Phénoménologie des distributions de partons généralisées grâce à la diffusion Compton profondément virtuelle

Hervé Dutrieux

université
PARIS-SACLAY

Phenomenology of generalised parton distributions from deeply virtual Compton scattering

Phénoménologie des distributions de partons généralisées grâce à la diffusion Compton profondément virtuelle

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n° 576 Particules, Hadrons, Énergie, Noyau, Instrumentation, Imagerie, Cosmos et Simulation (PHENIICS)

Spécialité de doctorat : Physique hadronique

Graduate School : Physique, *Référent* : Faculté des sciences d'Orsay

Thèse préparée au **Département de Physique Nucléaire** (Université Paris-Saclay, CEA) sous la direction de **Hervé MOUTARDE**, directeur de recherche au CEA

Thèse soutenue à Paris-Saclay, le 21 octobre 2022, par

Hervé DUTRIEUX

Composition du jury

Samuel WALLON Professeur, Université Paris-Saclay, IJCLab	Président
Krešimir KUMERIČKI Professeur, Université de Zagreb	Rapporteur & Examinateur
Lech SZYMANOWSKI Professeur, Centre national pour la recherche nucléaire, Varsovie	Rapporteur & Examinateur
Emilie MAURICE Maître de conférences, Institut Polytechnique de Paris	Examinatrice
Kornelija PASSEK-KUMERIČKI Chargée de recherches, Institut Ruđer Bošković	Examinatrice
Hervé MOUTARDE Directeur de recherche, Université Paris-Saclay, CEA	Directeur de thèse

Thèse de doctorat

NNT : 2022UPASP083

université PARIS-SACLAY

ÉCOLE DOCTORALE
Particules, hadrons, énergie et noyau : instrumentation, imagerie, cosmos et simulation (PHENIICS)

Titre : Phénoménologie des distributions de partons généralisées grâce à la diffusion Compton profondément virtuelle

Mots clés : Distributions de partons généralisées, problème de déconvolution, QCD, structure du nucléon, modélisation par réseau de neurones

Résumé : Les distributions de partons généralisées (GPD) contiennent une riche information sur la structure des hadrons. Elles décrivent notamment des distributions de quarks et de gluons tridimensionnelles ainsi que les distributions en énergie et en pression dans le milieu hadronique. Ces propriétés motivent un effort théorique et expérimental important, qui se concrétise notamment par la construction de nouvelles installations expérimentales à grande échelle comme le collisionneur électron-ion (EIC).
Les GPD sont étudiées expérimentalement au travers de processus exclusifs, dont notamment la diffusion Compton profondément virtuelle (DVCS) qui est considérée comme l'un des processus les mieux établis théoriquement pour accéder aux GPD. La relation entre les GPD et les données expérimentales DVCS est cependant complexe, et nécessite notamment de résoudre un problème de déconvolution. Nous présentons dans ce document la première étude systématique des caractéristiques de ce problème à l'ordre sous-dominant en perturbation. Nous introduisons la notion de "shadow distributions" comme un outil quantitatif pour mesurer la difficulté de la procédure de déconvolution. Il s'agit aussi d'un outil de modélisation intéressant pour effectuer des extractions de GPD tout en garantissant leurs propriétés théoriques correctes. Afin de réduire la dépendence de modèle, nous utilisons aussi des techniques de modélisation par réseaux de neurones.
Nous étudions en détail la possibilité d'extraire les propriétés mécaniques d'une manière moins dépendante de modèle que les études actuelles, et nous quantifions l'effet des futures installations envisagées à la fois sur l'incertitude expérimentale du DVCS et sur l'extraction des GPD par la procédure de déconvolution.

Title : Phenomenology of generalised parton distributions from deeply virtual Compton scattering

Keywords : Generalised parton distributions, deconvolution problem, QCD, nucleon structure, neural network modelling

Abstract : Generalised parton distributions (GPDs) contain a wealth of information about the structure of hadrons. In particular, they describe three-dimensional distributions of quarks and gluons as well as the energy and pressure distributions in the hadronic medium. These properties motivate a major theoretical and experimental effort, which is reflected in the construction of new large-scale experimental facilities such as the electron-ion collider (EIC).
GPDs are studied experimentally through exclusive processes, including in particular deeply virtual Compton scattering (DVCS) which is considered as one of the best theoretically established processes to access GPDs. The relationship between GPDs and experimental DVCS data is however complex, requiring in particular the solution of a deconvolution problem. In this paper we present the first systematic study of the characteristics of this problem at 1-loop in perturbation. We introduce the notion of shadow distributions as a quantitative tool to measure the difficulty of the deconvolution procedure. They represent also an interesting modelling tool to perform GPD extractions while guaranteeing their theoretically correct properties. To achieve lesser model dependence, we make use of neural networks modelling techniques.
We investigate in detail the possibility of extracting mechanical properties in a less model-dependent way than current studies, and quantify the effect of the possible future facilities on both the experimental uncertainty of the DVCS and on the extraction of GPDs by the deconvolution procedure.

Acknowledgments

First of all, I would like to express my gratitude to Hervé Moutarde, who provided a remarkable guidance to this work. He was always available to discuss scientific results, and offered countless insights and valuable advice despite the increasing weight of administrative tasks he had to bear over the years. I am particularly grateful to him for letting me explore personal research avenues significantly different from the Ph.D. subject for which he had initially recruited me. Beyond his scientific supervision, I will keep a moving memory of the warmth of the human relationship we built over my stay at DPhN, in spite of the many obstacles presented by the COVID pandemic.

I owe a great debt to Cédric Mezrag, who provided me with wonderful scientific counseling and made me feel welcomed in the lab as a friend and colleague more than a student. His kindness and willingness to answer any questions as long as necessary was extremely valuable to me. I will remain struck and extremely grateful for the energy he put into accompanying my search of postdoc positions at the end of this Ph.D.

My warmest thanks also go of course to Valerio Bertone, whose extreme precision and availability to answer any scientific question with the most remarkable promptitude were so precious to me. I would like to also thank Maxime Defurne, who kindly answered all my questions about experimental aspects of QCD, and who also saved me several times from the pain of evening RER rides. Francesco Bossù was kind enough to pretend to listen seriously to all my discussions on astronomy or gigantic infrastructure. I am grateful for his availability for any questions and his great uplifting spirit.

Overall, I feel extremely privileged to have been part of the LSN and DPhN, where it seems that everyone rivals each other in availability, open-mindedness and kindness in the service of one of the most interesting scientific subjects I can imagine. Of course, none of this could work without Franck Sabatié, whose scientific point of view I always felt honored to listen to during our lunches, and without the support of Danielle Coret and Isabelle Richard who greatly facilitate all the administrative steps.

I am very grateful to Lech Szymanowski and Krešimir Kumerički for having refereed my Ph.D. dissertation, and to Samuel Wallon, Kornelija Passek-Kumerički and Emilie Maurice for having accepted to be part of my jury. The defense and the discussions we had afterwards were truly impactful for me.

How can I not mention the remarkable friendships I developed with other students during this stay ? First of all with Aurore, with whom I shared an office for three years, and led so many interesting and random discussions, or improvised karaoke sessions. In spite of her calm and seriousness, she seems to have put up with my presence for three years, which must have required an infinite patience probably made easier by the remarkable antipodality of our working hours.

I would also like to thank Michael for all the great times we had together, from the depths of a jazzy oubliette in the center of Paris to his great culinary skills at the Thanksgiving party. The latter are only matched by Elie's twelve-egg omelette of which only he holds the deadly secret. I will remember fondly our games of ping-pong and table soccer, also with Andrea, Samy, Niv, Mostafa and many others.

Table of contents

Detailed summary

Generalised parton distributions

Factorisation theorems provide one of the few practical ways to characterise hadron structure from experimental data. They allow one to decompose scattering amplitudes in terms of parton distributions, which are non-perturbative objects containing universal information on the structure of the hadron target, and process-dependent coefficient functions computable in perturbation theory. Usual parton distribution functions (PDFs), which enter notably the description of deeply inelastic scattering, receive a probabilistic interpretation as the number density of partons (quarks and gluons) depending on the fraction of longitudinal momentum x they carry in a fast moving hadron. Exclusive processes – where all particles involved in the final state of the interaction are measured – like deeply virtual Compton scattering (DVCS) or deeply virtual meson production (DVMP) motivate the introduction of parton distributions of higher dimension, known as generalised parton distributions (GPDs). Contrary to PDFs, GPDs may be used to describe scattering amplitudes where a momentum transfer to the hadron target occurs. As a consequence, in addition to the longitudinal momentum x, they also depend on the longitudinal momentum transfer measured by ξ, and the total momentum transfer measured by t. PDFs are recovered from GPDs in the so-called forward limit, where no momentum is transferred to the hadron target ($\xi = t = 0$). GPDs also generalise elastic form factors which are obtained by integrating GPDs over x.

In addition to the three variables x, ξ and t, and similarly to PDFs, GPDs depend on a renormalisation scale μ^2 which arises from the necessity to eliminate the UV divergence of the field operators which formally define GPDs. The dependence of parton distributions on the renormalisation scale can be expressed in perturbation theory thanks to integro-differential equations known as evolution equations.

As generalisations of PDFs, GPDs also possess a probabilistic interpretation. They encode the correlation between the longitudinal momentum fraction x and the transverse position in the hadron with respect to the barycenter of longitudinal momentum. This cartography of the radial distribution of longitudinal momentum is known as hadron tomography, and it is one of the key motivations for the GPD study programme. Another crucial physical motivation is the remarkable possibility to express matrix elements of the energy momentum tensor in terms of gravitational form factors (GFFs) derived from GPDs. The GFFs in turn allow one to define distributions of energy and pressure inside the hadron matter. This subject has been at the center of a considerable theoretical and experimental interest in recent years, and several phenomenological extractions based on various available DVCS datasets with different modelling hypotheses have been published.

The DVCS deconvolution problem

Although GPDs are involved in several different experimental processes, DVCS has received most of the theoretical and experimental attention in recent years. Indeed, it has both the advantage of a significant cross-section – if necessary when taking into account its interference with the Bethe-Heitler process – and a relatively clean theoretical description

– compared for instance with DVMP which involves another non-perturbative function known as distribution amplitude. A factorisation theorem demonstrates that the DVCS amplitude can be parameterised by Compton form factors (CFFs) $\mathcal{F}$ computed from GPDs F in the following generic way

$$\mathcal{F}(\xi,t,Q^2) = \sum_a \int_{-1}^{1} \frac{\mathrm{d}x}{\xi} T^a\left(\frac{x}{\xi}, \frac{Q^2}{\mu^2}, \alpha_s(\mu^2)\right) \frac{F^a(x,\xi,t,\mu^2)}{x^{p_a}}\,, \tag{1}$$

where the sum runs over parton types (with $a = q$ for quarks and g for gluons), T^a is the DVCS coefficient function computable in perturbation theory, $p_a = 0$ of $a = q$ and 1 if $a = g$, α_s is the strong coupling and Q^2 is the virtuality of the photon mediating the interaction between the incoming lepton and hadron target. The question of whether the actual value of F^a can be extracted from the experimental knowledge of $\mathcal{F}$ is known as the DVCS deconvolution problem and is a central element of the study of this document. The question must be refined, as for instance the parity of T^q already implies that only the x-odd part of the GPD H^q, known as singlet quark GPD and denoted by

$$H^{q(+)}(x,\xi,t,\mu^2) = H^q(x,\xi,t,\mu^2) - H^q(-x,\xi,t,\mu^2)\,, \tag{2}$$

contributes to the integral of Eq. (1). Furthermore, the separation of the contributions of the various parton types is notoriously hard to perform. The question is made all the more difficult that GPDs must follow a number of theoretical constraints. Notably their Mellin moments in x must be polynomials in ξ due to Lorentz covariance :

$$\int_{-1}^{1} \mathrm{d}x\, x^n H^q(x,\xi,t,\mu^2) = \sum_{\substack{k=0 \\ k \text{ even}}}^{n+1} H^q_{n,k}(t,\mu^2)\xi^k\,. \tag{3}$$

It is actually equivalent for H^q to satisfy Eq. (3) and to write as an integral transform known as Radon transform

$$H^q(x,\xi,t,\mu^2) = \int \mathrm{d}\alpha\mathrm{d}\beta\, \delta(x-\beta-\alpha\xi)\Big[f^q(\beta,\alpha,t,\mu^2) + \xi\delta(\beta)D^q(\alpha,t,\mu^2)\Big]\,, \tag{4}$$

where the support of $f^q(\beta,\alpha)$ is restricted to $|\alpha| + |\beta| \leq 1$ and that of $D^q(\alpha)$ to $\alpha \in [-1,1]$. f^q is known as a double distribution (DD) and D^q as the q-contribution to the D-term. Although the x and ξ dependence of GPDs is entangled by the requirement of Eq. (3), the α and β dependence of DDs is unconstrained by that relation except for the support property.

A study of the analytical properties of scattering amplitudes with some assumptions on the behaviour of $\mathrm{Im}\,\mathcal{H}(\xi,t,Q^2)$ for $\xi \to 0$ allows one to derive dispersion relations that relate the real and imaginary parts of CFFs

$$\mathrm{Re}\,\mathcal{H}(\xi,t,Q^2) = \frac{1}{\pi}\mathrm{p.v.}\int_0^1 \mathrm{d}\xi'\,\mathrm{Im}\,\mathcal{H}(\xi',t,Q^2)\left(\frac{1}{\xi-\xi'} - \frac{1}{\xi+\xi'}\right) + C_H(t,Q^2)\,, \tag{5}$$

where C_H, independent of ξ, is known as the subtraction constant, and is related to the D-term by

$$C_H(t,Q^2) = \frac{2}{\pi}\int_1^\infty \sum_a \frac{\mathrm{d}\omega}{\omega^{p_a}}\,\mathrm{Im}\,T^a\left(\omega, \frac{Q^2}{\mu^2}, \alpha_s(\mu^2)\right)\int_{-1}^{1}\mathrm{d}z\,\frac{D^a(z,t,\mu^2)}{\omega - z}\,. \tag{6}$$

In the end, the information on GPDs contained in DVCS experimental data is exactly made of the imaginary part of Eq. (1) which constrains solely the β-odd part of the DD denoted by $f^{q(+)}$, and Eq. (6) which constrains solely the D-term.

At leading order (LO) in α_s, the imaginary part of Eq. (1) reads

$$\mathrm{Im}\,\mathcal{H}(\xi,t,Q^2) \overset{LO}{=} \pi \sum_q e_q^2 H^{q(+)}(\xi,\xi,t,\mu^2)\,, \tag{7}$$

$$= \pi \sum_q e_q^2 \int_{(\xi-1)/(1+\xi)}^{1} \mathrm{d}\alpha\, f^{q(+)}(\xi(1-\alpha),\alpha,t,\mu^2)\,, \tag{8}$$

and Eq. (6) yields

$$C^q(t,Q^2) \overset{LO}{=} 2\sum_q e_q^2 \int_{-1}^{1} \mathrm{d}z\, \frac{D^q(z,t,\mu^2)}{1-z}\,, \tag{9}$$

where it is customary to choose μ^2 close to Q^2. We will assume in the following that $\mu^2 = Q^2$ unless explicitly stated otherwise.

Since both Eqs. (8) and (9) involve integrals of the DD and the D-term where one variable is integrated out, it seems at first unlikely that even a perfect knowledge of DVCS experimental data over a large kinematic range will be enough to actually retrieve the β-odd part of the DD and the D-term. However, the knowledge of the dependence of these two objects on μ^2 thanks to evolution equations provides the theoretical argument that demonstrates the possibility to perform this extraction.

Quantifying the uncertainty of the deconvolution problem for the subtraction constant

The LO evolution of the D-term is conveniently expressed thanks to an expansion in terms of Gegenbauer polynomials :

$$D^q(z,t,\mu^2) = (1-z^2)\sum_{\text{odd } n}^{\infty} d_n^q(t,\mu^2) C_n^{(3/2)}(z)\,, \tag{10}$$

$$D^g(z,t,\mu^2) = \frac{3}{2}(1-z^2)^2\sum_{\text{odd } n}^{\infty} d_n^g(t,\mu^2) C_{n-1}^{(5/2)}(z)\,. \tag{11}$$

Then Eq. (9) gives

$$C_H(t,Q^2) \overset{LO}{=} 4\sum_q e_q^2 \sum_{\text{odd } n}^{\infty} d_n^q(t,\mu^2)\,. \tag{12}$$

The interest of this representation is that it is possible to write LO evolution equations as

$$d_n^q(t,\mu^2) \overset{LO}{=} \Gamma_n^{qq}(\mu^2,\mu_0^2) d_n^q(t,\mu_0^2) + \Gamma_n^{qg}(\mu^2,\mu_0^2) d_n^g(t,\mu_0^2)\,, \tag{13}$$

$$d_n^g(t,\mu^2) \overset{LO}{=} \Gamma_n^{gq}(\mu^2,\mu_0^2) d_n^q(t,\mu_0^2) + \Gamma_n^{gg}(\mu^2,\mu_0^2) d_n^g(t,\mu_0^2)\,. \tag{14}$$

The evolution operators Γ_n^{ab} dictate the behaviour of d_n^a as functions of μ^2. If the d_n^a form a linearly independent family of functions of μ^2, then the measurement of DVCS

experimental data on a range of various values Q^2 should give in principle the power to discriminate each d_n^q and reconstruct therefore the full D-term $D^q(\alpha, t, \mu^2)$ from the simple knowledge of $C_H(t, Q^2)$.

The assessment of the practical possibility to perform such extraction, both with respect to the current DVCS dataset and the expected impact of the electron ion collider (EIC), is one of the original contributions of this document. A model independent extraction of the D-term would require d_n^a for all values of $n \in \{1, 3, 5, ...\}$ to be obtained directly from experimental data. As a simplification, we first only allow d_1 to be fitted, and then d_1 and d_3 jointly. By studying the effect of adding a new free parameter in the representation of the D-term, we can quantify the possibility of a lesser biased extraction. We show that under a certain number of modelling assumptions that are used in state-of-the-art phenomenological extractions of the D-term, it is possible to estimate the uncertainty on d_3^q by

$$\Delta d_3^q \approx \frac{3}{4}\sigma \left(1 - \frac{\Gamma_3^{qq}(\mu^2_{\rm max}, \mu^2_{\rm min})}{\Gamma_1^{qq}(\mu^2_{\rm max}, \mu^2_{\rm min})}\right)^{-1} , \tag{15}$$

where $[\mu^2_{\rm min}, \mu^2_{\rm max}]$ represents the available range in Q^2 where precise DVCS subtraction constant data are collected, and σ their typical uncertainty. Eq. (15) highlights that the uncertainty linked to the extraction of the D-term comes from two main sources : uncertainty on the experimental data σ, and uncertainty due to the limited handle in Q^2 here quantified by how similar the evolution of the various d_n is on the available range in Q^2. It is indeed the fact that each d_n behaves differently under evolution which provides the theoretical leverage to solve the deconvolution problem.

We demonstrate how to reinterpret the uncertainty estimate of the deconvolution procedure (15) by introducing what we called "shadow" D-terms. These objects bring exactly no contribution to the subtraction constant $C_H(t, Q^2)$ at the chosen perturbative order for $Q^2 = \mu_0^2$, and their contribution remains below the value of σ on the available range of Q^2. We show that the maximal size of the shadow D-terms which belong to the allowed modelling space for the D-term provides a similar quantification of the uncertainty of the deconvolution procedure compared to (15).

DVCS phenomenology

We show that our estimate is able to give a fair account of the current phenomenology of DVCS. First we present in detail neural network modelling tools we use to give an account of uncertainty propagation in complicated analyses of experimental data. We discuss strategies for the training and validation of the networks and stress the impact of outlier removal. We compare this computationally intensive technique with traditional statistical tools like the Hessian matrix, and show in both cases how fits can accommodate new data points without the need for a lengthy recalculation.

We notably use one of these tools, known as Bayesian reweighting, to investigate the impact of a potential future experimental facility on the experimental uncertainty σ in Eq. (15). We show how data taking with a positron beam at the Thomas Jefferson National Accelerator Facility (JLab) could allow a significant reduction of uncertainty on $\mathrm{Re}\,\mathcal{H}$, which plays a crucial role in the subtraction constant uncertainty, thanks to the measurement of beam charge assymmetries. The assessment of the expected uncertainty reduction on $\mathrm{Re}\,\mathcal{H}$ under hypotheses detailed in the document is shown on Fig. 1.

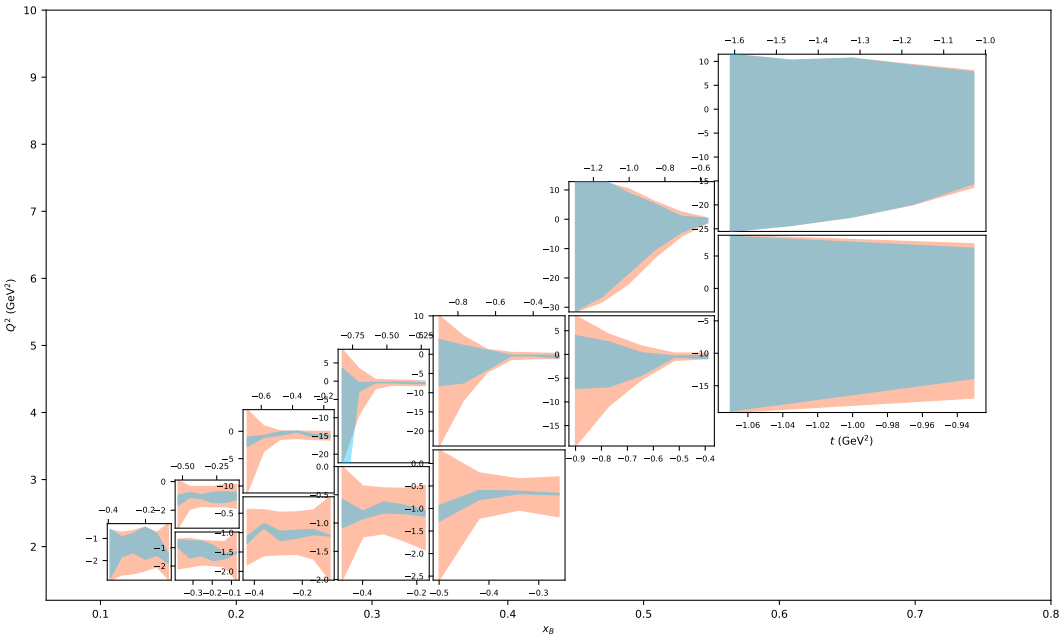

Figure 1 – 68% confidence regions for $\mathrm{Re}\,\mathcal{H}$ with the current DVCS dataset (orange band) and the expected reduced uncertainty thanks to the positron programme at JLab (blue band). For each of the thirteen (x_B, Q^2) bins $[\xi \approx x_B/(2-x_B)]$, the subgraph shows the results on the t bins.

To investigate the uncertainty in the deconvolution procedure linked to the leverage in Q^2 in Eq. (15), we lead a detailed numerical analysis of evolution operators Γ_n^{qq}. This study allows us to understand several intriguing features of our phenomenological extraction of the D-term : we are notably able to explain the $\sim$20 times increase in uncertainty when fitting jointly d_1 and d_3 compared to d_1 alone in our phenomenological extraction of the D-term with the current DVCS dataset, but also the lack of sensitivity to gluon contributions when they are generated radiatively from quark ones. We are furthermore able to predict that with an extended coverage in Q^2 provided by the EIC, a reduction of uncertainty by a factor $\sim$3 on the joint fit of d_1 and d_3 may be expected from the sole effect of the Q^2 range, not taking into account an effect of increased precision of the measurements σ.

We also present the first results of a phenomenological study of the subtraction constant led at next-to-leading (NLO) order on the current DVCS dataset, and conclude that most of the features identified at leading order remain relevant.

The deconvolution problem for double distributions

Having discussed in detail the uncertainty of the deconvolution procedure allowing to recover the D-term from the subtraction constant, we turn to the issue of extracting the DD from the imaginary part of the CFF. We use for this an extension of our concept of shadow D-terms, namely shadow DDs. These are DDs whose contribution to the CFF at the chosen perturbative order is exactly 0 at scale $Q^2 = \mu_0^2$. As in the case of shadow D-terms, the size of the contribution to the CFF of shadow DDs under evolution over the available Q^2 experimental range provides a measure of the deconvolution uncertainty.

We demonstrate in detail the existence of shadow DDs at NLO by using DDs which

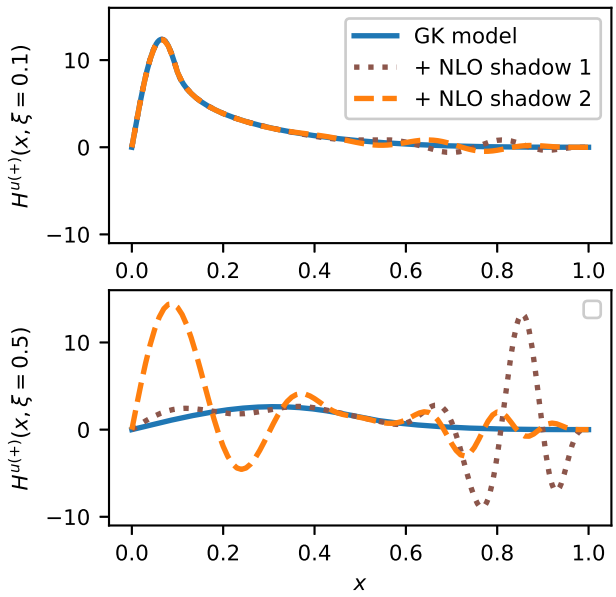

Figure 2 – Three GPDs depicted as functions of x for $\xi = 0.1$ and $\xi = 0.5$ whose NLO CFF at $\xi = 0.1$ differs by less than $\sim 10^{-5}$ although the GPDs themselves differ by $\sim$1.

are polynomials in α and β. We show how this choice allows us to build a series of linear systems with a convenient form, give explicitly the formulas required for the iterative solution of these systems, and provide explicit solutions. We also discuss some advantages and drawbacks of the use of polynomial DDs for the phenomenology of GPDs, and notably the covariant extension programme.

To clarify the effect of LO evolution on shadow DDs, we devote a detailed presentation to the difference between the expansion of the DVCS coefficient function in terms of leading logarithms and leading powers of α_s. Although we are working at NLO, and therefore shadow DDs strictly cancel the terms of order α_s^0 and α_s^1 of the DVCS coefficient function, the resummation of all terms of order $\alpha_s^k \log^k(Q^2/\mu^2)$ for $k \geq 2$ (leading logarithm expansion of the DVCS coefficient function) yields a non-vanishing contributions to the CFF under LO evolution. We provide therefore the first result at order α_s^1 with LO evolution showing that even on a range of $[1, 100]$ GeV^2, the three GPDs presented on Fig. 2 are impossible to tell apart with only DVCS data, as their CFFs are indistinguishable in typical experimental uncertainty. Practically, shadow DDs represent archetypal pitfalls for an unbiased extraction of GPDs : as the flexibility of the GPD model increases, and such shadow GPDs enter the chosen modelling space, the deconvolution uncertainty explodes.

New models of generalised parton distributions

On the goal towards lesser biased GPD phenomenology, neural network models, which are already used to represent CFFs extracted from experimental data, arise as a natural possibility. We demonstrate how it is possible to satisfy many theoretical constraints on GPDs by using a neural network representation of DDs which enforces by design the polynomiality of Mellin moments (see Eq. (3)), expected discrete symmetries, and

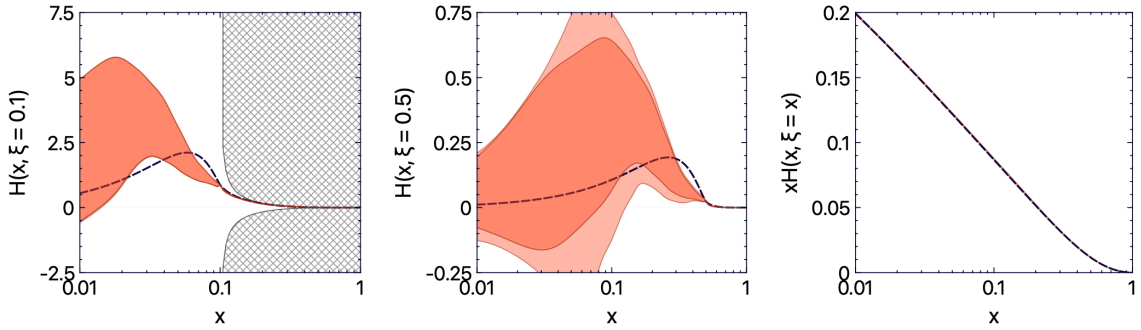

Figure 3 – Demonstration of uncertainty for a neural network model trained to reproduce the forward limit and LO CFF of a phenomenological model (Goloskokov-Kroll) while respecting polynomiality of Mellin moments and a simplified positivity constraint. Comparison between the neural network model (orange band) and the Goloskokov-Kroll model (black line) for $\xi = 0.1$ (left), $\xi = 0.5$ (center) and $\xi = x$ (right). The hatched bands represent the excluded zone by the positivity condition.

reproduces at the same time a given forward limit and LO CFF. An example of the result of such neural network model applied to a popular phenomenological model is demonstrated in Fig. 3. We design a specific DD representation to give an account of the uncertainty related to shadow DDs. We also give a special attention to positivity inequalities, which bring strong constraints on the shape of distributions in the large x region. We notably show how the training procedure allows the implementation of these constraints and stress their impact on the uncertainty of the fit.

We observe that shadow GPDs do not bring significant uncertainty on the small ξ region due to an implicit assumption of regularity of DDs. In the small ξ and x limit, we focus on a popular modelling proposal for GPDs based on the Shuvaev transform, which allows one to approximate small ξ GPDs entirely from their forward limit. This possibility is particularly enticing to constrain gluon PDFs at very low x from measurement of heavy vector meson production in ultra-peripheral collisions. We show how it is possible to reinterpret this modelling proposal thanks to a new evolution code APFEL++. As the validity of the procedure relies crucially on the idea that the large x region of the PDF at some low initial scale μ_0^2 controls the small x and ξ region at a large scale μ^2, we propose a modification of the method to introduce a quantification of its systematic uncertainty. Our quantification of uncertainty consists in an actual measurement of the dominance of the large x region at some initial scale to the evolution of the GPD. It allows us to establish some criteria on how small x and ξ, and how large the energy scale involved in the process should be for the reconstruction of the GPD from its forward limit to appear as a sound procedure.

Introduction

Quantum chromodynamics (QCD) provides the current standard description of the strong force, one of the three fundamental interactions in nature besides gravitation and the electro-weak interaction. Introduced in 1973 by Fritzsch, Leutwyler and Gell-Mann, it gives a formal frame to earlier ideas like Feynman's parton model or the constituent quarks of Gell-Mann and Zweig. QCD is a quantum field theory involving fermionic and bosonic fields – respectively called quarks [1] and gluons – which carry a quantum number known as colour. Because it is a non abelian gauge theory of colour symmetry SU(3) and strong coupling at low energy, the dynamics of QCD is significantly more difficult to elucidate than that of its eminent predecessor, quantum electrodynamics (QED). About 50 years after the introduction of QCD, many challenges remain to access a better understanding of the strong interaction. The fundamental issues of a precise description of confinement – which prevents the observation of isolated coloured objects – and the origin of the mass of nucleons – which stems mostly from spontaneous chiral symmetry breaking [2] – are key elements of a Millenium Prize Problem by the Clay Mathematical Institute entitled "Yang-Mills and mass gap". Other subjects of no lesser interest are for instance the description of nuclei in terms of partonic degrees of freedom, or the study of hadron matter under extreme conditions which prevailed a few microseconds after the Big Bang thanks to the quark-gluon plasma. A final subject of considerable importance in the realm of QCD is the description of the inner structure of hadrons, and notably nucleons. Protons and neutrons are certainly among the most concrete and down to earth objects of study for a physicist. Yet the precise dynamics of quarks and gluons inside such common objects remains largely to unravel. In parallel to the search for exotic physics, the quest for precision in the description of systems as ordinary as nucleons is now a clearly affirmed objective of the high energy physics community. Beyond its intrinsic interest, this precision is anyway necessary for searches of beyond the Standard Model physics, as the uncertainty in the description of hadron states is now a significant part of the experimental uncertainty at the LHC.

Hadron structure is non perturbative in nature due to the large value of the strong coupling at energies of the order of the nucleon mass. A major tool to probe hadron structure while keeping ingredients from perturbation theory is provided by factorisation theorems. They decompose processes in a perturbative and a non-perturbative parts. The latter is encoded in a universal object known as parton distribution which contains crucial information on the inner dynamics of the hadron. We will focus in this work on generalised parton distributions (GPDs) which were introduced more than two decades ago. They contain a remarkable multidimensional information on the structure of hadrons, providing insight on the long-standing question of the distribution of spin among the

1. Quark fields conceived as degrees of freedom of the QCD langrangian are conceptually quite different from the constituent quarks previously mentioned. The latter had mostly been conceived prior to the establishment of QCD as a mathematical encoding of the classification subtended by the Eightfold way introduced by Gell-Mann and Ne'eman. However, in the limit of interaction at low energy, some features of the QCD description of hadrons can be related to a description in terms of constituent quarks.

2. The explicit chiral symmetry breaking due to the coupling of quarks to the Higgs field contributes only at the percent level to the mass of nucleons

hadron constituents, giving a unique access to form factors of the energy-momentum tensor, and mapping the radial distribution of momentum in a fast-moving hadron. Our work mostly concerns a question whose formulation is very simple. What do factorised observables, in particular of deeply virtual Compton scattering (DVCS), tell us exactly on GPDs ? Although this question has been addressed several times since the introduction of GPDs, we find that it has never been studied in the systematic way that we will propose. Among the important contributions of this work are **shadow distributions**, which we introduce as a way to study precisely the possibility of reverting the convolution produced by factorisation. This issue, known as the **deconvolution problem**, will serve as a central theme of the second part of our presentation.

The structure of this document is as follows. In the first part, we will present a brief review of the theory of GPDs and revisit some aspects needed for the developments of the second part. References on our published and preliminary articles are summarised at the end of this introduction.

— The first chapter is devoted to an intuitive presentation of factorisation theorems, the definition of GPDs and their link to non-perturbative objects of different dimensions and the presentation of some selected physical motivations for the study of GPDs.

— The second chapter focuses on important properties of GPDs, like the consequences of Lorentz covariance, the representation in terms of Fock state expansions and a brief evocation of renormalisation which will be the subject of later developments in the document, thanks notably to our publication of a new LO evolution code in momentum space described in Ref. [1].

— The third chapter presents phenomenological aspects of GPDs, by studying their experimental access and some existing modelling strategies.

In the second part, we will put a lot more emphasis on the study of the extraction of GPDs, mostly from DVCS experimental data.

— The fourth chapter presents important aspects of statistical nature to understand our analyses of experimental data and modelling strategies. We apply these tools notably to discuss a study we have published in Refs. [2, 3] to provide a preliminary assessment of the impact of a positron beam at the Thomas Jefferson National Accelerator Facility (JLab) on some aspects of GPD physics.

— The fifth chapter deals in detail with the question of extracting a form factor of the energy-momentum tensor from DVCS data. We present a detailed analysis of the potential of current and future experimental data, and show how this extraction is a prototype of the deconvolution problem, highlighting particularly the role of evolution in its solution. This chapter contains several new results in addition to the presentation of our published phenomenological analysis of the proton mechanical properties in Ref. [4]. We notably demonstrate a new method to estimate the impact of future experimental facilities like the electron ion collider (EIC), introduce a discussion of what we have called shadow D-terms and present new results of a next-to-leading order extraction of the gravitational form factor which will be published in an article in preparation [5].

— The sixth chapter presents our derivation of shadow GPDs as a tool to give a quantitative answer to the full deconvolution problem at next-to-leading order. We show in detail the choices which have led us to the derivation we published in

Refs. [6, 7]. We give an emphasis on the effect of evolution on these objects thanks to the study of the requirements of consistency between DVCS coefficient functions and evolution operators and discuss their consequences for the phenomenology of GPDs.

— The seventh chapter builds upon our understanding of the contribution of shadow GPDs to the modelling uncertainty to propose new models of GPDs. An analytical unpublished model with the ability to exactly reproduce DVCS data at leading order as well as other enticing properties is first derived, and its advantages and limitations are discussed. We then present our modelling strategy of GPDs with neural networks published in Ref. [8]. We end our discussion by focusing on the question of the extrapolation to zero skewness of GPDs. This limit plays a specific role with respect to the deconvolution problem and shadow distributions. We present a reinterpretation of the validity of the Shuvaev transform thanks to the LO GPD evolution code published in Ref. [1] in momentum space. We propose a closely related but nonetheless different strategy so as to give an account of the uncertainty of the reconstruction of the low ξ dependence of GPDs from their forward limit. We discuss the phenomenological interest of our proposal in the context of photo-production of heavy vector mesons. This proposal will soon be submitted to publication [9].

Published articles

[2] H. Dutrieux, V. Bertone, H. Moutarde, P. Sznajder. "Impact of a positron beam at JLab on an unbiased determination of DVCS Compton form factors". *Eur. Phys. J. A* 57.8 (2021)

[3] A. Accardi et al. "An experimental program with high duty-cycle polarized and unpolarized positron beams at Jefferson Lab". *Eur. Phys. J. A* 57.8 (2021)

[4] H. Dutrieux, C. Lorcé, H. Moutarde, P. Sznajder, A. Trawiński, J. Wagner. "Phenomenological assessment of proton mechanical properties from deeply virtual Compton scattering". *Eur. Phys. J. C* 81.4 (2021)

[6] V. Bertone, H. Dutrieux, C. Mezrag, H. Moutarde, P. Sznajder. "Deconvolution problem of deeply virtual Compton scattering". *Phys. Rev. D* 103.11 (2021)

[8] H. Dutrieux, O. Grocholski, H. Moutarde, P. Sznajder. "Artificial neural network modelling of generalised parton distributions". *Eur. Phys. J. C* 82.3 (2022)

Article under review

[1] V. Bertone, H. Dutrieux, C. Mezrag, J.M. Morgado, H. Moutarde. "Revisiting evolution equations for generalised parton distributions". (2022)

Articles in preparation

[5] H. Dutrieux, T. Meisgny, C. Mezrag, H. Moutarde, P. Sznajder. "Proton internal pressure from deeply virtual Compton scattering on collider kinematics".

