (QCD) is the fundamental theory underlying calculations. Extraction of new physics signal from huge Standard Model background requires understanding of QCD at a remarkably accurate level.

The fundamental role in application of QCD play so-called factorization theorems, which allow to separate the short-distance perturbative physics from the long-distance nonperturbative effects. While the standard collinear factorization is a sufficiently good approximation for most inclusive processes in hadronic collisions, it leaves many problems unaddressed. A wide class of processes that require extension are those involving measurable internal transverse momenta of partons. A consistent theoretical treatment of such processes, depending on kinematic regime, is based on the Transverse Momentum Dependent (TMD) or High Energy factorization (HEF). In high-energy factorization, applicable when both momentum scale and energy scale involved in the scattering process are high, long-distance part is described by unintegrated gluon densities, which in addition to the longitudinal hadron momentum fraction, depend as well on the transversal momentum. As a consequence parton densities in high-energy factorization have to be convoluted with the hard process calculated with the initiating gluons being off-shell. In TMD factorization, the on-shell hard process is accompanied by several process-dependent TMD distributions, which leads to violation of factorization in hadron-hadron collisions.

In the present thesis, we study the phenomenology of TMD and HEF factorizations. The thesis is organized as follows. Section 1 is introductory and contains a number of preliminary topics. In particular, we describe Lagrangian of QCD and review evolution equations. We also present a short overview of collider phenomenology. In Section 2 we discuss factorization theorems. Section 3 describes the k_T-dependent distributions. In Section 4 we calculate TMD distributions for multiparton processes. Section 5 contains phenomenology of multijet processes. A final discussion is presented in Section 6, which is followed by six appendices.

The original work, presented in this Thesis, is based on the following publications

- **Single and double inclusive forward jet production at the LHC at $\sqrt{s} = 7$ and 13 TeV**
  M. Bury, M. Deak, K. Kutak and S. Sapeta, Phys. Lett. B 760 (2016)
- **Calculations with off-shell matrix elements, TMD parton densities and TMD parton showers**
  M. Bury, A. van Hameren, H. Jung, K. Kutak, S. Sapeta and M. Serino, Eur. Phys. J. C 78 (2018) no. 2
- **Single inclusive jet production and the nuclear modification ratio at very forward rapidity in proton-lead collisions with $\sqrt{s_{NN}} = 5.02$ TeV**
  M. Bury, H. van Haevermaet, A. van Hameren, P. van Mechelen, K. Kutak and M. Serino, Phys. Lett. B 780 (2018)
- **TMD gluon distributions for multiparton processes**
  M.Bury, P. Kotko and K.Kutak, Eur. Phys. J. C 79 (2019) no. 2

1 Introduction

1.1 Basics of QCD

Quantum Chromodynamics, formulated in 1973 [1–3], is the part of the Standard Model that describes strong interactions of quark and gluons. Formally, QCD is a non-abelian gauge theory based on the symmetry group SU(N_c), also called a Yang-Mills theory, with $N_c = 3$, and the degree of freedom associated with this gauge group referred to as color. Some of the most important experimental observations predicted by QCD are asymptotic freedom, confinement and scaling violation.

The SU(N_c)[1] group is a Lie group specified by $N_c^2 - 1 = 8$ generators T^a, which, together with the commutation relations

$$\left[T^a, T^b\right] = i f^{abc} T^c \,, \tag{1.1}$$

form the Lie algebra [4]. f^{abc} are called the structure constants of the group and $a, b, c = 1, \ldots, N_c^2 - 1$ (summation over repeated indices is implied). Among the representations of the SU(N_c) group, two of them are of particular importance. In the *fundamental representation*, generators are the $N_c \times N_c$ matrices $T^a(F) \equiv t^a = \lambda^a/2$, where λ^a are (generalized) Gell-Mann matrices [5]. In the *adjoint representation*, generators have the form of the $N_c^2 - 1 \times N_c^2 - 1$ matrices defined by the structure constants $T^a_{bc} = -i f^{abc}$. Quarks transform under the fundamental representation whereas gluons under the adjoint representation. For any representation R, one can construct a Casimir operator that commutes with the generators

$$\sum_a T^a_{ik}(R) T^a_{kj}(R) = C_R \delta_{ik} \,. \tag{1.2}$$

The value of the constant C_R in the fundamental representation is given by $C_F = \frac{N_c^2 - 1}{2N_c}$ and in the adjoint representation by $C_A = N_c$.

The QCD Lagrangian density is given by [4]

$$\mathcal{L}_{Classical} = \sum_{flavors} \bar{q}_i \left(i \not{D} - m\right)_{ij} q_j - \frac{1}{4} F^{a\,\mu\nu} F^a_{\mu\nu} \,, \tag{1.3}$$

where $F^a_{\mu\nu}$ is the field-strength tensor derived from the gauge field A^a_μ,

$$F^a_{\mu\nu} \equiv \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g_s f^{abc} A^b_\mu A^c_\nu \,. \tag{1.4}$$

$\not{D}$ is a symbolic notation for $\gamma_\mu D^\mu$, where γ_μ are the Dirac matrices satisfying the anticommutation relations $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$, with the metric $g^{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$. The covariant derivative takes the form

$$D_\mu = \partial_\mu + i g_s A^a_\mu T^a \,. \tag{1.5}$$

The fundamental parameters of QCD embedded in the Lagrangian are the masses of different quark flavors and the coupling constant g_s, which determines the strength of the interaction. It is useful to introduce also the quantities

$$\alpha_s = \frac{g_s^2}{4\pi} \quad \text{and} \quad \bar{\alpha}_\mathrm{s} = \frac{\mathrm{N_c}}{\pi} \alpha_\mathrm{s} \,. \tag{1.6}$$

Certain features of the theory can be inferred directly from analyzing the structure of the Lagrangian above. The last term of Eq. (1.4) gives rise to triplet and quartic gluon vertices, which distinguish QCD from Quantum Electrodynamics (QED). Gluons carry color and interact among themselves, unlike photons, which have no charge. These self-interactions lead ultimately to the properties of asymptotic freedom and the existence of jets. In the former case, gluon splittings are the source of an antiscreening effect, which manifests itself as the weakening of the coupling with decreasing distance. In the latter case, gluon splittings

[1] We leave N_c unspecified whenever possible.

play a major role in the creation of large number of particles, all originating from the same initial parton.

Another property of the theory that is evident from the Lagrangian, is gauge invariance. The theory is constructed in a manner that ensures it is invariant under local gauge transformations [6]

$$q(x) \to U(x)q(x)\,, \qquad \bar{q}(x) \to \bar{q}(x)U^{-1}(x)\,, \tag{1.7}$$

$$A_\mu(x) \to U(x)A_\mu(x)U^{-1}(x) + \frac{i}{g_s}\left[\partial_\mu U(x)\right]U^{-1}(x)\,, \tag{1.8}$$

where $U(x) = e^{i\alpha^a(x)t^a}$ and the $\alpha^a(x)$ are arbitrary real functions. Actually, the Lagrangian (1.3) can be derived directly from the gauge symmetry of the form above [7]. This property means that one can perform a redefinition of the color quark fields independently at every point in space-time without changing the physical content of the theory. Even though the quantization procedure explicitly breaks the invariance, its effect is still present in the fact that physical observables should be independent of the gauge choice used to construct the gluon propagator. The same is known to be true also for QED, however, what is distinctive in non-abelian theories, is that the field-strength tensor is not gauge invariant and it transforms like

$$F_{\mu\nu}(x) \to U(x)F_{\mu\nu}(x)U^{-1}(x)\,.$$

This is, again, because of the self-interaction of gluons. As a consequence, it is not possible to unambiguously define states with a definite number of gluons. Gauge transformation can turn a state with a quark and no gluons to a state with a quark and many gluons, so that the partonic picture of processes involving hadrons depends on it.

The Lagrangian (1.3) represents only the classical part of the theory and needs to be quantized. The quantization procedure requires addition of supplementary terms, namely

$$\mathcal{L}_{QCD} = \mathcal{L}_{Classical} + \mathcal{L}_{Gauge} + \mathcal{L}_{Ghost}\,. \tag{1.9}$$

The gauge-fixing term is necessary to avoid double counting of equivalent gluon-field configurations, so that the gluon propagator can be defined. Moreover, if the gauge-fixing term is of the covariant type, also the so-called ghost term is needed to cancel unphysical degrees of freedom. In axial gauges, however, ghosts are absent.

The main measurable quantities that we are interested in calculating in particle physics are decay widths and scattering cross sections. They both depend on the squared *matrix element* or *amplitude* for the transition between some initial and final states, which contains all the dynamical information describing interactions between the involved particles. Amplitudes can be evaluated as a perturbation series over the coupling constant, provided that the coupling is small enough. This theoretical framework is called perturbative QCD (pQCD). The perturbation series is conveniently represented using Feynman diagrams, consisting of particle propagators and interaction vertices, which, together with Feynman rules for translating diagrams into mathematical expressions, are extracted from the Lagrangian. In the region of large coupling, other formalisms are required, such as *e.g.* lattice QCD. In this approach, the theory is formulated in the Euclidean space-time, discretized into lattice with spacing a, with quark fields placed on sites, and gauge fields on the links between sites. The continuum theory is recovered by taking the limit of vanishing lattice spacing. An important feature of the lattice formulation of QCD, which leads to considerable simplifications, is that it preserves gauge invariance without any need for gauge-fixing. The building blocks for calculations are expectation values of various operators associated, for instance, with hadronic masses or quark - antiquark potential.

1.2 Renormalization and running coupling

The value of the coupling is essential from the standpoint of applicability of pQCD. Nevertheless, the property of running coupling is not straightforward from the Lagrangian formulation and in order to understand its origins, it is necessary to address the removal of ultraviolet divergences and the renormalization procedure. Calculations of amplitudes at higher orders involve loop diagrams containing momentum integrals, which are divergent as momentum goes to infinity (UV divergences). These integrals can be formally calculated, *e.g.* by introducing a cutoff parameter μ_R. Such procedure is called regularization and exists in several different versions. The obtained expressions consist of finite terms and terms which are infinite when the cutoff is removed. Divergent terms appear because the coupling g_s from the Lagrangian, called the *bare coupling*, is not a correct expansion parameter and needs to be redefined. The bare coupling absorbs divergent terms and gives finite, experimentally measured physical coupling. This procedure is called *renormalization*, and it needs to be applied also to the bare mass and bare fields from the Lagrangian. Ultimately, the renormalized Lagrangian splits into two parts. The first part has the form identical to the original Lagrangian, but with the physical (renormalized) quantities instead of bare, whereas the second part contains formally divergent counterterms. These counterterms cancel the divergences from the bare-like part, so that the perturbation series based on the renormalized Lagrangian gives finite results. The artifact of this procedure is the dependence of the Lagrangian on the renormalization scale μ_R and regularization method. Physical observables, however, cannot depend on μ_R, and this requirement leads to the *renormalization group equations*.

The renormalization group equation that determines the evolution of the coupling reads

$$\mu_R^2 \frac{\partial \alpha_s}{\partial \mu_R^2} = \beta(\alpha_s)\,. \tag{1.10}$$

The function $\beta(\alpha_s)$ introduced above has a perturbative expansion, and at one-loop order it reads [3, 8]

$$\beta(\alpha_s) = -\frac{(11N_c - 2N_f)\alpha_s^2}{12\pi}\,, \tag{1.11}$$

where N_f is the number of active flavors. We can see that for $N_f \leq 16$ this function is negative, which means that increasing the energy scale (or decreasing the distance scale) lowers the value of the renormalized coupling. This is a fundamental property of the theory, called asymptotic freedom, since due to the smallness of the coupling at high energies quarks behave effectively as free particles. In contrast, in QED the coupling grows with the scale. The formula (1.11) for the beta function has two contributions. The term proportional to the number of flavors comes from the contributions of fermion loops to the gluon propagator. This is an analogue of vacuum polarization in QED, which is associated with a screening effect. On the other hand, the term proportional to the number of colors comes from the contribution of gluon loops to the gluon propagator. The opposite effect of gluons as compared to fermions has been argued to be due to the gluons having spin 1 [9]. Although there is no straightforward intuitive explanation for this phenomenon, contribution from gluons dominates and leads to an effective antiscreening.

The change of the value of the coupling with the energy scale can be obtained from (1.10), provided we operate in perturbative region. At one-loop order we get

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu_R^2)}{1 + \alpha_s(\mu_R^2) b \ln(Q^2/\mu_R^2)}\,, \tag{1.12}$$

where $b = (11N_c - 2N_f)/12\pi$. In this expression, the parameters are the renormalization scale μ_R and the value of the coupling at that scale. They can be replaced for a single dimensionful parameter Λ_{QCD}, defined as the scale at which coupling calculated perturbatively would diverge. In terms of this parameter

$$\alpha_s(Q^2) = \frac{1}{b\,\ln(Q^2/\Lambda_{\mathrm{QCD}}^2)}\,. \tag{1.13}$$

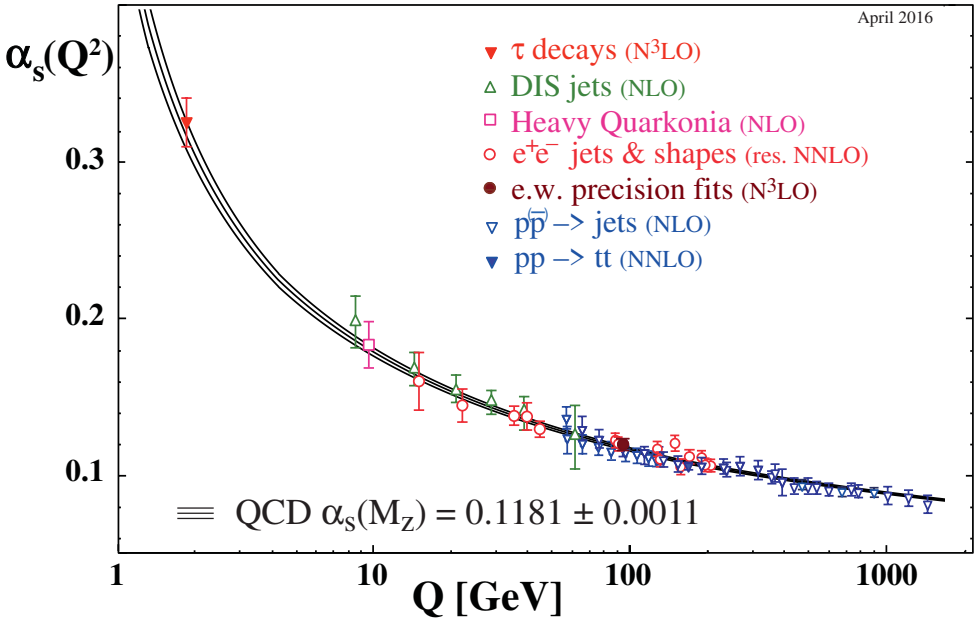


Figure 1: Summary of measurements of α_s as a function of the energy scale Q. The respective degree of QCD perturbation theory used in the extraction of α_s is indicated in brackets (NLO: next-to-leading order; NNLO: next-to-next-to leading order; res. NNLO: NNLO matched with resummed next-to-leading logs; N3LO: next-to-NNLO). Excerpted from [10].

Quantitatively, Λ_{QCD} determines the approximate scale where non-perturbative effects become important and thereby perturbation theory breaks down. Experimentally this scale is found to be approximately 200 MeV, however, its value depends on the precise definition (which involves *e.g.* the number of flavors and renormalization scheme). Perturbation theory is therefore reliable at energy scales $\gtrsim 1$ GeV.

The growth of the coupling for smaller scales predicted by (1.13) is consistent with experimental observation that quarks and gluons never appear as asymptotic states, only color singlet hadrons are detected. The $Q \lesssim \Lambda_{\mathrm{QCD}}$ region cannot be addressed by perturbative methods, however lattice studies show that when a quark and an antiquark are separated by a large distance, the force between them does not fall off with distance, while the potential energy grows linearly. Before quarks can be separated, the stored energy is large enough to create new quark-antiquark pair that combines with the original one, such that color neutrality is preserved. This is called *color confinement*.

Experimental measurements of the coupling constant confirm predictions calculated in the QCD framework. Fig. 1 shows a summary of measurements of α_s as a function of the energy scale Q. One can clearly see the asymptotic behavior, as predicted by the theory. World average value of the coupling at the mass of the Z boson is currently $\alpha_s(M_Z) = 0.1181$ [10]. The results indicate that already at scales of the order of a few GeV, the coupling is small enough to perform perturbative calculations.

1.3 Factorization and distribution functions

The properties of running coupling, described in the previous section, limit the application of pQCD primarily to the study of hard processes, where the exchanged transverse momentum is very large. The

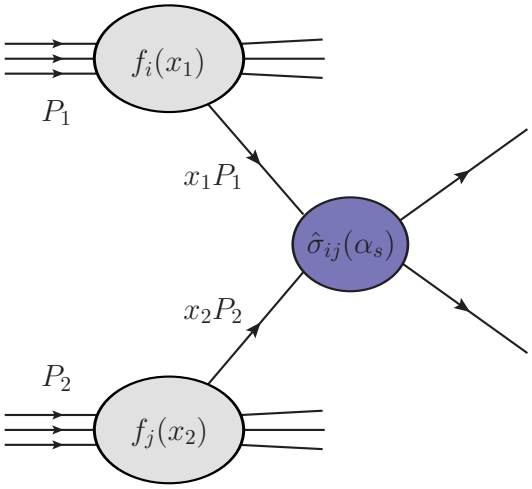


Figure 2: Schematic picture of factorized hadronic cross section.

fundamental degrees of freedom in perturbative calculations are then quarks and gluons rather than hadrons observed in experiments. It is therefore necessary to establish a systematic approach to relate calculations performed at the partonic level to the hadronic states. In the QCD improved parton model, the partonic level cross sections are convoluted with the corresponding probability of finding the interacting parton inside a hadron. In such a setup, the cross section for scattering of two hadrons with four-momenta P_1 and P_2 can be written as [6]

$$\sigma(P_1, P_2) = \sum_{i,j} \int dx_1 dx_2 f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2)\, \hat{\sigma}_{ij}(p_1, p_2, \alpha_s(\mu_R^2), Q^2/\mu_F^2, Q^2/\mu_R^2)\,, \tag{1.14}$$

which corresponds to the structure depicted in Fig. 2. The parton model picture is most conveniently formulated in the infinite momentum frame in which the hadron is moving very fast. In this frame the hadron is a collection of constituents moving almost parallel to each other and carrying a fraction x of hadron's momentum. The functions $f_i(x, \mu_F^2)$ are thus (collinear) *parton distribution functions* (PDFs) describing the probability to find a specific parton inside a hadron and $p_i = x_i P_i$ are momenta of partons entering short-distance cross section $\hat{\sigma}_{ij}$. Q denotes the characteristic scale of the hard scattering, *e.g.* the mass of an electroweak boson or the transverse momentum of a jet. In the perturbative regime, the short-distance cross section can be calculated as a series in the strong coupling [6]

$$\hat{\sigma} = \alpha_s^k(\mu_R^2) \sum_{m=0}^{n} \alpha_s^m(\mu_R^2)\, \hat{\sigma}^{(m)} \left(p_1, p_2, \frac{Q^2}{\mu_F^2}, \frac{Q^2}{\mu_R^2} \right)\,, \tag{1.15}$$

with the leading power k depending on the hard process. Beyond the leading order, the short-distance cross section receives contributions that contain soft (infrared) and collinear singularities (beside UV singularities resolved by renormalization, as discussed earlier). In general, soft divergences arise from the emission of particles with vanishing four-momentum, whereas collinear ones are related to the splitting of massless particles at small angles. We cannot distinguish soft emissions and collinear splittings from events in which these emissions and splittings are absent. It was shown, originally in QED by Bloch and Nordsieck [11] and later proven for QCD by Kinoshita, Lee and Nauenberg (KLN) [12,13], that the infrared and collinear (IRC) singularities cancel after performing summation over degenerate states. In QCD, KLN theorem requires summation over both initial and final state degeneracies and this cannot be done for hadrons in the initial state. When this is the case, singular contributions are factored out from the short-distance cross section

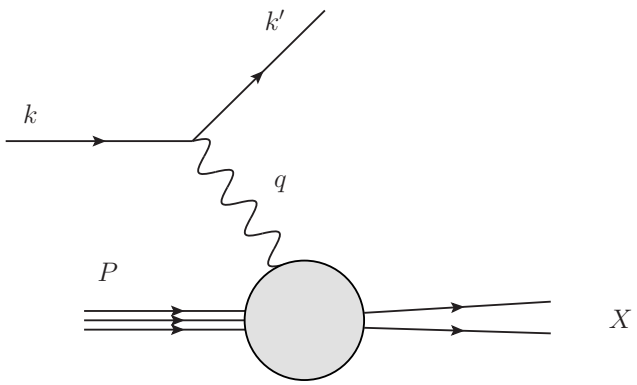


Figure 3: Kinematics of deep inelastic charged lepton-proton scattering.

and absorbed into PDFs. After this operation, the short-distance cross section becomes dependent only on high momentum transfers, hence short times and distances, and is not influenced by low-momentum scales. Furthermore, the factorization of singularities, order by order, renders both the short-distance cross section and the PDFs dependent on an arbitrary parameter, the factorization scale μ_F. It can be interpreted as a transverse momentum scale separating long- and short-distance physics. Partons emitted with transverse momentum below this scale are considered to be part of the hadron structure and are absorbed into a PDF, while the rest belongs to the short-distance cross section. These two components depend on the factorization scale in such a way that the resulting hadronic cross section is scale independent, which leads to evolution equations for parton distributions. Usually, factorization and renormalization scales are set to a single scale $\mu = \mu_R = \mu_F$, with the standard choice $\mu = Q$, the hard scattering scale.

The formula for hadronic cross section (1.14) should in principle include also the fragmentation functions, which describe the probability of producing a specific hadron from the hadronization of an outgoing parton. The main focus of this thesis are however inclusive cross sections, where summation over all possible hadronic final states is performed and fragmentation functions are absent.

Factorization theorems enable us to use QCD as a predictive calculational tool with controllable approximations. Some details of this very broad subject are presented in Section 2. We will see that the formula (1.14) is just the leading contribution of an expansion in terms of the inverse powers of the hard scale and that the probabilistic interpretation of the PDFs is not necessary. Their definition requires only the ability of factoring out the non-perturbative effects.

Collinear parton distribution functions are inherently non-perturbative and cannot be calculated from first principles. These functions are however universal in the sense that they are process-independent, *i.e.* the same distribution can be used for any hard process which involves hadrons. In practice, the distributions are measured in deep inelastic scattering (DIS) experiments and then used to calculate cross sections for hadron-hadron collisions. Since the probe in DIS, namely a photon or electroweak boson, is color-neutral and does not interact strongly, such experiments are perfect to determine parton distribution functions of a hadron. In the following, we present a brief description of this process.

In deep inelastic scattering a lepton scatters off a hadron via the exchange of an electroweak boson. Highly energetic boson penetrates deep into the target hadron and transfers large energy-momentum that knocks a quark out and causes hadron to break up inelastically. Let us focus on DIS of an electron scattering off a proton through the exchange of a virtual photon[2], as depicted in Fig. 3. An incident electron with four-momentum k^μ interacts with a proton with four-momentum P^μ, with the outgoing particles being the electron with four-momentum k'^μ and hadronic state X. The emitted photon carries four-momentum q^μ equal to the change of the electron momentum $q^\mu = k^\mu - k'^\mu$. Conventionally these kinematic variables are

[2]We will neglect the Z boson contribution for simplicity.

expressed in terms of the following Lorentz invariant quantities

$$Q^2 \equiv -q^2\,, \qquad x \equiv \frac{Q^2}{2P\cdot q}\,, \qquad y \equiv \frac{P\cdot q}{P\cdot k}\,, \tag{1.16}$$

where Q^2 is the virtuality of the photon, x is known as the Bjorken-x variable that measures the inelasticity of the process, $0 \leq x \leq 1$, with $x = 1$ corresponding to the elastic scattering, and y is the fraction of electron's energy transferred to the proton in the proton rest frame, $0 \leq y \leq 1$. Using the above invariants, and neglecting the proton mass, the DIS cross section can be parametrized in the following way

$$\frac{d^2\sigma}{dxdQ^2} = \frac{4\pi\alpha_{em}^2}{Q^4}\left[[1+(1-y)^2]F_1 + \frac{(1-y)}{x}(F_2 - 2xF_1)\right]\,, \tag{1.17}$$

where $F_i(x,Q^2)$ are the *structure functions*, which contain the information about the proton probed by the virtual photon. In the limit of $Q^2 \to \infty$, with x fixed, the structure functions satisfy a scaling behavior, called *Bjorken scaling* [14, 15], *i.e.* they depend only on the dimensionless variable x, which, if we neglect the transverse momentum of a parton, has a physical interpretation as the fraction of the proton momentum carried by the struck parton. Scaling implies that the constituents of the proton have no associated length scale and therefore are pointlike. Moreover, structure functions are related to each other, $F_2(x) = 2xF_1(x)$ (the Callan-Gross relation [16]), which is a consequence of quarks carrying spin 1/2. F_2 can be written in terms of the quark and antiquark densities, f_f and $f_{\bar{f}}$, and their couplings to the photon (electric charges e_f) as[3]

$$F_2(x,Q^2) = x\sum_f e_f^2\left[f_f(x,Q^2) + f_{\bar{f}}(x,Q^2)\right]\,, \tag{1.18}$$

where the sum runs over all quark flavors.

The inclusion of QCD corrections leads to the violation of the Bjorken scaling, as well as the Callan-Gross relation. Quarks can emit and absorb gluons, which subsequently can split into quark-antiquark pairs or gluon pairs. More and more of these partons are resolved as the virtuality of the photon increases, since the resolving power depends on the wavelength $\lambda = \frac{\hbar}{Q}$. When Q^2 increases, we observe a depletion of quarks at large x and a corresponding accumulation at smaller x. This is reflected in the shape of F_2. Fig. 4 shows F_2 as a function of Q^2 for different values of x.

1.4 Parton evolution

The parton distribution functions are non-perturbative objects, however their dependence on both arguments can be studied by perturbative methods and leads to a set of evolution equations. In order to derive these equations one needs to take into account corrections that come from considering multiple parton branchings. In practice, the calculations are performed with certain approximations that restrict the phase space of radiation and are valid in regions of x and Q^2 where the selected contributions dominate. Basically there are two regions. The first one is the collinear region, which gives logarithmic enhancements of the form $\alpha_s \ln(\frac{Q^2}{Q_0^2})$. The second one is the soft region, which gives logarithmic enhancements of the form $\alpha_s \ln(\frac{1}{x})$. The overlap of these two gives double logarithmic enhancements of the form $\alpha_s \ln(\frac{Q^2}{Q_0^2})\ln(\frac{1}{x})$. Each enhancement can be related to the phase space region where successive partons have strongly ordered transverse and/or longitudinal momenta [17]. Fig. 5 shows the $\ln(Q^2) - \ln(\frac{1}{x})$ plane and the regions of validity of different evolution equations.

In general, the individual contributions to the evolution equation can be represented by so-called ladder diagrams, as illustrated in Fig. 6 [19]. The evaluation of such diagram requires integrations of functions that describe the dynamics of the emitted partons (splitting functions) over internal momenta exchanged between the rungs. We work in an axial gauge, in which the gluon has only the two physical polarization states and

[3] In DIS factorization scheme; in other schemes such relation holds only at leading order perturbation theory.

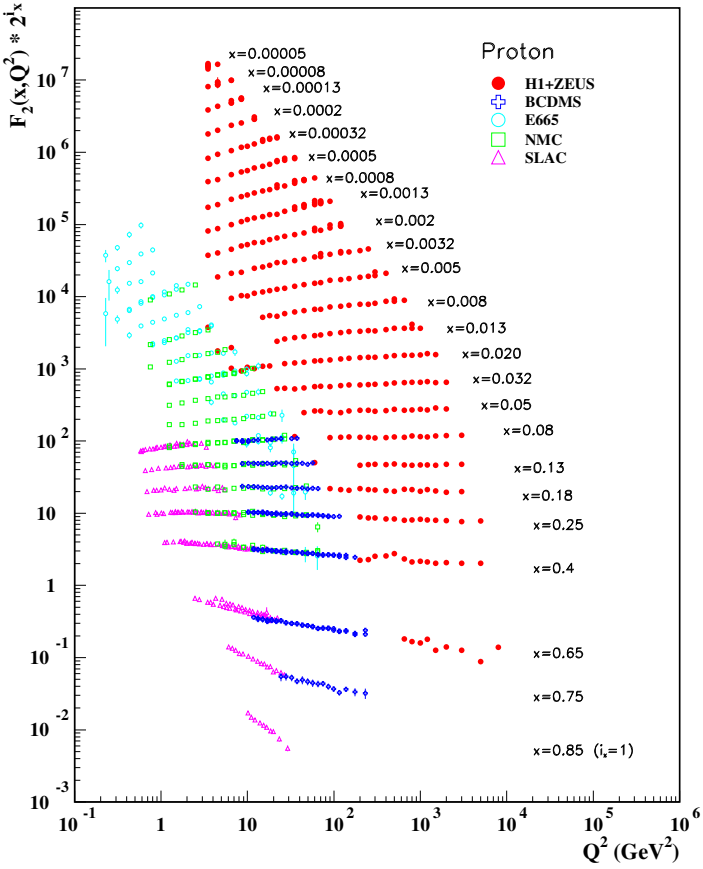


Figure 4: The proton structure function F_2 (for the purpose of plotting multiplied by the factor 2^{i_x}, where i_x is the number of the x bin). From [10].

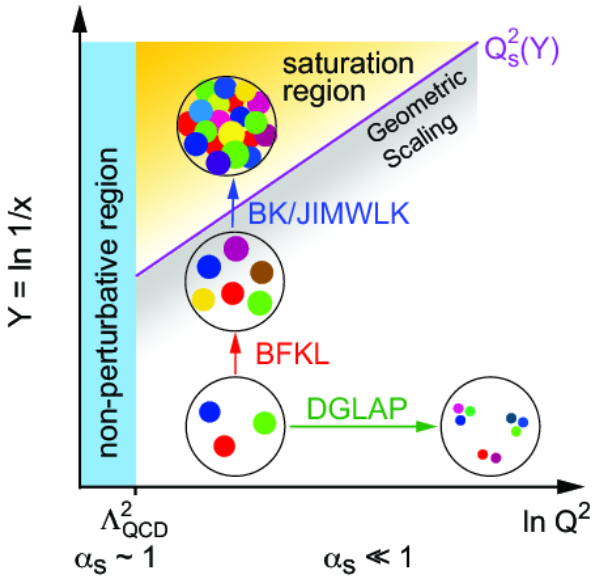


Figure 5: Diagram showing the QCD evolution of the partonic structure of the proton and the validity range for the different evolution equations [18].

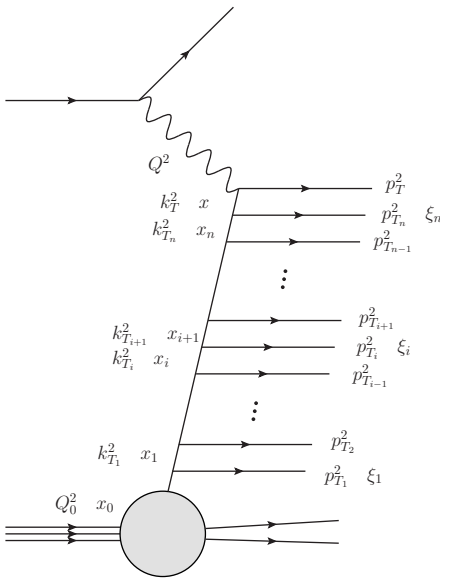


Figure 6: Ladder diagram with n rungs.

ghosts contributions are not needed, and in infinite momentum frame of the proton. The emitted partons (rungs of the ladder) carry the transverse momenta p_{T_i} and fractions of proton longitudinal momentum ξ_i, whereas for side rails these components are denoted respectively by k_{T_i} and x_i. Energy-momentum conservation implies that $x_i = x_{i+1} + \xi_{i+1}$ and thus $x_i > x_{i+1}$.

1.4.1 DGLAP

The approximation that leads to the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) equations [20–22] sums the ladder diagrams in which the transverse momenta of the side rails are strongly ordered, *i.e.* $Q^2 \gg k_{T_n}^2 \gg \ldots \gg k_{T_1}^2 \gg Q_0^2$. Then the virtualities of the partons in the side rails are very small in comparison to the virtuality of the parton that enters the interaction and can be neglected, so the nested integration over all n rungs can be performed

$$\left(\frac{\alpha_s}{2\pi}\right)^n \int_{Q_0^2}^{Q^2} \frac{dk_{T_n}^2}{k_{T_n}^2} \int_{Q_0^2}^{k_{T_n}^2} \frac{dk_{T_{n-1}}^2}{k_{T_{n-1}}^2} \cdots \int_{Q_0^2}^{k_{T_3}^2} \frac{dk_{T_2}^2}{k_{T_2}^2} \int_{Q_0^2}^{k_{T_2}^2} \frac{dk_{T_1}^2}{k_{T_1}^2} = \left(\frac{\alpha_s}{2\pi}\right)^n \frac{1}{n!} \ln^n \left(\frac{Q^2}{Q_0^2}\right) . \tag{1.19}$$

The result is an expression $\propto \left[\alpha_s \ln\left(\frac{Q^2}{Q_0^2}\right)\right]^n$ in which α_s is multiplied by a logarithm that can be large when the difference of scales is large. These logarithms balance the smallness of α_s. Since α_s decreases logarithmically with Q^2 and is compensated by logarithmically growing term in Q^2, in a perturbative expansion, all graphs with rungs up to $n = \infty$ must be summed. This is called the leading log approximation in $\ln\left(\frac{Q^2}{Q_0^2}\right)$, because each power of the strong coupling is accompanied by the same power of the logarithm. Such approximation holds when Q^2 is large, but x is not too small to produce also the large logarithms of $\frac{1}{x}$,

$$\alpha_s \ln(\frac{1}{x}) \ll \alpha_s \ln\left(\frac{Q^2}{Q_0^2}\right) \lesssim 1 . \tag{1.20}$$

The evolution equations in this approximation - the DGLAP equations for the gluon and quark densities

are

$$\begin{aligned}\frac{dq_f(x,Q^2)}{d\ln Q^2} =&\frac{\alpha_s}{2\pi}\int_x^1\frac{dz}{z}\left[q_f(z,Q^2)P_{qq}(\frac{x}{z})+g(z,Q^2)P_{qg}(\frac{x}{z})\right],\\ \frac{dg(x,Q^2)}{d\ln Q^2} =&\frac{\alpha_s}{2\pi}\int_x^1\frac{dz}{z}\left[\sum_f q_f(z,Q^2)P_{gq}(\frac{x}{z})+g(z,Q^2)P_{gg}(\frac{x}{z})\right],\end{aligned}\tag{1.21}$$

for each quark flavor f. They contain the Altarelli-Parisi splitting functions [21], $P_{ij}(z)$, which are interpreted as the probability that a parton of type j emits a collinear parton of type i, with the momentum fraction z of the parent parton. The (unregularized) splitting functions at LO are given by

$$\begin{aligned}P_{qq}(z) =&C_F\,\frac{1+z^2}{1-z}\,,\\ P_{qg}(z) =&T_R\left[z^2+(1-z)^2\right],\qquad T_R=\frac{1}{2}\,,\\ P_{gq}(z) =&C_F\,\frac{1+(1-z)^2}{z}\,,\\ P_{gg}(z) =&2C_A\left[\frac{z}{(1-z)}+\frac{1-z}{z}+z(1-z)\right].\end{aligned}\tag{1.22}$$

The DGLAP equations are coupled integro-differential equations that allow one to calculate quark and gluon densities for any value of Q^2 and $x>x_0$, provided that the densities are known at a particular value of Q_0^2 for $x>x_0$.

The DGLAP equations apply also in the overlapping region of large Q^2 and small x. In this region, besides strong k_T ordering, also strong ordering in x is required, $x\ll x_n\ll\cdots\ll x_1\ll x_0$. The large logarithmic terms that arise from the integration need to be summed and have the form $\propto\left[\alpha_s\ln\left(\frac{Q^2}{Q_0^2}\right)\ln(\frac{1}{x})\right]^n$. This is the double leading log (DLL) approximation relevant when

$$\left.\begin{matrix}\alpha_s\ln\left(\frac{Q^2}{Q_0^2}\right)\\ \alpha_s\ln(\frac{1}{x})\end{matrix}\right\}\ll\alpha_s\ln\left(\frac{Q^2}{Q_0^2}\right)\ln(\frac{1}{x})\lesssim 1\,.\tag{1.23}$$

At small-x, the gluon dominates, so in the first approximation quarks can be neglected. Then, the DGLAP equation can be written as

$$\frac{dg(x,Q^2)}{d\ln Q^2}=\frac{\alpha_s}{2\pi}\int_x^1\frac{dz}{z}P_{gg}\left(\frac{x}{z}\right)g(z,Q^2)\,,\tag{1.24}$$

which can be solved directly, *e.g.* with the help of the Mellin transform. To obtain the solution from the summation of ladder diagrams, let us look at a contribution from a single rung. At small-x, $P_{gg}(z)\approx\frac{2N_c}{z}$, see (1.22), so each rung produce a factor [23]

$$\int\frac{dx_{i-1}}{x_{i-1}}\int dk_{T_i}^2\left\{\frac{\alpha_s}{2\pi}\frac{1}{k_{T_i}^2}P_{gg}\left(\frac{x_i}{x_{i-1}}\right)\right\}\approx\frac{N_c\alpha_s}{\pi}\int\frac{dx_{i-1}}{x_{i-1}}\int\frac{dk_{T_i}^2}{k_{T_i}^2}\left(\frac{x_{i-1}}{x_i}\right).\tag{1.25}$$

The integrations over the transverse and longitudinal momenta can be performed separately. As before, from the strongly k_T-ordered region nested integrations give

$$\left(\frac{N_c\alpha_s}{\pi}\right)^n\int_{Q_0^2}^{Q^2}\frac{dk_{T_n}^2}{k_{T_n}^2}\int_{Q_0^2}^{k_{T_n}^2}\frac{dk_{T_{n-1}}^2}{k_{T_{n-1}}^2}\cdots\int_{Q_0^2}^{k_{T_2}^2}\frac{dk_{T_1}^2}{k_{T_1}^2}=(\bar\alpha_s)^n\,\frac{1}{n!}\ln^n\left(\frac{Q^2}{Q_0^2}\right),\tag{1.26}$$

whereas from the strongly x-ordered region, we get

$$\int_x^1\frac{dx_n}{x_n}\cdots\int_{x_2}^1\frac{dx_1}{x_1}\int_{x_1}^1\frac{dx_0}{x_0}x_0g(x_0,Q_0^2)=\frac{1}{n!}\ln^n\left(\frac{1}{x}\right)G_0\,,\tag{1.27}$$

where G_0 is the small-x limit of $xg(x,Q_0^2)$. The sum of the ladder diagrams that gives the DLL approximation for $xg(x,Q^2)$ is the sum of products of the above two terms. The summation gives a modified Bessel function, which using the large-argument asymptotics can be written as

$$xg(x,Q^2) = G_0 \sum_n \left(\frac{1}{n!}\right)^2 \left(\bar{\alpha}_s \ln\left(\frac{Q^2}{Q_0^2}\right) \ln\left(\frac{1}{x}\right)\right)^n \sim G_0 \exp\left(2\sqrt{\bar{\alpha}_s \ln\left(\frac{Q^2}{Q_0^2}\right) \ln\left(\frac{1}{x}\right)}\right) . \tag{1.28}$$

The solution that takes into account the running of the coupling is also known [24], and it reads

$$xg(x,Q^2) \approx xg(x,Q_0^2) \exp \sqrt{\frac{4N_c}{\pi b} \ln\left[\frac{\ln(Q^2/\Lambda_{\rm QCD}^2)}{\ln(Q_0^2/\Lambda_{\rm QCD}^2)}\right] \ln\left(\frac{1}{x}\right)} . \tag{1.29}$$

If the initial distribution of gluons is too steep at small-x, however, these solutions will not hold. The resulting gluon density shows a strong growth in the small-x region. Actually, it grows faster than any power of $\ln(\frac{1}{x})$.

Similarly to the renormalization group equations, DGLAP equations enable us to calculate the change of the distribution function with the scale. Nevertheless, the absolute value at a given scale cannot be determined without specifying the initial condition, which is not provided by the perturbative theory itself and has to be fitted to the data. The determination of PDFs has great practical importance and many groups, like *e.g.* MMHT2014 [25], NNPDF [26] and CTEQ-TEA [27], have been continuously working on improving their predictions as new experimental data become available. The usual procedure to determine PDF is the following. The quark and gluon distribution functions are parametrized at some initial scale $Q_0 \sim 1$ GeV by a general expression of the form

$$xf_f(x,Q_0^2) = x^{a_1}(1-x)^{a_2} P(x,\{a_3,\ldots,a_n\}) , \tag{1.30}$$

where P is some polynomial and a_i are parameters which need to be determined from the fit. These initial PDFs are then evolved, using DGLAP equations, to the scale Q_i which correspond to the i^{th} data point and used to calculate theoretical predictions and χ_i^2 function for that point. This operation is repeated for all experimental points and the total $\chi^2 = \sum_i \chi_i^2$ is calculated and then minimized to obtain the final values of parameters a_i. As an example, Fig. 7 shows the (unpolarized) parton distribution functions $xf(x,Q)$ for fixed values of $Q = 8$ GeV and $Q = 85$ GeV [27].

1.4.2 BFKL

In the region of small x and moderate Q^2, not large enough to reach the DLL regime, the leading $\ln(\frac{1}{x})$ terms need to be summed while keeping the full Q^2-dependence. This means that we have strongly ordered x without strong ordering of k_T, which need to be integrated over the full range. As a result we work with the *unintegrated* gluon distribution $\mathcal{F}(x,k_T^2)$ (since gluons dominate at small-x), related to the usual gluon distribution in the DLL limit via

$$xg(x,Q^2) = \int_0^{Q^2} dk_T^2 \mathcal{F}(x,k_T^2) . \tag{1.31}$$

The leading $\ln(\frac{1}{x})$ behavior can be also viewed as a sum of ladder diagrams, however, in this case, calculations are more complicated and the ladder diagrams are only an effective representation for an entire set of Feynman diagrams, which were originally summed by Balitsky, Fadin, Kuraev and Lipatov [29–31]. This approximation is valid in the region where

$$\alpha_s \ln\left(\frac{Q^2}{Q_0^2}\right) \ll \alpha_s \ln\left(\frac{1}{x}\right) \lesssim 1 . \tag{1.32}$$

The BFKL equation is an evolution equation in x for the unintegrated gluon distribution and at the leading order it is given by

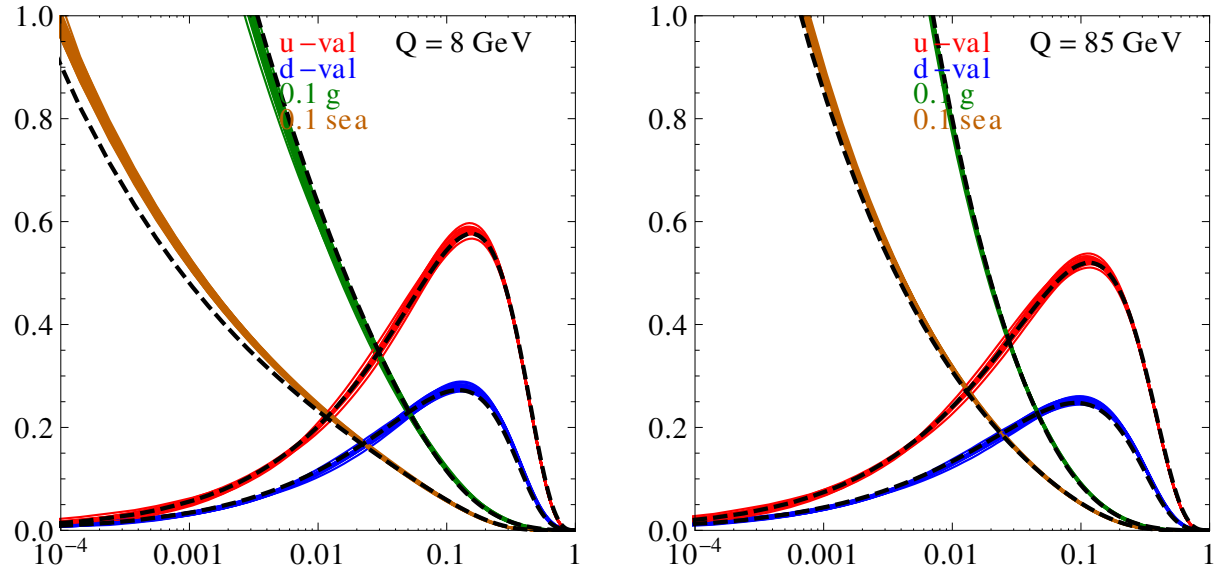


Figure 7: CT10NNLO parton distribution functions [27]. Each graph shows $xu_{\rm val} = x(u-\bar{u})$, $xd_{\rm val} = x(d-\bar{d})$, $0.1xg$ and $0.1xq_{\rm sea}$ as functions of x for a fixed value of $Q = 8$ (left) and $Q = 85$ GeV (right). The quark sea contribution is $q_{\rm sea} = 2(\bar{d}+\bar{u}+\bar{s})$. The dashed curves are the central CT10NLO fit [28].

$$\frac{\partial \mathcal{F}(x,k_T^2)}{\partial \ln(1/x)} = \bar{\alpha}_s \int_0^\infty \frac{dk_T'^2}{k_T'^2}\left[\frac{k_T'^2\mathcal{F}(x,k_T'^2) - k_T^2\mathcal{F}(x,k_T^2)}{|k_T'^2 - k_T^2|} + \frac{k_T^2\mathcal{F}(x,k_T^2)}{\sqrt{4k_T'^4 + k_T^4}}\right]. \tag{1.33}$$

One can also introduce dimensionless gluon density $f(x,k_T^2) = k_T^2\mathcal{F}(x,k_T^2)$, for which the BFKL equation reads

$$\frac{\partial f(x,k_T^2)}{\partial \ln(1/x)} = \bar{\alpha}_s k_T^2 \int_0^\infty \frac{dk_T'^2}{k_T'^2}\left[\frac{f(x,k_T'^2) - f(x,k_T^2)}{|k_T'^2 - k_T^2|} + \frac{f(x,k_T^2)}{\sqrt{4k_T'^4 + k_T^4}}\right]. \tag{1.34}$$

Once $\mathcal{F}(x,k_T^2)$ is known at some value of x_0 for all k_T^2, it can be calculated at smaller values of x. For fixed α_s, the BFKL equation can be solved analytically to give (in the saddle point approximation)

$$\mathcal{F}(x,k_T^2) \sim \left(\frac{x}{x_0}\right)^{-\lambda} \frac{\left(k_T^2\right)^{\frac{3}{2}}}{\sqrt{2\pi\lambda" \ln(x_0/x)}} \exp\left[\frac{-\ln^2(k_T^2/\bar{k}_T^2)}{2\lambda" \ln(x_0/x)}\right] \sim \left(\frac{x}{x_0}\right)^{-\lambda}, \tag{1.35}$$

with $\lambda = \frac{N_c\alpha_s}{\pi}\cdot 4\ln 2$ and $\lambda" = \frac{N_c\alpha_s}{\pi}\cdot 28\zeta(3)$ (the Riemann zeta function $\zeta(3) = 1.202$). $\bar{k}_T^2$ specifies the starting point of the evolution. We see the characteristic $x^{-\lambda}$ behavior, modulated by a $(\ln(1/x))^{-\frac{1}{2}}$ factor. Hence, the gluon density is expected to rise like a power of $1/x$ for decreasing x, faster than the DLL result (1.29). Actually, the DLL approximation is reproduced when we constrain the k_T's to be strongly ordered, *i.e.* $k_T^2 \gg k_T'^2$ in Eq. (1.33) [23]. If we take the typical value of $\alpha_s = 0.18$ then $\lambda \approx 0.5$, however it decreases when running of α_s and higher order corrections are included [32].

Another characteristic feature of the solutions of the BFKL equation, which is a consequence of no ordering in k_T, is a random walk in k_T of an individual evolution path, such that evolution to smaller x is accompanied by a diffusion in k_T. The expression (1.35) explicitly shows the diffusion pattern, given by a Gaussian distribution in $\ln k_T^2$ with a width that increases as $\sqrt{\ln(1/x)}$. k_T-diffusion poses a problem in the applicability of the BFKL equation, since k_T may diffuse into the infrared region ($k_T < Q_0$), where the equation is not expected to be valid. To avoid this, usually a lower cut-off k_0^2 is imposed on the k_T integration.

The BFKL equation does not properly describe the behavior of structure function F_2, as it predicts that λ

remains fixed instead of rising with Q^2 [33]. It was argued that this is because of not taking sufficient account of momentum conservation and nonperturbative effects from soft gluons [34, 35]. The inclusion of running coupling does not help and only makes the equation unstable [36–38]. The situation is even aggravated by the next-to-leading order correction [39], which turned out to be large and negative and led to negative cross sections. The root of the problem has been traced back to unresummed logarithms of Q^2 [40]. Several proposals to correct the BFKL equation have been made, like resummation [41, 42], imposing momentum conservation [43] or imposing perturbative stability [44]. The overview of these methods can be found in [33].

1.4.3 CCFM

Although the predictions obtained from the BFKL equation and the DGLAP equations are different, both approaches to describe the evolution of the gluon PDF do not contradict each other. They are simply organizing the double power expansion in logarithms in a different way. As we increase the accuracy of each approach, the predictions should approach each other.

The BFKL equation has a limited region of validity, determined by the size of the Q^2 logarithms it resums. For the LLA equation $\alpha_s \ln 1/x \sim 1$, while $\alpha_s \ln \frac{Q^2}{\Lambda^2_{\rm QCD}} \ll 1$. On the other hand, the DGLAP equation, which resums $\left(\alpha_s \ln \frac{Q^2}{\Lambda^2_{\rm QCD}}\right)^n$ terms is not expected to hold at very small values of x. For the cases in which both type of logarithms become sizable, it is important to have an unified way of evolving the DIS structure functions throughout the $x - Q^2$ plane. A theoretical framework which provides such a treatment has been developed by Ciafaloni-Catani-Fiorani-Marchesini (CCFM) [45–47]. It leads to an evolution equation, called the CCFM equation, which reduces to the BFKL equation in the leading $\ln 1/x$ approximation, and is equivalent to the DGLAP equation at moderate x (solutions of the CCFM equation can be found in [48]). The CCFM equation is based on the coherent branching of gluons along a ladder. The emissions are coherent in the sense that there is an angular ordering $\theta_1 < \ldots < \theta_n$ going downwards along the ladder, where θ_i is the polar angle that the i-th gluon forms with the original direction, the one of the first parton emitted by the parent hadron. Outside this kinematic region there is a destructive interference such that the multigluon contributions vanish to leading order.

According to the CCFM evolution equations, the emission of gluons ("ladders") during the initial cascade is only allowed in an angular-ordered region of phase space. The CCFM evolution with respect to the evolution variable $\bar{q}^2$ can be written as [49]

$$\bar{q}^2 \frac{d}{d\bar{q}^2} \frac{x\mathcal{A}(x, k_T^2, \bar{q}^2)}{\Delta_s(\bar{q}^2, Q_0^2)} = \int dz \frac{d\phi}{2\pi} \frac{\tilde{P}(z, (\bar{q}/z)^2, k_T^2)}{\Delta_s(\bar{q}^2, Q_0^2)} x' \mathcal{A}\left(x', k_T^{2'}, (\bar{q}/z)^2\right) , \tag{1.36}$$

where the splitting function $\tilde{P}(z, (\bar{q}/z)^2, k_T^2)$ is related to the two scales $\bar{q}$ and k_T. The introduced Sudakov form factor $\Delta_s(\bar{q}^2, Q_0^2)$ is simply the probability of evolving from Q_0^2 to $\bar{q}^2$ without branching. The unintegrated parton density, written in standard notation $\mathcal{A}\left(x', k_T^{2'}, (\bar{q}/z)^2\right)$ (identical to $g(x, Q^2)$ in the collinear DGLAP picture) describes the probability of finding a parton carrying a longitudinal momentum fraction x and transverse momentum fraction k_T at the factorization scale $\mu = \bar{q}$. The scale

$$\bar{q}_i \equiv \frac{|q_i|}{1 - z_i} , \tag{1.37}$$

is related to the angle of the emitted gluon θ_i, such that

$$\frac{\bar{q}_{i+1}}{\bar{q}_i} = z_i \frac{\theta_{i+1}}{\theta_i} , \tag{1.38}$$

and from using the angular ordering condition $\theta_{i+1} > \theta_i$ one obtains ordering condition for rescaled transverse momenta

$$\bar{q}_{i+1} > z_i \bar{q}_i . \tag{1.39}$$

At small x, the angular ordering does not provide any constraint on the transverse momenta along the gluon ladder, so that $\mathcal{A}$ becomes independent on the scale and one recovers BFKL equation.

1.4.4 BK

The steep growth of parton density is an unwanted feature, because it leads to the cross sections which, at large center-of-mass energy s behave like s^λ, and hence grow faster than allowed by the Froissart-Martin bound [50, 51]

$$\sigma_{tot}(s) \leq \text{const} \cdot \ln^2 s\,, \tag{1.40}$$

which means that unitarity is violated.

When the number of partons is too large their wave functions start to overlap and recombination effects should take place. These phenomena, known as *saturation* effects would limit the growth of the gluon PDF. The scale at which the hadron is so densely packed that the recombination effects are important is called the saturation scale $Q_s(x)$. Saturation scale being greater than Λ_{QCD} ensures that it can be understood within perturbative methods. Balitsky and Kovchegov [52,53] showed that in the large-N_c limit, there is a nonlinear generalization of the BFKL equation, the BK equation.

The BK equation for gluon number density $\Phi(x, k^2)$ reads [54]

$$\Phi(x,k^2) \;=\; \Phi_0(x,k^2) + \bar{\alpha}_s \int_{x/x_0}^1 \frac{dz}{z} \int_0^\infty \frac{dl^2}{l^2} \left[\frac{l^2\Phi(x/z,l^2) - k^2\Phi(x/z,k^2)}{|k^2-l^2|} + \frac{k^2\Phi(x/z,k^2)}{\sqrt{(4l^4+k^4)}} \right] - \frac{\bar{\alpha}_s}{\pi R^2} \int_{x/x_0}^1 \frac{dz}{z} \Phi^2(x/z,k^2)\,. \tag{1.41}$$

As we will later see, $\Phi(x, k^2)$ is one of the basic gluon distributions, called the Weizsäcker-Williams distribution and is related to the unintegrated (dipole) gluon density by

$$\mathcal{F}(x,k^2) = \frac{N_c}{4\alpha_s\pi^2} k^2 \nabla_k^2 \Phi(x,k^2)\,. \tag{1.42}$$

The parameter R in (1.41) controls the strength of the nonlinear term, and can be interpreted as a hadron's radius. The BK equation is a central tool for understanding the initial conditions in hadronic collisions when gluon density approaches unitarity limits.

1.5 Kinematics and acceptance

Let us look at a basic $2 \to 2$ process, in which two incoming particles with four-momenta p_1, p_2 scatter and produce a final state of two particles with four-momenta p_3, p_4. The transverse momentum p_T of a particle can be written as

$$|\mathbf{p}_T| = |\mathbf{p}'| \sin\theta = \sqrt{p_x^2 + p_y^2}\,, \tag{1.43}$$

with θ the angle with respect to the z axis. It is also customary to introduce the Mandelstam variables

$$s = (p_1 + p_2)^2 = (p_3 + p_4)^2\,, \tag{1.44}$$

$$t = (p_1 - p_3)^2 = (p_2 - p_4)^2\,, \tag{1.45}$$

$$u = (p_1 - p_4)^2 = (p_2 - p_3)^2\,, \tag{1.46}$$

satisfying the identity

$$s + t + u = \sum_{i=1}^4 m_i^2\,, \tag{1.47}$$

where m_i are the masses of the particles.

Another frequently used kinematic variable is the rapidity of a particle. It is defined as

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \,. \tag{1.48}$$

The rapidity transforms under a boost in the z direction with velocity β as

$$y \to y' = y - \coth \beta \,, \tag{1.49}$$

hence the shape of the rapidity distribution is invariant with respect to Lorentz boosts along the beam axis. In the high energy limit, or for massless particles, rapidity can be directly related to the angle θ

$$\eta \equiv y|_{m=0} = -\ln \tan \frac{\theta}{2} \,, \tag{1.50}$$

where η is called pseudorapidity of a particle. Particles that are produced at high rapidities have a very small angle θ and are very close to the beam pipes of the accelerator. This region of phase space is called the forward region.

It is useful to introduce the light-cone basis defined using two null four-vectors $n = (1, 0, 0, -1)/\sqrt{2}$ and $\tilde{n} = (1, 0, 0, 1)/\sqrt{2}$. They define the 'plus' and 'minus' components of a four-vector v: $v^+ = n \cdot v$, $v^- = \tilde{n} \cdot v$, so that the four-vector has a decomposition

$$v^\mu = v^+ \, \tilde{n}^\mu + v^- \, n^\mu + v_T^\mu \,. \tag{1.51}$$

The light-cone coordinates are $(v^+, v^-, \vec{v}_T)$, where the Euclidean transverse vector is defined (in canonical coordinates) as $v_T^\mu = (0, \vec{v}_T, 0)$.

1.6 Collider phenomenology

In the previous subsection we presented the description of hadron collisions from a theoretical point of view. Here we will introduce phenomenological models which are needed to describe the experimental data and which complement the factorization approach. Hadronic collisions are usually divided to different components that describe a certain stage in the dynamics of the collision and the Monte Carlo generators, like PYTHIA [56] or *Herwig* [57] are used to simulate these different stages. Fig. 8 represent a typical hadron-collider event. The central part of the event is provided by the hard process (the dark red blob in the figure), and can be calculated in fixed order perturbation theory. This stage of the simulation is managed by computations based on matrix elements, and they are provided by special programs called parton-level or matrix-element generators. The QCD evolution of the process from the hard scale to the hadronization scale (red in the figure) is described by parton showers, which model multiple QCD radiation in some approximation to exact perturbation theory [55]. At the hadronization scale the transition to colorless hadrons occurs (light green blobs). This stage is described by purely phenomenological fragmentation models, which are fitted to data. At the end, hadrons decay (dark green blobs) into particles observed in detectors. Another type of phenomena arises from interactions between hadron remnants, which may undergo secondary hard or semi-hard interactions. Such radiation is called the underlying event (represented by purple blob). Also the interactions between hadrons from the same bunches that cross during the collision may contaminate the results (pileup). The underlying event and pileup have mixed, perturbative and nonperturbative nature [58] and phenomenological models must be employed to describe them.

1.6.1 Parton showers

The partons participating in the hard scattering are accelerated and, just as electric charges emit QED radiation, colored partons will emit QCD radiation in the form of gluons. As gluons themselves carry color

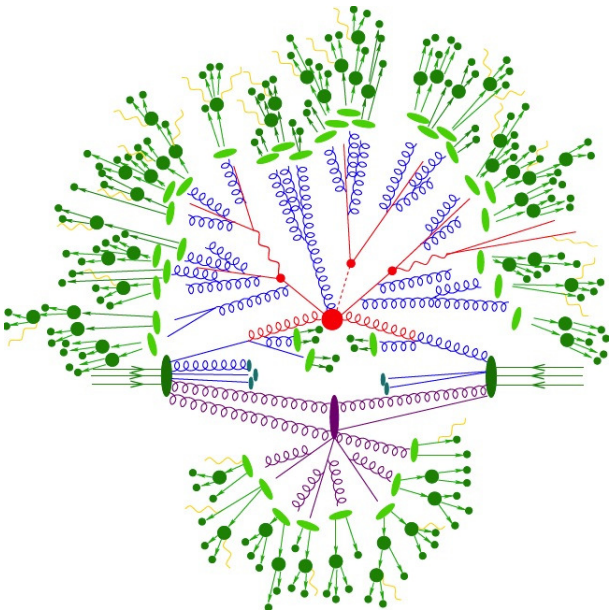


Figure 8: Schematic view of a hadron-hadron collision and its different components. The hard interaction (big red blob) is followed by the decay of the outgoing particles (small red blobs). Additional hard QCD radiation is produced (red) and a secondary interaction takes place (purple blob) before the final-state partons hadronize (light green blobs) and hadrons decay (dark green blobs). Photon radiation occurs at any stage (yellow). From [55].

charges, they emit further radiation, leading to parton showers. In principle, the showers represent higher-order corrections to the hard subprocess. When a parton with 4-momentum p^μ emits radiation, its virtuality $T^2 = -p^2$ changes [19]. The parameter $t = \ln(T^2/\Lambda_{\mathrm{QCD}})$ is the evolution parameter. A final-state shower develops from a parton outgoing from the hard process and its evolution start from the primary parton at high energy and a large virtuality, and loses energy and virtuality until it falls to some small scale at which splitting is terminated (time-like shower). An initial-state shower develops on an incoming parton of the hard process. The constituent parton starts at high energy and low virtuality and evolves to higher virtuality by emitting partons and losing energy (space-like shower). The showering of this parton terminates when it collide to initiate the hard process.

The probability $\mathcal{P}$ that a branching $i \to jk$ will occur with a small change in dt is given by the evolution equation [20–22]

$$\frac{d\mathcal{P}_{i\to jk}}{dt} = \int_0^1 dz \frac{\alpha_s(Q^2)}{2\pi} P_{i\to jk}(z)\,, \tag{1.52}$$

where the functions $P_{i\to jk}(z)$ are the Altarelli-Parisi splitting functions P_{ji}, given by Eq. (1.22).

1.6.2 Hadronization

The parton shower terminates at some low scale of the order of a few GeV, after which the colored partons recombine into final state color-neutral hadrons observed in experiments in a process called hadronization. As a nonperturbative phenomenon, it can be described only by empirical models and not from first principles. There are two hadronization models that are most widely used - the Lund string model [59] and the cluster-hadronization model [60]. The Lund model is based on the observation, from lattice simulations of QCD, that the potential energy of color sources (like quark-antiquark pair) increases linearly with their separation, corresponding to a distance-independent force of attraction. As a quark and antiquark produced in the collision move apart, a color flux tube is stretched between the pair and the stored energy inside this tube

increases linearly. At some point the energy stored is so high that the string breaks and form a new quark-antiquark pair, producing two color-singlet states. In cluster-hadronization model, the basic assumption is a local parton-hadron duality, *i.e.* the idea that quantum numbers on the hadron level follow very closely the flow of quantum numbers on the parton level [61]. Therefore, the clusters are considered as a kind of "hadron matter", carrying the quantum numbers of hadrons.

1.6.3 Jets

The high energy partons originating from the hard scattering are not directly observed by the detectors, since as they separate from the rest of the proton, the value of the coupling constant increases and the QCD force becomes stronger (confinement). This increases the probability of radiation of secondary partons, predominantly at small angles, until a point where a non-perturbative transition combines partons into bound states of color-neutral hadrons. The result is a collimated stream of hadrons, forming roughly conical shape, extending from the point of interaction into the detectors, called a jet. Collective energy and momentum of hadrons building a jet reflect the energy and momentum of the partons taking part in the hard scattering. To ensure such correspondence one needs a precise jet definition.

A complete jet definition consists of a jet algorithm, in conjunction with some parameters and a recombination scheme. Jet algorithm provide a set of rules for grouping particles into jets and the parameters indicate how close two particles must be in order to belong to the same jet. Recombination scheme specifies what momentum to assign to the combination of two particles, and currently a simple 4-momenta sum is almost exclusively used.

Jet algorithms fall into two broad categories: the *cone algorithms* and the *sequential-recombination algorithms*. The cone algorithms can be thought of as a top-down approach to jet finding, because they rely on the idea that QCD branching and hadronization does not change the event's energy flow to a great extent. The sequential recombination algorithms represent a bottom-up approach that repeatedly recombine the closest pair of particles. An important property that a jet algorithm must satisfy is the *infrared* and *collinear* (IRC) *safety*, which states that the set of hard jets that are found in the event should remain unchanged if one modifies the event by a collinear splitting or the addition of a soft emission.

The cone algorithms were historically first, with the Sterman-Weinberg algorithm [62] for e^+e^- , and they were later extensively used at hadron colliders, especially at the Tevatron [63]. Most of them were however plagued with the issues of the IRC unsafety [64]. The problems originated from the need to define seeds in order to start an iterative procedure to search for stable cones. Those seed were identified with final state particles. Such procedure is manifestly IRC-unsafe, as an emission of a soft or collinear parton changes the set of initial seeds, which in turn, for a non-negligible fraction of events, leads to a different set of the final-state jets. Resolution of this long-standing problem came with the Seedless Infrared-Safe Cone jet algorithm (SISCone) [65], where an efficient procedure for finding stable cones, without introducing initial seeds, was proposed.

The sequential recombination algorithms currently dominate in the jet measurements. They combine the closest particles according to a distance measure which can be generally written as

$$d_{ij} = \min\left(p_{T_i}^{2p}, p_{T_j}^{2p}\right)\frac{\Delta R_{ij}^2}{R^2}\,, \qquad d_{iB} = p_{T_i}^{2p}\,, \tag{1.53}$$

where d_{ij} is a distance between the particles i and j and d_{iB} is a distance between the particle i and the beam. The parameter R is called the *jet radius* and

$$\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2 \tag{1.54}$$

is the geometric distance between the particles in the rapidity-azimuthal angle plane. The parameter p defines specific algorithm from the sequential-recombination family: $p = 1$ for the k_T algorithm [66, 67], $p = 0$ for the Cambridge/Aachen (C/A) algorithm [68, 69], and $p = -1$ for the anti-k_T [70] algorithm.

The procedure of finding jets with the sequential-recombination algorithm consists of the following steps:

1. Compute distances between all pairs of final-state particles, d_{ij} , as well as the particle-beam distances, d_{iB}, using the measure from (1.53).

2. Find the smallest d_{ij} and the smallest d_{iB} in the sets of distances obtained above. If $d_{ij} < d_{iB}$, recombine the two particles, remove them from the list of final-state particles, and add the particle ij to that list. If $d_{iB} < d_{ij}$, call the particle i a jet and remove it from the list of particles.

3. Repeat the above procedure until there is no particles left.

Despite the same formula for distance measure used in all three algorithms, they have different properties. The k_T algorithm starts from clustering together low-p_T objects and successively accumulates particles around them. The C/A algorithm is insensitive to the transverse momenta of particles and build up jets by merging particles closest in the $y-\phi$ plane. The anti-k_T algorithm starts from accumulating particles around high-p_T objects, and stops when there are no particles within radius R around the hard center, producing circularly-shaped jets in $y-\phi$ plane [58].

2 Factorization theorems

Factorization is the basic concept that underlies the phenomenological studies of high-energy hadronic physics. QCD factorization theorems allow for the systematic separation of short- and long-distance effects in field theory. The most well-established form of factorization, the collinear factorization [71], briefly discussed qualitatively in the previous section, involves a single large energy scale which dominates the dynamics, and the PDFs are only a function of this scale and of the longitudinal momentum fraction x of the parton. All other relevant energy scales are assumed to be of the same order. Consequently, the processes which can be described using this framework are inclusive processes where it is necessary to sum over a large class of possible final states and all variables that are not of the same order as the hard scale need to be integrated over. Studies of less inclusive processes that depend on a second, smaller scale, showed the need for generalization of PDFs to include the dependence of transverse momentum, giving rise to transverse-momentum dependent (TMD) PDFs, and appropriate extension of collinear factorization into TMD factorization [72, 73]. TMD PDFs, in contrast to their collinear counterparts, are not universal and depend on the hard process under consideration. This is because their operator definitions require insertions of appropriate Wilson lines which resum collinear gluons related to initial and final state interactions. In the kinematic region of small x another type of factorization is needed, the so-called high-energy (HEF) or k_T-factorization [74–77]. It involves universal unintegrated parton distributions (usually unintegrated gluon distribution as gluon dominates at small x), which just like the TMDs depend on the transverse momentum, convoluted with off-shell matrix elements. At even smaller x, *i.e.* in the saturation regime, the high density of gluons allows for a semiclassical treatment of the color field, leading to an effective theory, the Color Glass Condensate (CGC) [78–81]. In the context of dilute-dense collisions, it was shown that by considering the appropriate limits both the TMD and k_T-factorization formulas can be derived from CGC, and a new formalism to interpolate between these two limits has been developed, called the small-x improved TMD factorization (iTMD) [82, 83].

The subject of factorization, despite being intensively studied since the very beginnings of QCD, is still a very active field of research with many open problems. Therefore, a comprehensive review of such a broad topic is beyond the scope of this work. This section is intended to briefly discuss various approaches mentioned above. We start from collinear factorization and consider the most important ingredients required to prove it. Then, the TMD and HEF factorization are introduced. We finish this section by describing factorization in forward jet production.

2.1 Collinear factorization

The first step in constructing a proof of a factorization theorem is to analyze all Feynman diagrams that contribute to the cross section. A general diagram involves integrals over loop momenta and we need to identify the integration regions that give dominant contributions to these integrals. The term dominant refers to the limit of small parameter Λ/Q, where Q is the hard scale of the process and Λ represents small kinematic quantities and the scale of nonperturbative QCD interactions. The dominant contribution is called the leading twist, while the terms suppressed as powers of Λ/Q are called higher twist. The leading contributions come from the integration regions near singularities in the integrand of the graph when all quantities of order Λ are set to zero [84]. However, only those singularities at which the integration contour is pinched between two or more singularities are relevant [85, 86], because otherwise the singularities can be avoided by a deformation of the integration contour. Such configurations of the loop momenta are called *pinch singular surfaces* (PSS). The most general leading PSS can be represented diagrammatically, with the subdiagrams characterized by the momenta of internal lines being either hard, soft or collinear to an external direction, as shown in Fig. 9(left) [58, 87]:

- Collinear subgraphs C_A, C_B describe the incoming partons and the corresponding beam remnants after the hard scattering. The typical momenta of particles in this region, in light-cone basis, scale as

$$p^\mu_{C_A} \sim (Q, \Lambda^2/Q, \Lambda)\,, \qquad p^\mu_{C_B} \sim (\Lambda^2/Q, Q, \Lambda)\,. \tag{2.1}$$

- Hard subgraph H, with the momenta
$$p_H^\mu \sim (Q, Q, Q)\,. \tag{2.2}$$
- Final-state jets $J_1, J_2, \ldots, J_N$, formed after the hard scattering, with momenta collinear to the ones that go out of the hard part.
- Soft subgraph S, with the momenta
$$p_S^\mu \sim (\Lambda, \Lambda, \Lambda)\,. \tag{2.3}$$

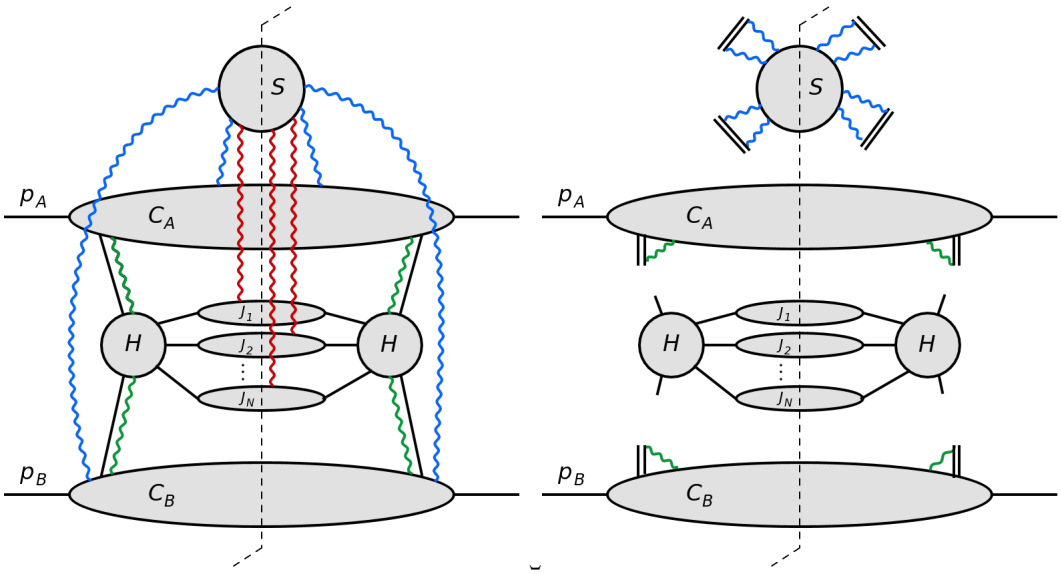


Figure 9: The most general leading pinch singular surface diagram (left) and factorized form of cross section (right) for hadron-hadron collision. From [58].

The subdiagrams above are in general connected by quark and gluon lines. It can be shown, however, that many of these connections contribute only to the higher twist [71], while the remaining must be resummed. The hadronic cross section factorizes, if at leading twist it can be written as a convolution

$$d\sigma = \sum_{i,j} f_{i/A} \otimes f_{j/B} \otimes H_{ij} \otimes S_{ij} \otimes J_1 \otimes \ldots J_N\,, \tag{2.4}$$

where $f_{i/A}$ and $f_{j/B}$ are the already introduced in previous sections collinear (integrated) PDFs, corresponding to the collinear subgraphs and the rest of the notation follows that of Fig. 9. The sets of the soft, S_{ij}, and the hard, H_{ij}, functions are process-dependent, and together form the perturbatively calculated short-distance cross section, as discussed in previous section, Eq. (1.15). In order to prove that factorization applies, one need to show that the expressions corresponding to Fig. 9 can be written in the form of Eq. 2.4. In gauge theories there are the following connections that can spoil the factorization [87], denoted by wavy lines in Fig. 9:

- Soft-gluon connections between the wide angle jets in the hard subdiagram (red lines).
- Soft-gluon connections between the collinear subgraphs C_A and C_B through the soft function S (blue lines).
- Longitudinally-polarized-gluon connections between the collinear subgraphs $C_{A,B}$ and the hard part H (green lines).

The factorization of the first and the second type connections, *i.e.* the soft gluons, can be proved by deformation of the integration contour in the complex plane of the soft gluon momenta, such that one of the longitudinal components dominates along the entire integration path. Subsequently, the non-abelian Ward identities [88, 89] are used (hence, gauge invariance) to show that the corresponding contribution either vanishes or factorizes. In some cases, however, this procedure cannot be applied, because the connecting gluons are pinched in so-called *Glauber region* [87, 90] or the *Coulomb region* [91–93]. The components of the gluon's four-momentum in this region are not comparable, as in Eq. (2.3), but the longitudinal components are much smaller than the transverse one and the necessary contour deformation is impossible. Contributions from that region must vanish in order for the factorization procedure to be successful. The proof that it happens is one of the most technical aspect involved in the proof of factorization. We shall not consider it in details here and only note that the factorization is rescued from the soft gluons pinched in the Glauber/Coulomb region by the sum over final states. It turns out that, after such a sum is performed, the soft gluon connections between the final state jets J_i vanish [87, 94]. These cancellations take place because of unitarity, as those interactions happen long after the hard process. The same situation holds for both the fully inclusive cross sections and the jet production cross sections, because the hard final state jets are well separated in space and they cannot meet again to produce another hard scattering.

In case of the soft connections between the collinear subgraphs C_A and C_B, the pinches in the Glauber region vanish after the sum over final states is performed, so that the deformation of the contour is possible. Then, the application of non-abelian Ward identities turn the $S - C_{A,B}$ connections into connections between the soft function S and the Wilson lines. Therefore, that part of the formula is factorized from the rest, as shown in Fig. 9(right) [58].

The last problem, of the longitudinally polarized gluons connecting H with C_A and C_B, can be solved by absorbing the longitudinal gluons into the parton distribution through the so-called $gauge - links$, using the eikonal propagators [94]. Then again, due to the non-abelian Ward identities, it can be shown that the gluons from C_A and C_B effectively connect to the Wilson line, represented as double line in Fig. 9(right) [58], and factorize from the rest of the diagram. Physically, this is because the collinear gluons cannot resolve any transverse structure of H, which appears to those gluons as a Wilson line.

The above procedure of absorbing longitudinal gluons into the parton distribution functions via the gauge-links leads to the following, gauge invariant definitions of the parton distribution functions [71,84,95] [84,95,96]

$$f(x)_{q/h} = \int \frac{d\xi^-}{2\pi} e^{-ixP^+\xi^-} \langle P|\bar{\psi}(0,0,\mathbf{0}_T)\mathcal{U}^n(0;\xi^-)\frac{\gamma^+}{2}\psi(0,\xi^-,\mathbf{0}_T)|P\rangle\,, \tag{2.5}$$

$$f(x)_{g/h} = \int \frac{d\xi^-}{2\pi P^+} e^{-ixP^+\xi^-} \langle P|F_a^{+j}(0,0,\mathbf{0}_T)\mathcal{U}^n_{ab}(0,\xi^-)F_b^{+j}(0,\xi^-,\mathbf{0}_T)|P\rangle\,, \tag{2.6}$$

where ψ is a Dirac field and F_a^{ij} is a gluon field-strength tensor. The object $\mathcal{U}^n(0;\xi^-)$ in the above expressions is called the $Wilson\ line$ (double line in Fig. 9). It resumms all exchanges of the longitudinal gluons between the hard part and the collinear part. A generic Wilson line joining x and y through a path C is defined as path-ordered exponential

$$\mathcal{U}_C(x;y) = \mathcal{P}\exp\left\{-ig\int_C dz_\mu \hat{A}^\mu(z)\right\}. \tag{2.7}$$

and is a matrix in color space. The Wilson line can be defined also in the adjoint representation, by replacing generators t^a by $(T^a)_{bc} = -if^{abc}$. In the definitions above, the path C runs along the minus direction with the endpoints $(0,0,\mathbf{0}_T)$ and $(0,\xi^-,\mathbf{0}_T)$. Insertion of this Wilson line between the quark and gluon fields in the distributions guarantees their gauge invariance.

2.2 TMD factorization

In collinear factorization it is assumed that only one longitudinal component of each incoming parton's momentum is kept (*e.g.* k^+ for parton moving in the plus direction), while the others are neglected, *c.f.*

Eq. (2.1). Such approximation however is not appropriate for a certain class of observables, which are directly sensitive to the transverse component of incoming parton's momentum. The typical example is the distribution of the transverse-momentum imbalance $q_T = |p_{T1} + p_{T2}|$ between two leptons produced in the Drell-Yan process or between two hardest jets in proton-proton scattering. When the two momenta are back-to-back in the transverse plane, q_T is very small and its value can be comparable with the transverse momentum k_T of the incoming parton. Therefore, neglecting k_T in that case would lead to a significant modification of the q_T distribution in the low-q_T region. A proper treatment of such processes requires introduction of the *transverse momentum dependent* (TMD) parton distributions.

In principle, extending the definitions of the distributions in Eq. (2.5) and (2.6) to include transverse momentum should not be difficult, all that is needed is to allow for a transverse separation of the two field operators and then perform the Fourier transform on the transverse coordinates [95]

$$f(x, k_T) = \int \frac{d\xi^- d^2\xi_T}{(2\pi)^3} e^{-ixP^+\xi^- + i\mathbf{k}_T\cdot\xi_T} \langle P|\phi^\dagger(0,0,\mathbf{0}_T)\mathcal{U}^{[C]}(0,\xi^-,\xi_T)\phi(0,\xi^-,\xi_T)|P\rangle\,, \tag{2.8}$$

where ϕ represents quark fields ψ or gluon field-strength tensor F^a. However, if the fields are allowed to be at different transverse coordinates, the Wilson line in transverse direction has to be included. Therefore, in general the object $\mathcal{U}^{[C]}$ is a gauge-link, composed of multiple Wilson lines in both the light-cone and the transverse directions

$$\mathcal{U}^n_{[a;b]} = \mathcal{P}\exp\left\{-ig\int_a^b dzn\cdot A(z)\right\}, \qquad \mathcal{U}^T_{[a;b]} = \mathcal{P}\exp\left\{-ig\int_a^b dz_T\cdot A_T(z)\right\}. \tag{2.9}$$

There is no obvious choice for the path of the gauge-links connecting the space-time points, and different choices would lead to different definitions. Therefore, a prescription for the gauge-link path is needed. The Wilson lines are usually thought of as a way of resumming contributions from initial or final state interactions. In the collinear case most of these contributions cancel out, but in the TMD case it leads to process-dependence of the distributions.

The general method to determine the appropriate gauge-links for arbitrary process is known [97]. The gauge-links are related to the resummation of gluon emissions collinear to one of the participating partons. This resummation procedure is quite general and depends on the color flow. The simplest example comes from comparing single inclusive deep inelastic scattering (SIDIS) and Drell-Yan process. In calculations of the quark distributions in SIDIS it is necessary to consider only final state interactions and therefore the gauge-link structure is "future pointing", with Wilson lines extending in the plus infinite longitudinal direction. In Drell-Yan the gauge-link is "past pointing" with Wilson lines extending to minus infinity in the longitudinal direction. Using the Wilson lines in the longitudinal and transverse directions, Eq. (2.9), the future pointing gauge-link $\mathcal{U}^{[+]}$ in SIDIS and the past pointing gauge-link $\mathcal{U}^{[-]}$ in Drell-Yan are given by

$$\mathcal{U}^{[\pm]} = \mathcal{U}^n_{[(0^-,\mathbf{0}_T);(\pm\infty^-,\mathbf{0}_T)]}\mathcal{U}^T_{[(\pm\infty^-,\mathbf{0}_T);(\pm\infty^-,\infty_T)]}\mathcal{U}^T_{[(\pm\infty^-,\infty_T);(\pm\infty^-,\xi_T)]}\mathcal{U}^n_{[(\pm\infty^-,\xi_T);(\xi^-,\xi_T)]}\,. \tag{2.10}$$

and the corresponding paths are depicted in Fig. 10. The transverse gauge-links at infinity have been shown to be essential to maintain the gauge invariance, even though their contribution vanish when choosing a gauge where the fields are zero at infinity. The detailed structure of the transverse link at infinity is chosen in that particular way for technical reasons [98].

Since we have different types of TMDs for different processes, the strict factorization property is lost due to the loss of universality of the TMDs. However, the TMDs for SIDIS and Drell-Yan are related by time-reversal and differ only by the change of sign for two of the TMDs [99], so that the loss of universality does not spoil the predictive power. For more complicated processes, like dijet production in hadronic collisions, new gauge-link structures appears, like for example Wilson loops $\mathcal{U}^{[\square]}$

$$\mathcal{U}^{[\square]} = \mathcal{U}^{[+]}\mathcal{U}^{[-]\dagger} = \mathcal{U}^{[-]}\mathcal{U}^{[+]\dagger}\,, \tag{2.11}$$

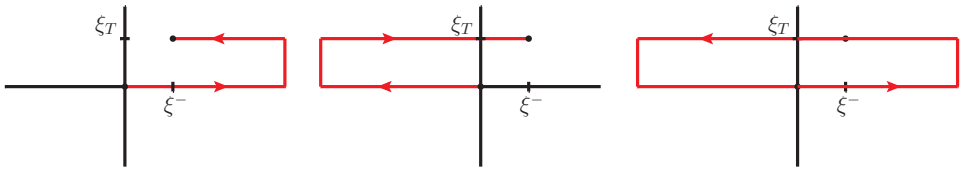


Figure 10: Path of future pointing gauge link $\mathcal{U}^{[+]}$ (left), past pointing gauge link $\mathcal{U}^{[-]}$ (middle) and closed link structure - Wilson loop $\mathcal{U}^{[\square]}$ (right).

shown in Fig. 10(right). As a result, the dijet production requires new types of TMDs, which are not reducible to those encountered in SIDIS or DY. Different TMDs appear even for different channels in dijet production, hence cross section does not factorize in the strong sense. The calculations of TMD structures for processes involving 3 and more colored partons will be the subject of Section 4 of this thesis.

It has been shown in Refs. [100,101] that the TMD factorization breaks down for hadron-hadron collisions. The problem originates from the fact that the soft gluon interactions can be resummed into gauge-links only when one hadron is considered at a time, as gluon emissions from different participants could interfere leading to lack of possibility to separate gauge-links. Factorization works however when the transverse momentum of only one of the partons incoming to hard scattering is relevant.

2.3 High Energy Factorization

For processes where the center-of-mass energy is much larger than any other scale, *i.e.* $s \gg Q$, High Energy factorization or k_T-factorization applies. HEF has been proposed in the context of heavy quark production [75, 102], which is another type of multi-scale problem. The key observation is that, at high energies, the dominant contribution to the cross section comes from exchanges of longitudinal gluons, whereas other terms are subdominant. Hence, the cross section formula can be factorized into an unintegrated gluon distribution function, undergoing BFKL evolution, which emits an off-shell (*i.e.* $k_T \neq 0$) gluon, and an off-shell matrix element. In Ref. [75] an effective procedure is derived, which guarantees gauge invariance of the off-shell amplitudes within a subclass of axial gauges.

In high energy factorization heavy quark pair is produced via the tree-level hard subprocess $g^*(k_A)g^*(k_B) \to Q\bar{Q}$ in the axial gauge. The initial-state gluons are off-shell, with momenta

$$k_A = x_A p_A + k_{T\,A}\,, \qquad\qquad k_B = x_B p_B + k_{T\,B}\,, \tag{2.12}$$

where p_A, p_B are the momenta of the incoming hadrons and $p_i \cdot k_{T\,i} = 0$. This form of the exchanged momenta is a consequence of the imposed high energy limit. The off-shell gluons have "polarization vectors" given by p_A and p_B, respectively. Due to this kinematics, the subproces given by ordinary Feynman diagrams is gauge invariant despite its off-shellness. The factorization formula for heavy quark production is written as (see. Fig. 11A) [103]

$$\begin{aligned} d\sigma_{AB\to Q\bar{Q}} = \int d^2k_{T\,A} \int \frac{dx_A}{x_A} \int d^2k_{T\,B} \int \frac{dx_B}{x_B} \\ \mathcal{F}_{g^*/A}(x_A, k_{T\,A}) \mathcal{F}_{g^*/B}(x_B, k_{T\,B})\, d\hat{\sigma}_{g^*g^*\to Q\bar{Q}}(x_A, x_B, k_{T\,A}, k_{T\,B})\,, \end{aligned} \tag{2.13}$$

where $d\hat{\sigma}_{g^*g^*\to Q\bar{Q}}$ is the partonic cross section and $\mathcal{F}_{g^*/A}$, $\mathcal{F}_{g^*/B}$ are unintegrated gluon distributions for hadrons A and B.

The approach presented above has been extended to jet production cross sections in hadron-hadron collisions (see *e.g.* [103] for overview). The problematic issue that needed to be solved comes from the processes with gluons in the final state, for which the ordinary method of calculating amplitudes does not give the gauge-invariant results. A few methods have been developed to calculate a gauge-invariant extension

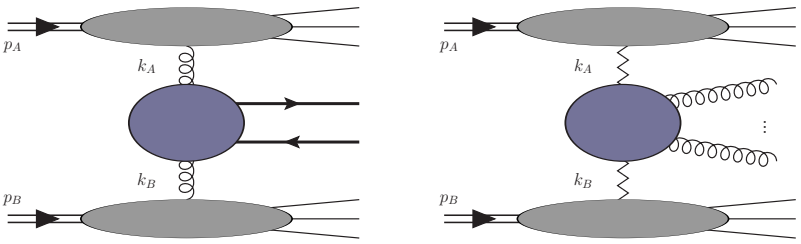


Figure 11: A) The k_T-factorization for inclusive heavy quark production; despite the fact that the gluons entering the central blob are off-shell, the subprocess is gauge invariant. B) For subprocesses with final state gluons, the gauge invariance requires the off-shell gluons to be replaced by the effective particles giving rise to multiple eikonal gluon exchanges between the blobs. From [103].

of such amplitudes [104–107], which are based on helicity techniques and allow for efficient implementation in computer programs. These gauge-invariant off-shell amplitudes correspond to a vertex that can be calculated from the Lipatov's effective action [108, 109] (see Fig. 11B).

2.4 Factorization in forward jet production

In this section we consider a subclass of dijet processes in which both jets are produced at large rapidities. This case corresponds to an asymmetric configuration, in which the two colliding objects are probed in different momentum ranges.

In the process of dijet production, schematically written as

$$A + B \mapsto a + b \to \text{jet}_1 + \text{jet}_2 + X\,, \tag{2.14}$$

the fractions of the longitudinal momenta of the incoming parton from the projectile, x_1, and from the target, x_2, can be expressed in terms of the rapidities and transverse momenta of produced jets as

$$x_1 = \frac{1}{\sqrt{s}}(p_{T1}e^{y_1} + p_{T2}e^{y_2})\,, \qquad\qquad x_2 = \frac{1}{\sqrt{s}}(p_{T1}e^{-y_1} + p_{T2}e^{-y_2})\,. \tag{2.15}$$

which in the limit $y_i \gg 0$ for forward rapidities effectively selects those fractions to be $x_1 \sim 1$ and $x_2 \ll 1$. As the number of gluons grows rapidly with decreasing momentum fraction, the forward production corresponds to dilute-dense collisions, where the projectile, probed at a high momentum fraction, hence appearing as dilute, can be described in terms of a collinear parton distribution. In contrast, the small-x gluons of the dense target nucleus are described by an unintegrated or transverse-momentum dependent distribution, which evolve toward small x according to nonlinear equations. Moreover, the momentum of the incoming gluon from the target, besides the longitudinal component, has in general a non-zero transverse component, k_T, which leads to imbalance of transverse momentum of the produced jets

$$k_T^2 = |\mathbf{p_{T1}} + \mathbf{p_{T2}}|^2 = p_{T1}^2 + p_{T2}^2 + 2p_{T1}p_{T2}\cos\Delta\phi\,. \tag{2.16}$$

The process of dijet production in the forward direction is a multi-scale problem. The highest scale is the collision energy $\sqrt{s}$, then the jets' transverse momenta p_{T1}, p_{T2} and, finally, the dijet imbalance (or equivalently, the transverse momentum of the off-shell gluon), given by Eq. (2.16), which can in principle be anywhere below the transverse momenta of the individual jets. The case with $k_T \sim p_{T1}, p_{T2}$, corresponds to a very small angle $\Delta\phi$, between the two forward jets, while if $k_T \ll p_{T1}, p_{T2}$, the two jets are almost back-to-back. The former is the domain of application of high energy factorization and the latter generalized TMD factorization.

2.4.1 HEF

In the context of the forward dijet production, the HEF formula takes a hybrid form, as we only need to consider the off-shell gluon effects in one of the colliding hadrons [110, 111]

$$\frac{d\sigma^{pA\to\text{dijets}+X}}{dy_1 dy_2 d^2p_{T1} d^2p_{T2}} = \frac{1}{16\pi^3 (x_1 x_2 s)^2} \sum_{a,c,d} x_1 f_{a/p}(x_1,\mu^2)\, |\overline{\mathcal{M}_{ag^*\to cd}}|^2 \mathcal{F}_{g/A}(x_2,k_T) \frac{1}{1+\delta_{cd}}\,. \tag{2.17}$$

This formula contains only a single unintegrated gluon distribution $\mathcal{F}_{g/A}$ (dipole distribution) which is determined from fits to DIS data. The hard factor (which, because we have only one TMD, is just a matrix element squared) depends on the transverse momentum, hence, we shall refer to it as an off-shell matrix element. The complete set of the off-shell matrix elements needed for hadroproduction of forward jets were calculated and analyzed in Refs. [110, 112]. As demonstrated in Ref. [82], the factorization formula (2.17) can be also derived for all channels from the CGC approach in the kinematic window $p_{T1}, p_{T2} \sim k_T \gg Q_s$. This limit corresponds to the dilute target approximation, hence, it should not be employed to study nonlinear effects in dense systems. It can however be used in the, so-called, geometric scaling region, where the linear approximation is still valid but saturation effects can be felt [111, 113, 114]. However, the HEF formula is not applicable in situations corresponding to $k_T \sim Q_s$. This deficiency is fixed by the improved TMD factorization framework.

2.4.2 Generalized TMD factorization

An approach valid in the regime where the transverse momentum imbalance between the outgoing particles, Eq. (2.16), is much smaller than their individual transverse momenta, is the generalized TMD factorization. Even though there is no TMD factorization theorem for jet production in hadron-hadron collisions, such a factorization can be established when only one of the colliding hadrons is described by a TMD gluon distribution. The TMD factorization formula for dijet production in dilute-dense collision reads [115]

$$\frac{d\sigma^{pA\to\text{dijets}+X}}{dy_1 dy_2 d^2p_{T1} d^2p_{T2}} = \frac{\alpha_s^2}{(x_1 x_2 s)^2} \sum_{a,c,d} x_1 f_{a/p}(x_1,\mu^2) \sum_i H^{(i)}_{ag\to cd} \mathcal{F}^{(i)}_{ag}(x_2,k_T) \frac{1}{1+\delta_{cd}}\,, \tag{2.18}$$

where several TMD gluon distributions $\mathcal{F}^{(i)}_{ag}$ with different operator definition are involved and convoluted with the hard factors $H^{(i)}_{ag\to cd}$. The hard factors were calculated in [115] as if the small-x_2 gluon was on-shell. The k_T dependence is only in the gluon distributions and there are eight of them, which in the limit of large N_c reduce to five that can be written down as convolutions of two fundamental distributions: the so-called *dipole distribution* and the *Weizäsacker-Williams distribution* (WW) (we will discuss these distributions in Section 4.4).

In Ref. [115] it was also found that the same factorization formula, Eq. (2.18) involving identical gluon distributions, can be derived from the CGC in the correlation limit (*i.e.* for nearly back-to-back dijet configurations). As a standard QCD framework is completely different than semi-classical CGC, the fact that the two lead to identical results should be regarded as a nontrivial check of the generalized TMD factorization formula.

If the k_T dependence of the hard factors inside (2.18) can be restored, a connection can be made between the HEF and TMD frameworks, providing a unified formulation which encompasses both the dilute and nearly back-to-back limit.

2.4.3 Improved TMD Factorization

A framework unifying the HEF formalism (applicable when $k_T \sim p_{T1}, p_{T2}$) and the generalized TMD formalism (applicable for $k_T \ll p_{T1}, p_{T2}$) was proposed in Ref. [82]. It can be regarded as a generalization of Ref. [115] to the case in which the k_T dependence is kept also in the hard factors. The latter were computed

with two independent methods: the original procedure of Refs. [75,102] as well as the color ordered amplitudes approach [116]. The improved factorization formula reads

$$\frac{d\sigma^{pA\to\text{dijets}+X}}{d^2P_Td^2k_Tdy_1dy_2}=\frac{\alpha_s^2}{(x_1x_2s)^2}\sum_{a,c,d}x_1f_{a/p}(x_1,\mu^2)\sum_{i=1}^{2}K^{(i)}_{ag^*\to cd}(P_T,k_T)\Phi^{(i)}_{ag\to cd}(x_2,k_T)\ \frac{1}{1+\delta_{cd}}\ , \tag{2.19}$$

where $K^{(i)}_{ag^*\to cd}$ and $\Phi^{(i)}_{ag\to cd}$ are the new hard factors and the new TMDs, replacing, respectively, $H^{(i)}_{ag\to cd}$ and $\mathcal{F}^{(i)}_{ag}$ from Eq. (2.18). As we see, $K^{(i)}_{ag^*\to cd}$ is a hard factor for an off-shell, hence k_T-dependent, incoming gluon. This formula involves two TMDs per each channel, thus, six altogether. Another improvement of Ref. [82] was a restoration of the full N_c dependence in the hard factors.

The formula (2.19) coincides with CGC expressions in two important limits. They both, in turn reduce to the TMD factorization formula when $Q_s\sim k_T\ll P_T$ and to the HEF formula when $Q_s\ll k_T\sim P_T$:

- The TMD factorization formula with k_T-dependent gluon distributions and on-shell matrix elements is obtained from (2.19) after simplifying $K^{(i)}_{ag^*\to cd}(P_T,k_T)$ into $K^{(i)}_{ag^*\to cd}(P_T,0)$
- The HEF formula with a single gluon TMD and off-shell matrix elements is obtained using the fact that up to power corrections, all the gluon TMDs coincide in the large-k_T limit:

$$\Phi^{(i)}_{ag\to cd}(x_2,k_T)\to\Phi_{g/A}(x_2,k_T)+\mathcal{O}(1/k_T^2)\,, \tag{2.20}$$

  and denoting

$$g_s^4\sum_{i=1}^{2}K^{(i)}_{ag^*\to cd}(P_T,k_T)=\overline{|\mathcal{M}_{ag^*\to cd}|}^2\,. \tag{2.21}$$

The improved TMD factorization (2.19) is valid in the limit $p_{T1},p_{T2}\gg Q_s$ for an arbitrary value of k_T, hence it provides a powerful framework for studies of the non-linear domain of QCD with hard jets. Preliminary results indicate differences with respect to the HEF formalism [117].

3 Unintegrated parton distribution functions

3.1 KMRW

Here we present the method of obtaining unintegrated parton distributions from collinear ones by the application of the KMRW procedure [118–120]. In this method, the k_T-dependent distributions are calculated from the DGLAP equation by taking into account only the contribution corresponding to a single real emission. The virtual contributions between the scales k_T and μ are resummed into a Sudakov factor, which describes the probability that there are no emissions.

The precise expressions for the unintegrated distributions read

$$\mathcal{F}_i(x, k_T^2, \mu^2) = \frac{\partial}{\partial k_T^2} \left[x f_i(x, k_T^2)\, \Delta_i(k_T^2, \mu^2) \right] , \tag{3.1}$$

with the Sudakov factors for quarks

$$\Delta_q(k_T^2, \mu^2) = \exp\left(- \int_{k_T^2}^{\mu^2} \frac{d\kappa_T^2}{\kappa_T^2} \frac{\alpha_S(\kappa_T^2)}{2\pi} \int_0^1 d\zeta\, P_{qq}(\zeta) \Theta(1 - z_M - \zeta) \right) , \tag{3.2}$$

and for gluons

$$\Delta_g(k_T^2, \mu^2) = \exp\left(- \int_{k_T^2}^{\mu^2} \frac{d\kappa_T^2}{\kappa_T^2} \frac{\alpha_S(\kappa_T^2)}{2\pi} \int_0^1 d\zeta\, [\, \zeta\, P_{gg}(\zeta) \Theta(1 - z_M - \zeta) \Theta(\zeta - z_M) + n_F P_{qg}(\zeta)\,] \right) . \tag{3.3}$$

Here n_F is the number of active quark-antiquark flavors into which the gluon may split, and, in what follows, we set $n_F = 5$. The infrared cutoff $z_M \equiv \frac{k_t}{\mu + k_t}$ arises because of the singular behavior of the splitting functions $P_{qq}(z)$ and $P_{gg}(z)$ at $z = 1$, which correspond to soft gluon emission.

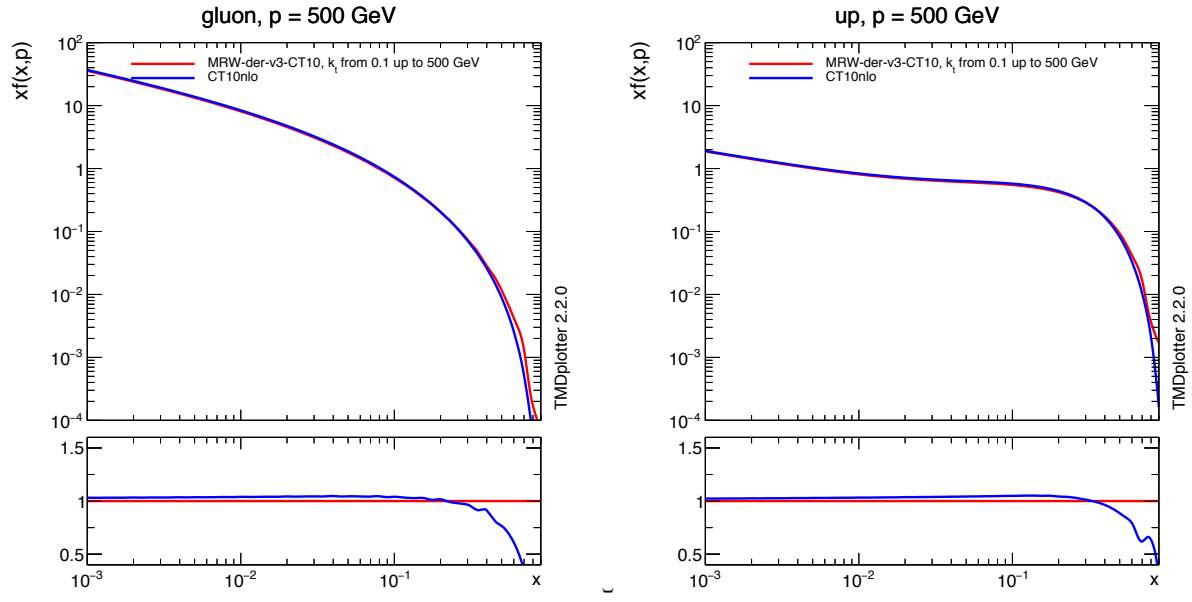


Figure 12: Comparison between the integrated TMD using the method of Ref. [120] and the underlying collinear CT10nlo gluon PDFs [28] at a scale $\mu = 500$ GeV for gluons (left) and u-quarks (right).

The unintegrated distributions are defined only for $k_T > \mu_0$, where $\mu_0 \sim 1$ GeV is the minimum scale for the the integrated (collinear) PDFs. In order to extend them to the region $k_T < \mu_0$, we tested three methods. One is to set the UPDF proportional to k_T, the second is to freeze the UPDF at $k_T = \mu_0$ and the third is

taken from Ref. [120] and is used here:

$$\mathcal{F}_i(x, k_T^2, \mu^2) = \frac{1}{\mu_0^2}\, x f_i(x, \mu_0^2)\, \Delta_i(\mu_0^2, \mu^2)\,. \tag{3.4}$$

The unintegrated distributions that will be applied to phenomenology (MRW-CT10nlo) are based on the CT10nlo collinear PDF set [28] including the appropriate running coupling α_s. In Fig. 12, we show a comparison of the original CT10 parton density with the TMDs constructed here, integrated over k_T up to the scale μ using the TMDplotter tool [121]. We observe reasonable agreement, except at large x, where the integration limits in the Sudakov form factor play a role. The large x region is, however, not relevant for the processes studied in this thesis.

In Fig. 13, we show the k_T dependence of the unintegrated distributions at the scale $\mu = 500$ GeV, for different values of x. One can clearly see the treatment of the non-perturbative region of $k_T < 1$ GeV. The discontinuity at small k_T comes from the matching procedure in Eq. (3.4). The grids of KMRW distributions as a function of x, k_T^2 and μ^2 were calculated using the C++ program MRWCALC [122].

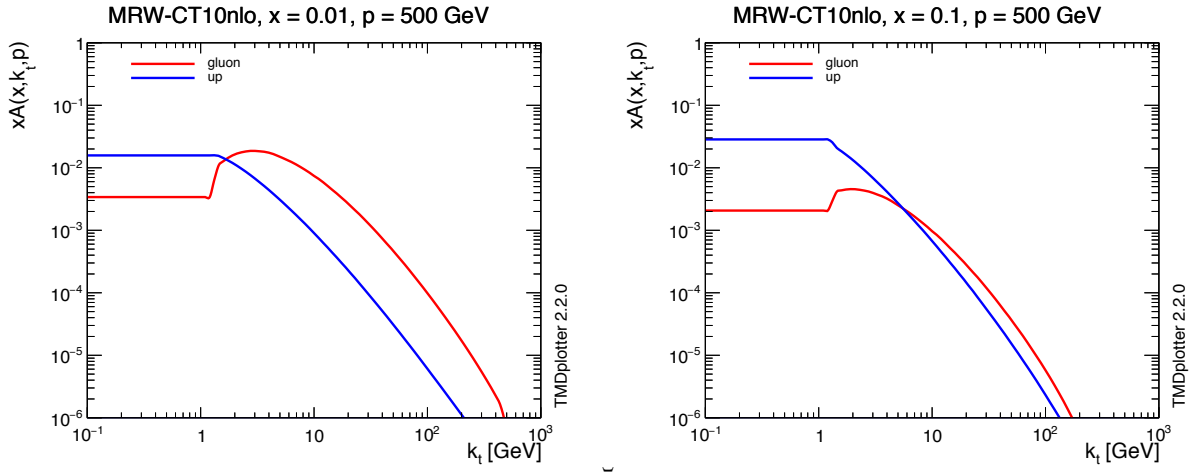


Figure 13: Transverse momentum distribution of the unintegrated PDFs at the scale $\mu = 500$ GeV for gluons and u-quarks at $x = 0.01$ (left) and $x = 0.1$ (right).

3.2 Unified BK/DGLAP

In this approach, the problem of instability of the NLO BFKL equation is resolved by the inclusion of a subset of higher order corrections, based on the formulation presented in [123]. The main correction is provided by the consistency constraint on emissions of real gluons. Other corrections come from running of the coupling and the nonsingular pieces of DGLAP splitting functions. The result is the improved BK equation for the unintegrated gluon density [124, 125]

$$\begin{aligned}\mathcal{F}(x,k^2) \;=\; & \mathcal{F}^{(0)}(x,k^2) \\ & + \frac{\alpha_s(k^2)N_c}{\pi}\int_x^1 \frac{dz}{z}\int_{k_0^2}^{\infty}\frac{dl^2}{l^2}\left\{\frac{l^2\mathcal{F}(\frac{x}{z},l^2)\,\theta(\frac{k^2}{z}-l^2)\;-\;k^2\mathcal{F}(\frac{x}{z},k^2)}{|l^2-k^2|}+\frac{k^2\mathcal{F}(\frac{x}{z},k^2)}{|4l^4+k^4|^{\frac{1}{2}}}\right\} \\ & + \frac{\alpha_s(k^2)}{2\pi k^2}\int_x^1 dz\left[\left(P_{gg}(z)-\frac{2N_c}{z}\right)\int_{k_0^2}^{k^2} dl^2\,\mathcal{F}\left(\frac{x}{z},l^2\right)+zP_{gq}(z)\Sigma\left(\frac{x}{z},k^2\right)\right] \\ & - \frac{2\alpha_s^2(k^2)}{R^2}\left[\left(\int_{k^2}^{\infty}\frac{dl^2}{l^2}\mathcal{F}(x,l^2)\right)^2+\mathcal{F}(x,k^2)\int_{k^2}^{\infty}\frac{dl^2}{l^2}\ln\left(\frac{l^2}{k^2}\right)\mathcal{F}(x,l^2)\right], \quad (3.5)\end{aligned}$$

where $z = x/x'$ and k and l are the transverse momenta of gluons in the side rails of the ladder. The kinematic constraint is introduced through the theta function. The DGLAP effects generated by the nonsingular part of the splitting function $P_{gg}(z)$ in the $z \to 0$ limit and contributions from the sea quarks are provided by the first and the second term in the third line of the equation, respectively, where $\Sigma(x,k^2)$ denotes the singlet quark distribution (sum of quark and antiquark distributions of all flavors). The nonlinear term (the last line of the equation), which supplies unitarity corrections, comes from the triple pomeron vertex [126] (it allows for the recombination of gluons). The strength of the nonlinear term is controlled by the parameter R, which stems from integration of the gluon density over the impact parameter b and has an interpretation of the proton radius. The input gluon distribution $\mathcal{F}_p^{(0)}(x,k^2)$ is given by

$$\mathcal{F}^{(0)}(x,k^2) = \frac{\alpha_s(k^2)}{2\pi k^2}\int_x^1 dz P_{gg}(z)\frac{x}{z}g\left(\frac{x}{z},k_0^2\right), \quad (3.6)$$

where $xg(x,k_0^2)$ is the integrated gluon distribution at the initial scale $k_0^2 = 1\,\mathrm{GeV}^2$, parametrized as

$$xg(x,1\mathrm{GeV}^2) = N(1-x)^{\beta}(1-Dx). \quad (3.7)$$

The parameters N, β, D, together with the proton radius R, were constrained with a fit to HERA data [127] in [111]. The obtained results gave a very good description of data, which corresponds to $\chi^2/\mathrm{ndof} = 1.73$ and the following values of the parameters: $N = 0.994$, $\beta = 18.6$, $D = -82.1$ and $R = 2.40\,\mathrm{GeV}^{-1}$. The gluon density resulting from this procedure will be referred to as *KS nonlinear*.

The nonlinear evolution equation (3.5) provides a framework that accounts for saturation of gluon density. The framework without saturation can be constructed in a straightforward way from a linearized version of this equation, obtained simply by dropping the last term. The linearized equation then becomes independent of R. An analogous fit to the HERA data was performed in [111], resulting in the following fit parameters at the minimal value of $\chi^2 = 1.51$: $N = 0.004$, $\beta = 26.7$ and $D = -51102$. This version predicts too strong rise of F_2 with x,especially at low values of Q^2, and it was concluded that some mechanism of damping the gluon density at low x and Q^2 is needed to describe HERA data in the full range of Q^2. The gluon density obtained from the linearized version of Eq. (3.5) will be referred to as *KS linear*.

The evolution equation (3.5) and its linearized version were used to determine the unintegrated gluon above the initial momentum scale, that is for $k^2 > 1\,\mathrm{GeV}^2$. Below this scale, the gluon density $\mathcal{F}(x,k^2)$ is constrained by the condition that it should match the evolved unintegrated gluon density at $k^2 = k_0^2$. Each version has different parametrization in that region, namely

$$\begin{aligned}&\text{KS nonlinear}: \quad \mathcal{F}(x,k^2) = k^2\mathcal{F}(x,k_0^2 = 1\,\mathrm{GeV}^2), \\ &\text{KS linear}: \quad \mathcal{F}(x,k^2) = \mathcal{F}(x,k_0^2 = 1\,\mathrm{GeV}^2).\end{aligned} \quad (3.8)$$

The evolution of the unintegrated gluon density in the case of a nucleus, $\mathcal{F}_A$, can be obtained through

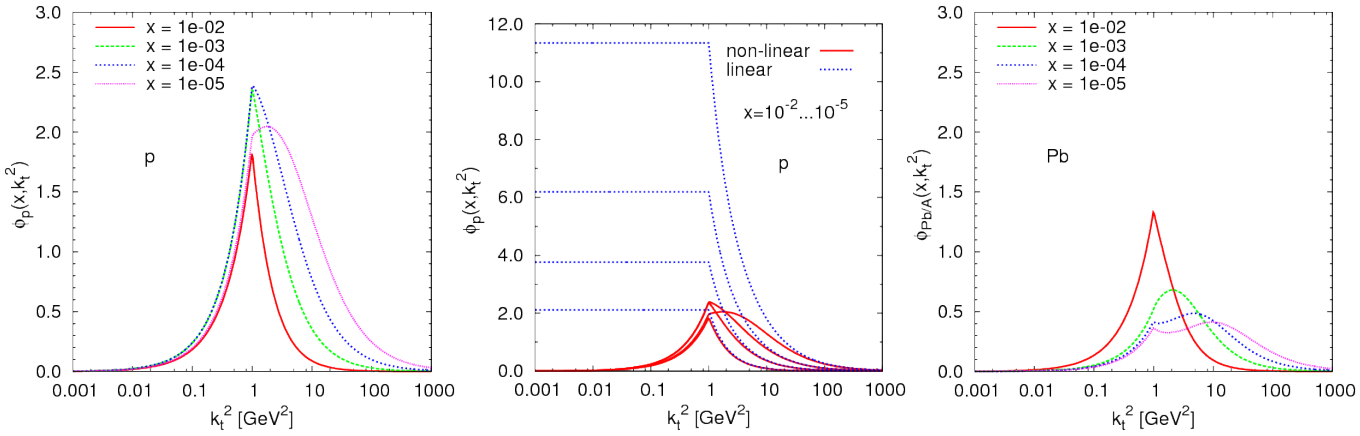


Figure 14: Left: KS nonlinear unintegrated gluon density in the proton. Middle: Comparison between KS nonlinear and KS linear gluon density. Right: KS nonlinear gluon density in the Pb nucleus ($A = 207$). From [111].

the following formal substitution in Eq. (3.5)

$$\frac{1}{R^2} \to \frac{A}{R_A^2}\,, \qquad\qquad R_A^2 = R^2 A^{2/3}\,,$$

where R_A is the nuclear radius and A is the mass number ($A = 208$ for Pb). The resulting nuclear gluon density is normalized to the number of nucleons in the nuclei, with the strength of the nonlinear term enhanced by $A^{1/3}$. Since the A-dependence is already included in the evolution (enters the nonlinear term through the impact parameter), the initial condition $\mathcal{F}^{(0)}(x,k^2)$ is the same for proton and nucleus.

The KS nonlinear unintegrated gluon density as a function of gluon transverse momentum, for several values of x, is shown in Fig. 14 (left). The sharp peak corresponds to the matching point k_0^2. One can see that with lower values of x, perturbatively generated maximum starts to emerge, which signals the presence of saturation. Fig. 14 (middle) compares KS nonlinear and KS linear. At large values of x and k_T^2 both distributions are similar, however the linear gluon grows much faster than the nonlinear one in low-k_T^2 region, especially for smaller values of x. Fig. 14 (right) shows the KS nonlinear gluon density in the Pb nucleus. It can be noticed that, due to stronger saturation effects in Pb, the gluon density is lower than that in the proton, and the maxima are shifted towards larger values of k_T^2, which corresponds to the larger saturation scale.

3.3 Unified BK/DGLAP with hard scale dependence

Another type of effect, that is beyond the BK equation, is color coherence, *i.e.* angular ordering of successive parton branchings. The inclusion of such effects leads to dependence of the gluon density on the scale of the hard process. A relatively straightforward framework to provide the hard scale dependence is based on the KMRW method described in the previous subsection, where the Sudakov effects [128] are factorized into form factors. In [129] this method was used to introduce the hard scale dependence to the KS linear and nonlinear unintegrated gluon densities. Here, we present a short outline of such construction.

The main assumption is that the hard-scale-dependent gluon density, $\mathcal{F}(x,k^2,\mu^2)$, after integration is equal to $\mathcal{F}(x,k^2)$. This ensures that the Sudakov form factor only modifies the shape of the gluon density, leaving the distribution unchanged at the inclusive level. The contribution when $k^2 > \mu^2$ is simply given by the unintegrated distribution. The above can be summarized by the formula [129]

$$\mathcal{F}(x,k^2,\mu^2) := \theta(\mu^2-k^2)\Delta(\mu^2,k^2)\frac{xg(x,\mu^2)}{xg_{hs}(x,\mu^2)}\mathcal{F}(x,k^2) + \theta(k^2-\mu^2)\mathcal{F}(x,k^2)\,, \tag{3.9}$$

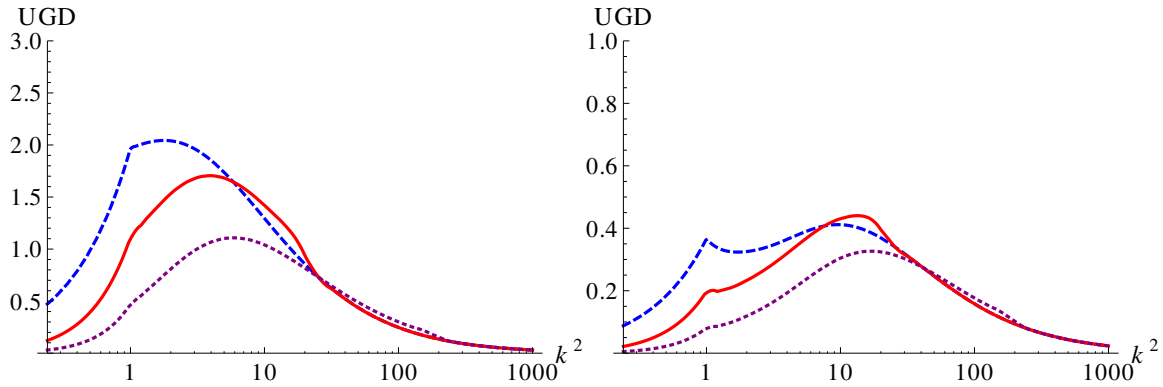


Figure 15: KS hardscale nonlinear unintegrated gluon density (UGD), evaluated at $x = 10^{-5}$ and several values of μ^2, for proton (left) and Pb (right). Hard scale $\mu^2 = 20\,\mathrm{GeV}^2$ (red continuous line), $\mu^2 = 200\,\mathrm{GeV}^2$ (purple dotted line), KS nonlinear at $x = 10^{-5}$ (blue dashed line). From [129].

where

$$xg_{hs}(x,\mu^2) = \int^{\mu^2} dk^2 \Delta(\mu^2, k^2)\mathcal{F}(x,k^2)\,, \qquad xg(x,\mu^2) = \int^{\mu^2} dk^2 \mathcal{F}(x,k^2)\,, \tag{3.10}$$

and the Sudakov form factor is given by Eq. (3.3), as in the KMRW method. After integration up to the hard scale in Eq. (3.9) and application of Eq. (3.10), the xg_{hs} terms cancel and the part proportional to $\theta(k^2 - \mu^2)$ drops. Hence, at the integrated level, the number of gluons does not change, the Sudakov form factor just makes the shape of the gluon density scale dependent.

The KS linear and KS nonlinear unintegrated gluon densities (described in the previous subsection) supplemented with hard scale dependence will be referred to as *KS hardscale linear* and *KS hardscale nonlinear*, respectively. Fig. 15 shows KS hardscale nonlinear gluon in proton and Pb for several values of μ^2, compared with KS nonlinear. As already discussed, $k_0^2 = 1\,\mathrm{GeV}^2$ is the initial scale, below which model extension is applied. The maximum of the distribution signals the emergence of the saturation scale. One can see that the hard-scale-dependent gluon density dominates regular gluon density in the regions where the hard scale is approaching k. The ratio of the unintegrated gluon density of lead to the unintegrated gluon density of proton, shown in Fig. 16, indicates that the gluon density of proton is more affected by Sudakov effects than Pb, since the ratio is smaller than one in a wider range of k. This is because the saturation effects in lead are larger and the suppression of the low k region is more significant already in hard scale independent gluon density.

The inclusion of the Sudakov effects is necessary in order to describe the LHC jet data at small x, but before reaching the saturation regime, as was shown in [114, 130, 131]. An alternative approach to introduce such effects incorporates them directly as a part of the evolution equation, *i.e.* at all steps in the evolution, and leads to the CCFM evolution, discussed in Section 1.4.3, and the nonlinear equation developed in [54,132,133]. However, due to its numerical complexity, the nonlinear equation has not yet been applied to phenomenology.

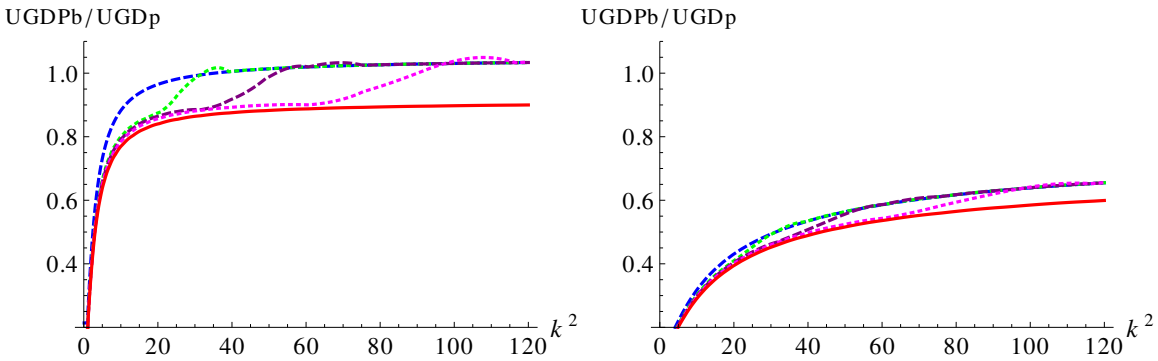


Figure 16: The ratio of the KS hardscale nonlinear unintegrated gluon density of lead (UGDPb) to unintegrated gluon density of proton (UGDp), evaluated at $x = 10^{-3}$ (left) and $x = 10^{-5}$ (right) for several values of μ^2. Hard scale $\mu^2 = 25\,\mathrm{GeV}^2$ (green dotted line), $\mu^2 = 45\,\mathrm{GeV}^2$ (purple dashed line), $\mu^2 = 80\,\mathrm{GeV}^2$ (magenta dotted line), $\mu^2 = 400\,\mathrm{GeV}^2$ (red continuous line), KS nonlinear (blue dashed line). From [129].

4 TMD gluon distributions for multiparton processes

The TMD factorization, as discussed in Section 2, in strict sense is applicable only to a few processes. Nevertheless, the basic objects appearing in the formalism, the TMD distributions, can be studied in the broader context. They are defined as the Fourier transforms of the hadronic matrix elements of bilocal field operators with non-light-like separation. To ensure the gauge invariance, the Wilson links connecting the two space-time points must be inserted. For the gauge invariance itself the shape of the links is not relevant. In the TMD factorization however, the shape of the links is determined by the hard process accompanying the TMD parton distribution. This happens because the collinear gluons (to the incoming hadron), which couple to various components of the hard process, have to be considered as a part of the nonperturbative wave function. They can be resummed into the Wilson links attached to each external leg by means of the Ward identity. Since the external legs are connected by certain color matrix, so are the pieces of Wilson links and this is how the process dependence enters [97, 134]. For simple processes like Drell–Yan pairs production, the color flow in the hard process is rather simple because of only two colored partons. Consequently, the resulting TMD parton distribution has also simple structure. On the contrary, for processes with several colored partons one gets multiple nonequivalent structures (including Wilson loops), which cannot be eliminated by a gauge choice.