[9] H. Dutrieux, M. Winn, V. Bertone. "When exclusive meets inclusive at low Bjorken-x_B : how to use exclusive measurements to constrain PDFs based on evolution equations".

Published proceeding

[7] V. Bertone, H. Dutrieux, C. Mezrag, H. Moutarde, P. Sznajder. "Shadow generalized parton distributions : a practical approach to the deconvolution problem of DVCS".

SciPost Phys. Proc. (8 2022)

0.1 . Notations

Recurring abbreviations by alphabetic order

BCA	: beam charge asymmetry
CFF	: Compton form factor
DD	: double distribution
DDVCS	: double deeply virtual Compton scattering
DIS	: deep inelastic scattering
DVCS	: deeply virtual Compton scattering
DVMP	: deeply virtual meson production
EFF	: elastic form factor
EIC	: electron ion collider
EMT	: energy momentum tensor
GFF	: gravitational form factor
GPD	: generalised parton distribution
HVMP	: heavy vector meson production
l.h.s. / r.h.s.	: left hand side / right hand side of an equation
LL / NLL	: leading logarithm / next-to-leading logarithm
LO / NLO	: leading order / next-to-leading order
MLP	: multi-layer perceptron
PDF	: parton distribution function
RDDA	: Radyushkin's double distribution Ansatz
RGE	: renormalisation group equation
TCS	: time-like Compton scattering

Symbols

In a general fashion, superscripts a or b denote the specialisation of a quantity to a parton of type a or b. For instance $H^a(x,\mu^2)$ is a GPD of parton of type a. Sometimes, we will specify it by replacing a with g for gluons and q for quarks, or even $u, d, s, c, ...$ if the quark flavour plays a specific role.

$A(t), B(t), C(t), \bar{C}(t), D_{GFF}(t)$: GFFs

$a(x,\mu^2)$	:	unpolarised PDF of parton of type a (for instance $u(x,\mu^2)$ for quark flavour u, $g(x,\mu^2)$ for gluons).
$\alpha_s(\mu^2)$	:	strong coupling constant
$C_H(t,Q^2)$	:	DVCS subtraction constant
$C_{H,k}[D^q, D^g](t,Q^2)$	:	DVCS subtraction constant computed from the D-terms D^q and D^g at N^{k}LO

$C_n^{(\alpha)}(.)$	:	Gegenbauer polynomial of order n ($\alpha = 3/2$ for quarks and $5/2$ for gluons)
$D^a(z,t,\mu^2)$	:	D-term of parton of type a
$d_n^a(t,\mu^2)$	:	coefficient of the expansion of $D^a(z,t,\mu^2)$ on Gegenbauer polynomials for $C_n^{(\alpha)}$
$\delta(.)$	:	Dirac delta distribution
δ_{jk}	:	Kronecker symbol, equals 1 if $j = k$ and 0 otherwise
Δ	:	four-momentum transfer to the hadron target
e_q	:	electric charge of the quark of flavour q
$f^a(\beta,\alpha,t,\mu^2)$	:	DD of parton of type a
$F_1(t), F_2(t)$	:	Dirac and Pauli EFFs
$\Gamma^{ab}(z/\xi,\xi/x;\mu^2,\mu_0^2)$	:	evolution operator of the GPD H^b at lower scale μ_0^2 to the GPD H^a at upper scale μ^2
$\Gamma^{ab,(k)}(z/\xi,\xi/x;\mu^2,\mu_0^2)$	:	evolution operator of the GPD truncated at perturbative order α_s^k
$\Gamma_n^{ab}(\mu^2,\mu_0^2)$	:	evolution operator of the $d_n^b(t,\mu^2)$ coefficients at lower scale μ_0^2 to the $d_n^a(t,\mu^2)$ coefficient at upper scale μ^2
$H^a, E^a, \widetilde{H}^a, \widetilde{E}^a$	:	GPDs of parton of type a
$H^{a(+)}/H^{a(-)}$	:	singlet / non-singlet components of the GPD
$\mathcal{H}, \mathcal{E}, \widetilde{\mathcal{H}}, \widetilde{\mathcal{E}}$	:	CFFs
M	:	hadron mass
μ^2	:	renormalisation scale
$\mathcal{O}_n^a(\xi,t,\mu^2)$	:	n-th conformal moment of H^a
Ω	:	support of the DDs (also called rhombus)
p_a	:	equals 1 if $a = g$, and 0 otherwise
Q^2	:	square of the lepton four-momentum transfer (usually photon virtuality)
t	:	square of four-momentum transfer to the hadron target Δ^2
$T_k^a(x/\xi,Q^2/\mu^2,\alpha_s(\mu^2))$	:	DVCS coefficient function computed at N^{k}LO for parton of type a
$\Theta(.)$	:	Heaviside step function
x	:	longitudinal momentum fraction
x_B	:	Bjorken's variable
ξ	:	skewness variable

Part I

Generalised parton distributions

1 - An introduction to generalised parton distributions

1.1 . An intuitive take on factorisation theorems

Generalised parton distributions were first introduced in the description of two exclusive experimental processes : deeply virtual Compton scattering (DVCS) [10, 11, 12] and deeply virtual meson production (DVMP) [13, 14] whose diagrammatic representations are given in Fig. 1.1. In both processes, a virtual photon γ^* with large space-like squared four-momentum $q^2 \ll -M^2$ is exchanged usually from an incoming lepton beam with a hadron target of mass M. Contrary to inclusive processes where any final state of the interaction is considered, exclusive processes require that the hadron remains intact in the final state. DVCS and DVMP differ in the additional particle produced in the interaction, a real photon for DVCS and a meson for DVMP. Both processes share however a deep similarity, in that they can be described by factorisation theorems involving GPDs [1] [15, 16, 17]. Ref. [18] provides a clear picture of the physical content of factorisation for inclusive processes and the interpretation of usual parton distribution functions (PDFs) as probability distributions. We also refer to the introduction of Ref. [19] for a very nice intuitive introduction to factorisation. We try here to give a very naive understanding of why deeply virtual processes allow us to access information about the inner structure of hadrons, and why exclusive processes give a more complete picture compared to inclusive ones. We will not dwell on the numerous technicalities of the proofs of factorisation theorems, and postpone a formal definition of GPDs to the next section.

When a deeply virtual photon interacts with a parton inside a hadron, the struck parton receives a large four-momentum transfer. It is usual to introduce $Q^2 \equiv -q^2 \gg M^2$ as the opposite of the squared four-momentum of the virtual photon, and the interaction of the struck parton with the rest of the hadron can be viewed diagrammatically as the exchange of partons of virtuality of the order of Q^2. Their contribution to the cross-section of the process is suppressed by powers of $1/Q$ as the number of exchanged

1. These theorems are demonstrated at leading twist (see later for an intuitive view of the twist expansion) for transverse polarised photons in DVCS and longitudinal polarised photons in DVMP, in specific kinematic regions. We will provide more details in Chapter 3 when coming back on experimental access to GPDs.

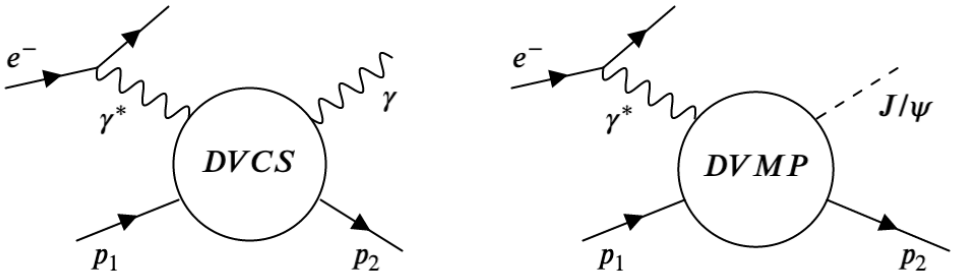

Figure 1.1 – Deeply virtual Compton scattering (left) and deeply virtual production of a J/ψ meson (right) for an incoming electron beam.

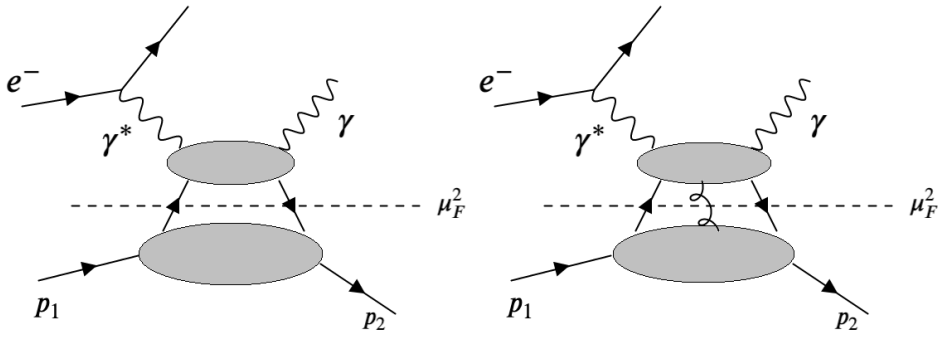

Figure 1.2 – Quark contributions to the factorisation of DVCS in the leading twist (left panel) and next-to-leading twist approximations (right panel). μ_F^2 is the factorisation scale. A similar plot can be produced with gluons instead of quarks.

partons increases [2]. This gives rise to a so-called twist expansion of the process (see for instance Ref. [20] for a precise definition of twist). We show on the left panel in Fig. 1.2 the diagrammatic representation of a leading twist contribution – that is twist-two in this context – to the description of DVCS, where the upper grey area representing the interaction of the struck parton with the virtual photon is separated from the lower grey area representing the rest of the consituents of the hadron as long as a large virtuality is transferred to the struck parton. The right panel of Fig. 1.2 shows a higher twist contribution, where an interaction between the struck parton carrying a large four-momentum and the rest of the hadron is mediated by a gluon.

The region where a large virtuality flows thanks to the interaction with the deeply virtual photon is called **hard part**, whereas the region where only small virtuality is encountered is called **soft part**. The separation between large and small virtuality is somewhat arbitrary and fixed by the choice of a **factorisation scale** μ_F^2. Technically, the factorisation scale arises as the renormalisation scale of the bilocal operator involved in the formal definition of GPDs given in Section 1.2. It corresponds roughly to an upper cut-off on the transverse momentum $k_\perp$ carried by partons within the hadron (see Ref. [18, 21] for instance for more details). Intuitively the renormalisation scale can be viewed as the energy resolution at which the system is described. As it increases, partons are splitted into multiple radiatively generated constituents. It is usual to choose the renormalisation / factorisation scale close to the actual virtuality involved in the process to reduce the necessity of resumming large logarithmic corrections due intuitively to the inadequacy between the resolution at which the system is described and the one at which it is experimentally probed. We will explore aspects of the resummation of logarithmic contributions in more detail in Section 6.3.1.

Thanks the remarkable property of **asymptotic freedom** of QCD [22, 23, 24], it is possible to describe the hard part of the scattering with a perturbative expansion in increasing orders of the strong coupling $\alpha_s(\mu_R^2)$, where μ_R^2 is the renomalisation scale. Fig. 1.3 shows the leading order (LO) and an example of next-to-leading order (NLO) contributions to the description of the hard part where the struck quark radiates a gluon loop. The perturbative expression of the hard part of the scattering is called **coefficient**

2. We do not consider the case of longitudinally polarised gluons whose contribution is contained in the Wilson line which is cancelled by an appropriate gauge fixing procedure.

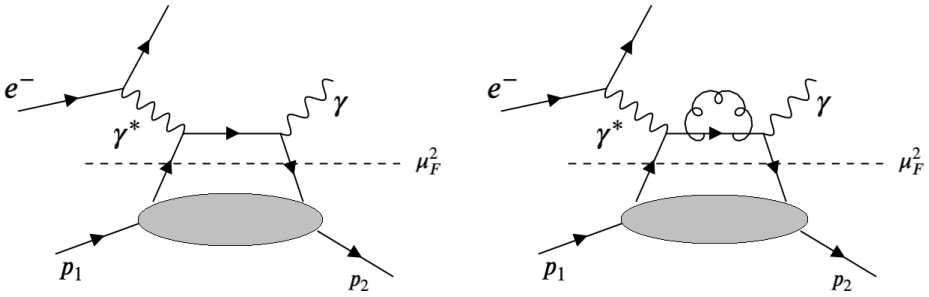

Figure 1.3 – Factorisation of DVCS in the leading twist description at leading order (left panel) and an example of a next-to-leading order contribution (right panel).

function. Different processes like DVCS and DVMP are described by different coefficient functions. As we have previously mentioned, the coefficient functions must depend on the factorisation scale to absorb – at least up to the perturbative order at which it is defined – the consequence of arbitrary variations of scale. Section 6.3.1 will notably derive the practical consequences of this requirement.

On the contrary, the soft part of the interaction, which contains the dynamics of the low energy QCD interaction in the hadron, does not let itself handled within the perturbative formalism. It is described by a non-perturbative parton distribution. Deeply inelastic scattering (DIS) [25] requires only the knowledge of usual PDFs. On the contrary, DVCS and DVMP require the introduction of GPDs. The reason for this refinement is the more complicated kinematic structure of the latter two processes. Whereas structure functions of DIS depend solely on Bjorken's variable x_B, Q^2 and angles characterising the trajectory of the lepton, the fact that the final state of exclusive processes is measured introduces a kinematic dependence on the transfer of four-momentum to the target. GPDs generalise the information contained in PDFs, which only describe the distribution of forward momentum x in the struck hadron, by also depending on the total four-momentum transfer t and its longitudinal component ξ, often called **skewness**.

Unlike the coefficient functions which depend on the process under consideration, parton distributions are universal objects in the sense that the same object is involved in the description of various processes. For instance, GPDs are not only involved in DVCS and DVMP, but also time-like Compton scattering (TCS) [26], double deeply virtual Compton scattering (DDVCS) [27, 28, 29], di-photon production [30, 31, 32] or photon-meson pair production [33, 34].

Already at the stage of this very informal introduction, the reader might wonder what will be the driving question of most of this work : does the combination of the soft and hard part of the scattering through the means of factorisation theorems allow one to actually recover from experimental data the GPD characterising the structure of the proton ? Before dealing with this question, it is however useful to formally introduce GPDs and understand what information they actually contain on hadron structure.

1.2 . Formal definition

We follow the notations of the review Ref. [35]. The formal definition of GPDs is conveniently expressed in light-cone coordinates, defined from the usual time-space coordinates

(z^0, z^1, z^2, z^3) with Minkowskian metric $\eta^{\mu\nu}$ of signature $(+,-,-,-)$ by

$$z^{\pm} = \frac{1}{\sqrt{2}}(z^0 \pm z^3) \quad \text{and} \quad z_\perp = (0, z^1, z^2, 0)\,. \tag{1.1}$$

GPDs are formally defined as integral transforms of bilocal matrix elements, where the quark and gluon fields composing the operator are separated by a light-like distance. For a spin-1/2 hadron, quark GPDs H^q and E^q are defined as

$$\begin{aligned}\frac{1}{2}\int \frac{\mathrm{d}z^-}{2\pi} e^{ixP^+z^-} \left\langle p_2 \middle| \bar{\psi}^q\left(-\frac{z}{2}\right)\gamma^+\psi^q\left(\frac{z}{2}\right)\middle| p_1\right\rangle_{z_\perp=0,z^+=0} \\ = \frac{1}{2P^+}\left(H^q(x,\xi,t)\bar{u}(p_2)\gamma^+u(p_1) + E^q(x,\xi,t)\bar{u}(p_2)\frac{i\sigma^{+\mu}\Delta_\mu}{2M}u(p_1)\right)\,,\end{aligned} \tag{1.2}$$

where p_1 and p_2 denote the incoming and outgoing hadron four-momenta, and we define

$$P \equiv \frac{1}{2}(p_1+p_2)\,, \qquad \Delta \equiv p_2 - p_1\,, \tag{1.3}$$

$$t \equiv \Delta^2\,, \qquad \xi \equiv -\frac{\Delta^+}{2P^+}\,. \tag{1.4}$$

We work in the light-cone gauge, where we do not need to consider Wilson lines which ensure the gauge invariance of the definition. To simplify, we have not made the renormalisation scale apparent until we actually discuss specifically this aspect in Section 2.4. ψ^q is a quark field of flavour q and the Dirac matrices are defined by $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}$, and $\gamma^+ = (\gamma^0+\gamma^3)/\sqrt{2}$. $\sigma^{\mu\nu} = \frac{i}{2}(\gamma^\mu\gamma^\nu - \gamma^\mu\gamma^\nu)$. The normalisation of the spinors is taken as $\bar{u}(p_2)u(p_1) = 2M\delta(\lambda_1,\lambda_2)$ where $\lambda_{1,2}$ are the respective helicities of the incoming and outgoing hadron states and M the hadron mass. The gluon GPDs H^g and E^g are likewise defined as

$$\begin{aligned}\frac{1}{P^+}\int \frac{\mathrm{d}z^-}{2\pi} e^{ixP^+z^-} \left\langle p_2 \middle| G^{+\mu}\left(-\frac{z}{2}\right)G^+_\mu\left(\frac{z}{2}\right)\middle| p_1\right\rangle_{z_\perp=0,z^+=0} \\ = \frac{1}{2P^+}\left(H^g(x,\xi,t)\bar{u}(p_2)\gamma^+u(p_1) + E^g(x,\xi,t)\bar{u}(p_2)\frac{i\sigma^{+\mu}\Delta_\mu}{2M}u(p_1)\right)\,,\end{aligned} \tag{1.5}$$

where $G^{\mu\nu}$ the gluon field strength. It is also possible to define polarised proton GPDs $\widetilde{H}$ and $\widetilde{E}$ by introducing a γ_5 operator in the matrix element. For instance for quarks,

$$\begin{aligned}\frac{1}{2}\int \frac{\mathrm{d}z^-}{2\pi} e^{ixP^+z^-} \left\langle p_2 \middle| \bar{\psi}^q\left(-\frac{z}{2}\right)\gamma^+\gamma_5\psi^q\left(\frac{z}{2}\right)\middle| p_1\right\rangle_{z_\perp=0,z^+=0} \\ = \frac{1}{2P^+}\left(\widetilde{H}^q(x,\xi,t)\bar{u}(p_2)\gamma^+\gamma_5u(p_1) + \widetilde{E}^q(x,\xi,t)\bar{u}(p_2)\frac{\gamma_5\Delta^+}{2M}u(p_1)\right)\,,\end{aligned} \tag{1.6}$$

For a spin-0 hadron, the situation is simplified since only the GPDs H and $\widetilde{H}$ are necessary to parameterise the Fourier transform of the non-local matrix element, giving immediately

$$H^q_{\text{spin }0}(x,\xi,t) = \frac{1}{2}\int \frac{\mathrm{d}z^-}{2\pi} e^{ixP^+z^-} \left\langle p_2 \middle| \bar{\psi}^q\left(-\frac{z}{2}\right)\gamma^+\psi^q\left(\frac{z}{2}\right)\middle| p_1\right\rangle_{z_\perp=0,z^+=0}\,. \tag{1.7}$$

For completeness, let us notice that all GPDs we have presented here are chiral-even, in the sense that the struck parton helicity is conserved. There exist also chiral-odd GPDs, often coined as transversity GPDs, which enter notably the description of exclusive processes with several produced particles in the final state (see for instance Ref. [36]). We will not consider these objects in this document.

Partonic interpretation

GPDs are functions of x, ξ and t. The latter two variables are defined in Eq. (1.4). We notice that t is the traditional Mandelstam variable equal to the square of the total transfer of four-momentum to the hadron, while ξ characterises the transfer of plus-momentum. In a frame where the hadron travels almost at light-speed, the plus-momentum p^+ becomes proportional to the forward momentum p^3 and to the energy p^0. It is therefore frequent to denote the plus-momentum as "longitudinal momentum" or "forward momentum" with an implicit reference to this frame. On the other hand, x is only indirectly defined in Eq. (1.2) as a factor preceding P^+ in the exponential weight $e^{ixP^+z^-}$ of the integral transform.

Using the context of exclusive processes which we introduced in the previous section, it is possible to give a partonic interpretation to GPDs. xP^+ is then shown to be the average plus-momentum of the struck parton, whereas $\Delta^+ = -2\xi P^+$ (1.4) is the plus-momentum transfer to the struck parton, or equivalently the hadron. We deduce then that $(x+\xi)P^+$ is the plus-momentum of the struck parton before interacting with the virtual photon, and $(x-\xi)P^+$ its plus-momentum before re-absorption in the hadron. It is shown in Ref. [37] that GPDs are defined for $(x,\xi) \in [-1,1]^2$. Depending on the relative values of x and ξ, several regions with significantly different properties can be identified for a quark GPD [3].

- If $x \geq |\xi|$, both $x+\xi$ and $x-\xi$ are positive, and the struck parton can be interpreted as a simple quark as in the right panel of Fig. 1.4.
- If on the contrary $x \leq -|\xi|$, then both $x+\xi$ and $x-\xi$ are negative, and the struck parton can be interpreted as an anti-quark with initial plus-momentum $\xi - x$ and final plus-momentum $-x-\xi$ as in the left panel of Fig. 1.4.
- Finally, if $-|\xi| \leq x \leq |\xi|$, $x+\xi \geq 0$ and $x-\xi \leq 0$, so the scattering can actually be interpreted as the annihilation of a pair of a quark and an anti-quark carrying respectively plus-momentum fractions $x+\xi$ and $\xi - x$ as in the central panel of Fig. 1.4.

Since gluons are their own anti-particles, the same reasoning shows that the gluon GPD is even in the variable x.

1.3 . The family tree of parton distributions

1.3.1 . Forward limit

As we have mentioned in Section 1.1, the fact that a non-vanishing four-momentum transfer t is received by the hadron target in DVCS and DVMP makes it necessary to describe the soft part of the interaction thanks to GPDs, and not ordinary PDFs. The link between the two distributions can be observed formally by setting $\Delta = 0$, known as

3. GPDs can be extended for $|\xi|$ larger than 1 provided that $|x| \leq |\xi|$. The extension is then known as generalised distribution amplitudes (GDAs) [38, 39, 40, 41].

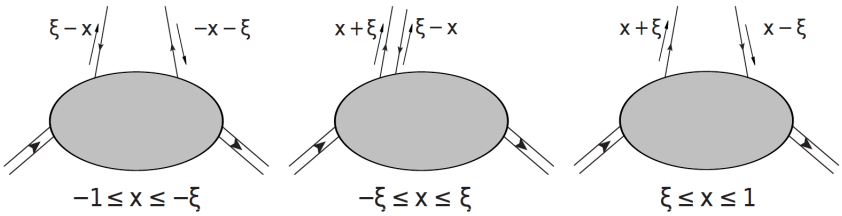

FIGURE 1.4 – Partonic interpretation of GPDs for $\xi \geq 0$, depending on the values of x. The grey area represents the soft part of the factorised description of exclusive processes – similar to the lower grey areas in Fig. 1.2 for instance. For $-1 \leq x \leq -\xi$, the struck parton can be interpreted as an anti-particle which is emitted and later re-absorbed by the hadron (left panel). For $-\xi \leq x \leq \xi$, the hard scattering can be interpreted as the annihilation of a particle anti-particle pair (central panel). For $\xi \leq x \leq 1$, the struck parton can be interpreted as an ordinary particle. Figure taken from Ref. [19].

forward limit in Eq. (1.2) : the obtained matrix element then exactly defines PDFs. We obtain the important result

$$H^q(x, \xi = 0, t = 0) = q(x)\Theta(x) - \bar{q}(-x)\Theta(-x)\,, \tag{1.8}$$
$$H^g(x, \xi = 0, t = 0) = xg(x)\Theta(x) - xg(-x)\Theta(-x)\,, \tag{1.9}$$

where Θ is the Heaviside step function, and $q, g(x)$ are the usual quark and gluon unpolarised PDFs and $\bar{q}(x)$ the unpolarised PDF of $\bar{q}$. Since PDFs are well-known from a wealth of experimental data, notably thanks to inclusive processes like DIS, the fact that they are the forward limit of GPDs is a crucial aspect of the modelling of the latter.

A similar relation involving polarised PDFs can be obtained for $\widetilde{H}$. On the contrary, due to the presence of a factor Δ_μ in front of the E distribution in Eq. (1.2), the forward limit $\Delta = 0$ does not bring any constraint on E. The situation is similar for $\widetilde{E}$. Nonetheless, values of E in this limit which are unaccessible from DIS plays an important role in the determination of the spin decomposition of hadrons, as we will show later.

1.3.2 . Link to elastic form factors

GPDs also generalise the usual elastic form factors (EFFs) F_1^q and F_2^q which parameterise the matrix element of the electromagnetic current at $z = 0$:

$$\langle p_2|\bar{\psi}^q(z)\gamma^\mu\psi^q(z)|p_1\rangle = \bar{u}(p_2)\left[F_1^q(t)\gamma^\mu + F_2^q(t)\frac{i\sigma^{\mu\nu}\Delta_\nu}{2M}\right]u(p_1)\,. \tag{1.10}$$

From $F_1^q(t)$ and $F_2^q(t)$ are defined the Dirac and Pauli form factors $F_1(t)$ and $F_2(t)$ defined by

$$F_1(t) = \sum_q e_q F_1^q(t) \quad \text{and} \quad F_2(t) = \sum_q e_q F_2^q(t)\,, \tag{1.11}$$

where e_q is the electric charge of the quark of flavour q. At $t = 0$, $F_1(0)$ gives the total electric charge of the hadron, and $F_2(0)$ its anomalous magnetic moment.

The link between GPDs and EFFs is obtained by taking the limit $z^- = 0$ in Eq. (1.2), which can be achieved by integrating H^q and E^q over their dependence on x. Indeed,

using the fact that for $\xi \in [-1,1]$, GPDs take non-zero values only for $x \in [-1,1]$,

$$\begin{aligned}
\bar{u}(p_2)\int_{-1}^{1}\mathrm{d}x\,&\left[H^q(x,\xi,t)\gamma^+ + E^q(x,\xi,t)\frac{i\sigma^{+\nu}\Delta_\nu}{2M}\right]u(p_1)\\
&= \int_{-\infty}^{\infty}\mathrm{d}x\,\bar{u}(p_2)\left[H^q(x,\xi,t)\gamma^+ + E^q(x,\xi,t)\frac{i\sigma^{+\nu}\Delta_\nu}{2M}\right]u(p_1) \qquad (1.12)\\
&= P^+\int_{-\infty}^{\infty}\mathrm{d}x\int\frac{\mathrm{d}z^-}{2\pi}e^{ixP^+z^-}\left\langle p_2\middle|\bar{\psi}^q\left(-\frac{z^-}{2}\right)\gamma^+\psi^q\left(\frac{z^-}{2}\right)\middle|p_1\right\rangle, \qquad (1.13)\\
&= P^+\int\frac{\mathrm{d}z^-}{2\pi}\left(\int_{-\infty}^{\infty}\mathrm{d}x\,e^{ixP^+z^-}\right)\left\langle p_2\middle|\bar{\psi}^q\left(-\frac{z^-}{2}\right)\gamma^+\psi^q\left(\frac{z^-}{2}\right)\middle|p_1\right\rangle, \qquad (1.14)
\end{aligned}$$

and the Fourier transform of 1 gives that

$$\int_{-\infty}^{\infty}\mathrm{d}x\,e^{ixP^+z^-} = 2\pi\delta(P^+z^-) = \frac{2\pi}{P^+}\delta(z^-)\,, \qquad (1.15)$$

where δ is the Dirac delta distribution [4], hence

$$\begin{aligned}
\bar{u}(p_2)\int_{-1}^{1}\mathrm{d}x\,&\left[H^q(x,\xi,t)\gamma^+ + E^q(x,\xi,t)\frac{i\sigma^{+\nu}\Delta_\nu}{2M}\right]u(p_1)\\
&= \int\mathrm{d}z^-\delta(z^-)\left\langle p_2\middle|\bar{\psi}^q\left(-\frac{z^-}{2}\right)\gamma^+\psi^q\left(\frac{z^-}{2}\right)\middle|p_1\right\rangle, \qquad (1.16)\\
&= \langle p_2|\bar{\psi}^q(0)\gamma^+\psi^q(0)|p_1\rangle\,. \qquad (1.17)
\end{aligned}$$

We deduce therefore that

$$\int_{-1}^{1}\mathrm{d}x\,H^q(x,\xi,t) = F_1^q(t) \quad \text{and} \quad \int_{-1}^{1}\mathrm{d}x\,E^q(x,\xi,t) = F_2^q(t) \qquad (1.18)$$

This relation has been extensively used to decribe the t-dependence of GPDs.

One can therefore understand why GPDs are called "generalised" since they encompass several well-known sources of non-perturbative information on hadron structure. There exist an even higher order generalisation of parton distributions, known as generalised transverse momentum dependent distributions (GTMDs) [42, 43] as shown on the "family tree" of parton distributions in Fig. 1.5. GTMDs corresponds to unintegrated GPDs over the parton transverse momentum $k_\perp$. Performing this integration amounts to setting $z_\perp = 0$ for GPDs, which we have implicitly done in Eq. (1.2) by only considering z^- in the definition of the matrix element.

1.4 . Why study generalised parton distributions ?

We select in this section two important physical motivations for the study of GPDs which have been at the center of a continued experimental and theoretical interest for

4. We used that P^+ is positive since it corresponds to the four-momentum of a massive particle : $(P^0)^2 - (P^1)^2 - (P^2)^2 - (P^3)^2 = M^2 \geq 0$, so $(P^0)^2 \geq (P^3)^2$, and since $P^0 \geq 0$, we obtain $P^0 \geq |P^3|$. Therefore both P^+ and P^- are positive. This is not true for a state which is not on its mass shell.

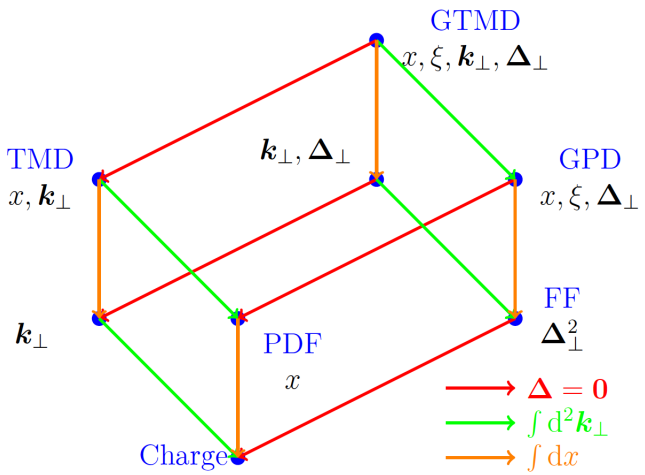

Figure 1.5 – Family tree of parton distributions. The link from GPDs to EFFs (here simply written as FFs) is obtained by integrating the x dependence, whereas the link from GPDs to PDFs is obtained thanks to the limit $\Delta = 0$. Figure taken from Ref. [44], itself inspired from Ref. [45].

several decades : the perspective of hadron tomography, and that of mapping the distributions of energy and momentum inside the hadron. This interest has been strongly reaffirmed in the context of the discussion of new facilities dedicated to nuclear experiments at the electron-ion collider (EIC) [46, 47], Chinese electron-ion collider (EIcC) [48, 49] and large hadron-electron collider (LHeC) [50].

1.4.1 . Hadron tomography : impact parameter distributions

Usual PDFs $a(x)$ possess a probabilistic interpretation as the number density of partons of type a carrying a fraction x of the plus-momentum of the hadron. As generalisations of PDFs, GPDs have an even more interesting probabilistic interpretation in the so-called **impact parameter representation**. At zero skewness $\xi = 0$, the Fourier transform of a GPD with respect to $\Delta_\perp$ is called an impact parameter distribution (IPD) [51, 52]

$$q(x, b_\perp) = \int \frac{\mathrm{d}^2\Delta_\perp}{(2\pi)^2}\, e^{-ib_\perp\cdot\Delta_\perp} H^q(x, 0, t = -\Delta_\perp^2)\,. \tag{1.19}$$

$b_\perp$ is the Fourier conjugate variable of the transverse momentum transfer $\Delta_\perp$. It can be interpreted as the transverse distance to the center of plus-momentum of the hadron $b_\perp^{(0)}$, which is the average of $b_{\perp,i}$ defined with respect to an arbitrary origin weighted by the plus-momentum fraction k_i^+ over all partons i making up the hadron :

$$b_\perp^{(0)} = \frac{\sum k_i^+ b_{\perp,i}}{\sum k_i^+}\,, \tag{1.20}$$

where the sum is running over all partons in the hadron. This formulation can be established rigorously in the wave function representation (see Section 2.3 for a brief presentation of this formalism). Since the incoming hadron carries a plus-momentum $(1+\xi)P^+$ whereas the outgoing one carries $(1-\xi)P^+$, the center of plus-momentum is shifted during the interaction if ξ is non zero. Ref. [52] shows precisely that this shift is proportional to ξ.

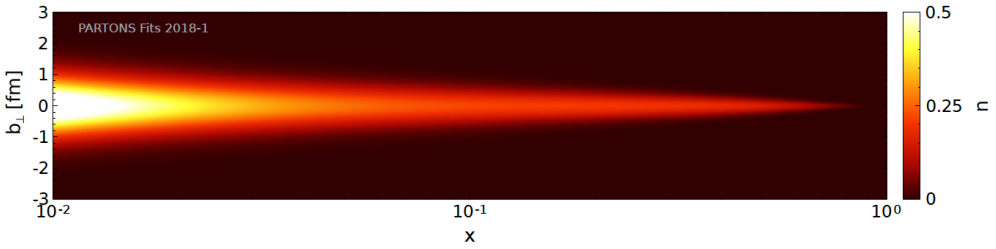

FIGURE 1.6 – Number density of u quarks as a function of x and $b_\perp$ in an unpolarised proton. Figure taken from Ref. [53].

When $\xi = 0$ on the contrary, the IPD of Eq. (1.19) possess a probabilistic interpretation as the number density of quarks with any helicity, plus-momentum fraction x and transverse distance $b_\perp$ from the center of plus-momentum. GPDs therefore allow access to a three-dimensional information on the structure of hadron, in an hybrid space with one dimension of momentum and two spatial dimensions.

The practical extraction of IPDs faces the challenges of accessing a large range of values in t. However, as we will discuss more in the description of experimental processes in Chapter 3, factorisation theorems are only valid for $|t|$ small with respect to Q^2, which restricts considerably the access to moderate values of $|t|$ of the order of several GeV^2. In addition, an extrapolation to $\xi = 0$ is necessary. We will focus on the question of the modelling of GPDs at small ξ in Section 7.3. Nonetheless, extractions from experimental data with constrained parameterisations have been attempted. For instance, we reproduce in Fig. 1.6 the result of the fit of IPDs led in Ref. [53].

1.4.2 . Properties of hadron matter : the energy-momentum tensor

A remarkable feature of GPDs is their relation to the QCD energy-momentum tensor (EMT). It allows not only to shed light on the long-standing puzzle of the nucleon spin decomposition which emerged 30 years ago with the EMC measurements [54], but can also be used to access information about the mechanical properties of hadrons [55, 56], like distributions of pressure inside the nucleon. The possibility of studying mechanical properties of partonic matter was first highlighted in Ref. [57].

In the case of a spin-1/2 hadron, the matrix element of the local gauge-invariant EMT operator can be parameterised in terms of five gravitational form factors (GFFs) $A^a(t)$, $B^a(t)$, $C^a(t)$, $\bar{C}^a(t)$ and $D^a_{GFF}(t)$ [5] as [58, 59, 60]

$$\langle p_2|T^{\mu\nu}_a(0)|\, p_1\rangle = \bar{u}(p_2)\Bigg\{\frac{P^\mu P^\nu}{M}\,A^a(t) + \frac{\Delta^\mu\Delta^\nu - \eta^{\mu\nu}\Delta^2}{M}\,C^a(t) + M\eta^{\mu\nu}\bar{C}^a(t) + \frac{P^{\{\mu}i\sigma^{\nu\}\rho}\Delta_\rho}{4M}\,[A^a(t) + B^a(t)] + \frac{P^{[\mu}i\sigma^{\nu]\rho}\Delta_\rho}{4M}\,D^a_{GFF}(t)\Bigg\}u(p_1)\,, \tag{1.21}$$

where $a^{\{\mu}b^{\nu\}} = a^\mu b^\nu + a^\nu b^\mu$ and $a^{[\mu}b^{\nu]} = a^\mu b^\nu - a^\nu b^\mu$. The connection between GPDs and

5. The choice of notation for the last GFF will become clearer when we introduce the completely independent notion of D-term in the next chapter, which we will denote by D^a.

GFFs is given for quarks by [35]

$$\int_{-1}^{1} \mathrm{d}x\, x H^q(x,\xi,t) = A^q(t) + 4\xi^2 C^q(t)\,, \tag{1.22}$$

$$\int_{-1}^{1} \mathrm{d}x\, x E^q(x,\xi,t) = B^q(t) - 4\xi^2 C^q(t)\,, \tag{1.23}$$

$$\sum_q \int_{-1}^{1} \mathrm{d}x\, \widetilde{H}^q(x,\xi,t) = -\sum_q D^q_{GFF}(t)\,, \tag{1.24}$$

and for gluons by

$$\int_{-1}^{1} \mathrm{d}x\, H^g(x,\xi,t) = A^g(t) + 4\xi^2 C^g(t)\,, \tag{1.25}$$

$$\int_{-1}^{1} \mathrm{d}x\, E^g(x,\xi,t) = B^g(t) - 4\xi^2 C^g(t)\,, \tag{1.26}$$

$$D^g_{GFF}(t) = 0\,. \tag{1.27}$$

Access to the GFF $\bar{C}^a$ is more intricate since it involves higher twist GPDs [61, 59, 62].