Even though the strict all-order factorization theorems fail for processes with more than two colored partons participating in the hard collision [101], in the nonlinear small-x regime the lowest order TMDs are of great phenomenological importance. In Ref. [115] a leading power limit of the expressions for dijet production in pA collisions within the CGC was studied. They found that the correlators of Wilson lines averaged over color sources according to the CGC theory correspond exactly to the TMD gluon distributions for $2 \to 2$ processes, provided the hadronic matrix elements are traded for the color source averages. Not only the correlators agree, but also the hard factors. Although it is not yet known whether this correspondence survives beyond the leading order, it opened new phenomenological opportunities to study with better theoretical control semi-hard jets in the gluon saturation domain [82, 135–138]. In particular, in Ref. [82] a beyond-leading-power extension of the TMD factorization for forward dijets in pA collisions was proposed, such that it coincides with the leading power of CGC in the dense nucleus regime, and with the all-power high energy factorization in the dilute nucleus limit (see subsection 2.4). The formal TMD factorization breaks because one is unable to define the separate correlators whilst more than two colored partons are present [101]. In the small-x approach for dilute-dense collisions, however, there is only one correlator with transverse separation. Therefore the complications that lead to the lack of possibility to separate Wilson links into TMD operators, formally do not appear here. Outside the small-x limit for dilute-dense collisions these results might also be useful: for example to access the factorization breaking effects.

Motivated by the phenomenological usability of the non-universal TMD gluon distributions, we will present explicit results for the operator structures for all 3 - 6 colored parton processes. We first describe the rules derived in [97] for calculation of a TMD operator structure in an arbitrary process and show some examples of how to apply them to calculate structures corresponding to several Feynman diagrams. Afterwards, we implement this prescription to color-ordered amplitudes and present the results of TMD structures appearing in multiparton processes.

4.1 Gauge-links in arbitrary processes

We start off by providing necessary definitions of the TMD distribution correlators. The same rules apply also for calculations of the TMD fragmentation correlators, but we will not consider them here. The TMD quark and gluon correlators are defined by the following matrix elements [139–141] [139–141]

$$\Phi_q(x, k_T) = \int \frac{d\xi^- \, d^2\xi_T}{(2\pi)^3} \, e^{ik\cdot\xi} \langle H|\psi_i^\dagger(0)\psi_j(\xi)|H\rangle \,, \tag{4.1}$$

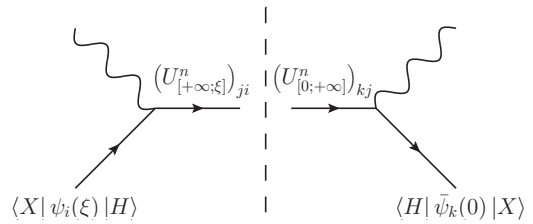


Figure 17: Gauge-link in quark distribution correlator in SIDIS.

$$\Phi_g(x,k_T) = \int \frac{d\xi^- d^2\xi_T}{(2\pi)^3}\, e^{ik\cdot\xi} \langle H|F_a^{\dagger\mu\nu}(0)F_b^{\rho\delta}(\xi)|H\rangle\,, \tag{4.2}$$

where $|H\rangle$ is a hadron state, ψ_i is color triplet quark field and $\hat{F}^{\mu\nu}(x) = F_a^{\mu\nu}(x)\,t^a$ is the field strength tensor (we use $\mathrm{Tr}\left(t^a t^b\right) = T_F\delta^{ab}$, $T_F = 1/2$ convention for the generators). Such bilocal products of field operators require insertion of Wilson lines to ensure gauge invariance. A generic Wilson line was defined in Eq. (2.7). The requirement of gauge invariance alone does not uniquely fix the integration path. The Wilson lines are obtained by resumming all diagrams describing the exchange of collinear gluons between the soft and hard parts [142], so the integration path is fixed by the hard part of the process.

As explained in Ref. [97], in order to obtain gauge-link in a particular correlator, all the gluon couplings to the hard must be summed, and at leading twist it involves gluons which are collinear to the hadron's momentum. Such resummation leads to the attachment of Wilson lines to each external leg of the hard part, except to the one which connects the hard part and the correlator under consideration. The type of the external parton determines what kind of Wilson line is attached (this is summarized in Table 1). In the final step all the Wilson lines are pulled through the color fields of the hard parts to the correlator, where they combine into the required gauge-link. The process-dependence of the gauge-link stems from this last step. In the final expression the transverse pieces of the gauge-link must be included.

In summary, the procedure to derive the full gauge-link for any process consists of the following steps:

1. Consider the diagram for an elementary squared amplitude and replace the free spinors of external partons by matrix elements given in the column 'fields in correlator' in Table 1.

2. Replace the color wave functions of the external partons (except the one which is connected to the correlator) by Wilson lines given in the column 'contribution to other gauge-links' in Table 1.

3. Pull the Wilson lines through the color structure of the hard part using the Fierz identity

$$t^a_{ij}t^a_{kl} = T_F\left(\delta_{il}\delta_{jk} - \frac{1}{N_c}\delta_{ij}\delta_{kl}\right)\,. \tag{4.3}$$

4. The resulting expression contains the desired gauge-link and the structure corresponding to the bare diagram. Divide this expression by the color structure of a bare diagram where the legs flowing to the correlator have open indices.

Let us now perform some calculations to illustrate the procedure. We will need to transform Wilson lines in the adjoint representation to the Wilson lines in the fundamental representation, which is done using the following relation

$$\left(\mathcal{U}^{[C]}(\eta;\xi)\right)_{ab} = \frac{1}{T_F}\,\mathrm{Tr}\left[t^a\mathcal{U}^{[C]}(\eta;\xi)t^b\mathcal{U}^{[C]\dagger}(\eta;\xi)\right]\,. \tag{4.4}$$

Distribution in SIDIS

The Feynman diagram for the semi-inclusive deep inelastic scattering is depicted in Fig. 17. In the first step we replace spinors of the incoming quark by the matrix elements $\langle X|\psi_i(\xi)|H\rangle$ and $\langle H|\bar{\psi}_k(0)|X\rangle$ in the

	'free' wave functions	fields in correlator	contribution to other gauge-links
incoming quark	$u_i(p)e^{ip\cdot\xi}$	$\langle X\|\delta_{ij}\psi_j(\xi)\|H\rangle e^{ip\cdot\xi}$	$\left(U^n_{[\xi;-\infty]}\right)_{ij}$
incoming antiquark	$\bar{v}_i(p)e^{ip\cdot\xi}$	$\langle X\|\bar{\psi}_j(\xi)\delta_{ji}\|H\rangle e^{ip\cdot\xi}$	$\left(U^n_{[-\infty;\xi]}\right)_{ji}$
incoming gluon	$\epsilon_a(p)e^{ip\cdot\xi}$	$\langle X\|\delta_{ab}F^{\mu\nu}_b(\xi)\|H\rangle e^{ip\cdot\xi}$	$\left(U^n_{[\xi;-\infty]}\right)_{ab}$
outgoing quark	$\bar{u}_i(k)e^{-ik\cdot\xi}$	$\langle hX\|\bar{\psi}_j(\xi)\delta_{ji}\|0\rangle e^{-ik\cdot\xi}$	$\left(U^n_{[+\infty;\xi]}\right)_{ji}$
outgoing antiquark	$v_i(k)e^{-ik\cdot\xi}$	$\langle hX\|\delta_{ij}\psi_j(\xi)\|0\rangle e^{-ik\cdot\xi}$	$\left(U^n_{[\xi;+\infty]}\right)_{ij}$
outgoing gluon	$\epsilon^*_a(k)e^{-ik\cdot\xi}$	$\langle hX\|F^{\mu\nu}_b(\xi)\delta_{ba}\|0\rangle e^{-ik\cdot\xi}$	$\left(U^n_{[+\infty;\xi]}\right)_{ba}$

	'free' wave functions	fields in correlator	contribution to other gauge-links
incoming quark	$\bar{u}_i(p)$	$\langle H\|\bar{\psi}_j(0)\delta_{ji}\|X\rangle$	$\left(U^n_{[-\infty;0]}\right)_{ji}$
incoming antiquark	$v_i(p)$	$\langle H\|\delta_{ij}\psi_j(0)\|X\rangle$	$\left(U^n_{[0;-\infty]}\right)_{ij}$
incoming gluon	$\epsilon_a(p)$	$\langle H\|F^{\mu\nu}_b(0)\delta_{ba}\|X\rangle$	$\left(U^n_{[-\infty;0]}\right)_{ba}$
outgoing quark	$u_i(k)$	$\langle 0\|\delta_{ij}\psi_j(0)\|hX\rangle$	$\left(U^n_{[0;+\infty]}\right)_{ij}$
outgoing antiquark	$\bar{v}_i(k)$	$\langle 0\|\bar{\psi}_j(0)\delta_{ji}\|hX\rangle$	$\left(U^n_{[+\infty;0]}\right)_{ji}$
outgoing gluon	$\epsilon^*_a(k)$	$\langle 0\|\delta_{ab}F^{\mu\nu}_b(0)\|hX\rangle$	$\left(U^n_{[0;+\infty]}\right)_{ab}$

Table 1: Fields and gauge-links entering hadron correlators. The standard free wave functions of the partons in the hard scattering amplitudes are given in the first column. The second column shows how the partons appear in the correlators. The third column gives the contributions of the external partons to the various gauge-links. The $U^n_{[a;b]}$ are the Wilson lines along the light-cone direction n. The upper table is for the diagram corresponding to the hard amplitude, and the lower table for the diagram corresponding to the conjugate amplitude. From [97].

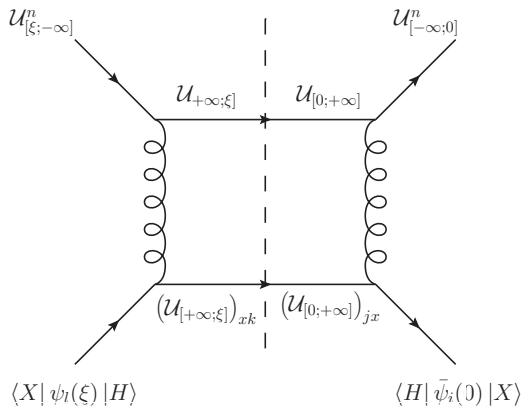


Figure 18: Diagram contributing to the quark-quark scattering.

amplitude and the conjugated amplitude, respectively. The spinors of the outgoing quark are replaced by the Wilson lines $(U^n_{[+\infty;\xi]})_{ji}$ and $(U^n_{[0;+\infty]})_{kj}$ (again, in the amplitude and the conjugated amplitude, respectively), as prescribed in Table 1. Therefore the expression for the quark correlator is

$$\Phi_q = \sum_X \int \frac{d\xi^- d^2\xi_T}{(2\pi)^3}\, e^{ik\cdot\xi} \langle H|\bar{\psi}_k(0)|X\rangle (U^n_{[0;+\infty]})_{kj} (U^n_{[+\infty;\xi]})_{ji} \langle X|\psi_i(\xi)|H\rangle\,, \tag{4.5}$$

and the Wilson lines give the gauge-link $\mathcal{U}^{[+]}$

$$(U^n_{[0;+\infty]})_{kj} (U^n_{[+\infty;\xi]})_{ji} = \left(\mathcal{U}^{[+]}\right)_{ki}\,. \tag{4.6}$$

After using the resolution of unity $\sum_X |X\rangle\langle X|$, the quark TMD correlator in SIDIS is

$$\Phi^{[+]}_q(x,k_T) = \int \frac{d\xi^- d^2\xi_T}{(2\pi)^3}\, e^{ik\cdot\xi} \langle H|\bar{\psi}(0)\mathcal{U}^{[+]}\psi(\xi)|H\rangle\,, \tag{4.7}$$

so we reproduced the well-known quark correlator with a future pointing Wilson line. In this process the hard part does not depend on color, so step 3 and 4 of the procedure do not apply.

Distribution in quark-quark scattering

Here we will calculate the quark distribution correlator for one of the quark-quark scattering channels, shown in Fig. 18. The expression for the quark correlator obtained by applying the steps 1 and 2 is

$$\langle H|\bar{\psi}_i(0)|X\rangle t^a_{ij} (\mathcal{U}^n_{[0;+\infty]})_{jx} (\mathcal{U}^n_{[+\infty;\xi]})_{xk} t^b_{kl} \langle X|\psi_l(\xi)|H\rangle \times (\mathcal{U}^n_{[-\infty;0]})_{mn} t^a_{no} (\mathcal{U}^n_{[0;+\infty]})_{oy} (\mathcal{U}^n_{[+\infty;\xi]})_{yp} t^b_{pq} (\mathcal{U}^n_{[\xi;-\infty]})_{qm} \tag{4.8}$$

$$\begin{aligned} &= \langle H|\bar{\psi}_i(0)|X\rangle t^a_{ij} \mathcal{U}^{[+]}_{jk} t^b_{kl} \langle X|\psi_l(\xi)|H\rangle \times (\mathcal{U}^n_{[-\infty;0]})_{mn} t^a_{no} \mathcal{U}^{[+]}_{op} t^b_{pq} (\mathcal{U}^n_{[\xi;-\infty]})_{qm} = \\ &= \langle H|\bar{\psi}_i(0)|X\rangle t^a_{ij} \mathcal{U}^{[+]}_{jk} t^b_{kl} \langle X|\psi_l(\xi)|H\rangle \times \mathrm{Tr}[\mathcal{U}^n_{[-\infty;0]} t^a \mathcal{U}^{[+]} t^b \mathcal{U}^n_{[\xi;-\infty]}] = \\ &= \langle H|\bar{\psi}_i(0)|X\rangle t^a_{ij} \mathcal{U}^{[+]}_{jk} t^b_{kl} \langle X|\psi_l(\xi)|H\rangle \times \mathcal{U}^{[-]\dagger}_{mn} t^a_{no} \mathcal{U}^{[+]}_{op} t^b_{pm}\,. \end{aligned} \tag{4.9}$$

Now we need to use the Fierz identity (4.3) to rewrite products of color matrices and gauge links

$$\begin{aligned}
&\mathcal{U}^{[+]}_{jk}\mathcal{U}^{[-]\dagger}_{mn}\mathcal{U}^{[+]}_{op}(t^a_{ij}t^a_{no})(t^b_{kl}t^b_{pm}) = \\
&= \mathcal{U}^{[+]}_{jk}\mathcal{U}^{[-]\dagger}_{mn}\mathcal{U}^{[+]}_{op}(T_F\delta_{io}\delta_{jn} - \frac{T_F}{N_c}\delta_{ij}\delta_{no})(T_F\delta_{km}\delta_{lp} - \frac{T_F}{N_c}\delta_{kl}\delta_{pm}) = \\
&= T_F^2\mathcal{U}^{[+]}_{jk}\mathcal{U}^{[-]\dagger}_{mn}\mathcal{U}^{[+]}_{op}(\delta_{io}\delta_{jn}\delta_{km}\delta_{lp} - \frac{1}{N_c}\delta_{io}\delta_{jn}\delta_{kl}\delta_{pm} - \frac{1}{N_c}\delta_{ij}\delta_{no}\delta_{km}\delta_{lp} + \frac{1}{N_c^2}\delta_{ij}\delta_{no}\delta_{kl}\delta_{pm}) = \\
&= T_F^2(\mathcal{U}^{[+]}_{jk}\mathcal{U}^{[-]\dagger}_{kj}\mathcal{U}^{[+]}_{il} - \frac{1}{N_c}\mathcal{U}^{[+]}_{nl}\mathcal{U}^{[-]\dagger}_{mn}\mathcal{U}^{[+]}_{im} - \frac{1}{N_c}\mathcal{U}^{[+]}_{im}\mathcal{U}^{[-]\dagger}_{mo}\mathcal{U}^{[+]}_{ol} + \frac{1}{N_c^2}\mathcal{U}^{[+]}_{il}\mathcal{U}^{[-]\dagger}_{mo}\mathcal{U}^{[+]}_{om})
\end{aligned} \tag{4.10}$$

Putting this back to the correlator we have

$$\begin{aligned}
&T_F^2\langle H|\bar\psi_i(0)|X\rangle\times \\
&\left(\mathcal{U}^{[+]}_{il}\,\mathrm{Tr}\left[\mathcal{U}^{[+]}\mathcal{U}^{[-]\dagger}\right] - \frac{1}{N_c}\mathcal{U}^{[+]}_{im}\mathcal{U}^{[-]\dagger}_{mn}\mathcal{U}^{[+]}_{nl} - \frac{1}{N_c}\mathcal{U}^{[+]}_{im}\mathcal{U}^{[-]\dagger}_{mo}\mathcal{U}^{[+]}_{ol} + \frac{1}{N_c^2}\mathcal{U}^{[+]}_{il}\,\mathrm{Tr}\left[\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}\right]\right) \\
&\qquad\qquad \times\langle X|\psi_l(\xi)|H\rangle = \qquad (4.11)
\end{aligned}$$

$$= T_F^2\langle H|\bar\psi(0)\left\{\frac{N_c^2+1}{N_c^2}\,\mathrm{Tr}\left(\mathcal{U}^{[\square]}\right)\mathcal{U}^{[+]} - \frac{2}{N_c}\mathcal{U}^{[\square]}\mathcal{U}^{[+]}\right\}\psi(\xi)|H\rangle. \tag{4.12}$$

The color factor for the bare diagram with open indices i and l is

$$t^a_{ij}t^b_{jl}t^a_{mn}t^b_{nm} = T_F^2\left(\delta_{in}\delta_{jm} - \frac{1}{N_c}\delta_{ij}\delta_{mn}\right)\left(\delta_{jm}\delta_{ln} - \frac{1}{N_c}\delta_{jl}\delta_{nm}\right) = \tag{4.13}$$

$$= T_F^2\left(N_c\delta_{il} - \frac{1}{N_c}\delta_{il} - \frac{1}{N_c}\delta_{il} + \frac{1}{N_c^2}N_c\delta_{il}\right) = T_F^2\frac{N_c^2-1}{N_c}\delta_{il}. \tag{4.14}$$

After dividing by this color factor and summing over X states we get the final answer

$$\Phi_q^{[\mathcal{U}]}(x,k_T) = \int\frac{d\xi^- d^2\xi_T}{(2\pi)^3}e^{ik\cdot\xi}\langle H|\bar\psi(0)\left\{\frac{N_c^2+1}{N_c^2-1}\frac{\mathrm{Tr}\left(\mathcal{U}^{[\square]}\right)}{N_c}\mathcal{U}^{[+]} - \frac{2}{N_c^2-1}\mathcal{U}^{[\square]}\mathcal{U}^{[+]}\right\}\psi(\xi)|H\rangle\,. \tag{4.15}$$

A few more examples, for gluon distribution correlators, are presented in Appendix A. For processes with larger number of partons the calculations become more involved and also the number of contributing Feynman diagrams grows rapidly. Therefore, we resort to the methods applied in calculations of amplitudes, *i.e.* the color decompositions.

In the rest of this section, we will be concerned with the *gluon* TMD distributions of the generic form

$$\mathcal{F}(x,k_T) = 2\int\frac{d\xi^- d^2\xi_T}{(2\pi)^3\,P^+}\,e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T}\;\langle P|\,\mathrm{Tr}\left\{\hat F^{i+}(0)\,\mathcal{U}_{C_1}\hat F^{i+}\left(\xi^+=0,\xi^-,\vec\xi_T\right)\mathcal{U}_{C_2}\right\}|P\rangle\;, \tag{4.16}$$

where $\mathcal{U}_{C_1}, \mathcal{U}_{C_2}$ are certain fundamental representation Wilson lines, multiplied by possible traces of Wilson loops. The exact shape of Wilson lines will depend on the hard process coupled to the TMD. Their calculation for multiple partons is the main goal of the present section.

4.2 Color decomposition

The calculation of the operator structure entering the TMD distributions is nicely systematized not by considering a particular diagrams, but rather by considering various color flows in the amplitude (squared)

under consideration. Such systematization is achieved by using gauge invariant decomposition of amplitudes into so-called color-ordered amplitudes (called also partial, or dual amplitudes). Here we are presenting only necessary definitions and properties, see *e.g.* [116] for a complete review.

We will consider n-parton processes with a gluon in the initial state

$$g\left(k_1\right)+b_n\left(k_n\right)\to b_2\left(k_2\right)+\cdots+b_{n-1}\left(k_{n-1}\right)\,, \tag{4.17}$$

where the partons b_i can be quarks or gluons (restricted by the flavor number conservation). The initial state gluon with momentum k_1 carries a fraction x of the parent hadron momentum $P^\mu=P^+\tilde{n}^\mu$ (in light-cone basis)

$$k_1^\mu\simeq xP^\mu+k_T^\mu\,. \tag{4.18}$$

Above, the minus component is suppressed as it is neglected in the hard part. The transverse component is also neglected within the leading twist collinear and TMD factorization. In the more general case, the gluon may be off-shell and a suitable redefinition of the hard process is required to maintain the gauge invariance (see *e.g.* [107, 143–146]). Even then, at least formally, the principles to obtain the TMD distributions still hold, therefore we shall not distinguish these situations here.

Let us start with pure gluonic tree-level amplitudes. For the sake of this section we assume that all the momenta are outgoing (later, it will become necessary to distinguish incoming and outgoing legs). The most standard decomposition reads

$$\mathcal{M}^{a_1\dots a_n}\left(k_1,\dots,k_n\right)=\sum_{\pi\in S_n/Z_n}\mathrm{Tr}\left(t^{a_{\pi(1)}}\dots t^{a_{\pi(n)}}\right)\,\mathcal{A}\left(\pi\left(1\right),\dots,\pi\left(n\right)\right)\,, \tag{4.19}$$

where the sum runs over all noncyclic permutations $\pi\in S_n/Z_n$ of an n-element set (S_n is the set of all permutations of n gluons, while Z_n is the subset of cyclic permutations). Three important properties of the above decomposition are: i) the partial amplitudes $\mathcal{A}$ are gauge invariant, ii) the partial amplitudes contain only planar diagrams; consequently the full amplitude squared satisfies $\left|\mathcal{M}\right|^2=C\sum_{S_{n-1}}\left|\mathcal{A}\left(1,\pi\left(2\right),\dots,\pi\left(n\right)\right)\right|^2+\mathcal{O}\left(1/N_c^2\right)$, with C being a color factor, iii) the amplitudes $\mathcal{A}$ satisfy so-called Ward identities: $\mathcal{A}\left(1,\dots,n\right)+\mathcal{A}\left(1,\dots,n,n-1\right)+\dots+\mathcal{A}\left(1,n,2,\dots\right)=0$ (and similar for other partial amplitudes). Because of the last property, sometimes more desirable is a decomposition which utilizes only $(n-2)!$ independent partial amplitudes, instead of $(n-1)!$ as in the fundamental-representation (4.19). Such decomposition uses the adjoint generators [147]:

$$\mathcal{M}^{a_1\dots a_n}\left(k_1,\dots,k_n\right)=\frac{1}{2}\sum_{\pi\in S_{n-2}}\left(T^{a_{\pi(2)}}\dots T^{a_{\pi(n-1)}}\right)_{a_1a_n}\,\mathcal{A}\left(1,\pi\left(2\right),\dots,\pi\left(n-1\right),n\right)\,, \tag{4.20}$$

with $\left(T^a\right)_{bc}=-if^{abc}$. The partial amplitudes above are the same as in the fundamental-representation decomposition.

Finally, let us recall the so-called color flow decomposition [148]. It will be useful especially for processes with quarks as it treats gluons and quarks on equal footing. The basic idea is to work with the gluon fields as the elements of the $\mathrm{SU}\left(N_c\right)$ algebra, *i.e.* matrices $\hat{A}^i_j\equiv A_a\left(t^a\right)^i_j$. That is, a gluon is characterized by a pair of fundamental and anti-fundamental representation indices $i,j=\{1,\dots,N_c\}$. In this representation, the amplitude can be decomposed as

$$\mathcal{M}^{i_1\dots i_n}_{j_1\dots j_n}\left(k_1,\dots,k_n\right)=2^{-n/2}\sum_{\pi\in S_{n-1}}\delta^{i_1}_{j_{\pi(2)}}\delta^{i_{\pi(2)}}_{j_{\pi(3)}}\delta^{i_{\pi(3)}}_{j_{\pi(4)}}\dots\delta^{i_{\pi(n)}}_{j_1}\mathcal{A}\left(1,\pi\left(2\right),\dots,\pi\left(n\right)\right)\,, \tag{4.21}$$

again with exactly the same partial amplitudes as in the other two representations.

For processes with quarks, there is also a fundamental representation [149] and for a process with one

quark-antiquark pair,

$$g\left(k_1\right) q\left(k_2\right) g\left(k_3\right) \ldots g\left(k_{n-1}\right) \bar{q}\left(k_n\right) \to \emptyset \,,$$

it reads

$$\mathcal{M}^{i_2^q a_1 a_3 \ldots a_{n-1}}_{j_n^{\bar{q}}}\left(k_1, \ldots, k_n\right) = \sum_{\pi \in S_{n-2}} \left(t^{a_{\pi(1)}} t^{a_{\pi(3)}} \ldots t^{a_{\pi(n-1)}}\right)^{i_2^q}_{j_n^{\bar{q}}} \mathcal{A}\left(2^q, \pi\left(1\right), \pi\left(3\right), \ldots, \pi\left(n-1\right), n^{\bar{q}}\right) \quad (4.22)$$

Above we put superscripts q, $\bar{q}$ to remind which indices belong to a quark (antiquark). However, for processes with quarks we will use the color flow decomposition, as it treats the quarks and gluons on equal footing, and is best for easy calculation of the TMD operator structures. The color flow decomposition for a process with one quark-antiquark pair is given by

$$\mathcal{M}^{i_1 i_2^q i_3 \ldots i_{n-1}}_{j_1 j_3 \ldots j_{n-1} j_n^{\bar{q}}}\left(k_1, \ldots, k_n\right) = 2^{-(n-2)/2} \sum_{\pi \in S_{n-2}} \delta^{i_2^q}_{j_{\pi(1)}} \delta^{i_{\pi(1)}}_{j_{\pi(3)}} \delta^{i_{\pi(3)}}_{j_{\pi(4)}} \ldots \delta^{i_{\pi(n-1)}}_{j_n^{\bar{q}}}$$
$$\mathcal{A}\left(2^q, \pi\left(1\right), \pi\left(3\right), \ldots, \pi\left(n-1\right), n^{\bar{q}}\right) . \quad (4.23)$$

and for a process with two quark-antiquark pairs,

$$g\left(k_1\right) q\left(k_2\right) \bar{q}\left(k_3\right) q\left(k_4\right) g\left(k_5\right) \ldots g\left(k_{n-1}\right) \bar{q}\left(k_n\right) \to \emptyset \,,$$

reads

$$\mathcal{M}^{i_1 i_2^q i_4^q i_5 \ldots i_{n-1}}_{j_1 j_3^{\bar{q}} \ldots j_{n-1} j_n^{\bar{q}}}\left(k_1, \ldots, k_n\right) = 2^{-(n-4)/2} \sum_{\pi \in S_{n-3}} \delta^{i_2^q}_{j^{\bar{q}}_{\{3}} \delta^{i_4^q\}}_{j_{\pi(1)}} \delta^{i_{\pi(1)}}_{j_{\pi(5)}} \delta^{i_{\pi(5)}}_{j_{\pi(6)}} \ldots \delta^{i_{\pi(n-1)}}_{j_n^{\bar{q}}}$$
$$\mathcal{A}\left(2^q, \pi\left(3^{\bar{q}}, 4^q\right), \pi\left(1\right), \pi\left(5\right), \ldots, \pi\left(n-1\right), n^{\bar{q}}\right)$$
$$- \frac{1}{N_c} \sum_{r \in \{1,5,\ldots,n-1\}} \sum_{\pi \in S_{n-4}} \left(\delta^{i_2^q}_{j_{\pi(1)}} \delta^{i_{\pi(1)}}_{j_{\pi(5)}} \ldots \delta^{i_{\pi(r)}}_{j_n^{\bar{q}}}\right) \left(\delta^{i_4^q}_{j_{\pi(r+1)}} \delta^{i_{\pi(r+1)}}_{j_{\pi(r+2)}} \ldots \delta^{i_{\pi(n-1)}}_{j_3^{\bar{q}}}\right)$$
$$\mathcal{A}\left(2^q, \pi\left(1\right), \ldots, \pi\left(r_i\right), n^{\bar{q}}, 4^q, \pi\left(r_{i+1}\right), \ldots, \pi\left(n-1\right), 3^{\bar{q}}\right) . \quad (4.24)$$

In the decomposition above, the first sum runs over all permutations of the $n-4$ gluons and a quark-antiquark pair (the curly brackets in deltas denote that the enclosed indices should be permuted together, according to the permutation π), while the second sum runs over various partitions of the two quark-antiquark pairs with gluon insertions. The second sum is genuinely suppressed by $1/N_c$ in case of distinct quark-antiquark pairs; for identical pairs subleading terms will contribute to both sums in the partial amplitudes. In the present work, we shall explicitly consider processes with up to 6 partons, thus we do not give decomposition for more quark-antiquark pairs.

When constructing the TMD operators, the initial and final states are treated differently, *i.e.* they are assigned different gauge links. Therefore, we have to adjust the color flow decomposition (4.21)-(4.24) to take into account the fact, that two legs are incoming (recall, that these decomposition are within the standard convention of all outgoing partons). This is fixed by making the replacement $i_1 \longleftrightarrow j_1$, $i_n \longleftrightarrow j_n$, as in our convention always the first and the last partons are incoming.

In a sense, there is a price for the simplicity of the color flow decomposition. Namely, to each final state gluon we have to apply the projector

$$\mathcal{P}^{ii'}_{jj'} = \delta^{ii'}\delta_{jj'} - \frac{1}{N_c}\delta^i_j\delta^{i'}_{j'} \,, \quad (4.25)$$

which removes the redundant degrees of freedom from the sum over colors. For pure gluon amplitude they are actually not needed, but must be applied to the quark amplitudes.

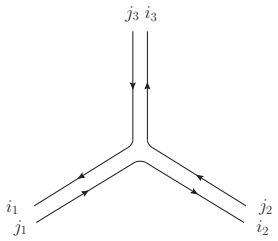
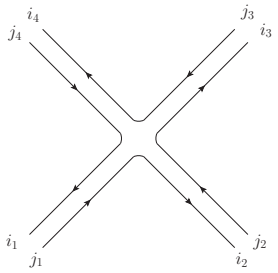
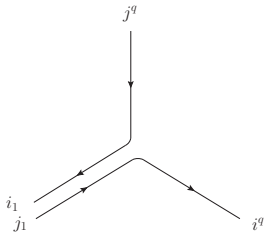
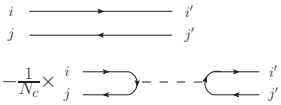
Triple gluon vertex	$\sim \delta^{i_2}_{j_1}\delta^{i_3}_{j_2}\delta^{i_1}_{j_3}$	
Four-gluon vertex	$\sim \delta^{i_2}_{j_1}\delta^{i_3}_{j_2}\delta^{i_4}_{j_3}\delta^{i_1}_{j_4}$	
Quark-gluon vertex	$\sim \delta^{i^q}_{j_1}\delta^{i_1}_{j^q}$	
Gluon propagator	$\sim \left(\delta^{ii'}\delta_{jj'} - \frac{1}{N_c}\delta^i_j\delta^{i'}_{j'}\right)$	

Table 2: Standard color flow Feynman rules for partial amplitudes. All momenta are outgoing. In the middle column we show the color part only.

4.3 Color flow Feynman rules for TMD operators

The color flow Feynman rules (see *e.g.* [148]) are useful for calculating color factors. It turns out that they are also very useful in the context of calculation of the structure of the TMD operators in (4.16), especially, when quarks are involved. We shall supplement the standard color flow rules for color-ordered diagrams (see Table 2) with a set of additional rules which are simple color flow representations of the rules derived in [97] for calculation of a TMD operator structure in an arbitrary process.

The original procedure, described in subsection 4.1, effectively leads to the following recipe. For each *final state*, we assign the gauge link $\mathcal{U}^{[+]}$, which joins the points 0 and ξ through the point in $+\infty$, defined in Eq. (2.10). In the case of gluons, the gauge link is to be defined in adjoint representation. The Wilson link replaces the deltas for color summation when the amplitude is squared: $\delta_{i'i} \to \left(\mathcal{U}^{[+]}\right)_{i'i}$ for quarks, $\delta^{jj'} \to \left(\mathcal{U}^{[+]\dagger}\right)^{jj'}$ for antiquarks and $\delta_{a'a} \to \left(\mathcal{U}^{[+]}\right)_{a'a}$ for gluons (here and in what follows $i, j, k, \ldots$ are fundamental color indices, while $a, b, c, \ldots$ are adjoint). For the *initial state* (not connected to the TMD gluon distribution), the resummation of the initial state interactions leads to the Wilson line extending to $-\infty$, Eq. (2.10). Similar to final states, one needs to replace the color deltas for initial states by the matrix elements of $\mathcal{U}^{[-]}$. The remaining initial state (connected to the TMD) is attached to $F^{i+}_a(\xi)$ in the amplitude and to $F^{i+}_{a'}(0)$ in the conjugate amplitude. The rest of the procedure is similar to calculating color factors: one extracts the color structure of the pertinent amplitude and makes all the contractions (here with Wilson lines and field strength tensors instead of deltas). In the end, one needs to divide out the color factor for a

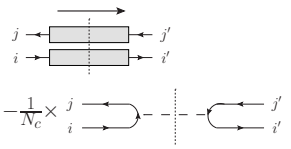
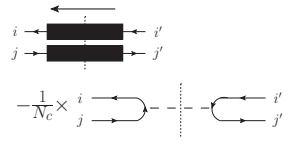
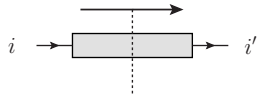
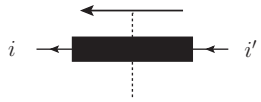
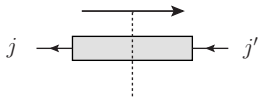
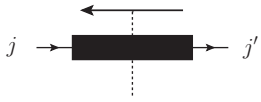
Outgoing gluon	$\left(\mathcal{U}^{[+]}\right)_{i'i}\left(\mathcal{U}^{[+]\dagger}\right)^{jj'} - \frac{1}{N_c}\delta_i^j\delta_{i'}^{j'}$	
Incoming gluon	$\left(\mathcal{U}^{[-]\dagger}\right)^{ii'}\left(\mathcal{U}^{[-]}\right)_{j'j} - \frac{1}{N_c}\delta_j^i\delta_{j'}^{i'}$	
Outgoing quark	$\left(\mathcal{U}^{[+]}\right)_{i'i}$	
Incoming quark	$\left(\mathcal{U}^{[-]\dagger}\right)^{ii'}$	
Outgoing antiquark	$\left(\mathcal{U}^{[+]\dagger}\right)^{jj'}$	
Incoming antiquark	$\left(\mathcal{U}^{[-]}\right)_{j'j}$	
Field strength operators	$2\left(\hat{F}^{+i}(\xi)\right)_i^j\left(\hat{F}^{+i}(0)\right)_{i'}^{j'}$	

Table 3: Color flow Feynman rules for the gauge links. The diagrams correspond to the cut lines, as denoted by the vertical dotted line. The routing in the color loops is clock-wise.

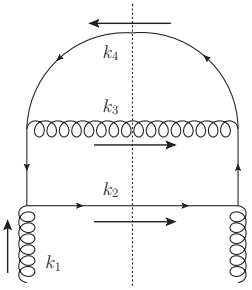
process without gauge links.

Passing to the color flow representation is straightforward. Nothing really is to be done for quarks and antiquarks. For gluons, we first need to make a connection of the adjoint Wilson line with the trace of fundamental-representation instances of the same Wilson line, and next project it onto the fundamental color quantum numbers with the help of the Fierz identity. All rules with graphical representation are collected in Table 3. The procedure of calculating the TMD operator structures is now reduced to considering all possible color flows and applying the rules. Although, in principle, we could consider all standard Feynman diagrams, draw them in the color flow representation and calculate TMD operator structures, fortunately, we do not need to do this. Instead we can just use the color flow decomposition described in Subsection 4.2. This will also ensure, that we work with gauge invariant sets from the start.

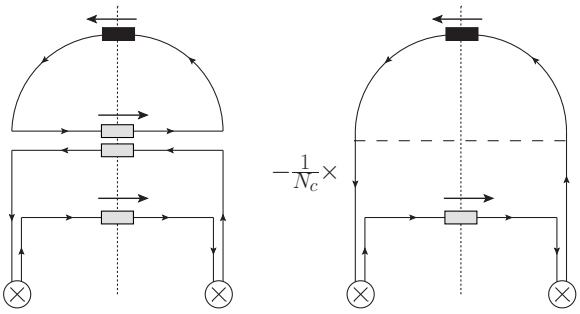
Below, we present some examples to better illustrate the procedure.

4.3.1 Examples

Let us first illustrate the usage of color flow Feynman rules to calculate the structure of the TMD operator for the following diagram:



This diagram contributes to the process $g(k_1)\,q(k_4) \to q(k_2)\,g(k_3)$ and represents the diagram squared and summed over final/initial colors (except insertions of the field operators). The arrows indicate whether the line is incoming/outgoing. Let us stress, that considering particular diagrams is not the way we will ultimately proceed; instead we will consider various color flows as defined in Eqs. (4.21)-(4.24). The structure of the TMD operator for this diagram was calculated in [97]. In the color flow representation we have to consider two diagrams:



The diagram with dashed line represents an exchange of the $U(1)$ gluon (a colorless gluon). To calculate the diagrams we simply look for the closed quark loops and make the trace of the objects appearing in the loop. The direction of the trace is clockwise. The dashed lines carry no color, thus they do not make any traces (they also always accompany $1/N_c$ factors). Note, we calculate only color part (with possible $\mathrm{SU}(N_c)$ matrix insertions) - we are not concerned with any kinematic factors. For the first diagram, we have

$$\mathrm{Tr}\left\{F(\xi)\mathcal{U}^{[+]\dagger}F(0)\mathcal{U}^{[+]}\right\}\mathrm{Tr}\left\{\mathcal{U}^{[\square]}\right\}, \tag{4.26}$$

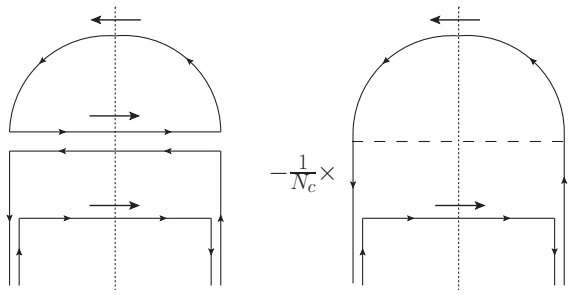
where the first trace corresponds to the bottom loop, the second to the top loop. Above, we defined the Wilson loop [97]

$$\mathcal{U}^{[\square]} = \mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}. \tag{4.27}$$

We also use shorthand notation $F(\xi) \equiv \hat{F}^{i+}\left(\xi^+ = 0, \xi^-, \vec{\xi}_T\right)$. The second diagram reads

$$-\frac{1}{N_c}\,\mathrm{Tr}\left(F(\xi)\mathcal{U}^{[-]\dagger}F(0)\mathcal{U}^{[+]}\right). \tag{4.28}$$

To get the final result, the sum of the two contributions must be divided by the sum of the color factors (without the Wilson lines), with open indices where the field operators are attached:



The part multiplying the open indices reads

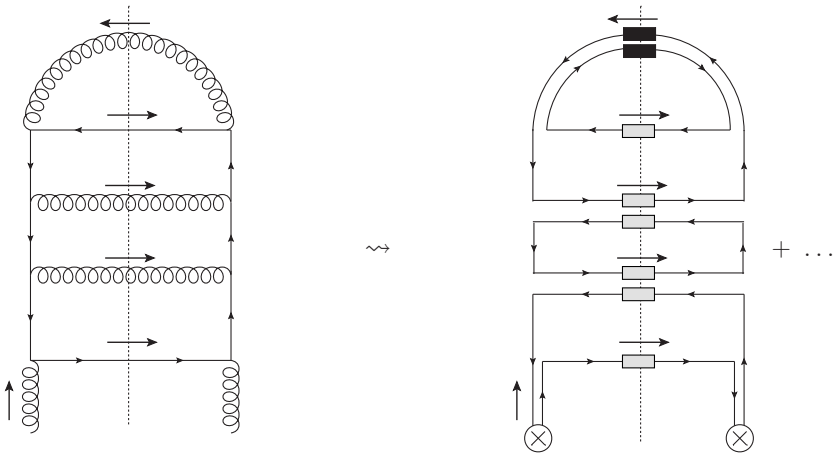
$$N_c - \frac{1}{N_c} = \frac{N_c^2 - 1}{N_c}\,. \tag{4.29}$$

Thus the TMD operator reads

$$\mathrm{Tr}\left\{F\left(\xi\right)\left[\frac{N_c^2}{N_c^2-1}\,\frac{\mathrm{Tr}\mathcal{U}^{[\square]}}{N_c}\,\mathcal{U}^{[+]\dagger} - \frac{1}{N_c^2-1}\mathcal{U}^{[-]\dagger}\right]F\left(0\right)\mathcal{U}^{[+]}\right\}, \tag{4.30}$$

which exactly agrees with the result quoted in [97].

As an illustration of a more complicated structure, let us consider an example contribution to the process $gg \to q\bar{q}gg$:



Applying the color flow rules gives immediately the operator structure for the leading color flow displayed on the r.h.s.:

$$N_c\mathrm{Tr}\left\{F\left(\xi\right)\mathcal{U}^{[+]\dagger}F\left(0\right)\mathcal{U}^{[+]}\right\}\,\mathrm{Tr}\mathcal{U}^{[\square]}\,\mathrm{Tr}\mathcal{U}^{[\square]\dagger}\,. \tag{4.31}$$

Above, the N_c factor comes from the second loop from the bottom, $\mathrm{Tr}\left\{\mathcal{U}^{[+]\dagger}\mathcal{U}^{[+]}\right\} = \mathrm{Tr}\,\mathbb{1} = N_c$.

To close this subsection, let us stress, that the problem of proliferation of color flow diagrams compared to ordinary diagrams, will not concern us at all. As mentioned, we shall use the color flow decomposition, which sets the color flow without need to consider particular diagrams.

4.4 The operator basis for arbitrary TMD gluon distribution

Using the color flow Feynman rules from the previous section we can easily determine all possible 'basis' operators, from which a TMD gluon distribution for arbitrary process can be constructed. Alternatively, one can think about 'basis' TMD gluon distributions.

Plenty of different operators already appear for processes with four colored partons considered in [97]. In order to find all of them, we use the following facts. First, there are at most two $\mathcal{U}^{[-]}$ Wilson lines. This is

the case for initial state gluons where $\mathcal{U}^{[-]}$ and $\mathcal{U}^{[-]\dagger}$ appear. Thus, we can built at most two Wilson loops (4.27), when they are looped with $\mathcal{U}^{[+]}$ or $\mathcal{U}^{[+]\dagger}$ (see the last example in Section 4.3). Second, any color flow loop will contribute trace of at most first power of $\mathcal{U}^{[\pm]}$, $\mathcal{U}^{[\pm]\dagger}$ (and $F(\xi)$, $F(0)$, or both), in addition to mentioned Wilson loops (at most $\mathcal{U}^{[\square]}$ and $\mathcal{U}^{[\square]\dagger}$). This is because for a color flow loop with many Wilson lines (contributed by many final states), most of the Wilson lines will collapse to unity, $\mathcal{U}^{[+]\dagger}\mathcal{U}^{[+]} = \mathbf{1}$, leaving only at most single instances of $\mathcal{U}^{[\pm]}$, $\mathcal{U}^{[\pm]\dagger}$, $\mathcal{U}^{[\square]}$, $\mathcal{U}^{[\square]\dagger}$.

Basing on the above, below we list all 'basis' TMD gluon distributions, from which an arbitrary TMD is given as a linear combination. We assume here, that the *correlators are real* valued functions.

$$\begin{aligned}\mathcal{F}_{qg}^{(1)}(x,k_T) &= 2\int \frac{d\xi^- d^2\xi_T}{(2\pi)^3 P^+} e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \left\langle \mathrm{Tr}\left[\hat{F}^{i+}(\xi)\mathcal{U}^{[-]\dagger}\hat{F}^{i+}(0)\mathcal{U}^{[+]}\right]\right\rangle \\ &= 2\int \frac{d\xi^- d^2\xi_T}{(2\pi)^3 P^+} e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \left\langle \mathrm{Tr}\left[\hat{F}^{i+}(\xi)\mathcal{U}^{[+]\dagger}\hat{F}^{i+}(0)\mathcal{U}^{[-]}\right]\right\rangle ,\end{aligned} \tag{4.32}$$

$$\begin{aligned}\mathcal{F}_{qg}^{(2)}(x,k_T) &= 2\int \frac{d\xi^- d^2\xi_T}{(2\pi)^3 P^+} e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \left\langle \frac{\mathrm{Tr}\left[\mathcal{U}^{[\square]}\right]}{N_c}\mathrm{Tr}\left[\hat{F}^{i+}(\xi)\mathcal{U}^{[+]\dagger}\hat{F}^{i+}(0)\mathcal{U}^{[+]}\right]\right\rangle \\ &= 2\int \frac{d\xi^- d^2\xi_T}{(2\pi)^3 P^+} e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \left\langle \frac{\mathrm{Tr}\left[\mathcal{U}^{[\square]\dagger}\right]}{N_c}\mathrm{Tr}\left[\hat{F}^{i+}(\xi)\mathcal{U}^{[+]\dagger}\hat{F}^{i+}(0)\mathcal{U}^{[+]}\right]\right\rangle ,\end{aligned} \tag{4.33}$$

$$\begin{aligned}\mathcal{F}_{qg}^{(3)}(x,k_T) &= 2\int \frac{d\xi^- d^2\xi_T}{(2\pi)^3 P^+} e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \left\langle \mathrm{Tr}\left[\hat{F}^{i+}(\xi)\mathcal{U}^{[+]\dagger}\hat{F}^{i+}(0)\mathcal{U}^{[\square]}\mathcal{U}^{[+]}\right]\right\rangle \\ &= 2\int \frac{d\xi^- d^2\xi_T}{(2\pi)^3 P^+} e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \left\langle \mathrm{Tr}\left[\hat{F}^{i+}(\xi)\mathcal{U}^{[\square]\dagger}\mathcal{U}^{[+]\dagger}\hat{F}^{i+}(0)\mathcal{U}^{[+]}\right]\right\rangle ,\end{aligned} \tag{4.34}$$

$$\begin{aligned}\mathcal{F}_{gg}^{(1)}(x,k_T) &= 2\int \frac{d\xi^- d^2\xi_T}{(2\pi)^3 P^+} e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \left\langle \frac{\mathrm{Tr}\left[\mathcal{U}^{[\square]\dagger}\right]}{N_c}\mathrm{Tr}\left[\hat{F}^{i+}(\xi)\mathcal{U}^{[-]\dagger}\hat{F}^{i+}(0)\mathcal{U}^{[+]}\right]\right\rangle \\ &= 2\int \frac{d\xi^- d^2\xi_T}{(2\pi)^3 P^+} e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \left\langle \frac{\mathrm{Tr}\left[\mathcal{U}^{[\square]}\right]}{N_c}\mathrm{Tr}\left[\hat{F}^{i+}(\xi)\mathcal{U}^{[+]\dagger}\hat{F}^{i+}(0)\mathcal{U}^{[-]}\right]\right\rangle ,\end{aligned} \tag{4.35}$$

$$\begin{aligned}\mathcal{F}_{gg}^{(2)}(x,k_T) &= 2\int \frac{d\xi^- d^2\xi_T}{(2\pi)^3 P^+} e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \frac{1}{N_c}\left\langle \mathrm{Tr}\left[\hat{F}^{i+}(\xi)\mathcal{U}^{[\square]\dagger}\right]\mathrm{Tr}\left[\hat{F}^{i+}(0)\mathcal{U}^{[\square]}\right]\right\rangle \\ &= 2\int \frac{d\xi^- d^2\xi_T}{(2\pi)^3 P^+} e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \frac{1}{N_c}\left\langle \mathrm{Tr}\left[\hat{F}^{i+}(\xi)\mathcal{U}^{[\square]}\right]\mathrm{Tr}\left[\hat{F}^{i+}(0)\mathcal{U}^{[\square]\dagger}\right]\right\rangle ,\end{aligned} \tag{4.36}$$

$$\mathcal{F}_{gg}^{(3)}(x,k_T) = 2\int \frac{d\xi^- d^2\xi_T}{(2\pi)^3 P^+} e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \left\langle \mathrm{Tr}\left[\hat{F}^{i+}(\xi)\mathcal{U}^{[+]\dagger}\hat{F}^{i+}(0)\mathcal{U}^{[+]}\right]\right\rangle , \tag{4.37}$$

$$\mathcal{F}_{gg}^{(4)}(x,k_T) = 2\int \frac{d\xi^- d^2\xi_T}{(2\pi)^3 P^+} e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \left\langle \mathrm{Tr}\left[\hat{F}^{i+}(\xi)\mathcal{U}^{[-]\dagger}\hat{F}^{i+}(0)\mathcal{U}^{[-]}\right]\right\rangle , \tag{4.38}$$

$$\mathcal{F}_{gg}^{(5)}(x,k_T) = 2\int \frac{d\xi^- d^2\xi_T}{(2\pi)^3 P^+} e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \left\langle \mathrm{Tr}\left[\hat{F}^{i+}(\xi)\mathcal{U}^{[\square]\dagger}\mathcal{U}^{[+]\dagger}\hat{F}^{i+}(0)\mathcal{U}^{[\square]}\mathcal{U}^{[+]}\right]\right\rangle , \tag{4.39}$$

$$\mathcal{F}_{gg}^{(6)}(x,k_T) = 2\int \frac{d\xi^- d^2\xi_T}{(2\pi)^3 P^+} e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \left\langle \frac{\mathrm{Tr}\left[\mathcal{U}^{[\square]}\right]}{N_c}\frac{\mathrm{Tr}\left[\mathcal{U}^{[\square]\dagger}\right]}{N_c}\mathrm{Tr}\left[\hat{F}^{i+}(\xi)\mathcal{U}^{[+]\dagger}\hat{F}^{i+}(0)\mathcal{U}^{[+]}\right]\right\rangle , \tag{4.40}$$

$$\begin{aligned}\mathcal{F}_{gg}^{(7)}\left(x,k_T\right) &= 2\int \frac{d\xi^- d^2\xi_T}{\left(2\pi\right)^3 P^+} e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \left\langle \frac{\mathrm{Tr}\left[\mathcal{U}^{[\square]}\right]}{N_c} \mathrm{Tr}\left[\hat{F}^{i+}\left(\xi\right)\mathcal{U}^{[\square]\dagger}\mathcal{U}^{[+]\dagger}\hat{F}^{i+}\left(0\right)\mathcal{U}^{[+]}\right]\right\rangle \\ &= 2\int \frac{d\xi^- d^2\xi_T}{\left(2\pi\right)^3 P^+} e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \left\langle \frac{\mathrm{Tr}\left[\mathcal{U}^{[\square]\dagger}\right]}{N_c} \mathrm{Tr}\left[\hat{F}^{i+}\left(\xi\right)\mathcal{U}^{[+]\dagger}\hat{F}^{i+}\left(0\right)\mathcal{U}^{[\square]}\mathcal{U}^{[+]}\right]\right\rangle .\end{aligned} \tag{4.41}$$

In the definitions above, the average should be understood as the hadronic matrix elements, *cf.* Eq. (4.16). Two new structures appear in addition to those known in the literature: $\mathcal{F}_{qg}^{(3)}$ and $\mathcal{F}_{gg}^{(7)}$.

Here the subscripts refer to a partonic process to which a given TMD distribution belongs - whether this is a pure gluonic process or a process with quarks [4]. This notation was first introduced in [115] in the context of the small-x limit and we stick to that notation in the present work.

The above set of basic TMD gluon distribution constitutes the basis for any TMD gluon distribution to be convoluted with a hard process. As discussed in Section 2, this assumption is meant to be used at small-x where it can be justified from the CGC effective theory. It is important to note, that it is the complete basis within the rules of [134] - it does not represent a basis for a gluon correlator with arbitrary gauge link structure. At least formally, the basis structures are independent; however, in the large-k_T limit, they start to be degenerate (or vanish).

The definitions above represent bare TMD distributions, which need to be renormalized. In the small-x regime with gluon saturation playing significant role, the evolution equations are nonlinear and thus more complicated than the equations at moderate x [151–153]. In addition, the program of obtaining the renormalization group evolution equations for all possible TMD operators is nowhere near the end. Hopefully, the correspondence of the small-x TMD gluon distributions and CGC correlators [115] allows for a treatment of evolution in the strict small-x limit using the Balitsky–Jalilian–Marian–Iancu–McLerran–Weigert–Leonidov–Kovner (B-JIMWLK) equations [154–160] following Ref. [136]. At small x, but in the linear regime, where the saturation scale is much smaller than the typical scale of the internal transverse momenta, it seems that the various TMD gluon distributions converge to one universal distribution, which may be identified with the unintegrated gluon distribution, the same for any color flow (see Section 3).

In saturation physics, the two TMD distributions with the most elementary Wilson line structure, *i.e.* $\mathcal{F}_{qg}^{(1)}$ and $\mathcal{F}_{gg}^{(3)}$, has been known as the dipole gluon distribution and the Weizsäcker-Williams (WW) gluon distribution, respectively [161]. The WW gluon distribution is calculated from the correlator of two classical gluon fields of relativistic hadrons (non-abelian Weizsäcker-Williams fields) and has a clear physical interpretation as the number density of gluons inside the hadron in light-cone gauge. The dipole distribution is defined as the Fourier transform of the color dipole cross section and does not have a clear partonic interpretation. The dipole density is probed in most QCD processes, like inclusive DIS, semi-inclusive DIS or Drell-Yan, thus it is quite well constrained from data. In contrast, the WW distribution so far has been only calculated from models, however, it can be probed in more complicated processes, like dijet production in γA and pA collisions [115].

Numerical implementation of the iTMD factorization, which will be performed in the next section, requires evaluation of the TMD gluon distributions. The dipole distribution $\mathcal{F}_{qg}^{(1)}$, often denoted $xG^{(2)}$, in the small-x limit can be related to the Fourier transform of the fundamental dipole amplitude $N_F(x,\mathbf{r})$, where $\mathbf{r}$ denote the transverse size of the dipole [115, 136]

$$\mathcal{F}_{qg}^{(1)}\left(x,k_T\right) = \frac{N_c}{\alpha_s\pi(2\pi)^3}\int d^2b\int d^2\mathbf{r}\; e^{-ik_T\cdot\mathbf{r}}\nabla_{\mathbf{r}}^2\; N_F(x,\mathbf{r}) \equiv xG^{(2)}(x,k_T)\,. \tag{4.42}$$

The amplitude N_F is defined through the CGC expectation value of the S-matrix, S_F, of a quark-antiquark dipole scattering off the dense target: $N_F(x,\mathbf{r}) = 1 - S_F(x,\mathbf{r})$ with $S_F(x,\mathbf{r}) = \left\langle \mathrm{Tr}\left[U(\mathbf{r})U^\dagger(\mathbf{0})\right]\right\rangle_x/N_c$ in terms of fundamental Wilson lines [135]. The dipole gluon distribution can then be written as

[4] The notation for the above TMD gluon distributions should not be confused with the double-TMD parton distributions (see *e.g.* [150]).

$$xG^{(2)}(x,k_T) = \frac{N_c\ k_T^2\ S_\perp}{2\pi^2\alpha_s} F(x,k_T)\,, \tag{4.43}$$

where $S_\perp$ denote the transverse area of the target and $F(x,k_T)$ is the Fourier transform of the fundamental dipole

$$F(x,k_T) = \int \frac{d^2\mathbf{r}}{(2\pi)^2} e^{-ik_T\cdot\mathbf{r}} S_F(x,\mathbf{r})\,. \tag{4.44}$$

Calculations of the other TMD distributions are more complicated. The WW distribution, usually denoted $xG^{(1)}$, can be obtained from quadrupole operator, and in general is not related to $F(x,k_T)$. However, the evaluation of the other TMD distributions can be simplified by using a mean-field approximation. In the so-called Gaussian approximation of the CGC [162–168], which assumes that all the color charge correlations in the target stay Gaussian throughout the evolution, and in the large-N_c limit, the WW gluon distribution can be expressed in terms of an adjoint dipole [135]

$$xG^{(1)}(x,k_T) = \frac{C_F}{2\alpha_s\pi^4}\int d^2b \int \frac{d^2\mathbf{r}}{\mathbf{r}^2}\ e^{-ik_T\cdot\mathbf{r}}\ \left[1 - S_A(x,\mathbf{r})\right]\ , \tag{4.45}$$

where $S_A(x,\mathbf{r})$ is an S-matrix for the scattering of a gluon dipole involving adjoint Wilson lines. The Gaussian approximation allows also to write $S_F = S_{BK}^{2C_F/C_A}$ and $S_A = S_{BK}^2$, where S_{BK} is the solution of the BK equation. At large N_c, $S_A(x,\mathbf{r}) = \left[S_F(x,\mathbf{r})\right]^2$, and one can write WW distribution in terms of the dipole distribution and the Fourier transform of the fundamental dipole (see [135] for more detailed calculations)

$$xG^{(1)}(x,k_T) = \frac{1}{2}\int_{k_T^2}^{\infty} dk_T'^2 \ln\left(\frac{k_T'^2}{k_T^2}\right)\int \frac{d^2q_T}{q_T^2}\ xG^{(2)}(x,q_T)F(x,k_T'-q_T)\,. \tag{4.46}$$

As was shown in [82] (which we will verify later in this section), in the large-N_c limit there are five gluon distributions entering the dijet cross section in TMD/iTMD factorization: $\mathcal{F}_{qg}^{(1)}$, $\mathcal{F}_{qg}^{(2)}$, $\mathcal{F}_{gg}^{(1)}$, $\mathcal{F}_{gg}^{(2)}$, $\mathcal{F}_{gg}^{(6)}$ (so the WW distribution is not directly one of them). In the Gaussian approximation and in the large-N_c limit these distributions can be expressed in terms of $xG^{(1)}$ and $xG^{(2)}$ [115]

$$\mathcal{F}_{qg}^{(1)}(x,k_T) = xG^{(2)}(x,q_T)\,, \tag{4.47}$$

$$\mathcal{F}_{qg}^{(2)}(x,k_T) = \int d^2q_T\, xG^{(1)}(x,q_T)F(x,k_T-q_T)\,, \tag{4.48}$$

$$\mathcal{F}_{gg}^{(1)}(x,k_T) = \int d^2q_T\, xG^{(2)}(x,q_T)F(x,k_T-q_T)\,, \tag{4.49}$$

$$\mathcal{F}_{gg}^{(2)}(x,k_T) = -\int d^2q_T \frac{(k_T-q_T)\cdot q_T}{q_T^2}\, xG^{(2)}(x,q_T)F(x,k_T-q_T)\,, \tag{4.50}$$

$$\mathcal{F}_{gg}^{(6)}(x,k_T) = \int d^2q_T d^2q_T'\, xG^{(1)}(x,q_T)F(x,q_T')F(x,k_T-q_T-q_T')\,. \tag{4.51}$$

Therefore, employing Eq. (4.43) and Eq. (4.46) they can be obtained from the solution of the BK equation in the momentum space $F(x,k_T)$. In Ref. [135] these TMD distributions were calculated using the KS gluon distribution [111], described in subsection 3.2, which provides directly the dipole gluon $xG^{(2)}(x,k_T)$. The resulting TMD distributions as a function of k_T are shown in Fig. 19. There is a small mismatch in their high-k_T behavior, due to the initial condition for the evolution in x in KS gluon.

4.5 Operator structures for multi-parton processes

In this subsection we present the calculations of TMD gluon distributions for the processes with 3-6 colored partons, together with their large N_c limit, which will be useful phenomenologically in short run.

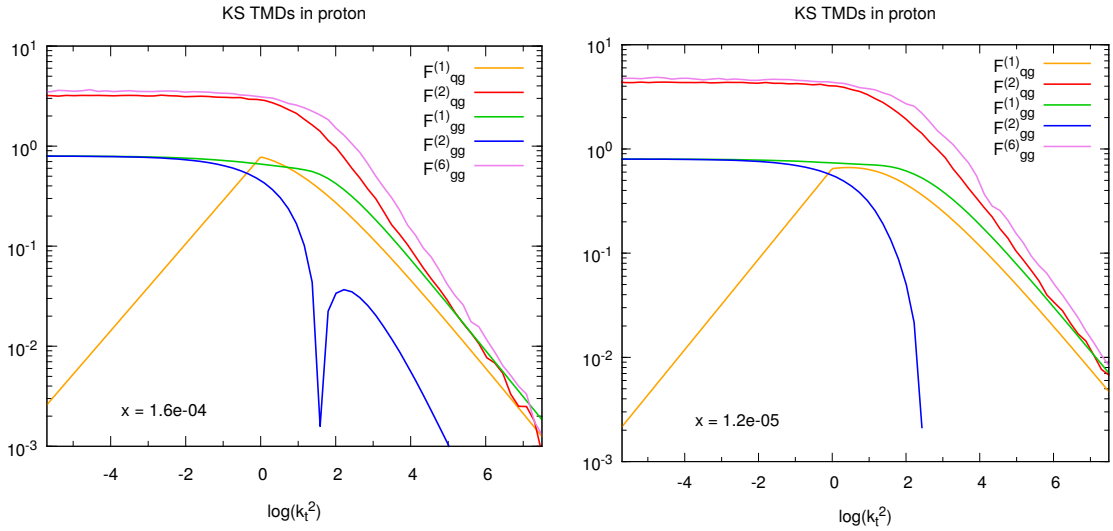


Figure 19: The KS gluon TMDs as a function of $\log(k_T^2/\mathrm{GeV}^2)$ for the proton (left) and the lead nucleus (right). Since $\mathcal{F}_{gg}^{(2)}$ goes negative, its absolute value is shown on the figures. From [135].

We will start with a derivation of 3 parton TMD operators, to demonstrate the procedure utilizing the color decomposition and to introduce the general notation we shall use for more complicated processes (the operator structures for 4 parton processes were first obtained in [97], using the standard Feynman diagram method, as shown in the previous subsections, and in [82] using the color decomposition). The results presented here were obtained using FORM program GTMDCALC [169].

4.5.1 3 partons

This is a good example to demonstrate the algorithm which is used for processes with more partons, since the decompositions consists of only one term, represented by a single diagram, depicted in Fig. 20.

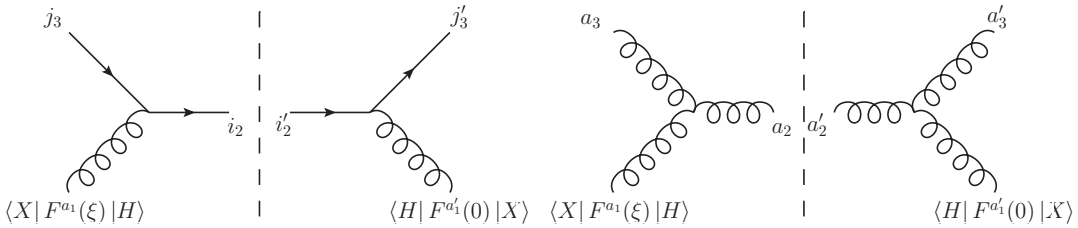


Figure 20: Diagrams contributing to single jet production.

$gq \to q$

The process

$$g(k_1)q(k_3) \to q(k_2)\,, \tag{4.52}$$

shown in Fig. 20(left), can be decomposed in both the fundamental basis, Eq. (4.22) and the color flow basis, Eq. (4.23). The final result is of course the same. The decomposition in the color flow basis is

$$\mathcal{M}^{i_1 i_2}_{j_1 j_3} = \frac{1}{\sqrt{2}}\delta^{i_2}_{j_1}\delta^{i_1}_{j_3}\mathcal{A}(2,1,3)\,, \tag{4.53}$$

while in the fundamental basis we have

$$\mathcal{M}_{j_3}^{i_2 a_1} = (t^{a_1})_{j_3}^{i_2}\, \mathcal{A}(2,1,3)\,. \tag{4.54}$$

As in this case we have only one permutation $\mathcal{A}(2,1,3)$, we will not introduce matrix notation here. The square of amplitude, summed over color is

$$(t^{a_1})_{j_3}^{i_2} \left(t^{a_1'}\right)_{i_2'}^{j_3'} \delta^{A_1 A_1'} \delta_{i_2'}^{i_2}\, \delta_{j_3}^{j_3'} = T_F \left(\delta_{i_2'}^{i_2}\delta_{j_3}^{j_3'} - \frac{1}{N_c}\delta_{j_3}^{i_2}\delta_{i_2'}^{j_3'}\right) \delta_{i_2'}^{i_2}\, \delta_{j_3}^{j_3'} = T_F\left(N_c^2 - 1\right)\,. \tag{4.55}$$

It is calculated starting from the fundamental decomposition in this case. In order to calculate the TMD operator structure, we need to insert the appropriate gauge links instead of deltas summing over colors, as reviewed in Section 4.3. In this case, we assign the future-pointing gauge-link $\mathcal{U}^{[+]}$ for the outgoing quark and the past-pointing gauge-link $\mathcal{U}^{[-]\dagger}$ for the incoming quark to obtain

$$\mathcal{M}_{j_3}^{i_2 a_1} \mathcal{M}_{i_2' a_1'}^{\dagger j_3'} \left(\mathcal{U}^{[+]}\right)_{i_2'}^{i_2} \left(\mathcal{U}^{[-]\dagger}\right)_{j_3}^{j_3'} F_{a_1}^{i+}(\xi)\, F_{a_1'}^{i+}(0)\,. \tag{4.56}$$

Since these Wilson lines are in the fundamental representation, we immediately obtain the following structure

$$\begin{aligned}(t^{A_1})_{j_3}^{i_2} \left(t^{A_1'}\right)_{i_2'}^{j_3'} \left(\mathcal{U}^{[+]}\right)_{i_2'}^{i_2} \left(\mathcal{U}^{[-]\dagger}\right)_{j_3}^{j_3'} F^{A_1'}(0) F^{A_1}(\xi) = \\ = (F(\xi))_{j_3}^{i_2} \left(\mathcal{U}^{[-]\dagger}\right)_{j_3}^{j_3'} (F(0))_{i_2'}^{j_3'} \left(\mathcal{U}^{[+]}\right)_{i_2'}^{i_2} = \mathrm{Tr}\left[F(\xi)\mathcal{U}^{[-]\dagger} F(0)\mathcal{U}^{[+]}\right]\,.\end{aligned} \tag{4.57}$$

The above result must be divided by the color factor of the squared amplitude, but without summation of the indices of the gluon for which we calculate the TMD operator, which is given by Eq. 4.55 divided by $\left(N_c^2 - 1\right)$. Therefore, the considered structure, divided by T_F gives

$$2\int \frac{d\xi^- d^2\xi_T}{(2\pi)^3 P^+}\, e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T} \left\langle \mathrm{Tr}\left[\hat{F}^{i+}(\xi)\mathcal{U}^{[-]\dagger}\hat{F}^{i+}(0)\mathcal{U}^{[+]}\right]\right\rangle = \mathcal{F}_{qg}^{(1)}(x, k_T)\,. \tag{4.58}$$

$gg \to g$

The process, represented in Fig. 20(right)

$$g(k_1)\, g(k_3) \to g(k_2)\,. \tag{4.59}$$

has the following decomposition

$$\mathcal{M} = T_{a_1 a_3}^{a_2} \mathcal{A}(1,2,3) = i f^{a_1 a_2 a_3} \mathcal{A}(1,2,3)\,. \tag{4.60}$$

Here we need to insert gauge-links in the adjoint representation, the $\mathcal{U}^{[+]}$ for the outgoing gluon and $\mathcal{U}^{[-]\dagger}$ for the incoming gluon, and thus we obtain

$$i f^{a_1 a_2 a_3} \left(-i f^{a_1' a_2' a_3'}\right) \left(\mathcal{U}^{[+]}\right)^{a_2' a_2} \left(\mathcal{U}^{[-]\dagger}\right)^{a_3 a_3'} F^{a_1'}(0) F^{a_1}(\xi)\,. \tag{4.61}$$

The Wilson lines in the adjoint representation are then transformed using Eq. (4.4 to the fundamental representation, which leads to

$$-\frac{F^{a_1'}(0) F^{a_1}(\xi)}{T_F^2}\, i f^{a_1 a_2 a_3}\, i f^{a_1' a_2' a_3'}\, \mathrm{Tr}\left[t^{a_2'}\mathcal{U}^{[+]} t^{a_2}\mathcal{U}^{[+]\dagger}\right]\, \mathrm{Tr}\left[t^{a_3}\mathcal{U}^{[-]\dagger} t^{a_3'}\mathcal{U}^{[-]}\right]\,. \tag{4.62}$$

Lets take care of the trace part first

$$\mathrm{Tr}\left[t^{a'_2}\mathcal{U}^{[+]}[t^{a_1},t^{a_3}]\mathcal{U}^{[+]\dagger}\right]\,\mathrm{Tr}\left[t^{a_3}\mathcal{U}^{[-]\dagger}[t^{a'_1},t^{a'_2}]\mathcal{U}^{[-]}\right]= \tag{4.63}$$

$$=t^{a'_2}_{ij}\mathcal{U}^{[+]}_{jk}\left(t^{a_1}_{kl}t^{a_3}_{lm}-t^{a_3}_{kl}t^{a_1}_{lm}\right)\mathcal{U}^{[+]\dagger}_{mi}\times t^{a_3}_{ab}\mathcal{U}^{[-]\dagger}_{bc}\left(t^{a'_1}_{cd}t^{a'_2}_{de}-t^{a'_2}_{cd}t^{a'_1}_{de}\right)\mathcal{U}^{[-]}_{ea}= \tag{4.64}$$

$$=\mathcal{U}^{[+]}_{jk}\mathcal{U}^{[+]\dagger}_{mi}\mathcal{U}^{[-]\dagger}_{bc}\mathcal{U}^{[-]}_{ea}\left(t^{a_1}_{kl}t^{a_3}_{lm}t^{a_3}_{ab}-t^{a_3}_{ab}t^{a_3}_{kl}t^{a_1}_{lm}\right)\left(t^{a'_1}_{cd}t^{a'_2}_{de}t^{a'_2}_{ij}-t^{a'_2}_{ij}t^{a'_2}_{cd}t^{a'_1}_{de}\right)= \tag{4.65}$$

$$\begin{aligned}=T_F^2\,\mathcal{U}^{[+]}_{jk}\mathcal{U}^{[+]\dagger}_{mi}\mathcal{U}^{[-]\dagger}_{bc}\mathcal{U}^{[-]}_{ea}\left[t^{a_1}_{kl}\left(\delta_{lb}\delta_{ma}-\frac{1}{N_c}\delta_{lm}\delta_{ab}\right)-\left(\delta_{al}\delta_{bk}-\frac{1}{N_c}\delta_{ab}\delta_{kl}\right)t^{a_1}_{lm}\right]\times\\ \times\left[t^{a'_1}_{cd}\left(\delta_{dj}\delta_{ei}-\frac{1}{N_c}\delta_{de}\delta_{ij}\right)-\left(\delta_{id}\delta_{jc}-\frac{1}{N_c}\delta_{ij}\delta_{cd}\right)t^{a'_1}_{de}\right]=\end{aligned} \tag{4.66}$$

$$\begin{aligned}=T_F^2\,\mathcal{U}^{[+]}_{jk}\mathcal{U}^{[+]\dagger}_{mi}\mathcal{U}^{[-]\dagger}_{bc}\mathcal{U}^{[-]}_{ea}\left[\left(t^{a_1}_{kb}\delta_{ma}-\frac{1}{N_c}t^{a_1}_{km}\delta_{ab}\right)-\left(t^{a_1}_{am}\delta_{bk}-\frac{1}{N_c}t^{a_1}_{km}\delta_{ab}\right)\right]\times\\ \times\left[\left(t^{a'_1}_{cj}\delta_{ei}-\frac{1}{N_c}t^{a'_1}_{ce}\delta_{ij}\right)-\left(t^{a'_1}_{ie}\delta_{jc}-\frac{1}{N_c}t^{a'_1}_{ce}\delta_{ij}\right)\right]=\end{aligned} \tag{4.67}$$

$$=T_F^2\,\mathcal{U}^{[+]}_{jk}\mathcal{U}^{[+]\dagger}_{mi}\mathcal{U}^{[-]\dagger}_{bc}\mathcal{U}^{[-]}_{ea}\left[t^{a_1}_{kb}\delta_{ma}-t^{a_1}_{am}\delta_{bk}\right]\left[t^{a'_1}_{cj}\delta_{ei}-t^{a'_1}_{ie}\delta_{jc}\right]= \tag{4.68}$$

$$=T_F^2\,\left[\mathcal{U}^{[+]}_{jk}t^{a_1}_{kb}\mathcal{U}^{[-]\dagger}_{bc}\mathcal{U}^{[+]\dagger}_{ai}\mathcal{U}^{[-]}_{ea}-\mathcal{U}^{[+]}_{jb}t^{a_1}_{am}\mathcal{U}^{[+]\dagger}_{mi}\mathcal{U}^{[-]\dagger}_{bc}\mathcal{U}^{[-]}_{ea}\right]\left[t^{a'_1}_{cj}\delta_{ei}-t^{a'_1}_{ie}\delta_{jc}\right]= \tag{4.69}$$

$$\begin{aligned}=T_F^2\,\left(t^{a_1}_{kb}\mathcal{U}^{[-]\dagger}_{bc}t^{a'_1}_{cj}\mathcal{U}^{[+]}_{jk}\mathcal{U}^{[+]\dagger}_{ae}\mathcal{U}^{[-]}_{ea}-t^{a_1}_{kb}\mathcal{U}^{[-]\dagger}_{bc}\mathcal{U}^{[+]}_{ck}\mathcal{U}^{[+]\dagger}_{ai}\mathcal{U}^{[-]}_{ea}t^{a'_1}_{ie}\right)-\\ -T_F^2\,\left(t^{a_1}_{am}\mathcal{U}^{[+]\dagger}_{me}\mathcal{U}^{[-]}_{ea}\mathcal{U}^{[-]\dagger}_{bc}t^{a'_1}_{cj}\mathcal{U}^{[+]}_{jb}-t^{a_1}_{am}\mathcal{U}^{[+]\dagger}_{mi}t^{a'_1}_{ie}\mathcal{U}^{[-]}_{ea}\mathcal{U}^{[-]\dagger}_{bc}\mathcal{U}^{[+]}_{cb}\right)=\end{aligned} \tag{4.70}$$

$$\begin{aligned}=T_F^2\,\left(\mathrm{Tr}\left[t^{a_1}\mathcal{U}^{[-]\dagger}t^{a'_1}\mathcal{U}^{[+]}\right]\mathrm{Tr}\left[\mathcal{U}^{[\square]\dagger}\right]-\mathrm{Tr}\left[t^{a_1}\mathcal{U}^{[\square]}\right]\mathrm{Tr}\left[t^{a'_1}\mathcal{U}^{[\square]\dagger}\right]\right)-\\ -T_F^2\,\left(\mathrm{Tr}\left[t^{a_1}\mathcal{U}^{[\square]\dagger}\right]\mathrm{Tr}\left[t^{a'_1}\mathcal{U}^{[\square]}\right]-\mathrm{Tr}\left[t^{a_1}\mathcal{U}^{[+]\dagger}t^{a'_1}\mathcal{U}^{[-]}\right]\mathrm{Tr}\left[\mathcal{U}^{[\square]}\right]\right)\end{aligned} \tag{4.71}$$

After restoring the fields we obtain

$$\begin{aligned}\left(\mathrm{Tr}\left[F(\xi)\mathcal{U}^{[-]\dagger}F(0)\mathcal{U}^{[+]}\right]\mathrm{Tr}\left[\mathcal{U}^{[\square]\dagger}\right]-\mathrm{Tr}\left[F(\xi)\mathcal{U}^{[\square]}\right]\mathrm{Tr}\left[F(0)\mathcal{U}^{[\square]\dagger}\right]\right)-\\ -\left(\mathrm{Tr}\left[F(\xi)\mathcal{U}^{[\square]\dagger}\right]\mathrm{Tr}\left[F(0)\mathcal{U}^{[\square]}\right]-\mathrm{Tr}\left[F(\xi)\mathcal{U}^{[+]\dagger}F(0)\mathcal{U}^{[-]}\right]\mathrm{Tr}\left[\mathcal{U}^{[\square]}\right]\right).\end{aligned} \tag{4.72}$$

There are four structures, however two of them are the complex conjugate of the other two. Expressed in the basis distributions, the result is

$$N_c\mathcal{F}^{(1)}_{gg}-N_c\mathcal{F}^{(2)}_{gg}\,, \tag{4.73}$$

which after dividing by the color factor of the square of amplitude (without N_c^2-1) gives the TMD distribution

$$\mathcal{F}^{(1)}_{gg}-\mathcal{F}^{(2)}_{gg}\,. \tag{4.74}$$

4.5.2 4 partons

As the color decomposition is most straightforward for pure gluonic amplitude, let us start with the process

$$g\left(k_1\right) g\left(k_4\right) \to g\left(k_2\right) g\left(k_3\right) \,. \tag{4.75}$$

For gluons, three color decompositions can be used: the fundamental (4.19), the color flow (4.21), and the adjoint (4.20). First two involve 6 partial amplitudes, while the last one, only two. As mentioned in Section 4.2, the 6 partial amplitudes are not independent, but their squares give the leading contribution in the large N_c limit - a property which we will use in Section 4.6. Here, we are interested in the full answer, thus we use the adjoint color decomposition (for processes with quarks we will use exclusively color flow decomposition). It reads

$$\mathcal{M}^{a_1a_2a_3a_4}\left(k_1,k_2,k_3,k_4\right) = \frac{1}{2}\left(T^{a_2}T^{a_3}\right)_{a_1a_4}\,\mathcal{A}(1,2,3,4) + \frac{1}{2}\left(T^{a_3}T^{a_2}\right)_{a_1a_4}\,\mathcal{A}(1,3,2,4)\,. \tag{4.76}$$

The square of the amplitude, summed over colors, can in general be written in a matrix form

$$|\mathcal{M}|^2 = \vec{\mathcal{A}}^{\,\dagger}\,\mathbf{C}\,\vec{\mathcal{A}}\,, \tag{4.77}$$

where $\mathbf{C}$ is the color matrix and $\vec{\mathcal{A}}$ is a column vector constructed from the partial amplitudes. For the present case

$$\mathbf{C} = \frac{1}{4}N_c^2N_A\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}\,, \tag{4.78}$$

and $\vec{\mathcal{A}}$ is a column vector constructed from the partial amplitudes, given in Table 4. In order to calculate the TMD operator structure, we need to insert the appropriate gauge links instead of deltas summing over colors (as reviewed in Section 4.3)

$$\mathcal{M}^{a_1a_2a_3a_4}\mathcal{M}^{\dagger}_{a_1'a_2'a_3'a_4'}\left(\mathcal{U}^{[+]}\right)^{a_2'a_2}\left(\mathcal{U}^{[+]}\right)^{a_3'a_3}\left(\mathcal{U}^{[-]\dagger}\right)^{a_4a_4'}F^{i+}_{a_1}\left(\xi\right)F^{i+}_{a_1'}\left(0\right)\,. \tag{4.79}$$

The Wilson lines in adjoint representation are transformed to the fundamental representation using Eq. (4.4. Next, the decomposition (4.76) is used to represent the above expression in the following general form

$$\vec{\mathcal{A}}^{\,\dagger}\,\boldsymbol{F}\,\vec{\mathcal{A}}\,, \tag{4.80}$$

where $\boldsymbol{F}$ is the matrix of the TMD operators containing implicitly the color factors of the hard process. In most cases, it is reasonable to keep these color factors together with the hard matrix elements. Thus, to avoid double counting, we divide the elements of $\boldsymbol{F}$ by the corresponding color factors of the square of the amplitude, but without the summation of indices where the field operators are attached (this corresponds to the elements of the matrix $\mathbf{C}$ (4.78) divided by N_A). This leads to the following definition of the TMD distribution matrix

$$\boldsymbol{\Phi} = 2\int\frac{d\xi^-d^2\xi_T}{\left(2\pi\right)^3P^+}\,e^{ixP^+\xi^- - i\vec{k}_T\cdot\vec{\xi}_T}\,\left\langle P\right|\boldsymbol{F}\oslash\left(\frac{1}{N_A}\mathbf{C}\right)\left|P\right\rangle\,, \tag{4.81}$$

where the symbol $\oslash$ represents the Hadamard division, *i.e.* the element-wise division: $\left(\boldsymbol{A}\oslash\boldsymbol{B}\right)_{ij} = \boldsymbol{A}_{ij}/\boldsymbol{B}_{ij}$. It may happen, for certain multiparticle processes, that some elements of the color matrix $\mathbf{C}$ vanish, but the corresponding elements of $\boldsymbol{F}$ are non-zero. In that case, we need to modify the above prescription. We shall come back to this point when discussing processes where this happens. An additional motivation to divide out the color factors from the TMD operators is that one could in principle use the results with matrix elements not represented in the color-ordered form.

With the above definitions, the cross section for a collinear parton a to scatter off a gluon with some

$g_1g_4 \to g_2g_3$	$g_1q_4 \to g_2q_3$	$g_1\bar{q}_4 \to g_2\bar{q}_3$	$g_1g_4 \to q_2\bar{q}_3$
$\begin{pmatrix} \mathcal{A}(1,2,3,4) \\ \mathcal{A}(1,3,2,4) \end{pmatrix}$	$\begin{pmatrix} \mathcal{A}(3,1,2,4) \\ \mathcal{A}(3,2,1,4) \end{pmatrix}$	$\begin{pmatrix} \mathcal{A}(4,1,2,3) \\ \mathcal{A}(4,2,1,1) \end{pmatrix}$	$\begin{pmatrix} \mathcal{A}(2,1,4,3) \\ \mathcal{A}(2,4,1,3) \end{pmatrix}$

Table 4: Definitions of the vector of partial amplitudes for all four-parton processes. The subscripts in the sub-process indication correspond to the momenta enumeration.

internal transverse momentum and producing certain number of colored partons, can be generically written as

$$d\sigma_{ag\to X} = \int \vec{\mathcal{A}}^{\dagger} \left(\boldsymbol{C} \circ \boldsymbol{\Phi}_{ag\to X}\right) \vec{\mathcal{A}} \, d\Gamma \,, \tag{4.82}$$

where $d\Gamma$ represents all pre-factors, phase space, and convolution in x and k_T. The symbol $\circ$ is the Hadamard (element-wise) multiplication, $(\boldsymbol{A} \circ \boldsymbol{B})_{ij} = \boldsymbol{A}_{ij}\boldsymbol{B}_{ij}$.

In the present example of four gluons, the TMD gluon distribution matrix reads

$$\boldsymbol{\Phi}_{gg\to gg} = \begin{pmatrix} \Phi_1 & \Phi_2 \\ \Phi_2 & \Phi_1 \end{pmatrix} , \tag{4.83}$$

with two independent TMD gluon distributions expressed in terms of the basis distributions:

$$\Phi_1 = \frac{1}{2N_c^2}\left(N_c^2\mathcal{F}_{\rm gg}^{(1)} - 2\mathcal{F}_{\rm gg}^{(3)} + \mathcal{F}_{\rm gg}^{(4)} + \mathcal{F}_{\rm gg}^{(5)} + N_c^2\mathcal{F}_{\rm gg}^{(6)}\right) , \tag{4.84}$$

$$\Phi_2 = \frac{1}{N_c^2}\left(N_c^2\mathcal{F}_{\rm gg}^{(2)} - 2\mathcal{F}_{\rm gg}^{(3)} + \mathcal{F}_{\rm gg}^{(4)} + \mathcal{F}_{\rm gg}^{(5)} + N_c^2\mathcal{F}_{\rm gg}^{(6)}\right) . \tag{4.85}$$

For more complicated processes with gluons it is useful to write the above equations in a matrix form:

$$\begin{pmatrix} \Phi_1 \\ \vdots \\ \Phi_k \end{pmatrix} = \mathbf{M} \begin{pmatrix} \mathcal{F}_{\rm gg}^{(1)} \\ \mathcal{F}_{\rm gg}^{(2)} \\ \vdots \\ \mathcal{F}_{\rm gg}^{(7)} \end{pmatrix} , \tag{4.86}$$

where $\mathbf{M}$ is a matrix with k rows and 7 columns. For the case of four gluons, this matrix reads

$$\mathbf{M}_{gg\to gg} = \begin{pmatrix} \frac{1}{2} & 0 & -\frac{1}{N_c^2} & \frac{1}{2N_c^2} & \frac{1}{2N_c^2} & \frac{1}{2} & 0 \\ 0 & 1 & -\frac{2}{N_c^2} & \frac{1}{N_c^2} & \frac{1}{N_c^2} & 1 & 0 \end{pmatrix} . \tag{4.87}$$

In a similar fashion, one can derive the matrices $\boldsymbol{\Phi}$ and $\mathbf{M}$ for other 4-parton channels. The only difference is that for processes with quarks, we always use the color flow decomposition of an amplitude. For the channel

$$g\,(k_1)\,q\,(k_4) \to g\,(k_2)\,q\,(k_3) \,, \tag{4.88}$$

we obtain

$$\boldsymbol{\Phi}_{gq\to gq} = \begin{pmatrix} \Phi_2 & \Phi_1 \\ \Phi_1 & \Phi_1 \end{pmatrix} , \tag{4.89}$$

with the Φ_i given in Table 5. For a similar process with an antiquark, we get

$$\boldsymbol{\Phi}_{g\bar{q}\to g\bar{q}} = \begin{pmatrix} \Phi_1 & \Phi_1 \\ \Phi_1 & \Phi_2 \end{pmatrix} . \tag{4.90}$$

$g_1g_4 \to g_2g_3$	$g_1g_4 \to q_2\bar{q}_3$
$\begin{pmatrix} \frac{1}{2} & 0 & -\frac{1}{N_c^2} & \frac{1}{2N_c^2} & \frac{1}{2N_c^2} & \frac{1}{2} \\ 0 & 1 & -\frac{2}{N_c^2} & \frac{1}{N_c^2} & \frac{1}{N_c^2} & 1 \end{pmatrix}$	$\begin{pmatrix} \frac{N_c^2}{N_A} & 0 & -\frac{1}{N_A} & 0 & 0 & 0 \\ 0 & -N_c^2 & 1 & 0 & 0 & 0 \end{pmatrix}$
$g_1q_4 \to g_2q_3$	$g_1\bar{q}_4 \to g_2\bar{q}_3$
$\begin{pmatrix} 1 & 0 \\ -\frac{1}{N_A} & \frac{N_c^2}{N_A} \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ -\frac{1}{N_A} & \frac{N_c^2}{N_A} \end{pmatrix}$

Table 5: Matrices $\mathbf{M}$ of structures appearing in four-parton processes. The subscripts in the sub-process indication correspond to the momenta enumeration.

Finally, for

$$g\left(k_1\right) g\left(k_4\right) \to q\left(k_2\right) \bar{q}\left(k_3\right) , \tag{4.91}$$

we have

$$\mathbf{\Phi}_{gg\to q\bar{q}} = \begin{pmatrix} \Phi_1 & \Phi_2 \\ \Phi_2 & \Phi_1 \end{pmatrix} . \tag{4.92}$$

The partial amplitude vectors $\vec{\mathcal{A}}$ for the above cases are listed in Table 4.

4.5.3 5 partons

The calculation of the TMD gluon distributions with 5 colored partons proceeds in the same fashion, but is technically more complicated. Also, a new feature appears. Certain color factors, building up the matrix $\mathbf{C}$, vanish for some processes. However, some of the corresponding TMD operators do not vanish (more precisely, we mean here corresponding elements of the $\mathbf{F}$ matrix). It is a special property of the TMD factorization: certain color flows would not contribute in the collinear factorization (where only the matrix $\mathbf{C}$ appears), but they do contribute if the TMD gluon distributions are considered. Thus we need to modify the definition of the TMD gluon distribution matrix $\mathbf{\Phi}$ (4.81) and the Eq. (4.82) for such processes. In both formulas, instead of the matrix $\mathbf{C}$ (which has zeros), we use the matrix $\mathbf{C}'$ with elements

$$\mathbf{C}'_{ij} = \begin{cases} \mathbf{C}_{ij} & \text{if } \mathbf{C}_{ij} \neq 0 \\ 1 & \text{if } \mathbf{C}_{ij} = 0 \end{cases} . \tag{4.93}$$

This is a simple way to extract the hard matrix element color factors only from those TMD operators, for which the color factor is nonzero. For reader's convenience, the color factors for 5 parton processes in the color-ordered-amplitude representation are collected in Appendix E (they were cross-checked with [170], [147]).

Below, we present the TMD gluon distribution matrices $\mathbf{\Phi}$ for various channels. The vectors $\vec{\mathcal{A}}$ of partial amplitudes, corresponding to the entries of the matrices $\mathbf{\Phi}$, are given in Table 6 in Appendix B. The TMD gluon distributions Φ_i building up these matrices, are expressed through the 'basis' distributions (4.32)-(4.41), as given by the $\mathbf{M}$ matrices listed in Table 7 (Appendix B). The $\mathbf{M}$ matrices for processes in which the incoming and outgoing quarks are replaced by the incoming and outgoing antiquarks are the same.

For the pure gluonic process,

$$g\left(k_1\right) g\left(k_5\right) \to g\left(k_2\right) g\left(k_3\right) g\left(k_4\right) , \tag{4.94}$$

we obtain

$$\mathbf{\Phi}_{gg\to ggg} = \begin{pmatrix} \Phi_1 & \Phi_2 & \Phi_2 & \Phi_3 & \Phi_3 & \Phi_4^* \\ \Phi_2 & \Phi_1 & \Phi_3 & \Phi_4^* & \Phi_2 & \Phi_3 \\ \Phi_2 & \Phi_3 & \Phi_1 & \Phi_2 & \Phi_4^* & \Phi_3 \\ \Phi_3 & \Phi_4^* & \Phi_2 & \Phi_1 & \Phi_3 & \Phi_2 \\ \Phi_3 & \Phi_2 & \Phi_4^* & \Phi_3 & \Phi_1 & \Phi_2 \\ \Phi_4^* & \Phi_3 & \Phi_3 & \Phi_2 & \Phi_2 & \Phi_1 \end{pmatrix} . \tag{4.95}$$

The Φ_i gluon distributions are listed in the first row of Table 7. This process has the property mentioned in the beginning of this subsection *i.e.* some color factors in the matrix $\mathbf{C}$ vanish while the corresponding TMD operators do not vanish. The entries for which the color factors are zero are marked with the asterix *.

For

$$g\,(k_1)\,g\,(k_5) \to q\,(k_2)\,\bar{q}\,(k_3)\,g\,(k_4)\ , \tag{4.96}$$

we get

$$\mathbf{\Phi}_{gg\to q\bar{q}g} = \begin{pmatrix} \Phi_1 & \Phi_2 & \Phi_2 & \Phi_3 & \Phi_3 & \Phi_4 \\ \Phi_2 & \Phi_2 & \Phi_5 & \Phi_6 & \Phi_3 & \Phi_3 \\ \Phi_2 & \Phi_5 & \Phi_2 & \Phi_3 & \Phi_6 & \Phi_3 \\ \Phi_3 & \Phi_6 & \Phi_3 & \Phi_2 & \Phi_5 & \Phi_2 \\ \Phi_3 & \Phi_3 & \Phi_6 & \Phi_5 & \Phi_2 & \Phi_2 \\ \Phi_4 & \Phi_3 & \Phi_3 & \Phi_2 & \Phi_2 & \Phi_1 \end{pmatrix} , \tag{4.97}$$

with the TMD gluon distributions given in the second row of Table 7.

For the process with initial state quark

$$g\,(k_1)\,q\,(k_5) \to g\,(k_2)\,g\,(k_3)\,q\,(k_4)\ , \tag{4.98}$$

or antiquark, we obtain, respectively

$$\mathbf{\Phi}_{gq\to ggq} = \begin{pmatrix} \Phi_1 & \Phi_2 & \Phi_3 & \Phi_4 & \Phi_5 & \Phi_4 \\ \Phi_2 & \Phi_1 & \Phi_5 & \Phi_4 & \Phi_3 & \Phi_4 \\ \Phi_3 & \Phi_5 & \Phi_3 & \Phi_4 & \Phi_6 & \Phi_4 \\ \Phi_4 & \Phi_4 & \Phi_4 & \Phi_4 & \Phi_4 & \Phi_4 \\ \Phi_5 & \Phi_3 & \Phi_6 & \Phi_4 & \Phi_3 & \Phi_4 \\ \Phi_4 & \Phi_4 & \Phi_4 & \Phi_4 & \Phi_4 & \Phi_4 \end{pmatrix} , \tag{4.99}$$

and

$$\mathbf{\Phi}_{g\bar{q}\to gg\bar{q}} = \begin{pmatrix} \Phi_4 & \Phi_4 & \Phi_4 & \Phi_4 & \Phi_4 & \Phi_4 \\ \Phi_4 & \Phi_4 & \Phi_4 & \Phi_4 & \Phi_4 & \Phi_4 \\ \Phi_4 & \Phi_4 & \Phi_3 & \Phi_3 & \Phi_6 & \Phi_5 \\ \Phi_4 & \Phi_4 & \Phi_3 & \Phi_1 & \Phi_5 & \Phi_2 \\ \Phi_4 & \Phi_4 & \Phi_6 & \Phi_5 & \Phi_3 & \Phi_3 \\ \Phi_4 & \Phi_4 & \Phi_5 & \Phi_2 & \Phi_3 & \Phi_1 \end{pmatrix} , \tag{4.100}$$

with the TMD gluon distributions given in the third row of Table 7. These matrices differ by the permutations of the entries, which has its origin in a slightly different color decomposition for quarks and antiquarks. Namely, the order of quark-antiquark lines (with the outgoing-momenta convention) is reversed in one case with respect to the other.

Finally, the processes with two quark-antiquark pairs, with incoming quark

$$g\,(k_1)\,q\,(k_5) \to q\,(k_2)\,\bar{q}\,(k_3)\,q\,(k_4)\ , \tag{4.101}$$

or antiquark, involve respectively

$$\mathbf{\Phi}_{gq\to q\bar{q}q} = \begin{pmatrix} \Phi_1 & 0 & \Phi_1 & \Phi_1 \\ 0 & \Phi_2 & \Phi_3 & \Phi_1 \\ \Phi_1 & \Phi_3 & \Phi_2 & 0 \\ \Phi_1 & \Phi_1 & 0 & \Phi_1 \end{pmatrix} , \tag{4.102}$$

and

$$\mathbf{\Phi}_{g\bar{q}\to q\bar{q}\bar{q}} = \begin{pmatrix} \Phi_1 & 0 & \Phi_1 & \Phi_1 \\ 0 & \Phi_2 & \Phi_1 & \Phi_3 \\ \Phi_1 & \Phi_1 & \Phi_1 & 0 \\ \Phi_1 & \Phi_3 & 0 & \Phi_2 \end{pmatrix} . \tag{4.103}$$

The TMD gluon distributions appearing in these matrices are listed in the fourth row of Table 7. Interestingly, for this case, not only some of the color factors vanish, but also the corresponding TMDs.

As the potential phenomenological application of the results (in short run) concerns rather the large N_c limit, we present the relevant matrices in this limit in Appendix D.

4.5.4 6 partons

Six parton processes do not involve new features with respect to five partons, except more channels and more involved calculations. The vectors $\vec{\mathcal{A}}$ of the partial amplitudes, and the $\mathbf{M}$ matrices are given in Tables 8 and 9-12 in Appendix C. The $\mathbf{M}$ matrices for processes in which the incoming and outgoing quarks are replaced by the incoming and outgoing antiquarks are the same. Below, we present results for the TMD gluon distribution matrices $\mathbf{\Phi}$ for all channels. The number of partial amplitudes necessitates the use of block matrices to compactify the notation.

For the six-gluon process,

$$g\,(k_1)\,g\,(k_6) \to g\,(k_2)\,g\,(k_3)\,g\,(k_4)\,g\,(k_5)\ , \tag{4.104}$$

the $\mathbf{\Phi}$ matrix is

$$\mathbf{\Phi}_{gg\to gggg} = \begin{pmatrix} T_1 & T_2 & T_3 & T_4 \\ T_2 & T_1 & T_5 & T_6 \\ T_3^\intercal & T_5 & T_1 & T_7 \\ T_4^\intercal & T_6^\intercal & T_7 & T_1 \end{pmatrix} , \tag{4.105}$$

where T_i are 6×6 block matrices given by equations C.1 - C.2 (C.1).

In the present case, we have two TMD operators, for which the color factor vanishes - Φ_4^* and Φ_8^* (we remind, that we mark these matrix elements with an asterix). The full list of the TMD gluon distributions is given in the Table 9.

Next, consider the process

$$g\,(k_1)\,g\,(k_6) \to q\,(k_2)\,\bar{q}\,(k_3)\,g\,(k_4)\,g\,(k_5)\ . \tag{4.106}$$

The TMD matrix reads

$$\mathbf{\Phi}_{gg\to q\bar{q}gg} = \begin{pmatrix} T_1 & T_2 & T_3 & T_4 \\ T_2^\intercal & T_5 & T_6 & T_7 \\ T_3^\intercal & T_6 & T_5 & T_8 \\ T_4^\intercal & T_7^\intercal & T_8^\intercal & T_9 \end{pmatrix} , \tag{4.107}$$

where the blocks gathered in equations C.3 - C.7 (C.7). The TMD gluon distributions are given in the Table 10.

For the process

$$g\,(k_1)\,q\,(k_6) \to g\,(k_2)\,g\,(k_3)\,g\,(k_4)\,q\,(k_5)\ , \tag{4.108}$$

the TMD matrix reads

$$\mathbf{\Phi}_{gq\to ggggq} = \begin{pmatrix} T_1 & T_2 & T_3 & T_4 \\ T_2^{\intercal} & T_5 & T_6 & T_7 \\ T_3^{\intercal} & T_6 & T_5 & T_8 \\ T_4^{\intercal} & T_7^{\intercal} & T_8 & T_5 \end{pmatrix} , \tag{4.109}$$

with the blocks expressed by equations C.8 - C.11 (C.3). The TMD distributions are in Table 11.

Similarly, for the process with the antiquark

$$g\left(k_1\right)\bar{q}\left(k_6\right) \to g\left(k_2\right)g\left(k_3\right)g\left(k_4\right)\bar{q}\left(k_5\right) , \tag{4.110}$$

we get

$$\mathbf{\Phi}_{g\bar{q}\to ggg\bar{q}} = \begin{pmatrix} T_1 & T_1 & T_1 & T_1 \\ T_1 & T_2 & T_3 & T_4 \\ T_1 & T_3 & T_2 & T_5 \\ T_1 & T_4^{\intercal} & T_5 & T_2 \end{pmatrix} , \tag{4.111}$$

with the blocks given by equations C.12 - C.14 (C.4). The TMD distributions are in Table 11.

Processes with two quark-antiquark pairs have smaller number of partial amplitudes. For the process

$$g\left(k_1\right)g\left(k_6\right) \to q\left(k_2\right)\bar{q}\left(k_3\right)q\left(k_4\right)\bar{q}\left(k_5\right) , \tag{4.112}$$

we obtain

$$\mathbf{\Phi}_{gg\to q\bar{q}q\bar{q}} = \begin{pmatrix} T_1 & T_2 \\ T_2 & T_1 \end{pmatrix} , \tag{4.113}$$

with only two blocks:

$$T_1 = \begin{pmatrix} \Phi_1 & 0 & \Phi_2 & \Phi_3 & 0 & \Phi_2 \\ 0 & \Phi_4 & 0 & 0 & \Phi_5 & 0 \\ \Phi_2 & 0 & \Phi_1 & \Phi_2 & 0 & \Phi_3 \\ \Phi_3 & 0 & \Phi_2 & \Phi_1 & 0 & \Phi_2 \\ 0 & \Phi_5 & 0 & 0 & \Phi_4 & 0 \\ \Phi_2 & 0 & \Phi_3 & \Phi_2 & 0 & \Phi_1 \end{pmatrix} , \quad T_2 = \begin{pmatrix} \Phi_1 & \Phi_3 & \Phi_1 & \Phi_3 & \Phi_1 & \Phi_3 \\ \Phi_3 & \Phi_6 & \Phi_1 & \Phi_1 & \Phi_6 & \Phi_3 \\ \Phi_1 & \Phi_1 & \Phi_1 & \Phi_3 & \Phi_3 & \Phi_3 \\ \Phi_3 & \Phi_1 & \Phi_3 & \Phi_1 & \Phi_3 & \Phi_1 \\ \Phi_1 & \Phi_6 & \Phi_3 & \Phi_3 & \Phi_6 & \Phi_1 \\ \Phi_3 & \Phi_3 & \Phi_3 & \Phi_1 & \Phi_1 & \Phi_1 \end{pmatrix} . \tag{4.114}$$

The TMD distributions are given in the Table 12.

For the process

$$g\left(k_1\right)q\left(k_6\right) \to g\left(k_2\right)q\left(k_3\right)\bar{q}\left(k_4\right)q\left(k_5\right) , \tag{4.115}$$

we have

$$\mathbf{\Phi}_{gq\to gq\bar{q}q} = \begin{pmatrix} T_1 & T_2 \\ T_2^{\intercal} & T_3 \end{pmatrix} , \tag{4.116}$$

with three different blocks

$$T_1 = \begin{pmatrix} \Phi_1 & 0 & \Phi_2 & \Phi_3 & 0 & \Phi_2 \\ 0 & \Phi_4 & \Phi_5^* & 0 & \Phi_3 & \Phi_6^* \\ \Phi_2 & \Phi_5^* & \Phi_4 & \Phi_3 & 0 & \Phi_4 \\ \Phi_3 & 0 & \Phi_3 & \Phi_3 & 0 & \Phi_3 \\ 0 & \Phi_3 & 0 & 0 & \Phi_3 & 0 \\ \Phi_2 & \Phi_6^* & \Phi_4 & \Phi_3 & 0 & \Phi_4 \end{pmatrix} , \quad T_2 = \begin{pmatrix} \Phi_1 & \Phi_3 & \Phi_1 & \Phi_3 & \Phi_1 & \Phi_3 \\ \Phi_2 & \Phi_3 & \Phi_1 & \Phi_2 & \Phi_7 & \Phi_3 \\ \Phi_2 & \Phi_3 & \Phi_1 & \Phi_2 & \Phi_2 & \Phi_3 \\ \Phi_3 & \Phi_3 & \Phi_3 & \Phi_3 & \Phi_3 & \Phi_3 \\ \Phi_3 & \Phi_3 & \Phi_3 & \Phi_3 & \Phi_3 & \Phi_3 \\ \Phi_2 & \Phi_3 & \Phi_3 & \Phi_7 & \Phi_2 & \Phi_3 \end{pmatrix} , \tag{4.117}$$

$$T_3 = \begin{pmatrix} \Phi_4 & 0 & \Phi_2 & \Phi_4 & \Phi_8^* & \Phi_3 \\ 0 & \Phi_3 & 0 & 0 & \Phi_3 & 0 \\ \Phi_2 & 0 & \Phi_1 & \Phi_2 & 0 & \Phi_3 \\ \Phi_4 & 0 & \Phi_2 & \Phi_4 & \Phi_9^* & \Phi_3 \\ \Phi_8^* & \Phi_3 & 0 & \Phi_9^* & \Phi_4 & 0 \\ \Phi_3 & 0 & \Phi_3 & \Phi_3 & 0 & \Phi_3 \end{pmatrix} . \tag{4.118}$$

Note, that in this process there appear both the vanishing structures for vanishing color factors and non-vanishing structures for vanishing color factors. The list of the TMD distributions is given in the Table 12. Similarly for the process with an antiquark, we get:

$$\mathbf{\Phi}_{g\bar{q}\to gq\bar{q}\bar{q}} = \begin{pmatrix} T_1 & T_2 \\ T_2 & T_3 \end{pmatrix} , \tag{4.119}$$

with

$$T_1 = \begin{pmatrix} \Phi_3 & 0 & \Phi_3 & \Phi_3 & 0 & \Phi_3 \\ 0 & \Phi_4 & \Phi_6^* & 0 & \Phi_3 & \Phi_5^* \\ \Phi_3 & \Phi_6^* & \Phi_4 & \Phi_2 & 0 & \Phi_4 \\ \Phi_3 & 0 & \Phi_2 & \Phi_1 & 0 & \Phi_2 \\ 0 & \Phi_3 & 0 & 0 & \Phi_3 & 0 \\ \Phi_3 & \Phi_5^* & \Phi_4 & \Phi_2 & 0 & \Phi_4 \end{pmatrix} , \; T_2 = \begin{pmatrix} \Phi_3 & \Phi_3 & \Phi_3 & \Phi_3 & \Phi_3 & \Phi_3 \\ \Phi_3 & \Phi_7 & \Phi_2 & \Phi_1 & \Phi_3 & \Phi_2 \\ \Phi_3 & \Phi_2 & \Phi_7 & \Phi_3 & \Phi_3 & \Phi_2 \\ \Phi_3 & \Phi_1 & \Phi_3 & \Phi_1 & \Phi_3 & \Phi_1 \\ \Phi_3 & \Phi_3 & \Phi_3 & \Phi_3 & \Phi_3 & \Phi_3 \\ \Phi_3 & \Phi_2 & \Phi_2 & \Phi_1 & \Phi_3 & \Phi_2 \end{pmatrix} , \tag{4.120}$$

$$T_3 = \begin{pmatrix} \Phi_3 & 0 & \Phi_3 & \Phi_3 & 0 & \Phi_3 \\ 0 & \Phi_4 & \Phi_9^* & 0 & \Phi_3 & \Phi_8^* \\ \Phi_3 & \Phi_9^* & \Phi_4 & \Phi_2 & 0 & \Phi_4 \\ \Phi_3 & 0 & \Phi_2 & \Phi_1 & 0 & \Phi_2 \\ 0 & \Phi_3 & 0 & 0 & \Phi_3 & 0 \\ \Phi_3 & \Phi_8^* & \Phi_4 & \Phi_2 & 0 & \Phi_4 \end{pmatrix} . \tag{4.121}$$

The large N_c limits of gluon distributions for 6 parton processes were gathered in Tables 14-17 in Appendix D. Additionally, we collect the color factors for all processes in Appendix E.

4.6 Large N_c analysis for arbitrary number of gluons

In this section, we shall utilize the color flow method described in subsection 4.3 to give the large N_c results for a process with n gluons

$$g(k_1)\, g(k_n) \to g(k_2) \dots g(k_{n-1}) . \tag{4.122}$$

We shall use the fact that the color flow decomposition (4.21) involves all $(n-1)!$ partial amplitudes which are the same as in the fundamental decomposition (4.19). Therefore, the leading N_c contribution is given by the partial amplitudes squared (the interference terms are subleading) [171]

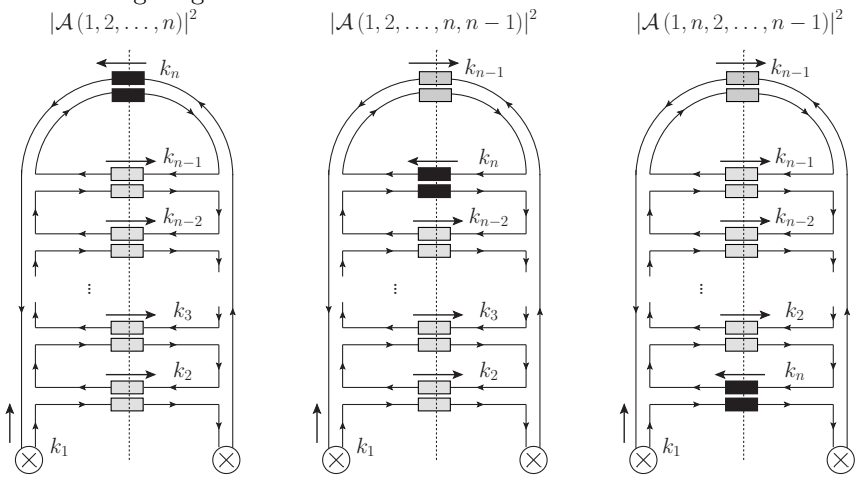
$$|\mathcal{M}|^2 = \mathcal{C} \sum_{\pi \in S_n/Z_n} \left\{ |\mathcal{A}(\pi(1), \dots, \pi(n))|^2 + \mathcal{O}\left(\frac{1}{N_c^2}\right) \right\} , \tag{4.123}$$

with $\mathcal{C}$ being some color coefficient. Note that, if we used the adjoint color decomposition to reduce the number of partial amplitudes, only to the linearly independent ones, as we did in the previous section, we would not be able to claim (4.123). Consequently, the general analysis of large N_c would be very difficult. Therefore, there is a trade off: switching to a general argumentation requires giving up the advantage of using minimal number of amplitudes. In practice, however, any partial amplitude can be easily calculated numerically, so the real loss is not so big.

Based on the above, the idea is to calculate first the diagonal elements of matrix $\mathbf{\Phi}$, as they will definitely

contribute in the large N_c limit. This would be the final answer, if there is no enhancement of powers of N_c^2 for some of the non-diagonal elements. In fact, as we shall see, the enhancement indeed occurs, but still the TMD gluon distribution appearing off the diagonal is numerically small.

Let us start with calculating the diagonal elements of the TMD gluon distribution matrix $\mathbf{\Phi}$. It is sufficient to consider only the following diagrams:



The first diagram from the left corresponds to the partial amplitude squared $|\mathcal{A}(1,2,\ldots,n)|^2$ and the TMD operator reads (after dividing by the corresponding color factor)

$$\frac{N_c^{n-3}}{N_c^{n-2}}\mathrm{Tr}\left\{F(\xi)\mathcal{U}^{[-]\dagger}F(0)\mathcal{U}^{[+]}\right\}\mathrm{Tr}\mathcal{U}^{[\square]\dagger} \rightsquigarrow \mathcal{F}_{gg}^{(1)}\,, \tag{4.124}$$

i.e. it corresponds to the TMD $\mathcal{F}_{gg}^{(1)}$, Eq. (4.35). However, any permutation of the following $(n-2)$ final state legs will give the same contribution, thus, the set

$$\left\{|\mathcal{A}(1,\pi(2),\pi(3),\ldots,\pi(n-1),n)|^2\,\middle|\,\pi\in S_{n-2}\right\} \rightsquigarrow \mathcal{F}_{gg}^{(1)}\,. \tag{4.125}$$

The second diagram, corresponding to $|\mathcal{A}(1,2,\ldots,n,n-1)|^2$, gives

$$\frac{N_c^{n-4}}{N_c^{n-2}}\mathrm{Tr}\left\{F(\xi)\mathcal{U}^{[+]\dagger}F(0)\mathcal{U}^{[+]}\right\}\mathrm{Tr}\mathcal{U}^{[\square]\dagger}\mathrm{Tr}\mathcal{U}^{[\square]} \rightsquigarrow \mathcal{F}_{gg}^{(6)}\,. \tag{4.126}$$

Not only any permutation of final states will give the same result, but also any diagram with leg k_n permuted with $\{3,\ldots,n-2\}$. Thus

$$\begin{aligned}&\left\{|\mathcal{A}(1,\pi(2),\ldots,\pi(n-2),n,\pi(n-1))|^2\,\middle|\,\pi\in S_{n-2}\right\}\\ &\qquad\cup\left\{|\mathcal{A}(1,\pi(2),\ldots,\pi(n-3),n,\pi(n-2),\pi(n-1))|^2\,\middle|\,\pi\in S_{n-2}\right\}\\ &\qquad\cdots\cup\left\{|\mathcal{A}(1,\pi(2),n,\pi(3),\ldots,\pi(n-2),\pi(n-1))|^2\,\middle|\,\pi\in S_{n-2}\right\}\\ &\qquad\qquad\rightsquigarrow \mathcal{F}_{gg}^{(6)}\,.\end{aligned} \tag{4.127}$$

Finally, the third diagram, gives complex conjugate of the operator in (4.124), thus also $\mathcal{F}_{gg}^{(1)}$, because of our assumption of the reality of the correlators. We get therefore

$$\left\{|\mathcal{A}(1,n,\pi(2),\pi(3),\ldots,\pi(n-1))|^2\,\middle|\,\pi\in S_{n-2}\right\} \rightsquigarrow \mathcal{F}_{gg}^{(1)}\,. \tag{4.128}$$

Now let us put together the above results, using the matrix notation as in Section 4.5. Let us define the

partial amplitude vector so that it preserves the block structure emerging above

$$\vec{\mathcal{A}} = \begin{pmatrix} \mathcal{A}(\widehat{1}, 2, \ldots, n-2, n-1, \widehat{n}) \\ \vdots \\ \mathcal{A}(\widehat{1}, 2, 3, \ldots, n-2, \widehat{n}, n-1) \\ \vdots \\ \mathcal{A}(\widehat{1}, 2, 3, \ldots, \widehat{n}, n-2, n-1) \\ \vdots \\ \mathcal{A}(\widehat{1}, 2, \widehat{n}, 3, \ldots, n-2, n-1) \\ \vdots \\ \mathcal{A}(\widehat{1}, \widehat{n}, 3, \ldots, n-2, n-1) \\ \vdots \end{pmatrix}, \tag{4.129}$$

where we have used hats to denote momenta with fixed position in a given group (the actual ordering in each group doesn't matter). For this choice of the vector $\vec{\mathcal{A}}$, the diagonal contribution to the matrix $\mathbf{\Phi}$ at large N_c reads

$$\mathbf{\Phi}_{\text{diag}} = \begin{pmatrix} T_1 & 0 & 0 \\ 0 & T_2 & 0 \\ 0 & 0 & T_1 \end{pmatrix}, \tag{4.130}$$

where

$$T_1 = \mathcal{F}_{gg}^{(1)} \mathbf{1}_{(n-2)!}, \quad T_2 = \mathcal{F}_{gg}^{(6)} \mathbf{1}_{(n-3)(n-2)!}\,. \tag{4.131}$$

For example, for $n = 4$, we have explicitly

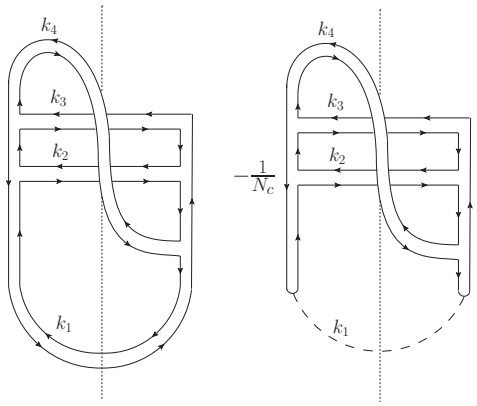
$$\mathbf{\Phi}_{\text{diag}} = \begin{pmatrix} \mathcal{F}_{gg}^{(1)} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathcal{F}_{gg}^{(1)} & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathcal{F}_{gg}^{(6)} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathcal{F}_{gg}^{(6)} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathcal{F}_{gg}^{(1)} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathcal{F}_{gg}^{(1)} \end{pmatrix}. \tag{4.132}$$

Now, let us consider the nondiagonal elements. As said above, these elements will be convoluted with partial amplitudes (interference terms) whose color factors are suppressed by at least $1/N_c^2$ (to say it differently, the non-diagonal elements of the color matrix $\mathbf{C}$ in (4.82), if it is calculated in fundamental color decomposition, are subleading of at least $1/N_c^2$). Therefore, they do not contribute in large N_c, unless some off-diagonal TMD gluon distribution is enhanced by at least N_c^2. This still might not be enough, but is a sign that a careful analysis has to be carried.

The most suspicious non-diagonal elements are those which correspond to the color flow diagrams with the least number of loops. This is slightly counter-intuitive, but we have to keep in mind that, by definition, we divide the color factors out of the TMD (there are no vanishing color factors for gluons in the color flow representation, unlike for the adjoint representation). Thus, the enhancement may happen if the diagrams with Wilson lines have much more loops than the pure color factor diagrams. It is best to illustrate this by an explicit example. Consider a 4-gluon process and the following interference term:

$$\mathcal{A}(1, 2, 3, 4)\, \mathcal{A}^*(1, 4, 2, 3)\,. \tag{4.133}$$

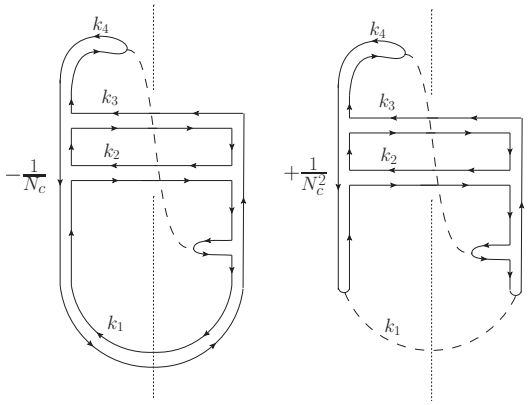
We have the following two leading diagrams for the color factor:



The $U(1)$ colorless propagator for the k_1 leg stems from the projectors that have to be inserted for the final state gluons, *cf.* (4.25). Recall, that in general these color factor diagrams have to be divided by N_A to get the color factor with 'open indices'. The first, Möbius-loop-like diagram, cancels with the second:

$$\frac{1}{N_A}\left(N_c^2 - \frac{1}{N_c}N_c^3\right) = 0\,. \tag{4.134}$$

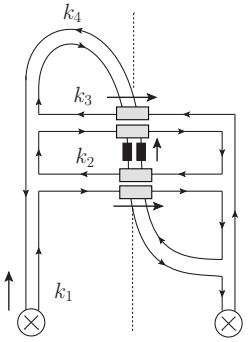
(We called the first diagram 'Möbius-loop-like diagram' because one of the internal loops shares its border with the external loop.) Similar cancellation happens for the diagrams, where the $U(1)$ gluon appears for legs k_2 and k_3. The sub-leading diagrams are those where $U(1)$ colorless gluon is k_4, *i.e.* it crosses the other legs:



In this case, we get

$$\frac{1}{N_A}\left(-\frac{1}{N_c}N_c^3 + \frac{1}{N_c^2}N_c^2\right) = -1\,. \tag{4.135}$$

Now, let us look at the leading diagram for the TMD operator:



It reads

$$N_c \mathrm{Tr}\left\{F\left(\xi\right)\mathcal{U}^{[\square]}\right\}\mathrm{Tr}\left\{F\left(0\right)\mathcal{U}^{[\square]\dagger}\right\} = N_c^2 \mathcal{F}_{gg}^{(2)}\,. \tag{4.136}$$

Dividing by the leading color factor (divided by N_A) we get finally the following non-diagonal element of the $\mathbf{\Phi}$ matrix in the large N_c limit:

$$\left(\mathbf{\Phi}\right)_{15} = -N_c^2 \mathcal{F}_{gg}^{(2)}\,. \tag{4.137}$$

As the color factor for $\mathcal{A}\left(1,2,3,4\right)\mathcal{A}^*\left(1,4,3,2\right)$ is suppressed with respect to diagonal elements by $1/N_c^2$, the TMD gluon distribution $-\mathcal{F}_{gg}^{(2)}$ indeed contributes in the large N_c limit. In a similar manner, but considering much more complicated diagrams with maximal number of crossed lines, one can deduce, that this will be always the case for some non-diagonal elements for any multi gluon process. For example, after a similar but tedious calculation for 5 gluon process, we find that the dominant non-diagonal element is $-N_c^4\mathcal{F}_{gg}^{(2)}/4$. We will always get the $\mathcal{F}_{gg}^{(2)}$ TMD gluon distribution, because of the Möbius-loop-like structure, which gives the two traces appearing in the definition (4.36).

While perhaps it is possible to derive the answer for the non-diagonal leading N_c elements for any n, let us note that $\mathcal{F}_{gg}^{(2)}$ gives numerically rather small contribution to the cross section, compared to the other gluon distributions [172,173]. Indeed, it vanishes very quickly with k_T, so that it is small for transverse momenta around the saturation scale. Moreover, it does not survive the collinear limit. Therefore, in possible phenomenological studies of multigluon production, it is safe to set

$$\mathbf{\Phi}_{gg\to g\ldots g} = \mathbf{\Phi}_{\mathrm{diag}}\,. \tag{4.138}$$

The study of large N_c limit for multiparton processes with quarks, and for gluons without the approximation described above, is left for a separate work.

5 Phenomenology of multi-jet processes

In this section we present the phenomenology of multi-jet processes. We will start however with the description of the initial-state parton shower that is based on unintegrated parton distributions. The implementation of this parton shower completes the framework that is used to perform calculations in High Energy factorization.

5.1 Initial State Parton Shower based on unintegrated parton distributions

The parton shower, which is described here, follows consistently the parton evolution of the unintegrated distributions. By this we mean that the splitting functions P_{ab}, the order in α_s, the scale in the calculation of α_s, as well as the kinematic restrictions applied are identical in both the parton shower and the evolution of the parton densities.

The backward evolution method, as now common in Monte Carlo event generators, is applied for the initial state parton shower, evolving from the large scale of the matrix-element process backwards down to the scale of the incoming hadron. However, in contrast to the conventional parton shower, which generates a transverse momentum of the initial state partons during the backward evolution, the transverse momentum of the initial partons of the hard scattering process is fixed by the unintegrated distribution and the parton shower does not change the kinematics. The transverse momenta during the cascade follow the behavior of the unintegrated PDF. The hard scattering process is obtained directly using the off-shell matrix element calculations The partonic configuration is stored in the form of an LHE (Les Houches Event) text file, but now including the transverse momenta of the incoming partons. This LHE files are input to the shower and hadronization interface of CASCADE [174, 175] (new version `2.4.x`) for the transverse momentum dependent shower where events in HEPMC [176] format are produced. LHE and HEPMC are standardized interfaces for communication between matrix element generators and event generators.

The backward evolution of the initial state parton shower follows very closely the description in [174,175, 177,178]. The evolution scale μ is selected from the hard scattering process, with $\mu^2 = \hat{p}_T^2$ or $\mu^2 = Q_T^2 + \hat{s}$ for an evolution in virtuality or angular ordering, with $\hat{p}_T$ being the transverse momentum of the hard process, Q_T being the vectorial sum of the initial state transverse momenta and s being the invariant mass of the subprocess.