Ji's sum rule

It can be shown that the total angular momentum carried by each flavour of quark J^q and gluons J^g verifies for a spin-1/2 hadron [63, 12]

$$J^q = \frac{1}{2}(A^q(0) + B^q(0))\,, \quad \text{and} \quad J^g = \frac{1}{2}(A^g(0) + B^g(0))\,. \tag{1.28}$$

We deduce therefore from Eqs. (1.22) to (1.26) the decomposition of the total spin of the hadron, known as Ji's sum rule

$$\frac{1}{2} = \sum_q J^q + J^g \tag{1.29}$$

$$= \sum_q \frac{1}{2}\int_{-1}^{1} \mathrm{d}x\, x(H^q(x,\xi,0) + E^q(x,\xi,0)) + \frac{1}{2}\int_{-1}^{1} \mathrm{d}x\, H^g(x,\xi,0) + E^g(x,\xi,0)\,. \tag{1.30}$$

The fact that this relation is independent of ξ will be investigated in depth in the next chapter. Note however that the precise decomposition between quark flavours and gluons is renormalisation scale and scheme dependent. This sum rule is not without reminding of the one observed by usual PDFs :

$$1 = \sum_{a=q,\bar{q},g} \int_0^1 \mathrm{d}x\, x a(x)\,. \tag{1.31}$$

Thanks to the probabilistic interpretation of PDFs, this sum rule merely states that when adding the fractional plus-momentum over all partons, we obtain the full plus-momentum of the hadron.

Mechanical properties

We note by boldface characters spatial three-vectors. In the Breit frame where $\boldsymbol{P} = \mathbf{0}$ and $t = -\boldsymbol{\Delta}^2$, Fourier transforms of GFFs with respect to $\boldsymbol{\Delta}$ allow us to measure how energy and momentum are distributed inside the hadron [57, 56, 64]. Denoting by $r = |\boldsymbol{r}|$ the radial coordinate, one can define for instance the radial energy $\varepsilon_a(r)$ distribution in the Breit frame as a Fourier transform of the GFFs $A^a(t)$, $B^a(t)$, $C^a(t)$ and $\bar{C}^a(t)$:

$$\varepsilon_a(r) = M \int \frac{\mathrm{d}^3\boldsymbol{\Delta}}{(2\pi)^3}\, e^{-i\boldsymbol{\Delta}\cdot\boldsymbol{r}} \left\{ A^a(t) + \bar{C}^a(t) + \frac{t}{4M^2}\Big[B^a(t) - 4C^a(t)\Big] \right\} . \tag{1.32}$$

Among several distributions of radial mechanical properties which can be defined, the pressure anisotropy $s_a(r)$ has attracted attention due to the fact that it does not depend on $\bar{C}a(t)$. In theory, it can therefore be completely extracted from leading twist GPDs :

$$s_a(r) = -\frac{4M}{r^2} \int \frac{\mathrm{d}^3\boldsymbol{\Delta}}{(2\pi)^3}\, e^{-i\boldsymbol{\Delta}\cdot\boldsymbol{r}}\, \frac{t^{-1/2}}{M^2} \frac{\mathrm{d}^2}{\mathrm{d}t^2}\big[t^{5/2}\, C^a(t)\big] . \tag{1.33}$$

We will study in more depth this distribution in Section 5.2.4 when considering a phenomenological extraction of $C^a(t)$.

2 - Further properties

2.1 . General behaviour

Domain

We have already mentioned that the definition domain of GPDs consists of $(x, \xi) \in [-1, 1]^2$ [37]. However, as ξ is the related to the plus component of the total four-momentum transfer Δ, it is also bound kinematically by the value of t according to

$$|\xi| \leq \frac{\sqrt{-t}}{\sqrt{-t + 4M^2}} \,. \tag{2.1}$$

In practice, this bound is significantly constraining the available range in ξ, since for $t = -0.2$ GeV^2, a typical value considered in JLab DVCS kinematics, $|\xi|$ is limited to less than 0.23 for the nucleon.

Parity

The study of the effect of time reversal on field operators led for instance in Ref. [21] demonstrates that for spin-0 or 1/2 hadrons most GPDs [1] verify

$$F(x, -\xi, t) = F(x, \xi, t) \,, \tag{2.2}$$

notably for $F = H, E, \widetilde{H}, \widetilde{E}$. Added to the fact that hermicity constraints yield

$$F^*(x, \xi, t) = F(x, -\xi, t) \,, \tag{2.3}$$

where * denotes complex conjugation, one obtains that GPDs are real-valued [2].

The fact that gluons are their own anti-particles implies that gluons GPDs H^g and E^g are even in x, whereas $\widetilde{H}^g$ and $\widetilde{E}^g$ are odd. No such constraint appears on the quark GPDs. It is useful to separate their x-odd and x-even parts, which exhibit different behaviours with respect to the charge operator. We will coin **singlet** GPD $F^{q(+)}$ the part which couples with the gluon distribution, and **non-singlet** GPD $F^{q(-)}$ the part which remains independent from the gluon distribution. For $F = H, E$,

$$F^{q(+)}(x, \xi, t) = F^q(x, \xi, t) - F^q(-x, \xi, t) \,, \tag{2.4}$$

$$F^{q(-)}(x, \xi, t) = F^q(x, \xi, t) + F^q(-x, \xi, t) \,, \tag{2.5}$$

whereas for $\widetilde{F} = \widetilde{H}, \widetilde{E}$,

$$\widetilde{F}^{q(+)}(x, \xi, t) = \widetilde{F}^q(x, \xi, t) + \widetilde{F}^q(-x, \xi, t) \,, \tag{2.6}$$

$$\widetilde{F}^{q(-)}(x, \xi, t) = \widetilde{F}^q(x, \xi, t) - \widetilde{F}^q(-x, \xi, t) \,. \tag{2.7}$$

1. For the quark GPDs in a spin-1/2 hadron, Ref. [21] demonstrates that only the transversity GPD $\widetilde{E}_T^q$ is an odd function of ξ. The situation is more complicated for hadrons of larger spin as the number of GPDs proliferates.

2. $\widetilde{E}_T$ is also real-valued owing to the fact that $(\widetilde{E}_T^q)^*(x, \xi, t) = -\widetilde{E}_T^q(x, -\xi, t)$.

Limit $x \to 1$

In the limit where $x \to 1$, the active parton carries alone most of the hadron plus-momentum, so is itself the center of plus-momentum as defined in Eq. (1.20). Ref. [65], following the idea that in that limit, hard gluon exchanges dominate the structure functions [66] and allow a calculation in perturbative QCD, demonstrates notably that for the pion

$$H^q(x,\xi,t) \sim \frac{(1-x)^2}{1-\xi^2}\,, \tag{2.8}$$

and for the nucleon,

$$H^q_\pi(x,\xi,t) \sim \frac{(1-x)^3}{(1-\xi^2)^2}\,, \tag{2.9}$$

$$E^q_\pi(x,\xi,t) \sim \frac{(1-x)^5}{(1-\xi^2)^2}\,. \tag{2.10}$$

It is remarkable that there is no t dependence in these formulas.

2.2 . Polynomiality of Mellin moments

We have so far explored the link of $\int \mathrm{d}x\, H^q$ with the EFF F_1^q (1.18), and of $\int \mathrm{d}x\, xH^q$ with the GFFs A^q and C^q (1.22). Likewise, $\int \mathrm{d}x\, H^g$ has been related to the GFFs A^g and C^g (1.25). Similar relations have been established for the GPD E, and could be derived as well for $\widetilde{H}$ and $\widetilde{E}$. We refer to reviews such as [35, 21, 67] for complete sets of such relations. These are particular cases of Mellin moments of GPDs, defined by

$$H^a_n(\xi,t) = \int_{-1}^{1} \mathrm{d}x\, x^n H^a(x,\xi,t)\,. \tag{2.11}$$

In a similar way to the demonstration of Eq. (1.17), it is possible to relate these moments to matrix elements of twist-two local operators. For instance for a spin-1/2 hadron,

$$\begin{aligned}\bar{u}(p_2) \int_{-1}^{1} \mathrm{d}x\, x^n \left[H^q(x,\xi,t)\gamma^+ + E^q(x,\xi,t)\frac{i\sigma^{+\nu}\Delta_\nu}{2M}\right] u(p_1) \\ = P^+ \int \frac{\mathrm{d}z^-}{2\pi}\left(\int_{-\infty}^{\infty} \mathrm{d}x\, x^n e^{ixP^+z^-}\right)\left\langle p_2\middle|\bar{\psi}^q\left(-\frac{z^-}{2}\right)\gamma^+\psi^q\left(\frac{z^-}{2}\right)\middle|p_1\right\rangle,\end{aligned} \tag{2.12}$$

and using this time the fact that the Fourier transform of x^n yields

$$\int_{-\infty}^{\infty} \mathrm{d}x\, x^n e^{ixP^+z^-} = 2\pi i^n \delta^{(n)}(P^+z^-) = \frac{2\pi i^n}{(P^+)^{n+1}}\left(\frac{\partial}{\partial z^-}\right)^n\Bigg|_{z^-=0}, \tag{2.13}$$

where $\delta^{(n)}$ is the n-th derivative of the Dirac delta distribution [3], we find

$$\begin{aligned}\bar{u}(p_2) \int_{-1}^{1} \mathrm{d}x\, x^n \left[H^q(x,\xi,t)\gamma^+ + E^q(x,\xi,t)\frac{i\sigma^{+\nu}\Delta_\nu}{2M}\right] u(p_1) \\ = \frac{1}{(P^+)^n}\left\langle p_2\middle|\left(i\frac{\partial}{\partial z^-}\right)^n \bar{\psi}^q(0)\,\gamma^+\psi^q(0)\middle|p_1\right\rangle.\end{aligned} \tag{2.14}$$

3. We have used that $\delta^{(n)}(ax) = \mathrm{sgn}(a)\delta^{(n)}(x)/a^{n+1}$, $P^+ \geq 0$ and $\delta^{(n)}(x) = (-1)^n(\partial^n/\partial x^n)|_{x=0}$. $\mathrm{sgn}(a) = |a|/a$ designates the sign of a.

Eq. (2.14) demonstrates that Mellin moments of GPDs are form factors of some local twist-two operators. A precise study of the requirements imposed by Lorentz covariance on these operators shows that the form factor decomposition can be written as specific combinations of the four-vectors P and Δ with t dependent coefficients $a^q_{n,i}(t)$, $b^q_{n,i}(t)$ and $c^q_n(t)$ [35] :

$$\left\langle p_2 \middle| \left(i\frac{\partial}{\partial z^-}\right)^n \bar{\psi}^q(0)\,\gamma^+\psi^q(0) \middle| p_1 \right\rangle = \bar{u}(p_2)\Bigg[\gamma^+ \sum_{\substack{i=0\\ i\text{ even}}}^{n} a^q_{n,i}(t)(\Delta^+)^i(P^+)^{n-i} + \frac{i\sigma^{+\nu}\Delta_\nu}{2M}\sum_{\substack{i=0\\ i\text{ even}}}^{n} b^q_{n,i}(t)(\Delta^+)^i(P^+)^{n-i} + \frac{\mathrm{mod}(n,2)}{M}c^q_n(t)(\Delta^+)^{n+1}\Bigg]u(p_1)\,, \tag{2.15}$$

where $\mathrm{mod}(n,2)$ equals 1 if n is odd, and 0 if n is even. Finally, with the factor $1/(P^+)^n$ of Eq. (2.14), $\xi = -\Delta^+/(2P^+)$ and the use of the Gordon identity [68] to treat the final term of Eq. (2.15)

$$\frac{P^+}{M}\bar{u}(p_2)u(p_1) = \bar{u}(p_2)\left[\gamma^+ - i\frac{\sigma^{+\nu}\Delta_\nu}{2M}\right]u(p_1)\,, \tag{2.16}$$

we obtain

$$\int_{-1}^{1}\mathrm{d}x\,x^n H^q(x,\xi,t) = \sum_{\substack{i=0\\ i\text{ even}}}^{n} a^q_{n,i}(t)(2\xi)^i + \mathrm{mod}(n,2)c^q_n(t)(2\xi)^{n+1}\,, \tag{2.17}$$

$$\int_{-1}^{1}\mathrm{d}x\,x^n E^q(x,\xi,t) = \sum_{\substack{i=0\\ i\text{ even}}}^{n} b^q_{n,i}(t)(2\xi)^i - \mathrm{mod}(n,2)c^q_n(t)(2\xi)^{n+1}\,. \tag{2.18}$$

This remarkable property is called **polynomiality**. Deeply rooted in Lorentz covariance, it states that the n-th Mellin moments of H^q and E^q are polynomials in ξ of order $n+1$ at most if n is odd, and of order n at most if n is even. We recover in particular that the 0-th Mellin moments of H^q and E^q are independent of ξ, which was noticed in (1.18).

As the coefficients of the term of order $n+1$ are opposite in sign for H^q and E^q, n-th Mellin moments of $H^q + E^q$ are only polynomials in ξ of order n. A similar derivation can be obtained for polarised GPDs $\widetilde{H}^q$ and $\widetilde{E}^q$, without any term of order $n+1$ because the study of their form factor decomposition does not make the last term of Eq. (2.15) appear. Gluon GPDs obey similar relations, except the polynomial in ξ is of order up to $n+2$.

2.2.1 . The double distribution representation

The polynomiality of Mellin moments imposes an interesting constraint on the mutual dependence on x and ξ of GPDs. A mathematical result by Hertle [69] demonstrates that it is equivalent for the n-th Mellin moments integrated over x to be polynomials in ξ of order n and for the GPD to belong to the image of the Radon transform [70, 71]. The theorem can be applied directly for the case of $H^q + E^q$, $\widetilde{H}^q$ or $\widetilde{E}^q$. On the contrary, for H^q and E^q, it is first necessary to isolate the problematic power ξ^{n+1} which may appear for n odd.

This can be done by defining a function [4], known as the D-**term**, such that

$$\int_{-1}^{1} \mathrm{d}\alpha\, \alpha^n D^q(\alpha,t) = \mathrm{mod}(n,2) 2^{n+1} c_n^q(t)\,. \tag{2.19}$$

As all its even moments are 0, $D^q(\alpha,t)$ is an α-odd function, and by design its support is limited to $\alpha \in [-1,1]$. Then the function defined by $H^q(x,\xi,t) - \mathrm{sgn}(\xi)\Theta(1 - |x|/|\xi|)D^q(x/\xi,t)$, where $\mathrm{sgn}(\xi) = |\xi|/\xi$ is the sign of ξ, verifies

$$\begin{aligned}
&\int_{-1}^{1} \mathrm{d}x\, x^n \left[H^q(x,\xi,t) - \mathrm{sgn}(\xi)\Theta\left(1 - \frac{|x|}{|\xi|}\right) D^q\left(\frac{x}{\xi},t\right)\right] \\
&\qquad = \int_{-1}^{1} \mathrm{d}x\, x^n H^q(x,\xi,t) - \mathrm{sgn}(\xi) \int_{-|\xi|}^{|\xi|} \mathrm{d}x\, x^n D^q\left(\frac{x}{\xi},t\right)\,, &(2.20)\\
&\qquad = \int_{-1}^{1} \mathrm{d}x\, x^n H^q(x,\xi,t) - \xi^{n+1} \int_{-1}^{1} \mathrm{d}\alpha\, \alpha^n D^q\left(\alpha,t\right)\,, &(2.21)\\
&\qquad = \int_{-1}^{1} \mathrm{d}x\, x^n H^q(x,\xi,t) - \xi^{n+1} \mathrm{mod}(n,2) 2^{n+1} c_n^q(t)\,, &(2.22)\\
&\qquad = \sum_{\substack{i=0 \\ i \text{ even}}}^{n} a_{n,i}^q(t)(2\xi)^i\,, &(2.23)
\end{aligned}$$

where we used successively Eqs. (2.19) and (2.17) in the two final lines. The subtraction of $\mathrm{sgn}(\xi)\Theta(1-|x|/|\xi|)D^q(x/\xi,t)$ allows therefore to satisfy the conditions of Hertle's theorem even for H^q and E^q. Taking the example of H^q, there exists a function $f^q(\beta,\alpha,t)$, coined **double distribution** (DD), such that $H^q(x,\xi,t) - \mathrm{sgn}(\xi)\Theta(1 - |x|/|\xi|)D^q(x/\xi,t)$ is the Radon transform of $f^q(\beta,\alpha,t)$:

$$H^q(x,\xi,t) - \mathrm{sgn}(\xi)\Theta\left(1 - \frac{|x|}{|\xi|}\right) D^q\left(\frac{x}{\xi},t\right) = \int \mathrm{d}\beta \mathrm{d}\alpha\, \delta(x - \beta - \alpha\xi) f^q(\beta,\alpha,t)\,, \tag{2.24}$$

which can be put concisely under the form

$$H^q(x,\xi,t) = \int \mathrm{d}\beta \mathrm{d}\alpha\, \delta(x - \beta - \alpha\xi) \left[f^q(\beta,\alpha,t) + \xi\delta(\beta) D^q(\alpha,t)\right]\,. \tag{2.25}$$

The DD formalism was introduced independently in Refs. [10, 15]. Initially proposed as an alternative way compared to GPDs to parameterise the hadron matrix element of Eq. (1.2), its link to GPDs via the Radon transform was later formalised in Ref. [73]. The D-term was originally introduced Ref. [74] from the point of view of tensorial structures in the DD.

A similar DD can be constructed for GPD E^q, with the exception that the D-term is opposite. A DD representation can likewise be built for $H^q + E^q$, $\widetilde{H}^q$ and $\widetilde{E}^q$ without any contribution of a D-term. On the contrary, gluon DDs involve a D-term.

4. Whether or not there exists a unique function determined by the value of its integer moments is a delicate question. A general answer is known for under the assumption that the function to reconstruct is positive – see for instance the Hausdorff moment problem [72], but this assumption does not hold in our case. If the D-term is well behaved enough, we might hope that our "definition" is satisfactory.

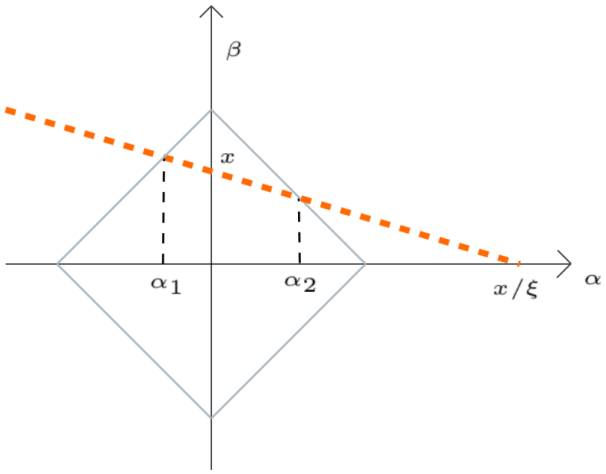

Figure 2.1 – The grey square represents the rhombus Ω on which DDs are defined. The orange dotted line corresponds to $x = \beta + \alpha\xi$. Integration of the DD along this line yields the contribution of the DD to the GPD through the Radon transform of Eq. (2.25). The line intersects the edges of the rhombus at $\alpha_1 = (x-1)/(1+\xi)$ and $\alpha_2 = (1-x)/(1-\xi)$ for the case $x > \xi > 0$ which we have depicted here.

2.2.2 . Properties of double distributions

Domain

As a consequence of the fact that GPDs have a domain of definition $(x, \xi) \in [-1, 1]^2$, DDs are non-vanishing only on a so-called **rhombus** Ω defined by

$$\Omega = \{(\beta, \alpha) \mid |\alpha| + |\beta| \leq 1\}\,. \tag{2.26}$$

This region is represented by the grey square on Fig. 2.1.

The Radon transform of Eq. (2.25) corresponds to a line integration on the DD. One such line is shown on Fig. 2.1 in the case $x > \xi > 0$. The kinematic region $x > |\xi|$, which we have interpreted in Section 1.2 as corresponding to the emission and re-absorption of a quark, only probes the DD for $\beta > 0$. On the contrary, the kinematic region $x < -|\xi|$ probes the DD only for $\beta < 0$ and corresponds to the partonic interpretation of GPDs in terms of emission and reabsorption of an anti-quark.

The region $-|\xi| < x < |\xi|$, in which GPDs probe quark anti-quark pairs, corresponds to Radon integration lines which cross the line $\beta = 0$. It is interesting to note that lines with $|\xi|$ larger than 1 while respecting $-|\xi| < x < |\xi|$ have an intersection with the rhombus. The DD can therefore describe an object which extends the definition domain of GPDs beyond $|\xi| = 1$. These objects are known as generalised distribution amplitudes (GDAs) [38, 39, 40, 41].

Parity

Inherited directly from the ξ parity and real-valuedness of the GPD H^q, the DD f^q is even in α and real-valued as well. Singlet and non-singlet components of the DD can

likewise be defined as

$$f^{q(+)}(\beta,\alpha,t) = f^q(\beta,\alpha,t) - f^q(-\beta,\alpha,t) \tag{2.27}$$

$$f^{q(-)}(\beta,\alpha,t) = f^q(\beta,\alpha,t) + f^q(-\beta,\alpha,t)\,. \tag{2.28}$$

For the DDs associated to other GPDs than H^q, simple modifications according to the parity of the GPD must be applied.

Moments of DDs

We have specifically introduced the D-term as a way to encompass the terms with a dependence of order ξ^{n+1} in the n-th Mellin moment of the GPD. Subtracting the D-term allowed us to introduce the DD f^q thanks to the property of the Radon transform. We know therefore that f^q will only contribute to the terms of order ξ^i up to $i = n$ in the polynomial expansion of Mellin moments of the GPD. Let us verify quickly that it is indeed the case :

$$\int_{-1}^{1} \mathrm{d}x\, x^n H^q(x,\xi,t) = \int_{-1}^{1} \mathrm{d}x\, x^n \int_\Omega \mathrm{d}\beta\mathrm{d}\alpha\, \delta(x-\beta-\alpha\xi)\left[f^q(\beta,\alpha,t) + \xi\delta(\beta)D^q(\alpha,t)\right]\,, \tag{2.29}$$

$$= \int_\Omega \mathrm{d}\beta\mathrm{d}\alpha\, (\beta+\alpha\xi)^n \left[f^q(\beta,\alpha,t) + \xi\delta(\beta)D^q(\alpha,t)\right]\,, \tag{2.30}$$

$$= \sum_{i=0}^{n} \xi^i \binom{n}{i} \int_\Omega \mathrm{d}\beta\mathrm{d}\alpha\, \alpha^i\beta^{n-i} f^q(\beta,\alpha,t) + \xi^{n+1}\int_{-1}^{1} \mathrm{d}\alpha\, \alpha^n D^q(\alpha,t)\,. \tag{2.31}$$

The fact that only even powers of ξ are actually non vanishing in Eq. (2.31) is a consequence of the correct parity of DDs.

The coefficients $a^q_{n,i}(t)$ mentioned in Eqs. (2.15) and (2.17) can then be expressed in terms of twice integrated moments of the DD as

$$2^i a^q_{n,i}(t) = \binom{n}{i} \int_\Omega \mathrm{d}\beta\mathrm{d}\alpha\, \alpha^i\beta^{n-i} f^q(\beta,\alpha,t)\,. \tag{2.32}$$

Coming back to the relation between the GPD H^q and the GFFs A^q and C^q stated in Eq. (1.22), we can deduce from Eq. (2.31) that C^q is linked to the D-term thanks to

$$C^q(t) = \frac{1}{4}\int_{-1}^{1} \mathrm{d}\alpha\, \alpha D^q(\alpha,t)\,, \tag{2.33}$$

A similar result can be obtained for gluons under the form

$$C^g(t) = \frac{1}{4}\int_{-1}^{1} \mathrm{d}\alpha\, D^g(\alpha,t)\,. \tag{2.34}$$

We will use extensively this relation in Chapter 5.

As a conclusion, the DD representation allows by design to satisfy the important property of polynomiality of Mellin moments of GPDs : although it implies a complex interdependence of x and ξ at the level of the GPD $H(x,\xi,t)$, it does not yield any constraint on the interdependence of β and α at the level of the DD. We stress that **admitting a representation in terms of DDs is equivalent to satisfying the polynomiality property derived from Lorentz covariance**.

2.2.3 . Schemes of double distributions

We have dealt with the issue of contributions of order ξ^{n+1} in the n-th Mellin moment of GPDs by encompassing them all in the D-term. However, as was pointed out in Ref. [73], there are other ways to handle this situation, giving rise to different schemes (or representations) of double distributions. The scheme we have introduced so far, with a DD f^q and a D-term, is often named the Polyakov-Weiss representation owing to the contribution of the two authors to the introduction of the D-term [74].

It is also possible to deal with terms of order ξ^{n+1} by considering the Mellin moments of $(H^q(x,\xi,t) - H^q(x,0,t))/\xi$ instead of $H^q(x,\xi,t) - \mathrm{sgn}(\xi)\Theta(1-|x|/|\xi|)D^q(x/\xi,t)$ as we have done so far. Then using the coefficients $a^q_{n,i}(t)$ and $c^q_n(t)$ (2.17),

$$\int_{-1}^{1} \mathrm{d}x\, x^n \left(\frac{H^q(x,\xi,t) - H^q(x,0,t)}{\xi} \right) = \frac{1}{\xi}\int_{-1}^{1} \mathrm{d}x\, x^n H^q(x,\xi,t) - \frac{1}{\xi}\int_{-1}^{1} \mathrm{d}x\, x^n H^q(x,0,t)\,, \tag{2.35}$$

$$= \frac{1}{\xi} a^q_{n,0} + \sum_{\substack{i=2 \\ i \text{ even}}}^{n} a^q_{n,i}(t) 2^i \xi^{i-1} + \mathrm{mod}(n,2) c^q_n(t) 2^{n+1}\xi^n - \frac{1}{\xi} a^q_{n,0}(t)\,. \tag{2.36}$$

Therefore, the Mellin moments of $(H^q(x,\xi,t) - H^q(x,0,t))/\xi$ are polynomials in ξ of order n, so it is possible to introduce a different DD g^q_T [75] such that

$$H^q(x,\xi,t) = H^q(x,0,t) + \xi \int_\Omega \mathrm{d}\beta \mathrm{d}\alpha\, \delta(x-\beta-\alpha\xi) g^q_T(\beta,\alpha,t)\,. \tag{2.37}$$

This scheme has the very nice feature of making explicit the t-dependent forward limit of the GPD, which we know to be equal to usual PDFs when $t = 0$. Eq. (2.37) can be rewritten as

$$H^q(x,\xi,t) = \int_\Omega \mathrm{d}\beta \mathrm{d}\alpha\, \delta(x-\beta-\alpha\xi) \left[\delta(\alpha) H^q(\beta,0,t) + \xi g^q_T(\beta,\alpha,t) \right]\,. \tag{2.38}$$

Another popular way to define a DD representation, often called the Pobylitsa representation, is to take [76]

$$H^q(x,\xi,t) = (1-x) \int_\Omega \mathrm{d}\beta \mathrm{d}\alpha\, \delta(x-\beta-\alpha\xi) f^q_P(\beta,\alpha,t)\,. \tag{2.39}$$

One could wonder what is the interest of having many different ways to represent the same object. Each scheme highlights a different feature of the GPD which makes it more convenient to work in one scheme or the other depending on the objective of the study.

For instance, most of our work in Chapters 5 and 6 will be performed in the Polyakov-Weiss representation because we will be strongly interested in the D-term : in Chapter 5 due to its relation to the GFF $C^a(t)$, and in Chapter 6 due to dispersion relations that relate the real and imaginary parts of Compton form factors (see Chapter 3). However, some results in Chapter 6 will be obtained in the Pobylitsa gauge to compare them with a publication. In Chapter 7, we will actually develop our own DD gauge only suited for the singlet quark GPD due to specific modelling challenges we will be facing.

An interested reader will find more details in Ref. [77] on how to navigate from one scheme to the other, as well as examples of how a very simple model in one scheme can become singular and complicated in another.

2.3 . Wave function representation and positivity bounds

2.3.1 . GPDs representation as an overlap of light-front wave functions

Light-front quantisation allows one to decompose a hadron state on a Fock basis [78]. For a nucleon for instance,

$$|p\rangle = \sum_{\beta} \Phi_{\beta}^{qqq} |qqq\rangle + \sum_{\beta} \Phi_{\beta}^{qqq,q\bar{q}} |qqq, q\bar{q}\rangle + ... \tag{2.40}$$

where the expansion contains Fock states with an ever larger number of partons summed over the relevant quantum numbers collectively denoted by β. Φ denotes light-front wave functions (LFWFs) [79], non-perturbative objects which depend on the momenta and quantum numbers of partons contained in the associated Fock state. Precisely, the LFWF of a Fock state with N particles, which we will denote Φ_{β}^{N} in the following, depends on the longitudinal momentum fractions $(x_1, ..., x_N)$ of the constituents which verify $\sum_{i=1}^{N} x_i = 1$, and on their transverse momentum $(k_{\perp,1}, ...k_{\perp,N})$ satisfying $\sum_{i=1}^{N} k_{\perp,i} = p_{\perp}$.

The link between LFWFs and GPDs is derived in detail in Ref. [80]. The presentation of this result requires the introduction of many rather cumbersome notations. For the sake of legibility, we will therefore adopt a simplified presentation. In the regions $x > |\xi|$ (respectively $x < -|\xi|$) where GPDs can be interpreted as the emission and re-absorption of a particle (respectively anti-particle), they can be written as an overlap between LFWFs with the same number of particles in the initial and final states. For instance for $x > |\xi|$, a generic representation of that overlap is given by

$$H^q(x,\xi,t) - \frac{\xi^2}{1-\xi^2} E^q(x,\xi,t) = \sum_{N} \left(16\pi^3\sqrt{1-\xi^2}\right)^{1-N} \sum_{\beta} \sum_{j=1}^{N} \delta_{f_j q} \int \left[\prod_{i=1}^{N} \mathrm{d}x_i \, \mathrm{d}^2 k_{\perp,i}\right] \times \delta\left(1 - \sum_{i=1}^{N} x_i\right) \delta^{(2)}\left(p_{\perp} - \sum_{i=1}^{N} k_{\perp,i}\right) \delta(x - x_j) \left(\Phi_{\beta}^{N}(r_{out})\right)^* \Phi_{\beta}^{N}(r_{in}) \,. \tag{2.41}$$

The right hand side (r.h.s.) is summed over the number of particles in the Fock state N and the active quark j whose flavour is denoted by f_j. An integration over the longitudinal and transverse momentum of all constituents of the Fock state – denoted by i – is performed. $\delta_{f_j q}$ denotes the Kronecker symbol. $\Phi_{\beta}^{N}(r)$ is a LFWF associated to a Fock state with N particles and quantum numbers β evaluated at the momentum configuration r encompassing the momentum of its N constituents. We refer to Ref. [80] for the precise definition of the frames and momenta involved in this expression. Owing to the fact that only LFWFs corresponding to a similar number of particles overlap, it is possible to truncate the Fock expansion in an unambiguous way.

One the contrary, in the region $-|\xi| < x < |\xi|$, the partonic interpretation of GPDs as the annihilation of a parton anti-parton pair means that the number of particles is not conserved between the initial and final states of the interaction. The GPD then writes as overlap of LFWFs with $N+1$ partons in the initial state and $N-1$ in the final. The structure in terms of quantum numbers is more complicated since it must take into account the quantum numbers of both the active particle and antiparticle. The GPD

overlap writes generically as

$$H^q(x,\xi,t) - \frac{\xi^2}{1-\xi^2}E^q(x,\xi,t) = \sum_N 16\pi^3(1-\xi)\left(16\pi^3\sqrt{1-\xi^2}\right)^{-N}\sum_{\beta,\beta'}\sum_{j,j'=1}^{N+1}\frac{\Xi}{\sqrt{n_j n_{j'}}}$$
$$\times\int\left[\prod_{i=1,i\neq j'}^{N+1}\mathrm{d}x_i\,\mathrm{d}^2k_{\perp,i}\right]\delta\left(1-\xi-\sum_{i=1,i\neq j,j'}^{N+1}x_i\right)\delta^{(2)}\left(\frac{\Delta_\perp}{2}-\sum_{i=1,i\neq j,j'}^{N+1}k_{\perp,i}\right)$$
$$\times\,\delta(x-x_j)\left(\Phi^{N-1}_{\beta'}(r_{out})\right)^*\Phi^{N+1}_{\beta}(r_{in})\,. \tag{2.42}$$

We hid in the simple character Ξ a large multiplication of Kronecker symbols intended to guarantee the correct flavours, colours and helicities both for the active and spectator partons. The active quark is denoted by j and anti-quark by j'. n_j (respectively $n_{j'}$) is the number of quarks (respectively anti-quarks) in the initial proton wave function. As the overlap writes in this kinematic region between LFWFs with $N+1$ and $N-1$ partons, it is not clear how to truncate the Fock expansion in a consistent way with the truncation performed in the previous region $|x|\geq|\xi|$. This gives rise to a difficulty when employing models based on truncated overlap of LFWFs notably discussed in Ref. [81, 82]. The method only provides suitable results for $|x|>|\xi|$. A way to go beyond this difficulty is presented in Section 3.4.3, and further touched again in Section 6.2.2.

Similar relations to those briefly sketched here are demonstrated in Ref. [80] for polarised and gluon GPDs.

2.3.2 . Positivity inequalities

In the forward limit $\Delta=0$ where GPDs yield PDFs, there is no asymmetry between the incoming and outgoing proton states. Therefore, $r_{out}=r_{in}=r$. Since $\xi=0$, we are always in the kinematic region $|x|\geq|\xi|$, and the overlap of Eq. (2.41) involves

$$\left(\Phi^N_\beta(r_{out})\right)^*\Phi^N_\beta(r_{in}) = |\Phi^N_\beta(r)|^2\,. \tag{2.43}$$

As LFWFs are quantum weights of states in a Fock basis, their squared modulus can be classically interpreted as the probability that the hadron is found in this Fock state. Then summing and integrating $\delta(x-x_j)|\Phi^N_\beta(r)|^2$ over all other variables in Eq. (2.41) amounts to extracting the marginal distribution of the probability of interacting with an active quark carrying a forward momentum x, which is exactly the traditional probabilistic interpretation of PDFs.

As observed for instance in Ref. [83], the representation of GPDs in terms of overlap of LFWFs has formally the structure of a scalar product in the Hilbert space of LFWFs. We have noticed in Eq. (2.43) that the norm of LFWFs is linked to usual PDFs. Positivity conditions can then be derived from the Cauchy-Schwartz inequality which relates a scalar product to the product of norms of the vectors [5]. Refs. [84, 85] complemented by [80] give

$$\left|H^q(x,\xi,t)-\frac{\xi^2}{1-\xi^2}E^q(x,\xi,t)\right| \leq \sqrt{q\left(x_1\right)q\left(x_2\right)\frac{1}{1-\xi^2}}\,, \tag{2.44}$$

where

$$x_1=\frac{x+\xi}{1+\xi}\quad\text{and}\quad x_2=\frac{x-\xi}{1-\xi}\,. \tag{2.45}$$

5. $|\langle u|v\rangle|\leq||u||\cdot||v||$

The equivalent bound for gluon GPDs is

$$\left|H^g(x,\xi,t) - \frac{\xi^2}{1-\xi^2}E^g(x,\xi,t)\right| \leq \sqrt{g\left(x_1\right) g\left(x_2\right) \frac{x^2-\xi^2}{1-\xi^2}}\,. \tag{2.46}$$

A more constraining relation is derived in Ref. [86] as

$$\left|H^q(x,\xi,t) - \frac{\xi^2}{1-\xi^2}E^q(x,\xi,t)\right|^2 + \frac{\sqrt{t_0-t}}{2M\sqrt{1-\xi^2}}\left|E^q(x,\xi,t)\right|^2 \leq \frac{q\left(x_1\right) q\left(x_2\right)}{1-\xi^2}\,, \tag{2.47}$$

where

$$t_0 = -\frac{4\xi^2M^2}{1-\xi^2}\,. \tag{2.48}$$

t_0 is the maximal value of t (that is $t \leq t_0 \leq 0$) which is kinematically allowed at a fixed value of ξ following Eq. (2.1). Not only is Eq. (2.47) a stronger bound compared to Eq. (2.44), but it yields the additional result on the E^q distribution alone

$$\frac{\sqrt{t_0-t}}{2M}|E^q(x,\xi,t)| \leq \sqrt{q\left(x_1\right) q\left(x_2\right)}\,. \tag{2.49}$$

Ref. [87] demonstrates how to derive stronger inequalities involving polarised and transversity PDFs, and derives a bound on the GPD H^q alone as

$$|H^q(x,\xi,t)| \leq \sqrt{\left(1-\frac{t_0\xi^2}{t_0-t}\right)\frac{q(x_1)q(x_2)}{1-\xi^2}}\,. \tag{2.50}$$

Refs. [88, 87] show that positivity bounds are much more general than the few examples we have mentioned so far, and form a vast class of inequalities that can be obtained by the mean of constructing arbitrary states and requiring that their norm in the Hilbert is positive. Working with LFWFs is **a way to systematically fulfill all these inequalities at once**. We will demonstrate how the GPD models we develop in Chapter 7 based on the DD formalism can accommodate some of these positivity constraints, which will represent an important tool to counter the uncertainty arising from the shadow GPDs introduced in Chapter 6.

2.4 . Evolution of generalised parton distributions

As we mentioned briefly at the beginning of Chapter 1, GPDs depend additionally on a renormalisation scale, which arises from the necessity to cure UV divergences of the bilocal operator defining GPDs. In practice, it is possible to make the scale dependence emerge from the UV regulator ϵ through the relation

$$\frac{H^a(x,\xi,t,\mu^2)}{x^{p_a}} = \sum_b \int_{-1}^{1}\frac{\mathrm{d}z}{|z|}\,Z^{ab}\left(\frac{z}{\xi},\frac{\xi}{x},\alpha_s(\mu^2),\epsilon\right)\frac{\hat{H}^b(z,\xi,t,\epsilon)}{z^{p_b}}\,, \tag{2.51}$$

where the sum is performed over parton types b to account for the possibility of mixing of GPDs under evolution, and $p_a = 1$ if $a = g$ and 0 otherwise[6]. In Ref. [1], we give

6. Thanks to the scaling by x^{-p_a}, the forward limit ($\xi = 0, t = 0$) of $x^{-p_a}H^a(x,\xi,t)$ is exactly $a(x)$ even in the gluon case. This scaling allows one to write Z^{ab}, and later the splitting kernels, as functions of the two ratios z/ξ and ξ/x instead of the three variables x, ξ, z.

more details on the determination of Z^{ab} which is obtained by studying the poles of the operator in dimensional regularisation in the $\overline{MS}$ renormalisation scheme. Using the fact that $\hat{H}$ has no dependence on the scale, the differentiation of Eq. (2.51) with respect to $\log(\mu^2)$ yields the renormalisation group equation (RGE), which writes generically at LO in α_s as

$$\frac{1}{x^{p_a}}\frac{\partial H^a(x,\xi,t,\mu^2)}{\partial \log(\mu^2)} = \alpha_s(\mu^2)\sum_b \int_x^1 \frac{\mathrm{d}z}{\xi} K^{ab,(0)}\left(\frac{z}{\xi},\frac{\xi}{x}\right)\frac{H^b(z,\xi,t,\mu^2)}{z^{p_b}}\,. \tag{2.52}$$

The $K^{qb,(0)}$ distributions are known as **splitting kernels**, where $^{(0)}$ denotes the LO expansion. To fix the notations, we will denote the RGE of α_s as [89]

$$\frac{\partial \alpha_s(\mu^2)}{\partial \log(\mu^2)} = \beta(\alpha_s(\mu^2)) = -b_0\alpha_s^2(\mu^2) - b_1\alpha_s^3(\mu^2) - \dots\,, \tag{2.53}$$

where notably $b_0 = \beta_0/(4\pi)$, $\beta_0 \equiv 11 - 2n_f/3$ and n_f is the number of active flavours.