Starting with the hard scale $\mu = \mu_i$, the parton shower algorithm searches for the next scale μ_{i-1} at which a resolvable branching occurs. This scale μ_{i-1} is selected from the Sudakov form factor Δ_S making use of the unintegrated densities $\mathcal{F}_a(x', k_t', \mu')$ which depend on the longitudinal momentum fraction $x' = xz$ of parton a, its transverse momentum k_T' probed at a scale μ' (see also [174]). The Sudakov form factor Δ_S for the backward evolution is given by (see Fig. 21 left):

$$\Delta_S(x, \mu_i, \mu_{i-1}) = \exp\left[- \int_{\mu_{i-1}}^{\mu_i} \frac{d\mu'}{\mu'} \frac{\alpha_s(\tilde{\mu}')}{2\pi} \sum_a \int dz P_{a\to bc}(z) \frac{x' \mathcal{F}_a(x', k_T', \mu')}{x \mathcal{F}_b(x, k_T, \mu')} \right], \tag{5.1}$$

which describes the probability that the parton b remains at x with the transverse momentum k_T when evolving from μ_i to $\mu_{i-1} < \mu$. Please note that the argument in α_s is $\tilde{\mu}'$, and depends on the ordering condition as discussed later.[5]

In the parton shower language, the selection of the next branching comes from solving the Sudakov form factor Eq. (5.1) for μ_{i-1}. However, solving the integrals in Eq. (5.1) numerically for every branching would be very time consuming. Instead, the veto algorithm [177,179] is applied. Veto algorithm allows to generate a probability distribution $F(\mu)$ using simpler over-estimating function $G(\mu)$. After the value of μ is generated, according to $G(\mu)$, with $\mu < \mu_{\max}$, it is either kept, with probability $\frac{dF(\mu)/d\mu}{dG(\mu)/d\mu}$, or a new value is generated,

[5] In Eq. (5.1) ordering in μ is assumed, if angular ordering, as in CCFM, is applied then the ratio of parton densities would change to $\frac{x' \mathcal{F}_a(x', k_T', \mu'/z)}{x \mathcal{F}_b(x, k_T, \mu')}$ as discussed in [174].

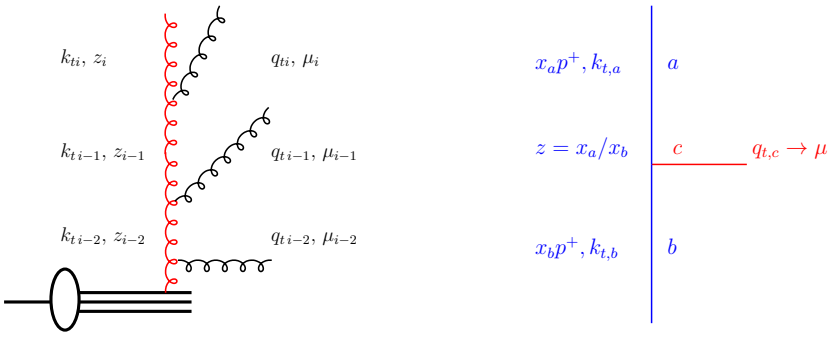


Figure 21: Left: Schematic view of a parton branching process. Right: Branching process $b \to a + c$.

but this time with $\mu_{\max}$ set to be equal to the previously generated value. The selection of μ_{i-1} and the branching splitting z_{i-1} follows the standard methods [177].

The splitting function P_{ab}, as well as the argument $\tilde{\mu}$ in the calculation of α_s is chosen exactly as used in the evolution of the parton density. In a parton shower one treats "resolvable" branchings, defined via a cut in $z < z_M$ in the splitting function (see Eq. (3.3)) to avoid the singular behavior of the terms $\frac{1}{1-z}$, and branchings with $z > z_M$ are regarded as "non-resolvable" and are treated similarly as virtual corrections: they are included in the Sudakov form factor Δ_S.

The longitudinal momentum fraction $x_{i-1} = \frac{x_i}{z_{i-1}}$ is calculated by generating z_{i-1} according to the splitting function. With z_{i-1} and μ_{i-1} all variables needed for a collinear parton shower are obtained.

The calculation of the transverse momentum k_T is sketched in Fig. 21 (right). The transverse momentum $q_{T\,i}$ can be obtained by giving a physical interpretation to the evolution scale μ_i (see Fig. 21 right), and $q_{T\,i}$ can be calculated in case of angular ordering (μ is associated with the angle of the emission) in terms of the angle Θ of the emitted parton wrt the beam directions $q_{T,c} = (1-z)E_b \sin\Theta$:

$$\mathbf{q}_{T,i}^2 = (1-z)^2 \mu_i^2 \,. \tag{5.2}$$

Once the transverse momentum of the emitted parton q_T is known, the transverse momentum of the propagating parton can be calculated from

$$\mathbf{k}_{T\,i-1} = \mathbf{k}_{T\,i} + \mathbf{q}_{T\,i-1} \tag{5.3}$$

with a uniformly distributed azimuthal angle ϕ is assumed for the vector components of $\mathbf{k}$ and $\mathbf{q}$.

The whole procedure is iterated until one reaches a scale $\mu_{i-1} < q_0$ with q_0 being a cut-off parameter, which can be chosen to be the starting evolution scale of the unintegrated distribution. However, it turns out that during the backward evolution the transverse momentum k_T can reach large values, even for small scales μ_{i-1}, because of the random ϕ distribution. On average the transverse momentum decreases, and it is of advantage to continue the parton shower evolution to a scale $q_0 \sim \Lambda_{\rm QCD} \sim 0.3$ GeV, to allow enough emissions to share the transverse momenta generated.

5.2 Single jet production

The single inclusive jet production is a process which can directly probe partonic content of the proton without a need for large corrections from fragmentation functions. What makes it interesting is the possibility to apply the appropriate formula already at the leading order in high energy factorization. This is to be contrasted with collinear factorization, where the $2 \to 1$ emission vertex vanishes identically and one has to account for higher order corrections, either at fixed order of α_s, or with a parton shower.

The single inclusive jet production process can be schematically written as

$$A + B \mapsto a + b \to \text{jet} + X \,, \tag{5.4}$$

where A and B are the colliding hadrons, each of which provides a parton, respectively a and b. X corresponds to undetected, real radiation and the beam remnants from the hadrons A and B are understood in the above equation.

The longitudinal kinematic variables read

$$x_1 = \frac{1}{\sqrt{s}}\, p_{T,\text{jet}}\, e^{y_{\text{jet}}} \,, \qquad\qquad x_2 = \frac{1}{\sqrt{s}}\, p_{T,\text{jet}}\, e^{-y_{\text{jet}}} \,, \tag{5.5}$$

where $s = (p_A + p_B)^2$ is the total squared energy of the colliding hadrons, while y_{jet} and p_{jet} are the rapidity and transverse momentum of the leading final state jet, respectively.

The hybrid, high energy factorization formula for the process (5.4), justified for configurations with $x_1 \gg x_2$, at the leading $\ln(1/x)$ accuracy[6], reads [182]

$$\frac{d\sigma}{dy_{\text{jet}} dp_{T,\text{jet}}} = \frac{1}{2}\frac{\pi\, p_{T,\text{jet}}}{(x_1 x_2 s)^2} \sum_{a,b,c} \overline{|\mathcal{M}_{ab^*\to c}|}^2 x_1 f_{a/A}(x_1,\mu^2)\, \mathcal{F}_{b/B}(x_2, p_{T,\text{jet}}^2, \mu^2) \,, \tag{5.6}$$

where $\mathcal{F}$ is a generic notation for the unintegrated parton density, which depends in general also on the hard scale μ. The factorization formula neglects contribution of gauge links, discussed in the previous sections. It can be justified in the region where saturation effects are rather small. The matrix elements $|\mathcal{M}_{ab^*\to c}|^2$ can be obtained via application of the helicity-based formalism [104, 105, 183] for off-shell partons or the *parton reggezation approach* [184]. The following channels contribute to the single jet production in HEF approach

$$g g^* \to g \,, \qquad q g^* \to q \,, \qquad g q^* \to q \,, \qquad \bar{q} q^* \to g \,. \tag{5.7}$$

The explicit expressions for the corresponding matrix elements are collected in Appendix F.

We now turn to predictions for the transverse momentum spectra of the single inclusive forward jets at the LHC. The calculations were performed at the center-of-mass energies of $\sqrt{s} = 7$ and 13 TeV. The event selection was applied by requiring a leading jet with $p_{t,\text{jet}} > 35\,\text{GeV}$ in the rapidity window of $3.2 < |y_{\text{jet}}| < 4.7$, following the cuts used in the CMS analyses of Refs. [185, 186]. For the on-shell partons we used the distribution from the CT10 NLO set [28]. For the off-shell partons, $\mathcal{F}_{b/B}(x_2, p_{T,\text{jet}}^2, \mu^2)$, we chose the following set of distributions, described in Section 3:

- KS nonlinear,
- KS linear,
- KShardscale nonlinear,
- KShardscale linear,
- DLC2016, which is the KMRW method with the double logarithm approximation applied for Sudakov form factor [187].

All HEF predictions in this and in the following section were obtained with the FORWARD program [188]. The code implements the hybrid high energy factorization for the single and double jet production and it is capable of using both gluon and quark off-shell parton distributions.

As we see from the above list, all parametrizations, except that of DLC2016, provide only off-shell gluons and neglect off-shell quarks by assuming that their relative contribution is much smaller. The DLC2016 distributions provide the full set of partons and this gives us unique opportunity to verify this assumption.

[6] At the NLO level, there exists an extension of this formula for single particle production [180, 181].

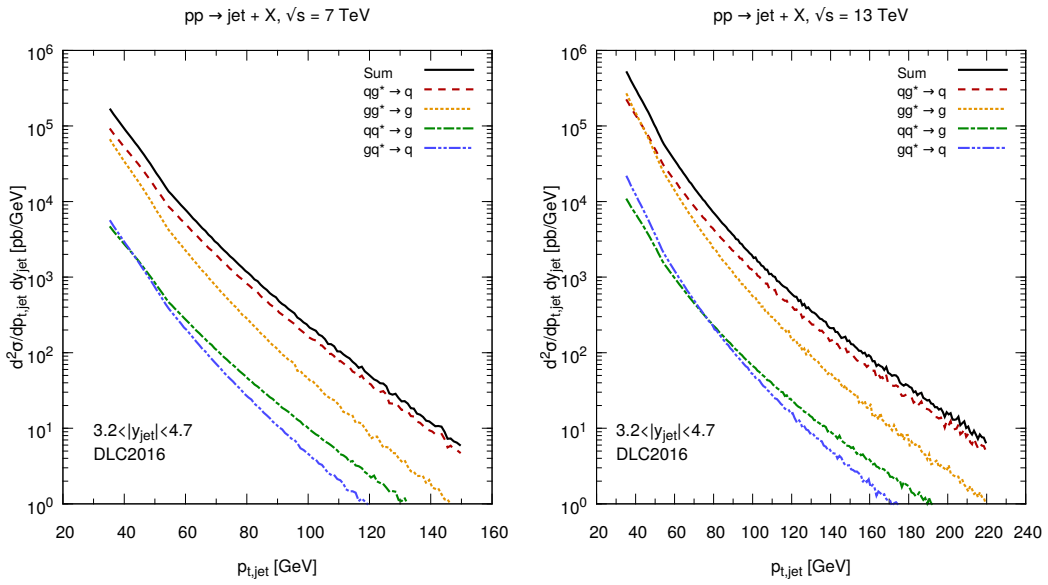


Figure 22: Single inclusive forward jet production. Comparison between different channels contributing to the spectrum of jet's transverse momentum. The results use DLC2016 [187] parametrization for the off-shell partons.

In Fig. 22 we show predictions for various contributions to the single inclusive jet production obtained using Eq. (5.6) with the DLC2016 off-shell partons. It is evident that the off-shell quark contribution can be indeed effectively neglected and we can proceed just with the off-shell gluons in the initial state. The second interesting observation is that the $qg^* \to q$ channel gives larger contribution than $gg^* \to g$ at high transverse momentum while the two channels contribute comparably at lower $p_{T,\mathrm{jet}}$.

As the off-shell quark contributions are negligible, it is justified to use all of the off-shell gluon sets listed above as an input for predictions of single inclusive jet spectra. The corresponding results are shown in Fig. 23, where the upper panel shows the absolute distributions, whereas the two lower panels show theory-to-data ratio. We observe good compatibility of the predictions and the 7 and 13-TeV CMS data [185, 186] across a range of unintegrated gluon distributions. We believe that this is a consequence of the TMD factorization applied to low-x physics [115], which states that the same gluon density (if saturation effects are negligible in the considered phase space region) is to be used for the F_2 structure function and for the single inclusive gluon production.

5.2.1 Saturation

In this section we present calculations of the differential cross section for the production of single inclusive jets as a function of transverse momentum and energy within the rapidity interval $5.2 < y < 6.6$ in proton-lead collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The chosen rapidity range corresponds to the acceptance of the CASTOR calorimeter, installed only on one side of the nominal interaction point ($-6.6 < \eta < -5.2$) of the CMS experiment [189], which has collected proton-proton and proton-lead collision data at the LHC at various center-of-mass energies. In our setup, the phase space is defined such that in proton-lead collisions the proton hosting the high-x initial state parton moves towards the negative rapidity hemisphere, where the CASTOR detector resides. This allows to probe the structure of the lead ion, which moves away from CASTOR, and hosts a very low-x parton. All results are presented in the center-of-mass frame and need to be boosted to the laboratory frame for comparisons with data measurements.

All samples used in this section are obtained with the KATIE Monte Carlo event generator, which can

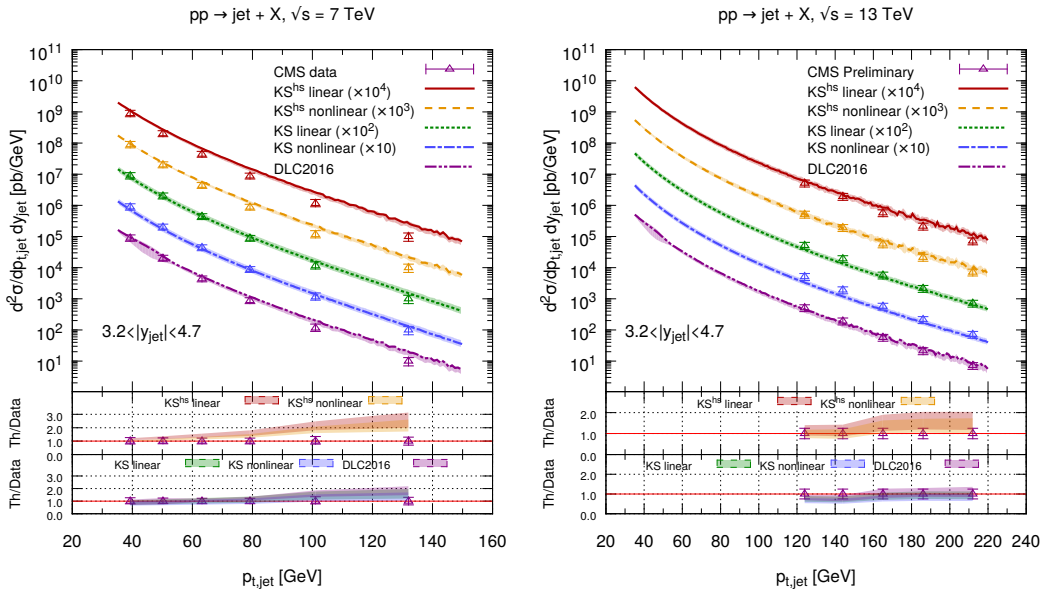


Figure 23: Single inclusive forward jet production. Comparison of predictions for the transverse momentum distributions of a jet with CMS data at 7 [185] and 13 TeV [186]. The bands correspond to different gluon distributions used for calculations. The width of the bands comes from varying the factorization and renormalization scales by factors $\frac{1}{2}$ and 2 around the central value equal to $\mu_F = \mu_R = p_{T,\mathrm{jet}}$. In case of KShardscale for 13 TeV, because of limited size of the available grid, we could vary the scale only by the factors $\frac{1}{2}$ and $\frac{3}{2}$. For better visibility, data and predictions with various unintegrated gluons were multiplied by factors 10^n, with $n = 0, \ldots, 4$.

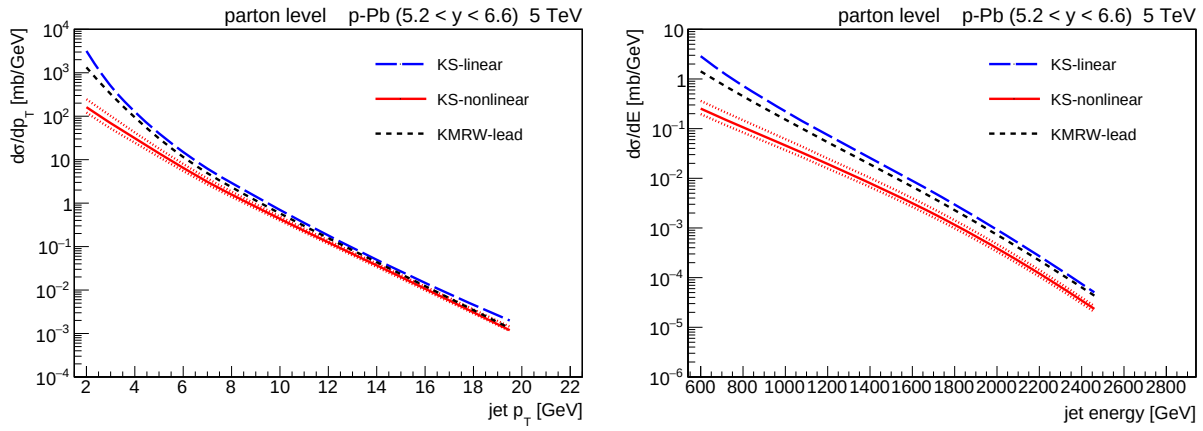


Figure 24: Parton level predictions of KaTie for KS-linear, KS-nonlinear and KMRW-lead gluon densities. Differential jet cross sections as a function of jet p_{T} (left) and energy (right) are presented for proton-lead interactions. The fine dotted lines represent the upper and lower uncertainty in the nonlinear suppression factor.

produce tree-level matrix element calculations in HEF [190]. In this particular case, $gg^* \to g$ and $qg^* \to q$ processes at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV are generated with the initial gluon being off-shell. The minimum p_T of the final-state jet is set to 1.4 GeV. The renormalization and factorization scale is set to be the transverse momentum of the final state jet. Finally, the aforementioned CT10 NLO is used as collinear PDF for the on-shell parton, and the specific KS-linear, KS-nonlinear and KMRW-lead unintegrated gluon PDFs for the off-shell parton. The resulting samples are analyzed with Rivet [191] where the jets are clustered with FastJet [192] using the anti-k_{T} algorithm with distance parameter equal to 0.5 [70]. To obtain the hadron level samples, the output from KaTie is fed to the Cascade 2.4.13 Monte Carlo event generator [174] that employs an adopted version of the Lund string fragmentation model (as used in Pythia) to account for hadronization.

Transverse momentum and energy spectra The observable that conveniently reveals the partonic dynamics in which we are interested is the inclusive jet transverse momentum spectrum: in particular, when the jet is produced in the forward direction, its p_{T} has to be sufficiently small for it to potentially carry information about the saturation phenomenon. In Fig. 24 (left) we plot the parton level cross section as a function of p_{T}, while in Fig. 24 (right) it is presented as a function of the energy of the jet. Predictions obtained using KaTie with KS-linear, KS-nonlinear and KMRW-lead parton densities are compared. The result with the KS-nonlinear PDF is obtained when the nonlinear suppression term is multiplied by a parameter equal to 0.75. Since this is a free parameter that needs input from data, additional scenarios, where the nonlinear suppression term is multiplied by 0.5 and 1.0 are included as fine dotted curves in Fig. 24 (and all other following figures in this subsection). Although this choice is arbitrary, it does indicate the sensitivity of the observable if the nonlinear suppression is altered by $\pm 33\%$. We see that, as expected, at lower p_{T} and lower energy the saturation effects can be substantial, since the KS-nonlinear result is suppressed as compared to KS-linear. In addition we see that KMRW-lead, which follows from DGLAP evolution equations and does not account for nonlinear effects during the evolution but accounts for nuclear shadowing [193], is very close to KS-linear. In Fig. 25, we add hadronization effects by combining the output of KaTie with Cascade (following methods developed in [194]), which leads to a significant decrease of the overall cross section and slightly changes the shape of the distributions as well. Nevertheless, a clear difference in the predictions remains at lower p_{T} and energy values, which makes the observable suitable for comparisons with measurements from data corrected to particle level.

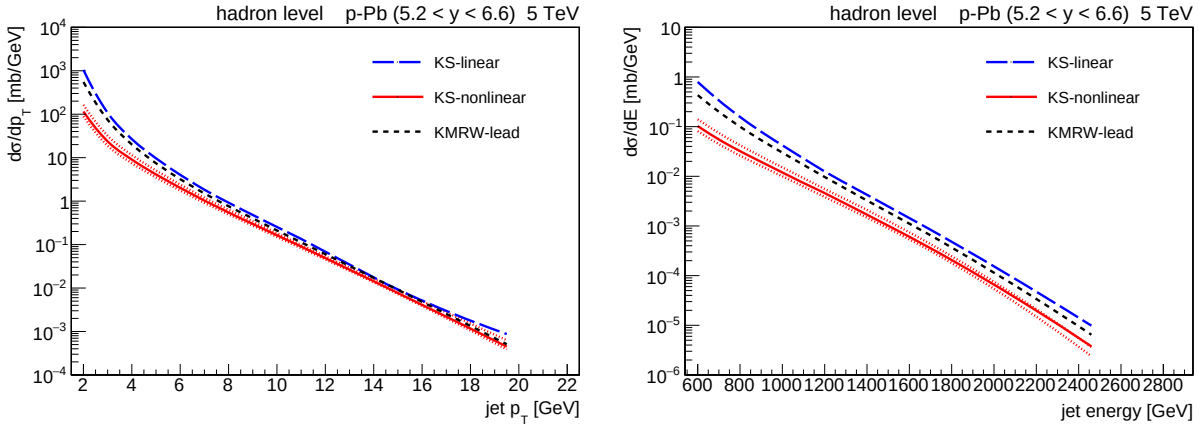


Figure 25: Hadron level predictions of KaTie+Cascade for KS-linear, KS-nonlinear and KMRW-lead gluon densities. Differential jet cross sections as a function of jet p_{T} (left) and energy (right) are presented for proton-lead interactions. The fine dotted lines represent the upper and lower uncertainty in the nonlinear suppression factor.

5.2.2 Nuclear modification ratio

In order to quantify the strength of nonlinearities as one goes from proton-proton to proton-lead, it is convenient to calculate a quantity called the nuclear modification ratio R_{pA}. For a generic observable $\mathcal{O}$, it is defined as

$$R_{pA} = \frac{\sigma_{pPb}(\mathcal{O})}{A\,\sigma_{pp}(\mathcal{O})}\,. \tag{5.8}$$

On one hand, in absence of any effects like saturation in the low-x evolution equations, the ratio would be just consistent with unity (since one considers gluon dominated observables). On the other hand, if saturation effects are present, they would be visible to us in the nuclear modification ratio as a deviation from unity within some range, for instance in the transverse momentum spectra of the measured jets. In the low-x approach, the suppression is directly linked to a denser partonic system (and therefore larger contribution of the nonlinear term) as one goes from proton to lead (since, roughly, $Q_s \sim A^{1/3}$). In the DGLAP approach the possible suppression is due to shadowing, which is implemented by fitting parton densities to data without accounting for any additional dynamical effects that may happen when going from a proton to a nucleus[7].

In Figs. 26 and 27 we compare parton and hadron level predictions of the nuclear modification ratios as obtained using KS-nonlinear and KMRW-lead parton densities. It shows a significant suppression for KS-nonlinear at low values of p_{T}, which indicates that the saturation of the gluon density in lead is large compared to the saturation in the proton. At high values of p_{T} the results obtained with KS-nonlinear and KMRW-lead converge, which shows that nonlinear suppression is negligible in this region. Both do not converge to unity however, which can imply that other suppression effects coming from the nuclei are present. The results obtained using the KMRW-lead gluon density exhibits a different, more constant, behavior as it does not include saturation effects. However, the nuclear shadowing can be substantial, even at $p_{\mathrm{T}} = 10$ GeV, since at the considered energies the nuclear PDF is probed at $x = 10^{-5}$.

The nuclear modification ratio as a function of the jet energy shows a similar behavior, although the effects are somewhat washed out, leading to an overall different slope and normalization of the KS-nonlinear and KMRW-lead predictions. Looking at the fine dotted curves in Figs. 26 and 27, it is seen that the uncertainty due to saturation is large, indicating a high sensitivity of these observables to saturation effects. Even though

[7] In general, in absence of any nuclear effects and saturation, the small deviation from unity is due to the difference between proton and neutron PDFs that contribute to a nuclear PDF.

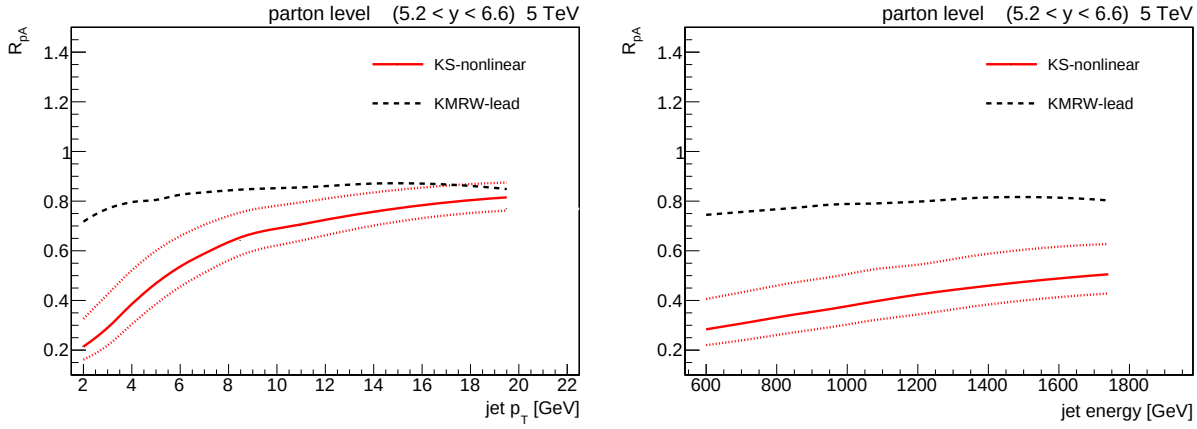


Figure 26: Parton level predictions of the nuclear modification ratio R_{pA} as function of the jet p_{T} (left) and jet energy (right). The fine dotted lines represent the upper and lower uncertainty in the nonlinear suppression factor.

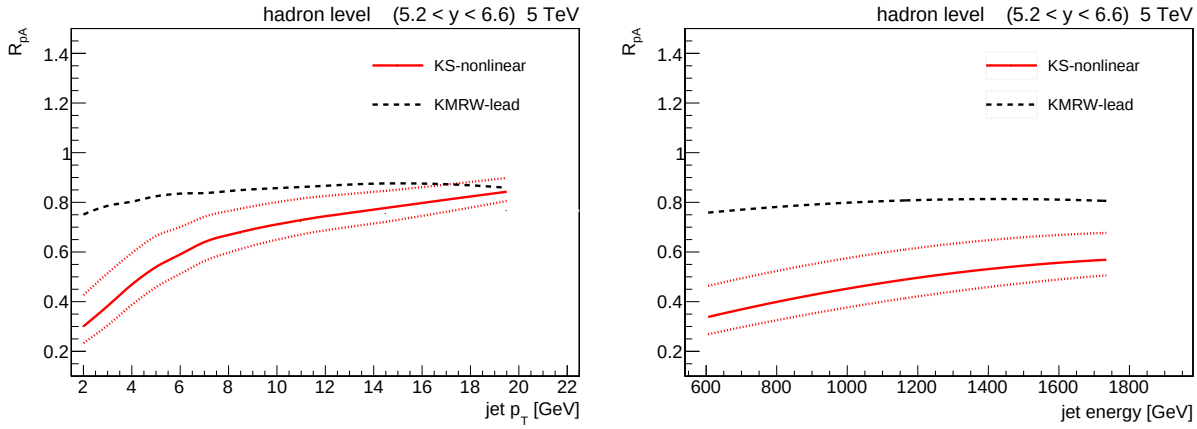


Figure 27: Hadron level predictions of the nuclear modification ratio R_{pA} as function of the jet p_{T} (left) and jet energy (right). The fine dotted lines represent the upper and lower uncertainty in the nonlinear suppression factor.

the variations can be significant, there is always a clear difference with respect to KMRW-lead, which does not include such effects. Therefore a measurement of the nuclear modification factor of forward, low p_{T} jets is ideal to disentangle the linear effects from the nonlinear and to constrain the amount of suppression in the cross section.

In order to better understand this suppression of the cross section, we also plot the ratio of the unintegrated gluon densities for lead ions (UGDPb) and protons (UGDp) as a function of $k_{\perp}^2$ in Fig. 28 (left), evaluated at different values of x. We see that at larger values, $x = 10^{-3}$, the suppression is much reduced and the ratio converges quickly to unity. At very low values, $x = 10^{-5}$, we see a similar behavior as reported before. This is consistent, since the production of low p_{T} jets within $5.2 < y < 6.6$ reaches values of x as low as 10^{-6}.

In addition, Fig. 28 (right) shows parton level predictions of R_{pA} for jets within the default rapidity range, $5.2 < y < 6.6$, and for jets that are produced more centrally in $4.0 < y < 5.0$, for both KS-nonlinear and KMRW-lead parton densities. This confirms the dependence of the nonlinear behavior on the rapidity regions probed in a collision, and shows that parton densities that do not incorporate nonlinear dynamics, such as KMRW-lead, are less sensitive to rapidity selection.

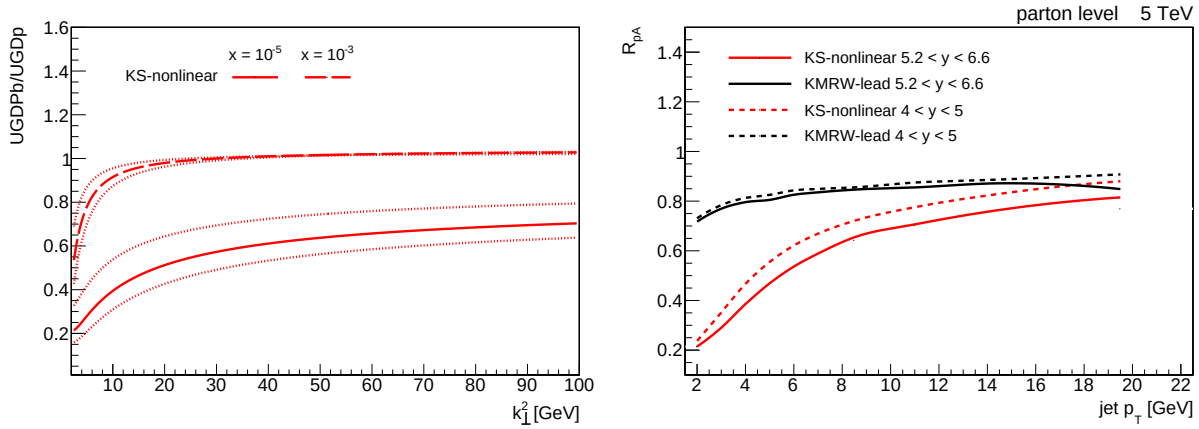


Figure 28: Ratio of KS-nonlinear unintegrated gluon densities for lead and proton evaluated at different values of x. The fine dotted lines represent the upper and lower uncertainty in the nonlinear suppression factor (left). Parton level comparison of the nuclear modification ratio R_{pA} for jets within $5.2 < y < 6.6$ and, more centrally, in $4.0 < y < 5.0$ for both KS-nonlinear and KMRW-lead (right).

5.3 High Energy Factorization for dijets

5.3.1 SPS and DPS contributions

Dijets can be produced either in a *single-parton scattering* (SPS)

$$A + B \mapsto a + b \to \text{jet}_1 + \text{jet}_2 + X\,, \tag{5.9}$$

where the partons a and b interact through a $2 \to 2$ process, or in a *double parton scattering* (DPS)

$$A + B \mapsto a_1 + b_1 + a_2 + b_2 \to \text{jet}_1 + \text{jet}_2 + X\,, \tag{5.10}$$

in which each of the incoming hadrons provides two-parton pairs $a_i + b_i$, which in turn undergo two $2 \to 1$ scatterings. The DPS can be thought of as the single inclusive jet production process of Eq. (5.4) squared.

5.3.2 Results within HEF formalism

In general, in order to comply with the state of the art of theoretical development, description of the SPS process needs corrections from the improved TMD factorization [82], as it gets contribution from the so-called quadrupole configurations of color glass condensate (CGC) states and the latter are important in the non-linear domain.

In the study here, however, we focus on the region of azimuthal distance between the two leading jets, $\Delta\phi$, where the bulk of linear and nonlinear KS densities [111] give comparable results [113] and our aim is just to quantify the potential corrections coming from other physical effects like DPS contributions and final-state parton shower. Encouraged by the good description of the single inclusive jet production, presented in Section 5.2, we aim at evaluation of the DPS contribution to inclusive dijet production in order to assess its relative impact with respect to SPS.

The formula for the SPS contribution to the forward dijet cross section reads [110, 111]

$$\frac{d\sigma_{\text{SPS}}}{dy_1 dy_2 dp_{T\,1} dp_{T\,2} d\Delta\phi} = \frac{p_{T\,1} p_{T\,2}}{8\pi^2 (x_1 x_2 s)^2} \sum_{a,c,d} x_1 f_{a/A}(x_1, \mu^2)\, |\overline{\mathcal{M}_{ag^* \to cd}}|^2 \mathcal{F}_{g/B}(x_2, k_T^2) \frac{1}{1 + \delta_{cd}}\,, \tag{5.11}$$

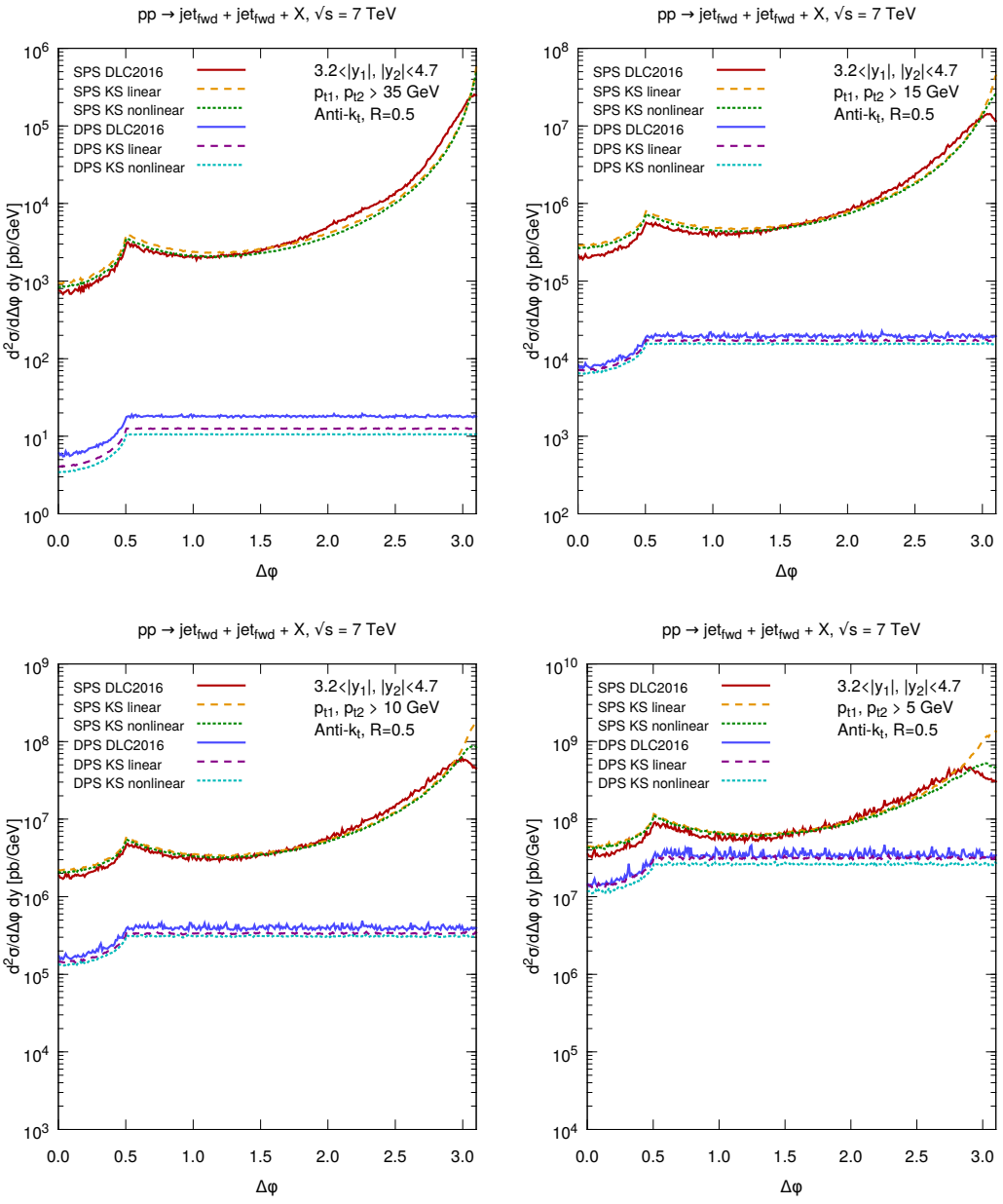


Figure 29: SPS and DPS contribution to forward dijet production for various cuts on transverse momentum of a jet at the LHC at $\sqrt{s} = 7$ TeV.

where

$$x_1 = \frac{1}{\sqrt{s}} \left(|p_{T\,1}| e^{y_1} + |p_{T\,2}| e^{y_2} \right) , \qquad x_2 = \frac{1}{\sqrt{s}} \left(|p_{T\,1}| e^{-y_1} + |p_{T\,2}| e^{-y_2} \right) , \tag{5.12}$$

are the fractions of incoming particles' momenta carried by the partons participating in the hard interaction and

$$k_T^2 = |\mathbf{p_{T\,1}} + \mathbf{p_{T\,2}}|^2 = p_{T\,1}^2 + p_{T\,2}^2 + 2 p_{T\,1} p_{T\,2} \cos \Delta\phi \tag{5.13}$$

is an imbalance of the transverse momentum of the two leading jets, which, in the HEF formalism is equal to the off-shellness of the incoming gluon. The two leading jets are separated in the transverse plane by the angle $\Delta\phi$.

The expressions for the matrix elements can be found in Refs. [110,111], while the parton densities at the approximation we are working with are of the same kind as for the single inclusive jet production discussed in subsection 5.2. Our aim is now to identify and quantify potential corrections to the HEF framework encapsulated in Eq. (5.2).

One of the above may come from double parton scattering [187, 195–200]. In general, the cross section for DPS involves parton density functions which take into account correlations of partons inside the hadrons before the hard scattering [201–204]. However, recent study of Ref. [205] shows that a factorized assumption for DPS is largely valid at high scales ($Q^2 > 10^2$ GeV^2). Following this observation, we can therefore write

$$\frac{d\sigma_{\rm DPS}}{dy_1 d^2p_{T\,1} dy_2 d^2p_{T\,2}} = \frac{1}{\sigma_{\rm effective}} \frac{d\sigma}{dy_1 d^2p_{T\,1}} \frac{d\sigma}{dy_2 d^2p_{T\,2}} , \tag{5.14}$$

where $\sigma_{\rm effective} = 15\,$mb, based on the recent measurement of the LHCb [206, 207] collaboration, which confirmed previous results of D0 [208] and CDF [209]. The explicit expression for the DPS contribution in the factorized approximation reads

$$\begin{aligned}
\frac{d\sigma_{\rm DPS}}{dy_1 dy_2 dp_{T\,1} dp_{T\,2} d\Delta\phi} = {} & \frac{1}{\sigma_{\rm eff}} \frac{\pi}{8} \frac{p_{T\,1}}{(x_1 x_2 s)^2} \frac{p_{T\,2}}{(\bar{x}_1 \bar{x}_2 s)^2} \\
& \times \left(\overline{|\mathcal{M}_{gg^* \to g}|}^2 x_1 f_{g/A}(x_1) + \sum_{i=1}^{n_f} \overline{|\mathcal{M}_{qg^* \to q}|}^2 x_1 f_{q(i)/A}(x_1) \right) \\
& \times \left(\overline{|\bar{\mathcal{M}}_{gg^* \to g}|}^2 \bar{x}_1 \bar{f}_{g/A}(\bar{x}_1) + \sum_{i=1}^{n_f} \overline{|\bar{\mathcal{M}}_{qg^* \to q}|}^2 \bar{x}_1 \bar{f}_{q(i)/A}(\bar{x}_1) \right) \\
& \times \mathcal{F}_{g^*/B}(x_2, p_{T\,1}^2) \bar{\mathcal{F}}_{g^*/B}(\bar{x}_2, p_{T\,2}^2) \theta(1 - x_1 - \bar{x}_1) \theta(1 - x_2 - \bar{x}_2) ,
\end{aligned} \tag{5.15}$$

where, in order to be compatible with the SPS formula (5.11), we introduced an auxiliary azimuthal angle between the final state jets, $\Delta\phi$. The notation follows that of Eq. (5.11) except that now, each of the incoming particles provides a pair of partons, whose energy fractions are given by x_1 and $\bar{x}_1$ for hadron A, and x_2 and $\bar{x}_2$ for hadron B. The theta functions guarantee that a pair of partons from a single hadron does not carry more than 100% of the hadron's energy.

The DPS contributions are in general expected to be strong in the low-p_T region of phase space. In order to quantify the role of DPS in forward-forward dijet production, we have calculated the DPS contribution to the azimuthal-angle dependence. Of course, we expect that, in the approximation of Eq. (5.14), where the correlations between incoming partons from different pairs are neglected, the contribution will be just of a pedestal type, thus only changing the overall normalization.

In Fig. 29 we show the SPS and the DPS contributions to the azimuthal angle distributions for various cuts on the hardest jet's transverse momentum, set respectively at 35, 15, 10 and 5 GeV. We see that the relative contribution of DPS increases with lowering the transverse momentum jet cut, but it is always significantly smaller than SPS at the experimentally relevant value of 35 GeV. We have checked that the picture looks very similar at 13 TeV.

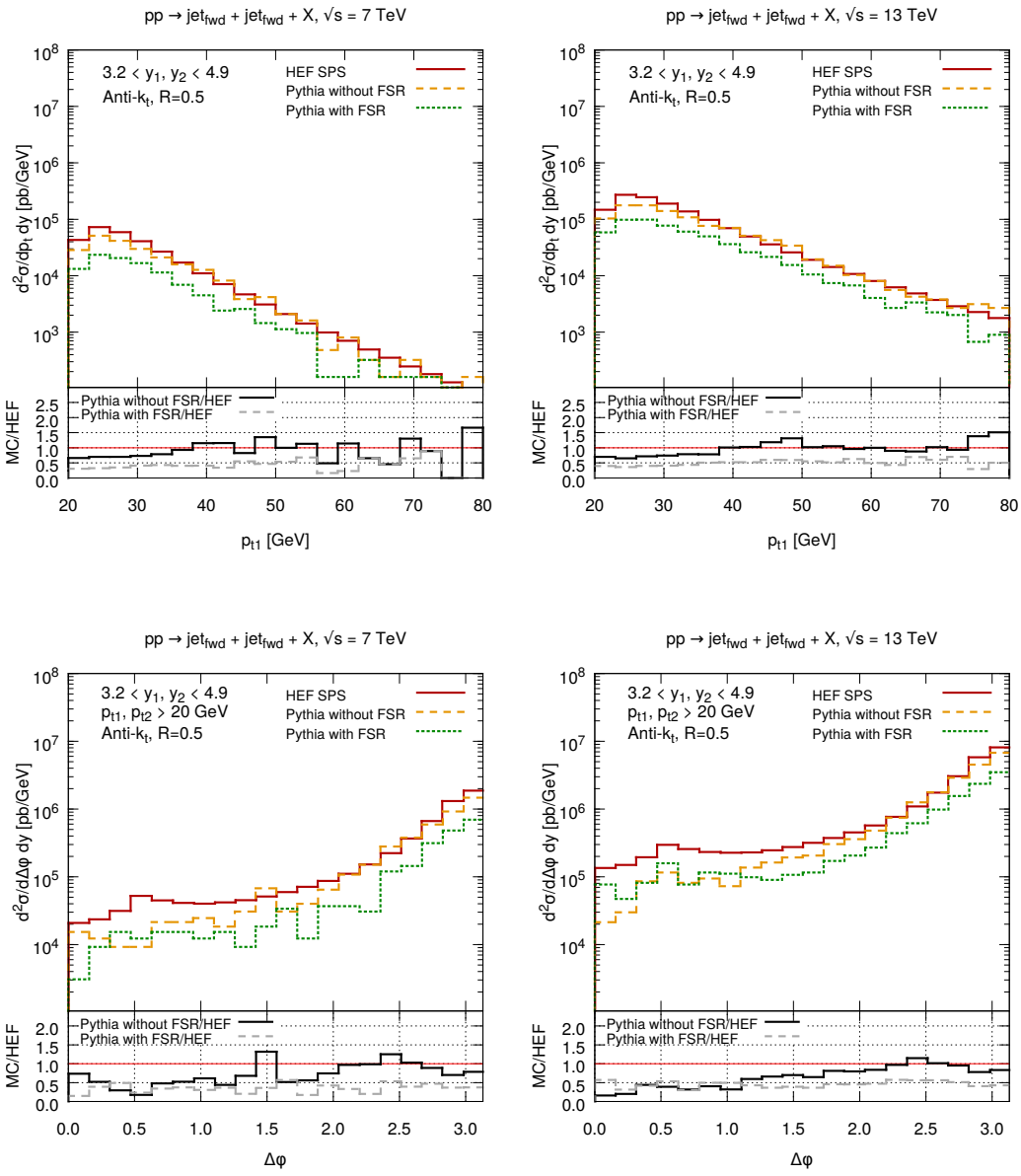


Figure 30: Transverse momentum distribution and azimuthal decorrelation in forward dijet production. Comparison of predictions from high energy factorization, Eq. (5.11) with DLC2016 unintegrated gluon [187], and PYTHIA MC generator. We checked that including MPI has negligible effect on these distributions.

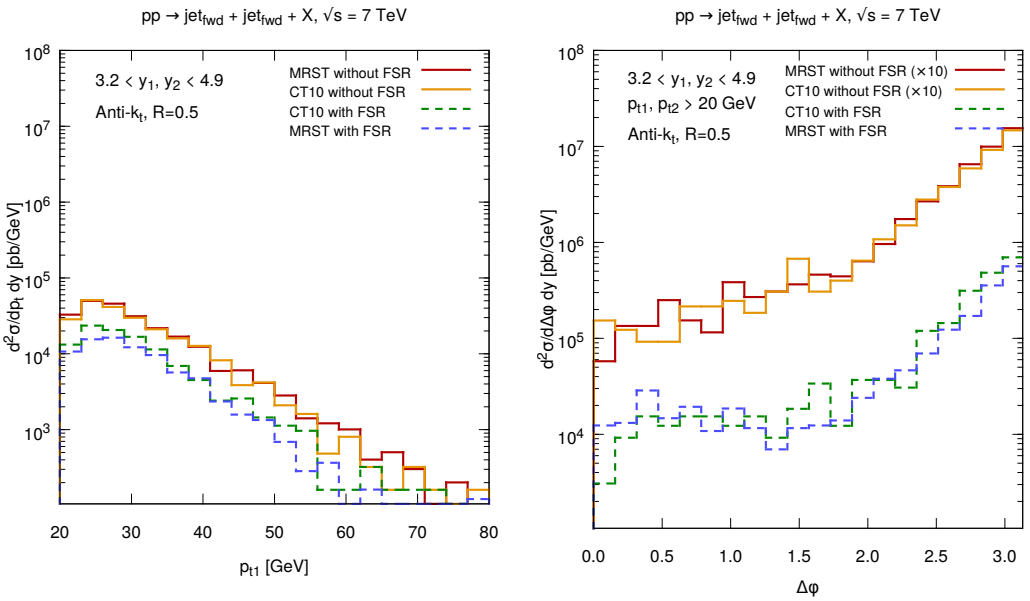


Figure 31: Comparison of results for forward dijet spectra obtained with different PDFs used in PYTHIA MC generator: MRST NLO [210] and CT10 NLO [28].

5.3.3 HEF vs. collinear factorization

We now turn to the comparison of our HEF predictions for the forward dijet production with the results obtained within collinear factorization[8]. To produce the latter, we have used the PYTHIA 8.2 MC generator. To cluster parton level particles into jets, the anti-k_T algorithm [70] was used, as implemented in the FASTJET package [192]. The cuts on the jets' momenta were set to $p_{T,1,2} > 20$ GeV and the rapidity window $3.2 < y_1, y_2 < 4.9$ was imposed to select the jets. The jet radius was set to $R = 0.5$, as in the HEF calculation discussed in preceding subsection.

PYTHIA was set up to generate events including the leading partonic sub-processes. For the comparison with HEF, the CT10 NLO PDFs [28] were used. Runs were performed at two proton-proton collision energies: 7 and 13 TeV. For each energy, two sets of MC data were produced, distinguished by the final state radiation (FSR) option turned on or off.

The comparison between the HEF and the collinear factorization results in shown in Fig. 30. We see that, in the case in which FSR radiation is turned off, the two formalisms agree quite well in description of the p_T spectra in the whole range of considered values. In the case of the azimuthal angle distributions, shown in the lower panel of Fig. 30, the results agree in the region of large and moderate angles and differ in the region of small angles. We attribute the latter to the effect of different treatment of singularities in the two frameworks. In HEF, matrix element diverges as the two outgoing partons become collinear, see Ref. [82]. This divergence is regularized by the jet algorithm, which is responsible for the kink around $\Delta\phi = 0.5$, seen in the lower panel of Fig. 30. On the other hand, in the case of PYTHIA, the shape of the distribution is a result of initial state radiation generated via parton shower matched with the collinear matrix element. Since the collinear matrix elements have different singularity structure than the HEF matrix elements this leads to different results at small $\Delta\phi$ [82, 110].

In Fig. 31 we show a comparison between distributions obtained with PYTHIA but using two different PDF sets. As we see, both sets give similar results, hence, the qualitative differences between PYTHIA and HEF, seen in Fig. 30, cannot be attributed to a choice of PDFs.

[8]The preliminary estimate of the result in HEF at 7 TeV has been performed in [211].

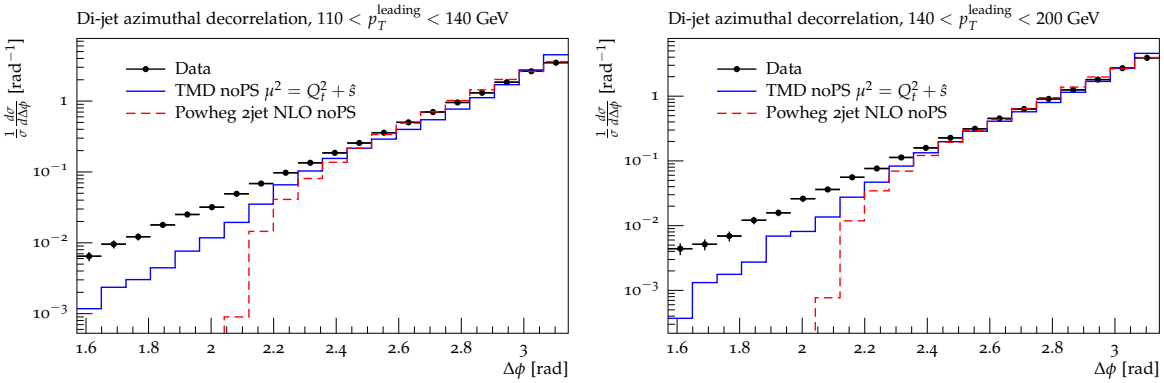


Figure 32: $\Delta\phi$ distribution for high p_{T} dijet production [212]. The solid (blue) histogram shows the prediction using off-shell $2 \to 2$ matrix elements with unintegrated parton densities, the dashed (red) line is a 3-parton configuration obtained with POWHEG . Both predictions are without parton shower and hadronization.

In Fig. 30, we also observe that turning on the FSR in PYTHIA leads to change in normalization of both the p_T and $\Delta\phi$ distributions. The spectra decrease by factor ~ 2 for moderate and large p_T, as well as $\Delta\phi$, values. Low-p_T and low-$\Delta\phi$ parts of the distributions are almost not affected by FSR. The observed difference in normalization of the transverse momentum spectra can be explained by the energy loss of the leading hard parton that happens readily via FSR parton shower emissions. For a significant fraction of events, this leads to the situation in which the parton originating from the hard collision splits into two partons separated by an angle sufficient to produce two lower-p_T jets. This mechanism takes the high-p_T events from the tail of the spectrum without FSR and moves them to the region below the jet cut. Hence, they effectively do not contribute to the observables shown in Fig. 30.

Finally, we mention that we have checked explicitly that the picture of Fig. 30 persists if PYTHIA events are supplemented with multi-parton interactions (MPI). Hence, forward dijet production in the collinear factorization framework is weakly sensitive to MPIs, which is consistent with the negligible effect of DPS in HEF, which we demonstrated in Section 5.2.

5.3.4 Predictions including transverse momentum dependent parton showers

Let us now move to predictions for dijet production that include initial state parton showers. The calculations were performed using off-shell matrix elements of $2 \to 2$ QCD processes and the unintegrated distributions obtained with KMRW procedure, described in Section 3. The results of the parton level calculation are fed via LHE files to the shower and hadronization interface of CASCADE [174, 175] for the transverse momentum dependent shower where the events in HEPMC format are stored for further processing as via RIVET [191]. The results were compared to the LHC data at $\sqrt{s} = 7$ TeV [212] in the central region. The events containing two leading jets, each with $p_T > 30$ GeV and rapidity $|y| < 1.1$, were selected and assigned into one of two regions, based on the largest transverse momentum $p_{\mathrm{T}}^{\text{leading}}$ in the event, *i.e.* $110 < p_{\mathrm{T}}^{\text{leading}} < 140$ and $140 < p_{\mathrm{T}}^{\text{leading}} < 200$ GeV.

First, we show parton level results of azimuthal decorrelations of high p_{T} dijet production. In Fig. 32 we compare predictions obtained from our calculation (without parton shower) with the one from POWHEG dijet (without parton shower). One can observe reasonable agreement between both parton-level calculations at high $\Delta\phi$. The POWHEG prediction shows a sharp drop at $\Delta\phi = 2\pi/3$, which is the kinematic limit for a 3 parton configuration. The prediction using transverse momentum dependent shower shows a smooth distribution to smaller values of $\Delta\phi$ which is typical for a configuration where more partons are radiated in the initial state. The distribution of our prediction depends entirely on the shape of the unintegrated distributions. Thus, with a precise determination of the unintegrated PDFs, we expect the $\Delta\phi$ distribution to be well described, without any tuning and without any adjustment of additional parameters.

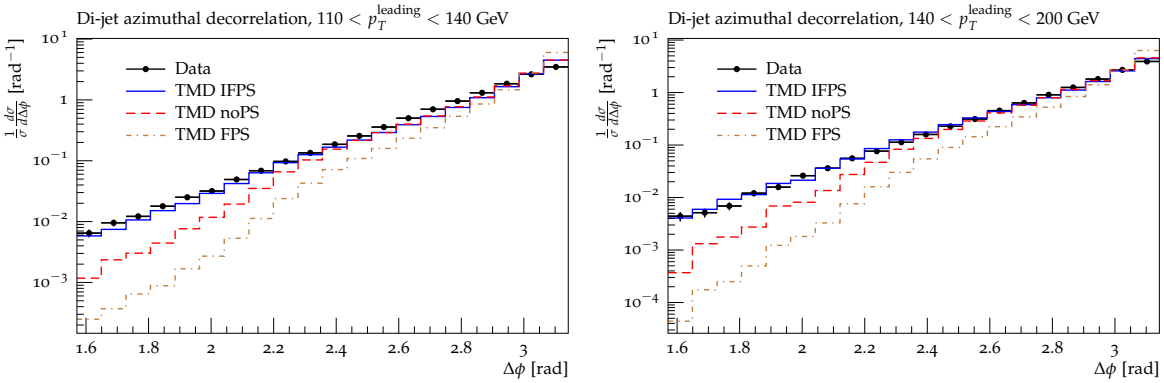


Figure 33: $\Delta\phi$ distribution of high p_{T} dijet events for different regions of $p_{\mathrm{T}}^{\mathrm{leading}}$: without parton shower (noPS, dashed red line), with final state parton shower (FPS, dashed-dotted brown line), with initial transverse momentum dependent shower and final state parton shower (IFPS, blue solid line). The factorization scale $\mu^2 = Q_T^2 + \hat{s}$ was chosen.

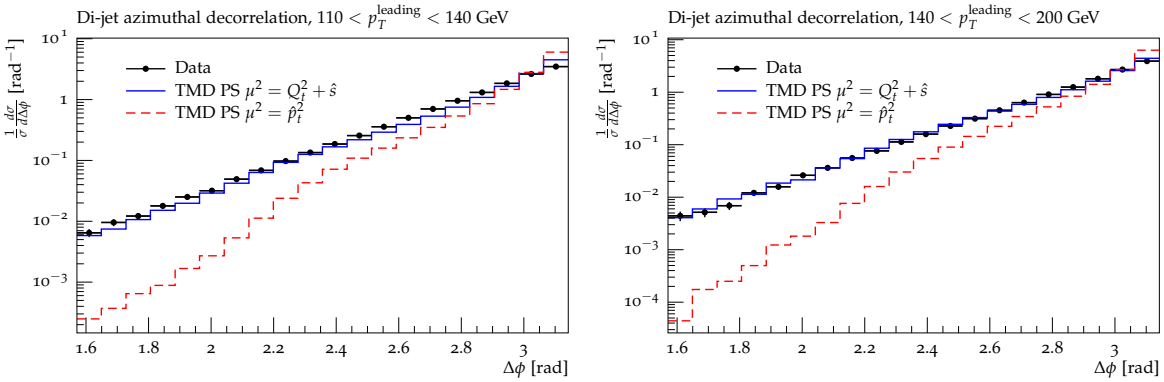


Figure 34: $\Delta\phi$ distribution as measured by [212] for different regions of $p_{\mathrm{T}}^{\mathrm{leading}}$. The data are compared with predictions using off-shell $2 \to 2$ matrix elements with unintegrated parton densities, an initial state TMD parton shower, conventional final state parton shower and hadronization. Shown are predictions for two different choices of the factorization scale, as discussed in the text.

In Fig. 32 we show the advantage in using unintegrated parton densities compared to a fixed order collinear calculation: due to the resummation of multiple parton emissions in the unintegrated parton density, the phase space for multi-jet production is covered, as seen in the tail to small $\Delta\phi$. Of course, the experimental measurement is different from a purely 2-parton final state, even using unintegrated distributions, since the jet clustering is based on multiple partons (hadrons). In Fig. 33 we show a comparison of the prediction using unintegrated PDFs with and without initial state transverse momentum dependent parton showering and including final state parton shower and hadronization (taken from PYTHIA [56]), with a final state parton shower scale of $\mu_{fps} = 2\hat{p}_T$ being the average transverse momentum of the outgoing matrix element partons. While, even without the parton shower, a tail towards small $\Delta\phi$ is observed, the simulation of the parton shower, both initial transverse momentum dependent and final state parton shower contributes to the shape of the distribution and brings it close to the measurement.

In Fig. 34 we show predictions for the azimuthal decorrelation $\Delta\phi$ for high p_{T} dijets for different regions of $p_{\mathrm{T}}^{\mathrm{leading}}$ using unintegrated parton densities with off-shell matrix elements, parton shower and hadronization in comparison with measurements at $\sqrt{s} = 7$ TeV, in pp collisions, at the LHC [212]. We show predictions for two different factorization scales: $\mu^2 = Q_T^2 + \hat{s}$, where Q_T is the vectorial sum of the initial state transverse momenta and $\sqrt{\hat{s}}$ is the invariant mass of the partonic subsystem and $\mu^2 = \hat{p_T}^2$. The first scale choice is

motivated by angular ordering (see Ref. [213]), the second one is the conventional scale choice. The scale choice motivated from angular ordering describes the measurements significantly better than the conventional one.

It is important to note, that there are no free parameters left: once the unintegrated parton density is determined, the initial state parton shower follows exactly the unintegrated parton distribution. The unintegrated parton distribution is the essential ingredient in the present calculation, and a precise determination of the parton distribution over a wide range in x, k_{T} and scale μ is an important topic. First steps towards a precision determination of the unintegrated densities from HERA measurements have been performed in Ref. [214, 215].