In the limit where $\xi = 0$, Eq. (2.52) yields the RGE of usual PDFs [7], known as the Dokshitzer – Gribov – Lipatov – Altarelli – Parisi (DGLAP) equation [90, 91, 92]. On the contrary, in the limit $\xi \to 1$, GPDs evolve according to the Efremov – Radyushkin – Brodsky – Lepage (ERBL) RGE [93, 94], which is also responsible for the evolution of distribution amplitudes. In particular, the D-term evolves exactly following the ERBL equation.

Observing that both the forward limit and the D-term follow their own evolution equations, we can conclude that if either the forward limit or the D-term are 0 at one scale, then they remain null at any scale. This statement is true at any perturbative order of evolution. We will make a significant use of this result in Chapters 5 and 6.

More generally, evolution is completely compatible with the polynomiality property at any order, as it can be expressed at the level of DDs themselves [85]. On the contrary, the preservation of the positivity property (2.44) is a more delicate question. Ref. [88] argues that positivity constraints on GPDs are stable under LO evolution to higher scales : if they are verified at one scale, then they are as well under upward evolution. However, in general, renormalisation includes subtractions which can violate naive positivity bounds. It is known for instance that some renormalisation schemes produce PDFs with negative values for low renormalisation scales [95, 96], which is generally not considered as a problem as long as physical cross-sections computed from these objects are positive. As positivity bounds on GPDs arise from the same reasoning as positivity of PDFs, Ref. [88] underlines that the possibility of violation of positivity for PDFs implies a similar possibility of violation of positivity bounds for GPDs. Ref. [97] clarifies the physical domain where the positivity bounds are expected to hold. Let us notice that the question of positivity of PDFs is still at the center of a lot of attention as testified by recent contradictory takes on the subject – see notably Ref. [96].

Conformal moments

The LO splitting kernels were determined already in seminal papers of the field, like Refs. [12, 10, 15, 98]. Next-to-leading order corrections were computed in Refs. [99, 100,

7. Splitting kernels are independent of t at leading twist, so the GPD at $\xi = 0$ and $t \neq 0$ still evolves exactly in the same manner as the PDF.

101] and recently confirmed [102]. Three-loop computations are also available in the non-singlet sector [103].

Many of these derivations were performed in the space of so-called **conformal moments** owing to the fact that LO evolution in QCD preserves conformal symmetry [104]. Conformal moment of GPDs can be defined as

$$\mathcal{O}_n^q(\xi,t,\mu^2) = \frac{\Gamma(n+1)\Gamma(3/2)}{2^n\Gamma(n+3/2)}\,\xi^n \int_{-1}^{1} \mathrm{d}x\, C_n^{(3/2)}\left(\frac{x}{\xi}\right) H^q(x,\xi,t,\mu^2)\,, \tag{2.54}$$

$$\mathcal{O}_n^g(\xi,t,\mu^2) = \frac{\Gamma(n)\Gamma(5/2)}{2^{n-1}\Gamma(n+3/2)}\,\xi^{n-1} \int_{-1}^{1} \mathrm{d}x\, C_{n-1}^{(5/2)}\left(\frac{x}{\xi}\right) H^g(x,\xi,t,\mu^2)\,, \tag{2.55}$$

where $\Gamma(n) = (n-1)!$ denotes the Gamma function which extends the concept of factorial and $C_n^{(\alpha)}$ are Gegenbauer polynomials [8] with $\alpha = 3/2$ for quarks and $\alpha = 5/2$ for gluons. The peculiar normalisation in the definition of Eq. (2.54) is chosen so that in the limit where $\xi = 0$, conformal moments coincide with Mellin moments of the PDF :

$$\mathcal{O}_n^a(\xi=0,t,\mu^2) = \int_{-1}^{1} \mathrm{d}x\, x^n a(x,\mu^2)\,. \tag{2.56}$$

Conformal moments are in principle a particularly suitable way to study the evolution of GPDs because moments associated to different values of n do not mix under LO evolution [105]. For n even, the quark conformal moments factorise their ξ and μ dependence following

$$\mathcal{O}_n^q(\xi,t,\mu^2) = \mathcal{O}_n^q(\xi,t,\mu_0^2)\left(\frac{\alpha_s(\mu^2)}{\alpha_s(\mu_0^2)}\right)^{2\gamma_n/\beta_0}\,, \tag{2.57}$$

whereas the gluon conformal moments are strictly 0 due to the parity in x of H^g. The exponents γ_n, often known as **anomalous dimensions**, are remarkably independent of ξ [9]. It means that they are in particular the same governing the solution of the DGLAP equation ($\xi = 0$) and the ERBL equation ($\xi = 1$). We postpone a precise definition of these exponents to Eq. (5.9), where the reason why we will need to study their numerical values will appear more clearly.

On the contrary, for n odd, some mixing occurs between quark and gluon conformal moments, but only strictly for the same values of n. More precisely, flavour non singlet components obtained as the difference of two GPDs associated to different quark flavours evolve in the same way as Eq. (2.57) :

$$\mathcal{O}_n^{q_1}(\xi,t,\mu^2) - \mathcal{O}_n^{q_2}(\xi,t,\mu^2) = \left[\mathcal{O}_n^{q_1}(\xi,t,\mu_0^2) - \mathcal{O}_n^{q_2}(\xi,t,\mu_0^2)\right]\left(\frac{\alpha_s(\mu^2)}{\alpha_s(\mu_0^2)}\right)^{2\gamma_n/\beta_0}\,, \tag{2.58}$$

whereas the flavour singlet component, summed over all quark flavours, mixes with gluons

8. Gegenbauer polynomials $C_n^{(\alpha)}(x)$ are polynomial in x of order n which are orthogonal on $[-1,1]$ with respect to the weight function $(1-x^2)^{\alpha-1/2}$. They generalise two other families of orthogonal polynomials : the Legendre and Chebyshev polynomials.

9. This observation is linked to the fact that in the $\overline{MS}$ scheme, anomalous dimensions of local operators are fixed independently of incoming or outgoing states.

in the following manner [35] :

$$\frac{1}{n_f}\sum_q \mathcal{O}_n^q(\xi,t,\mu^2) = \mathcal{O}_n^+(\xi,t,\mu_0^2)\left(\frac{\alpha_s(\mu^2)}{\alpha_s(\mu_0^2)}\right)^{2\gamma_n^+/\beta_0} + \mathcal{O}_n^-(\xi,t,\mu_0^2)\left(\frac{\alpha_s(\mu^2)}{\alpha_s(\mu_0^2)}\right)^{2\gamma_n^-/\beta_0}, \tag{2.59}$$

$$\mathcal{O}_n^g(\xi,t,\mu^2) = a_n^+\mathcal{O}_n^+(\xi,t,\mu_0^2)\left(\frac{\alpha_s(\mu^2)}{\alpha_s(\mu_0^2)}\right)^{2\gamma_n^+/\beta_0} + a_n^-\mathcal{O}_n^-(\xi,t,\mu_0^2)\left(\frac{\alpha_s(\mu^2)}{\alpha_s(\mu_0^2)}\right)^{2\gamma_n^-/\beta_0}, \tag{2.60}$$

where $\mathcal{O}_n^\pm(\xi,t,\mu_0^2)$ are defined straightforwardly by evaluating Eqs. (2.59) and (2.60) at $\mu^2=\mu_0^2$:

$$\frac{1}{n_f}\sum_q \mathcal{O}_n^q(\xi,t,\mu_0^2) = \mathcal{O}_n^+(\xi,t,\mu_0^2) + \mathcal{O}_n^-(\xi,t,\mu_0^2), \tag{2.61}$$

$$\mathcal{O}_n^g(\xi,t,\mu_0^2) = a_n^+\mathcal{O}_n^+(\xi,t,\mu_0^2) + a_n^-\mathcal{O}_n^-(\xi,t,\mu_0^2). \tag{2.62}$$

The values of coefficients $a_n^\pm$ and anomalous dimensions $\gamma_n^\pm$ were first derived in Refs. [106, 107]. We will provide an alternative matrix formulation of LO evolution in the space of conformal moments for n odd in Chapter 5, where the precise values of $a_n^\pm$ and $\gamma_n^\pm$ will be detailed, and an approximate solution of the evolution derived.

At higher perturbative order, the property of non-mixing of conformal moments with different values of n is no longer true in the $\overline{MS}$ scheme. A GPD evolution code based on conformal moments is proposed in Ref. [108] at NLO in the $\overline{MS}$ scheme, and NNLO in a specific choice of renormalisation scheme called the conformal scheme. Let us notice however that the exact reconstruction of GPDs from conformal moments is a technical procedure, which requires the introduction of formal resummations or the analytical continuation of the moments to complex values of n. Two analytical techniques, the Mellin-Barnes integral [108, 109] and the Shuvaev transform [110, 111] have been proposed to reconstruct exactly the GPD from its conformal moments. Both were shown to be equivalent in Ref. [112]. If no analytical form of the moments is known, the GPD reconstruction can be attempted numerically by an expansion on a polynomial basis. One of the early studies of this kind performed using an expansion of GPDs on Gegenbauer polynomials [10] reported in Ref. [114] to have used 80 polynomials. Due to the oscillatory behaviour of the polynomials, it was necessary to perform exact calculations and interpolate the results by a smooth function. The review led in Ref. [35] likewise mentions the necessity of tens if not hundreds of polynomials, and severe numerical issues for x close to ξ or ξ too small where GPDs may present fast variations in x.

For these reasons, and considering the peculiar shadow distributions whose evolution we will be interested in, we preferred to use an evolution code we published recently in Ref. [1]. It solves the LO RGE directly in the (x,ξ) space thanks to a numerical method (Runge-Kutta method), bypassing any need for a reconstruction from the space of conformal moments. We have led extensive tests to verify the excellent numerical preservation of polynomiality of Mellin moments, as well as the correct reduction to the DGLAP and

10. Formally, it is not needed to use Gegenbauer polynomials to expand the GPD, and other bases of orthogonal polynomials have been used. Ref. [113] uses 70 terms computed with Legendre polynomials.

ERBL kernels in the $\xi \to 0$ and $\xi \to 1$ limits. Although it is not the first evolution code in the (x, ξ) space, it offers significantly improved numerical properties compared to the no-longer maintained Vinnikov code [115]. In addition to providing the results of the evolution of shadow GPDs, our (x, ξ) space code will allow us to revisit in a modelling proposal of the behaviour of GPDs at small x and ξ inspired from the Shuvaev transform in Chapter 7.

3 - Experimental sensitivity and modelling

As we have already mentioned in Chapter 1, GPDs were introduced in the context of the description of DVCS and DVMP thanks to factorisation theorems. We have skimmed intuitively the content of these theorems, stating that they allow us to represent a scattering process in terms of a hard part characterised by a coefficient function and a soft part involving a parton distribution. In this chapter, we will present in a more systematic way the concrete realisation of the factorisation theorems, the experimental access they provide to GPDs, as well as some modelling tools which have been developed for phenomenological studies.

3.1 . Deeply virtual Compton scattering

As depicted schematically on the left panel of Fig. 1.1 in the case of an electron, DVCS corresponds to the initial and final states $lh \to l\gamma h$, where l is a lepton, γ a real photon and h the hadron target. DVCS is therefore measured thanks to **photon lepto-production** events. There exists however another process with the same initial and final states, known as Bethe-Heitler (BH). A diagrammatic representation of a contribution to the BH is given in Fig. 3.1. The incoming electron is elastically scattered by the hadron target and radiates a photon. The description of this process involves only EFFs.

DVCS and BH interfere coherently, so the cross-section of lepto-production of a photon ($lh \to l\gamma h$) writes as the sum of the squared amplitudes of the two processes $|\mathcal{T}_{DVCS}|^2$ and $|\mathcal{T}_{BH}|^2$ with the addition of an interference term $\mathcal{I}$. The differential cross-section can therefore be written (see Ref. [116] and references therein) as :

$$\frac{\mathrm{d}^5\sigma}{\mathrm{d}x_B\,\mathrm{d}|t|\,\mathrm{d}Q^2\,\mathrm{d}\phi\,\mathrm{d}\phi_S} \propto |\mathcal{T}_{DVCS}|^2 + |\mathcal{T}_{BH}|^2 + \mathcal{I}\,, \tag{3.1}$$

$$\mathcal{I} = \mathcal{T}^*_{DVCS}\mathcal{T}_{BH} + \mathcal{T}_{DVCS}\mathcal{T}^*_{BH}\,, \tag{3.2}$$

where * denotes complex conjugation. The differential cross-section depends on two variables we have already introduced : t, the square of the total four-momentum transfer to the hadron (1.4) and $Q^2 \equiv -q^2$ defined as the opposite of the square of the four-momentum of the photon emitted by the incoming lepton. In addition, the cross-section

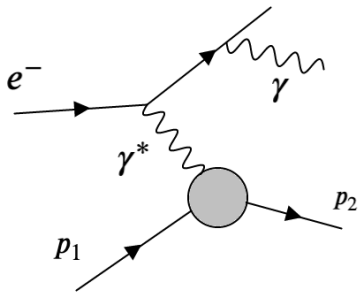

Figure 3.1 – A contribution to the Bethe-Heitler process.

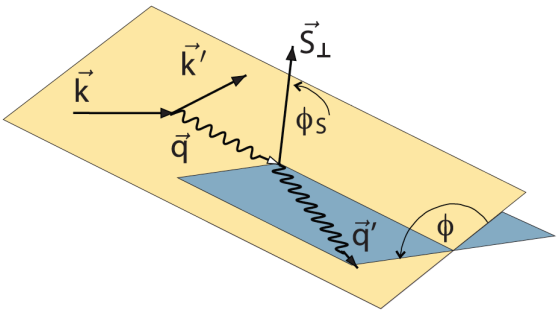

Figure 3.2 – Kinematics of a photon lepto-production event in the hadron rest frame (Trento notations [117]). The incoming and outgoing leptons are denoted by $\vec{k}$ and $\vec{k'}$, the exchanged photon by $\vec{q}$ and the real photon in the final state by $\vec{q'}$. The incoming and outgoing lepton trajectories define the leptonic plane, and the recoil hadron and real photon define the hadronic plane. In this reference system, ϕ is the angle between the leptonic and hadronic planes. ϕ_S is the angle between the leptonic plane and the transverse polarisation of the hadron in case of a polarised target. Figure taken from Ref. [118].

depends on previously not introduced variables : Bjorken's variable x_B and angles ϕ and ϕ_S. Bjorken's variable is defined by

$$x_B \equiv \frac{Q^2}{2p_1 \cdot q}\,, \tag{3.3}$$

which can be related to the skewness variable $\xi = -\Delta^+/(2P^+)$ (1.4) of GPDs by noting that

$$\xi = \frac{p_1^+ - p_2^+}{p_1^+ + p_2^+} \quad \text{so} \quad p_2^+ = p_1^+ \frac{1-\xi}{1+\xi}\,, \tag{3.4}$$

hence the approximate relation [1]

$$x_B \approx \frac{2\xi}{1+\xi}\left(1 - \frac{t}{Q^2}\right)\,. \tag{3.5}$$

The angular dependence of the DVCS cross-section in the hadron rest frame is defined in Fig. 3.2 : ϕ is the angle between the leptonic and hadronic planes, and ϕ_S is the angle between the leptonic plane and the transverse polarisation of the hadron in case of a polarised target. Dependence on ϕ is often expressed by a decomposition on Fourier harmonics. For instance, up to twist-three contributions and corrections suppressed by $1/Q$ [119],

$$|\mathcal{T}_{DVCS}|^2 = \Gamma_{DVCS}(x_B, Q^2, t)\left[c_0^{DVCS} + \sum_{n=1}^{2}\left(c_n^{DVCS}\cos(n\phi) + s_n^{DVCS}\sin(n\phi)\right)\right]\,, \tag{3.6}$$

1. The requirement that the photon in the final state is on-shell yields $(q-\Delta)^2 = 0$, that is $-Q^2 + t + 2p_1 \cdot q - 2p_2 \cdot q = 0$, or equivalently $-1 + \frac{t}{Q^2} + \frac{1}{x_B} - \frac{2p_2 \cdot q}{Q^2} = 0$. Using that $p_2^+ = p_1^+ \frac{1-\xi}{1+\xi}$ (3.4) and assuming that most of the momentum transfer is in the forward direction, that is $p^+ \gg p^-, p_\perp$, we obtain the approximation $p_2 \approx p_1 \frac{1-\xi}{1+\xi}$. Hence $-1 + \frac{t}{Q^2} + \frac{1}{x_B} - \frac{1}{x_B}\frac{1-\xi}{1+\xi} \approx 0$ which yields immediately Eq. (3.5).

where $\Gamma_{DVCS}(x_B, Q^2, t)$ is a known kinematic prefactor [2]. Pure BH term $|\mathcal{T}_{BH}|^2$ and BH-DVCS interference $\mathcal{I}$ are likewise frequently expressed in terms of Fourier harmonics, although the situation is made more difficult by the presence of an additional ϕ dependence originating from lepton propagators in the BH amplitude [119] :

$$|\mathcal{T}_{BH}|^2 = \frac{\Gamma_{BH}(x_B, Q^2, t)}{P_1(\phi)P_2(\phi)}\left[c_0^{BH} + \sum_{n=1}^{2} c_n^{BH}\cos(n\phi) + s_1^{BH}\sin(\phi)\right], \tag{3.7}$$

$$\mathcal{I} = \frac{\Gamma_I(x_B, Q^2, t)}{P_1(\phi)P_2(\phi)}\left[c_0^{I} + \sum_{n=1}^{3}\left(c_n^{I}\cos(n\phi) + s_n^{I}\sin(n\phi)\right)\right]. \tag{3.8}$$

The terms $1/(P_1(\phi)P_2(\phi))$ coming from the lepton propagators make the expansion in Fourier harmonics of $|\mathcal{T}_{BH}|^2$ and $\mathcal{I}$ infinite, and the practical recovery of the c_n and s_n coefficients is therefore more complicated.

Recent expressions for coefficients c_n and s_n are found in Refs. [120]. Pure BH harmonics let themselves write simply in terms of the EFFs F_1 and F_2. As elastic scattering provides a precise enough source of information on EFFs for the needs of DVCS measurements, we may consider that the pure BH contribution is well-known theoretically. On the contrary, pure DVCS term and BH-DVCS interference harmonics bring a new source of information on the hadron inner structure. They are usually expressed in terms of **Compton form factors** (CFFs).

3.1.1 . Compton form factors

CFFs enter linearly the description of the BH-DVCS interference term, and quadratically that of the pure DVCS term. At leading twist, they are expressed as convolutions of GPDs with coefficient functions thanks to a factorisation theorem [16]. The CFF $\mathcal{H}$ associated to the GPD H can be expressed as

$$\mathcal{H}(\xi, t, Q^2) = \sum_a \int_{-1}^{1} \frac{\mathrm{d}x}{\xi}\, T^a\left(\frac{x}{\xi}, \frac{Q^2}{\mu^2},\ \alpha_s(\mu^2)\right)\frac{H^a(x, \xi, t, \mu^2)}{x^{p_a}}, \tag{3.9}$$

$$\equiv \sum_a T^a\left(\frac{x}{\xi}, \frac{Q^2}{\mu^2},\ \alpha_s(\mu^2)\right) \otimes H^a(x, \xi, t, \mu^2) \tag{3.10}$$

where $p_a = 1$ if $a = g$ and 0 otherwise, and we have introduced the notation $\otimes$ to subtend the integration on the first variable of GPDs with correct x^{p_a} factor. T^a is the coefficient function, a complex distribution computable in perturbation theory. We denote by T_k^a its expansion up to order α_s^k. At LO, the coefficient function has no scale dependence and reads

$$\frac{1}{\xi}T_0^q\left(\frac{x}{\xi}\right) = e_q^2\left(\frac{1}{\xi - x - i\epsilon} - \frac{1}{\xi + x - i\epsilon}\right). \tag{3.11}$$

The formulation with $i\epsilon$ is often used, but for many practical purposes, it is preferable to compute directly the limit $\epsilon \to 0^+$. Owing to the Sokhotski-Plemelj formula, it is

2. There is still a dependence in $\phi - \phi_S$ in coefficients c_n and s_n, reason why some authors use $\varphi = \phi - \phi_S$ or equivalent definitions as angular dependence instead of ϕ_S (see for instance Ref. [119]). We stick however to the Trento conventions. Ref. [116] explains precisely how to switch to the alternative BMK conventions.

straightforward at LO, and

$$T_0^q(\omega) = e_q^2\left(\text{p.v.}\left[\frac{1}{1-\omega}-\frac{1}{1+\omega}\right]+i\pi[\delta(1-\omega)-\delta(1+\omega)]\right), \tag{3.12}$$

$$T_0^g = 0\,, \tag{3.13}$$

where p.v. denotes Cauchy's principal value. Finally, the convolution with the GPD yields for the imaginary part

$$\text{Im}\, T_0^q\left(\frac{x}{\xi}\right)\otimes H^q(x,\xi,t,\mu^2) = \pi e_q^2\left[H^q(\xi,\xi,t,\mu^2)-H^q(-x,\xi,t,\mu^2)\right], \tag{3.14}$$

$$= \pi e_q^2 H^{q(+)}(\xi,\xi,t,\mu^2)\,, \tag{3.15}$$

$$\tag{3.16}$$

and for the real part

$$\text{Re}\, T_0^q\left(\frac{x}{\xi}\right)\otimes H^q(x,\xi,t,\mu^2) = e_q^2\ \text{p.v.}\int_{-1}^{1}\text{d}x\left[\frac{1}{\xi-x}-\frac{1}{\xi+x}\right]H^q(x,\xi,t,\mu^2)\,, \tag{3.17}$$

$$= e_q^2\ \text{p.v.}\int_{0}^{1}\text{d}x\left[\frac{1}{\xi-x}-\frac{1}{\xi+x}\right]H^{q(+)}(x,\xi,t,\mu^2)\,, \tag{3.18}$$

where we have used the explicit x parity of the DVCS coefficient function to restrict integration on $[0,1]$. It is a general feature that, as the DVCS process is C-even, its coefficient functions have a definite parity in terms of x and $T^a\otimes H^{a(-)}=0$. Therefore it is already clear that DVCS does not provide any information on non-singlet GPDs.

Due to the presence of Cauchy's principal value in Eq. (3.18), this expression is still a bit cumbersome to evaluate numerically. Subtracting the diagonal $H^{q(+)}(\xi,\xi,t,\ \mu^2)$ increases the regularity of the integrand, yielding an expression which can be computed straightforwardly without principal value for $\xi>0$:

$$\begin{aligned}&\text{Re}\, T_0^q\left(\frac{x}{\xi}\right)\otimes H^q(x,\xi,t,\mu^2)\\ &= -e_q^2\int_0^1\text{d}x\left[\frac{H^{q(+)}(x,\xi)-H^{q(+)}(\xi,\xi)}{x-\xi}+\frac{H^{q(+)}(x,\xi)}{\xi+x}\right]-e_q^2H^{q(+)}(\xi,\xi)\log\left(\frac{1-\xi}{\xi}\right),\end{aligned} \tag{3.19}$$

where we have omitted (t,μ^2) dependence in the right hand side for brevity and used that for $\xi>0$,

$$\text{p.v.}\int_0^1\frac{\text{d}x}{x-\xi} = \log\left(\frac{1-\xi}{\xi}\right). \tag{3.20}$$

The coefficient functions T^a expressed at NLO have been computed for instance in Refs. [121, 17, 122, 123, 124, 29]. Two loops corrections have been recently published in Refs. [125, 126]. The general structure of the perturbative expansion of the coefficient function will be studied in detail in Section 6.3.1. We show in Appendix A.3 how to derive expressions without $i\epsilon$ on a more complicated example compared to the simple LO one, as such expressions will be useful both in Chapters 5 and 6.

Most studies of DVCS have been performed at LO, taking advantage of the remarkable direct access to the diagonal of the GPD provided by the imaginary part of the LO coefficient function (3.15), and using the fact that the limited coverage in Q^2 of the current experimental DVCS dataset does not require a careful study of Q^2 dependence. NLO studies of DVCS have been pioneered in Ref. [108]. Future facilities with an extended coverage in Q^2, which we will detail in Section 3.3, make studies at least at NLO necessary. We will therefore develop a full NLO approach of the question of whether or not it is possible to recover $H^{a(+)}$ from the knowledge of $T^a \otimes H^a$, issue known as the **deconvolution problem**.

DVCS observables

Before even trying to invert the convolution of Eq. (3.10), extracting CFFs from experimental data is already a challenge in itself. Experimentalists have the possibility to measure different kinds of observables : notably total or differential cross-sections and asymmetries. The latter are defined as the ratio of cross-sections, and offer an intrinsic advantage since some systematic uncertainties linked to the normalisation of the cross-sections cancel out. As an example of asymmetry, let us mention the beam spin asymmetry (BSA) measured for longitudinally polarised electrons and unpolarised hadrons defined by [127]

$$A_{LU}^{-}(x_B, t, Q^2, \phi) = \frac{\mathrm{d}^4\sigma^{\rightarrow,-}(x_B, t, Q^2, \phi) - \mathrm{d}^4\sigma^{\leftarrow,-}(x_B, t, Q^2, \phi)}{\mathrm{d}^4\sigma^{\rightarrow,-}(x_B, t, Q^2, \phi) + \mathrm{d}^4\sigma^{\leftarrow,-}(x_B, t, Q^2, \phi)} \,. \tag{3.21}$$

The superscript $^-$ denotes the fact that the asymmetry is measured with a negatively charged beam, the $_{LU}$ subscript characterises the polarisations of the beam (L for longitudinal) and target (U for unpolarised), and the superscripts $^{\rightarrow}$, $^{\leftarrow}$ denote the helicity of the beam.

The Fourier harmonics of A_{LU}^{-} are defined by

$$A_{LU}^{-,\cos n\phi} = \frac{1}{\pi} \int_0^{2\pi} \mathrm{d}\phi \, \cos(n\phi) A_{LU}^{-}(\phi) \,, \tag{3.22}$$

$$A_{LU}^{-,\sin n\phi} = \frac{1}{\pi} \int_0^{2\pi} \mathrm{d}\phi \, \sin(n\phi) A_{LU}^{-}(\phi) \,. \tag{3.23}$$

The parameterisation in terms of CFFs of the BH-DVCS interference contribution to $A_{LU}^{-,\sin\phi}$ can be expressed at leading twist and leading order in $1/Q$ [128]

$$A_{LU,\mathcal{I}}^{-,\sin\phi} \propto \mathrm{Im}\left[F_1\mathcal{H} + \xi(F_1+F_2)\widetilde{\mathcal{H}} - \frac{t}{4M^2}F_2\mathcal{E}\right] . \tag{3.24}$$

Different observables – including different Fourier harmonics of the same asymmetry – are sensitive to different combinations of CFFs and provide therefore various handles to perform CFF extraction from experimental data. Refs. [129, 116] summarise expressions of various observables in terms of CFFs. The question of determining which observables to measure in which kinematic regions is central to increase the experimental knowledge on CFFs. We will develop this aspect in Section 4.4, where we present another observable known as beam charge asymmetry, and gauge the impact of its potential measurement at JLab on the knowledge of $\mathrm{Re}\,\mathcal{H}$.

We have overlooked in this very brief presentation many subtleties that render the extraction of CFFs from experimental data particularly challenging. Notably, $1/Q$ power corrections which we have neglected may be quite large considering the low values of Q^2 at which most data are obtained, which forces to consider higher twist CFFs. Corrections in t/Q^2 and M/Q^2 are presented for instance in Refs. [130, 131, 132] and exploited in Ref. [133].

Two main strategies of extraction of CFFs from experimental data have been developed. The first consists in so-called **local fits** (see for instance Refs. [134, 135, 136, 137, 138, 139, 140, 141]). Independently for each kinematic bin, the CFFs are fitted as free parameters on the experimental data. This method introduces therefore no bias on the general form of the CFF apart from the chosen framework of general approximation, like twist truncation, neglection of some CFFs, assumption of dominance of either BH or either DVCS, ... Local fits have however no ability to predict the result of measurements in previously unexplored kinematics. On the other hand, **global fits** (see for instance Refs. [135, 137, 108, 109, 53]) assume a functional shape of the CFFs and adjust it using data on all available kinematics at once. The model may be built at the level of the CFF, or directly at the level of the GPD. We will explore in Section 3.4 several distinct strategies to model GPDs. The global fit method allows one to predict measurements in unprobed kinematics, but at the cost of introducing some level of model dependence. Efforts to reduce this bias have been led, notably by the introduction of neural network parameterisations of CFFs [142, 143]. We will present this aspect in detail in Section 4.2.

The difficulty of CFF extractions, which require to manipulate a large number of observables, as well as expansions in twist and orders of α_s, has motivated the development of integrated softwares dedicated to the phenomenology of GPDs, such as PARTONS [144] (`https://partons.cea.fr`) and GeParD (`https://gepard.phy.hr`). The study of the deconvolution problem of DVCS will show how multi-channel analyses – that is including processes different from DVCS – is crucial to extract GPDs, making the pertinence of these integrated software approaches all the more stringent.

3.1.2 . Dispersion relations

A remarkable description of CFFs is inherited from the study of their analytical properties [145, 108, 146]. The ξ dependences of the real and imaginary parts of CFFs are related through the following dispersion relation with one subtraction [3]

$$\mathrm{Re}\,\mathcal{H}(\xi,t,Q^2) = \frac{1}{\pi}\mathrm{p.v.}\int_0^1 \mathrm{d}\xi'\,\mathrm{Im}\,\mathcal{H}(\xi',t,Q^2)\left(\frac{1}{\xi-\xi'}-\frac{1}{\xi+\xi'}\right) + C_H(t,Q^2)\,. \tag{3.25}$$

$C_H(t,Q^2)$ is known as the **subtraction constant** as it is independent of ξ. The derivation of Eq. (3.25) finds its roots in the general analytical properties of scattering amplitudes, notably causality and unitarity. In particular, it does not depend on the formalism of perturbative QCD and the development of dispersion relations applied to hadron scattering

3. The number of necessary subtractions is dictated by the behaviour of the CFF at high-energy, which corresponds to the limit $\xi \to 0$. Commonly advocated Regge trajectories give that the CFF behaves as $\xi^{-\alpha}$ with $1 < \alpha < 2$, so one subtraction is thought to be enough and yields a ξ independent term $C_H(t,Q^2)$. If a higher value of α is encountered, more subtractions need to be introduced (see Refs. [146, 108] for more details), and the derived dispersion relation will involve terms of the form $C_H^{[k]}(t,Q^2)/\xi^k$ with $k > 0$ even, which are therefore no longer constant with respect to ξ.

dates back to the 1950s[4]. Ref. [148] underlines that since the value of ξ is bounded by kinematic considerations (2.1), the integral in Eq. (3.25) is partly performed in an unphysical region where the CFF simply cannot be measured. In practice, it means that $\mathrm{Im}\,\mathcal{H}$ must be continued in some manner, for instance thanks to a model of its ξ dependence, to extract the subtraction constant. In spite of the difficulty linked to the unphysical large values of ξ, let us mention that the real part of the CFF contains an information on the imaginary part integrated from $\xi' = 0$ to the upper kinematic bound, which represents an opportunity to constrain the imaginary part at small values of ξ which are difficult to reach in experiments.

Ref. [146] observes that the coefficient functions obey similar dispersion relations thanks to their interpretation as hard parton scattering. The authors demonstrate that the consistency of the dispersion relations at the level of the CFF and the coefficient function, combined with the requirements of Lorentz covariance on GPDs expressed through the form of Eq. (2.25), implies that, for a coefficient function T_k^a computed at order α_s^k, the subtraction constant reads

$$C_H(t, Q^2) = \sum_a C_{H,k}^a[D^a](t, Q^2, \mu^2) + \mathcal{O}(\alpha_s^{k+1})\,, \tag{3.26}$$

where

$$C_{H,k}^a[D^a](t, Q^2, \mu^2) = \frac{2}{\pi}\int_1^\infty \frac{\mathrm{d}\omega}{\omega^{p_a}}\,\mathrm{Im}\,T_k^a\left(\omega, \frac{Q^2}{\mu^2}, \alpha_s(\mu^2)\right)\int_{-1}^1 \mathrm{d}z\,\frac{D^a(z, t, \mu^2)}{\omega - z}\,. \tag{3.27}$$

To differentiate the experimental value of the subtraction constant $C_H(t, Q^2)$ defined in Eq. (3.25) from the computation of the contribution of the D-term to the subtraction at $\mathrm{N}^k\mathrm{LO}$, we introduce the notation

$$C_{H,k}[D^q, D^g](t, Q^2) \equiv \sum_a C_{H,k}^a[D^a](t, Q^2, \mu^2)\,. \tag{3.28}$$

Therefore,

$$C_H(t, Q^2) = C_{H,k}[D^q, D^g](t, Q^2) + \mathcal{O}(\alpha_s^{k+1})\,. \tag{3.29}$$

At LO in $\alpha_s(\mu^2)$, $T_0^g = 0$ and $\mathrm{Im}\,T_0^q(\omega) = \pi e_q^2[\delta(1-\omega) - \delta(1+\omega)]$ (3.13), so

$$C_{H,0}[D^q, D^g](t, Q^2) = 2\sum_q e_q^2 \int_{-1}^1 \mathrm{d}z\,\frac{D^q(z, t, \mu^2)}{1 - z}\,. \tag{3.30}$$

The derivation of an NLO quark contribution is demonstrated in Appendix A.3. The subtraction constant represents a channel for the specific extraction of the D-term which we will exploit in detail in Chapter 5. On the contrary, the imaginary part of the CFF allows the access to the pure DD f^a contribution, whose exact nature we will study in Chapter 6.

4. See for instance the interesting discussion of Ref. [147] where it is heuristically argued why the real part of a scattering can be obtained as an integral of the imaginary part. It is also demonstrated mathematically that "strict causality" amounts for the scattering amplitude to be the boundary value on a complex half-plane of an analytical function of its complexified kinematic dependence. The $i\epsilon$ prescription allows one to identify how the boundary value should be approached from the complex plane and is deeply related to the requirement of causality.

3.2 . Sensitivity to other processes

Although DVCS is widely considered as the golden channel to extract GPDs thanks to its relative theoretical cleanness, we have already mentioned in Chapter 1 that GPDs are universal objects which enter the description of other exclusive processes. As DVCS will occupy most of our attention in this document, we only briefly mention a few of these additional channels, and some properties relevant to our study.

Time-like Compton scattering

Time-like Compton scattering (TCS) [26] is a close relative to DVCS since it corresponds to a similar process, except the incoming photon is real, and outgoing photon is virtual. The virtual photon usually decays in a lepton pair in the final state. Due to its proximity with DVCS, TCS can be described in terms of CFFs with coefficient functions which are notably detailed in Refs. [29, 149].

The similarity of the two processes allows one to extend to TCS theoretical developments performed for DVCS. The slight differences in coefficient functions provide a first test of universality of GPDs : can objects derived from DVCS describe equally well TCS observables ? The first experimental measurements have been recently obtained at JLab [150] and pave the way to this important test of universality.

Deeply virtual meson production

DVMP, whose schematic depiction is given on the right panel of Fig. 1.1, played an important role in the introduction of GPDs. However, because mesons are bound states of QCD, DVMP involves further theoretical refinements compared to DVCS or TCS. Its factorisation [14], proved for longitudinally polarised photons, involves not only GPDs and coefficient functions known at NLO [151], but also another non-perturbative object : the distribution amplitude (DA). The equivalent of a CFF for DVMP, called transition form factor $\mathcal{T}$ (TFF), reads generically at twist-two as

$$\mathcal{T}(\xi,t,Q^2) = \int_0^1 \mathrm{d}u \int_{-1}^1 \mathrm{d}x\, T_{DVMP}\left(x,\xi,u,\frac{Q^2}{\mu^2},\alpha_s(\mu^2)\right) H(x,\xi,t,\mu^2)\phi(u,\mu^2)\,, \tag{3.31}$$

where $\phi(u,\mu^2)$ is the leading-twist meson DA and T_{DVMP} a coefficient function which now depends jointly on x,ξ and u. At LO, the coefficient function T_{DVMP} can be separated into $T_{DVMP}(x,\xi,u) = T_0^{(1)}(u)T^{(2)}(x,\xi)$ under the form [151]

$$\mathcal{T}(\xi,t,Q^2) \overset{LO}{=} \left[\int_0^1 \mathrm{d}u\, T_0^{(1)}(u)\phi(u,\mu^2)\right]\left[\int_{-1}^1 \mathrm{d}x\, T^{(2)}\left(x,\xi\right) H(x,\xi,t,\mu^2)\right], \tag{3.32}$$

$$\propto \alpha_s(\mu^2)\left[\int_0^1 \mathrm{d}u\, \frac{\phi(u,\mu^2)}{1-u}\right]\left[\int_{-1}^1 \mathrm{d}x \left(\frac{1}{\xi-x-i\epsilon}\right) H^{(\pm)}(x,\xi,t,\mu^2)\right], \tag{3.33}$$

where we have only shown a generic quark contribution for simplicity. We refer to Ref. [151] for a detailed presentation of the results at LO and NLO. One will notice that the x-dependent LO DVMP coefficient function is exactly the same as the LO DVCS one. The TFF is sensitive to singlet GPDs for longitudinally polarised light vector mesons,

and to non-singlet GPDs for pseudo-scalar mesons, providing therefore a complementary knowledge on GPDs compared to DVCS.