5.4 iTMD factorization for three jets

In this subsection, we study inclusive three jet production using HEF and iTMD factorizations. Schematically, this process can be written as

$$A + B \mapsto a + b \to \mathrm{jet}_1 + \mathrm{jet}_2 + \mathrm{jet}_3 + X\,, \tag{5.16}$$

with the longitudinal momentum fractions of the initiating partons parametrized with

$$x_1 = \sum_i \frac{|\vec{p}_{T\,i}|}{\sqrt{s}} e^{y_i}\,, \qquad x_2 = \sum_i \frac{|\vec{p}_{T\,i}|}{\sqrt{s}} e^{-y_i}\,, \tag{5.17}$$

where $\vec{p}_{T\,i}$ and y_i are the transverse momenta and rapidities of the jets. Here we only present a preliminary result, with only the gluon-gluon channel included. The cross section for the hybrid HEF reads [216]

$$d\sigma_{AB\to X} = \int \frac{d^2k_{T\,B}}{\pi} \int \frac{dx_2}{x_2} \int dx_1 f_{g/A}(x_1,\mu^2)\, |\overline{\mathcal{M}_{gg^*\to ggg}}|^2 \mathcal{F}_{g^*/B}(x_2, k_{T\,B})\,, \tag{5.18}$$

where $k_{T\,B}$ is the transverse momentum of the incoming off-shell gluon. As was discussed in Section 4, in iTMD factorization, each color flow channel has different distribution. Therefore, the matrix element squared, written using a vector of partial amplitudes $|\mathcal{M}|^2 = \vec{\mathcal{A}}^\dagger\, \mathbf{C}\, \vec{\mathcal{A}}$, together with unintegrated gluon distribution $\mathcal{F}_{g^*/B}(x_2, k_{T\,B})$, is replaced by the TMD distribution matrix, defined through Eq. (4.81). For the gluon channel only, we need to insert the result given by Eq. (4.95). Contributions with quarks would require unknown for the current moment distributions $\mathcal{F}^{(3)}_{qg}$ (4.34). In the large-N_c limit, the calculation of the gluon channel for three jet production requires the same gluon distributions that appear in the dijet production, *i.e.* $\mathcal{F}^{(1)}_{gg}$, $\mathcal{F}^{(2)}_{gg}$, $\mathcal{F}^{(6)}_{gg}$. In Ref. [135] these TMD distributions were calculated in the Gaussian approximation using the KS gluon distribution [111] , as described in subsection 4.4, and we use them to perform calculations of cross sections.

The calculations were performed for the center of mass energy $\sqrt{s} = 7$ TeV, with the following kinematic cuts restricting the phase space to interesting events. First, the transverse momentum cut for the final jets

$$|\vec{p}_{T_i}| > p_{T\,\mathrm{cut}}\,, \quad i = 1, 2, 3\,, \tag{5.19}$$

were set to $p_{T\,\mathrm{cut}} = 20$ GeV, also the following ordering of momenta was imposed, $|\vec{p}_{T\,1}| > |\vec{p}_{T\,2}| > |\vec{p}_{T\,3}|$. The rapidities were restricted to the forward region

$$3.2 \le y_i \le 4.9\,. \tag{5.20}$$

Further restriction is imposed by jet definition, and the usual anti-k_T clustering algorithm is applied with $R = 0.5$. The HEF calculations were performed using KaTie with CT10nlo collinear gluon distribution and KS nonlinear unintegrated gluon distribution, and the appropriate reweighting was performed by replacing the color matrix and KS distribution by the TMD distribution matrix, in the large-N_c limit given in the first

row in Table 13. The estimations of the theoretical uncertainty are obtained by varying the factorization and renormalization scales by factors $\frac{1}{2}$ and 2 around the central value equal to the average transverse momentum of the jets

$$\mu = \frac{p_{T\,1} + p_{T\,2} + p_{T\,3}}{3} \,. \tag{5.21}$$

We present the results for azimuthal decorrelations, *i.e.* the azimuthal angle between the hardest and the softest jet

$$\phi_{13} = |\phi_1 - \phi_3| \,, \quad \phi_{13} \in [0, 2\pi) \,, \tag{5.22}$$

and the jet unbalanced transverse momentum

$$\Delta p_T = |\vec{p}_{T\,1} + \vec{p}_{T\,2} + \vec{p}_{T\,3}| \,, \tag{5.23}$$

which corresponds to the transverse momentum of the off-shell gluon, $\Delta p_T = \left|\vec{k}_{T\,B}\right|$. The results are shown in Fig. 35. The iTMD predictions dominate in the whole range of considered values, both for ϕ_{13} and Δp_T. We can see, however, that that the difference between the two formalisms is smaller at smaller angles and higher Δp_T. The results of iTMD are larger especially at higher angles and smaller Δp_T, as one would naively expect.

These results are only preliminary, and further calculations are needed. As a first step, it is necessary to check if the inclusion of the remaining channels change the above picture and what predictions would be obtained for pA collisions. The potential correction might also come from the linearly polarized gluons [217].

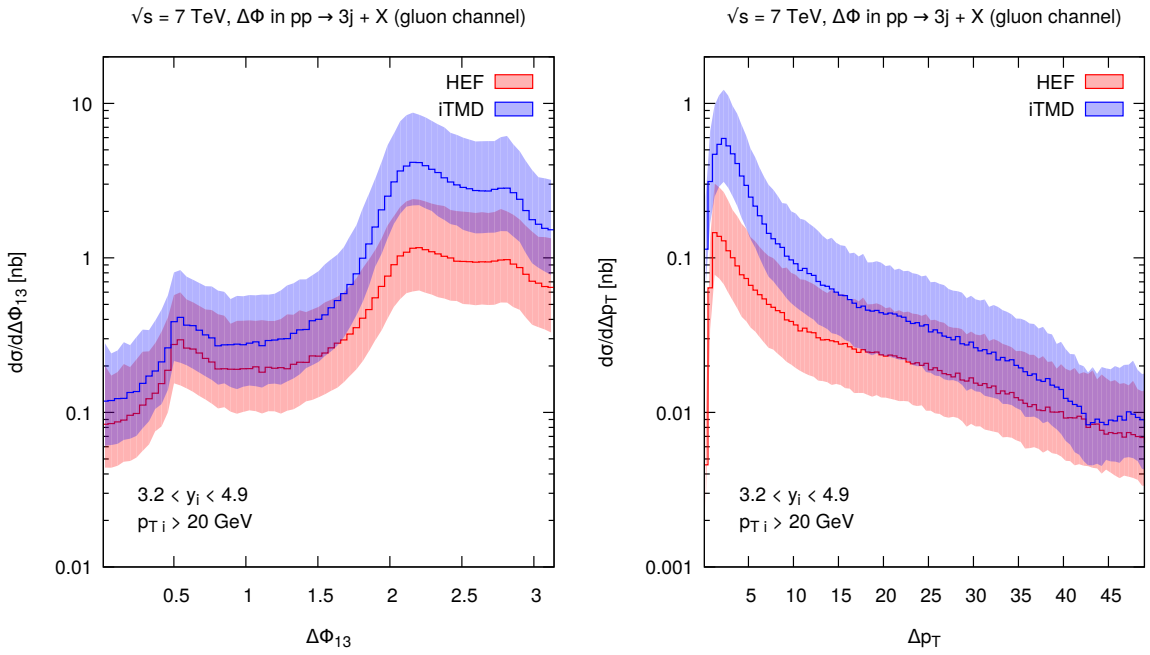


Figure 35: Differential cross section in difference of the azimuthal angles between the hardest and the softest jet (left) and differential cross section in the unbalanced p_T (right) for forward jets. The band represents the theoretical uncertainty due to the scale variation.

6 Summary

In this thesis we studied transverse-momentum dependent factorization frameworks in QCD. After presenting numerous introductory topics that describe the theoretical background and defining the components of the framework necessary for phenomenological analyses in Sections 1-3, we calculate transverse-momentum dependent gluon distributions for multiparton processes in Section 4, and describe the processes of inclusive one-, two- and three-jet production in Section 5.

In Section 4 we calculated the TMD gluon distributions for processes with five and six colored partons. So far, in the literature, processes with four partons were considered. At leading order, our results are sufficient to calculate three and four jet production in the gluon saturation regime, provided the two new-found basic TMD gluon distributions are determined. This can be done using the B-JIMWLK equation, as in [136]. Instead of calculating the structure of the operators for particular Feynman diagrams, we used the color decomposition, which is the most efficient way of dealing with the multi-particle QCD amplitudes. In particular, for processes with quarks, we used the color flow decomposition, which treats quarks and gluons on equal footing. In addition, we formulated straightforward color flow Feynman rules for the gauge links, which allow immediate derivation of the TMD operator for a given color flow. The color flow Feynman rules are particularly convenient for large N_c analysis. We found a general answer for a process with an arbitrary number of gluons, however in a certain approximation motivated by numerical studies of TMD gluon distributions. In the large N_c limit, we find that only two structures contribute, for any number of legs.

Before presenting multijet phenomenology in Section 5, the construction of initial-state parton shower based on unintegrated parton distributions is described. Afterwards, we start with single inclusive jet production. We demonstrated that the HEF framework describes well the single inclusive jet production and the main uncertainty comes from the unintegrated parton distributions. In this context, we have observed that the contribution from the off-shell quarks is negligible. Subsequently, we calculated transverse momentum and energy spectra of single inclusive forward jets in proton-lead collisions, in the rapidity region of $5.2 < y < 6.6$, which corresponds to the CASTOR detector acceptance at the CMS experiment. The HEF calculations have been performed using the KaTie Monte Carlo event generator supplemented with KS-linear, KS-nonlinear, and KMRW-lead parton densities, and interfaced to the CASCADE Monte Carlo event generator in order to account for hadronization effects. We observe that the energy and transverse momentum spectra of KS-linear are overall consistent with KMRW-lead spectra. The nonlinear dynamics as encoded in KS-nonlinear distributions predicts a suppression of the cross section for values of p_T smaller than 8 GeV. The energy spectrum, which can be measured in the CASTOR calorimeter, is affected by the nonlinear effects in the whole range. We also calculated nuclear modification ratios that measure the change of the dynamics as one increases the nuclear mass number. A clear difference between linear and nonlinear evolutions is observed with decreasing p_T or x, as expected from saturation effects. In order to improve on this, one needs formal calculations with higher order accuracy both of the hard matrix elements [218], as well as to account for first principles calculations of resummed higher order corrections to the gluon density including effects like: collinear resummation [219], running coupling and quarks contribution relevant at moderate and large x [214, 220]. This is also because the NLO corrections [221] introduce instabilities, as pointed out in [222]. Furthermore, a gluon density valid in the whole kinematic regime would increase the predictivity of the theory. Progress in the latter direction can be achieved once a program of calculating transverse-momentum dependent splitting functions in k_T-factorization is completed [223].

The analysis of inclusive dijet production started from calculating contributions to azimuthal angle distributions coming from double parton scattering, constructed from "squaring" the single jet production process in the HEF formalism. We shown that, for typical experimental cuts used in inclusive dijet production processes, the double parton scattering effects can be safely neglected. Next, we performed comparisons of predictions to dijet production calculated using HEF and collinear factorization. We observed that the effect of the final state radiation is not negligible and it leads to a change of normalization of differential distributions

in forward dijet production. Further improvements come from including initial-state transverse-momentum dependent parton shower, based on unintegrated parton distributions. We calculated the azimuthal decorrelation of high-p_T dijets and found very good agreement with the measurement. The implementation of unintegrated parton distribution functions together with the off-shell matrix element in calculations covers already a larger phase space than is accessible in collinear higher order calculations. Including initial-state transverse-momentum dependent parton showers together with the conventional final state parton showers gives a remarkably good description of the measurements, which opens the floor for a rich phenomenology at the LHC.

The combined set of tools, consisting of KaTie [190] parton-level event generator and Cascade [174] Monte Carlo generator, forms a complete framework to perform calculations in high energy factorization. KaTie produces matrix elements calculated in HEF and stores them in an LHE file, similar to the conventional LHE files, but now containing also the transverse momenta of the initial partons. The LHE files are read in by Cascade, which provide a newly developed initial-state transverse momentum dependent parton shower, as well as optional final-state parton shower and hadronization. Cascade produces events in HEPMC format [176], which can be further processed, *e.g.* using Rivet [191]. The set of unintegrated parton distributions, which include all flavors and is valid over a wide range in x, k_T and μ, was determined with KMRW approach using mrwcalc [122] and implemented in [121].

Finally, the calculations of TMD gluon distribution, performed in Section 4, were used to compare the predictions for three jet production in HEF and Improved TMD factorization.

A Examples

A.1 Distribution in quark-gluon scattering

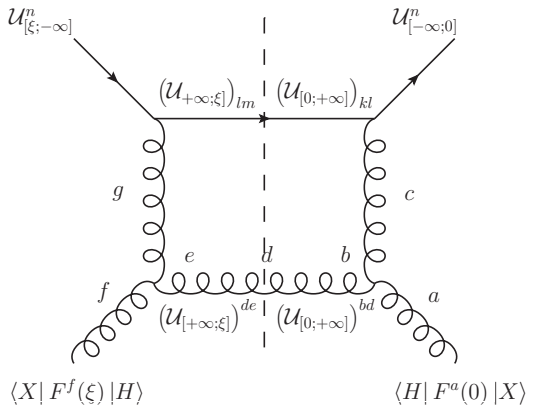


Figure 36: Diagram contributing to quark-gluon scattering.

The expression for gluon distribution correlator in the quark-gluon scattering channel shown in Fig. 36 is

$$\langle H|\, F^a(0)\,|X\rangle\, if^{abc}\left(\mathcal{U}_{[0;+\infty]}\right)^{bd}\left(\mathcal{U}_{[+\infty;\xi]}\right)^{de} if^{efg}\,\langle X|\, F^f(\xi)\,|H\rangle\,\times \\ \left(\mathcal{U}_{[-\infty;0]}\right)_{ij} t^c_{jk}\left(\mathcal{U}_{[0;+\infty]}\right)_{kl}\left(\mathcal{U}_{[+\infty;\xi]}\right)_{lm} t^g_{mn}\left(\mathcal{U}_{[\xi;-\infty]}\right)_{ni} =$$

$$= if^{abc}if^{efg}\,\langle H|\, F^a(0)\,|X\rangle\left(\mathcal{U}^{[+]}\right)^{be}\langle X|\, F^f(\xi)\,|H\rangle\times \mathrm{Tr}\left(\mathcal{U}^{[-]\dagger}t^c\mathcal{U}^{[+]}t^g\right).$$

Denoting $\langle H|\, F^a(0)\,|X\rangle\,\langle X|\, F^f(\xi)\,|H\rangle$ by $(\phi_g)^{af}$ and using Eq. (4.4) to transform Wilson line in adjoint representation to the fundamental representation

$$\left(\mathcal{U}^{[+]}\right)^{be} = \frac{1}{T_F}\,\mathrm{Tr}\left(t^b\mathcal{U}^{[+]}t^e\mathcal{U}^{[+]\dagger}\right)$$

we get to

$$\frac{1}{T_F}\,(\phi_g)^{af}\,if^{abc}if^{efg}\,\mathrm{Tr}\left(t^b\mathcal{U}^{[+]}t^e\mathcal{U}^{[+]\dagger}\right)\mathrm{Tr}\left(\mathcal{U}^{[-]\dagger}t^c\mathcal{U}^{[+]}t^g\right).$$

Next, we use commutation property of color matrices $[t^a,t^b]=if^{abc}t^c$ to get rid of structure constants

$$\frac{1}{T_F}\,(\phi_g)^{af}\,\mathrm{Tr}\left([t^a,t^c]\mathcal{U}^{[+]}[t^f,t^g]\mathcal{U}^{[+]\dagger}\right)\mathrm{Tr}\left(\mathcal{U}^{[-]\dagger}t^c\mathcal{U}^{[+]}t^g\right) =$$

$$= \frac{1}{T_F}\,(\phi_g)^{af}\left[\mathcal{U}^{[-]\dagger}_{ij}t^c_{jk}\mathcal{U}^{[+]}_{kl}t^g_{li}\right]\left[t^a_{mn}t^c_{no}\mathcal{U}^{[+]}_{op}t^f_{pq}t^g_{qr}\mathcal{U}^{[+]\dagger}_{rm} - t^a_{mn}t^c_{no}\mathcal{U}^{[+]}_{op}t^g_{pq}t^f_{qr}\mathcal{U}^{[+]\dagger}_{rm} - \right. \\ \left. - t^c_{mn}t^a_{no}\mathcal{U}^{[+]}_{op}t^f_{pq}t^g_{qr}\mathcal{U}^{[+]\dagger}_{rm} + t^c_{mn}t^a_{no}\mathcal{U}^{[+]}_{op}t^g_{pq}t^f_{qr}\mathcal{U}^{[+]\dagger}_{rm}\right] =$$

$$= \underbrace{\frac{1}{T_F}\left(\phi_g\right)^{af}\left[\mathcal{U}^{[-]\dagger}_{ij}\mathcal{U}^{[+]}_{kl}\mathcal{U}^{[+]}_{op}\mathcal{U}^{[+]\dagger}_{rm}\right]}_{[X]}\left[t^a_{mn}t^c_{jk}t^c_{no}t^f_{pq}t^g_{li}t^g_{qr} - t^a_{mn}t^c_{jk}t^c_{no}t^g_{li}t^g_{pq}t^f_{qr} - \right.$$

$$\left. - t^c_{jk}t^c_{mn}t^a_{no}t^f_{pq}t^g_{li}t^g_{qr} + t^c_{jk}t^c_{mn}t^a_{no}t^g_{li}t^g_{pq}t^f_{qr}\right].$$

Let us now take care of products of color matrices. We rewrite the expression in the second bracket using Fierz identity

$$T_F^2\Big[\underbrace{t^a_{mn}t^f_{pq}\left(\delta_{jo}\delta_{kn} - \frac{1}{N_c}\delta_{jk}\delta_{no}\right)\left(\delta_{lr}\delta_{iq} - \frac{1}{N_c}\delta_{li}\delta_{qr}\right)}_{(1)} -$$

$$\underbrace{- t^a_{mn}t^f_{qr}\left(\delta_{jo}\delta_{kn} - \frac{1}{N_c}\delta_{jk}\delta_{no}\right)\left(\delta_{lq}\delta_{ip} - \frac{1}{N_c}\delta_{li}\delta_{pq}\right)}_{(2)} -$$

$$\underbrace{- t^a_{no}t^f_{pq}\left(\delta_{jn}\delta_{km} - \frac{1}{N_c}\delta_{jk}\delta_{mn}\right)\left(\delta_{lr}\delta_{iq} - \frac{1}{N_c}\delta_{li}\delta_{qr}\right)}_{(3)} +$$

$$\underbrace{+ t^a_{no}t^f_{qr}\left(\delta_{jn}\delta_{km} - \frac{1}{N_c}\delta_{jk}\delta_{mn}\right)\left(\delta_{lq}\delta_{ip} - \frac{1}{N_c}\delta_{li}\delta_{pq}\right)}_{(4)}\Big]$$

$$(1) = t^a_{mn}t^f_{pq}\left(\delta^{jo}\delta^{kn}\delta^{lr}\delta^{iq} - \frac{1}{N_c}\delta^{jo}\delta^{kn}\delta^{li}\delta^{qr} - \frac{1}{N_c}\delta^{jk}\delta^{no}\delta^{lr}\delta^{iq} + \frac{1}{N_c^2}\delta^{jk}\delta^{no}\delta^{li}\delta^{qr}\right)$$

$$(2) = t^a_{mn}t^f_{qr}\left(\delta_{jo}\delta_{kn}\delta_{lq}\delta_{ip} - \frac{1}{N_c}\delta_{jo}\delta_{kn}\delta_{li}\delta_{pq} - \frac{1}{N_c}\delta_{jk}\delta_{no}\delta_{lq}\delta_{ip} + \frac{1}{N_c^2}\delta_{jk}\delta_{no}\delta_{li}\delta_{pq}\right)$$

$$(3) = t^a_{no}t^f_{pq}\left(\delta_{jn}\delta_{km}\delta_{lr}\delta_{iq} - \frac{1}{N_c}\delta_{jn}\delta_{km}\delta_{li}\delta_{qr} - \frac{1}{N_c}\delta_{jk}\delta_{mn}\delta_{lr}\delta_{iq} + \frac{1}{N_c^2}\delta_{jk}\delta_{mn}\delta_{li}\delta_{qr}\right)$$

$$(4) = t^a_{no}t^f_{qr}\left(\delta_{jn}\delta_{km}\delta_{lq}\delta_{ip} - \frac{1}{N_c}\delta_{jn}\delta_{km}\delta_{li}\delta_{pq} - \frac{1}{N_c}\delta_{jk}\delta_{mn}\delta_{lq}\delta_{ip} + \frac{1}{N_c^2}\delta_{jk}\delta_{mn}\delta_{li}\delta_{pq}\right)$$

Multiplying by the part that was left out ($T_F^2[X]$) we have

$$T_F^2[X](1) = T_F\left(\phi_g\right)^{af}\left[\mathcal{U}^{[-]\dagger}_{ij}\mathcal{U}^{[+]}_{kl}\mathcal{U}^{[+]}_{op}\mathcal{U}^{[+]\dagger}_{rm}\right]t^a_{mn}t^f_{pq}\big(\delta_{jo}\delta_{kn}\delta_{lr}\delta_{iq} -$$

$$- \frac{1}{N_c}\delta_{jo}\delta_{kn}\delta_{li}\delta_{qr} - \frac{1}{N_c}\delta_{jk}\delta_{no}\delta_{lr}\delta_{iq} + \frac{1}{N_c^2}\delta_{jk}\delta_{no}\delta_{li}\delta_{qr}\big) =$$

$$= T_F\left(\phi_g\right)^{af}\left[\mathcal{U}^{[-]\dagger}_{qo}\mathcal{U}^{[+]}_{op}t^f_{pq}\mathcal{U}^{[+]}_{nr}\mathcal{U}^{[+]\dagger}_{rm}t^a_{mn} - \frac{1}{N_c}\mathcal{U}^{[-]\dagger}_{io}\mathcal{U}^{[+]}_{op}t^f_{pq}\mathcal{U}^{[+]\dagger}_{qm}t^a_{mn}\mathcal{U}^{[+]}_{ni} - \right.$$

$$\left. - \frac{1}{N_c}\mathcal{U}^{[-]\dagger}_{qk}\mathcal{U}^{[+]}_{kl}\mathcal{U}^{[+]\dagger}_{lm}t^a_{mn}\mathcal{U}^{[+]}_{np}t^f_{pq} + \frac{1}{N_c^2}\mathcal{U}^{[-]\dagger}_{ik}\mathcal{U}^{[+]}_{ki}t^a_{mn}\mathcal{U}^{[+]}_{np}t^f_{pq}\mathcal{U}^{[+]\dagger}_{qm}\right] =$$

$$T_F\left(\phi_g\right)^{af}\Big[\operatorname{Tr}\left(\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}t^f\right)\underbrace{\operatorname{Tr}\left(\mathcal{U}^{[+]}\mathcal{U}^{[+]\dagger}t^a\right)}_{=\operatorname{Tr}(t^a)=0} - \frac{1}{N_c}\operatorname{Tr}\left(\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}t^a\mathcal{U}^{[+]}\right) -$$

$$- \frac{1}{N_c}\operatorname{Tr}\left(\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}\mathcal{U}^{[+]\dagger}t^a\mathcal{U}^{[+]}t^f\right) + \frac{1}{N_c^2}\operatorname{Tr}\left(\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}\right)\operatorname{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right)\Big] =$$

$$= T_F\left(\phi_g\right)^{af}\left[-\frac{1}{N_c}\,\mathrm{Tr}\left(t^a\mathcal{U}^{[+]}\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right)-\frac{1}{N_c}\,\mathrm{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[-]\dagger}\right)+\frac{1}{N_c^2}\,\mathrm{Tr}\left(\mathcal{U}^{[\square]}\right)\mathrm{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right)\right]$$

$$T_F^2[X](2) = T_F\left(\phi_g\right)^{af}\left[\mathcal{U}_{ij}^{[-]\dagger}\mathcal{U}_{kl}^{[+]}\mathcal{U}_{op}^{[+]}\mathcal{U}_{rm}^{[+]\dagger}\right]t^a_{mn}t^f_{qr}\big(\delta_{jo}\delta_{kn}\delta_{lq}\delta_{ip}-$$
$$-\frac{1}{N_c}\delta_{jo}\delta_{kn}\delta_{li}\delta_{pq}-\frac{1}{N_c}\delta_{jk}\delta_{no}\delta_{lq}\delta_{ip}+\frac{1}{N_c^2}\delta_{jk}\delta_{no}\delta_{li}\delta_{pq}\big)=$$

$$= T_F\left(\phi_g\right)^{af}\left[\mathcal{U}_{io}^{[-]\dagger}\mathcal{U}_{nq}^{[+]}t^f_{qr}\mathcal{U}_{oi}^{[+]}\mathcal{U}_{rm}^{[+]\dagger}t^a_{mn}-\frac{1}{N_c}\mathcal{U}_{io}^{[-]\dagger}\mathcal{U}_{ni}^{[+]}\mathcal{U}_{oq}^{[+]}t^f_{qr}\mathcal{U}_{rm}^{[+]\dagger}t^a_{mn}-\right.$$
$$\left.-\frac{1}{N_c}\mathcal{U}_{ik}^{[-]\dagger}\mathcal{U}_{kq}^{[+]}t^f_{qr}\mathcal{U}_{ni}^{[+]}\mathcal{U}_{rm}^{[+]\dagger}t^a_{mn}+\frac{1}{N_c^2}\mathcal{U}_{ik}^{[-]\dagger}\mathcal{U}_{ki}^{[+]}\mathcal{U}_{nq}^{[+]}t^f_{qr}\mathcal{U}_{rm}^{[+]\dagger}t^a_{mn}\right]=$$

$$= T_F\left(\phi_g\right)^{af}\left[\,\mathrm{Tr}\left(\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}\right)\mathrm{Tr}\left(\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}t^a\right)-\frac{1}{N_c}\,\mathrm{Tr}\left(\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}t^a\mathcal{U}^{[+]}\right)-\right.$$
$$\left.-\frac{1}{N_c}\,\mathrm{Tr}\left(\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}t^a\mathcal{U}^{[+]}\right)+\frac{1}{N_c^2}\,\mathrm{Tr}\left(\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}\right)\mathrm{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right)\right]=$$

$$= T_F\left(\phi_g\right)^{af}\left[\frac{N_c^2+1}{N_c^2}\,\mathrm{Tr}\left(\mathcal{U}^{[\square]}\right)\mathrm{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right)-\frac{2}{N_c}\,\mathrm{Tr}\left(t^a\mathcal{U}^{[+]}\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right)\right]$$

$$T_F^2[X](3) = T_F\left(\phi_g\right)^{af}\left[\mathcal{U}_{ij}^{[-]\dagger}\mathcal{U}_{kl}^{[+]}\mathcal{U}_{op}^{[+]}\mathcal{U}_{rm}^{[+]\dagger}\right]t^a_{no}t^f_{pq}\big(\delta_{jn}\delta_{km}\delta_{lr}\delta_{iq}-$$
$$-\frac{1}{N_c}\delta_{jn}\delta_{km}\delta_{li}\delta_{qr}-\frac{1}{N_c}\delta_{jk}\delta_{mn}\delta_{lr}\delta_{iq}+\frac{1}{N_c^2}\delta_{jk}\delta_{mn}\delta_{li}\delta_{qr}\big)$$

$$= T_F\left(\phi_g\right)^{af}\left[\mathcal{U}_{in}^{[-]\dagger}t^a_{no}\mathcal{U}_{mr}^{[+]}\mathcal{U}_{op}^{[+]}t^f_{pi}\mathcal{U}_{rm}^{[+]\dagger}-\frac{1}{N_c}\mathcal{U}_{in}^{[-]\dagger}t^a_{no}\mathcal{U}_{mi}^{[+]}\mathcal{U}_{op}^{[+]}t^f_{pq}\mathcal{U}_{qm}^{[+]\dagger}-\right.$$
$$\left.-\frac{1}{N_c}\mathcal{U}_{ik}^{[-]\dagger}\mathcal{U}_{kr}^{[+]}\mathcal{U}_{op}^{[+]}t^f_{pi}\mathcal{U}_{rn}^{[+]\dagger}t^a_{no}+\frac{1}{N_c^2}\mathcal{U}_{ik}^{[-]\dagger}\mathcal{U}_{ki}^{[+]}\mathcal{U}_{op}^{[+]}t^f_{pq}\mathcal{U}_{qn}^{[+]\dagger}t^a_{no}\right]=$$

$$= T_F\left(\phi_g\right)^{af}\left[\underbrace{\mathrm{Tr}\left(\mathcal{U}^{[+]}\mathcal{U}^{[+]\dagger}\right)}_{=\mathrm{Tr}(\mathbb{1})=N_c}\mathrm{Tr}\left(\mathcal{U}^{[-]\dagger}t^a\mathcal{U}^{[+]}t^f\right)-\frac{1}{N_c}\,\mathrm{Tr}\left(\mathcal{U}^{[-]\dagger}t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\mathcal{U}^{[+]}\right)-\right.$$
$$\left.-\frac{1}{N_c}\,\mathrm{Tr}\left(\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}\mathcal{U}^{[+]\dagger}t^a\mathcal{U}^{[+]}t^f\right)+\frac{1}{N_c^2}\,\mathrm{Tr}\left(\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}\right)\mathrm{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right)\right]=$$

$$= T_F\left(\phi_g\right)^{af}\left[\frac{N_c^2-2}{N_c}\,\mathrm{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[-]\dagger}\right)+\frac{1}{N_c^2}\,\mathrm{Tr}\left(\mathcal{U}^{[\square]}\right)\mathrm{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right)\right]$$

$$T_F^2[X](4) = T_F\left(\phi_g\right)^{af}\left[\mathcal{U}_{ij}^{[-]\dagger}\mathcal{U}_{kl}^{[+]}\mathcal{U}_{op}^{[+]}\mathcal{U}_{rm}^{[+]\dagger}\right]t^a_{no}t^f_{qr}\big(\delta_{jn}\delta_{km}\delta_{lq}\delta_{ip}-$$
$$-\frac{1}{N_c}\delta_{jn}\delta_{km}\delta_{li}\delta_{pq}-\frac{1}{N_c}\delta_{jk}\delta_{mn}\delta_{lq}\delta_{ip}+\frac{1}{N_c^2}\delta_{jk}\delta_{mn}\delta_{li}\delta_{pq}\big)$$

$$= T_F\left(\phi_g\right)^{af}\left[\mathcal{U}^{[-]\dagger}_{in}t^a_{no}\mathcal{U}^{[+]}_{mq}t^f_{qr}\mathcal{U}^{[+]}_{oi}\mathcal{U}^{[+]\dagger}_{rm} - \frac{1}{N_c}\mathcal{U}^{[-]\dagger}_{in}t^a_{no}\mathcal{U}^{[+]}_{mi}\mathcal{U}^{[+]}_{oq}t^f_{qr}\mathcal{U}^{[+]\dagger}_{rm} - \right.$$
$$\left. - \frac{1}{N_c}\mathcal{U}^{[-]\dagger}_{ik}\mathcal{U}^{[+]}_{kq}t^f_{qr}\mathcal{U}^{[+]}_{oi}\mathcal{U}^{[+]\dagger}_{rn}t^a_{no} + \frac{1}{N_c^2}\mathcal{U}^{[-]\dagger}_{ik}\mathcal{U}^{[+]}_{ki}\mathcal{U}^{[+]}_{oq}t^f_{qr}\mathcal{U}^{[+]\dagger}_{rn}t^a_{no}\right] =$$

$$= T_F\left(\phi_g\right)^{af}\left[\operatorname{Tr}\left(\mathcal{U}^{[-]\dagger}t^a\mathcal{U}^{[+]}\right)\underbrace{\operatorname{Tr}\left(\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right)}_{=\operatorname{Tr}(t^f)=0} - \frac{1}{N_c}\operatorname{Tr}\left(\mathcal{U}^{[-]\dagger}t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\mathcal{U}^{[+]}\right) - \right.$$
$$\left. - \frac{1}{N_c}\operatorname{Tr}\left(\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}t^a\mathcal{U}^{[+]}\right) + \frac{1}{N_c^2}\operatorname{Tr}\left(\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}\right)\operatorname{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right)\right] =$$

$$= T_F\left(\phi_g\right)^{af}\left[-\frac{1}{N_c}\operatorname{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[-]\dagger}\right) - \frac{1}{N_c}\operatorname{Tr}\left(t^a\mathcal{U}^{[+]}\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right) + \frac{1}{N_c^2}\operatorname{Tr}\left(\mathcal{U}^{[\square]}\right)\operatorname{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right)\right]$$

$$T_F^2\left([X](1) - [X](2)\right) = T_F\left(\phi_g\right)^{af}\left[\frac{1}{N_c}\operatorname{Tr}\left(t^a\mathcal{U}^{[+]}\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right) - \frac{1}{N_c}\operatorname{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[-]\dagger}\right) - \right.$$
$$\left. - \operatorname{Tr}\left(\mathcal{U}^{[\square]}\right)\operatorname{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right)\right]$$

$$T_F^2\left([X](4) - [X](3)\right) = T_F\left(\phi_g\right)^{af}\left[\left(-N_c + \frac{1}{N_c}\right)\operatorname{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[-]\dagger}\right) - \frac{1}{N_c}\operatorname{Tr}\left(t^a\mathcal{U}^{[+]}\mathcal{U}^{[-]\dagger}\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right)\right]$$

The final formula is

$$T_F^2\left([X](1) - [X](2) - [X](3) + [X](4)\right) = T_F\left(\phi_g\right)^{af}\left[-N_c\operatorname{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[-]\dagger}\right) - \operatorname{Tr}\left(\mathcal{U}^{[\square]}\right)\operatorname{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right)\right] =$$

$$= -2T_F N_c\left[\frac{1}{2}\operatorname{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[-]\dagger}\right) + \frac{1}{2}\frac{\operatorname{Tr}\left(\mathcal{U}^{[\square]}\right)}{N_c}\operatorname{Tr}\left(t^a\mathcal{U}^{[+]}t^f\mathcal{U}^{[+]\dagger}\right)\right]\left(\phi_g\right)^{af}$$

The color structure of the bare diagram with open indices a and f is

$$if^{abc}t^c_{ij}if^{fbg}t^g_{ji} = -f^{abc}f^{fbg}\underbrace{\operatorname{Tr}\left(t^c t^g\right)}_{T_F\delta_{cg}} = -T_F f^{abc}f^{fbg}\delta_{cg} = -T_F\underbrace{f^{abc}f^{fbc}}_{N_c\delta^{af}} = -T_F N_c\delta^{af}$$

This result agrees with Ref. [97], however the color factor for the bare diagram which they have is $-2T_F N_c$, and we have $-T_F N_c$. There is a difference in conventions by the factor of $1/T_F$ (we have to multiply our final result by T_F). The final expression for the gluon correlator is

$$\Phi_g^{[\mathcal{U}]}(x, k_T) = \int\frac{d\xi^- d^2\xi_T}{(2\pi)^3}e^{ik\cdot\xi}\langle H|\frac{1}{2}\operatorname{Tr}\left\{F(0)\mathcal{U}^{[+]}F(\xi)\mathcal{U}^{[-]\dagger}\right\} + \frac{1}{2}\frac{\operatorname{Tr}\left(\mathcal{U}^{[\square]}\right)}{N_c}\operatorname{Tr}\left\{F(0)\mathcal{U}^{[+]}F(\xi)\mathcal{U}^{[+]\dagger}\right\}|H\rangle\,.$$

A.2 Distribution in quark-gluon scattering - mixed diagram

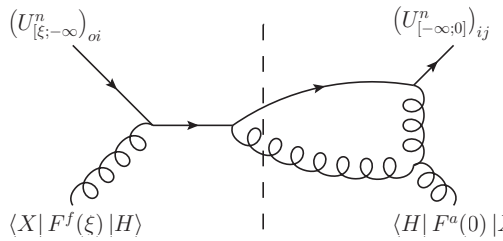


Figure 37: Diagram contributing to quark-gluon scattering (mixed channel).

Due to a large number of indices, the capital letters will denote the indices in the adjoint representation. The expression for the gluon correlator in the channel shown in Fig. 37 reads

$$\langle H|F^A(0)|X\rangle if^{ABC}\left(\mathcal{U}_{[0;+\infty]}\right)^{CX}\left(\mathcal{U}_{[+\infty;\xi]}\right)^{XD}\langle X|F^E(\xi)|H\rangle\times$$
$$\times\left(\mathcal{U}_{[-\infty;0]}\right)_{xi}t^B_{ij}\left(\mathcal{U}_{[0;+\infty]}\right)_{jy}\left(\mathcal{U}_{[+\infty;\xi]}\right)_{yk}t^D_{kl}t^E_{lm}\left(\mathcal{U}_{[\xi;-\infty]}\right)_{mx}=$$

$$=\Phi^{AE}if^{ABC}\left(\mathcal{U}^{[+]}\right)^{CD}\mathrm{Tr}\left(t^B\mathcal{U}^{[+]}t^Dt^E\mathcal{U}^{[-]\dagger}\right)=$$

Now we transform $\left(\mathcal{U}^{[+]}\right)^{BE}$ to fundamental representation using (4.4)

$$=\frac{\Phi^{AE}}{T_F}if^{ABC}\,\mathrm{Tr}\left(t^C\mathcal{U}^{[+]}t^D\mathcal{U}^{[+]\dagger}\right)\mathrm{Tr}\left(t^B\mathcal{U}^{[+]}t^Dt^E\mathcal{U}^{[-]\dagger}\right)=$$

$$=\frac{\Phi^{AE}}{T_F}\,\mathrm{Tr}\left([t^A,t^B]\mathcal{U}^{[+]}t^D\mathcal{U}^{[+]\dagger}\right)\mathrm{Tr}\left(t^B\mathcal{U}^{[+]}t^Dt^E\mathcal{U}^{[-]\dagger}\right)=$$

$$=\frac{\Phi^{AE}}{T_F}\left(t^A_{no}t^B_{op}-t^B_{no}t^A_{op}\right)\mathcal{U}^{[+]}_{pq}t^D_{qr}\mathcal{U}^{[+]\dagger}_{rn}\times t^B_{ij}\mathcal{U}^{[+]}_{jk}t^D_{kl}t^E_{lm}\mathcal{U}^{[-]\dagger}_{mi}=$$

$$=\frac{\Phi^{AE}}{T_F}\mathcal{U}^{[+]}_{jk}t^E_{lm}\mathcal{U}^{[-]\dagger}_{mi}\mathcal{U}^{[+]}_{pq}\mathcal{U}^{[+]\dagger}_{rn}\left(t^A_{no}t^B_{ij}t^B_{op}-t^B_{ij}t^B_{no}t^A_{op}\right)t^D_{kl}t^D_{qr}=$$

$$=\frac{\Phi^{AE}}{T_F}\mathcal{U}^{[+]}_{jk}t^E_{lm}\mathcal{U}^{[-]\dagger}_{mi}\mathcal{U}^{[+]}_{pq}\mathcal{U}^{[+]\dagger}_{rn}\left[t^A_{no}T_F\left(\delta_{ip}\delta_{jo}-\frac{1}{N_c}\delta_{ij}\delta_{op}\right)-t^A_{op}T_F\left(\delta_{io}\delta_{jn}-\frac{1}{N_c}\delta_{ij}\delta_{no}\right)\right]T_F\left(\delta_{kr}\delta_{lq}-\frac{1}{N_c}\delta_{kl}\delta_{qr}\right)=$$

$$=T_F\Phi^{AE}\mathcal{U}^{[+]}_{jk}t^E_{lm}\mathcal{U}^{[-]\dagger}_{mi}\left[\left(t^A_{nj}\delta_{ip}-\underline{\frac{1}{N_c}t^A_{np}\delta_{ij}}\right)-\left(t^A_{ip}\delta_{jn}-\underline{\frac{1}{N_c}t^A_{np}\delta_{ij}}\right)\right]\left(\mathcal{U}^{[+]}_{pl}\mathcal{U}^{[+]\dagger}_{kn}-\frac{1}{N_c}\mathcal{U}^{[+]}_{pq}\mathcal{U}^{[+]\dagger}_{qn}\delta_{kl}\right)=$$

$$=T_F\Phi^{AE}\mathcal{U}^{[+]}_{jk}t^E_{lm}\mathcal{U}^{[-]\dagger}_{mi}\left[t^A_{nj}\delta_{ip}-t^A_{ip}\delta_{jn}\right]\left(\mathcal{U}^{[+]}_{pl}\mathcal{U}^{[+]\dagger}_{kn}-\frac{1}{N_c}\mathcal{U}^{[+]}_{pq}\mathcal{U}^{[+]\dagger}_{qn}\delta_{kl}\right)=$$

$$=T_F\Phi^{AE}\mathcal{U}^{[+]}_{jk}t^E_{lm}\mathcal{U}^{[-]\dagger}_{mi}\left[\left(\mathcal{U}^{[+]}_{il}\mathcal{U}^{[+]\dagger}_{kn}t^A_{nj}-\frac{1}{N_c}\underbrace{U^{[+]}_{iq}\mathcal{U}^{[+]\dagger}_{qn}}\delta_{kl}t^A_{nj}\right)-\left(\mathcal{U}^{[+]}_{pl}\mathcal{U}^{[+]\dagger}_{kj}t^A_{ip}-\frac{1}{N_c}\underbrace{\mathcal{U}^{[+]}_{pq}\mathcal{U}^{[+]\dagger}_{qj}}\delta_{kl}t^A_{ip}\right)\right]=$$

$$=T_F\Phi^{AE}\mathcal{U}^{[+]}_{jk}t^E_{lm}\mathcal{U}^{[-]\dagger}_{mi}\left[\left(\mathcal{U}^{[+]}_{il}\mathcal{U}^{[+]\dagger}_{kn}t^A_{nj}-\underline{\frac{1}{N_c}t^A_{ij}\delta_{kl}}\right)-\left(t^A_{ip}\mathcal{U}^{[+]}_{pl}\mathcal{U}^{[+]\dagger}_{kj}-\underline{\frac{1}{N_c}t^A_{ij}\delta_{kl}}\right)\right]=$$

$$=T_F\Phi^{AE}\left[\underbrace{t^A_{nj}\mathcal{U}^{[+]}_{jk}\mathcal{U}^{[+]\dagger}_{kn}}\mathcal{U}^{[+]}_{il}t^E_{lm}\mathcal{U}^{[-]\dagger}_{mi}-t^A_{ip}\mathcal{U}^{[+]}_{pl}t^E_{lm}\mathcal{U}^{[-]\dagger}_{mi}\underbrace{\mathcal{U}^{[+]\dagger}_{kj}\mathcal{U}^{[+]}_{jk}}\right]=$$

$$=-N_cT_F\Phi^{AE}\,\mathrm{Tr}\left[t^A\mathcal{U}^{[+]}t^E\mathcal{U}^{[-]\dagger}\right]=$$

$$= -N_c T_F \langle H| \operatorname{Tr}\left[F(0)\mathcal{U}^{[+]}F(\xi)\mathcal{U}^{[-]\dagger}\right]|H\rangle = -N_c T_F \langle H| \operatorname{Tr}\left[F(\xi)\mathcal{U}^{[-]\dagger}F(0)\mathcal{U}^{[+]}\right]|H\rangle$$

The color factor for the bare diagram is

$$if^{ABC}t^B_{ij}t^C_{jk}t^D_{ki} = t^B_{ij}\left[t^A,t^B\right]_{jk}t^D_{ki} = t^B_{ij}\left(t^A_{jm}t^B_{mk} - t^B_{jm}t^A_{mk}\right)t^D_{ki} =$$
$$= t^A_{jm}t^D_{ki}t^B_{ij}t^B_{mk} - t^A_{mk}t^D_{ki}t^B_{ij}t^B_{jm} = t^A_{jm}t^D_{ki}T_F\left(\delta_{ik}\delta_{jm} - \frac{1}{N_c}\delta_{ij}\delta_{mk}\right) - t^A_{mk}t^D_{ki}T_F\left(\delta_{im}\delta_{jj} - \frac{1}{N_c}\delta_{ij}\delta_{jm}\right) =$$
$$= -\frac{T_F}{N_c}\operatorname{Tr}\left(t^A t^D\right) - N_c T_F \operatorname{Tr}\left(t^A t^D\right) + \frac{T_F}{N_c}\operatorname{Tr}\left(t^A t^D\right) =$$
$$= -N_c T_F \operatorname{Tr}\left(t^A t^D\right) = -N_c T_F^2 \delta^{AD}$$

The important thing to watch is the direction of color indices in the structure constants. In the amplitude the indices are taken in clockwise direction, whereas in conjugated amplitude with anticlockwise direction. The final answer is

$$\Phi_g^{[\mathcal{U}]}(x,k_T) = \int \frac{d\xi^- d^2\xi_T}{(2\pi)^3}e^{ik\cdot\xi}\frac{1}{T_F}\langle H| \operatorname{Tr}\left[F(\xi)\mathcal{U}^{[-]\dagger}F(0)\mathcal{U}^{[+]}\right]|H\rangle\,.$$

A.3 Distribution in gluon-gluon scattering

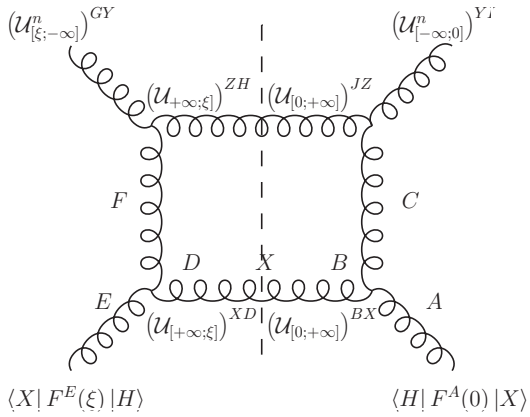


Figure 38: Diagram contributing to gluon-gluon scattering.

The expression for the gluon correlator in the gluon-gluon scattering channel shown in Fig. 38 reads

$$\langle H|F^A(0)|X\rangle if^{ACB}\left(\mathcal{U}_{[0;+\infty]}\right)^{BX}\left(\mathcal{U}_{[+\infty;\xi]}\right)^{XD}if^{DEF}\langle X|F^E(\xi)|H\rangle\times$$
$$\times\left(\mathcal{U}_{[-\infty;0]}\right)^{YI}if^{IJC}\left(\mathcal{U}_{[0;+\infty]}\right)^{JZ}\left(\mathcal{U}_{[+\infty;\xi]}\right)^{ZH}if^{GHF}\left(\mathcal{U}_{[\xi;-\infty]}\right)^{GY} =$$

$$= \Phi^{AE}if^{ACB}\left(\mathcal{U}^{[+]}\right)^{BD}if^{DEF}if^{IJC}\left(\mathcal{U}^{[+]}\right)^{JH}if^{GHF}\left(\mathcal{U}^{[-]\dagger}\right)^{GI} =$$
$$= \frac{\Phi^{AE}}{T_F^3}if^{ACB}\operatorname{Tr}\left(t^B\mathcal{U}^{[+]}t^D\mathcal{U}^{[+]\dagger}\right)if^{DEF}if^{IJC}\operatorname{Tr}\left(t^J\mathcal{U}^{[+]}t^H\mathcal{U}^{[+]\dagger}\right)if^{GHF}\operatorname{Tr}\left(t^G\mathcal{U}^{[-]\dagger}t^I\mathcal{U}^{[-]}\right) =$$
$$= \frac{\Phi^{AE}}{T_F^3}\operatorname{Tr}\left([t^A,t^C]\mathcal{U}^{[+]}[t^E,t^F]\mathcal{U}^{[+]\dagger}\right)\operatorname{Tr}\left(t^J\mathcal{U}^{[+]}t^H\mathcal{U}^{[+]\dagger}\right)\operatorname{Tr}\left([t^H,t^F]\mathcal{U}^{[-]\dagger}[t^J,t^C]\mathcal{U}^{[-]}\right) =$$

$$= \frac{\Phi^{AE}}{T_F^3} \left(t_{ij}^A t_{jk}^C - t_{ij}^C t_{jk}^A\right) \mathcal{U}_{kl}^{[+]} \left(t_{lm}^E t_{mn}^F - t_{lm}^F t_{mn}^E\right) \mathcal{U}_{ni}^{[+]\dagger} t_{op}^J \mathcal{U}_{pq}^{[+]} t_{qr}^H \mathcal{U}_{ro}^{[+]\dagger} \left(t_{st}^H t_{tu}^F - t_{st}^F t_{tu}^H\right) \mathcal{U}_{uv}^{[-]\dagger} \left(t_{vw}^J t_{wx}^C - t_{vw}^C t_{wx}^J\right) \mathcal{U}_{xs}^{[-]} =$$

$$= \frac{\Phi^{AE}}{T_F^3} \mathcal{U}_{kl}^{[+]} \mathcal{U}_{ni}^{[+]\dagger} \mathcal{U}_{pq}^{[+]} \mathcal{U}_{ro}^{[+]\dagger} \mathcal{U}_{uv}^{[-]\dagger} \mathcal{U}_{xs}^{[-]} \left(t_{ij}^A t_{jk}^C - t_{ij}^C t_{jk}^A\right) \left(t_{lm}^E t_{mn}^F - t_{lm}^F t_{mn}^E\right) t_{qr}^H \left(t_{st}^H t_{tu}^F - t_{st}^F t_{tu}^H\right) t_{op}^J \left(t_{vw}^J t_{wx}^C - t_{vw}^C t_{wx}^J\right) =$$

$$= \frac{\Phi^{AE}}{T_F^3} \mathcal{U}_{kl}^{[+]} \mathcal{U}_{ni}^{[+]\dagger} \mathcal{U}_{pq}^{[+]} \mathcal{U}_{ro}^{[+]\dagger} \mathcal{U}_{uv}^{[-]\dagger} \mathcal{U}_{xs}^{[-]} \left(t_{ij}^A t_{jk}^C - t_{ij}^C t_{jk}^A\right) \left(t_{lm}^E t_{mn}^F - t_{lm}^F t_{mn}^E\right) \times$$
$$\times \left[t_{tu}^F T_F \left(\delta_{qt}\delta_{rs} - \frac{1}{N_c}\delta_{qr}\delta_{st}\right) - t_{st}^F T_F \left(\delta_{qu}\delta_{rt} - \frac{1}{N_c}\delta_{qr}\delta_{tu}\right)\right] \times$$
$$\times \left[t_{wx}^C T_F \left(\delta_{ow}\delta_{pv} - \frac{1}{N_c}\delta_{op}\delta_{vw}\right) - t_{vw}^C T_F \left(\delta_{ox}\delta_{pw} - \frac{1}{N_c}\delta_{op}\delta_{wx}\right)\right] =$$

$$= \frac{\Phi^{AE}}{T_F} \mathcal{U}_{kl}^{[+]} \mathcal{U}_{ni}^{[+]\dagger} \mathcal{U}_{pq}^{[+]} \mathcal{U}_{ro}^{[+]\dagger} \mathcal{U}_{uv}^{[-]\dagger} \mathcal{U}_{xs}^{[-]} \left(t_{ij}^A t_{jk}^C - t_{ij}^C t_{jk}^A\right) \left(t_{lm}^E t_{mn}^F - t_{lm}^F t_{mn}^E\right) \times$$
$$\times \left[t_{qu}^F \delta_{rs} - \underline{\frac{1}{N_c} t_{su}^F \delta_{qr}} - t_{sr}^F \delta_{qu} + \underline{\frac{1}{N_c} t_{su}^F \delta_{qr}}\right] \left[t_{ox}^C \delta_{pv} - \underline{\frac{1}{N_c} t_{vx}^C \delta_{op}} - t_{vp}^C \delta_{ox} + \underline{\frac{1}{N_c} t_{vx}^C \delta_{op}}\right] =$$

$$= \frac{\Phi^{AE}}{T_F} \mathcal{U}_{kl}^{[+]} \mathcal{U}_{ni}^{[+]\dagger} \mathcal{U}_{pq}^{[+]} \mathcal{U}_{ro}^{[+]\dagger} \mathcal{U}_{uv}^{[-]\dagger} \mathcal{U}_{xs}^{[-]} \left(t_{ij}^A t_{jk}^C - t_{ij}^C t_{jk}^A\right) \left[t_{ox}^C \delta_{pv} - t_{vp}^C \delta_{ox}\right] \left(t_{lm}^E t_{mn}^F - t_{lm}^F t_{mn}^E\right) \left[t_{qu}^F \delta_{rs} - t_{sr}^F \delta_{qu}\right] =$$
$$= \frac{\Phi^{AE}}{T_F} \mathcal{U}_{kl}^{[+]} \mathcal{U}_{ni}^{[+]\dagger} \mathcal{U}_{pq}^{[+]} \mathcal{U}_{ro}^{[+]\dagger} \mathcal{U}_{uv}^{[-]\dagger} \mathcal{U}_{xs}^{[-]} \left[t_{ij}^A t_{jk}^C t_{ox}^C \delta_{pv} - t_{ij}^A t_{jk}^C t_{vp}^C \delta_{ox} - t_{ij}^C t_{jk}^A t_{ox}^C \delta_{pv} + t_{ij}^C t_{jk}^A t_{vp}^C \delta_{ox}\right] \times$$
$$\times \left[t_{lm}^E t_{mn}^F t_{qu}^F \delta_{rs} - t_{lm}^E t_{mn}^F t_{sr}^F \delta_{qu} - t_{lm}^F t_{mn}^E t_{qu}^F \delta_{rs} + t_{lm}^F t_{mn}^E t_{sr}^F \delta_{qu}\right] =$$

$$= T_F \Phi^{AE} \mathcal{U}_{kl}^{[+]} \mathcal{U}_{ni}^{[+]\dagger} \mathcal{U}_{pq}^{[+]} \mathcal{U}_{ro}^{[+]\dagger} \mathcal{U}_{uv}^{[-]\dagger} \mathcal{U}_{xs}^{[-]} \times$$
$$\times \left[t_{ij}^A \delta_{pv} \left(\delta_{jx}\delta_{ko} - \frac{1}{N_c}\delta_{jk}\delta_{ox}\right) - t_{ij}^A \delta_{ox} \left(\delta_{jp}\delta_{kv} - \frac{1}{N_c}\delta_{jk}\delta_{vp}\right) - t_{jk}^A \delta_{pv} \left(\delta_{ix}\delta_{jo} - \frac{1}{N_c}\delta_{ij}\delta_{ox}\right) + t_{jk}^A \delta_{ox} \left(\delta_{ip}\delta_{jv} - \frac{1}{N_c}\delta_{ij}\delta_{vp}\right)\right]$$
$$\times \left[t_{lm}^E \delta_{rs} \left(\delta_{mu}\delta_{nq} - \frac{1}{N_c}\delta_{mn}\delta_{qu}\right) - t_{lm}^E \delta_{qu} \left(\delta_{mr}\delta_{ns} - \frac{1}{N_c}\delta_{mn}\delta_{sr}\right) -\right.$$
$$\left. - t_{mn}^E \delta_{rs} \left(\delta_{lu}\delta_{mq} - \frac{1}{N_c}\delta_{lm}\delta_{qu}\right) + t_{mn}^E \delta_{qu} \left(\delta_{lr}\delta_{ms} - \frac{1}{N_c}\delta_{lm}\delta_{sr}\right)\right] =$$

$$= T_F \Phi^{AE} \mathcal{U}_{kl}^{[+]} \mathcal{U}_{ni}^{[+]\dagger} \mathcal{U}_{pq}^{[+]} \mathcal{U}_{ro}^{[+]\dagger} \mathcal{U}_{uv}^{[-]\dagger} \mathcal{U}_{xs}^{[-]} \times$$
$$\times \left[t_{ix}^A \delta_{pv}\delta_{ko} - t_{ip}^A \delta_{ox}\delta_{kv} - t_{ok}^A \delta_{pv}\delta_{ix} + t_{vk}^A \delta_{ox}\delta_{ip}\right] \left[t_{lu}^E \delta_{rs}\delta_{nq} - t_{lr}^E \delta_{qu}\delta_{ns} - t_{qn}^E \delta_{rs}\delta_{lu} + t_{sn}^E \delta_{qu}\delta_{lr}\right] =$$

$$= T_F \Phi^{AE} \mathcal{U}_{kl}^{[+]} \mathcal{U}_{ni}^{[+]\dagger} \times \left[\mathcal{U}_{up}^{[-]\dagger} t_{ix}^A \mathcal{U}_{xs}^{[-]} \delta_{ko} - \mathcal{U}_{uk}^{[-]\dagger} \mathcal{U}_{os}^{[-]} t_{ip}^A - \mathcal{U}_{up}^{[-]\dagger} \mathcal{U}_{is}^{[-]} t_{ok}^A + \mathcal{U}_{uv}^{[-]\dagger} \mathcal{U}_{os}^{[-]} t_{vk}^A \delta_{ip}\right] \times$$
$$\times \left[\mathcal{U}_{pn}^{[+]} \mathcal{U}_{so}^{[+]\dagger} t_{lu}^E - \mathcal{U}_{pu}^{[+]} \mathcal{U}_{ro}^{[+]\dagger} t_{lr}^E \delta_{ns} - \mathcal{U}_{pq}^{[+]} \mathcal{U}_{so}^{[+]\dagger} t_{qn}^E \delta_{lu} + \mathcal{U}_{pu}^{[+]} \mathcal{U}_{lo}^{[+]\dagger} t_{sn}^E\right] =$$

$$= T_F \Phi^{AE} \times \left[\mathcal{U}^{[-]\dagger}_{up} t^A_{ix} \mathcal{U}^{[-]}_{xs} \mathcal{U}^{[+]}_{ol} - \underbrace{\mathcal{U}^{[-]\dagger}_{uk} \mathcal{U}^{[+]}_{kl}} \mathcal{U}^{[-]}_{os} t^A_{ip} - \mathcal{U}^{[-]\dagger}_{up} \mathcal{U}^{[-]}_{is} t^A_{ok} \mathcal{U}^{[+]}_{kl} + \mathcal{U}^{[-]\dagger}_{uv} \mathcal{U}^{[-]}_{os} t^A_{vk} \mathcal{U}^{[+]}_{kl} \delta_{ip} \right] \times$$
$$\times \left[\underbrace{\mathcal{U}^{[+]}_{pn} \mathcal{U}^{[+]\dagger}_{ni}} \mathcal{U}^{[+]\dagger}_{so} t^E_{lu} - \mathcal{U}^{[+]}_{pu} \mathcal{U}^{[+]\dagger}_{ro} t^E_{lr} \mathcal{U}^{[+]\dagger}_{si} - \mathcal{U}^{[+]}_{pq} \mathcal{U}^{[+]\dagger}_{so} t^E_{qn} \mathcal{U}^{[+]\dagger}_{ni} \delta_{lu} + \mathcal{U}^{[+]}_{pu} \mathcal{U}^{[+]\dagger}_{lo} t^E_{sn} \mathcal{U}^{[+]\dagger}_{ni} \right] =$$

$$= T_F \Phi^{AE} \times \left[\mathcal{U}^{[-]\dagger}_{up} t^A_{ix} \mathcal{U}^{[-]}_{xs} \mathcal{U}^{[+]}_{ol} - \mathcal{U}^{[\square]}_{ul} \mathcal{U}^{[-]}_{os} t^A_{ip} - \mathcal{U}^{[-]\dagger}_{up} \mathcal{U}^{[-]}_{is} t^A_{ok} \mathcal{U}^{[+]}_{kl} + \mathcal{U}^{[-]\dagger}_{uv} \mathcal{U}^{[-]}_{os} t^A_{vk} \mathcal{U}^{[+]}_{kl} \delta_{ip} \right] \times$$
$$\times \left[\delta_{pi} \mathcal{U}^{[+]\dagger}_{so} t^E_{lu} - \mathcal{U}^{[+]}_{pu} \mathcal{U}^{[+]\dagger}_{ro} t^E_{lr} \mathcal{U}^{[+]\dagger}_{si} - \mathcal{U}^{[+]}_{pq} \mathcal{U}^{[+]\dagger}_{so} t^E_{qn} \mathcal{U}^{[+]\dagger}_{ni} \delta_{lu} + \mathcal{U}^{[+]}_{pu} \mathcal{U}^{[+]\dagger}_{lo} t^E_{sn} \mathcal{U}^{[+]\dagger}_{ni} \right] =$$