A closely related process is heavy vector meson production (HVMP) where the hard scale of the process is not provided by the virtuality of the incoming photon, but the mass of the produced meson. One can therefore consider incoming real or quasi-real photons, as for instance in ultra-peripheral collisions [152]. As HVMP is the subject of a significant theoretical and experimental interest due to the possibility of constraining gluon PDFs, we will focus more on this process in Section 7.3.

Processes with a richer kinematic structure

The three processes we have mentioned so far can be described in terms of form factors (either CFFs or TFFs) which depend on (ξ, t, Q^2). As GPDs are functions of (x, ξ, t, μ^2), one variable is "missing" in the kinematic structure of the form factors. In practice, the x variable is integrated out in the convolutions of Eqs. (3.10) and (3.31). We will observe in Chapter 6 that this renders the deconvolution problem – that is the extraction of GPDs from experimental data – particularly difficult.

Several other processes involving GPDs with a richer kinematic structure have been studied. First, allowing the outgoing photon of DVCS to be virtual while keeping the incoming one also virtual yields double DVCS (DDVCS) [27, 28, 29]. DVCS and TCS are therefore limiting cases of this more general process. Noting q' the four-momentum of the outgoing virtual photon and keeping q for the incoming virtual photon and $p_{1,2}$ for the hadron four-momenta, it is possible to define specifically for DDVCS

$$\xi' = -\frac{(q+q')^2}{4(p_1+p_2)\cdot(q+q')} \quad \text{and} \quad \xi = -\frac{(p_2-p_1)\cdot(q+q')}{2(p_1+p_2)\cdot(q+q')}\,. \tag{3.34}$$

In the DVCS limit, $\xi' = \xi$, while for TCS, $\xi' = -\xi$. DDVCS form factors depend on (ξ', ξ, t, Q^2), and the imaginary part of the convolution with the LO coefficient function writes [29]

$$\begin{aligned}
&\mathrm{Im}\, T^q_{0,DDVCS}(x,\xi') \otimes H^q(x,\xi,t,\mu^2) \\
&\qquad = \mathrm{Im} \int_{-1}^{1} \mathrm{d}x\, e_q^2 \left(\frac{1}{\xi' - x - i\epsilon} - \frac{1}{\xi' + x - i\epsilon} \right) H^q(x,\xi,t,\mu^2)\,, \qquad (3.35) \\
&\qquad = \pi e_q^2 H^{q(+)}(\xi',\xi,t,\mu^2)\,. \qquad (3.36)
\end{aligned}$$

Therefore, the imaginary part of the DDVCS form factor at LO gives a direct access to values of the GPD outside of the diagonal without the need for any deconvolution. Let us note however that, because the outgoing virtual photon is time-like, $|\xi'| \leq |\xi|$, so we can only probe a limited region of the GPD. An analogous strategy with respect to the one we will present in Section 3.4.2 may be developed in this situation to recover the full GPD based on the requirements of Lorentz covariance.

Other channels with a richer kinematic structure compared to DVCS, TCS and DVMP are processes where several particles are produced in the final state. For instance, the production of a pair of photons [30, 31, 32] or a photon-meson pair [33, 34] have been suggested. However, these processes, as well as DDVCS, are more challenging experimentally and their plausible impact on GPD extraction is still being studied.

Lattice-QCD inputs on parton distributions

Although computations on the lattice do not belong to experimental measurements, the progress in numerical simulations deserves a mention. This progress is first characterised by an increase in computational power and an improvement of algorithms which allow one to carry out simulations at the physical value of parameters like the pion mass (see for instance Ref. [153]). But it is also and maybe more importantly a profound conceptual change. For a long time, lattice-QCD was understood to only be able to compute values of Mellin moments of parton distributions, accessible from local matrix elements in Euclidean space. In practice only the first three moments could even be computed due to divergence and noise issues [154].

The developments of new techniques to match Euclidean correlators to light-cone ones has however completely changed the picture. It is now possible to extract x-dependent parton distributions owing to a matching procedure quite analogous to the standard factorisation of experimental cross-sections. Two approaches are most commonly met : the quasi-distribution formalism [155] and the pseudo-distribution one [156]. A recent review on the subject may be found in Ref. [157]. The precise control of systematic uncertainty, both in the computation of matrix elements and in the matching procedure to the light-cone, is still being actively investigated. GPDs represent in addition further complications compared to usual PDFs, so experimental data are bound to remain the major source of information on GPDs still for some years.

3.3 . Status of available DVCS data and future experimental projects

The first measurements of DVCS started about 20 years ago, at the HERA accelarator (DESY, Germany) and at JLab in the US. At HERA, the H1 [158] and ZEUS [159] collaborations used a collider set-up, giving access to the smallest values of x_B and largest values of Q^2 where DVCS was measured so far. Values as small as $x_B \sim 10^{-4}$ and as large as $Q^2 \sim 100$ GeV^2 were probed as shown on Fig. 3.3, unfortunately with only with a rather low statistics. Having access to both an electron and a positron beam, they published beam charge asymmetries [160] (see Section 4.4) as well as cross-sections dominated by the pure DVCS contribution at small values of x_B.

At HERA also, the HERMES collaboration [161] used a fixed target which could be longitudinally or transversely polarised. They published a complete set of asymmetries for x_B in the range 0.04 to 0.1 with Q^2 up to about 7 GeV^2.

At JLab, the CLAS [162, 163] and Hall A [133, 164] collaborations also used a fixed target, although they worked at a larger value of x_B of the order of 0.1 to 0.5 and even lower Q^2 up to about 4 GeV^2. With the upgrade of JLab 12 GeV, an increased kinematic region with better experimental precision is expected in the coming years.

At the CERN, the COMPASS collaboration [47] with a fixed target provides an intermediate range in x_B between HERMES and H1 / ZEUS data, with the possibility of using both positively and negatively charged muons.

For now, most of the statistically constraining data for global fits of DVCS is obtained at rather low values of Q^2 with the bulk of points ranging from 1.5 to 4 GeV^2. It is well understood that obtaining data on a larger range in Q^2 is crucial, and we will demonstrate in detail the importance of this matter in Chapters 5 and 6. Obtaining data with a higher luminosity than HERA at very low x_B offers also very interesting physical perspectives

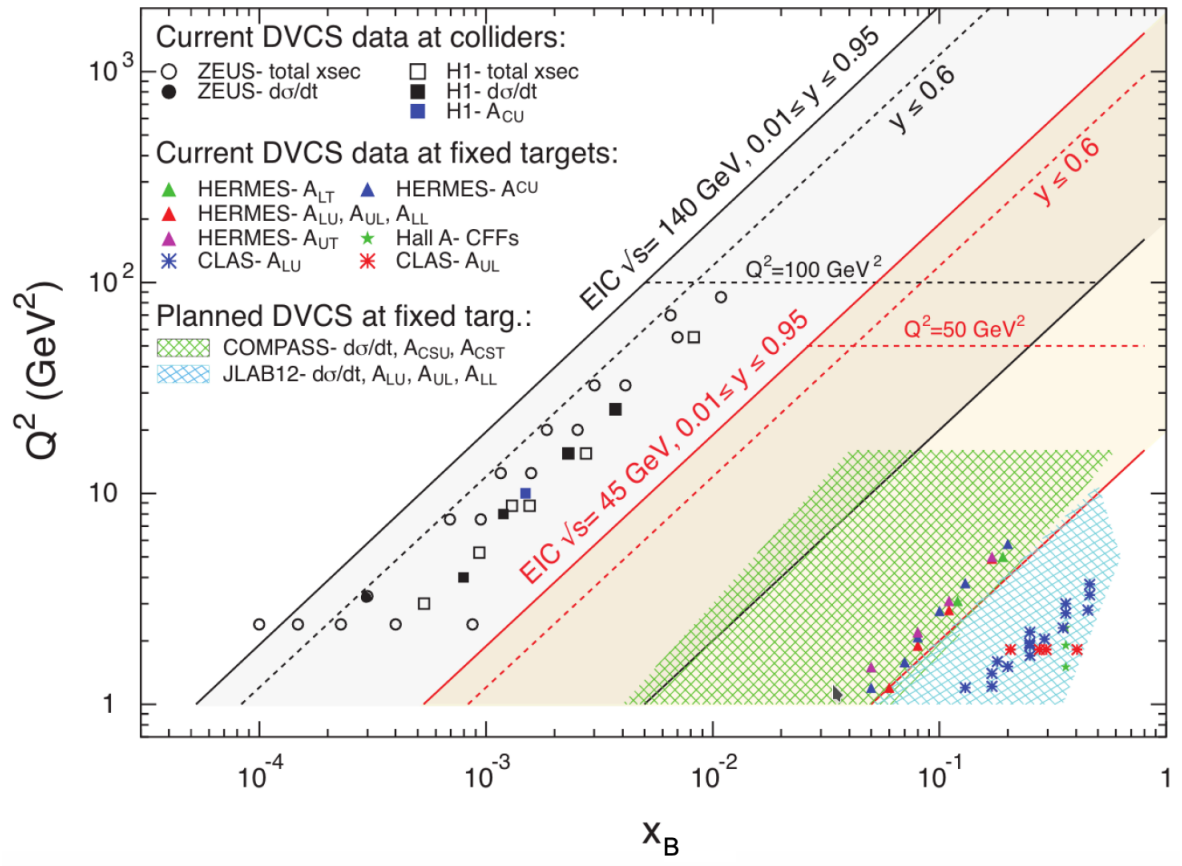

Figure 3.3 – Summary of the current and planned DVCS measurements in terms of x_B and Q^2. A few measurements are missing, without significantly changing the picture, as this figure was also used in the 2021 EIC yellow report. Figure taken from Ref. [46].

which we will mention more precisely in Section 7.3. Overall, an increased precision of experimental data is necessary to progress on the physical motivations for the study of GPDs, like hadron tomography (Section 1.4.1) and the extraction of mechanical properties (Section 1.4.2).

The future electron ion collider (EIC) [46, 47], which has already received a serious governmental commitment in the US, is at the center of a lot of attention thanks to its promise of a high luminosity coverage over an extended region at relatively small x_B and large Q^2 as depicted on Fig. 3.3. When deriving numerical estimates in this work, we will frequently have the EIC in mind and a coverage in Q^2 roughly estimated from 1 to 50 or 100 GeV^2. Let us mention also competing projects with the hadron-electron collider (LHeC) [50] and the Chinese electron-ion collider (EIcC) [48, 49].

3.4 . Models of generalised parton distributions

Due to the scarcity of experimental measurements and the difficulty of the overall procedure of extraction of GPDs from experimental data, models have played a very important role in phenomenology. We choose to only present three selected strategies of modelling in this section, and refer to Ref. [116] for a more detailed presentation of the subject. The three strategies we will present, one based on double distributions, one on

conformal moments and one on the inversion of the Radon transform, will all be revisited in the next chapters. We will present new models of double distributions in Section 7.1 and 7.2, where we intend to provide a lesser model dependence and a better flexibility. In Section 7.3, we will study a proposal of modelling GPDs at small ξ based on the Shuvaev transform for conformal moments, reinterpret its validity and justify why we feel that an alternative view on the matter is relevant. Finally we will propose a different strategy for the inversion of the Radon transform in Section 6.2.

3.4.1 . RDDA-based models

We remind that the DD formalism is equivalent to satisfying the requirements of Lorentz covariance on the polynomiality of Mellin moments of GPDs. A simple way to combine DDs and the excellent knowledge of PDFs is known as Radyushin's double distribution Ansatz (RDDA), introduced in Ref. [165]. In this Ansatz, the DD $f^q(\beta, \alpha, t)$ is modelled as

$$f^q(\beta, \alpha, t) = \pi_N(\beta, \alpha) q(\beta, t)\,, \tag{3.37}$$

where $q(x, t)$ produces the usual PDF $q(x)$ in the limit where $t = 0$, and π_N is called a profile function. Eq. (2.25) gives that

$$H^q(x, 0, t) = \int_{|x|-1}^{1-|x|} \mathrm{d}\alpha\, f^q(x, \alpha, t)\,, \tag{3.38}$$

$$= q(x, t) \int_{|x|-1}^{1-|x|} \mathrm{d}\alpha\, \pi_N(x, \alpha)\,, \tag{3.39}$$

so to produce the expected forward limit at $t = 0$, the profile function must be normalised in the following way :

$$\int_{|\beta|-1}^{1-|\beta|} \mathrm{d}\alpha\, \pi_N(\beta, \alpha) = 1\,. \tag{3.40}$$

A straightforward way to satisfy Eq. (3.40) is to propose

$$\pi_N(\beta, \alpha) = \frac{\Gamma(N + 3/2)}{\sqrt{\pi}\Gamma(N + 1)} \frac{((1 - |\beta|)^2 - \alpha^2)^N}{(1 - |\beta|)^{2N+1}}\,. \tag{3.41}$$

The fact that the α dependence of the RDDA is merely controlled by a single parameter N produces a rather inflexible modelling. In fact, the model converges quickly as N increases as demonstrated in Refs. [166, 167], so the effective freedom is limited.

The RDDA is however the basis of two very popular phenomenological models : the Vanderhaeghen - Guichon - Guidal (VGG) model [168, 169, 170, 171], and the Goloskokov - Kroll (GK) model [172, 173, 174]. The main difference between the two resides in the implementation of the t dependence of $q(x, t)$. VGG uses the generic form $q(x, t) \propto q(x)x^{-\eta(1-x)t}$ for the unpolarised GPD H^q, whereas GK uses $q(x, t) \propto q(x)x^{-\eta t}e^{\nu t}$. The exponents η and ν are fixed to approximately reproduce the large t behaviour of the EFF $F_1(t)$.

Although initially tailored on DVMP experimental data, the GK model produces a good agreement in a LO description of DVCS experimental data on a wide kinematic range as demonstrated in Ref. [129]. However, Ref. [175] argues that predictions obtained with RDDA models are generally too large at small x_B, and that the good agreement pointed out in Ref. [129] is coincidental.

3.4.2 . Partial wave expansion

Another popular approach to the modelling of GPDs is to work in the space of conformal moments introduced in Section 2.4. The relation between GPDs and conformal moments can be represented in terms of a formally divergent infinite polynomial expansion

$$H^q(x,\xi,t) = \sum_{j=0}^{\infty} \frac{1}{\xi} \left(-\frac{2}{\xi}\right)^j \frac{\Gamma(j+5/2)}{\Gamma(3/2)\Gamma(j+3)} \left(1-\frac{x^2}{\xi^2}\right) C_j^{(3/2)}\left(-\frac{x}{\xi}\right) \mathcal{O}_j^q(\xi,t)\,, \tag{3.42}$$

where the $C_j^{(3/2)}$ are Gegenbauer polynomials, and $\mathcal{O}_j^q(\xi,t)$ the conformal moments defined in Eq. (2.54). We have already mentioned in Section 2.4 that numerical studies of reconstruction of GPDs from conformal moments based on polynomial expansions face difficulties due to the large number of terms required to faithfully reconstruct the GPD, but analytical models can be used.

The simplest model consists probably in choosing $\mathcal{O}_j^q(\xi,t)$ to be ξ independent with the same t dependence for all j. This leads to an interesting proposal of model at small ξ – with some shortcomings – which we will look into in detail in Section 7.3.

A less trivial way to model the ξ and t dependence of the conformal moments is obtained in the formalism of SO(3) partial wave expansion in the t-channel [39, 176, 177, 108] :

$$\mathcal{O}_j^q(\xi,t) = \sum_{J=J_{\min}}^{j+1} \mathcal{O}_j^J(t)\xi^{j+1-J}\hat{d}^J(\xi)\,, \tag{3.43}$$

where the summation is performed over the t-channel angular momentum J. The $\hat{d}^J(\xi)$ functions derive from the Wigner rotation matrix and can be written in terms of Jacobi polynomials.

At $t=0$, the leading partial wave amplitude $\mathcal{O}_j^{j+1}(t=0)$ is the j-th Mellin moment of the PDF

$$q_j \equiv \mathcal{O}_j^{j+1}(t=0) = \int_0^1 \mathrm{d}x\, x^j q(x)\,. \tag{3.44}$$

The SO(3) expansion can be envisioned as the exchange of mesons carrying a total angular momentum J (see Refs. [177, 108, 109] for more details). A simplified proposal of dependence at small ξ is obtained by modelling this exchange of mesons by Regge trajectories :

$$\mathcal{O}_j^q(\xi,t) = q_j(1+s_2\xi^2+s_4\xi^4+...)\frac{1+j-\alpha}{1+j-\alpha-\alpha' t}\left(1-\frac{t}{m^2}\right)^{-p}\,, \tag{3.45}$$

where α and α' parameterise the Regge trajectory and p and m control the strength of the coupling between the hadron target and the meson. More explanations on the physical hypotheses leading to the model of Eq. (3.45) are found in Ref. [116] for instance.

3.4.3 . The covariant extension

After having broadly discussed modelling strategies based on double distributions and conformal moments, let us briefly mention a third distinct path of modelling, based on LFWFs introduced in Section 2.3. We have already noticed that, depending on the kinematic region under consideration, the representation of GPDs as an overlap of LFWFs

involves either an overlap of LFWFs with the same number of particles for $|x| > |\xi|$ (2.41), or an overlap of LFWFs with differing number of particles in the region $|x| < |\xi|$ (2.42). In the latter case, the mismatch in the number of particles in the Fock states triggers an ambiguity on the Fock space truncation : assuming a truncation at N particles has been implemented in the $|x| > |\xi|$ region, how should we proceed for $|x| < |\xi|$? Should we only keep overlap between at most N and $N-2$ particles ? Or are overlap between $N+1$ and $N-1$ particles also allowed ? There does not seem to be any natural good answer.

As the Fock space truncation does not produce ambiguities in the region $|x| > |\xi|$, a technique proposed in Ref. [77] consists in modelling the GPD in that region from LFWFs, and using afterwards the requirements of Lorentz covariance to perform the **covariant extension** of the GPD to the region $|x| < |\xi|$. In practice, a DD is extracted thanks to the inverse Radon transform from the knowledge of the GPD in the region $|x| > |\xi|$, and is then used to compute the value of the GPD for $|x| < |\xi|$ [5].

This procedure allows therefore to keep at the same time the advantages of the LFWF overlap representation, which ensures the respect of the complicated set of positivity constraints in the $|x| > |\xi|$ region, and the advantages of the DD representation, which guarantees correct properties with respect to Lorentz covariance.

A practical implementation of the inversion of the Radon transform is presented in Refs. [77, 179] thanks to a finite element method. The DD domain in (β, α) is discretised in a mesh with several hundred nodes, and a linear system relating the values of the DD on the nodes of the mesh to the targeted GPD in the $|x| > |\xi|$ region is built and then inverted. We will provide more technical details in Section 6.2, where we will propose an alternative strategy to the finite element method. Indeed, the formalism of DDs defined as polynomials in α and β which we will develop for a completely separate objective can be applied to obtain a covariant extension with very satisfactory results. The modelling based on DDs which we will later expose in Section 7.2 follows from the same general idea than the covariant expansion we have sketched here. However, instead of starting from positivity bounds (LFWF formalism) and implementing Lorentz covariance (DD formalism), we will follow the inverse path, starting from a DD and showing how it can satisfy positivity bounds thanks to the shadow distributions we introduce in Chapters 5 and 6.

5. Formally, it is demonstrated in Ref. [77] from a theorem in Ref. [178] that the knowledge of the GPD in the region $|x| > |\xi|$ is sufficient to exactly reconstruct the DD f^q, up to an ambiguity on the $\beta = 0$ line which is not probed by Radon integration lines in that kinematic region. In particular, the D-term, which writes in the DD formalism as $\delta(\beta) D^q(\alpha, t)$ (2.25), cannot be reconstructed, as well as any $\delta(\beta)$ singularity included directly in the DD f^q. This means that once the DD has been extracted, some new modelling choices are likely necessary to eliminate the ambiguity, but they will be exactly compatible with polynomiality of Mellin moments and positivity bounds.

Part II

Extraction of generalised parton distributions

4 - Neural network parameterisation of Compton form factors

As a prerequisite to the various analyses and modelling efforts led in this thesis, it is necessary to familiarise ourselves with uncertainty propagation from an experimental dataset to a generic model, and the general principles of neural network modelling. In this chapter, we notably demonstrate on a custom example the advantage of the replica method over the traditional Hessian method to estimate uncertainty. The last part of this chapter is devoted to the question of adding new experimental data to an existing fit. This issue is of significant interest since complex models like neural networks call for lengthy optimisation procedures that frequently require tens of hours if not days of computing on clusters. Assessing broadly but reliably the impact of a new dataset without needing the time and computing power of a full refit is therefore useful to perform impact studies and identify which observables and kinematic regions would bring the most striking constraints on extractions. Considering the number of available observables for experimentalists, and the projects of future facilities which would explore different kinematic domains, reliable impact studies have come at the forefront of hadron physics. We first demonstrate results on the addition of a new datapoint to an existing fit in the linear case, before focusing on the more general Bayesian reweighting procedure. We finally present the results of the study led in Ref. [2] where we assess the impact of a positron beam at JLab on the experimental uncertainty of Compton form factors.

4.1 . Modelling uncertainty with neural networks

4.1.1 . Some general aspects of uncertainty modelling

The most straightforward way to assess the uncertainty on a quantity measured on various kinematics is to propose a functional form of the relation between the quantity of interest and the kinematic variables, and fit its free parameters to the experimental data. The fit consists in finding the optimal parameters of the model so as to minimise a loss function like the least squares estimate, or maximise the likelihood of the observed dataset. The actual uncertainty on the extraction is then derived from the behaviour of the loss function around its extremum. Intuitively, if the likelihood is flat near its maximum, nearby parameters are equally likely, and the uncertainty is large. It is clear that this procedure to assess experimental uncertainty suffers from several sources of bias, notably the choice of model and loss function, and the practical procedure to study the behaviour of the loss function around its minimum. We will call the bias introduced by modelling choices **epistemic uncertainty** following the terminology described for instance in Ref. [180]. On the contrary, the intrinsic randomness contained in the fitted data also translates into an uncertainty in the extraction, which we will refer to as **aleatoric uncertainty**. We first present results concerning the aleatoric uncertainty in the case of a linear model, before discussing the general strategy we used to evaluate uncertainty in non-linear cases. Discussion of epistemic uncertainty will be led in the next section.

Uncertainty quantification in the linear Gaussian case

It is useful to study briefly the case where the fitted model is linear in its parameters and the experimental uncertainty is normally distributed. A linear model can be written in general as

$$f_A(X) = X^T A = \sum_{k=1}^{n} a_k x_k \,, \tag{4.1}$$

where $A = (a_k)$ is the vector of free parameters of length n and $X = (x_k)$ a vector of functions of the kinematic variables. X^T denote the transposition of matrix X. To clarify the notations, let us take the example of a model which depends on a kinematic variable μ through a linear combination of $\sin(\mu)$ and $\cos(\mu)$:

$$f(\mu) = a\sin(\mu) + b\cos(\mu)\,. \tag{4.2}$$

a and b are free parameters. Then the model $f(\mu)$ can be represented as $f_A(X)$ with our previous notations by the vector of free parameters $A = \begin{pmatrix} a \\ b \end{pmatrix}$ and the vector of functions of the kinematic variables $X = \begin{pmatrix} \sin(\mu) \\ \cos(\mu) \end{pmatrix}$.

Our goal is to estimate the aleatoric uncertainty on parameters A from a set of m experimental measurements $F = (f_i)$ for $1 \leq i \leq m$. The functions of the kinematic variables on which the measurements are obtained are stored in a matrix C with m lines and n columns. Using again our previous example, let us assume we want to fit a dataset made of two points $(\mu_1 = 0, f_1 = 1)$ and $(\mu_2 = 1, f_2 = 1)$. Then $F = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is the vector of measurements, and $C = \begin{pmatrix} \sin(\mu_1) & \cos(\mu_1) \\ \sin(\mu_2) & \cos(\mu_2) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \sin(1) & \cos(1) \end{pmatrix}$ is the matrix containing information on the probed kinematics. As we have exactly the same number of measurements as free parameters, there exist a set of parameters such that the model reproduces exactly the vector of measurements F. This set of parameters, which we will note $A^{(0)}$, is obtained by solving a simple 2×2 linear system. Observing the structure of the matrix C demonstrates that this system writes $CA^{(0)} = F$. Therefore

$$A^{(0)} = C^{-1}F = \begin{pmatrix} 1/\sin(1) - 1/\tan(1) \\ 1 \end{pmatrix} . \tag{4.3}$$

In general, we do not expect our model to be able to reproduce exactly the central values of the measurements. We are furthermore interested in understanding how uncertainty on the measurement vector F is translated into uncertainty on the best fit parameters. Let us therefore assume that F follows a multivariate normal distribution of dimension m with covariance matrix Ω, which gives an account of the correlations among uncertainty. Since we assume the experimental uncertainty to be normally distributed, minimising the generalised least squares loss function – that is the least squares with full account of correlations among fitted data introduced historically in Ref. [181] – also amounts to maximising the likelihood that the observed realisations result from a normal distribution. The generalised least squares loss function writes

$$\chi^2(A) = (F - CA)^T \, \Omega^{-1} \, (F - CA) \,. \tag{4.4}$$

$\chi^2(A)$ is a quadratic form with respect to A, and the optimal value of A to minimise $\chi^2(A)$, which we note $A^{(0)}$, is obtained by cancelling the gradient of the loss function :

$$\left(\frac{\partial\chi^2\left(A\right)}{\partial a_k}\right) = -2C^T\Omega^{-1}(F - CA)\,. \tag{4.5}$$

Provided the matrix $C^T\Omega^{-1}C$ is invertible, the general solution is obtained as [1]

$$A^{(0)} = (C^T\Omega^{-1}C)^{-1}C^T\Omega^{-1}F\,. \tag{4.6}$$

This expression is sufficient to derive the uncertainty on $A^{(0)}$. Direct calculation with the assumption of normal distribution of F gives that $A^{(0)}$ follows a multivariate normal distribution of covariance matrix $\Sigma = (C^T\Omega^{-1}C)^{-1}$. One will notice that the central value of F plays no role on the uncertainty of $A^{(0)}$, as a direct consequence of the linearity of our model in its parameters.

In the simple linear case that we are focusing on right now, $\chi^2(A)$ contains all the necessary information to recover the uncertainty on $A^{(0)}$ without the need for the closed form solution of the optimisation (4.6). Indeed, the Hessian matrix

$$\left(\frac{\partial^2\chi^2(A)}{\partial a_i\partial a_j}\right) = 2C^T\Omega^{-1}C = 2\Sigma^{-1} \tag{4.7}$$

is exactly equal to twice the inverse of the covariance matrix of the optimal parameter extraction. The **Hessian method** [182, 89] can be adapted to more complicated cases where a closed form solution relating $A^{(0)}$ to F is not available. Assuming the loss function behaves in the vicinity of $A^{(0)}$ in a quadratic way, the covariance matrix on $A^{(0)}$ will be approximated as

$$\left[\frac{1}{2}\frac{\partial^2\chi^2(A^{(0)})}{\partial a_i\partial a_j}\right]^{-1}\,. \tag{4.8}$$

As mentioned before, we recover the intuitive notion that if the likelihood is almost flat near its maximum – so with a small second derivative – the uncertainty on the parameter extraction is large. The Hessian method is still mostly satisfactory if experimental uncertainties are not normally distributed since the generalised least squares estimate is the best linear unbiased estimator [181]. A more striking limitation occurs for a model which is not linear in terms of its kinematic variables, for instance because the loss function might well exhibit several significant local minima. Such is particularly the case in parton distributions studies, since many models incorporate a dependence on the kinematic variable x as $(1-x)^\alpha x^\beta$ where α, β are parameters to fit, to account for the phenomenological behaviour of PDFs at end points. Furthermore, as mentioned in Ref. [142] for instance, it is frequent that the Hessian matrix is poorly conditioned, with eigenvalues varying in

1. Coming back to our simple example, one can verify easily that C is invertible, so $(C^T\Omega^{-1}C)^{-1}C^T\Omega^{-1} = C^{-1}\Omega C^{-T}C^T\Omega^{-1} = C^{-1} = \begin{pmatrix} -1/\tan(1) & 1/\sin(1) \\ 1 & 0 \end{pmatrix}$. This gives immediately the result of Eq. (4.3) when mutiplying by F. We observe in addition the general result that Ω plays no role in the determination of the best fit value if the matrix C is invertible. It corresponds indeed to having exactly the same number of free parameters and independent experimental constraints, so the best fit goes exactly through the central values of the measurements. However, the uncertainty on the best fit will of course still depend on the uncertainty of the measurements contained in Ω.

a very large range causing numerical difficulties. Finally, when fitting together datasets obtained by different experimental collaborations, it is not unlikely that the datasets are in tension with one another due to unevaluated systematic uncertainties. For this reason, uncertainties computed with the Hessian method are often increased by a factor 5 to 10, known as tolerance parameter T [183]. Recent fits in the field of PDF phenomenology have used adaptive values of the tolerance parameter for each eigenvector of the Hessian matrix, see Ref. [184] for a review of modern techniques in fits of PDFs.

The replica method

A more general method to evaluate aleatoric uncertainty is based on Monte-Carlo techniques, and consists in creating artificial experimental datasets by drawing new values of the measurements in their respective experimental distributions.The statistical and systematic uncertainties are taken into account with their correlations. For each artificial dataset, the optimal value of parameters is obtained to minimise the loss function. One obtains thus a sampling of the joint distribution of the model parameters, which is absolutely not constrained to remain close to a normal distribution, and can exhibit arbitrarily complex interdependence. Each fit on an artificial experimental dataset will be called a replica. The replicas form a sampling of the probability distribution of the quantity of interest in the functional space determined by the form of the model. In simple cases, the final uncertainty on the quantity at a given kinematic configuration is typically defined as the standard deviation computed over a large number of replicas, typically 100 or more decided on a case-by-case basis.

Fig. 4.1 shows an example that we have designed specifically to highlight how the replica method can provide results that would be out of reach for traditional methods which only study the behaviour of the loss function in the vicinity of its minimum. We give ourselves a simple, yet realistic model of PDF inspired from the traditional fitting method described in Ref. [184] :

$$x_B f(x_B) = a x_B^b (1 - x_B)^c (1 + d x_B)\,, \tag{4.9}$$

where a, b, c and d are free parameters. Let us then assume we have at our disposal an experimental dataset of eleven measurements. Ten of them are spaced regularly in logarithmic space between $x_B = 0.01$ and 0.4 with uncorrelated relative uncertainty of 15%. This corresponds to typical measurements in terms of kinematic coverage and experimental uncertainty for DVCS at JLab. One additional measurement is obtained at a much lower value of $x_B = 5e-4$ with uncorrelated relative uncertainty of 37%. The first ten measurements have central values that correspond exactly to model (4.9) for $a = 1.9$, $b = 0.5$, $c = 3$ and $d = 2$, in accordance with the general behaviour of u valence quark PDFs in the PDF4LHC15 NNLO fit at 10 GeV^2 [185]. However, the eleventh measurement clearly differs from these parameter values. We show on the left panel of Fig. 4.1 the result of the fit of model (4.9) on 100 different artificial datasets for the generalised least squares loss function (4.4). A remarkable feature is the separation of the replica bundle in two clusters at small x_B of approximately equal size. This branching is particularly obvious in the distribution of parameter b depicted on the right panel of the figure. Depending on the actual value drawn for the last isolated measurement in each artificial dataset, the strain it exerts on the overall fit is either dominant and drives the entire small x_B behaviour of

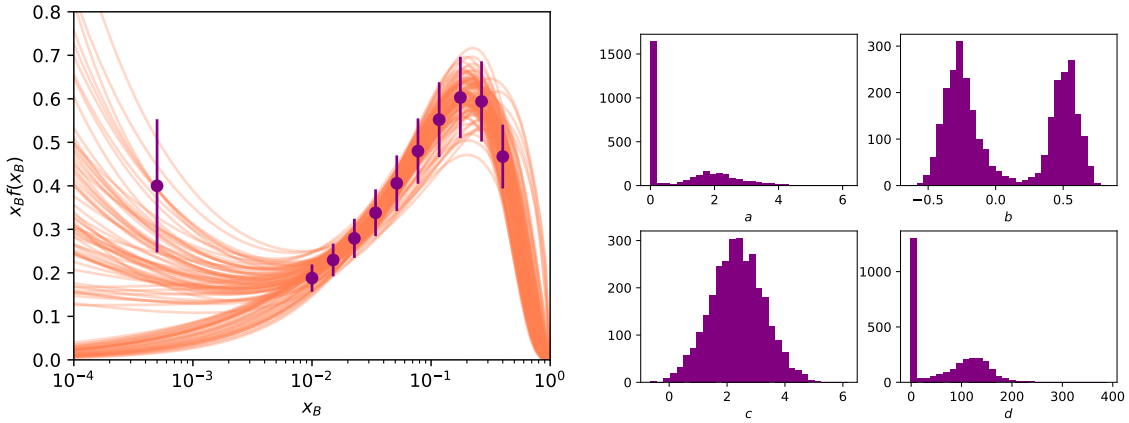

Figure 4.1 – (left) The purple points represent the experimental dataset with its uncertainty. The orange curves are 100 replicas fitted with model (4.9) on artificial datasets drawn according to the experimental uncertainty. – (right) We represent the distribution of values of the parameters of model (4.9) for 3000 replicas. It is especially clear on parameter b (upper right), which drives the behaviour of the model in the small x_B limit, that the replica bundle is separated in two clusters approximately equally populated.

the replica, or is insufficient to significantly alter the tendency entailed by the first ten measurements.

Since the distribution of parameters diverges significantly from a normal distribution to the point that it has several modes, the traditional Hessian method, which focuses on the study of χ^2 around its global minimum, will fail to give a precise account of uncertainty. If the Hessian method was applied unceremoniously to such a case, it would simply hide the nature of the problem and underestimate the total uncertainty. However, the χ^2 contains information on this separation in clusters, under the form of local minima. Since the χ^2 is a function of four parameters a, b, c and d in this example, it is unpractical to visualise the overall shape of the function. We use a dimensionality reduction tool, the t-distributed stochastic neighbor embedding (t-SNE) algorithm [186], to bring the study back to a more manageable two dimensions. As a consequence, the four dimensional space of parameters is embedded in a two dimensional representation via a non-linear transformation. Although non-linearity means that it is difficult to interpret the meaning of the final space with t-SNE, the objective is to preserve clusters, so that points that are close to each other in higher dimension remain close through dimensionality reduction [2]. We will use this tool to identify the regions corresponding to significant local minima in the space of parameters (a, b, c, d) and try to understand what kind of information on the branching of the probability distributions of the parameters the χ^2 contains.

In practice, on the left panel of Fig. 4.2, we first produce a four-dimensional grid spanning a large region of parameters which we expect to encompass all significant local minima of the χ^2. Out of 160000 points of the grid, we only select 1055 such that the

2. More precisely, given points in a high-dimensional space, the algorithm constructs a probability distribution on pairs of these points such that the probability associated to a pair of points which are close in high dimension is much larger than that associated to a pair of distant points. Then another distribution is built on pairs of points in the low dimensional space, and the algorithm minimises the Kullback-Leibler divergence between the two probability distributions by means of an iterative gradient-descent. We refer the reader to Ref. [186] for more details on this popular dimensionality reduction tool.

associated χ^2 is less than 30. This amounts to selecting a few separate regions of the four dimensional grid which contain the strongest local minima of the χ^2. The chosen parameters are the result of a series of trials, as is necessary when studying a complex function with many local minima and no clear theoretical property. The threshold on χ^2 chosen at 30 guarantees that we only consider parameters that at least remotely reproduce the fitted data, while we checked that increasing the region spanned by the grid does not reveal new local minima. Then we apply to the 1055 selected points in four dimension the t-SNE algorithm to visualise their characteristics in two dimensions. As seen on the left panel, it is possible to identify broadly four regions where notably small χ^2 are encountered. We denote them by 1 to 4, and highlight by a star the position of the local minimum in each region. The star in the first region represents the global minimum of χ^2. The parameters associated to each star are summarised in the following table :

Parameters	a	b	c	d	Value of χ^2
Global minimum (region 1)	0.0089	-0.38	2.3	300	2.3
Local minimum (region 2)	0.70	0.28	3.4	11	5.7
Local minimum (region 3)	2.6	0.58	2.3	0.0	6.4
Local minimum (region 4)	1.1	0.36	2.3	2.9	7.1

If one remembers that the first ten points in the fitted dataset were generated with $a = 1.9$, $b = 0.5$, $c = 3$ and $d = 2$, the global minimum singles out itself as being particularly remote from these values, unlike the other three local minima. This is visible on the right panel of Fig. 4.2 where we plot the curves associated to each minimum. We observe the same behaviour of branching of the parameter distributions at small x_B as the one we put forth thanks to the replica method in Fig. 4.1. This demonstrates that a study of the χ^2 not limited to the sole global minimum contains significant information for a complicated uncertainty extraction without the need of the replica method. Since the latter requires repeated global minimisations of the χ^2 for various artificially generated experimental dataset, it can become very computationally intensive. It is in particular the case if the model is complex and strongly non-linear – as neural networks can be – and if the loss function behaves poorly. It may then become **more interesting from a computational point of view to focus on the extraction of a few local minima of the χ^2 rather than repeatedly looking for its global minimum for varying experimental data**. We found that it was in general conceptually simpler and still manageable in terms of computing power to estimate our uncertainties thanks to the replica method. However, our neural networks models in Chapter 7 where we implement a simplified positivity constraints take of the order of a full day of computation to be trained. Should the computation time keep increasing, a more sophisticated approach in the fashion we have described here might prove necessary.

We have mostly concerned ourselves so far with aleatoric uncertainty emerging from experiment, neglecting epistemic uncertainty. However, in the absence of strong theoretical motivations guiding the choice of a model, epistemic uncertainty comes at the forefront. Fitting collaborations in the field of PDFs have increasingly relied on data-driven approaches allowed by neural network modelling [187, 188, 189, 190, 191]. As we will see, neural network architectures are very flexible, reducing the amount of prior knowledge introduced in the modelling and therefore the epistemic uncertainty. Additionally, there are practical tools specific to neural networks to assess to some extend epistemic uncertainty.

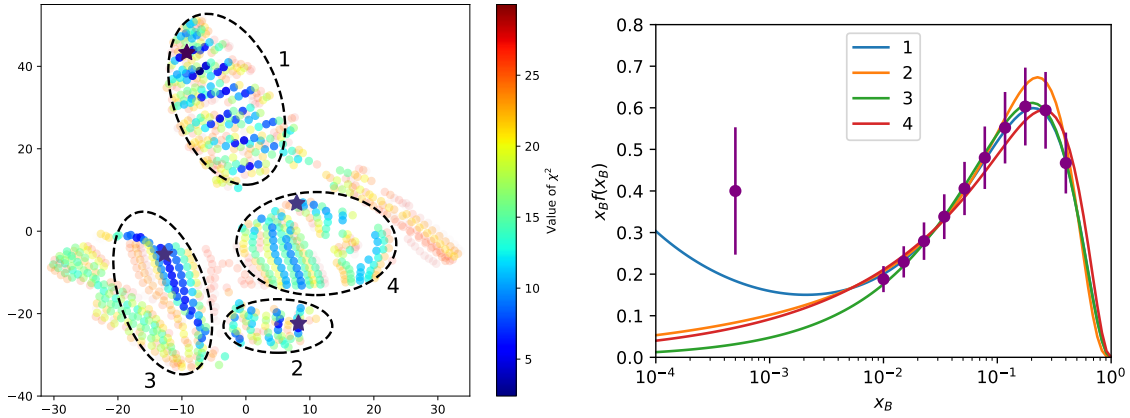

Figure 4.2 – (left) Embedding in a two dimensional space of the regions of the four dimensional parameter space where $\chi^2 < 30$. Four clusters appear. In each cluster, the local minimum of χ^2 is depicted by a star. The star of cluster 1 is the global minimum of χ^2. – (right) Representation of the model corresponding to each star. The best fitting model, model 1, is substantially different from the other three in its small x_B behaviour.

4.1.2 . Neural networks

The general idea of artificial neural networks can be traced back at least to 1957 with Rosenblatt's perceptron [192] – a single neuron whose output is 0 if the weighted sum of its input lies below a threshold, and 1 otherwise. Development of **back-propagation** algorithms in the 1970s and 1980s (see for instance Ref. [193] where the term back-propagation is coined, although the general principles of the technique predates the paper by about two decades) extended the applicability of the concept by allowing to train multilayer perceptrons on non linearly separable datasets (that is datasets whose convex hulls are disjoint). Both back-propagation and multilayer perceptrons will be detailed in the following. Overwhelming interest in neural networks in recent years is linked to deep learning, the training of neural networks with a large number of layers and free parameters on very large datasets. For instance, the deep convolutional neural network AlexNet [194] which revolutionised the field of image recognition in 2012 possesses more than 62 million free parameters, and the ImageNet database contains more than 14 million annotated images. Deep learning was only made possible recently, thanks to the availability of computing power provided by graphics processing units (GPUs) and the emergence of large datasets in the age of social media, although grounding theoretical works on convolutional neural networks dates back to the 1980s for instance [195]. The enthusiasm for the field has led to a proliferation of architectures and training procedures, and it is out of the scope of this document to give a general overview of the subject.

Architecture of multilayer perceptrons

For the purpose of our study, we only used the simple architecture of the multilayer perceptron (MLP) which we detail now. An MLP can be considered as a complicated function taking n features as input, and producing m values as output. Fig. 4.3 shows a typical MLP with $n = 3$ and $m = 2$.

The information is processed sequentially layer after layer. The internal layers of the network – that is neither the input nor the output layers – are called **hidden** layers. Let

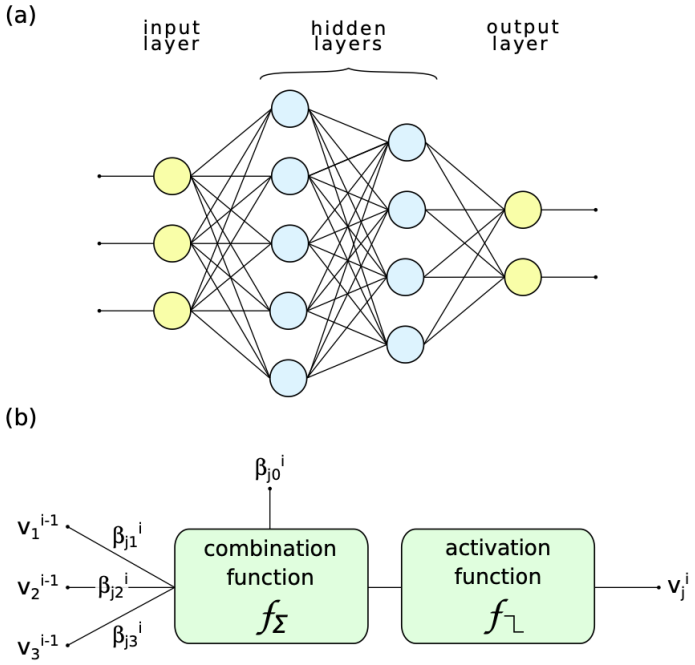

Figure 4.3 – (a) A multilayer perceptron with two hidden layers. – (b) General principle of an individual neuron. Figure taken from Ref. [143].

us note v_k^i the output of the k-th neuron in the i-th layer. The first layer simply distributes the input features to the next layer, so v_k^1 are the input features. v_k^{i-1} is distributed to the j-th neuron of layer i via a connection of **weight** $\beta_{j,k}^i$. Each neuron computes the weighted sum of its inputs, adds a specific **bias** $\beta_{j,0}^i$ and feeds the result to an **activation function** σ^i. The image through the activation function determines the output v_j^i. The operation of one neuron can be summarised as

$$v_j^i = \sigma^i \left(\beta_{j,0}^i + \sum_k \beta_{j,k}^i \, v_k^{i-1} \right) . \tag{4.10}$$

The successive layers of the network act as nested compositions of activation functions and weighted summations. The free parameters of the network are the weights and biases $\beta_{j,k}^i$ which are optimised during a training procedure. The number of layers, neurons in each layers and the choice of activation functions σ are usually considered as modelling hypotheses, although some advanced training procedures allow one to adapt these choices during the training.

Two main advantages justify the use of neural networks to model the complex relation existing between kinematic variables and interesting quantities like CFFs or parton distributions themselves. First, they provide abstract parameterisations that are not driven by a strong physical *a priori* on the relation. In this sense, the model is essentially data-driven, and not theory-driven. However this can only work if the model is flexible enough so that the specific architecture choices do not practically act as a strong prior bias. The flexibility is the second advantage of this modelling. It can be expressed mathematically by universal approximation theorems, demonstrated for their first variant in Ref. [196] and notably extended in Ref. [197]. Let us rephrase the universal approximation theorem

for arbitrary width under a form which is better suited for our presentation :

> **Universal approximation theorem for MLPs of arbitrary width (Hornik, [197])** For every $(n, m) \in \mathbb{N}^2$, compact K of $\mathbb{R}^n$, continuous function f from K to $\mathbb{R}^m$ and $\epsilon > 0$, there exists $k \in \mathbb{N}$ and an MLP with exactly one hidden layer of k neurons and a linear activation function in its output layer whose output $g(x)$ verifies
>
> $$\sup_{x \in K} ||g(x) - f(x)|| < \epsilon \,, \tag{4.11}$$
>
> *if and only if* the activation function of the hidden layer is not polynomial.

It is interesting to observe that the key element to the universal approximation theorem is the non polynomiality of the activation function in the hidden layer. Traditional activation functions are the sigmoid $x \mapsto 1/(1 + \exp(-\lambda x))$ and the rectified linear unit (ReLU) $x \mapsto x\Theta(x)$ where $\Theta(x)$ designates the Heaviside step function.

Training

Once the architecture of the network has been defined, the determination of the free parameters is performed through a training procedure. For our modelling purposes, we used **supervised learning**, since we know the expected output that the neural network should produce for a given set of kinematic variables. Several types of supervised learning procedures are available. The most frequent one consists in defining a loss function to measure the error between the current output of the network and its expected output. Then the gradient of the loss function depending on the free parameters of the network is computed. The gradient calculation can be optimised by taking into account the fact that the network consists in a nested composition of activation functions. Thanks to the derivation chain rule, the gradient is computed iteratively layer after layer starting from the final layer in a procedure known as back-propagation. Once the gradient is computed, the parameters are modified in its opposite direction so as to decrease the value of the loss function. Back-propagation is used to obtain a pre-training of the neural network modelling of CFFs described in the next section. This very basic minimisation strategy faces the risk of converging to a local minimum of the loss function. To prevent it, it is customary to split the training dataset into several batches, and to iteratively apply gradient descent while regularly changing the training batch. Another issue which prevented for a long time the training of very deep neural networks is the vanishing or exploding gradient problem (see for instance Ref. [198]) : several activation functions, like the sigmoid $x \mapsto 1/(1 + \exp(-x))$, have derivatives which are systematically much lower than 1. Then the derivation chain rule shows that the gradient becomes exponentially suppressed for parameters in the initial layers compared to the final layers, due to repeated products of numbers smaller than 1. Likewise, an exploding gradient problem can occur in situations where all gradients would be much larger than 1.

A popular alternative to gradient descent is provided by **genetic algorithms** [199], which are particularly suited for exploring spaces with no clear underlying structure and many potential local minima. We will mostly use genetic algorithms for the training of our neural network models. They are also particularly suited to implement conditions on the

output of the network that do not let themselves write easily in terms of χ^2 loss functions. It is the case for instance in Section 7.2 where we will enforce positivity conditions on the output of a neural network model of GPDs published in Ref. [8]. We initially build a number of candidates, which are randomly generated configurations of free parameters. At each step of the training, called a training **epoch**, a given fraction of the best candidates with respect to the loss function are selected. The winners of this selection form the group A. The candidates that are not selected are discarded, and new candidates are generated to maintain their total number at a constant level by two means : **cross-over** (group B) and **mutation** (group C). The cross-over procedure consists in building new candidates by mixing features of candidates of group A. The mutation procedure builds new candidates by randomly varying some features of candidates of group A. The analogy with biological evolution justifies the name of this algorithm. The selection of group A is an avatar of natural selection of the fittest individuals in a population. They are given the opportunity to reproduce, therefore mixing their traits and producing descendants in group B. They also suffer random mutations of their genes, producing group C. Mutation is key to allow the exploration of the space of parameters, but must not be too frequent at the risk of transforming the genetic algorithm into a simple random search. Cross-over is crucial to the speed of convergence, by allowing successful parameters to be propagated quickly to a large number of candidates. Genetic algorithms are heuristic optimisation procedures whose precise parameters are often a matter of empirical trials.

Regularisation

Regardless of the actual training procedure, it is important to try to prevent over-fitting, that is an excessive sensitivity of the trained parameters to anecdotical features of fitted data. When confronted with new data, the model is then notably less precise, even in the range of training data. Several techniques are used to prevent over-fitting, known as **regularisation methods**. The most frequent is to optimise the network only on a fraction of available data known as **training set**, usually of the order of 80% of all data. The remaining 20% of the dataset is known as **validation set**. Training is continued as long as the goodness of fit measured by the loss function on the validation set seems to increase, but the error on the validation set is never used directly in the optimisation procedure.

An alternative to prevent excessive sensitivity of the network to its training data is to disturb it regularly by de-activating randomly small parts of the network during training. During their de-activation phase, the values of the frozen parameters are put to zero, and their previous values restored once they are re-activated. This method, known as **drop-out** [200, 201], amounts in practice to training in parallel many different architectures of the network. Once training is satisfactory, drop-out can still be applied during standard use of the model. The different architectures probed during evaluation give different output values, producing an account of epistemic uncertainty. If all neurons are activated during evaluation, the model outputs an superposition of all the different architectures it was trained upon, increasing the robustness of its predictions. We have notably used drop-out to regularise our neural network model of GPDs in Section 7.2 [8].

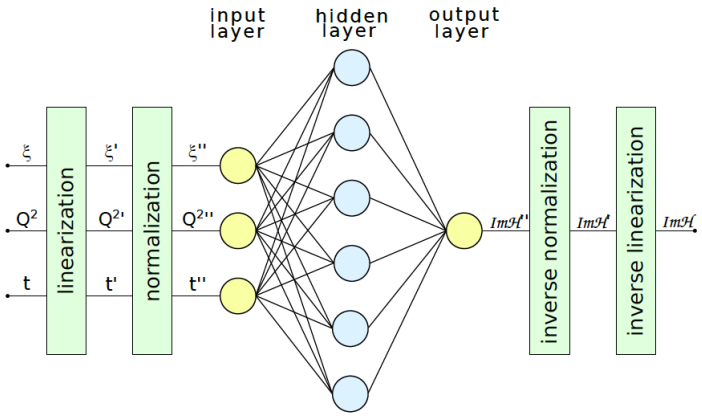

Figure 4.4 – Architecture of the neural network modelling the real or imaginary part of a Compton form factor. Figure taken from Ref. [143].

4.2 . State of world DVCS data in neural network analyses

The first use of neural network modelling in the field of generalised parton distributions was performed in 2011 by Kumerički, Müller and Schäfer in Ref. [142]. They fitted the CFF $\mathcal{H}$ on HERMES data for DVCS on unpolarised protons. In 2019, Moutarde, Sznajder and Wagner fitted most of the world DVCS data, that is 2624 x_B, t, Q^2, ϕ) data points collected over 17 years from about 30 observables at 6 experimental facilities in Ref. [143]. This study led prior to the contributions of this thesis is however the basis for most of the subsequent statistical analyses and deserves therefore a presentation. Other neural network parameterisations of DVCS data have been performed in Refs. [202, 203].

Out of the twelve CFFs entering the description of DVCS, only $\mathcal{H}$, $\mathcal{E}$, $\tilde{\mathcal{H}}$ and $\tilde{\mathcal{E}}$ associated to the four leading twist chiral-even GPDs H, E, $\tilde{H}$ and $\tilde{E}$ are taken into account, and the eight real and imaginary parts of these CFFs are modelled independently by eight neural networks such as the one depicted in Fig. 4.4. Each neural network takes as inputs the three kinematic variables (ξ, Q^2, t) which have previously undergone a specific pre-treatment detailed below. The output is a single real number which also undergoes a post-treatment. The hidden layer consists of 6 neurons with sigmoid activation functions. This number was determined by observing that it gave a satisfactory goodness of fit on a benchmark with Goloskokov and Kroll's (GK) model.

The pre-treatment consists in what Fig. 4.4 calls linearisation [3], that is taking the logarithm of ξ and Q^2, and a normalisation intended to linearly project input values in the $[-1, 1]$ interval. The post-treatment consists of an inverse normalisation to increase the range of values, and a so-called inverse linearisation which divides the output by ξ. The pre- and post-treatments considerably increase performance, by regularising the behaviour of the CFFs at small ξ and bringing input values in the range where the sigmoid function is most sensitive to avoid inputs that directly saturate the network.

Overall, the eight networks representing the real and imaginary parts of the four CFFs have 248 free parameters. The training is performed in two steps. First, a local extraction of CFFs is performed for each kinematic bin. The neural network is trained on

3. It is a linearisation in the sense that it attempts to spread linearly the kinematic values of the data points to facilitate the work of the neural network.

this local CFF extraction with back-propagation of a simple least squares loss function. This can be considered as a pre-training to initially adjust the parameters of the network. Then proper training is performed on real experimental measurements of observables via a genetic algorithm with a population of 1000 candidates. Since observables write as complicated non-linear functions of CFFs, the loss function is not a straightforward least squares between the output of the network and an expected value, but requires the intermediate step of evaluating observables from the output of the network. The use of a genetic algorithm allows in particular to evade the issue of computing the gradient of this complicated loss function.

The left panel of Fig. 4.5 shows the result of the fit for 101 replicas of the imaginary part of the CFF $\mathcal{H}$ as a function of ξ for some value of t and Q^2. It is customary to summarise this information under the form of confidence intervals. This is only a partial representation of the amount of information actually contained in the replicas – for instance, all replicas that go down at large ξ could go up at small ξ and vice versa, and confidence intervals would not reflect it – but it gives an account of the local uncertainty on the quantity of interest. On the right panel of Fig. 4.5, we show in light blue the confidence interval obtained naively as [mean - standard deviation, mean + standard deviation] for the distribution of replicas at each value of ξ. Under the hypothesis of a normal distribution of replicas, it should amount to the 68% confidence interval. However, this interval is extremely sensitive to the presence of outliers.

Several techniques are known to increase the robustness of confidence interval extractions. One of them is called **three sigma rule** [204]. Separately for each kinematic value, the replicas distant from the mean of the replica distribution by more than three standard deviations are considered as outliers and discarded. Mind that the replica might be discarded as outlier in some kinematic regions, but kept in others. The procedure is repeated with the remaining replicas and their newly computed mean and standard deviation until no outliers are suppressed. The band at the considered kinematic value is then again defined by the interval [mean - standard deviation, mean + standard deviation].

An alternative method which does not require the identification and removal of outliers is known as **median absolute deviation** (MAD) and traces back at least to Gauss (see for instance Ref. [205] for a more recent take on the subject). The central value of the interval is given by the median of the replica distribution, which we note median(X_i), and its spread depends on the median of the absolute distance of replicas to median(X_i) :

$$\frac{\text{median}\left(\left|X_i - \text{median}(X_i)\right|\right)}{\Phi^{-1}(3/4)}\,, \tag{4.12}$$

where Φ is the cumulative distribution function of a standard normal distribution, Φ^{-1} designates the reciprocal – and not the inverse – and $\Phi^{-1}(3/4) \approx 0.67$. The use this normalisation factor is an implicit assumption that the replica distribution is normally distributed. The estimator can be applied to other assumed distributions by choosing modifying the normalisation factor. Because the estimator uses the median and absolute distance, unlike the mean and squared distance as for the standard deviation, it is naturally less affected by outliers. We demonstrate on the right panel of Fig. 4.5 that the three sigma rule and the MAD give very comparable results, except for the very large ξ region where the three sigma rule gives sensibly larger uncertainties. Both robust estimators of

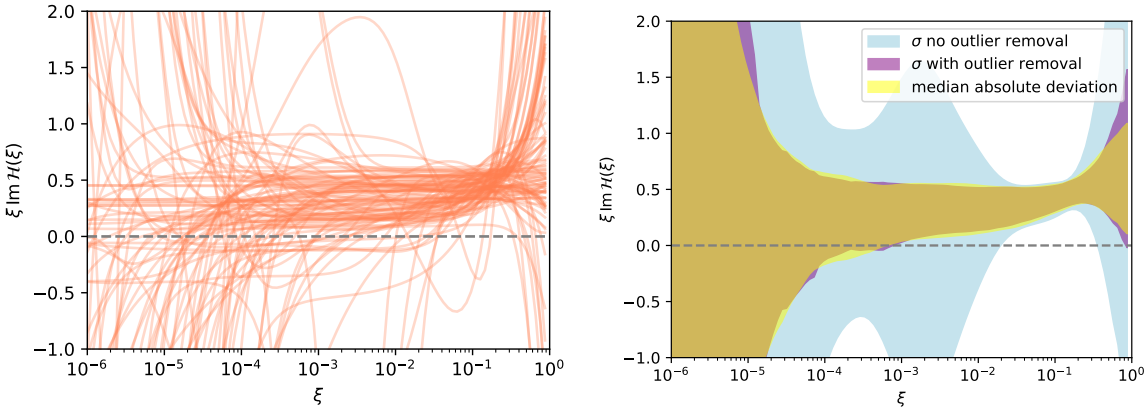

Figure 4.5 – (left) 101 replicas of $\xi \mathrm{Im}\,\mathcal{H}(\xi, t = -0.3\ \mathrm{GeV}^2, Q^2 = 2\ \mathrm{GeV}^2)$ – (right) Confidence intervals derived with the naive method [mean - standard deviation, mean + standard deviation] (blue band), the three sigma rule where outliers have been suppressed independently at each value of ξ before the same interval as before is computed (purple band) and the mean absolute deviation estimator (yellow band).

uncertainty are considerably tighter than the naive version except in the region where the CFF are the most constrained, around $\xi = 0.1$. It is therefore paramount to use robust estimators. In the following, we use the three sigma rule unless specified otherwise.

Fig. 4.6 shows the result of the fit for both the real and imaginary parts of the CFF $\mathcal{H}$ (grey band), and compares it to an extraction on an almost similar experimental dataset, but with a model depending on a small number of parameters described in Ref. [53] (blue band). The data-driven flexible parameterisation produces much larger uncertainties than the contrained parameterisation. Indeed, neural networks bring a much lesser model bias, so the variance of the model is much larger as a traditional case of variance-bias tradeoff. Said otherwise, the epistemic uncertainty associated with the constrained parameterisation, which is not accounted for in the blue band, is much larger than the one of the neural network model.

In some kinematic regions, particularly for ξ smaller than 10^{-4} and larger than 0.5, the uncertainty of the neural network extraction is very large due to the scarcity of available experimental data. Future experiments are expected to bring a much desired experimental precision in this region. We have worked on a proposal for a positron beam at JLab which could also bring a complementary knowledge on CFFs compared to the one showed in Ref. [143]. To assess the impact of these experiments, we could simulate plausible experimental results and train again the neural networks. However, the training is quite lengthy and computationally expensive. It is therefore interesting to be able to quantify broadly but reliably the impact of the addition of new experimental data on a previously completed fit. We develop interesting aspects surrounding this question in the following section.

4.3 . Assessing the impact of new measurements

Following the general structure of Section 4.1, we start by presenting the result of the effect of adding a single new point to a linear fit with normally distributed uncertainties.

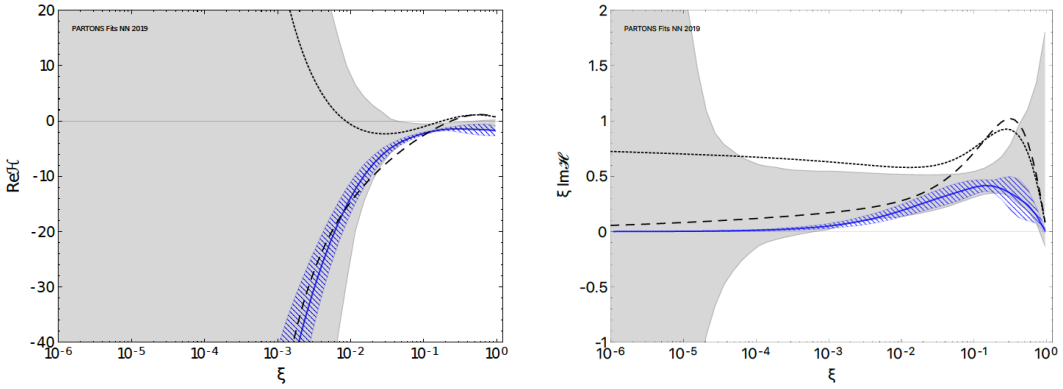

FIGURE 4.6 – Real (left) and imaginary (right) parts of the CFF $\mathcal{H}$ as a function of ξ for $t = -0.3$ GeV^2 and $Q^2 = 2$ GeV^2. The blue band denotes the result of a fit on an almost similar dataset, but with a model with a small number of parameters [53]. The grey band is the data-driven extraction produced by the neural networks. The dotted lines are the CFFs obtained from two phenomenological GPD models, GK and VGG. Figure taken from Ref. [143].

We briefly study the possible generalisations of the linear results to non-linear models. Under some criteria that we establish, we provide therefore a fast qualitative estimate of the constraining power of the addition of a new measure on the knowledge of a quantity of interest. We introduce then the general method of Bayesian reweighting suited for non-linear cases and uncertainties evaluated through the replica method. This paves the way to the precise discussion of the use of the reweighting technique in our impact study related to a positron beam at JLab.

4.3.1 . Adding a new point to a linear fit

Using the notations of Section 4.1.1 on the linear fit with the generalised least squares loss function, we denote by $f_A(X) = X^T A$ (4.1) a linear model evaluated on the kinematic configuration X with parameters A. When the model is fitted on experimental data with uncertainty, A becomes a random variable, and we note $\langle f_A(X) \rangle$ the central value of the fitted model at X and $\Delta f_A(X)$ its standard deviation giving an account of the aleatoric uncertainty. Let us assume that we have already performed a fit on experimental data and obtained a vector of best fit parameters $A^{(0)}$ (4.6).

We now add to the current fit a new uncorrelated datapoint obtained at the kinematic configuration $X^{(1)}$ with central value f_1 and normally distributed experimental uncertainty σ_1. The new vector of best fit parameters is noted $A^{(1)}$. We demonstrate in Appendix A.1 that a simple relation relates the central value and standard deviation of the initial fit at $X^{(1)}$ – that is $\langle f_{A^{(0)}}(X^{(1)}) \rangle$ and $\Delta f_{A^{(0)}}(X^{(1)})$ – to the same quantities for the new fit – $\langle f_{A^{(1)}}(X^{(1)}) \rangle$ and $\Delta f_{A^{(1)}}(X^{(1)})$. We are not aware of any other derivation of this result, which gives an interesting insight on the behaviour of linear fits when confronted to new data. We partially presented this relation in Ref. [206], although it was only obtained for models with a polynomial dependence of order 3 in a unique kinematic variable, whereas the new result is fully general.

Introducing

$$\lambda = \left[1 + \frac{(\Delta f_{A^{(0)}}(X^{(1)}))^2}{\sigma_1^2} \right]^{-1} , \tag{4.13}$$

we demonstrate that

$$\begin{cases} \langle f_{A^{(1)}}(X^{(1)})\rangle = \lambda\langle f_{A^{(0)}}(X^{(1)})\rangle + (1-\lambda) f_1\,, \\ (\Delta f_{A^{(1)}}(X^{(1)}))^2 = \lambda(\Delta f_{A^{(0)}}(X^{(1)}))^2\,. \end{cases} \tag{4.14}$$

It is therefore possible to predict the effect at $X = X^{(1)}$ of adding a new uncorrelated point to any linear fit thanks only to the knowledge of the current central value and uncertainty of the model at that point and the central value and uncertainty of the added point. Practically, it means that the final model distribution at $X = X^{(1)}$ can be **predicted graphically without any knowledge of the actual model that is fitted, except that it is linear !** However, the effect of the newly added point for $X \neq X^{(1)}$ can be arbitrarily complicated, and will depend on the actual model.

Considering the generality of this result, we would like to estimate its validity in the non-linear case, for instance for the PDF model presented in Eq. (4.9). Fig. 4.7 allows us to compare the actual distributions computed via the replica method to the predictions of the linear estimate (4.14). Successively at each point of the plane, we assume that a new data point is added to the original dataset of eleven measurements, with always the same uncertainty $\sigma_1 = 0.1$. The colour at that point shows the discrepancy between the linear estimate and the actual replica result. The plot allows therefore to identify regions where the linear estimate is satisfactory, and where it fails. The results for the mean of the distribution on the left panel are excellent in the strongly experimentally constrained region where the dense ten data points are located. There, the linear estimate of the mean agrees with the replica method at the percent level. Agreement quickly fades away at small and very large x_B, except in a narrow region approximately in the middle of the initial replica bundle. The results are qualitatively the same for the estimation of the standard deviation on the right panel, although the agreement is overall worse. We led in Ref. [206] several other tests and concluded empirically that the linear estimate gives satisfactory results when the replica bundle after inclusion of the new data point is approximately normally distributed at the new kinematic. The situation is notably met if (1) the new data point is considerably more precise than the previous knowledge at that kinematic, so the post-fit distribution of the replicas is close to the uncertainty of the new data point itself, (2) the replicas are already strongly constrained by previous measurements in the region and have already approximately a normal distribution, or (3) the new point is added close to the maximum of the previous replica distribution. When one of these criteria is fulfilled, the estimate (4.14) allows one to quickly measure the physical interest of probing this region to better constrain the quantity of interest.

Although the local impact of the addition of a new experimental measurement can be empirically assessed in the vicinity of the new kinematic provided one of the previously listed conditions is met, the linear estimate does not allow one to predict long range effects on the fit. Replicas not only allow the calculation of the local uncertainty at a given kinematic, but they also form a sampling of the probability distribution of the quantity of interest in the functional space determined by the form of the model. Therefore, they contain information on the long range dependence of the fit that we would like to use in the assessment of the impact of a new measurement.

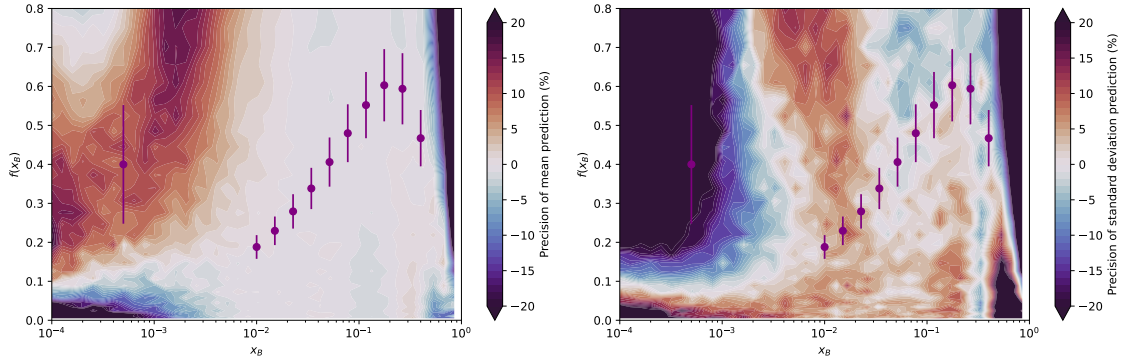

Figure 4.7 – Comparison between the prediction of the linear estimate (4.14) of the mean (left) and standard deviation (right) of model (4.9), and the actual values obtained by the replica method. The results are given as percentage of difference with the result of the replica method as reference. The purple points represent the original dataset of eleven measurements. At each point of plane, we successively simulate the addition of a new data point with uncertainty $\sigma_1 = 0.1$. For the left panel, if the linear prediction of the mean after inclusion of the new point is larger than the actual value given by replicas, the point is coloured red, and blue if the linear prediction is smaller than the actual value. Same on the right panel for the standard deviation.

4.3.2 . Bayesian reweighting

Given a set of replicas, Bayesian reweighting assesses the impact of a new dataset by attributing a weight to each replica measuring its compatibility with the new measurements. Strongly incompatible replicas receive a tiny weight, which amounts in effect to discarding them. The demonstration of the validity of this principle and the formula for the weights can be found for instance in Ref. [207]. For the k-th replica, the reweighting coefficient ω_k is defined by

$$w_k = \frac{1}{Z}\,(\chi^2_k)^{(N-1)/2}\exp(-\chi^2_k/2)\,, \tag{4.15}$$

where Z is a normalisation constant determined after all w_k are computed so that $\sum_k w_k = 1$, N stands for the number of points in the new dataset, and χ^2_k is the generalised least squares loss function measuring the goodness of fit of the k-th replica with the new dataset :

$$\chi^2_k = (Y - Y_k)^T \Omega^{-1} (Y - Y_k)\,, \tag{4.16}$$

where Y is the vector of central values of the new measurements, Ω its covariance matrix and Y_k the vector of values of the k-th replica at the kinematic configurations of the new dataset. For Bayesian reweighting to give sensible results, the initial set of replicas should be dense enough in the region where the new dataset is added, so that many replicas are roughly compatible with the new measurements. Otherwise, due to the sharp decrease of Eq. (4.15) as χ^2 increases, Bayesian reweighting quickly ends up solely selecting the least incompatible replica, and attributes it a weight $w \approx 1 - \epsilon$. To prevent this issue and keep track of the statistical relevance of the procedure, it is possible to compute the Shannon entropy [208] of the weights and use it to define an effective number of replicas post-reweighting by

$$N_{\text{eff}} = \exp\left(-\sum_{k=1}^{N_{\text{rep}}} w_k \log(w_k)\right)\,, \tag{4.17}$$

where N_{rep} is the number of replicas. N_{eff} can be interpreted as the number of replicas that still carry information on the distribution in functional space after the new measurements are included. The constraining power of the new dataset is measured by the ratio between N_{rep} and N_{eff}.

After Bayesian weights have been computed according to Eq. (4.15), it is possible to characterise the weighted distribution by its weighted mean and unbiased variance estimators, defined by [209]

$$\langle f(X)\rangle_{\text{reweighted}} = \sum_{k=1}^{N_{\text{rep}}} w_k f_k(X)\,, \tag{4.18}$$

$$(\Delta f(X)_{\text{unbiased}})^2 = \frac{1}{1-\sum_{k=1}^{N_{\text{rep}}} w_k^2} \sum_{k=1}^{N_{\text{rep}}} w_k (f_k(X) - \langle f(X)\rangle_{\text{reweighted}})^2\,, \tag{4.19}$$

where $f_k(X)$ is the value of the k-th replica at kinematic X, and N_{rep} is the number of replicas. The initial factor in the unbiased variance estimator corrects the bias due to the fact that the weighted mean estimator is correlated to the data samples. The situation is more familiar in the unweighted case, where the uncorrected variance estimator

$$\frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} (f_k(X) - \langle f(X)\rangle)^2 \tag{4.20}$$

is biased and systematically underestimates the actual variance by a factor $(1 - 1/N_{\text{rep}})$. Bessel's correction

$$\frac{1}{N_{\text{rep}}-1} \sum_{k=1}^{N_{\text{rep}}} (f_k(X) - \langle f(X)\rangle)^2 \tag{4.21}$$

produces an unbiased estimator of the variance. The difference is only significant if N_{rep} or N_{eff} are small respectively in the unweighted and weighted cases. It is interesting to notice that as N_{eff} is a real number unlike N_{rep}, we can study the limit $N_{\text{eff}} \to 1$ for a configuration of weights w_k. For instance, if the weights favour overwhelmingly the first replica and discard equally all others, that is $w_1 = 1 - (N_{\text{rep}} - 1)\epsilon$ and $w_k = \epsilon$ for $2 \leq k \leq N_{\text{rep}}$, the unbiased variance estimator converges when $\epsilon \to 0$ to

$$\frac{1}{2(N_{\text{rep}}-1)} \sum_{k=2}^{N_{\text{rep}}} (f_k(X) - f_1(X))^2\,, \tag{4.22}$$

which is half of the unbiased variance estimator in the unweighted case with $f_1(X)$ as empirical mean. This shows how even for an extremely constraining new dataset, the prior knowledge represented by the unweighted distribution of replicas is still driving significantly the final result.

Finally, taking the square-root of an unbiased variance estimator introduces a new layer of bias because of the non-linearity of the square-root function. Therefore, the standard deviation estimator derived from Eq. (4.19) is biased by a slight systematic underestimation, but there is no simple way to correct this bias systematically that is independent of the precise distribution of the variance estimator.

4.4 . Impact study of a positron beam at JLab

Bayesian reweighting techniques are commonly used in the field of PDF phenomenology (see for instance Ref. [210]) to assess the impact of new data points, including lattice QCD inputs on PDF fits [211]. To the best of our knowledge, the impact study led in Ref. [2] is the first where that technique was applied to a neural network extraction of CFFs.

So far, most of the knowledge on CFFs has been obtained from DVCS with a beam of negatively charged leptons. Only HERMES and COMPASS experiments allowed to collect some observables with positively charged beams (see experimental review of DVCS datasets in Section 3.3). The possibility to change experimental settings like the charge or polarisation of colliding particles is particularly interesting since it allows one to separate contributions of various terms in the photon lepto-production cross-section, and therefore provides different sensitivity to various CFFs (see Section 3.1). For instance, the unpolarised beam charge asymmetry (BCA) is defined as

$$A_C(x_B, t, Q^2, \phi) = \frac{\mathrm{d}^4\sigma^+(x_B, t, Q^2, \phi) - \mathrm{d}^4\sigma^-(x_B, t, Q^2, \phi)}{\mathrm{d}^4\sigma^+(x_B, t, Q^2, \phi) + \mathrm{d}^4\sigma^-(x_B, t, Q^2, \phi)}\,, \tag{4.23}$$

where $\mathrm{d}^4\sigma^\pm$ denotes the differential cross-section of lepto-production of a single photon averaged over the polarisations of both the lepton and hadron for either positively or negatively charged beam particles. The BCA is particularly sensitive to the real part of the CFF $\mathcal{H}$ as presented in Ref. [129], via for instance its $\cos\phi$ Fourier harmonics at leading twist and leading order under the assumption of dominance of the BH-DVCS interference term over the pure DVCS one :

$$A_C^{\cos\phi} \propto \mathrm{Re}\left[F_1\mathcal{H} + \xi(F_1 + F_2)\tilde{\mathcal{H}} - \frac{t}{4M^2}F_2\mathcal{E}\right]\,. \tag{4.24}$$

BCA is therefore considered as a particularly constraining measurement for $\mathrm{Re}\,\mathcal{H}$, which as we will see in the next chapter is currently the source of a large uncertainty in the extraction of proton mechanical properties.

The possibility of operating a positron beam at JLab [212, 3] yields the desirable perspective of collecting precise measurements of BCA in the valence region, whereas the few measurements of HERMES and COMPASS were obtained on restricted kinematics where sea quarks are expected to dominate. We investigate the foreseen impact of these measurements with the CLAS12 spectrometer [213]. To this purpose, we first generate BCA pseudo-data while giving a specific attention to their plausible experimental uncertainty. We then reweight the replicas of the neural network parameterisation of CFFs of Ref. [143] and summarise the observed features.

4.4.1 . Pseudo-data generation

We select 1656 kinematic bins spanning the (x_B, t, Q^2, ϕ) region based on a technical proposal from a working group of the CLAS Collaboration [4] for the beam operating at energy 10.6 GeV. The bins are constructed by first selecting 13 bins in (x_B, Q^2) shown on Fig. 4.8, and further dividing them along t and ϕ while respecting the condition $-t < 0.2Q^2$ which was used to discard measurements in the neural network fit of CFFs.