$$= T_F \Phi^{AE} \Bigg\{ \mathcal{U}^{[-]\dagger}_{up} t^A_{ix} \mathcal{U}^{[-]}_{xs} \mathcal{U}^{[+]}_{ol} \left[\delta_{pi} \mathcal{U}^{[+]\dagger}_{so} t^E_{lu} - \mathcal{U}^{[+]}_{pu} \mathcal{U}^{[+]\dagger}_{ro} t^E_{lr} \mathcal{U}^{[+]\dagger}_{si} - \mathcal{U}^{[+]}_{pq} \mathcal{U}^{[+]\dagger}_{so} t^E_{qn} \mathcal{U}^{[+]\dagger}_{ni} \delta_{lu} + \mathcal{U}^{[+]}_{pu} \mathcal{U}^{[+]\dagger}_{lo} t^E_{sn} \mathcal{U}^{[+]\dagger}_{ni} \right] -$$
$$- \mathcal{U}^{[\square]}_{ul} \mathcal{U}^{[-]}_{os} t^A_{ip} \left[\delta_{pi} \mathcal{U}^{[+]\dagger}_{so} t^E_{lu} - \mathcal{U}^{[+]}_{pu} \mathcal{U}^{[+]\dagger}_{ro} t^E_{lr} \mathcal{U}^{[+]\dagger}_{si} - \mathcal{U}^{[+]}_{pq} \mathcal{U}^{[+]\dagger}_{so} t^E_{qn} \mathcal{U}^{[+]\dagger}_{ni} \delta_{lu} + \mathcal{U}^{[+]}_{pu} \mathcal{U}^{[+]\dagger}_{lo} t^E_{sn} \mathcal{U}^{[+]\dagger}_{ni} \right] -$$
$$- \mathcal{U}^{[-]\dagger}_{up} \mathcal{U}^{[-]}_{is} t^A_{ok} \mathcal{U}^{[+]}_{kl} \left[\delta_{pi} \mathcal{U}^{[+]\dagger}_{so} t^E_{lu} - \mathcal{U}^{[+]}_{pu} \mathcal{U}^{[+]\dagger}_{ro} t^E_{lr} \mathcal{U}^{[+]\dagger}_{si} - \mathcal{U}^{[+]}_{pq} \mathcal{U}^{[+]\dagger}_{so} t^E_{qn} \mathcal{U}^{[+]\dagger}_{ni} \delta_{lu} + \mathcal{U}^{[+]}_{pu} \mathcal{U}^{[+]\dagger}_{lo} t^E_{sn} \mathcal{U}^{[+]\dagger}_{ni} \right] +$$
$$+ \mathcal{U}^{[-]\dagger}_{uv} \mathcal{U}^{[-]}_{os} t^A_{vk} \mathcal{U}^{[+]}_{kl} \delta_{ip} \left[\delta_{pi} \mathcal{U}^{[+]\dagger}_{so} t^E_{lu} - \mathcal{U}^{[+]}_{pu} \mathcal{U}^{[+]\dagger}_{ro} t^E_{lr} \mathcal{U}^{[+]\dagger}_{si} - \mathcal{U}^{[+]}_{pq} \mathcal{U}^{[+]\dagger}_{so} t^E_{qn} \mathcal{U}^{[+]\dagger}_{ni} \delta_{lu} + \mathcal{U}^{[+]}_{pu} \mathcal{U}^{[+]\dagger}_{lo} t^E_{sn} \mathcal{U}^{[+]\dagger}_{ni} \right] \Bigg\} =$$

$$= T_F \Phi^{AE} \Bigg\{ \Big[\mathcal{U}^{[-]\dagger}_{ui} t^A_{ix} \mathcal{U}^{[-]}_{xl} t^E_{lu} - \mathrm{Tr}\left(\mathcal{U}^{[\square]} \right) \mathrm{Tr}\left(t^A_{ix} \mathcal{U}^{[-]}_{xs} \mathcal{U}^{[+]\dagger}_{si} \right) \underbrace{\mathrm{Tr}\left(\mathcal{U}^{[+]}_{ol} t^E_{lr} \mathcal{U}^{[+]\dagger}_{ro} \right)} -$$
$$- t^A_{ix} \mathcal{U}^{[-]}_{xu} \mathcal{U}^{[\square]}_{uq} t^E_{qn} \mathcal{U}^{[+]\dagger}_{ni} + N_c \, \mathrm{Tr}\left(\mathcal{U}^{[\square]}_{up} \right) t^A_{ix} \mathcal{U}^{[-]}_{xs} t^E_{sn} \mathcal{U}^{[+]\dagger}_{ni} \Big] -$$
$$- \Big[\mathrm{Tr}\left(\mathcal{U}^{[\square]\dagger} \right) \mathrm{Tr}\left(t^A \right) t^E_{lu} \mathcal{U}^{[\square]}_{ul} - \mathcal{U}^{[\square]\dagger}_{oi} t^A_{ip} \mathcal{U}^{[+]}_{pu} \mathcal{U}^{[\square]}_{ul} t^E_{lr} \mathcal{U}^{[+]\dagger}_{ro} -$$
$$- \mathrm{Tr}\left(\mathcal{U}^{[\square]} \right) \mathrm{Tr}\left(\mathcal{U}^{[\square]\dagger} \right) t^A_{ip} \mathcal{U}^{[+]}_{pq} t^E_{qn} \mathcal{U}^{[+]\dagger}_{ni} + t^A_{ip} \mathcal{U}^{[+]}_{pu} \mathcal{U}^{[\square]}_{ul} \mathcal{U}^{[+]\dagger}_{lo} \mathcal{U}^{[-]}_{os} t^E_{sn} \mathcal{U}^{[+]\dagger}_{ni} \Big] -$$
$$- \Big[t^A_{ok} \mathcal{U}^{[+]}_{kl} t^E_{ls} \mathcal{U}^{[+]\dagger}_{so} - t^A_{ok} \mathcal{U}^{[+]}_{kl} t^E_{lr} \mathcal{U}^{[+]\dagger}_{ro} \mathrm{Tr}\left(\mathcal{U}^{[\square]} \right) \mathrm{Tr}\left(\mathcal{U}^{[\square]\dagger} \right) -$$
$$- t^A_{ok} \mathcal{U}^{[\square]}_{kp} \mathcal{U}^{[+]}_{pq} t^E_{qn} \mathcal{U}^{[+]\dagger}_{ni} \mathcal{U}^{[\square]\dagger}_{io} + \mathrm{Tr}\left(\mathcal{U}^{[\square]} \right) \mathrm{Tr}\left(t^A \right) \mathcal{U}^{[-]}_{is} t^E_{sn} \mathcal{U}^{[+]\dagger}_{ni} \Big] +$$
$$+ \Big[N_c \, \mathrm{Tr}\left(\mathcal{U}^{[\square]\dagger} \right) t^A_{vk} \mathcal{U}^{[+]}_{kl} t^E_{lu} \mathcal{U}^{[-]\dagger}_{uv} - t^A_{ok} \mathcal{U}^{[+]}_{kl} t^E_{lr} \mathcal{U}^{[+]\dagger}_{ro} -$$
$$- \mathcal{U}^{[-]\dagger}_{lv} t^A_{vk} \mathcal{U}^{[+]}_{kl} \mathrm{Tr}\left(\mathcal{U}^{[\square]\dagger} \right) \mathrm{Tr}\left(t^E \right) + t^A_{vk} \mathcal{U}^{[+]}_{kl} \mathcal{U}^{[\square]\dagger}_{ls} t^E_{sn} \mathcal{U}^{[-]\dagger}_{nv} \Big] \Bigg\} =$$

$$= T_F \Phi^{AE} \Big\{ \operatorname{Tr}\left(t^E \mathcal{U}^{[-]\dagger} t^A \mathcal{U}^{[-]}\right) - \operatorname{Tr}\left(t^E \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[-]} \underbrace{\mathcal{U}^{[\square]}}\right) + N_c \operatorname{Tr}\left(\mathcal{U}^{[\square]}\right) \operatorname{Tr}\left(t^E \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[-]}\right)$$
$$+ \operatorname{Tr}\left(t^E \mathcal{U}^{[\square]\dagger} \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[\square]} \mathcal{U}^{[+]}\right) + \operatorname{Tr}\left(\mathcal{U}^{[\square]}\right) \operatorname{Tr}\left(\mathcal{U}^{[\square]\dagger}\right) \operatorname{Tr}\left(t^E \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[+]}\right) - \operatorname{Tr}\left(t^E \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[+]}\right)$$
$$- \operatorname{Tr}\left(t^E \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[+]}\right) + \operatorname{Tr}\left(t^E \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[+]}\right) \operatorname{Tr}\left(\mathcal{U}^{[\square]}\right) \operatorname{Tr}\left(\mathcal{U}^{[\square]\dagger}\right) + \operatorname{Tr}\left(t^E \mathcal{U}^{[\square]\dagger} \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[\square]} \mathcal{U}^{[+]}\right)$$
$$+ N_c \operatorname{Tr}\left(\mathcal{U}^{[\square]\dagger}\right) \operatorname{Tr}\left(t^E \mathcal{U}^{[-]\dagger} t^A \mathcal{U}^{[+]}\right) - \operatorname{Tr}\left(t^E \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[+]}\right) + \operatorname{Tr}\left(t^E \mathcal{U}^{[-]\dagger} t^A \mathcal{U}^{[-]}\right) \Big\} =$$

$$= T_F \Phi^{AE} \Big\{ \operatorname{Tr}\left(t^E \mathcal{U}^{[-]\dagger} t^A \mathcal{U}^{[-]}\right) - \underline{\operatorname{Tr}\left(t^E \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[+]}\right)} + N_c \operatorname{Tr}\left(\mathcal{U}^{[\square]}\right) \operatorname{Tr}\left(t^E \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[-]}\right)$$
$$+ 2 \operatorname{Tr}\left(t^E \mathcal{U}^{[\square]\dagger} \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[\square]} \mathcal{U}^{[+]}\right) + \underline{\operatorname{Tr}\left(\mathcal{U}^{[\square]}\right) \operatorname{Tr}\left(\mathcal{U}^{[\square]\dagger}\right) \operatorname{Tr}\left(t^E \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[+]}\right)} - \underline{\operatorname{Tr}\left(t^E \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[+]}\right)}$$
$$- \underline{\operatorname{Tr}\left(t^E \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[+]}\right)} + \underline{\operatorname{Tr}\left(t^E \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[+]}\right) \operatorname{Tr}\left(\mathcal{U}^{[\square]}\right) \operatorname{Tr}\left(\mathcal{U}^{[\square]\dagger}\right)}$$
$$+ N_c \operatorname{Tr}\left(\mathcal{U}^{[\square]\dagger}\right) \operatorname{Tr}\left(t^E \mathcal{U}^{[-]\dagger} t^A \mathcal{U}^{[+]}\right) - \underline{\operatorname{Tr}\left(t^E \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[+]}\right)} + \operatorname{Tr}\left(t^E \mathcal{U}^{[-]\dagger} t^A \mathcal{U}^{[-]}\right) \Big\} =$$

$$= T_F \Phi^{AE} \Big\{ 2 \operatorname{Tr}\left(t^E \mathcal{U}^{[-]\dagger} t^A \mathcal{U}^{[-]}\right) + \operatorname{Tr}\left(t^E \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[+]}\right) \left(2 \operatorname{Tr}\left(\mathcal{U}^{[\square]}\right) \operatorname{Tr}\left(\mathcal{U}^{[\square]\dagger}\right) - 4\right) + N_c \operatorname{Tr}\left(\mathcal{U}^{[\square]}\right) \operatorname{Tr}\left(t^E \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[-]}\right)$$
$$+ 2 \operatorname{Tr}\left(t^E \mathcal{U}^{[\square]\dagger} \mathcal{U}^{[+]\dagger} t^A \mathcal{U}^{[\square]} \mathcal{U}^{[+]}\right)$$
$$+ N_c \operatorname{Tr}\left(\mathcal{U}^{[\square]\dagger}\right) \operatorname{Tr}\left(t^E \mathcal{U}^{[-]\dagger} t^A \mathcal{U}^{[+]}\right) \Big\} =$$

$$= T_F \langle H | \Big\{ 2 \operatorname{Tr}\left(F(\xi) \mathcal{U}^{[\square]\dagger} \mathcal{U}^{[+]\dagger} F(0) \mathcal{U}^{[\square]} \mathcal{U}^{[+]}\right) + 2 \operatorname{Tr}\left(F(\xi) \mathcal{U}^{[+]\dagger} F(0) \mathcal{U}^{[+]}\right) \left(\operatorname{Tr}\left(\mathcal{U}^{[\square]}\right) \operatorname{Tr}\left(\mathcal{U}^{[\square]\dagger}\right) - 2\right)$$
$$+ N_c \operatorname{Tr}\left(\mathcal{U}^{[\square]}\right) \operatorname{Tr}\left(F(\xi) \mathcal{U}^{[+]\dagger} F(0) \mathcal{U}^{[-]}\right) + + \operatorname{Tr}\left(F(\xi) \mathcal{U}^{[-]\dagger} F(0) \left\{ N_c \operatorname{Tr}\left(\mathcal{U}^{[\square]\dagger}\right) \mathcal{U}^{[+]} + 2 \mathcal{U}^{[-]} \right\}\right) \Big\} | H \rangle \,.$$

Color structure for the bare diagram (using $f^{ABC} = -f^{BAC}$ and $f^{ACD} f^{BCD} = N_c \delta^{AB}$):

$$i f^{ABC} i f^{BEF} i f^{HCI} i f^{HIF} = f^{ABC} f^{BEF} f^{CIH} f^{FHI} = -f^{ABC} f^{BEF} f^{CIH} f^{FIH} =$$
$$- N_c f^{ABC} f^{BEF} \delta^{CF} = -N_c f^{ABC} f^{BEC} = N_c f^{ABC} f^{EBC} = N_c^2 \delta^{AE} \,.$$

Next we divide by N_c^2 (from bare diagram) and multiply by T_F (from difference of conventions). T_F is set to $\frac{1}{2}$, so as a result we divide the above formula by $4N_c^2$ and the final result is

$$\Phi_g^{[\mathcal{U}]} \propto \langle H | \frac{1}{2N_c^2} \operatorname{Tr}\left[F(\xi) \mathcal{U}^{[\square]\dagger} \mathcal{U}^{[+]\dagger} F(0) \mathcal{U}^{[\square]} \mathcal{U}^{[+]}\right] + \frac{1}{2} \operatorname{Tr}\left[F(\xi) \mathcal{U}^{[+]\dagger} F(0) \mathcal{U}^{[+]}\right] \left(\frac{\operatorname{Tr}\left[\mathcal{U}^{[\square]}\right]}{N_c} \frac{\operatorname{Tr}\left[\mathcal{U}^{[\square]\dagger}\right]}{N_c} - \frac{2}{N_c^2}\right) +$$
$$+ \frac{1}{2} \operatorname{Tr}\left[F(\xi) \mathcal{U}^{[+]\dagger} F(0) \left\{ \frac{1}{2} \frac{\operatorname{Tr}\left[\mathcal{U}^{[\square]}\right]}{N_c} \mathcal{U}^{[-]} \right\}\right] + \frac{1}{2} \operatorname{Tr}\left[F(\xi) \mathcal{U}^{[-]\dagger} F(0) \left\{ \frac{1}{2} \frac{\operatorname{Tr}\left[\mathcal{U}^{[\square]\dagger}\right]}{N_c} \mathcal{U}^{[+]} + \frac{1}{N_c^2} \mathcal{U}^{[-]} \right\}\right] | H \rangle \,.$$

B Operator structures for 5 parton processes

$g_1g_5 \to g_2g_3g_4$	$g_1g_5 \to q_2\bar{q}_3g_4$	$g_1q_5 \to g_2g_3q_4$	$g_1\bar{q}_5 \to g_2g_3\bar{q}_4$
$\begin{pmatrix} \mathcal{A}(1,2,3,4,5) \\ \mathcal{A}(1,2,4,3,5) \\ \mathcal{A}(1,3,2,4,5) \\ \mathcal{A}(1,3,4,2,5) \\ \mathcal{A}(1,4,2,3,5) \\ \mathcal{A}(1,4,3,2,5) \end{pmatrix}$	$\begin{pmatrix} \mathcal{A}(2,1,4,5,3) \\ \mathcal{A}(2,1,5,4,3) \\ \mathcal{A}(2,4,1,5,3) \\ \mathcal{A}(2,4,5,1,3) \\ \mathcal{A}(2,5,1,4,3) \\ \mathcal{A}(2,5,4,1,3) \end{pmatrix}$	$\begin{pmatrix} \mathcal{A}(4,1,2,3,5) \\ \mathcal{A}(4,1,3,2,5) \\ \mathcal{A}(4,2,1,3,5) \\ \mathcal{A}(4,2,3,1,5) \\ \mathcal{A}(4,3,1,2,5) \\ \mathcal{A}(4,3,2,1,5) \end{pmatrix}$	$\begin{pmatrix} \mathcal{A}(5,1,2,3,4) \\ \mathcal{A}(5,1,3,2,4) \\ \mathcal{A}(5,2,1,3,4) \\ \mathcal{A}(5,2,3,1,4) \\ \mathcal{A}(5,3,1,2,4) \\ \mathcal{A}(5,3,2,1,4) \end{pmatrix}$
$g_1q_5 \to q_2\bar{q}_3q_4$		$g_1\bar{q}_5 \to q_2\bar{q}_3\bar{q}_4$	
$\begin{pmatrix} \mathcal{A}(2,3,4,1,5) \\ \mathcal{A}(2,1,3,4,5) \\ \mathcal{A}(2,5,4,1,3) \\ \mathcal{A}(2,1,5,4,3) \end{pmatrix}$		$\begin{pmatrix} \mathcal{A}(2,3,5,1,4) \\ \mathcal{A}(2,1,3,5,4) \\ \mathcal{A}(2,4,5,1,3) \\ \mathcal{A}(2,1,4,5,3) \end{pmatrix}$	

Table 6: Definitions of the vector of partial amplitudes $\vec{\mathcal{A}}$ for all five-parton processes. The subscripts in the sub-process indication correspond to the momenta enumeration.

$g_1g_5 \to g_2g_3g_4$	$\begin{pmatrix} \frac{1}{2N_c^2}+\frac{1}{4} & -\frac{1}{N_c^2} & -\frac{1}{N_c^2} & 0 & 0 & \frac{3}{4} & \frac{1}{2N_c^2} \\ \frac{1}{N_c^2} & -\frac{2}{N_c^2} & -\frac{1}{N_c^2} & -\frac{1}{2N_c^2} & -\frac{1}{2N_c^2} & 1 & \frac{1}{N_c^2} \\ \frac{4}{N_c^2} & -\frac{2}{N_c^2} & -\frac{1}{N_c^2} & \frac{1}{N_c^2} & -\frac{2}{N_c^2} & 1 & -\frac{2}{N_c^2} \\ N_c & -\frac{1}{4}N_c\left(N_c^2+2\right) & \frac{N_c}{4} & 0 & -\frac{3N_c}{4} & 0 & -\frac{N_c}{2} \end{pmatrix}$
$g_1g_5 \to q_2\bar{q}_3g_4$	$\begin{pmatrix} -\frac{N_c^2}{N_A^2} & 0 & -\frac{1}{N_A} & 0 & 0 & \frac{N_c^4}{N_A^2} & 0 \\ \frac{N_c^2}{N_A} & 0 & -\frac{1}{N_A} & 0 & 0 & 0 & 0 \\ 0 & -N_c^2 & 1 & 0 & 0 & 0 & 0 \\ 0 & -\frac{N_c^2}{F} & \frac{1}{F} & 0 & \frac{N_c^2}{F} & 0 & 0 \\ -N_c^2 & 0 & 1 & N_c^2 & 0 & 0 & 0 \\ 0 & -\frac{N_c^2}{F} & \frac{1}{F} & \frac{N_c^2}{F} & 0 & 0 & 0 \end{pmatrix}$
$g_1q_5 \to g_2g_3q_4$	$\begin{pmatrix} \frac{1}{N_A^2} & \frac{DN_c^2}{N_A^2} & 0 \\ -\frac{F}{N_A} & \frac{2N_c^2}{N_A} & 0 \\ -\frac{1}{N_A} & \frac{N_c^2}{N_A} & 0 \\ 1 & 0 & 0 \\ 1 & -N_c^2 & N_c^2 \\ \frac{1}{F} & 0 & \frac{N_c^2}{F} \end{pmatrix}$
$g_1q_5 \to q_2\bar{q}_3q_4$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Table 7: Matrices $\mathbf{M}$ of structures appearing in the five-parton processes ($D = N_c^2 - 2$, $F = N_c^2 + 1$). The subscripts in the sub-process indication correspond to the momenta enumeration.

C Operator structures for 6 parton processes

$g_1g_6 \to g_2g_3g_4g_5$	$g_1g_6 \to q_2\bar{q}_3g_4g_5$	$g_1q_6 \to g_2g_3g_4q_5$	$g_1\bar{q}_6 \to g_2g_3g_4\bar{q}_5$
$\mathcal{A}(1,2,3,4,5,6)$	$\mathcal{A}(2,1,4,5,6,3)$	$\mathcal{A}(5,1,2,3,4,6)$	$\mathcal{A}(6,1,2,3,4,5)$
$\mathcal{A}(1,2,3,5,4,6)$	$\mathcal{A}(2,1,4,6,5,3)$	$\mathcal{A}(5,1,2,4,3,6)$	$\mathcal{A}(6,1,2,4,3,5)$
$\mathcal{A}(1,2,4,3,5,6)$	$\mathcal{A}(2,1,5,4,6,3)$	$\mathcal{A}(5,1,3,2,4,6)$	$\mathcal{A}(6,1,3,2,4,5)$
$\mathcal{A}(1,2,4,5,3,6)$	$\mathcal{A}(2,1,5,6,4,3)$	$\mathcal{A}(5,1,3,4,2,6)$	$\mathcal{A}(6,1,3,4,2,5)$
$\mathcal{A}(1,2,5,3,4,6)$	$\mathcal{A}(2,1,6,4,5,3)$	$\mathcal{A}(5,1,4,2,3,6)$	$\mathcal{A}(6,1,4,2,3,5)$
$\mathcal{A}(1,2,5,4,3,6)$	$\mathcal{A}(2,1,6,5,4,3)$	$\mathcal{A}(5,1,4,3,2,6)$	$\mathcal{A}(6,1,4,3,2,5)$
$\mathcal{A}(1,3,2,4,5,6)$	$\mathcal{A}(2,4,1,5,6,3)$	$\mathcal{A}(5,2,1,3,4,6)$	$\mathcal{A}(6,2,1,3,4,5)$
$\mathcal{A}(1,3,2,5,4,6)$	$\mathcal{A}(2,4,1,6,5,3)$	$\mathcal{A}(5,2,1,4,3,6)$	$\mathcal{A}(6,2,1,4,3,5)$
$\mathcal{A}(1,3,4,2,5,6)$	$\mathcal{A}(2,4,5,1,6,3)$	$\mathcal{A}(5,2,3,1,4,6)$	$\mathcal{A}(6,2,3,1,4,5)$
$\mathcal{A}(1,3,4,5,2,6)$	$\mathcal{A}(2,4,5,6,1,3)$	$\mathcal{A}(5,2,3,4,1,6)$	$\mathcal{A}(6,2,3,4,1,5)$
$\mathcal{A}(1,3,5,2,4,6)$	$\mathcal{A}(2,4,6,1,5,3)$	$\mathcal{A}(5,2,4,1,3,6)$	$\mathcal{A}(6,2,4,1,3,5)$
$\mathcal{A}(1,3,5,4,2,6)$	$\mathcal{A}(2,4,6,5,1,3)$	$\mathcal{A}(5,2,4,3,1,6)$	$\mathcal{A}(6,2,4,3,1,5)$
$\mathcal{A}(1,4,2,3,5,6)$	$\mathcal{A}(2,5,1,4,6,3)$	$\mathcal{A}(5,3,1,2,4,6)$	$\mathcal{A}(6,3,1,2,4,5)$
$\mathcal{A}(1,4,2,5,3,6)$	$\mathcal{A}(2,5,1,6,4,3)$	$\mathcal{A}(5,3,1,4,2,6)$	$\mathcal{A}(6,3,1,4,2,5)$
$\mathcal{A}(1,4,3,2,5,6)$	$\mathcal{A}(2,5,4,1,6,3)$	$\mathcal{A}(5,3,2,1,4,6)$	$\mathcal{A}(6,3,2,1,4,5)$
$\mathcal{A}(1,4,3,5,2,6)$	$\mathcal{A}(2,5,4,6,1,3)$	$\mathcal{A}(5,3,2,4,1,6)$	$\mathcal{A}(6,3,2,4,1,5)$
$\mathcal{A}(1,4,5,2,3,6)$	$\mathcal{A}(2,5,6,1,4,3)$	$\mathcal{A}(5,3,4,1,2,6)$	$\mathcal{A}(6,3,4,1,2,5)$
$\mathcal{A}(1,4,5,3,2,6)$	$\mathcal{A}(2,5,6,4,1,3)$	$\mathcal{A}(5,3,4,2,1,6)$	$\mathcal{A}(6,3,4,2,1,5)$
$\mathcal{A}(1,5,2,3,4,6)$	$\mathcal{A}(2,6,1,4,5,3)$	$\mathcal{A}(5,4,1,2,3,6)$	$\mathcal{A}(6,4,1,2,3,5)$
$\mathcal{A}(1,5,2,4,3,6)$	$\mathcal{A}(2,6,1,5,4,3)$	$\mathcal{A}(5,4,1,3,2,6)$	$\mathcal{A}(6,4,1,3,2,5)$
$\mathcal{A}(1,5,3,2,4,6)$	$\mathcal{A}(2,6,4,1,5,3)$	$\mathcal{A}(5,4,2,1,3,6)$	$\mathcal{A}(6,4,2,1,3,5)$
$\mathcal{A}(1,5,3,4,2,6)$	$\mathcal{A}(2,6,4,5,1,3)$	$\mathcal{A}(5,4,2,3,1,6)$	$\mathcal{A}(6,4,2,3,1,5)$
$\mathcal{A}(1,5,4,2,3,6)$	$\mathcal{A}(2,6,5,1,4,3)$	$\mathcal{A}(5,4,3,1,2,6)$	$\mathcal{A}(6,4,3,1,2,5)$
$\mathcal{A}(1,5,4,3,2,6)$	$\mathcal{A}(2,6,5,4,1,3)$	$\mathcal{A}(5,4,3,2,1,6)$	$\mathcal{A}(6,4,3,2,1,5)$

$g_1g_6 \to q_2\bar{q}_3q_4\bar{q}_5$	$g_1q_6 \to g_2q_3\bar{q}_4q_5$	$g_1\bar{q}_6 \to g_2q_3\bar{q}_4\bar{q}_5$
$\mathcal{A}(2,3,4,1,6,5)$	$\mathcal{A}(3,4,5,1,2,6)$	$\mathcal{A}(3,4,6,1,2,5)$
$\mathcal{A}(2,1,3,4,6,5)$	$\mathcal{A}(3,1,4,5,2,6)$	$\mathcal{A}(3,1,4,6,2,5)$
$\mathcal{A}(2,1,6,3,4,5)$	$\mathcal{A}(3,1,2,4,5,6)$	$\mathcal{A}(3,1,2,4,6,5)$
$\mathcal{A}(2,3,4,6,1,5)$	$\mathcal{A}(3,4,5,2,1,6)$	$\mathcal{A}(3,4,6,2,1,5)$
$\mathcal{A}(2,6,3,4,1,5)$	$\mathcal{A}(3,2,4,5,1,6)$	$\mathcal{A}(3,2,4,6,1,5)$
$\mathcal{A}(2,6,1,3,4,5)$	$\mathcal{A}(3,2,1,4,5,6)$	$\mathcal{A}(3,2,1,4,6,5)$
$\mathcal{A}(2,5,4,1,6,3)$	$\mathcal{A}(3,6,5,1,2,4)$	$\mathcal{A}(3,5,6,1,2,4)$
$\mathcal{A}(2,1,5,4,6,3)$	$\mathcal{A}(3,1,6,5,2,4)$	$\mathcal{A}(3,1,5,6,2,4)$
$\mathcal{A}(2,1,6,5,4,3)$	$\mathcal{A}(3,1,2,6,5,4)$	$\mathcal{A}(3,1,2,5,6,4)$
$\mathcal{A}(2,5,4,6,1,3)$	$\mathcal{A}(3,6,5,2,1,4)$	$\mathcal{A}(3,5,6,2,1,4)$
$\mathcal{A}(2,6,5,4,1,3)$	$\mathcal{A}(3,2,6,5,1,4)$	$\mathcal{A}(3,2,5,6,1,4)$
$\mathcal{A}(2,6,1,5,4,3)$	$\mathcal{A}(3,2,1,6,5,4)$	$\mathcal{A}(3,2,1,5,6,4)$

Table 8: Definition of the vector of partial amplitudes $\vec{\mathcal{A}}$ for all six-parton processes. The subscripts in the sub-process indication correspond to the momenta enumeration.

$g_1 g_6 \to g_2 g_3 g_4 g_5$

$$\begin{pmatrix}
\frac{1}{4N_c^2}+\frac{1}{8} & 0 & -\frac{F}{N_c^4} & \frac{1}{2N_c^4} & \frac{1}{2N_c^4} & \frac{1}{2N_c^2}+\frac{7}{8} & \frac{1}{4N_c^2} \\
0 & \frac{1}{N_c^2} & -\frac{N_c^2+2}{N_c^4} & \frac{1}{N_c^4} & \frac{1}{N_c^4} & \frac{1}{N_c^2}+1 & 0 \\
\frac{3}{N_c^2} & \frac{1}{N_c^2} & -\frac{N_c^2+2}{N_c^4} & \frac{1}{N_c^4} & \frac{1}{N_c^4} & \frac{1}{N_c^2}+1 & -\frac{3}{N_c^2} \\
\frac{N_c^2}{2} & \frac{3N_c^2}{4} & -\frac{1}{2} & \frac{1}{8}\left(N_c^2+2\right) & \frac{1}{8}\left(N_c^2+2\right) & \frac{N_c^2}{4} & -N_c^2 \\
-\frac{1}{N_c^2} & \frac{4}{N_c^2} & -\frac{N_c^2+4}{N_c^4} & \frac{N_c^2+4}{2N_c^4} & \frac{N_c^2+4}{2N_c^4} & \frac{2}{N_c^2}+1 & -\frac{1}{N_c^2} \\
\frac{6}{N_c^2} & \frac{4}{N_c^2} & -\frac{N_c^2+4}{N_c^4} & \frac{N_c^2+2}{N_c^4} & \frac{2}{N_c^4} & \frac{2}{N_c^2}+1 & -\frac{8}{N_c^2} \\
\frac{4}{N_c^2} & \frac{4}{N_c^2} & -\frac{N_c^2+4}{N_c^4} & \frac{2}{N_c^4} & \frac{N_c^2+2}{N_c^4} & \frac{2}{N_c^2}+1 & -\frac{6}{N_c^2} \\
\frac{N_c^2}{4} & \frac{3N_c^2}{4} & -\frac{1}{2} & \frac{1}{4} & \frac{F}{4} & \frac{N_c^2}{4} & -\frac{3N_c^2}{4} \\
\frac{8}{N_c^2+12} & -\frac{2}{N_c^2+12} & -\frac{N_c^2+8}{N_c^2(N_c^2+12)} & \frac{N_c^2+4}{N_c^2(N_c^2+12)} & \frac{2\left(N_c^2+2\right)}{N_c^2(N_c^2+12)} & \frac{N_c^2+4}{N_c^2+12} & -\frac{2}{N_c^2+12} \\
\frac{1}{3} & 0 & -\frac{2}{3N_c^2} & \frac{1}{3N_c^2} & \frac{1}{3}\left(\frac{1}{N_c^2}+1\right) & \frac{1}{3} & 0 \\
0 & \frac{1}{12}\left(N_c^2+2\right) & -\frac{N_c^2+8}{12N_c^2} & \frac{1}{3N_c^2} & \frac{1}{3N_c^2}+\frac{7}{12} & \frac{1}{3} & \frac{1}{6}
\end{pmatrix}$$

Table 9: Matrices $\mathbf{M}$ of structures appearing in the six-parton processes (part I) ($F = N_c^2+1$). The subscripts in the sub-process indication correspond to the momenta enumeration.

$g_1 g_6 \to q_2 \bar{q}_3 g_4 g_5$

$$\begin{pmatrix}
\frac{N_c^2}{N_A^3} & 0 & -\frac{1}{N_A} & 0 & 0 & \frac{DN_c^4}{N_A^3} & 0 \\
-\frac{N_c^2}{N_A^2} & 0 & -\frac{1}{N_A} & 0 & 0 & \frac{N_c^4}{N_A^2} & 0 \\
-\frac{FN_c^2}{N_A^2} & 0 & -\frac{1}{N_A} & 0 & 0 & \frac{2N_c^4}{N_A^2} & 0 \\
\frac{N_c^2}{N_A} & 0 & -\frac{1}{N_A} & 0 & 0 & -\frac{N_c^4}{N_A} & \frac{N_c^4}{N_A} \\
\frac{N_c^2}{N_A} & 0 & -\frac{1}{N_A} & 0 & 0 & 0 & 0 \\
\frac{N_c^2}{N_c^4-1} & 0 & -\frac{1}{N_A} & 0 & 0 & 0 & \frac{N_c^4}{N_c^4-1} \\
0 & -N_c^2 & 1 & 0 & 0 & 0 & 0 \\
0 & -\frac{N_c^2}{F} & \frac{1}{F} & 0 & \frac{N_c^2}{F} & 0 & 0 \\
-N_c^2 & 0 & 1 & N_c^2 & 0 & 0 & 0 \\
0 & -\frac{N_c^2}{F} & \frac{1}{F} & \frac{N_c^2}{F} & 0 & 0 & 0 \\
-N_c^2 & -N_c^4 & F & 0 & 0 & 0 & 0 \\
\frac{N_c^4}{L} & \frac{N_c^2}{L} & -\frac{F}{L} & 0 & 0 & 0 & 0 \\
0 & \frac{N_c^2}{K} & -\frac{F}{K} & 0 & -\frac{N_c^2}{K} & 0 & \frac{N_c^4}{K} \\
-FN_c^2 & 0 & 1 & N_c^2 & 0 & N_c^4 & 0 \\
-\frac{N_c^2}{F} & -\frac{N_c^4}{F} & 1 & \frac{N_c^2}{F} & 0 & 0 & 0 \\
\frac{N_c^4}{K} & \frac{N_c^2}{K} & -\frac{F}{K} & -\frac{N_c^2}{K} & 0 & 0 & 0 \\
0 & -\frac{FN_c^2}{3N_c^2+1} & \frac{F}{3N_c^2+1} & \frac{N_c^2}{3N_c^2+1} & \frac{N_c^2}{3N_c^2+1} & 0 & 0 \\
\frac{N_c^2}{L} & 0 & -\frac{1}{L} & \frac{DN_c^2}{L} & 0 & 0 & 0 \\
0 & \frac{N_c^2}{K} & -\frac{1}{K} & \frac{DN_c^2}{K} & 0 & 0 & 0 \\
\frac{FN_c^2}{K} & 0 & -\frac{F}{K} & -\frac{2N_c^2}{K} & 0 & 0 & 0 \\
0 & -\frac{FN_c^2}{3N_c^2+1} & \frac{F}{3N_c^2+1} & \frac{2N_c^2}{3N_c^2+1} & 0 & 0 & 0 \\
0 & \frac{N_c^2}{K} & -\frac{1}{K} & 0 & \frac{DN_c^2}{K} & 0 & 0 \\
0 & -\frac{FN_c^2}{3N_c^2+1} & \frac{F}{3N_c^2+1} & 0 & \frac{2N_c^2}{3N_c^2+1} & 0 & 0 \\
0 & \frac{N_c^2}{L} & -\frac{F}{L} & 0 & 0 & 0 & \frac{N_c^4}{L}
\end{pmatrix}$$

Table 10: Matrices **M** of structures appearing in the six-parton processes (part II) ($D = N_c^2 - 2$, $F = N_c^2 + 1$, $K = N_c^4 - 2N_c^2 - 1$, $L = N_c^4 - N_c^2 - 1$). The subscripts in the sub-process indication correspond to the momenta enumeration.

$g_1 g_6 \to g_2 g_3 g_4 q_5$
$\begin{pmatrix} -\frac{1}{N_A^3} & \frac{N_c^2\left(N_c^4-3N_c^2+3\right)}{N_A^3} & 0 \\ \frac{F}{N_A^2} & \frac{N_c^2\left(N_c^2-3\right)}{N_A^2} & 0 \\ \frac{K}{N_A} & -\frac{N_c^2\left(N_c^2-3\right)}{N_A} & 0 \\ \frac{3N_c^2+1}{1-N_c^4} & \frac{N_c^2\left(N_c^2+3\right)}{N_c^4-1} & 0 \\ \frac{1}{N_A^2} & \frac{DN_c^2}{N_A^2} & 0 \\ -\frac{F}{N_A} & \frac{2N_c^2}{N_A} & 0 \\ -\frac{1}{N_A} & \frac{N_c^2}{N_A} & 0 \\ 1 & 0 & 0 \\ 1 & -N_c^2 & N_c^2 \\ -\frac{1}{N_A} & \frac{2N_c^2}{N_A} & -\frac{N_c^2}{N_A} \\ F & -2N_c^2 & N_c^2 \\ -\frac{1}{L} & \frac{N_c^2}{L} & \frac{DN_c^2}{L} \\ 1 & DN_c^2 & -DN_c^2 \\ 1 & -\frac{2N_c^2}{F} & \frac{2N_c^2}{F} \\ 1 & -\frac{N_c^2}{F} & \frac{N_c^2}{F} \\ -\frac{F}{K} & \frac{FN_c^2}{K} & -\frac{2N_c^2}{K} \\ \frac{1}{F} & 0 & \frac{N_c^2}{F} \\ \frac{1}{1-N_c^4} & \frac{N_c^2}{N_A} & \frac{N_c^2}{1-N_c^4} \\ \frac{1}{F} & -\frac{N_c^2}{F} & \frac{2N_c^2}{F} \\ -\frac{1}{K} & 0 & \frac{DN_c^2}{K} \\ -\frac{1}{K} & \frac{N_c^2}{K} & \frac{N_c^2\left(N_c^2-3\right)}{K} \\ \frac{F}{3N_c^2+1} & 0 & \frac{2N_c^2}{3N_c^2+1} \\ -\frac{F}{K} & \frac{N_c^4}{K} & -\frac{N_c^2}{K} \end{pmatrix}$

Table 11: Matrices $\mathbf{M}$ of structures appearing in the six-parton processes (part III) ($D = N_c^2-2$, $F = N_c^2+1$, $K = N_c^4-2N_c^2-1$, $L = N_c^4-N_c^2-1$). The subscripts in the sub-process indication correspond to the momenta enumeration.

$g_1 g_6 \to q_2 \bar{q}_3 q_4 \bar{q}_5$	$g_1 g_6 \to g_2 q_3 \bar{q}_4 q_5$
$\begin{pmatrix} \frac{N_c^2}{N_A} & 0 & -\frac{1}{N_A} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -N_c^2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{N_A} & 0 & 0 & \frac{N_c^2}{N_A} & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{1}{N_A} & 0 & 0 & 0 & \frac{N_c^2}{N_A} \end{pmatrix}$	$\begin{pmatrix} -\frac{1}{N_A} & \frac{N_c^2}{N_A} & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{4N_c^2} & -\frac{1}{4N_c^2} & 0 \\ 0 & -\frac{1}{4N_c^2} & \frac{1}{4N_c^2} \\ 0 & \frac{N_c^2}{N_A} & -\frac{1}{N_A} \\ \frac{1}{4} & -\frac{1}{4} & 0 \\ 0 & -\frac{1}{4} & \frac{1}{4} \end{pmatrix}$

Table 12: Matrices $\mathbf{M}$ of structures appearing in the six-parton processes (part IV). The subscripts in the sub-process indication correspond to the momenta enumeration.

C.1 $g(k_1)\,g(k_6) \to g(k_2)\,g(k_3)\,g(k_4)\,g(k_5)$

$$T_1 = \begin{pmatrix} \Phi_1 & \Phi_2 & \Phi_2 & \Phi_3 & \Phi_3 & \Phi_4^* \\ \Phi_2 & \Phi_1 & \Phi_3 & \Phi_4^* & \Phi_2 & \Phi_3 \\ \Phi_2 & \Phi_3 & \Phi_1 & \Phi_2 & \Phi_4^* & \Phi_3 \\ \Phi_3 & \Phi_4^* & \Phi_2 & \Phi_1 & \Phi_3 & \Phi_2 \\ \Phi_3 & \Phi_2 & \Phi_4^* & \Phi_3 & \Phi_1 & \Phi_2 \\ \Phi_4^* & \Phi_3 & \Phi_3 & \Phi_2 & \Phi_2 & \Phi_1 \end{pmatrix}, \quad T_2 = \begin{pmatrix} \Phi_2 & \Phi_5 & \Phi_3 & \Phi_6 & \Phi_7 & \Phi_8^* \\ \Phi_5 & \Phi_2 & \Phi_7 & \Phi_8^* & \Phi_3 & \Phi_6 \\ \Phi_3 & \Phi_7 & \Phi_4^* & \Phi_8^* & \Phi_9 & \Phi_{10} \\ \Phi_6 & \Phi_8^* & \Phi_8^* & \Phi_{10} & \Phi_{10} & \Phi_{11} \\ \Phi_7 & \Phi_3 & \Phi_9 & \Phi_{10} & \Phi_4^* & \Phi_8^* \\ \Phi_8^* & \Phi_6 & \Phi_{10} & \Phi_{11} & \Phi_8^* & \Phi_{10} \end{pmatrix}, \tag{C.1}$$

$$T_3 = \begin{pmatrix} \Phi_3 & \Phi_7 & \Phi_4^* & \Phi_8^* & \Phi_9 & \Phi_{10} \\ \Phi_6 & \Phi_8^* & \Phi_8^* & \Phi_{10} & \Phi_{10} & \Phi_{11} \\ \Phi_2 & \Phi_5 & \Phi_3 & \Phi_6 & \Phi_7 & \Phi_8^* \\ \Phi_5 & \Phi_2 & \Phi_7 & \Phi_8^* & \Phi_3 & \Phi_6 \\ \Phi_8^* & \Phi_6 & \Phi_{10} & \Phi_{11} & \Phi_8^* & \Phi_{10} \\ \Phi_7 & \Phi_3 & \Phi_9 & \Phi_{10} & \Phi_4^* & \Phi_8^* \end{pmatrix}, \quad T_4 = \begin{pmatrix} \Phi_6 & \Phi_8^* & \Phi_8^* & \Phi_{10} & \Phi_{10} & \Phi_{11} \\ \Phi_3 & \Phi_7 & \Phi_4^* & \Phi_8^* & \Phi_9 & \Phi_{10} \\ \Phi_8^* & \Phi_6 & \Phi_{10} & \Phi_{11} & \Phi_8^* & \Phi_{10} \\ \Phi_7 & \Phi_3 & \Phi_9 & \Phi_{10} & \Phi_4^* & \Phi_8^* \\ \Phi_2 & \Phi_5 & \Phi_3 & \Phi_6 & \Phi_7 & \Phi_8^* \\ \Phi_5 & \Phi_2 & \Phi_7 & \Phi_8^* & \Phi_3 & \Phi_6 \end{pmatrix},$$

$$T_5 = \begin{pmatrix} \Phi_4^* & \Phi_8^* & \Phi_3 & \Phi_7 & \Phi_{10} & \Phi_9 \\ \Phi_8^* & \Phi_{10} & \Phi_6 & \Phi_8^* & \Phi_{11} & \Phi_{10} \\ \Phi_3 & \Phi_6 & \Phi_2 & \Phi_5 & \Phi_8^* & \Phi_7 \\ \Phi_7 & \Phi_8^* & \Phi_5 & \Phi_2 & \Phi_6 & \Phi_3 \\ \Phi_{10} & \Phi_{11} & \Phi_8^* & \Phi_6 & \Phi_{10} & \Phi_8^* \\ \Phi_9 & \Phi_{10} & \Phi_7 & \Phi_3 & \Phi_8^* & \Phi_4^* \end{pmatrix}, \quad T_6 = \begin{pmatrix} \Phi_8^* & \Phi_{10} & \Phi_6 & \Phi_8^* & \Phi_{11} & \Phi_{10} \\ \Phi_4^* & \Phi_8^* & \Phi_3 & \Phi_7 & \Phi_{10} & \Phi_9 \\ \Phi_{10} & \Phi_{11} & \Phi_8^* & \Phi_6 & \Phi_{10} & \Phi_8^* \\ \Phi_9 & \Phi_{10} & \Phi_7 & \Phi_3 & \Phi_8^* & \Phi_4^* \\ \Phi_3 & \Phi_6 & \Phi_2 & \Phi_5 & \Phi_8^* & \Phi_7 \\ \Phi_7 & \Phi_8^* & \Phi_5 & \Phi_2 & \Phi_6 & \Phi_3 \end{pmatrix},$$

$$T_7 = T_2^M\,. \tag{C.2}$$

T_2^M denotes a mirror reflection of the matrix T_2 with respect to the anti-diagonal, which can be written as a similarity transformation

$$T_2^M = JT_2J\,,$$

with

$$J = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The same relation holds between block matrices for color factors (Eq. E.4). Nonetheless, for convenience we list explicitly elements of T_7 matrix:

$$T_7 = \begin{pmatrix} \Phi_{10} & \Phi_8^* & \Phi_{11} & \Phi_{10} & \Phi_6 & \Phi_8^* \\ \Phi_8^* & \Phi_4^* & \Phi_{10} & \Phi_9 & \Phi_3 & \Phi_7 \\ \Phi_{11} & \Phi_{10} & \Phi_{10} & \Phi_8^* & \Phi_8^* & \Phi_6 \\ \Phi_{10} & \Phi_9 & \Phi_8^* & \Phi_4^* & \Phi_7 & \Phi_3 \\ \Phi_6 & \Phi_3 & \Phi_8^* & \Phi_7 & \Phi_2 & \Phi_5 \\ \Phi_8^* & \Phi_7 & \Phi_6 & \Phi_3 & \Phi_5 & \Phi_2 \end{pmatrix}.$$

C.2 $g(k_1)\,g(k_6) \to q(k_2)\,\bar{q}(k_3)\,g(k_4)\,g(k_5)$

$$T_1 = \begin{pmatrix} \Phi_1 & \Phi_2 & \Phi_3 & \Phi_4 & \Phi_5 & \Phi_5 \\ \Phi_2 & \Phi_2 & \Phi_4 & \Phi_6 & \Phi_5 & \Phi_5 \\ \Phi_3 & \Phi_4 & \Phi_1 & \Phi_2 & \Phi_5 & \Phi_5 \\ \Phi_4 & \Phi_6 & \Phi_2 & \Phi_2 & \Phi_5 & \Phi_5 \\ \Phi_5 & \Phi_5 & \Phi_5 & \Phi_5 & \Phi_5 & \Phi_5 \\ \Phi_5 & \Phi_5 & \Phi_5 & \Phi_5 & \Phi_5 & \Phi_5 \end{pmatrix}, \quad T_2 = \begin{pmatrix} \Phi_2 & \Phi_5 & \Phi_5 & \Phi_7 & \Phi_7 & \Phi_8 \\ \Phi_5 & \Phi_5 & \Phi_9 & \Phi_{10} & \Phi_7 & \Phi_7 \\ \Phi_4 & \Phi_{11} & \Phi_5 & \Phi_7 & \Phi_{12} & \Phi_{13} \\ \Phi_{14} & \Phi_{15} & \Phi_{15} & \Phi_{16} & \Phi_{16} & \Phi_{17} \\ \Phi_9 & \Phi_9 & \Phi_{18} & \Phi_{19} & \Phi_{10} & \Phi_{10} \\ \Phi_{15} & \Phi_9 & \Phi_{20} & \Phi_{21} & \Phi_{10} & \Phi_{16} \end{pmatrix}, \tag{C.3}$$

$$T_3 = \begin{pmatrix} \Phi_4 & \Phi_{11} & \Phi_5 & \Phi_7 & \Phi_{12} & \Phi_{13} \\ \Phi_{14} & \Phi_{15} & \Phi_{15} & \Phi_{16} & \Phi_{16} & \Phi_{17} \\ \Phi_2 & \Phi_5 & \Phi_5 & \Phi_7 & \Phi_7 & \Phi_8 \\ \Phi_5 & \Phi_5 & \Phi_9 & \Phi_{10} & \Phi_7 & \Phi_7 \\ \Phi_{15} & \Phi_9 & \Phi_{20} & \Phi_{21} & \Phi_{10} & \Phi_{16} \\ \Phi_9 & \Phi_9 & \Phi_{18} & \Phi_{19} & \Phi_{10} & \Phi_{10} \end{pmatrix}, \quad T_4 = \begin{pmatrix} \Phi_7 & \Phi_7 & \Phi_8 & \Phi_{22} & \Phi_{13} & \Phi_{23} \\ \Phi_7 & \Phi_7 & \Phi_8 & \Phi_8 & \Phi_{24} & \Phi_{13} \\ \Phi_7 & \Phi_7 & \Phi_{13} & \Phi_{23} & \Phi_8 & \Phi_{22} \\ \Phi_7 & \Phi_7 & \Phi_{24} & \Phi_{13} & \Phi_8 & \Phi_8 \\ \Phi_7 & \Phi_7 & \Phi_7 & \Phi_7 & \Phi_7 & \Phi_7 \\ \Phi_7 & \Phi_7 & \Phi_7 & \Phi_7 & \Phi_7 & \Phi_7 \end{pmatrix}, \tag{C.4}$$

$$T_5 = \begin{pmatrix} \Phi_2 & \Phi_5 & \Phi_5 & \Phi_7 & \Phi_7 & \Phi_8 \\ \Phi_5 & \Phi_5 & \Phi_9 & \Phi_{10} & \Phi_7 & \Phi_7 \\ \Phi_5 & \Phi_9 & \Phi_5 & \Phi_7 & \Phi_{10} & \Phi_7 \\ \Phi_7 & \Phi_{10} & \Phi_7 & \Phi_5 & \Phi_9 & \Phi_5 \\ \Phi_7 & \Phi_7 & \Phi_{10} & \Phi_9 & \Phi_5 & \Phi_5 \\ \Phi_8 & \Phi_7 & \Phi_7 & \Phi_5 & \Phi_5 & \Phi_2 \end{pmatrix}, \quad T_6 = \begin{pmatrix} \Phi_6 & \Phi_{15} & \Phi_5 & \Phi_7 & \Phi_{16} & \Phi_{24} \\ \Phi_{15} & \Phi_{20} & \Phi_9 & \Phi_{10} & \Phi_{21} & \Phi_{16} \\ \Phi_5 & \Phi_9 & \Phi_5 & \Phi_7 & \Phi_{10} & \Phi_7 \\ \Phi_7 & \Phi_{10} & \Phi_7 & \Phi_5 & \Phi_9 & \Phi_5 \\ \Phi_{16} & \Phi_{21} & \Phi_{10} & \Phi_9 & \Phi_{20} & \Phi_{15} \\ \Phi_{24} & \Phi_{16} & \Phi_7 & \Phi_5 & \Phi_{15} & \Phi_6 \end{pmatrix}, \tag{C.5}$$

$$T_7 = \begin{pmatrix} \Phi_{10} & \Phi_{16} & \Phi_7 & \Phi_8 & \Phi_{17} & \Phi_{13} \\ \Phi_{10} & \Phi_{10} & \Phi_7 & \Phi_7 & \Phi_{16} & \Phi_{12} \\ \Phi_{19} & \Phi_{21} & \Phi_{10} & \Phi_7 & \Phi_{16} & \Phi_7 \\ \Phi_{18} & \Phi_{20} & \Phi_9 & \Phi_5 & \Phi_{15} & \Phi_5 \\ \Phi_9 & \Phi_9 & \Phi_5 & \Phi_5 & \Phi_{15} & \Phi_{11} \\ \Phi_9 & \Phi_{15} & \Phi_5 & \Phi_2 & \Phi_{14} & \Phi_4 \end{pmatrix}, \quad T_8 = \begin{pmatrix} \Phi_{16} & \Phi_{10} & \Phi_{17} & \Phi_{13} & \Phi_7 & \Phi_8 \\ \Phi_{10} & \Phi_{10} & \Phi_{16} & \Phi_{12} & \Phi_7 & \Phi_7 \\ \Phi_{21} & \Phi_{19} & \Phi_{16} & \Phi_7 & \Phi_{10} & \Phi_7 \\ \Phi_{20} & \Phi_{18} & \Phi_{15} & \Phi_5 & \Phi_9 & \Phi_5 \\ \Phi_9 & \Phi_9 & \Phi_{15} & \Phi_{11} & \Phi_5 & \Phi_5 \\ \Phi_{15} & \Phi_9 & \Phi_{14} & \Phi_4 & \Phi_5 & \Phi_2 \end{pmatrix}, \tag{C.6}$$

$$T_8^{\intercal} = T_2^M\,, \qquad T_9 = \begin{pmatrix} \Phi_5 & \Phi_5 & \Phi_5 & \Phi_5 & \Phi_5 & \Phi_5 \\ \Phi_5 & \Phi_5 & \Phi_5 & \Phi_5 & \Phi_5 & \Phi_5 \\ \Phi_5 & \Phi_5 & \Phi_2 & \Phi_2 & \Phi_6 & \Phi_4 \\ \Phi_5 & \Phi_5 & \Phi_2 & \Phi_1 & \Phi_4 & \Phi_3 \\ \Phi_5 & \Phi_5 & \Phi_6 & \Phi_4 & \Phi_2 & \Phi_2 \\ \Phi_5 & \Phi_5 & \Phi_4 & \Phi_3 & \Phi_2 & \Phi_1 \end{pmatrix} = T_1^M\,. \tag{C.7}$$

C.3 $g(k_1)\,q(k_6) \to g(k_2)\,g(k_3)\,g(k_4)\,q(k_5)$

$$T_1 = \begin{pmatrix} \Phi_1 & \Phi_2 & \Phi_2 & \Phi_3 & \Phi_3 & \Phi_4 \\ \Phi_2 & \Phi_1 & \Phi_3 & \Phi_4 & \Phi_2 & \Phi_3 \\ \Phi_2 & \Phi_3 & \Phi_1 & \Phi_2 & \Phi_4 & \Phi_3 \\ \Phi_3 & \Phi_4 & \Phi_2 & \Phi_1 & \Phi_3 & \Phi_2 \\ \Phi_3 & \Phi_2 & \Phi_4 & \Phi_3 & \Phi_1 & \Phi_2 \\ \Phi_4 & \Phi_3 & \Phi_3 & \Phi_2 & \Phi_2 & \Phi_1 \end{pmatrix}, \quad T_2 = \begin{pmatrix} \Phi_5 & \Phi_6 & \Phi_7 & \Phi_8 & \Phi_9 & \Phi_8 \\ \Phi_6 & \Phi_5 & \Phi_9 & \Phi_8 & \Phi_7 & \Phi_8 \\ \Phi_{10} & \Phi_{11} & \Phi_7 & \Phi_8 & \Phi_{12} & \Phi_8 \\ \Phi_{13} & \Phi_{14} & \Phi_{15} & \Phi_8 & \Phi_{16} & \Phi_8 \\ \Phi_{11} & \Phi_{10} & \Phi_{12} & \Phi_8 & \Phi_7 & \Phi_8 \\ \Phi_{14} & \Phi_{13} & \Phi_{16} & \Phi_8 & \Phi_{15} & \Phi_8 \end{pmatrix}, \tag{C.8}$$

$$T_3 = \begin{pmatrix} \Phi_{10} & \Phi_{11} & \Phi_7 & \Phi_8 & \Phi_{12} & \Phi_8 \\ \Phi_{13} & \Phi_{14} & \Phi_{15} & \Phi_8 & \Phi_{16} & \Phi_8 \\ \Phi_5 & \Phi_6 & \Phi_7 & \Phi_8 & \Phi_9 & \Phi_8 \\ \Phi_6 & \Phi_5 & \Phi_9 & \Phi_8 & \Phi_7 & \Phi_8 \\ \Phi_{14} & \Phi_{13} & \Phi_{16} & \Phi_8 & \Phi_{15} & \Phi_8 \\ \Phi_{11} & \Phi_{10} & \Phi_{12} & \Phi_8 & \Phi_7 & \Phi_8 \end{pmatrix}, \quad T_4 = \begin{pmatrix} \Phi_{13} & \Phi_{14} & \Phi_{15} & \Phi_8 & \Phi_{16} & \Phi_8 \\ \Phi_{10} & \Phi_{11} & \Phi_7 & \Phi_8 & \Phi_{12} & \Phi_8 \\ \Phi_{14} & \Phi_{13} & \Phi_{16} & \Phi_8 & \Phi_{15} & \Phi_8 \\ \Phi_{11} & \Phi_{10} & \Phi_{12} & \Phi_8 & \Phi_7 & \Phi_8 \\ \Phi_5 & \Phi_6 & \Phi_7 & \Phi_8 & \Phi_9 & \Phi_8 \\ \Phi_6 & \Phi_5 & \Phi_9 & \Phi_8 & \Phi_7 & \Phi_8 \end{pmatrix}, \tag{C.9}$$

$$T_5 = \begin{pmatrix} \Phi_5 & \Phi_6 & \Phi_7 & \Phi_8 & \Phi_9 & \Phi_8 \\ \Phi_6 & \Phi_5 & \Phi_9 & \Phi_8 & \Phi_7 & \Phi_8 \\ \Phi_7 & \Phi_9 & \Phi_7 & \Phi_8 & \Phi_{17} & \Phi_8 \\ \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 \\ \Phi_9 & \Phi_7 & \Phi_{17} & \Phi_8 & \Phi_7 & \Phi_8 \\ \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 \end{pmatrix}, \quad T_6 = \begin{pmatrix} \Phi_{18} & \Phi_{19} & \Phi_7 & \Phi_8 & \Phi_{20} & \Phi_8 \\ \Phi_{19} & \Phi_{21} & \Phi_9 & \Phi_8 & \Phi_{22} & \Phi_8 \\ \Phi_7 & \Phi_9 & \Phi_7 & \Phi_8 & \Phi_{17} & \Phi_8 \\ \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 \\ \Phi_{20} & \Phi_{22} & \Phi_{17} & \Phi_8 & \Phi_{23} & \Phi_8 \\ \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 \end{pmatrix}, \tag{C.10}$$

$$T_7 = \begin{pmatrix} \Phi_{19} & \Phi_{21} & \Phi_9 & \Phi_8 & \Phi_{22} & \Phi_8 \\ \Phi_{18} & \Phi_{19} & \Phi_7 & \Phi_8 & \Phi_{20} & \Phi_8 \\ \Phi_{20} & \Phi_{22} & \Phi_{17} & \Phi_8 & \Phi_{23} & \Phi_8 \\ \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 \\ \Phi_7 & \Phi_9 & \Phi_7 & \Phi_8 & \Phi_{17} & \Phi_8 \\ \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 \end{pmatrix}, \quad T_8 = \begin{pmatrix} \Phi_{21} & \Phi_{19} & \Phi_{22} & \Phi_8 & \Phi_9 & \Phi_8 \\ \Phi_{19} & \Phi_{18} & \Phi_{20} & \Phi_8 & \Phi_7 & \Phi_8 \\ \Phi_{22} & \Phi_{20} & \Phi_{23} & \Phi_8 & \Phi_{17} & \Phi_8 \\ \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 \\ \Phi_9 & \Phi_7 & \Phi_{17} & \Phi_8 & \Phi_7 & \Phi_8 \\ \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 \end{pmatrix}. \tag{C.11}$$

C.4 $g(k_1)\,\bar{q}(k_6) \to g(k_2)\,g(k_3)\,g(k_4)\,\bar{q}(k_5)$

$$T_1 = \begin{pmatrix} \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 \\ \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 \\ \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 \\ \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 \\ \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 \\ \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 & \Phi_8 \end{pmatrix}, \quad T_2 = \begin{pmatrix} \Phi_7 & \Phi_7 & \Phi_7 & \Phi_7 & \Phi_7 & \Phi_7 \\ \Phi_7 & \Phi_7 & \Phi_7 & \Phi_7 & \Phi_7 & \Phi_7 \\ \Phi_7 & \Phi_7 & \Phi_5 & \Phi_5 & \Phi_{18} & \Phi_{10} \\ \Phi_7 & \Phi_7 & \Phi_5 & \Phi_1 & \Phi_{10} & \Phi_2 \\ \Phi_7 & \Phi_7 & \Phi_{18} & \Phi_{10} & \Phi_5 & \Phi_5 \\ \Phi_7 & \Phi_7 & \Phi_{10} & \Phi_2 & \Phi_5 & \Phi_1 \end{pmatrix}, \tag{C.12}$$

$$T_3 = \begin{pmatrix} \Phi_{17} & \Phi_{17} & \Phi_9 & \Phi_9 & \Phi_{20} & \Phi_{12} \\ \Phi_{17} & \Phi_{23} & \Phi_9 & \Phi_{15} & \Phi_{22} & \Phi_{16} \\ \Phi_9 & \Phi_9 & \Phi_6 & \Phi_6 & \Phi_{19} & \Phi_{11} \\ \Phi_9 & \Phi_{15} & \Phi_6 & \Phi_2 & \Phi_{13} & \Phi_3 \\ \Phi_{20} & \Phi_{22} & \Phi_{19} & \Phi_{13} & \Phi_{21} & \Phi_{14} \\ \Phi_{12} & \Phi_{16} & \Phi_{11} & \Phi_3 & \Phi_{14} & \Phi_4 \end{pmatrix}, \quad T_4 = \begin{pmatrix} \Phi_{17} & \Phi_{23} & \Phi_9 & \Phi_{15} & \Phi_{22} & \Phi_{16} \\ \Phi_{17} & \Phi_{17} & \Phi_9 & \Phi_9 & \Phi_{20} & \Phi_{12} \\ \Phi_{20} & \Phi_{22} & \Phi_{19} & \Phi_{13} & \Phi_{21} & \Phi_{14} \\ \Phi_{12} & \Phi_{16} & \Phi_{11} & \Phi_3 & \Phi_{14} & \Phi_4 \\ \Phi_9 & \Phi_9 & \Phi_6 & \Phi_6 & \Phi_{19} & \Phi_{11} \\ \Phi_9 & \Phi_{15} & \Phi_6 & \Phi_2 & \Phi_{13} & \Phi_3 \end{pmatrix}, \tag{C.13}$$

$$T_5 = \begin{pmatrix} \Phi_{23} & \Phi_{17} & \Phi_{22} & \Phi_{16} & \Phi_9 & \Phi_{15} \\ \Phi_{17} & \Phi_{17} & \Phi_{20} & \Phi_{12} & \Phi_9 & \Phi_9 \\ \Phi_{22} & \Phi_{20} & \Phi_{21} & \Phi_{14} & \Phi_{19} & \Phi_{13} \\ \Phi_{16} & \Phi_{12} & \Phi_{14} & \Phi_4 & \Phi_{11} & \Phi_3 \\ \Phi_9 & \Phi_9 & \Phi_{19} & \Phi_{11} & \Phi_6 & \Phi_6 \\ \Phi_{15} & \Phi_9 & \Phi_{13} & \Phi_3 & \Phi_6 & \Phi_2 \end{pmatrix}. \tag{C.14}$$

D Large N_c limit for the TMD gluon distributions

For reader's convenience we list the large N_c expansions of the results presented in Section 4.5.