4. Eric Voutier, private communication

The neural network fit, whose determination we have detailed in Section 4.2, takes the form of 101 sets of replicas, each set containing the four CFFs $\mathcal{H}$, $\mathcal{E}$, $\tilde{\mathcal{H}}$ and $\tilde{\mathcal{E}}$. For a kinematic bin centered at (x_B, t, Q^2, ϕ) and the k-th set of CFF replicas, we compute the expected number $N_k^{\pm}(x_B, t, Q^2, \phi)$ of single photon lepto-production events respectively for a positron and electron beams of luminosity $0.6 \times 10^{35}\text{cm}^{-2}.\text{s}^{-1}$ with 40 days of data taking each. $N_k^{\pm}(x_B, t, Q^2, \phi)$ is the product of the duration of data taking, luminosity, differential cross-section $\text{d}^4\sigma^{\pm}(x_B, t, Q^2, \phi)$ and bin phase space volume. We assume that the variations of the cross-section over the bin phase space volume are negligible, and a perfect detector acceptance and efficiency. This computation is made straightforward by the inclusion of the replicas in the PARTONS software which allows us to evaluate many DVCS observables.

Once the values of the number of events $N_k^{\pm}(x_B, t, Q^2, \phi)$ are computed in each bin and for each set of replicas, we determine the BCA thanks to (4.23) :

$$A_{C,k}(x_B, t, Q^2, \phi) = \frac{N_k^+(x_B, t, Q^2, \phi) - N_k^-(x_B, t, Q^2, \phi)}{N_k^+(x_B, t, Q^2, \phi) + N_k^-(x_B, t, Q^2, \phi)} \,. \tag{4.25}$$

We omit the arguments (x_B, t, Q^2, ϕ) in the following for concision of the notations. The central value of the pseudo-data is taken as the average over all replicas of $A_{C,k}$, which we denote $\langle A_C \rangle$.

We also need to determine the experimental uncertainty of our pseudo-data. Assuming the $N_k^{\pm}$ are uncorrelated, their statistical uncertainty is $\Delta N_k^{\pm} = \sqrt{N_k^{\pm}}$, so the statistical uncertainty on $A_{C,k}$ can be expressed by the usual uncertainty propagation formula :

$$\left(\Delta A_{C,k}^{stat}\right)^2 = \left(\frac{\partial A_{C,k}}{\partial N_k^+}\right)^2 \left(\Delta N_k^+\right)^2 + \left(\frac{\partial A_{C,k}}{\partial N_k^-}\right)^2 \left(\Delta N_k^-\right)^2 \,, \tag{4.26}$$

$$= \left(\frac{2N_k^-}{(N_k^+ + N_k^-)^2}\right)^2 N_k^+ + \left(\frac{2N_k^+}{(N_k^+ + N_k^-)^2}\right)^2 N_k^- \,, \tag{4.27}$$

$$= \frac{4N_k^+ N_k^-}{(N_k^+ + N_k^-)^3} \,, \tag{4.28}$$

$$= \frac{1 - A_C^2}{N_k^+ + N_k^-} \,. \tag{4.29}$$

Therefore, it would only seem logical to take as statistical uncertainty of our pseudo-data

$$\Delta A_C^{stat} = \sqrt{\frac{1 - \langle A_C \rangle^2}{\langle N^+ + N^- \rangle}} \,, \tag{4.30}$$

where $\langle N^+ + N^- \rangle$ is the average over all replicas of the number of events in the bin. However, the neural network fit is essentially unconstrained in some regions of the probed (x_B, t, Q^2, ϕ) phase space due to a current lack of experimental measurements. In particular, the neural network does not implement – by design – the general theoretical expectation that CFFs and cross-sections should decrease at large $|t|$ values, and produces therefore cross-sections in this region that we consider abnormally large in view of our prior knowledge. This would lead to an unreasonable expectation for the statistical precision of our pseudo-data [5]. For this reason, $N^+ + N^-$ is evaluated in each bin thanks to the physically motivated CFF global fit of Ref. [53], and the result shown on Fig. 4.8.

5. More precisely, the fact that in an unconstrained region, the neural network of Ref. [143] produces

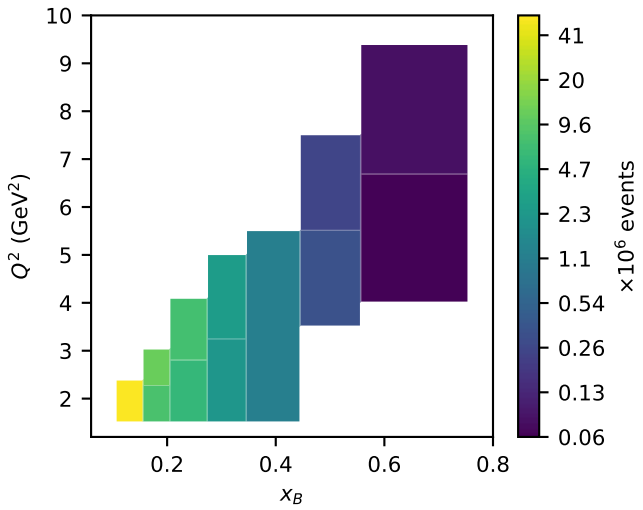

Figure 4.8 – Total number of single photon lepto-production events integrated over the t and ϕ bins for 80 days of operation of at beam energy 10.6 GeV and luminosity $0.6 \times 10^{35}\text{cm}^{-2}.\text{s}^{-1}$. Perfect detector acceptance and efficiency is assumed. Computed with the physically motivated CFF global fit of Ref. [53]. Figure taken from Ref. [2].

Since we use the physically motivated model of Ref. [53] to compute $N^+ + N^-$, we could use it as well to determine A_C. However, we prefer to keep the central value of our pseudo-data as $\langle A_C \rangle$ averaged over neural network replicas to ensure that our new data points are maximally compatible with the current neural network fit, while benefitting from reasonable uncertainties. Indeed, as we mentioned previously, Bayesian reweighting only gives reliable results when many replicas are compatible with the new dataset, and selecting central values of the simulated experimental points that are in tension with the replicas would reduce considerably the statistical relevance of the procedure.

An uncorrelated relative systematic uncertainty of 3% is furthermore assumed following Ref. [214], and added in quadrature to the statistical uncertainty for simplicity in this first sensitivity study. The total uncertainty on A_C is therefore obtained in each bin as uncorrelated

$$\Delta A_C = \sqrt{0.03^2 \langle A_C \rangle^2 + \frac{1 - \langle A_C \rangle^2}{N^+ + N^-}} \,. \tag{4.31}$$

Finally, to simulate actual experimental data taking, we smear the central values of our measurements according to ΔA_C.

large uncertainties approximately normally distributed should be considered as a modelling bias. Let us show the problem on a simple example. Observables are non linear functions of CFFs, which can for instance involve quantities like $1/(1+\mathcal{H}^2)$. It is particularly the case for timelike Compton scattering. In the limit where the uncertainty on $\mathcal{H}$ is represented by a normal distribution of infinite uncertainty, it becomes increasingly likely that $\mathcal{H}$ is large overall. Therefore, it is also increasingly likely that $1/(1+\mathcal{H}^2)$ is close to 0. We are reaching an apparent paradox, where the larger the uncertainty on $\mathcal{H}$, the more precisely the observable is known. The issue boilds down to (1) the non-linearity of the relation between the observable and the CFF means that distances and what they actually measure are severely distorted when switching between the spaces of observables and CFFs and (2) representing unconstrained CFFs by large standard deviation is making an implicit choice of measure, the usual distance in $\mathbb{R}$, which is a significant modelling bias.

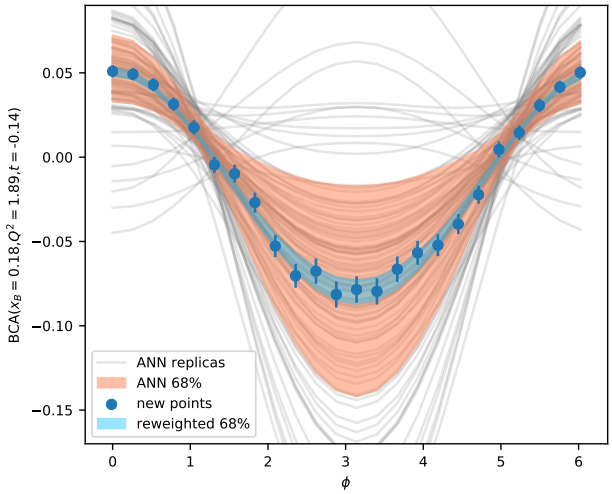

Figure 4.9 – BCA value computed on 24 bins in ϕ for ($x_B = 0.18, t = -0.14$ GeV^2, $Q^2 = 1.89$ GeV^2). The grey lines represent the 101 replicas obtained from the neural network fit and the orange band the associated 68% confidence interval. The blue points are the generated pseudo-data with their uncertainties. The blue band is the reweighted replica distribution. Figure taken from Ref. [2].

4.4.2 . Reweighting and discussion

Fig. 4.9 shows 24 bins in ϕ spanning $[0, 2\pi]$ for ($x_B = 0.18, t = -0.14$ GeV^2, $Q^2 = 1.89$ GeV^2). The value of the BCA computed for each set of CFF replicas is shown by the light grey curves, and the current uncertainty on BCA derived from this fit is shown as the orange 68% confidence interval. In each of the 24 bins in ϕ, a pseudo-data is generated according to the principles we detailed before. The assumed experimental uncertainty, determined by Eq. (4.31), is much smaller than the current uncertainty on the BCA represented by the orange band. The fact that pseudo-data are expected to be much more precise than the current knowledge of the BCA is particularly true in the large $|t|$ and middle (x_B, Q^2) region as well as the small $|t|$ and small (x_B, Q^2) region. On the contrary, the uncertainty of new points will probably not be very constraining in the intermediate $|t|$ region, because CFFs are already well constrained there, or large (x_B, Q^2) region because of the low expected statistics. In some bins of that region, only a few hundred events are expected.

The blue band of Fig. 4.9 shows the reweighted BCA replica distribution computed thanks to Eq. (4.19) after Bayesian reweighting on the new dataset of 24 points has been performed. The effective number of replicas (4.17) has been reduced from $N_{\text{rep}} = 101$ to $N_{\text{eff}} = 8$. For several (x_B, t, Q^2) bins, the effective number of replicas after reweighting on the 24 ϕ bins drops as small as 2. It means the pseudo-data are so precise compared to the actual knowledge of the BCA that all replicas except the couple least bad fits are essentially discarded. The estimation of the reweighted distribution in this context is of course very lacunar, so in an effort to make the results more robust and less dependent on the precise central value of the pseudo-data, we average the weighted mean and standard deviation on 300 successive smearings of the pseudo-data according to their assumed central value and uncertainty. The smearing results in a slight increase of the reweighted uncertainty in the case of particularly low N_{eff}.

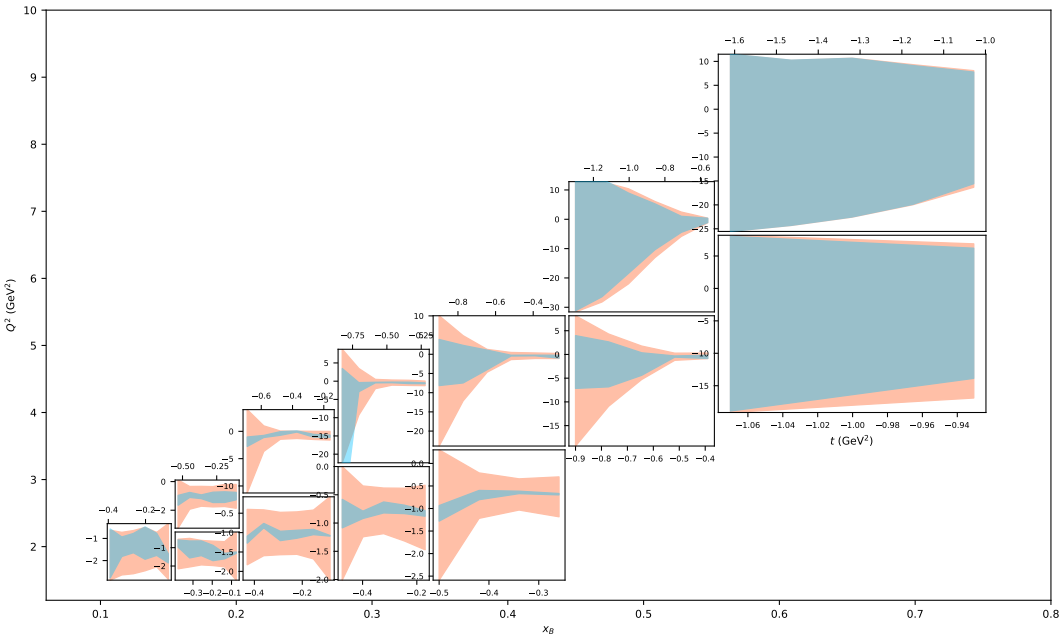

Figure 4.10 – 68% confidence regions for $\mathrm{Re}\,\mathcal{H}$ resulting from the initial neural network fit (orange band) and the local reweighting of the ϕ dependence of BCA pseudo-data (blue band). For each of the thirteen (x_B, Q^2) bins, the subgraph shows the results on the t bins. Figure taken from Ref. [2].

The reweighting of the ϕ dependence of the BCA performed independently on each (x_B, t, Q^2) bin allows us to attribute independent weights to the CFF replicas in each of these bins. We present on Fig. 4.10 the resulting weighted distribution of the real part of the CFF $\mathcal{H}$. Due to the fact that we consider each bin independently, we do not use the information encoded in the overall structure of the CFF replicas in the (x_B, t, Q^2) space. For this reason, we call this procedure a local reweighting. The results are nonetheless striking and show a first assessment of the remarkable refinement on the knowledge of $\mathrm{Re}\,\mathcal{H}$ that could be produced by the BCA observable measured at CLAS12.

An attempt at a global reweighting along the entire dataset of 1656 pseudo-data is bound to fail due to the relatively small number of replicas available. Already in some bins, the simple reweighting of the ϕ-dependence on 24 pseudo-data gave rise to a very low number of effective replicas. The amount of constraint brought by the large dataset would inevitably result in a final effective number of replicas of 1. It is however possible to partly circumvent the issue by considering the results of the local reweighting as an experimental-like input on CFFs themselves. Precisely, we will consider the 68% confidence interval obtained independently on each (x_B, t, Q^2) bin as a normally distributed input on CFFs, and further reweight the replica distribution along these new inputs. That way, we actually factorise the reweighting procedure along successive variables. We show on Fig. 4.11 the reweighting on the 6 t bins in a (x_B, Q^2) bins. Taking into account correlations along the t variables results in a further reduction of the uncertainty on CFFs.

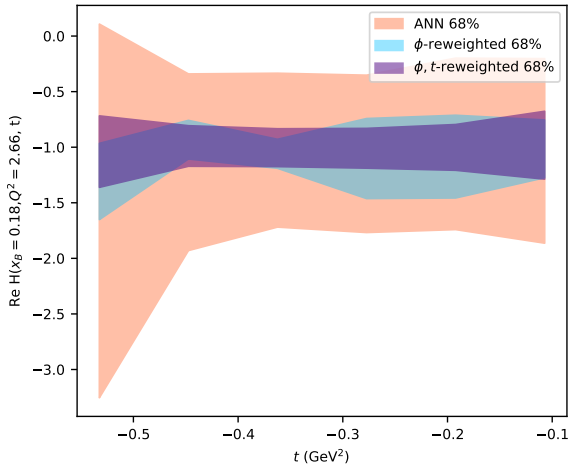

Figure 4.11 – $\mathrm{Re}\,\mathcal{H}(t)$ for $(x_B = 0.18, Q^2 = 2.66\ \mathrm{GeV}^2)$. The orange band is the 68% confidence interval stemming from the neural network fit. The blue band is the result of the local reweighting on the ϕ dependence of BCA pseudo-data. The purple band is a further reweighting using the blue band as an experimental-like input on CFFs. Figure taken from Ref. [2].

Conclusion

The statistical tools presented in this chapter constitute the foundation of the analyses and modelling efforts of this thesis : linear fits, uncertainty estimation via the replica method, neural networks and their training, Bayesian reweighting. We have briefly studied the interest and limitations of linear estimates for non-linear models. First we showed that a study of the χ^2 around its global minimum inspired from the linear case fails to give an account of uncertainty in some situations like a multimodal distribution of the fitted parameters. We highlighted nonetheless that a study of the local minima of the χ^2 contains a lot of information even in these complicated cases. We also demonstrated that the linear view on the effect of adding a new point to a fit can prove quite useful for non-linear models provided some general criteria are fulfilled. Finally we presented our impact study of the addition of observables linked to a positron beam at JLab on the uncertainty in CFF neural networks extractions. Our results show the vast increase in precision that could be achieved, notably on $\mathrm{Re}\,\mathcal{H}$, thanks to this new experimental setup. Reducing the experimental uncertainty on CFFs will appear as a crucial issue in the following chapter where we focus on the extraction of physical properties of hadron matter from the DVCS channel.

5 - Extracting a gravitational form factor from DVCS data

As presented in Section 1.4.2, one of the major motivations for the study of generalised parton distributions is their link to the hadron EMT and the possibility to compute distributions of hadron matter properties like energy or internal pressure. Among the five gravitational form factors (GFFs) A^a, B^a, C^a, $\bar{C}_a$ and D^a_{GFF} that parameterise the EMT proton matrix element (1.21), C^a has attracted most attention owing to the specific experimental sensitivity achievable through the DVCS channel. Section 3.1.2 presents the dispersive formalism of DVCS, which allows one to single out from CFFs the subtraction constant giving information on the GPD D-term $D^a(z,t,\mu^2)$. We remind that the GFF C^a can be computed from the D-term as (see Eqs. (2.33) and (2.34))

$$C^a(t,\mu^2) = \frac{1}{4}\int_{-1}^{1} \mathrm{d}z\, z^{1-p_a} D^a(z,t,\mu^2)\,, \tag{5.1}$$

where $p_a = 1$ if $a = g$ and 0 otherwise. On the other hand, the GFFs A^a, B^a and D^a_{GFF} depend on a Mellin moments of the full GPDs (see Eqs. (1.22) to (1.26)), but are insensitive to the D-term. The determination of these GFFs will therefore come as a by-product of our discussion of the full deconvolution problem in the next chapter.

The first experimental extraction of the GFF C^a from a DVCS dataset was performed in 2018 in Ref. [139]. Their approach relied on a parametric fit of CFFs inspired by the KM model [109, 116]. Shortly after, a data-driven approach to the DVCS subtraction constant using neural network parameterisations of CFFs to reduce the modelling bias was published in Ref. [202]. The first result of Ref. [139] headed in the direction of a very large internal pressure inside the proton, possibly even larger than that modelled at the center of neutron stars and clearly incompatible with 0. The subsequent study with more flexible modelling gave in turn extremely large uncertainties, to the point that the subtraction constant, and consequently the gravitational form factor C^a, appear compatible with 0.

We published in 2021 our own data-driven extraction of the GFF C^a from the world DVCS dataset in Ref. [4]. The neural network study of CFFs performed in Ref. [143], which we presented in Section 4.2 allowed the extraction of the DVCS subtraction constant. Based on this result, we devote a significant attention to the question of the extraction of the D-term and GFF C^a from the subtraction constant, exploring various schemes and their limitations. As part of the general program to extract GPDs from experimental data – known as deconvolution problem – the step of extracting simply the D-term from the subtraction constant is particularly important. It is a simpler problem, in the sense that the D-term is just a function of $(z = x/\xi, t, \mu^2)$ whereas GPDs are functions of (x, ξ, t, μ^2). This means that issues related to the sparsity of experimental data and the overall complexity of the extraction are less stringent. However all the ingredients of the generic problem of deconvolution of GPDs are already present at this stage.

5.1 . D-term extraction as a deconvolution problem

We remind from the discussion of dispersion relations in Section 3.1.2 that the DVCS subtraction constant is related to the D-term through Eq. (3.27), which notably yields at LO :

$$C_{H,0}[D^q, D^g](t, Q^2) = 2 \sum_q e_q^2 \int_{-1}^{1} \mathrm{d}z \, \frac{D^q(z, t, \mu^2)}{1 - z} \,. \tag{5.2}$$

Extracting the D-term at some perturbative order k from the knowledge of the experimental subtraction constant $C_H(t, Q^2)$ defined in Eq. (3.25) consists in finding the functions D^a such that $C_{H,k}[D^q, D^g](t, Q^2)$ fits best $C_H(t, Q^2)$. This raises however the issue of uniqueness of the solution to this optimisation problem. At LO, characterising uniquely $D^q(z, t, \mu^2)$ from the knowledge of $\int \mathrm{d}z \, D^q(z, t, \mu^2)/(1-z)$ in the absence of strong modelling assumptions might seem a hopeless task at first, since one of the variables has been integrated out. There is hope that the extraction of the GFF $C^a(t, \mu^2)$ might be easier, since it is related to the D-term through an integral on z (5.1) as the subtraction constant, so intuitively it contains less information than the full D-term. However, at this stage, in the absence of a straightforward relation between Eqs. (5.2) and (5.1), it is unclear whether there is a definitive advantage in focusing on the extraction of the GFF alone, or the D-term as a whole. We will clarify this aspect in the following.

The issue of extracting a function, or one of its non-trivial features, which has been integrated against a kernel is called a deconvolution problem. We will now demonstrate that for the LO evolution, the theoretical solution to this problem is obtained by observing that the μ^2 dependence of the D-term is already known.

5.1.1 . LO evolution of the D-term

We have detailed in Eqs. (2.58) to (2.60) the solution of the LO evolution for conformal moments of the GPD. We also remind that the D-term evolves independently from the rest of GPD, following exactly the ERBL evolution equation [93, 94]. We can therefore express the evolution of the D-term $D^a(z, t, \mu^2)$ in terms of its own conformal moments. More precisely, we expand the D-term on a basis of Gegenbauer polynomials by introducing the coefficients $d_n^{q,g}(t, \mu^2)$:

$$D^q(z, t, \mu^2) = (1 - z^2) \sum_{\text{odd } n} d_n^q(t, \mu^2) C_n^{(3/2)}(z) \,, \tag{5.3}$$

$$D^g(z, t, \mu^2) = \frac{3}{2}(1 - z^2)^2 \sum_{\text{odd } n} d_n^g(t, \mu^2) C_{n-1}^{(5/2)}(z) \,. \tag{5.4}$$

Unlike the solution of evolution given in Section 2.4 or in Ref. [4] for instance, we find it preferable to recast the evolution from initial scale μ_0^2 to final scale μ^2 at LO in α_s in a matrix system formulation, which allows us to convince ourselves more easily of the linear and invertible nature of evolution and to solve readily the system. We use a simplifying assumption concerning the behaviour of heavy quarks. We neglect the possibility of intrinsic heavy quark content of the proton at low μ^2, and we do not include corrections in the ERBL equation to take into account heavy quark masses. Instead, apart from the three light quarks which are always considered, heavy flavours are progressively activated once μ^2 is larger than the squared mass of the heavy quark m_q^2 with threshold condition $d_n^q(t, m_q^2) = 0$. This means that heavy quark contributions are not actual free parameters,

but entirely radiatively generated by lighter quarks and gluon splittings. If μ^2 and μ_0^2 belong to an interval where no activation threshold is crossed, evolution takes the following form for n odd and n_f active quark flavours [1] :

$$\begin{pmatrix} d_n^g(t,\mu^2) \\ d_n^{q_1}(t,\mu^2) \\ d_n^{q_2}(t,\mu^2) \\ \vdots \\ d_n^{q_{n_f}}(t,\mu^2) \end{pmatrix} = \left(\begin{array}{c|c|c} \frac{1}{a_n^- - a_n^+} & \frac{a_n^+}{(a_n^+ - a_n^-)n_f} \cdots \frac{a_n^+}{(a_n^+ - a_n^-)n_f} & \\ \frac{1}{a_n^+ - a_n^-} & \frac{a_n^-}{(a_n^- - a_n^+)n_f} \cdots \frac{a_n^-}{(a_n^- - a_n^+)n_f} & \\ \hline 0 & 1 & \\ \vdots & \vdots & -I_{n_f-1} \\ 0 & 1 & \end{array}\right)^{-1}$$

$$\times \left(\begin{array}{c|c|c} \frac{\Gamma_n^-(\mu^2,\mu_0^2)}{a_n^- - a_n^+} & \frac{\Gamma_n^-(\mu^2,\mu_0^2)a_n^+}{(a_n^+ - a_n^-)n_f} \cdots \frac{\Gamma_n^-(\mu^2,\mu_0^2)a_n^+}{(a_n^+ - a_n^-)n_f} & \\ \frac{\Gamma_n^+(\mu^2,\mu_0^2)}{a_n^+ - a_n^-} & \frac{\Gamma_n^+(\mu^2,\mu_0^2)a_n^-}{(a_n^- - a_n^+)n_f} \cdots \frac{\Gamma_n^+(\mu^2,\mu_0^2)a_n^-}{(a_n^- - a_n^+)n_f} & \\ \hline 0 & \Gamma_n^0(\mu^2,\mu_0^2) & \\ \vdots & \vdots & -\Gamma_n^0(\mu^2,\mu_0^2)\ I_{n_f-1} \\ 0 & \Gamma_n^0(\mu^2,\mu_0^2) & \end{array}\right) \begin{pmatrix} d_n^g(t,\mu_0^2) \\ d_n^{q_1}(t,\mu_0^2) \\ d_n^{q_2}(t,\mu_0^2) \\ \vdots \\ d_n^{q_{n_f}}(t,\mu_0^2) \end{pmatrix} . \tag{5.6}$$

I_{n_f-1} designates the identity matrix of rank $n_f - 1$,

$$\Gamma_n^0(\mu^2,\mu_0^2) = \left(\frac{\alpha_s(\mu^2)}{\alpha_s(\mu_0^2)}\right)^{2\gamma_n/\beta_0} , \tag{5.7}$$

$$\Gamma_n^\pm(\mu^2,\mu_0^2) = \left(\frac{\alpha_s(\mu^2)}{\alpha_s(\mu_0^2)}\right)^{2\gamma_n^\pm/\beta_0} , \tag{5.8}$$

We detail now the anomalous dimensions γ_n [41] whose definition we had postponed in Section 2.4 since we will need to study their precise values in the following :

$$\gamma_n = \gamma_{QQ}(n) , \tag{5.9}$$

$$\gamma_n^\pm = \frac{1}{2}\left(\gamma_{QQ}(n) + \gamma_{GG}(n) \pm \sqrt{(\gamma_{QQ}(n) - \gamma_{GG}(n))^2 + 4\gamma_{QG}(n)\gamma_{GQ}(n)}\right) , \tag{5.10}$$

$$\gamma_{QQ}(n) = C_F\left(\frac{1}{2} - \frac{1}{(n+1)(n+2)} + 2\sum_{k=2}^{n+1}\frac{1}{k}\right) , \tag{5.11}$$

$$\gamma_{QG}(n) = -n_f T_F \frac{n^2+3n+4}{n(n+1)(n+2)} , \tag{5.12}$$

$$\gamma_{GQ}(n) = -2C_F \frac{n^2+3n+4}{(n+1)(n+2)(n+3)} , \tag{5.13}$$

$$\gamma_{GG}(n) = \frac{2}{3}n_f T_F + C^a\left(\frac{1}{6} - \frac{2}{n(n+1)} - \frac{2}{(n+2)(n+3)} + 2\sum_{k=2}^{n+1}\frac{1}{k}\right) . \tag{5.14}$$

1. The careful reader may wonder why the matrix formulation seems to give a particular role to the quark flavor q_1. It is because it serves as a "pivot" for the flavour non singlet evolution equations

$$d_n^{q_i}(t,\mu^2) - d_n^{q_j}(t,\mu^2) = \Gamma_n^0(\mu^2,\mu_0^2)\left[d_n^{q_i}(t,\mu_0^2) - d_n^{q_j}(t,\mu_0^2)\right] , \tag{5.5}$$

which are all expressed using $q_i = q_1$. All flavours are therefore evolved with respect to the first one. However, when the matrix product is computed, all flavours play a symmetric role, as shown in Eq. (5.17).

Here $C_F = 4/3$, $T_F = 1/2$ and $C^a = 3$. Finally

$$a_n^{\pm} = 2\frac{n_f}{n}\frac{\gamma_n^{\pm} - \gamma_n}{\gamma_{QG}(n)}\,. \tag{5.15}$$

The first term of Eq. (5.6) is invertible since $\gamma_{QG}(n) < 0$ and therefore $a_n^- > a_n^+$. Keeping in mind we work with the assumption of no activation threshold crossed between μ_0^2 and μ^2, the explicit solution of the system is found as

$$d_n^g(t,\mu^2) = \left[\frac{\Gamma_n^- a_n^- - \Gamma_n^+ a_n^+}{a_n^- - a_n^+}\right] d_n^g(t,\mu_0^2) + \frac{a_n^- a_n^+}{n_f(a_n^- - a_n^+)}(\Gamma_n^+ - \Gamma_n^-)\sum_{j=1}^{n_f} d_n^{q_j}(t,\mu_0^2)\,, \tag{5.16}$$

$$\begin{aligned} d_n^{q_k}(t,\mu^2) = \Gamma_n^0 d_n^{q_k}(t,\mu_0^2) + \left[\frac{\Gamma_n^- - \Gamma_n^+}{a_n^- - a_n^+}\right] d_n^g(t,\mu_0^2) \\ + \frac{1}{n_f}\left(-\Gamma_n^0 + \frac{\Gamma_n^+ a_n^- - \Gamma_n^- a_n^+}{a_n^- - a_n^+}\right)\sum_{j=1}^{n_f} d_n^{q_j}(t,\mu_0^2)\,, \end{aligned} \tag{5.17}$$

where we have omitted the arguments (μ^2, μ_0^2) of the Γ_n functions for conciseness. This formulation allows us to immediately notice that the μ dependence of parameters $d_n^a(t,\mu^2)$ of the Gegenbauer expansion of the D-term is obtained simply by a linear combinations of the three functions $\Gamma_n^0(\mu^2,\mu_0^2)$ and $\Gamma_n^{\pm}(\mu^2,\mu_0^2)$.

If evolution from μ_0^2 to μ^2 crosses a heavy quark activation threshold $m_{q_{n_f}}^2$, we assume that the heavy quark contribution is exactly 0 at that threshold, so

$$\begin{pmatrix} d_n^g(t,\mu^2) \\ d_n^{q_1}(t,\mu^2) \\ d_n^{q_2}(t,\mu^2) \\ \vdots \\ d_n^{q_{n_f-1}}(t,\mu^2) \\ d_n^{q_{n_f}}(t,\mu^2) \end{pmatrix} = \Gamma_{n,n_f}\left(\mu^2, m_{q_{n_f}}^2\right) \begin{pmatrix} d_n^g(t,m_{q_{n_f}}^2) \\ d_n^{q_1}(t,m_{q_{n_f}}^2) \\ d_n^{q_2}(t,m_{q_{n_f}}^2) \\ \vdots \\ d_n^{q_{n_f-1}}(t,m_{q_{n_f}}^2) \\ 0 \end{pmatrix}, \tag{5.18}$$

$$= \Gamma_{n,n_f}\left(\mu^2, m_{q_{n_f}}^2\right) \begin{pmatrix} \Gamma_{n,n_f-1}\left(m_{q_{n_f}}^2,\mu_0^2\right) \\ \hline 0 \quad \cdots \quad 0 \end{pmatrix} \begin{pmatrix} d_n^g(t,\mu_0^2) \\ d_n^{q_1}(t,\mu_0^2) \\ d_n^{q_2}(t,\mu_0^2) \\ \vdots \\ d_n^{q_{n_f-1}}(t,\mu_0^2) \end{pmatrix}. \tag{5.19}$$

By lowering μ_0^2 below the squared charm mass, it is possible to express all parameters $d_n^a(t,\mu^2)$ as functions of the four free parameters $d_n^g(t,\mu_0^2)$ and $d_n^q(t,\mu_0^2)$ where $q = u,d,s$. It is straightforward to derive that the μ dependence of $d_n^a(t,\mu^2)$ is obtained by a linear combination of the three functions $\Gamma_n^0(\mu^2, m_{q_{n_f}}^2)$ and $\Gamma_n^{\pm}(\mu^2, m_{q_{n_f}}^2)$.

5.1.2 . Theoretical extraction of the D-term

The LO subtraction constant $C_{H,0}[D^q, D^g]$ (5.2) can be expressed in terms of the Gegenbauer parameters d_n^q (5.3) in the following way

$$C_{H,0}[D^q, D^g](t,Q^2) = 2\sum_q e_q^2 \sum_{\text{odd } n} d_n^q(t,\mu^2)\int_{-1}^{1} \mathrm{d}z\,(1+z)C_n^{(3/2)}(z)\,. \tag{5.20}$$

Noting that the Gegenbauer polynomials are coefficients of the formal series

$$\sum_{n=0}^{\infty} C_n^{(\alpha)}(z) t^n = (1 - 2zt + t^2)^{-\alpha}\,, \tag{5.21}$$

one obtains that

$$\sum_{n=0}^{\infty} \int_{-1}^{1} \mathrm{d}z\,(1+z) C_n^{(3/2)}(z) t^n = \int_{-1}^{1} \mathrm{d}z\,\frac{1+z}{(1-2zt+t^2)^{3/2}}\,, \tag{5.22}$$

$$= \frac{2}{1-t}\,, \tag{5.23}$$

$$= \sum_{n=0}^{\infty} 2t^n\,, \tag{5.24}$$

hence the identification

$$\int_{-1}^{1} \mathrm{d}z\,(1+z) C_n^{(3/2)}(z) = 2\,, \tag{5.25}$$

which yields

$$C_{H,0}[D^q, D^g](t, Q^2) = 4 \sum_q e_q^2 \sum_{\text{odd } n} d_n^q(t, \mu^2)\,. \tag{5.26}$$

As a consequence of this expression, the possibility of extracting unambiguously the LO D-term from the knowledge of the subtraction constant translates directly into the question of uniqueness of the decomposition of a function in an (infinite) sum of $d_n^q(t, \mu^2)$. If two different series $d_n^q(t, \mu^2)$ and $\bar{d}_n^q(t, \mu^2)$ produce the same results when summed along Eq. (5.26), then even a perfect experimental knowledge of the subtraction constant with no uncertainty will not be able to distinguish between the two. In that case, due to the linearity of the problem, there would even be an infinite number of solutions. On the contrary, if the μ^2 dependences of the $d_n^q(t, \mu^2)$ terms are linearly independent, then the uniqueness of the extraction can in theory be achieved. We need therefore to study the effect of evolution on the sum of $d_n^q(t, \mu^2)$.

Using the expression of Eq. (5.17), the solution of the D-term evolution from μ_0^2 to μ^2 with no threshold crossing gives

$$\begin{aligned}\sum_q e_q^2 \sum_{\text{odd } n} d_n^q(t, \mu^2) &= \sum_{\text{odd } n} \Gamma_n^0(\mu^2, \mu_0^2) \left[I_n - \frac{n_e}{n_f} J_n \right] \\ &+ \Gamma_n^-(\mu^2, \mu_0^2) \left[\frac{n_e}{a_n^- - a_n^+} K_n - \frac{n_e}{n_f} \frac{a_n^+}{a_n^- - a_n^+} J_n \right] - \Gamma_n^+(\mu^2, \mu_0^2) \left[\frac{n_e}{a_n^- - a_n^+} K_n - \frac{n_e}{n_f} \frac{a_n^-}{a_n^- - a_n^+} J_n \right],\end{aligned} \tag{5.27}$$

where we have introduced

$$n_e = \sum_q e_q^2\,, \tag{5.28}$$

$$I_n(t, \mu_0^2) = \sum_q e_q^2 d_n^q(t, \mu_0^2)\,, \tag{5.29}$$

$$J_n(t, \mu_0^2) = \sum_q d_n^q(t, \mu_0^2)\,, \tag{5.30}$$

$$K_n(t, \mu_0^2) = d_n^g(t, \mu_0^2)\,, \tag{5.31}$$

and omitted the argument (t, μ_0^2) of I_n, J_n and K_n for brevity. We do not show expressions for one or more threshold crossings whose complexity brings no theoretical change ; it is straightforward to check that the general form of Eq. (5.27) is preserved. The general μ^2 dependence of $C_{H,0}[D^q, D^g]$ remains the same, and its dependence on D^a is equally contained in the variables I_n, J_n and K_n, although they appear with different linear combinations. We remind that the $\Gamma_n^0(\mu^2)$ and $\Gamma_n^\pm(\mu^2)$ functions are proportional to $\alpha_s(\mu^2)^\gamma$ (5.7)-(5.8) for some exponent γ. If all distinct Γ functions possess different γ exponents, they are linearly independent functions of μ^2, meaning that the bracketed expressions after the Γ functions in Eq. (5.27) are uniquely determined by the knowledge of the functional dependence of the subtraction constant on Q^2. Then it is easy to see that the values of I_n, J_n and K_n are also uniquely determined by the invertible system

$$\begin{pmatrix} 1 & -\frac{n_e}{n_f} & 0 \\ 0 & -\frac{n_e}{n_f}\frac{a_n^+}{a_n^- - a_n^+} & \frac{n_e}{a_n^- - a_n^+} \\ 0 & \frac{n_e}{n_f}\frac{a_n^-}{a_n^- - a_n^+} & -\frac{n_e}{a_n^- - a_n^+} \end{pmatrix} \begin{pmatrix} I_n(t, \mu_0^2) \\ J_n(t, \mu_0^2) \\ K_n(t, \mu_0^2) \end{pmatrix} . \tag{5.32}$$

If an activation threshold is crossed, the above system is different and includes notably $\Gamma_n(m_q^2, \mu_0^2)$ factors, but it makes no theoretical change to the discussion. I_n, J_n and K_n for all n is the actual entire information on the D-term which can be retrieved from the perfect experimental knowledge of the DVCS subtraction constant in a LO analysis. Since there are four free parameters d_n^u, d_n^d, d_n^s and d_n^g to recover with our simplified heavy quark schemes, it is therefore not possible to separate the values of d_n^q for each flavour. However, a limited flavour separation is theoretically possible with the assumption $d_n^u = d_n^d$ and a free d_n^s and d_n^g, which provides three free variables to recover from the three quantities I_n, J_n and K_n. Notice that this separation is made possible by the fact that I_n and J_n are not proportional thanks to the different electric charges of quarks [2]. The most common assumption is however that $d_n^u = d_n^d = d_n^s$, that is equal light flavour contributions. This prior bias makes the extraction more robust, because the system is over-constrained with only two free parameters for three independent equations. The least squares solution can be used to determine the appropriate values.