$g_1 g_5 \to g_2 g_3 g_4$	$\begin{pmatrix} \frac{1}{4} & 0 & 0 & 0 & 0 & \frac{3}{4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -\frac{1}{4}N_c^3 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$
$g_1 g_5 \to q_2 \bar{q}_3 g_4$	$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -N_c^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ -N_c^2 & 0 & 0 & N_c^2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$
$g_1 q_5 \to g_2 g_3 q_4$	$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 2 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & -N_c^2 & N_c^2 \\ 0 & 0 & 1 \end{pmatrix}$
$g_1 q_5 \to q_2 \bar{q}_3 q_4$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Table 13: The large N_c limit of the matrices $\mathbf{M}$ from Table 7.

$g_1 g_6 \to g_2 g_3 g_4 g_5$
$\begin{pmatrix} \frac{1}{8} & 0 & 0 & 0 & 0 & \frac{7}{8} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \frac{N_c^2}{2} & \frac{3N_c^2}{4} & 0 & \frac{1}{8}N_c^2 & \frac{1}{8}N_c^2 & \frac{N_c^2}{4} & -N_c^2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \frac{N_c^2}{4} & \frac{3N_c^2}{4} & 0 & 0 & \frac{1}{4}N_c^2 & \frac{N_c^2}{4} & -\frac{3N_c^2}{4} \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & \frac{1}{12}N_c^2 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

Table 14: The large N_c limit of the matrices $\mathbf{M}$ from Table 9.

$g_1 g_6 \to q_2 \bar{q}_3 g_4 g_5$
$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & -N_c^2 & N_c^2 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & -N_c^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ -N_c^2 & 0 & 0 & N_c^2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -N_c^4 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ -N_c^4 & 0 & 0 & 0 & 0 & N_c^4 & 0 \\ 0 & -N_c^2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{3}N_c^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{3}N_c^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -\frac{1}{3}N_c^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

Table 15: The large N_c limit of the matrices **M** from Table 10.

$g_1 q_6 \to g_2 g_3 g_4 q_5$
$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ N_c^2 & -N_c^2 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ -1 & 2 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & -N_c^2 & N_c^2 \\ 0 & 2 & -1 \\ N_c^2 & -2N_c^2 & N_c^2 \\ 0 & 0 & 1 \\ 0 & N_c^4 & -N_c^4 \\ 1 & -2 & 2 \\ 1 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & 1 & 0 \end{pmatrix}$

Table 16: The large N_c limit of the matrices $\mathbf{M}$ from Table 11.

$g_1 g_6 \to q_2 \bar{q}_3 q_4 \bar{q}_5$	$g_1 q_6 \to g_2 q_3 \bar{q}_4 q_5$
$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -N_c^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{4} & -\frac{1}{4} & 0 \\ 0 & -\frac{1}{4} & \frac{1}{4} \end{pmatrix}$

Table 17: The large N_c limit of the matrices $\mathbf{M}$ from Table 12.

E Color matrices

Below, we list the color factors for five and six parton processes. The convention for the enumerating of the rows and columns, *i.e.* the order of the partial amplitudes are the same as in Section 4.5. These color factors agree with [147, 170], after a suitable permutation of partial amplitudes is done.

Let us remind, that the actual color factors to be used in factorization formula together with the TMD matrices, are defined in Eq. (4.93). That is, the zero matrix elements have to be replaced by one.

$\emptyset \to ggggg$

$$\mathbf{C} = \frac{1}{4} N_c^3 N_A \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{2} & 1 & \frac{1}{4} & 0 & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & 1 & \frac{1}{2} & 0 & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{1}{2} & 1 & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} & 1 & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix} \tag{E.1}$$

$\emptyset \to q\bar{q}ggg$

$$\mathbf{C} = \frac{1}{8} \frac{N_A}{N_c^2} \begin{pmatrix} N_A^2 & -N_A & -N_A & 1 & 1 & F \\ -N_A & N_A^2 & 1 & F & -N_A & 1 \\ -N_A & 1 & N_A^2 & -N_A & F & 1 \\ 1 & F & -N_A & N_A^2 & 1 & -N_A \\ 1 & -N_A & F & 1 & N_A^2 & -N_A \\ F & 1 & 1 & -N_A & -N_A & N_A^2 \end{pmatrix}, \tag{E.2}$$

with $F = N_c^2 + 1$.

$\emptyset \to q\bar{q}r\bar{r}g$

$$\mathbf{C} = \frac{1}{2} N_A \begin{pmatrix} \frac{1}{N_c} & 0 & -\frac{1}{N_c} & -\frac{1}{N_c} \\ 0 & \frac{1}{N_c} & -\frac{1}{N_c} & -\frac{1}{N_c} \\ -\frac{1}{N_c} & -\frac{1}{N_c} & N_c & 0 \\ -\frac{1}{N_c} & -\frac{1}{N_c} & 0 & N_c \end{pmatrix} \tag{E.3}$$

$\emptyset \to gggggg$

$$\mathbf{C} = \frac{1}{4} N_c^4 N_A \begin{pmatrix} C_1 & C_2 & C_3 & C_4 \\ C_2 & C_1 & C_5 & C_6 \\ C_3^\intercal & C_5 & C_1 & C_7 \\ C_4^\intercal & C_6^\intercal & C_7 & C_1 \end{pmatrix}, \tag{E.4}$$

where

$$C_1 = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{2} & 1 & \frac{1}{4} & 0 & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & 1 & \frac{1}{2} & 0 & \frac{1}{4} \\ \frac{1}{4} & 0 & \frac{1}{2} & 1 & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} & 1 & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}, \qquad C_2 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & 0 & \frac{1}{4} & \frac{1}{8} \\ \frac{1}{4} & \frac{1}{8} & 0 & 0 & a + \frac{1}{8} & a \\ \frac{1}{8} & 0 & 0 & a & a & a \\ \frac{1}{8} & \frac{1}{4} & a + \frac{1}{8} & a & 0 & 0 \\ 0 & \frac{1}{8} & a & a & 0 & a \end{pmatrix}, \tag{E.5}$$

$$C_3 = \begin{pmatrix} \frac{1}{4} & \frac{1}{8} & 0 & 0 & a + \frac{1}{8} & a \\ \frac{1}{8} & 0 & 0 & a & a & a \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & 0 & \frac{1}{4} & \frac{1}{8} \\ 0 & \frac{1}{8} & a & a & 0 & a \\ \frac{1}{8} & \frac{1}{4} & a + \frac{1}{8} & a & 0 & 0 \end{pmatrix}, \qquad C_4 = \begin{pmatrix} \frac{1}{8} & 0 & 0 & a & a & a \\ \frac{1}{4} & \frac{1}{8} & 0 & 0 & a + \frac{1}{8} & a \\ 0 & \frac{1}{8} & a & a & 0 & a \\ \frac{1}{8} & \frac{1}{4} & a + \frac{1}{8} & a & 0 & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & 0 & \frac{1}{4} & \frac{1}{8} \end{pmatrix}, \tag{E.6}$$

$$C_5 = \begin{pmatrix} 0 & 0 & \frac{1}{4} & \frac{1}{8} & a & a+\frac{1}{8} \\ 0 & a & \frac{1}{8} & 0 & a & a \\ \frac{1}{4} & \frac{1}{8} & \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{8} \\ \frac{1}{8} & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{4} \\ a & a & 0 & \frac{1}{8} & a & 0 \\ a+\frac{1}{8} & a & \frac{1}{8} & \frac{1}{4} & 0 & 0 \end{pmatrix}, \qquad C_6 = \begin{pmatrix} 0 & a & \frac{1}{8} & 0 & a & a \\ 0 & 0 & \frac{1}{4} & \frac{1}{8} & a & a+\frac{1}{8} \\ a & a & 0 & \frac{1}{8} & a & 0 \\ a+\frac{1}{8} & a & \frac{1}{8} & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & \frac{1}{8} & \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{8} \\ \frac{1}{8} & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{4} \end{pmatrix}, \tag{E.7}$$

$$C_7 = \begin{pmatrix} a & 0 & a & a & \frac{1}{8} & 0 \\ 0 & 0 & a & a+\frac{1}{8} & \frac{1}{4} & \frac{1}{8} \\ a & a & a & 0 & 0 & \frac{1}{8} \\ a & a+\frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{1}{4} \\ \frac{1}{8} & \frac{1}{4} & 0 & \frac{1}{8} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{8} & \frac{1}{8} & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}, \tag{E.8}$$

with $a = \frac{3}{2N_c^2}$.

$\emptyset \to q\bar{q}gggg$

$$\mathbf{C} = \frac{1}{16}\frac{N_A}{N_c^3} \begin{pmatrix} C_1 & C_2 & C_3 & C_4 \\ C_2 & C_1 & C_5 & C_6 \\ C_3^\intercal & C_5 & C_1 & C_7 \\ C_4^\intercal & C_6^\intercal & C_7 & C_1 \end{pmatrix}, \tag{E.9}$$

where

$$C_1 = \begin{pmatrix} N_A^3 & -N_A^2 & -N_A^2 & N_A & N_A & N_c^4-1 \\ -N_A^2 & N_A^3 & N_A & N_c^4-1 & -N_A^2 & N_A \\ -N_A^2 & N_A & N_A^3 & -N_A^2 & N_c^4-1 & N_A \\ N_A & N_c^4-1 & -N_A^2 & N_A^3 & N_A & -N_A^2 \\ N_A & -N_A^2 & N_c^4-1 & N_A & N_A^3 & -N_A^2 \\ N_c^4-1 & N_A & N_A & -N_A^2 & -N_A^2 & N_A^3 \end{pmatrix}, \tag{E.10}$$

$$C_2 = \begin{pmatrix} -N_A^2 & N_A & N_A & -1 & -1 & -N_c^2-1 \\ N_A & -N_A^2 & -1 & -N_c^2-1 & N_A & -1 \\ N_A & -1 & N_c^4-1 & -N_c^2-1 & L & K \\ -1 & -N_c^2-1 & -N_c^2-1 & K & K & -3N_c^2-1 \\ -1 & N_A & L & K & N_c^4-1 & -N_c^2-1 \\ -N_c^2-1 & -1 & K & -3N_c^2-1 & -N_c^2-1 & K \end{pmatrix}, \tag{E.11}$$

$$C_3 = \begin{pmatrix} N_A & -1 & N_c^4-1 & -N_c^2-1 & L & K \\ -1 & -N_c^2-1 & -N_c^2-1 & K & K & -3N_c^2-1 \\ -N_A^2 & N_A & N_A & -1 & -1 & -N_c^2-1 \\ N_A & -N_A^2 & -1 & -N_c^2-1 & N_A & -1 \\ -N_c^2-1 & -1 & K & -3N_c^2-1 & -N_c^2-1 & K \\ -1 & N_A & L & K & N_c^4-1 & -N_c^2-1 \end{pmatrix}, \tag{E.12}$$

$$C_4 = \begin{pmatrix} -1 & -N_c^2-1 & -N_c^2-1 & K & K & -3N_c^2-1 \\ N_A & -1 & N_c^4-1 & -N_c^2-1 & L & K \\ -N_c^2-1 & -1 & K & -3N_c^2-1 & -N_c^2-1 & K \\ -1 & N_A & L & K & N_c^4-1 & -N_c^2-1 \\ -N_A^2 & N_A & N_A & -1 & -1 & -N_c^2-1 \\ N_A & -N_A^2 & -1 & -N_c^2-1 & N_A & -1 \end{pmatrix}, \tag{E.13}$$

$$C_5 = \begin{pmatrix} N_c^4-1 & -N_c^2-1 & N_A & -1 & K & L \\ -N_c^2-1 & K & -1 & -N_c^2-1 & -3N_c^2-1 & K \\ N_A & -1 & -N_A^2 & N_A & -N_c^2-1 & -1 \\ -1 & -N_c^2-1 & N_A & -N_A^2 & -1 & N_A \\ K & -3N_c^2-1 & -N_c^2-1 & -1 & K & -N_c^2-1 \\ L & K & -1 & N_A & -N_c^2-1 & N_c^4-1 \end{pmatrix}, \tag{E.14}$$

$$C_6 = \begin{pmatrix} -N_c^2-1 & K & -1 & -N_c^2-1 & -3N_c^2-1 & K \\ N_c^4-1 & -N_c^2-1 & N_A & -1 & K & L \\ K & -3N_c^2-1 & -N_c^2-1 & -1 & K & -N_c^2-1 \\ L & K & -1 & N_A & -N_c^2-1 & N_c^4-1 \\ N_A & -1 & -N_A^2 & N_A & -N_c^2-1 & -1 \\ -1 & -N_c^2-1 & N_A & -N_A^2 & -1 & N_A \end{pmatrix}, \tag{E.15}$$

$$C_7 = \begin{pmatrix} K & -N_c^2-1 & -3N_c^2-1 & K & -1 & -N_c^2-1 \\ -N_c^2-1 & N_c^4-1 & K & L & N_A & -1 \\ -3N_c^2-1 & K & K & -N_c^2-1 & -N_c^2-1 & -1 \\ K & L & -N_c^2-1 & N_c^4-1 & -1 & N_A \\ -1 & N_A & -N_c^2-1 & -1 & -N_A^2 & N_A \\ -N_c^2-1 & -1 & -1 & N_A & N_A & -N_A^2 \end{pmatrix}, \tag{E.16}$$

We find the following symmetry transformations

$$C_3 = C_6^{M^\intercal}, \qquad C_7 = C_2^M, \tag{E.17}$$

where

$$A^M = JAJ, \tag{E.18}$$

with

$$J = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{E.19}$$

$\emptyset \to q\bar{q}r\bar{r}gg$

$$\mathbf{C} = \frac{1}{4} N_A \begin{pmatrix} C_1 & C_2 \\ C_2 & C_3 \end{pmatrix}, \tag{E.20}$$

where

$$C_1 = \begin{pmatrix} \frac{N_A}{N_c^2} & 0 & \frac{1}{N_c^2} & -\frac{1}{N_c^2} & 0 & \frac{1}{N_c^2} \\ 0 & \frac{N_A}{N_c^2} & 0 & 0 & \frac{1}{N_c^2} & 0 \\ \frac{1}{N_c^2} & 0 & \frac{N_A}{N_c^2} & \frac{1}{N_c^2} & 0 & -\frac{1}{N_c^2} \\ -\frac{1}{N_c^2} & 0 & \frac{1}{N_c^2} & \frac{N_A}{N_c^2} & 0 & \frac{1}{N_c^2} \\ 0 & \frac{1}{N_c^2} & 0 & 0 & \frac{N_A}{N_c^2} & 0 \\ \frac{1}{N_c^2} & 0 & -\frac{1}{N_c^2} & \frac{1}{N_c^2} & 0 & \frac{N_A}{N_c^2} \end{pmatrix}, \tag{E.21}$$

$$C_2 = \begin{pmatrix} -\frac{N_A}{N_c^2} & \frac{1}{N_c^2} & -\frac{N_A}{N_c^2} & \frac{1}{N_c^2} & -\frac{N_A}{N_c^2} & \frac{1}{N_c^2} \\ \frac{1}{N_c^2} & -\frac{N_A}{N_c^2} & -\frac{N_A}{N_c^2} & -\frac{N_A}{N_c^2} & -\frac{N_A}{N_c^2} & \frac{1}{N_c^2} \\ -\frac{N_A}{N_c^2} & -\frac{N_A}{N_c^2} & -\frac{N_A}{N_c^2} & \frac{1}{N_c^2} & \frac{1}{N_c^2} & \frac{1}{N_c^2} \\ \frac{1}{N_c^2} & -\frac{N_A}{N_c^2} & \frac{1}{N_c^2} & -\frac{N_A}{N_c^2} & \frac{1}{N_c^2} & -\frac{N_A}{N_c^2} \\ -\frac{N_A}{N_c^2} & -\frac{N_A}{N_c^2} & \frac{1}{N_c^2} & \frac{1}{N_c^2} & -\frac{N_A}{N_c^2} & -\frac{N_A}{N_c^2} \\ \frac{1}{N_c^2} & \frac{1}{N_c^2} & \frac{1}{N_c^2} & -\frac{N_A}{N_c^2} & -\frac{N_A}{N_c^2} & -\frac{N_A}{N_c^2} \end{pmatrix}, \tag{E.22}$$

$$C_3 = N_c^2 C_1 \,. \tag{E.23}$$

F Matrix elements for $2 \to 1$ processes

The matrix elements squared for three-parton processes, used in the calculation of the single inclusive forward jet distributions in Section 5.2 read

- $gg^* \to g$

$$\overline{|\mathcal{M}_{gg^*\to g}|}^2 = 4g_s^2 \frac{C_A}{N_c^2-1} \frac{(k\cdot q)^2}{k_T^2}\,,$$

- $qg^* \to q$

$$\overline{|\mathcal{M}_{qg^*\to q}|}^2 = 4g_s^2 \frac{C_F}{N_c^2-1} \frac{(k\cdot q)^2}{k_T^2}\,,$$

- $gq^* \to q$

$$\overline{|\mathcal{M}_{gq^*\to q}|}^2 = g_s^2 \frac{C_F}{N_c^2-1}(k\cdot q)\,,$$

- $\bar{q}q^* \to g$

$$\overline{|\mathcal{M}_{\bar{q}q^*\to g}|}^2 = g_s^2 \frac{C_F}{N_c}(k\cdot q)\,.$$

In the above, k and q are the momenta of the off-shell and on-shell partons, respectively. k_T is the transverse component of the off-shell momentum, g_s is a strong coupling and C_i is a colour factor of the emitter: C_F for a quark and C_A for a gluon.

References

[1] H. Fritzsch, M. Gell-Mann and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, *Phys. Lett.* **47B** (1973) 365–368. 8

[2] S. Weinberg, *Nonabelian Gauge Theories of the Strong Interactions*, *Phys. Rev. Lett.* **31** (1973) 494–497. 8

[3] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, *Phys. Rev. Lett.* **30** (1973) 1343–1346. 8, 10

[4] T. Muta, *Foundations of quantum chromodynamics. Second edition*, *World Sci. Lect. Notes Phys.* **57** (1998) 1–409. 8

[5] H. Georgi, *Lie algebras in particle physics*, *Front. Phys.* **54** (1999) 1–320. 8

[6] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and collider physics*, *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* **8** (1996) 1–435. 9, 12

[7] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995. 9

[8] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, *Phys. Rev. Lett.* **30** (1973) 1346–1349. 10

[9] N. K. Nielsen, *ASYMPTOTIC FREEDOM AS A SPIN EFFECT*, *Am. J. Phys.* **49** (1981) 1171. 10

[10] Particle Data Group collaboration, M. Tanabashi et al., *Review of Particle Physics*, *Phys. Rev.* **D98** (2018) 030001. 11, 15

[11] F. Bloch and A. Nordsieck, *Note on the radiation field of the electron*, *Phys. Rev.* **52** (Jul, 1937) 54–59. 12

[12] T. Kinoshita, *Mass singularities of Feynman amplitudes*, *J. Math. Phys.* **3** (1962) 650–677. 12

[13] T. D. Lee and M. Nauenberg, *Degenerate Systems and Mass Singularities*, *Phys. Rev.* **133** (mar, 1964) B1549–B1562. 12

[14] J. D. Bjorken, *Asymptotic Sum Rules at Infinite Momentum*, *Phys. Rev.* **179** (1969) 1547–1553. 14

[15] R. P. Feynman, *Photon-hadron interactions*, . 14

[16] C. G. Callan, Jr. and D. J. Gross, *High-energy electroproduction and the constitution of the electric current*, *Phys. Rev. Lett.* **22** (1969) 156–159. 14

[17] G. Dissertori, I. G. Knowles and M. Schmelling, *High energy experiments and theory*. 2003. 14

[18] C. Marquet, *Open questions in QCD at high parton density*, *Nucl. Phys.* **A904-905** (2013) 294c–301c, [1212.3482]. 15

[19] M. Kuhlen, *QCD and the Hadronic Final State in Deep Inelastic Scattering at HERA*. PhD thesis, Munich, Max Planck Inst., 1997. hep-ph/9712505. 14, 23

[20] V. Gribov and L. Lipatov, *Deep inelastic e p scattering in perturbation theory*, *Sov.J.Nucl.Phys.* **15** (1972) 438–450. 16, 23

[21] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl.Phys.* **B126** (1977) 298. 16, 17, 23

[22] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, *Sov. Phys. JETP* **46** (1977) 641–653. 16, 23

[23] A. D. Martin, *Structure functions and small x physics*, in *21st International Meeting on Fundamental Physics: Physics at HERA Madrid, Spain, May 9-15, 1993*, pp. 0250–292, 1993. 17, 19

[24] A. De Rujula, S. L. Glashow, H. D. Politzer, S. B. Treiman, F. Wilczek and A. Zee, *Possible NonRegge Behavior of Electroproduction Structure Functions*, *Phys. Rev.* **D10** (1974) 1649. 18

[25] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, *Eur. Phys. J.* **C75** (2015) 204, [1412.3989]. 18

[26] NNPDF collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040, [1410.8849]. 18

[27] J. Gao, M. Guzzi, J. Huston, H.-L. Lai, Z. Li, P. Nadolsky et al., *CT10 next-to-next-to-leading order global analysis of QCD*, *Phys. Rev.* **D89** (2014) 033009, [1302.6246]. 18, 19

[28] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, Others et al., *New parton distributions for collider physics*, *Phys. Rev. D* **82** (oct, 2010) 074024, [1007.2241]. 19, 34, 35, 70, 80

[29] L. N. Lipatov, *Reggeization of the Vector Meson and the Vacuum Singularity in Nonabelian Gauge Theories*, *Sov. J. Nucl. Phys.* **23** (1976) 338–345. 18

[30] E. A. Kuraev, L. N. Lipatov and V. S. Fadin, *The Pomeranchuk Singularity in Nonabelian Gauge Theories*, *Sov. Phys. JETP* **45** (1977) 199–204. 18

[31] I. I. Balitsky and L. N. Lipatov, *The Pomeranchuk Singularity in Quantum Chromodynamics*, *Sov. J. Nucl. Phys.* **28** (1978) 822–829. 18

[32] M. Ciafaloni, *K factorization and next-to-leading anomalous dimensions*, in *Proceedings, Ringberg Workshop on New Trends in HERA Physics 1997: Ringberg Castle, Tegernsee, Germany, May 25-30, 1997*, pp. 62–70, 1997. hep-ph/9709390. 19

[33] R. D. Ball and P. V. Landshoff, *The Challenge of small x*, *J. Phys.* **G26** (2000) 672–682, [hep-ph/9912445]. 20

[34] J. C. Collins and P. V. Landshoff, *Exploring the Lipatov equation*, *Phys. Lett.* **B276** (1992) 196–202. 20

[35] J. Bartels, H. Lotter and M. Vogt, *A Numerical estimate of the small k(T) region in the BFKL pomeron*, *Phys. Lett.* **B373** (1996) 215–222, [hep-ph/9511399]. 20

[36] L. P. A. Haakman, O. V. Kancheli and J. H. Koch, *The BFKL pomeron with running coupling constant: How much of its hard nature survives?*, *Nucl. Phys.* **B518** (1998) 275–302, [hep-ph/9707262]. 20

[37] Y. V. Kovchegov and A. H. Mueller, *Running coupling effects in BFKL evolution*, *Phys. Lett.* **B439** (1998) 428–436, [hep-ph/9805208]. 20

[38] N. Armesto, J. Bartels and M. A. Braun, *On the second order corrections to the hard pomeron and the running coupling*, *Phys. Lett.* **B442** (1998) 459–469, [hep-ph/9808340]. 20

[39] V. Fadin and L. Lipatov, *BFKL pomeron in the next-to-leading approximation*, *Phys. Lett. B* **429** (jun, 1998) 127–134, [hep-ph/9802290]. 20

[40] G. P. Salam, *A Resummation of large subleading corrections at small x*, *JHEP* **07** (1998) 019, [hep-ph/9806482]. 20

[41] M. Ciafaloni, D. Colferai and G. P. Salam, *Renormalization group improved small x equation*, *Phys. Rev.* **D60** (1999) 114036, [hep-ph/9905566]. 20

[42] M. Ciafaloni, D. Colferai and G. P. Salam, *A collinear model for small x physics*, *JHEP* **10** (1999) 017, [hep-ph/9907409]. 20

[43] G. Altarelli, R. D. Ball and S. Forte, *Resummation of singlet parton evolution at small x*, *Nucl. Phys.* **B575** (2000) 313–329, [hep-ph/9911273]. 20

[44] R. D. Ball and S. Forte, *The Small x behavior of Altarelli-Parisi splitting functions*, *Phys. Lett.* **B465** (1999) 271–281, [hep-ph/9906222]. 20

[45] M. Ciafaloni, *Coherence Effects in Initial Jets at Small q**2 / s*, *Nucl. Phys.* **B296** (1988) 49. 20

[46] S. Catani, F. Fiorani and G. Marchesini, *QCD Coherence in Initial State Radiation*, *Phys. Lett.* **B234** (1990) 339. 20

[47] S. Catani, F. Fiorani and G. Marchesini, *Small x Behavior of Initial State Radiation in Perturbative QCD*, *Nucl. Phys.* **B336** (1990) 18. 20

[48] J. Kwiecinski, A. D. Martin and P. J. Sutton, *The gluon distribution at small x obtained from a unified evolution equation*, *Phys. Rev.* **D52** (1995) 1445–1458, [hep-ph/9503266]. 20

[49] G. Marchesini, *QCD coherence in the structure function and associated distributions at small x*, *Nucl. Phys.* **B445** (1995) 49–80, [hep-ph/9412327]. 20

[50] M. Froissart, *Asymptotic behavior and subtractions in the Mandelstam representation*, *Phys. Rev.* **123** (1961) 1053–1057. 21

[51] A. Martin, *Unitarity and high-energy behavior of scattering amplitudes*, *Phys. Rev.* **129** (1963) 1432–1436. 21

[52] I. Balitsky, *Operator expansion for high-energy scattering*, *Nucl. Phys. B* **463** (1996) 99–157, [hep-ph/9509348]. 21

[53] Y. V. Kovchegov, *Small x F(2) structure function of a nucleus including multiple pomeron exchanges*, *Phys. Rev.* **D60** (1999) 034008, [hep-ph/9901281]. 21

[54] K. Kutak, *Resummation in nonlinear equation for high energy factorizable gluon density and its extension to include coherence*, *J. High Energy Phys.* **2012** (dec, 2012) 33, [1206.5757]. 21, 38

[55] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007, [0811.4622]. 22, 23

[56] T. Sjöstrand, S. Mrenna and P. Skands, *PYTHIA 6.4 physics and manual*, *J. High Energy Phys.* **2006** (may, 2006) 026–026, [hep-ph/0603175]. 22, 82

[57] M. Bähr, S. Gieseke, M. A. Gigg, D. Grellscheid, K. Hamilton, O. Latunde-Dada et al., *Herwig++ physics and manual*, *Eur. Phys. J. C* **58** (dec, 2008) 639–707, [0803.0883]. 22

[58] S. Sapeta, *QCD and Jets at Hadron Colliders*, *Prog. Part. Nucl. Phys.* **89** (2016) 1–55, [1511.09336]. 22, 25, 26, 27, 28

[59] B. Andersson, G. Gustafson, G. Ingelman and T. Sjostrand, *Parton Fragmentation and String Dynamics*, *Phys. Rept.* **97** (1983) 31–145. 23

[60] T. D. Gottschalk, *A Realistic Model for e+ e- Annihilation Including Parton Bremsstrahlung Effects*, *Nucl. Phys.* **B214** (1983) 201–222. 23

[61] F. Siegert, *Monte-Carlo event generation for the LHC*. PhD thesis, Durham U., 2010. 24

[62] G. F. Sterman and S. Weinberg, *Jets from Quantum Chromodynamics*, *Phys.Rev.Lett.* **39** (1977) 1436. 24

[63] S. Ellis, J. Huston, K. Hatakeyama, P. Loch and M. Tonnesmann, *Jets in hadron-hadron collisions*, *Prog.Part.Nucl.Phys.* **60** (2008) 484–551, [0712.2447]. 24

[64] G. P. Salam, *Towards Jetography*, *Eur.Phys.J.* **C67** (2010) 637–686, [0906.1833]. 24

[65] G. P. Salam and G. Soyez, *A Practical Seedless Infrared-Safe Cone jet algorithm*, *JHEP* **0705** (2007) 086, [0704.0292]. 24

[66] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, *Nucl. Phys.* **B406** (1993) 187–224. 24

[67] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev.* **D48** (1993) 3160–3166, [hep-ph/9305266]. 24

[68] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001, [hep-ph/9707323]. 24

[69] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross-sections in deep inelastic scattering*, in *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany, 1998-1999*, 1998. hep-ph/9907280. 24

[70] M. Cacciari, G. P. Salam and G. Soyez, *The anti-kT jet clustering algorithm*, *J. High Energy Phys.* **2008** (apr, 2008) 063–063, [0802.1189]. 24, 73, 80

[71] J. Collins, *Foundations of perturbative QCD*, vol. 32. Cambridge Univ. Press, 2011. 26, 27, 28

[72] Y. Dokshitzer, D. Dyakonov and S. Troyan, *Hard processes in quantum chromodynamics*, *Phys. Rep.* **58** (feb, 1980) 269–395. 26

[73] J. Collins, D. E. Soper and G. Sterman, *Transverse momentum distribution in Drell-Yan pair and W and Z boson production*, *Nucl. Phys. B* **250** (jan, 1985) 199–224. 26

[74] L. V. Gribov, E. M. Levin and M. G. Ryskin, *Semihard processes in QCD*, *Phys. Rep.* **100** (nov, 1983) 1–150. 26

[75] S. Catani, M. Ciafaloni and F. Hautmann, *High-energy factorization and small x heavy flavor production*, *Nucl. Phys.* **B366** (1991) 135–188. 26, 30, 33

[76] S. Catani and F. Hautmann, *High-energy factorization and small-x deep inelastic scattering beyond leading order*, *Nucl. Phys. B* **427** (oct, 1994) 475–524, [hep-ph/9405388]. 26

[77] J. Collins and R. Ellis, *Heavy-quark production in very high energy hadron collisions*, *Nucl. Phys. B* **360** (aug, 1991) 3–30. 26

[78] L. McLerran and R. Venugopalan, *Green's function in the color field of a large nucleus*, *Phys. Rev. D* **50** (1994) 2225–2233, [9402335]. 26

[79] L. D. McLerran, *The Color glass condensate and small x physics: Four lectures*, *Lect. Notes Phys.* **583** (2002) 291–334, [hep-ph/0104285]. 26

[80] F. Gelis, E. Iancu, J. Jalilian-Marian and R. Venugopalan, *The Color Glass Condensate*, *Annu. Rev. Nucl. Part. Sci.* **60** (nov, 2010) 463–489, [1002.0333]. 26

[81] Y. V. Kovchegov and E. Levin, *Quantum chromodynamics at high energy*, vol. 33. Cambridge University Press, 2012. 26

[82] P. Kotko, K. Kutak, C. Marquet, E. Petreska, S. Sapeta and A. van Hameren, *Improved TMD factorization for forward dijet production in dilute-dense hadronic collisions*, *J. High Energy Phys.* **2015** (sep, 2015) 106, [1503.03421]. 26, 32, 33, 40, 53, 54, 76, 80

[83] T. Altinoluk, R. Boussarie and P. Kotko, *Interplay of the CGC and TMD frameworks to all orders in kinematic twist*, *JHEP* **05** (2019) 156, [1901.01175]. 26

[84] J. C. Collins, D. E. Soper and G. Sterman, *Factorization of Hard Processes in QCD*, *Adv. Ser. Direct. High Energy Phys.* **5** (sep, 2004) 1–91, [0409313]. 26, 28

[85] G. F. Sterman, *Mass Divergences in Annihilation Processes. 1. Origin and Nature of Divergences in Cut Vacuum Polarization Diagrams*, *Phys. Rev.* **D17** (1978) 2773. 26

[86] G. F. Sterman, *Mass Divergences in Annihilation Processes. 2. Cancellation of Divergences in Cut Vacuum Polarization Diagrams*, *Phys. Rev.* **D17** (1978) 2789. 26

[87] J. C. Collins, D. E. Soper and G. Sterman, *Factorization for short distance hadron-hadron scattering*, *Nucl. Phys. B* **261** (jan, 1985) 104–142. 26, 27, 28

[88] J. C. Taylor, *Ward Identities and Charge Renormalization of the Yang-Mills Field*, *Nucl. Phys.* **B33** (1971) 436–444. 28

[89] A. A. Slavnov, *Ward Identities in Gauge Theories*, *Theor. Math. Phys.* **10** (1972) 99–107. 28

[90] G. T. Bodwin, S. J. Brodsky and G. P. Lepage, *Initial State Interactions and the Drell-Yan Process*, *Phys.Rev.Lett.* **47** (1981) 1799. 28

[91] J. C. Collins and G. F. Sterman, *Soft Partons in QCD*, *Nucl. Phys.* **B185** (1981) 172. 28

[92] S. Catani, D. de Florian and G. Rodrigo, *Space-like (versus time-like) collinear limits in QCD: Is factorization violated?*, *JHEP* **1207** (2012) 026, [1112.4405]. 28

[93] J. R. Forshaw, M. H. Seymour and A. Siodmok, *On the Breaking of Collinear Factorization in QCD*, *JHEP* **11** (2012) 066, [1206.6363]. 28

[94] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization of Hard Processes in QCD*, *Adv.Ser.Direct.High Energy Phys.* **5** (1988) 1–91, [hep-ph/0409313]. 28

[95] J. C. Collins and D. E. Soper, *Parton distribution and decay functions*, *Nucl. Phys. B* **194** (jan, 1982) 445–492. 28, 29

[96] G. Curci, W. Furmanski and R. Petronzio, *Evolution of parton densities beyond leading order*, *Nucl. Phys. B* **175** (nov, 1980) 27–92. 28

[97] C. J. Bomhof, P. J. Mulders and F. Pijlman, *The construction of gauge-links in arbitrary hard processes*, *Eur. Phys. J. C* **47** (jul, 2006) 147–162, [hep-ph/0601171]. 29, 40, 41, 42, 47, 49, 50, 54, 90

[98] A. V. Belitsky, X. Ji and F. Yuan, *Final state interactions and gauge invariant parton distributions*, *Nucl. Phys. B* **656** (2003) 165–198, [0208038]. 29

[99] J. C. Collins, *Leading twist single transverse-spin asymmetries: Drell-Yan and deep inelastic scattering*, *Phys.Lett.* **B536** (2002) 43–48, [hep-ph/0204004]. 29

[100] J. Collins and J.-W. Qiu, *<math display="inline"> <msub> <mi>k</mi> <mi>T</mi> </msub> </math> factorization is violated in production of high-transverse-momentum particles in hadron-hadron collisions*, *Phys. Rev. D* **75** (jun, 2007) 114014, [0705.2141]. 30

[101] T. C. Rogers and P. J. Mulders, *No Generalized TMD-Factorization in Hadro-Production of High Transverse Momentum Hadrons*, *Phys.Rev.* **D81** (2010) 094006, [1001.2977]. 30, 40

[102] S. Catani, M. Ciafaloni and F. Hautmann, *GLUON CONTRIBUTIONS TO SMALL x HEAVY FLAVOR PRODUCTION*, *Phys. Lett.* **B242** (1990) 97. 30, 33

[103] P. Kotko, W. Słomiński and D. Toton, *Unintegrated Gluon Distributions for Forward Jets at the LHC*, *Acta Phys. Pol. B* **46** (apr, 2015) 1527, [1504.00823]. 30, 31

[104] A. van Hameren, P. Kotko and K. Kutak, *Helicity amplitudes for high-energy scattering*, *JHEP* **01** (2013) 078, [1211.0961]. 31, 70

[105] A. van Hameren, P. Kotko and K. Kutak, *Multi-gluon helicity amplitudes with one off-shell leg within high energy factorization*, *JHEP* **12** (2012) 029, [1207.3332]. 31, 70

[106] A. Van Hameren, *BCFW recursion for off-shell gluons*, *J. High Energy Phys.* **2014** (jul, 2014) 138, [1404.7818]. 31

[107] P. Kotko, *Wilson lines and gauge invariant off-shell amplitudes*, *J. High Energy Phys.* **2014** (jul, 2014) 128, [1403.4824]. 31, 45

[108] L. Lipatov, *Gauge invariant effective action for high energy processes in QCD*, *Nucl. Phys. B* **452** (oct, 1995) 369–397, [hep-ph/9502308]. 31

[109] E. Antonov, I. Cherednikov, E. Kuraev and L. Lipatov, *Feynman rules for effective Regge action*, *Nucl. Phys. B* **721** (aug, 2005) 111–135, [0411185]. 31

[110] M. Deak, F. Hautmann, H. Jung and K. Kutak, *Forward jet production at the Large Hadron Collider*, *J. High Energy Phys.* **2009** (sep, 2009) 121–121, [0908.0538]. 32, 76, 78, 80

[111] K. Kutak and S. Sapeta, *Gluon saturation in dijet production in p-Pb collisions at the Large Hadron Collider*, *Phys. Rev. D* **86** (nov, 2012) 094043, [1205.5035]. 32, 36, 37, 53, 76, 78, 83

[112] M. Deak, F. Hautmann, H. Jung and K. Kutak, *Forward-Central Jet Correlations at the Large Hadron Collider*, 1012.6037. 32

[113] A. van Hameren, P. Kotko, K. Kutak, C. Marquet and S. Sapeta, *Saturation effects in forward-forward dijet production in p+Pb collisions*, *Phys. Rev. D* **89** (may, 2014) 094014, [1402.5065]. 32, 76

[114] A. van Hameren, P. Kotko, K. Kutak and S. Sapeta, *Small-x dynamics in forward-central dijet decorrelations at the LHC*, *Phys. Lett.* **B737** (2014) 335–340, [1404.6204]. 32, 38

[115] F. Dominguez, C. Marquet, B.-W. Xiao and F. Yuan, *Universality of unintegrated gluon distributions at small x*, *Phys. Rev. D* **83** (may, 2011) 105005, [1101.0715]. 32, 40, 52, 53, 71

[116] M. L. Mangano and S. J. Parke, *Multi-parton amplitudes in gauge theories*, *Phys. Rep.* **200** (feb, 1991) 301–367, [hep-th/0509223]. 33, 45

[117] A. van Hameren, P. Kotko, K. Kutak, C. Marquet, E. Petreska and S. Sapeta, *Forward di-jet production in p+Pb collisions in the small-x improved TMD factorization framework*, 1607.03121. 33

[118] M. A. Kimber, A. D. Martin and M. G. Ryskin, *Unintegrated parton distributions and prompt photon hadroproduction*, *Eur. Phys. J.* **C12** (2000) 655–661, [hep-ph/9911379]. 34

[119] M. A. Kimber, A. D. Martin and M. G. Ryskin, *Unintegrated parton distributions*, *Phys. Rev.* **D63** (2001) 114027, [hep-ph/0101348]. 34

[120] A. D. Martin, M. G. Ryskin and G. Watt, *NLO prescription for unintegrated parton distributions*, *Eur. Phys. J.* **C66** (2010) 163–172, [0909.5529]. 34, 35

[121] F. Hautmann, H. Jung, M. Krämer, P. J. Mulders, E. R. Nocera, T. C. Rogers et al., *TMDlib and TMDplotter: library and plotting tools for transverse-momentum-dependent parton distributions*, *Eur. Phys. J.* **C74** (2014) 3220, [1408.3015]. 35, 86

[122] M. Bury, "MRWCALC." code on request. 35, 86

[123] J. Kwiecinski, A. D. Martin and A. M. Stasto, *A Unified BFKL and GLAP description of F2 data*, *Phys. Rev.* **D56** (1997) 3991–4006, [hep-ph/9703445]. 35

[124] K. Kutak and J. Kwiecinski, *Screening effects in the ultra-high energy neutrino interactions*, *Eur. Phys. J. C* **29** (aug, 2003) 521–530, [hep-ph/0303209]. 35

[125] K. Kutak and A. M. Staśto, *The unintegrated gluon distribution from the modified BK equation*, *Eur. Phys. J. C* **41** (jun, 2005) 343–351, [hep-ph/0408117]. 35

[126] J. Bartels and M. Wusthoff, *The Triple Regge limit of diffractive dissociation in deep inelastic scattering*, *Z. Phys.* **C66** (1995) 157–180. 36

[127] H1, ZEUS collaboration, F. D. Aaron et al., *Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA*, *JHEP* **01** (2010) 109, [0911.0884]. 36

[128] V. V. Sudakov, *Vertex parts at very high-energies in quantum electrodynamics*, *Sov.Phys.JETP* **3** (1956) 65–71. 37

[129] K. Kutak, *Hard scale dependent gluon density, saturation and forward-forward dijet production at the LHC*, *Phys. Rev.* **D91** (2015) 034021, [1409.3822]. 37, 38, 39

[130] A. van Hameren, P. Kotko and K. Kutak, *Resummation effects in the forward production of Z_0+jet at the LHC*, *Phys. Rev.* **D92** (2015) 054007, [1505.02763]. 38

[131] A. van Hameren, P. Kotko, K. Kutak and S. Sapeta, *Broadening and saturation effects in dijet azimuthal correlations in p-p and p-Pb collisions at $\sqrt{\mathbf{s}} = 5.02$ TeV*, *Phys. Lett.* **B795** (2019) 511–515, [1903.01361]. 38

[132] K. Kutak, K. Golec-Biernat, S. Jadach and M. Skrzypek, *Nonlinear equation for coherent gluon emission*, *J. High Energy Phys.* **2012** (feb, 2012) 117, [1111.6928]. 38

[133] K. Kutak, *Nonlinear extension of the CCFM equation*, *Proceedings, 20th Int. Work. Deep. Scatt. Relat. Subj. (DIS 2012) Bonn, Ger. March 26-30, 2012* (jun, 2012) 565–568, [1206.1223]. 38

[134] C. Bomhof and P. Mulders, *Non-universality of transverse momentum dependent parton distribution functions*, *Nucl. Phys. B* **795** (may, 2008) 409–427, [0709.1390]. 40, 52

[135] A. van Hameren, P. Kotko, K. Kutak, C. Marquet, E. Petreska and S. Sapeta, *Forward di-jet production in p+Pb collisions in the small-x improved TMD factorization framework*, *JHEP* **12** (2016) 034, [1607.03121]. 40, 52, 53, 54, 83

[136] C. Marquet, E. Petreska and C. Roiesnel, *Transverse-momentum-dependent gluon distributions from JIMWLK evolution*, *JHEP* **10** (2016) 065, [1608.02577]. 40, 52, 85

[137] P. Kotko, K. Kutak, S. Sapeta, A. M. Stasto and M. Strikman, *Estimating nonlinear effects in forward dijet production in ultra-peripheral heavy ion collisions at the LHC*, *Eur. Phys. J.* **C77** (2017) 353, [1702.03063]. 40

[138] C. Marquet, C. Roiesnel and P. Taels, *Linearly polarized small-x gluons in forward heavy-quark pair production*, *Phys. Rev.* **D97** (2018) 014004, [1710.05698]. 40

[139] D. E. Soper, *The Parton Model and the Bethe-Salpeter Wave Function*, *Phys. Rev.* **D15** (1977) 1141. 40

[140] D. E. Soper, *Partons and Their Transverse Momenta in QCD*, *Phys. Rev. Lett.* **43** (1979) 1847. 40

[141] J. C. Collins, D. E. Soper and G. F. Sterman, *Does the Drell-Yan Cross-section Factorize?*, *Phys. Lett.* **109B** (1982) 388–392. 40

[142] D. Boer and P. J. Mulders, *Color gauge invariance in the Drell-Yan process*, *Nucl. Phys.* **B569** (2000) 505–526, [hep-ph/9906223]. 41

[143] A. Van Hameren, P. Kotko and K. Kutak, *Multi-gluon helicity amplitudes with one off-shell leg within high energy factorization*, *J. High Energy Phys.* **2012** (2012) , [1207.3332]. 45

[144] A. Van Hameren, P. Kotko and K. Kutak, *Helicity amplitudes for high-energy scattering*, *J. High Energy Phys.* **2013** (2013) . 45

[145] A. van Hameren and M. Serino, *BCFW recursion for TMD parton scattering*, *J. High Energy Phys.* **2015** (jul, 2015) 10. 45

[146] A. van Hameren, *Calculating off-shell one-loop amplitudes for k_T-dependent factorization: a proof of concept*, 1710.07609. 45

[147] V. Del Duca, L. Dixon and F. Maltoni, *New color decompositions for gauge amplitudes at tree and loop level*, *Nucl. Phys. B* **571** (apr, 2000) 51–70, [9910563]. 45, 59, 106

[148] F. Maltoni, K. Paul, T. Stelzer and S. Willenbrock, *Color flow decomposition of QCD amplitudes*, *Phys. Rev.* **D67** (2003) 014026, [hep-ph/0209271]. 45, 47

[149] L. J. Dixon, *Calculating scattering amplitudes efficiently*, in *QCD and beyond. Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics, TASI-95, Boulder, USA, June 4-30, 1995*, pp. 539–584, 1996. hep-ph/9601359. 45

[150] M. G. A. Buffing, M. Diehl and T. Kasemets, *Transverse momentum in double parton scattering: factorisation, evolution and matching*, *JHEP* **01** (2018) 044, [1708.03528]. 52

[151] I. Balitsky and A. Tarasov, *Rapidity evolution of gluon TMD from low to moderate x*, *JHEP* **10** (2015) 017, [1505.02151]. 52

[152] I. Balitsky and A. Tarasov, *Gluon TMD in particle production from low to moderate x*, *JHEP* **06** (2016) 164, [1603.06548]. 52

[153] Y. V. Kovchegov and M. D. Sievert, *Valence Quark Transversity at Small x*, *Phys. Rev.* **D99** (2019) 054033, [1808.10354]. 52

[154] I. Balitsky, *Operator expansion for high-energy scattering*, *Nucl. Phys.* **B463** (1996) 99–160, [hep-ph/9509348]. 52

[155] J. Jalilian-Marian, A. Kovner, A. Leonidov and H. Weigert, *The BFKL equation from the Wilson renormalization group*, *Nucl. Phys. B* **504** (1997) 415–431, [9701284]. 52

[156] J. Jalilian-Marian, A. Kovner, A. Leonidov and H. Weigert, *The Wilson renormalization group for low x physics: Towards the high density regime*, *Phys. Rev.* **D59** (1998) 014014, [hep-ph/9706377]. 52

[157] E. Iancu, A. Leonidov and L. D. McLerran, *Nonlinear gluon evolution in the color glass condensate. 1.*, *Nucl. Phys.* **A692** (2001) 583–645, [hep-ph/0011241]. 52

[158] E. Iancu, A. Leonidov and L. D. McLerran, *The Renormalization group equation for the color glass condensate*, *Phys. Lett.* **B510** (2001) 133–144, [hep-ph/0102009]. 52

[159] E. Ferreiro, E. Iancu, A. Leonidov and L. McLerran, *Nonlinear gluon evolution in the color glass condensate. 2.*, *Nucl. Phys.* **A703** (2002) 489–538, [hep-ph/0109115]. 52

[160] H. Weigert, *Unitarity at small Bjorken x*, *Nucl. Phys.* **A703** (2002) 823–860, [hep-ph/0004044]. 52

[161] D. Kharzeev, Y. V. Kovchegov and K. Tuchin, *Cronin effect and high-pT suppression in pA collisions*, *Phys. Rev. D* **68** (nov, 2003) 094013, [hep-ph/0307037]. 52

[162] C. Marquet, *Forward inclusive dijet production and azimuthal correlations in p(A) collisions*, *Nucl. Phys.* **A796** (2007) 41–60, [0708.0231]. 53

[163] H. Fujii, F. Gelis and R. Venugopalan, *Quark pair production in high energy pA collisions: General features*, *Nucl. Phys. A* **780** (dec, 2006) 146–174, [hep-ph/0603099]. 53

[164] Y. V. Kovchegov, J. Kuokkanen, K. Rummukainen and H. Weigert, *Subleading-N(c) corrections in non-linear small-x evolution*, *Nucl. Phys.* **A823** (2009) 47–82, [0812.3238]. 53

[165] C. Marquet and H. Weigert, *New observables to test the Color Glass Condensate beyond the large-N_c limit*, *Nucl. Phys.* **A843** (2010) 68–97, [1003.0813]. 53

[166] A. Dumitru, J. Jalilian-Marian, T. Lappi, B. Schenke and R. Venugopalan, *Renormalization group evolution of multi-gluon correlators in high energy QCD*, *Phys. Lett. B* **706** (dec, 2011) 219–224, [1108.4764]. 53

[167] E. Iancu and D. N. Triantafyllopoulos, *JIMWLK evolution in the Gaussian approximation*, *JHEP* **04** (2012) 025, [1112.1104]. 53

[168] M. Alvioli, G. Soyez and D. N. Triantafyllopoulos, *Testing the Gaussian Approximation to the JIMWLK Equation*, *Phys. Rev.* **D87** (2013) 014016, [1212.1656]. 53

[169] M. Bury, "GTMDCALC." code on request. 54

[170] J. G. M. Kuijf, *Multiparton production at hadron colliders*. PhD thesis, Leiden U., 1991. 59, 106

[171] M. Mangano, S. Parke and Z. Xu, *Duality and multi-gluon scattering*, *Nucl. Phys. B* **298** (mar, 1988) 653–672. 63

[172] A. van Hameren, P. Kotko, K. Kutak, C. Marquet, E. Petreska and S. Sapeta, *Forward di-jet production in p+Pb collisions in the small-x improved TMD factorization framework*, *J. High Energy Phys.* **2016** (dec, 2016) 34, [1607.03121]. 67

[173] C. Marquet, E. Petreska and C. Roiesnel, *Transverse-momentum-dependent gluon distributions from JIMWLK evolution*, *J. High Energy Phys.* **2016** (oct, 2016) 65. 67

[174] H. Jung, S. Baranov, M. Deak, A. Grebenyuk, F. Hautmann, M. Hentschinski et al., *The CCFM Monte Carlo generator CASCADE Version 2.2.03*, *Eur. Phys. J. C* **70** (dec, 2010) 1237–1249, [1008.0152]. 68, 73, 81, 86

[175] H. Jung, *The CCFM Monte Carlo generator CASCADE*, *Comput. Phys. Commun.* **143** (2002) 100–111, [hep-ph/0109102]. 68, 81

[176] M. Dobbs and J. B. Hansen, *The HepMC C++ Monte Carlo event record for High Energy Physics*, *Comput. Phys. Commun.* **134** (2001) 41–46. 68, 86

[177] M. Bengtsson, T. Sjostrand and M. van Zijl, *Initial State Radiation Effects on W and Jet Production*, *Z. Phys.* **C32** (1986) 67. 68, 69

[178] H. Jung and G. P. Salam, *Hadronic final state predictions from CCFM: The Hadron level Monte Carlo generator CASCADE*, *Eur. Phys. J.* **C19** (2001) 351–360, [hep-ph/0012143]. 68

[179] S. Platzer and M. Sjodahl, *The Sudakov Veto Algorithm Reloaded*, *Eur. Phys. J. Plus* **127** (2012) 26, [1108.6180]. 68

[180] G. A. Chirilli, B.-W. Xiao and F. Yuan, *One-loop Factorization for Inclusive Hadron Production in pA Collisions in the Saturation Formalism*, *Phys. Rev. Lett.* **108** (2012) 122301, [1112.1061]. 70

[181] T. Altinoluk, N. Armesto, G. Beuf, A. Kovner and M. Lublinsky, *Single-inclusive particle production in proton-nucleus collisions at next-to-leading order in the hybrid formalism*, *Phys. Rev.* **D91** (2015) 094016, [1411.2869]. 70

[182] A. Dumitru, A. Hayashigaki and J. Jalilian-Marian, *The color glass condensate and hadron production in the forward region*, *Nucl. Phys. A* **765** (feb, 2006) 464–482, [hep-ph/0506308]. 70

[183] A. van Hameren, K. Kutak and T. Salwa, *Scattering amplitudes with off-shell quarks*, *Phys. Lett. B* **727** (nov, 2013) 226–233, [1308.2861]. 70

[184] M. A. Nefedov, V. A. Saleev and A. V. Shipilova, *Dijet azimuthal decorrelations at the LHC in the parton Reggeization approach*, *Phys. Rev. D* **87** (may, 2013) 094030, [1304.3549]. 70

[185] CMS collaboration, S. Chatrchyan et al., *Measurement of the inclusive production cross sections for forward jets and for dijet events with one forward and one central jet in pp collisions at* $\sqrt{s} = 7$ *TeV*, *JHEP* **1206** (2012) 036, [1202.0704]. 70, 71, 72

[186] CMS collaboration, C. Collaboration, *Measurement of the double-differential inclusive jet cross section at* $\sqrt{s} = 13$ *TeV*, . 70, 71, 72

[187] K. Kutak, R. Maciula, M. Serino, A. Szczurek and A. van Hameren, *Four-jet production in single- and double-parton scattering within high-energy factorization*, *JHEP* **04** (2016) 175, [1602.06814]. 70, 71, 78, 79

[188] S. Sapeta and M. Bury, "FORWARD." code on request. 70

[189] CMS collaboration, S. Chatrchyan et al., *The CMS Experiment at the CERN LHC*, *JINST* **3** (2008) S08004. 71

[190] A. van Hameren, *KaTie : For parton-level event generation with* k_T*-dependent initial states*, *Comput. Phys. Commun.* **224** (2018) 371–380, [1611.00680]. 73, 86

[191] A. Buckley, J. Butterworth, L. Lonnblad, D. Grellscheid, H. Hoeth, J. Monk et al., *Rivet user manual*, *Comput. Phys. Commun.* **184** (2013) 2803–2819, [1003.0694]. 73, 81, 86

[192] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J.* **C72** (2012) 1896, [1111.6097]. 73, 80

[193] N. Armesto, *Nuclear shadowing*, *J. Phys.* **G32** (2006) R367–R394, [hep-ph/0604108]. 73

[194] M. Bury, A. van Hameren, H. Jung, K. Kutak, S. Sapeta and M. Serino, *Calculations with off-shell matrix elements, TMD parton densities and TMD parton showers*, *Eur. Phys. J.* **C78** (2018) 137, [1712.05932]. 73

[195] J. R. Gaunt and W. J. Stirling, *Double Parton Scattering Singularity in One-Loop Integrals*, *JHEP* **06** (2011) 048, [1103.1888]. 78

[196] S. Bansal et al., *Progress in Double Parton Scattering Studies*, in *Workshop on Multi-Parton Interactions at the LHC (MPI @ LHC 2013) Antwerp, Belgium, December 2-6, 2013*, 2014. 1410.6664. 78

[197] A. van Hameren, R. Maciula and A. Szczurek, *Single-parton scattering versus double-parton scattering in the production of two $c\bar{c}$ pairs and charmed meson correlations at the LHC*, *Phys. Rev.* **D89** (2014) 094019, [1402.6972]. 78

[198] R. Maciula and A. Szczurek, *Double-parton scattering contribution to production of jet pairs with large rapidity separation at the LHC*, *Phys. Rev.* **D90** (2014) 014022, [1403.2595]. 78

[199] R. Maciuła and A. Szczurek, *Searching for and exploring double-parton scattering effects in four-jet production at the LHC*, *Phys. Lett.* **B749** (2015) 57–62, [1503.08022]. 78

[200] *Proceedings of the Sixth International Workshop on Multiple Partonic Interactions at the Large Hadron Collider*, 2015. 78

[201] M. Diehl, D. Ostermeier and A. Schafer, *Elements of a theory for multiparton interactions in QCD*, *JHEP* **03** (2012) 089, [1111.0910]. 78

[202] B. Blok, Yu. Dokshitzer, L. Frankfurt and M. Strikman, *Perturbative QCD correlations in multi-parton collisions*, *Eur. Phys. J.* **C74** (2014) 2926, [1306.3763]. 78

[203] B. Blok, Yu. Dokshitzer, L. Frankfurt and M. Strikman, *Origins of Parton Correlations in Nucleon and Multi-Parton Collisions*, 1206.5594. 78

[204] B. Blok, Yu. Dokshitser, L. Frankfurt and M. Strikman, *pQCD physics of multiparton interactions*, *Eur. Phys. J.* **C72** (2012) 1963, [1106.5533]. 78

[205] K. Golec-Biernat, E. Lewandowska, M. Serino, Z. Snyder and A. M. Stasto, *Constraining the double gluon distribution by the single gluon distribution*, *Phys. Lett.* **B750** (2015) 559–564, [1507.08583]. 78

[206] LHCb collaboration, R. Aaij et al., *Observation of J/ψ pair production in pp collisions at $\sqrt{s} = 7TeV$*, *Phys. Lett.* **B707** (2012) 52–59, [1109.0963]. 78

[207] LHCb collaboration, R. Aaij et al., *Observation of double charm production involving open charm in pp collisions at $\sqrt{s} = 7$ TeV*, *JHEP* **06** (2012) 141, [1205.0975]. 78

[208] D0 collaboration, V. M. Abazov et al., *Double parton interactions in γ+3 jet events in pp^- bar collisions $\sqrt{s} = 1.96$ TeV.*, *Phys. Rev.* **D81** (2010) 052012, [0912.5104]. 78

[209] CDF collaboration, F. Abe et al., *Double parton scattering in $\bar{p}p$ collisions at $\sqrt{s} = 1.8TeV$*, *Phys. Rev.* **D56** (1997) 3811–3832. 78

[210] A. Martin, W. Stirling and R. Thorne, *MRST partons generated in a fixed-flavour scheme*, *Phys. Lett. B* **636** (may, 2006) 259–264, [0603143]. 80

[211] B. Ducloue, L. Szymanowski and S. Wallon, *Evaluating the double parton scattering contribution to Mueller-Navelet jets production at the LHC*, *Phys. Rev.* **D92** (2015) 076002, [1507.04735]. 80

[212] CMS collaboration, V. Khachatryan et al., *Dijet Azimuthal Decorrelations in pp Collisions at $\sqrt{s} = 7$ TeV*, *Phys. Rev. Lett.* **106** (2011) 122003, [1101.5029]. 81, 82

[213] H. Jung, *k(t) factorization and CCFM: The Solution for describing the hadronic final states: Everywhere?*, *Mod. Phys. Lett.* **A19** (2004) 1–18, [hep-ph/0311249]. 83

[214] F. Hautmann, H. Jung, A. Lelek, V. Radescu and R. Zlebcik, *Collinear and TMD Quark and Gluon Densities from Parton Branching Solution of QCD Evolution Equations*, *JHEP* **01** (2018) 070, [1708.03279]. 83, 85

[215] F. Hautmann, H. Jung, A. Lelek, V. Radescu and R. Zlebcik, *Soft-gluon resolution scale in QCD evolution equations*, *Phys. Lett.* **B772** (2017) 446–451, [1704.01757]. 83

[216] A. van Hameren, P. Kotko and K. Kutak, *Three jet production and gluon saturation effects in p-p and p-Pb collisions within high-energy factorization*, *Phys. Rev.* **D88** (2013) 094001, [1308.0452]. 83

[217] P. J. Mulders and J. Rodrigues, *Transverse momentum dependence in gluon distribution and fragmentation functions*, *Phys. Rev.* **D63** (2001) 094021, [hep-ph/0009343]. 84

[218] E. Iancu, A. H. Mueller and D. N. Triantafyllopoulos, *CGC factorization for forward particle production in proton-nucleus collisions at next-to-leading order*, *JHEP* **12** (2016) 041, [1608.05293]. 85

[219] E. Iancu, J. D. Madrigal, A. H. Mueller, G. Soyez and D. N. Triantafyllopoulos, *Resumming double logarithms in the QCD evolution of color dipoles*, *Phys. Lett.* **B744** (2015) 293–302, [1502.05642]. 85

[220] S. P. Baranov, H. Jung, A. V. Lipatov and M. A. Malyshev, *Associated production of Z bosons and b-jets at the LHC in the combined k_T + collinear QCD factorization approach*, *Eur. Phys. J.* **C77** (2017) 772, [1708.07079]. 85

[221] I. Balitsky and G. A. Chirilli, *Next-to-leading order evolution of color dipoles*, *Phys. Rev.* **D77** (2008) 014019, [0710.4330]. 85

[222] T. Lappi and H. Mäntysaari, *Direct numerical solution of the coordinate space Balitsky-Kovchegov equation at next to leading order*, *Phys. Rev.* **D91** (2015) 074016, [1502.02400]. 85

[223] M. Hentschinski, A. Kusina, K. Kutak and M. Serino, *TMD splitting functions in k_T factorization: the real contribution to the gluon-to-gluon splitting*, *Eur. Phys. J.* **C78** (2018) 174, [1711.04587]. 85