As we have seen, the fact that I_n, J_n and K_n can theoretically be extracted from the subtraction constant depends on the fact that all γ exponents associated to the various Γ functions are actually always distinct from one another. Owing to the slightly complicated form of the exponents presented in Eqs. (5.9)-(5.14), it is however not completely obvious that it is the case. To demonstrate that γ_n and γ_n^- can never be equal, we first consider that in the limit $n \to \infty$,

$$\gamma_n \sim 2C_F \log(n) , \tag{5.33}$$

$$\gamma_{GG}(n) \sim 2C^a \log(n) , \tag{5.34}$$

$$\gamma_{QG}(n)\gamma_{GQ}(n) \sim \frac{2n_f C_F T_F}{n^2} . \tag{5.35}$$

Since $C^a > C_F$, $\gamma_{GG}(n)$ is asymptotically larger than γ_n, and $\gamma_{QG}(n)\gamma_{GQ}(n)$ is negligible

2. This limited flavour separation might be interesting since flavours u and d are closely related thanks to the (imperfect) isospin symmetry SU(2). The flavour symmetry SU(3) relating u, d and s is significant less stringent.

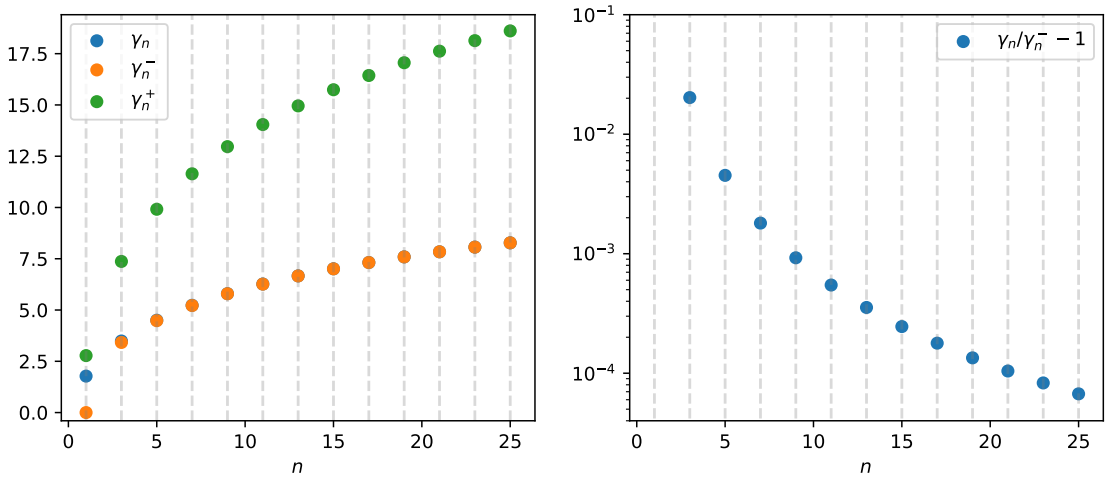

Figure 5.1 – (left) Values of γ_n and $\gamma_n^\pm$ for $n_f = 3$ and $n = 1$ to 25 – (right) Relative discrepancy between γ_n and γ_n^-. At $n = 3$, the two are already in agreement at the level of 2%.

compared $(\gamma_{GG}(n) - \gamma_n)^2$, so

$$\gamma_n^- = \frac{1}{2}\left(\gamma_n + \gamma_{GG}(n) - |\gamma_{GG}(n) - \gamma_n|\sqrt{1 + \frac{4\gamma_{QG}(n)\gamma_{GQ}(n)}{(\gamma_{GG}(n) - \gamma_n)^2}}\right) \tag{5.36}$$

$$= \gamma_n - \frac{\gamma_{QG}(n)\gamma_{GQ}(n)}{\gamma_{GG}(n) - \gamma_n} + \mathcal{O}\left(\frac{\gamma_{QG}(n)\gamma_{GQ}(n)}{\gamma_{GG}(n) - \gamma_n}\right)^2, \tag{5.37}$$

$$= \gamma_n - \frac{n_f C_F T_F}{(C^a - C_F)n^2 \log(n)} + \mathcal{O}\left(\frac{1}{n^2 \log^2(n)}\right). \tag{5.38}$$

Eq. (5.38) proves that as n increases, γ_n^- is asymptotically below γ_n by a distance of order $1/(n^2 \log(n))$. Since γ_{n-1} and γ_n are asymptotically separated by $2C_F/n$, it also means γ_n^- is too close to γ_n to possibly be equal to γ_{n-1}. Therefore, asymptotically, $\gamma_{n-1} < \gamma_n^- < \gamma_n$. Numerically, the asymptotic regime starts as soon as $n = 3$ as demonstrated on Fig. 5.1, with $\gamma_3 = 3.49$ and $\gamma_3^- = 3.42$.

As for γ_n^+, it varies as $\gamma_n C^a/C_F$, so grows so quickly compared to γ_n and γ_n^- that the question of equality of a value of γ_n^+ with some γ_k or γ_k^- is hardly a practical issue. It would take an extraction taking into account $n = 47$ terms in the Gegenbauer expansion of the D-term for γ_n to be even close to γ_5^+ for instance. In practice, the rapid growth of γ_n^+ associated to the limited number of extracted coefficients allows us to consider in the following that γ exponents associated to different Γ functions are formally distinct.

However, Eq. (5.38) also demonstrates how dangerously close γ_n^- is from γ_n. Since in practice $\gamma_n^- \approx \gamma_n$ for $n \geq 3$, we have equivalently $\Gamma_n^0 \approx \Gamma_n^-$ and we loose the ability to independently determine I_n, J_n and K_n from the three different coefficients of each term in n of Eq. (5.27). Indeed, there are now only two equations available to determine three parameters. We stress that what we are actually demonstrating is that, although the problem of extracting the I_n, J_n and K_n parameters admits formally a unique solution, it is numerically ill-defined, so that even with perfect experimental knowledge of the subtraction constant on an interval in Q^2, no reliable independent extraction of the I_n, J_n and K_n coefficients can be performed at even small values of n. The root of the problem

is that with a Gegenbauer expansion of p terms of the D-term, we are proposing a linear model with $3p$ Γ functions to fit the subtraction constant, but the family of Γ functions, although free which guarantees uniqueness of the decomposition, is comprised of numerically extremely similar functions [3]. Of course, the addition of experimental uncertainty will only make the problem worse. We delay however the study of actual experimental uncertainty to Section 5.2.

Conclusion

To summarise the characteristics of the deconvolution problem of the D-term in a LO study, we note that it is mathematically possible to uniquely determine I_n, J_n and K_n (5.29)-(5.31) at some point t from a perfect experimental knowledge of the subtraction constant $C_{H,LO}$ at the same t and an interval on Q^2. The fact that the subtraction constant is measured on an interval in Q^2 allows formally the discrimination between each terms in μ^2 contributing to Eq. (5.27). This is as close to a determination of the D-term as can be achieved considering it is not possible to separate flavours from the DVCS subtraction constant alone. However, due to the extreme numerical proximity of γ_n^- and γ_n, it is not practically possible to determine I_n, J_n and K_n for $n \geq 3$, but only a linear relation relating them. The addition of heavy quark contributions solely through evolution makes no formal difference on the extraction since it does not change the number of free parameters or the μ dependence.

Assumptions or external information coming from other dataset on flavour separation allows one to progress further. Early studies of the use of deeply virtual meson production data have demonstrated the potential of a joint analysis [215]. For instance, the assumption $d_n^u = d_n^d = d_n^s$ allows one to perform theoretically a full D-term extraction. Practically, by linking directly I_n to J_n and reducing the number of parameters to extract, this assumption evades the numerical issue with $\gamma_n^- \approx \gamma_n$. However, as n increases, γ_{n-1} gets in turn closer to γ_n [4], so the numerical stability of the extraction worsens anyway.

The larger the interval in Q^2 on which the subtraction constant is known, the more robust the extraction will be, since for instance the approximation $\Gamma_n^0 \approx \Gamma_n^-$ will be less accurate over a large range of scales, giving more breath to separate I_n, J_n and K_n.

5.1.3 . Theoretical extraction of the GFF C

In terms of Gegenbauer parameters d_n^q, the GFF C^q (5.1) writes

$$C^q(t,\mu^2) = \frac{1}{4}\sum_{\text{odd } n} d_n^q(t,\mu^2)\int_{-1}^{1} \mathrm{d}z\,(1-z^2)zC_n^{(3/2)}(z)\,, \tag{5.39}$$

$$= \frac{1}{12}\sum_{\text{odd } n} d_n^q(t,\mu^2)\int_{-1}^{1} \mathrm{d}z\,(1-z^2)C_1^{(3/2)}(z)C_n^{(3/2)}(z)\,. \tag{5.40}$$

The Gegenbauer polynomials $C_n^{(\alpha)}(z)$ are orthogonal for the $L^2([-1,1])$ inner product with weight $(1-z^2)^{\alpha-1/2}$, so only the term $n=1$ gives a non-vanishing contribution to

3. An ideal situation would be that the Γ functions are for instance orthogonal for the $L^2([\mu_0^2,\mu_{\max}^2])$ inner product, so that each new vector brings intrinsically different information, and not extremely redundant information as is the case now. It is a pity nature chose these values for the anomalous dimensions.

4. Since $\gamma_n \sim 2C_F\log(n)$ (5.33), $\gamma_n - \gamma_{n-1} \sim 2C_F/n$

the sum of Eq. (5.40), and

$$C^q(t,\mu^2) = \frac{1}{5} d_1^q(t,\mu^2)\,. \tag{5.41}$$

A similar calculation with Gegenbauer polynomials $C_n^{(5/2)}(z)$ gives

$$C^g(t,\mu^2) = \frac{1}{5} d_1^g(t,\mu^2)\,. \tag{5.42}$$

The sum over all types of partons defines the full GFF $C(t)$ which is scale-independent as an observable form factor of the energy momentum tensor :

$$C(t) = \sum_q C^q(t,\mu^2) + C^g(t,\mu^2) = \frac{1}{5}\left(\sum_q d_1^q(t,\mu^2) + d_1^g(t,\mu^2)\right)\,, \tag{5.43}$$

$$= \frac{1}{5}\left(B_1(t,\mu^2) + C_1(t,\mu^2)\right)\,, \tag{5.44}$$

following the notations introduced in Eqs. (5.29)-(5.31). As we have demonstrated in the previous section, B_1 and C_1 can theoretically and practically be extracted from a perfect experimental knowledge of the subtraction constant in a LO study at the same value of t and on an interval in Q^2. Since γ_1 and $\gamma_1^{\pm}$ do not interfere with other exponents of very close value which would bring a significant numerical uncertainty, the difficulties we highlighted in the previous section for $n \geq 3$ are less relevant here. This shows therefore that indeed, extraction of the GFF is more robust than that of the D-term.

We finish by noting that, since all values of the γ exponents are strictly positive except $\gamma_1^- = 0$ and $\alpha_s(\mu^2)$ tends to 0 as μ^2 tends to infinity, Eqs. (5.7) and (5.8) readily show that all $\Gamma(\mu^2,\mu_0^2)$ functions tend to 0 in this limit except $\Gamma_1^-(\mu^2,\mu_0^2)$ which equals 1. Therefore, Eq. (5.27) gives

$$\sum_q e_q^2 \sum_{\text{odd } n} d_n^q(t,\mu^2) \overset{\mu^2\to\infty}{\longrightarrow} \frac{n_e}{a_1^- - a_1^+} C_1(t,\mu_0^2) - \frac{n_e}{n_f}\frac{a_1^+}{a_1^- - a_1^+} B_1(t,\mu_0^2)\,, \tag{5.45}$$

and computing

$$a_1^- = 2\frac{C_F}{T_F}\,, \tag{5.46}$$

$$a_1^+ = -n_f\,, \tag{5.47}$$

we obtain

$$\sum_q e_q^2 \sum_{\text{odd } n} d_n^q(t,\mu^2) \overset{\mu^2\to\infty}{\longrightarrow} \frac{n_e}{2\frac{C_F}{T_F} + n_f}\left(C_1(t,\mu_0^2) + B_1(t,\mu_0^2)\right) = \frac{5n_e}{2\frac{C_F}{T_F} + n_f} C(t)\,. \tag{5.48}$$

As we mentioned before, obtaining values of the subtraction constant on a large Q^2 range is beneficial for the numerical quality of the extraction of I_n, J_n and K_n by reducing the similarity of the fitted Γ functions, but we see here that the large Q^2 limit gives an immediate access to the GFF $C(t)$ without any need for a practical deconvolution.

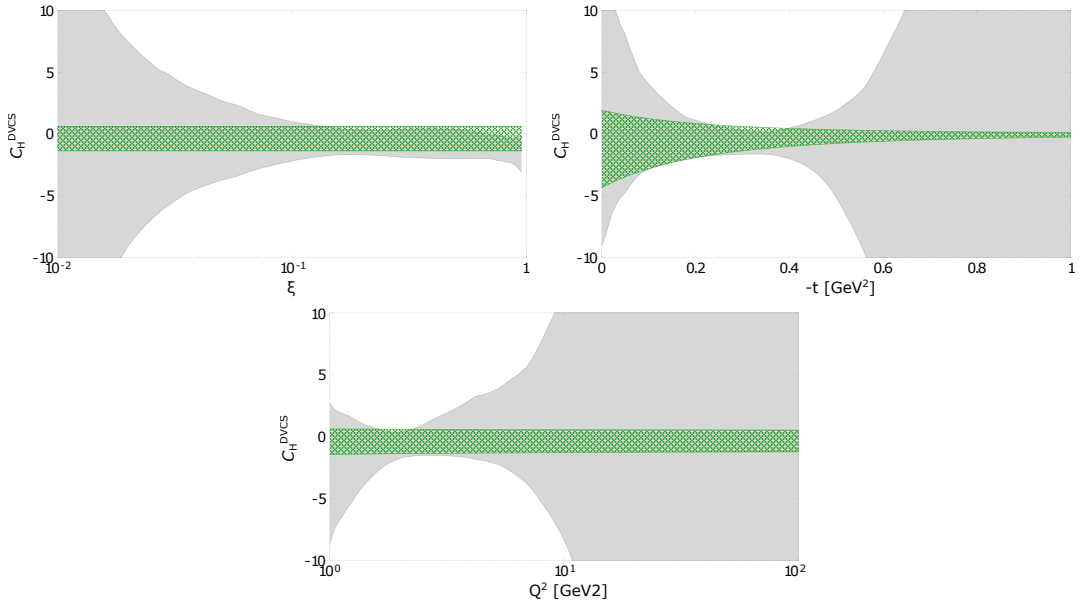

Figure 5.2 – Extraction of the DVCS subtraction constant C_H related to CFF $\mathcal{H}$ from the neural network parameterisation of Ref. [143] (grey band) and subsequent fit by the simplest Ansatz with just one free parameter detailed in Section 5.2.1 (green band) as a function of ξ for $(t = -0.3\ \text{GeV}^2, Q^2 = 2\ \text{GeV}^2)$ (top left), a function of $-t$ for $(\xi = 0.2, Q^2 = 2\ \text{GeV}^2)$ (top right) and a function of Q^2 for $(\xi = 0.2, t = -0.3\ \text{GeV}^2)$ (bottom). Figure taken from Ref. [4].

5.2 . Practical extraction from experimental data at leading order

We have mostly focused so far on the uniqueness and theoretical amplification of uncertainty of the extraction, without mentioning the additional uncertainty coming from experimental knowledge about the subtraction constant. This is the objective of this section, where we notably present our results on the extraction of the GFF from Ref. [4].

The analysis of most of the world DVCS dataset led in Ref. [143] provides a data-driven extraction of CFFs under the form of 101 neural network replicas. We have detailed in Section 4.2 the general features of this extraction. The real and imaginary parts of the CFFs are modelled independently, allowing to compute 101 replicas of the subtraction constant $C_H(t, Q^2)$ related to CFF $\mathcal{H}$ following Eq. (3.25). For practical reasons, the integral on the imaginary part of the CFF is only evaluated for $\xi' > 10^{-6}$, and not $\xi' > 0$, but we estimate that it generates an approximation of the order of 1% only for the values of ξ corresponding to actual DVCS measurements included in the neural network fit of the CFFs. The results are shown on Fig. 5.2 by the grey band, representing the 68% confidence interval with iterative removal of outliers lying more than three standard deviation from the mean, as described in more details in Section 4.2.

The subtraction constant $C_H(t, Q^2)$ is theoretically supposed to be independent of ξ, but since the real and imaginary parts of CFF $\mathcal{H}$ were extracted independently, it is expected that some variation along ξ occurs. We note that on Fig. 5.2, the confidence interval as a function of ξ is approximately horizontal in the region $\xi > 0.1$ at $t = -0.3$ GeV^2 and $Q^2 = 2$ GeV^2 which is where the bulk of JLab experimental measurements are located. At lower values of ξ, the lack of strong experimental constraints prevents the expected ξ independence. Similarly, the small and large $|t|$ and large Q^2 regions are

essentially unconstrained, explaining the very large confidence intervals.

One major feature of this extraction of the subtraction constant is that it is compatible with 0. This results to a large extent from the fact that the real part of CFF $\mathcal{H}$ is poorly constrained by current experimental data, as seen for instance on the left panel of Fig. 4.6. This highlights the interest of a better knowledge of this quantity, which could be achieved for instance by measuring unpolarised BCA observables with a positron beam at JLab as we showed in Section 4.4. That the subtraction constant is compatible with 0 for a flexible parameterisation of CFFs is also a feature of the neural network study of Ref. [202].

Because of the large experimental uncertainty, no clear trend appears in the subtraction constant at large $|t|$ or large Q^2. We know from the discussion on pseudo-data generation for reweighting performed in Section 4.4.1 that using the replicas at large $|t|$ introduces a source of bias in the analysis, because it favours large values in this region, while knowledge of elastic form factors at large $|t|$ would rather hint at small and decreasing values. Furthermore, we demonstrated in the previous section that the Q^2 dependence is instrumental to perform an actual extraction of the GFF or D-term. That the subtraction constant is so poorly known as soon as Q^2 is larger than a few GeV^2 is already preventing any possibility of a serious unbiased extraction of either quantities in the current state of experimental knowledge. Unlike Ref. [202], we wish nonetheless to proceed further and decide to introduce a physically motivated Ansatz at this stage.

5.2.1 . Simple Ansatz

We first start by the simplest possible Ansatz, with solely one free parameter. This is a for instance the assumption made in Ref. [139]. We limit the expansion of the D-term (5.3)-(5.4) to only $n = 1$ and all terms with $n \geq 3$ are assumed to be zero. Thanks to the properties of the evolution of the D-term which we detailed in Section 5.1.1, this property is true at all scale. We further assume that all light flavours have equal contributions, that is $d_1^u = d_1^d = d_1^s \equiv d_1^{uds}$. The gluon contribution is not let free, but generated radiatively from the threshold condition

$$d_1^g(t, 0.1 \text{ GeV}^2) = 0\,. \tag{5.49}$$

For $\mu^2 > 0.1$ GeV^2, $d_1^g(t, \mu^2)$ takes non vanishing values thanks to evolution. This assumption is frequently met in the computation of various parton distribution functions from quark models, and does not prevent an analysis of the existing DVCS data. We study furthermore later the sensitivity on this choice of threshold. Considering most of the DVCS data are obtained for Q^2 larger than the squared mass of the charm quark $m_c^2 \approx 1.64$ GeV^2, we allow a radiatively generated charm contribution, from threshold

$$d_1^c(t, m_c^2) = 0\,. \tag{5.50}$$

Finally, the t-dependence of d_1^{uds} is fixed by assuming the factorised form

$$d_1^{uds}(t, \mu^2) = d_1^{uds}(\mu^2) \left(1 - \frac{t}{\Lambda^2}\right)^{-\alpha}\,, \tag{5.51}$$

where $\Lambda = 0.8$ GeV and $\alpha = 3$. The value of Λ is motivated by the chiral quark-soliton model [55, 216], while that for α ensures a realistic shape of the pressure distribution at large t [64]. Since the dependence of d_1^{uds} on μ^2 is determined by evolution equations, the only free parameter in this model can be taken as the value of $d_1^{uds}(\mu_0^2)$ at some scale μ_0^2, so one real number.

We fit the model independently on each replica of the subtraction constant. The fit is performed by finding the optimal value of the free parameter $d_1^{uds}(\mu_0^2)$ minimising the following least squares for each replica i :

$$\chi_i^2(d_1^{uds}(\mu_0^2)) = \sum_{j=1}^{N_{\text{pts}}} \left(\frac{C_{H,i}(\xi_j, t_j, Q_j^2) - \text{model}_{d_1^{uds}(\mu_0^2)}(\xi_j, t_j, Q_j^2)}{\sigma_{C_H}(\xi_j, t_j, Q_j^2)} \right)^2 , \tag{5.52}$$

where (ξ_j, t_j, Q_j^2) represent the $N_{\text{pts}} = 277$ kinematics of all actual measurements included in the CFF neural network fit, $C_{H,i}(\xi_j, t_j, Q_j^2)$ is the value of the i-th replica of the subtraction constant at the probed kinematics, $\sigma_{C_H}(\xi_j, t_j, Q_j^2)$ the uncertainty on the subtraction constant computed from the robust estimate of the dispersion of the replica distribution, and $\text{model}_{d_1^{uds}(\mu_0^2)}$ the value produced by our model for the free parameter $d_1^{uds}(\mu_0^2)$. We only use values of kinematics which correspond to actual measurements included in the CFF fit to minimise the introduction of bias by fitting in unconstrained regions. The result of the fit is 101 replicas of our model, whose confidence interval is represented as the green bands in Fig. 5.2. Independence on ξ is now a direct feature of the model, and the t dependence is entirely driven by the form of Eq. (5.51). As for the Q^2 dependence, it is obtained by solving the LO evolution equations, and it appears that on a scale from $Q^2 = 1$ GeV^2 to 100 GeV^2, the standard deviation at $\xi = 0.2$ and $t = -0.3$ GeV^2 of the replicas is only reduced from 1.05 to 0.85 (see lower panel of Fig. 5.2). Evolution effects are therefore overall quite tame on d_1^q, which is partly a consequence of the fact that $\gamma_1^- = 0$.

The results of the fit are the following, expressed at reference scale $\mu_0^2 = 2$ GeV^2. Only $d_1^{uds}(\mu_0^2)$ is actually a free parameter, the other contributions being radiatively generated.

Parameter	Value
$d_1^{uds}(\mu_0^2)$	-0.5 ± 1.2
$d_1^c(\mu_0^2)$	-0.002 ± 0.005
$d_1^g(\mu_0^2)$	-0.6 ± 1.6

The charm quark contribution comes out as negligible, and as expected, all contributions are compatible with 0.

The value at which the radiative gluon threshold is fixed (5.49), which is 0.1 GeV^2 for the result presented above, has a significant impact on the extraction of $d_1^g(\mu_0^2)$. Fig. 5.4 shows that choosing 1 GeV^2 as a threshold for gluon contributions results in a reduction of the amplitude of gluon contributions at 2 GeV^2 by a factor 8. However, the value of $d_1^{uds}(\mu_0^2)$ is remarkably insensitive to that value. Therefore, the extraction of $\sum_q d_1^q(t, \mu^2)$ is robust with respect to this threshold, but not the GFF $C(t) = \sum_q d_1^q(t, \mu^2) + d_1^g(t, \mu^2)$ (5.44). Insensitivity of the quark contribution to the gluon threshold will be fully understood in the upcoming discussions. Fig 5.3 shows the comparison of our extraction with several other analyses, including phenomenological extractions, lattice-QCD predictions and theoretical models.

5.2.2 . New terms in the Gegenbauer expansion

Among several assumptions that can be relaxed, the most interesting to our discussion is that only $n = 1$ contributions were considered so far. The chiral quark-soliton model

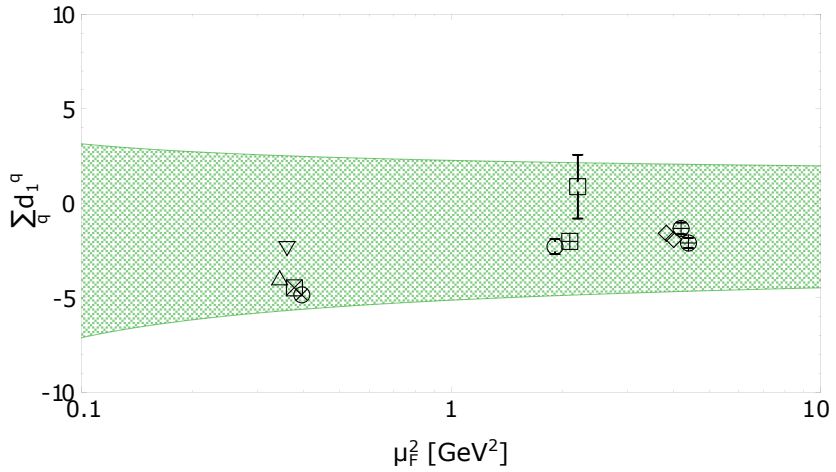

Figure 5.3 – $\sum_q d_1^q(t=0,\mu_F^2)$ obtained for different extractions. The reader interested in extractions of this quantity from lattice and theoretical models will find in Ref. [4] where this figure was borrowed the references to the studies corresponding to each marker. In the context of our presentation, let us only mention the three experimental extractions depicted. The green band represents the result of our study with the simplest Ansatz with just one free parameter. ○ at $\mu_F^2 = 2$ GeV^2 is the 2018 experimental extraction relying on a physically motivated parameterisation of CFFs [139]. □ at $\mu_F^2 = 2.2$ GeV^2 is the data-driven study with neural network [202]. The uncertainty associated to ○ is tiny, unlike □ which is compatible with 0. Our study has even larger uncertainty.

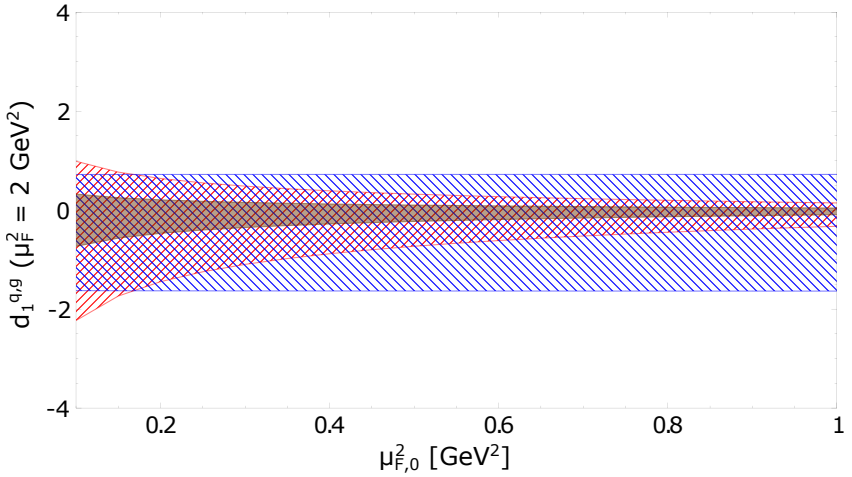

Figure 5.4 – Values of $d_1^{uds}(2\text{ GeV}^2)$ (blue hatched band), $d_1^g(2\text{ GeV}^2)$ (red hatched band) and $d_1^c(2\text{ GeV}^2)$ (brown solid band) depending on the threshold at which gluon contribution is fixed to 0. The charm contribution has been amplified by a factor 100. $d_1^{uds}(2\text{ GeV}^2)$ has no apparent sensitivity to the gluon threshold. Figure taken from Ref. [4].

(χQSM) [171] gives an insight on the behaviour of the Gegenbauer expansion :

$$d_1^{u+d}(0, 0.36 \text{ GeV}^2) \approx -4.0\,, \tag{5.53}$$
$$d_3^{u+d}(0, 0.36 \text{ GeV}^2) \approx -1.2\,, \tag{5.54}$$
$$d_5^{u+d}(0, 0.36 \text{ GeV}^2) \approx -0.4\,. \tag{5.55}$$

Each new term is roughly three times smaller than the previous. When extrapolating this tendency, terms larger than $n = 1$ therefore amount to about 30% of the value of the subtraction constant at LO for this model. Fitting with only the $n = 1$ contribution corresponds therefore to a significant inclusion in d_1^a of contributions that should belong to d_3^a, d_5^a, ... Not only are higher n contributions possibly significant from a numerical point of view, they also play an important theoretical role. We demonstrated in the beginning of this chapter how evolution is supposed to allow us to identify B_1 and C_1 independently from contributions of higher values of n. We presented leverage in Q^2 as a mean to perform deconvolution in a model independent way. Let us therefore allow a free contribution d_3^{uds}, while keeping radiatively generated gluon and charm contributions for both $n = 1$ and $n = 3$. That there is only one free parameter for each value of n discards the theoretical concerns we exposed about the inability to properly identify I_n, J_n and K_n. The result of the fit with the two free parameters d_1^{uds} and d_3^{uds} is the following at $\mu_0^2 = 2$ GeV^2 :

Parameter	Value
$d_1^{uds}(\mu_0^2)$	11 ± 25
$d_3^{uds}(\mu_0^2)$	-11 ± 26
$d_1^c(\mu_0^2)$	-0.05 ± 0.1
$d_3^c(\mu_0^2)$	-0.004 ± 0.01
$d_1^g(\mu_0^2)$	15 ± 34
$d_3^g(\mu_0^2)$	-1.8 ± 3.9

Compared to the previous fit where we had obtained $d_1^{uds}(\mu_0^2) = -0.5 \pm 1.2$, uncertainty on d_1^{uds} has been increased by a factor 20. Fig. 5.5 shows the joint distribution of $d_1^{uds}(0.1$ GeV$^2)$ and $d_3^{uds}(0.1$ GeV$^2)$ over the 101 replicas. With a correlation coefficient of -0.997, this distribution essentially amounts to a linear relation between d_1^{uds} and d_3^{uds}. We will now show how we can understand and predict the general features of these results.

Our objective is to derive an estimate of the standard deviation of the free parameters (d_1^{uds}, d_3^{uds}) depending on the available range of evolution, to broadly assess the impact of future experimental facilities like the EIC on the extraction of hadron mechanical properties without going through the generation a large pseudo-dataset and fitting it. To this purpose, we first present a simplification of evolution to make the problem more tractable analytically. Thanks to the independence of evolution operators on t at leading twist, all this derivation will be performed implicitly at a fixed value of t – in practice, $t = 0$ where all fitted values were extracted in the previous tables.

Simplified evolution

In the simple case where d_n^{uds} is the only free parameter for each n, since evolution equations are linear in d_n^{uds}, they can be exactly summarised as

$$d_n^{uds}(\mu^2) = \Gamma_n^{qq}(\mu^2, \mu_0^2)\, d_n^{uds}(\mu_0^2) \tag{5.56}$$

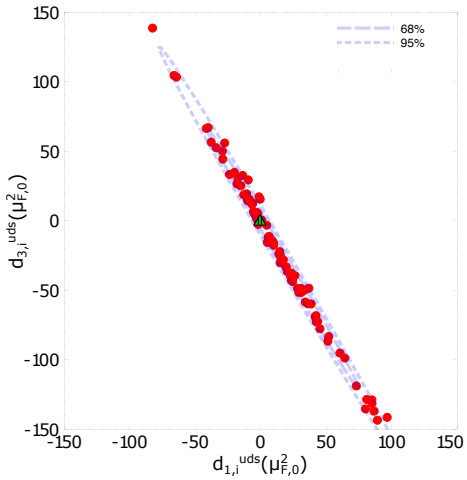

Figure 5.5 – Joint distribution of $d_1^{uds}(0.1\ \text{GeV}^2)$ and $d_3^{uds}(0.1\ \text{GeV}^2)$ over 101 replicas of the model where the two quantities are the only free parameters. The ellipses show the 68% and 95% confidence regions. Figure taken from Ref. [4].

for some functions Γ_n^{qq} [5]. Taking into account threshold crossings and radiative generation of gluons and heavy flavours is straightforward from a theoretical point of view, but complicates significantly all explicit derivations. Considering that the charm content obtained in the fit with the simplest Ansatz was of the order of 0.1% of the sum of the contributions of light flavours, we will consider that $d_n^c = 0$ at all scale. However, we will keep $n_f = 4$ in evolution equations since most of the fitted data are above charm production threshold, while verifying that $n_f = 3$ or 4 makes actually little difference on our results. Finally, we will likewise ignore gluon contributions. A formal discussion of why radiative gluon contributions can be ignored will be led in Section 5.2.4, but let us quickly justify it so far by the fact that varying the threshold where gluon contributions start to be radiatively generated from 0.1 GeV^2 to 1 GeV^2 made no difference on the fitted value of $d_1^{uds}(2\ \text{GeV}^2)$, indicating that in this regime, d_1^g brings no significant contribution to the subtraction constant. With these simplifying assumptions, Eq. (5.17) gives

$$\Gamma_n^{qq}(\mu^2,\mu_0^2) = \frac{1}{4}\,\Gamma_n^0(\mu^2,\mu_0^2) + \frac{3}{4}\,\frac{\Gamma_n^+(\mu^2,\mu_0^2)a_n^- - \Gamma_n^-(\mu^2,\mu_0^2)a_n^+}{a_n^- - a_n^+}\,. \tag{5.57}$$

Construction of an estimate of the uncertainty of d_3^{uds} depending on the lever arm in evolution

Fig. 5.6 shows the values of $\Gamma_1^{qq}(\mu^2,\mu_0^2)$ and $\Gamma_3^{qq}(\mu^2,\mu_0^2)$ for $\mu_0^2 = 2$ GeV^2 obtained thanks to Eq. (5.57) on the range $\mu^2 \in [1.5\ \text{GeV}^2, 10\ \text{GeV}^2]$. The latter contains almost all data on which the neural network fit of CFFs has been performed. It is striking that both functions are numerically very close to one another. We propose in Appendix A.2 a demonstration of the effect of performing a linear fit with two parameters whose associated functional dependences are numerically very close. The generic result is that, as the

5. Let us stress that Eq. (5.56) is not an approximation but an exact consequence of linearity as long as the postulate that only d_n^{uds} is a free parameter is valid. This means that the operator Γ_n^{qq} defined in this way has a complicated dependence on Γ_n^0, $\Gamma_n^\pm$ and all the radiative thresholds.

difference between the two functional dependences decreases like ϵ [6], the uncertainty on the fitted parameters increases as $1/\epsilon$, and their correlation coefficient converges quickly as $-1+\mathcal{O}(\epsilon^2)$. We understand therefore that the strong numerical proximity of $\Gamma_1^{qq}(\mu^2,\mu_0^2)$ and $\Gamma_3^{qq}(\mu^2,\mu_0^2)$ over the range in μ^2 probed by experimental data contributes to the explosion of uncertainty when fitting jointly $d_1^{uds}(\mu_0^2)$ and $d_3^{uds}(\mu_0^2)$, as well as the very large anti-correlation between the two [7].

If on the contrary, Γ_1^{qq} and Γ_3^{qq} had been extremely dissimilar, we could have expected that the fitted parameters $d_1^{uds}(\mu_0^2)$ and $d_3^{uds}(\mu_0^2)$ would be almost uncorrelated. Then using that

$$C_{H,0}[D^q,D^g](Q^2) = \frac{8}{3}\left(d_1^{uds}(Q^2)+d_3^{uds}(Q^2)\right), \tag{5.58}$$

$$= \frac{8}{3}\left(d_1^{uds}(\mu_0^2)\Gamma_1^{qq}(Q^2,\mu_0^2)+d_3^{uds}(\mu_0^2)\Gamma_3^{qq}(Q^2,\mu_0^2)\right), \tag{5.59}$$

assuming that $\mathrm{cov}(d_1^{uds}(\mu_0^2),d_3^{uds}(\mu_0^2))=0$, and noting $\Delta d_{1,3}^{uds}$ the standard deviation of $d_{1,3}^{uds}(\mu_0^2)$, we could write

$$\left(\Delta d_1^{uds}\right)^2\left(\Gamma_1^{qq}(Q^2,\mu_0^2)\right)^2+\left(\Delta d_3^{uds}\right)^2\left(\Gamma_3^{qq}(Q^2,\mu_0^2)\right)^2=\left(\frac{3}{8}\right)^2\sigma_{C_H}^2(Q^2), \tag{5.60}$$

where $\sigma_{C_H}(Q^2)$ is the experimental uncertainty on the subtraction constant at the implicit fixed value of t we use in all this derivation. Eq. (5.60) would yield for instance the simple inequality on the uncertainty of d_3^{uds}

$$\Delta d_3^{uds} \leq \frac{3}{8}\inf\frac{\sigma_{C_H}(Q^2)}{|\Gamma_3^{qq}(Q^2,\mu_0^2)|}, \tag{5.61}$$

where the inferior bound is computed on the range in Q^2 where experimental data are available. However, the very large correlation between d_1^{uds} and d_3^{uds} makes Eq. (5.60) grossly inexact.

To circumvent this issue, inspired by the demonstration of Appendix A.2, we introduce a function f_2 defined as a linear combination of Γ_1^{qq} and Γ_3^{qq} by

$$f_2(\mu^2,\mu_0^2) = \frac{1}{\epsilon}\Gamma_3^{qq}(\mu^2,\mu_0^2) - \frac{1-\epsilon}{\epsilon}\Gamma_1^{qq}(\mu^2,\mu_0^2) \tag{5.62}$$

for a given ϵ. Replacing Γ_3^{qq} by f_2, we can propose an alternative model of the subtraction constant – which is completely equivalent to the one of Eq. (5.59) since f_2 is a linear combination of Γ_1^{qq} and Γ_3^{qq} – as

$$C_{H,0}[D^q,D^g](t,Q^2) = \frac{8}{3}\left(\alpha\Gamma_1^{qq}(Q^2,\mu_0^2)+\beta f_2(Q^2,\mu_0^2)\right). \tag{5.63}$$

6. This difference might be defined in many different ways, we are merely interested in its general scaling properties.

7. Similarity of the functional dependences of the model is not the only cause of strong correlations among fitted parameters, another obvious cause being strong correlations in the fitted data themselves. We will come back to this aspect later, but let us consider so far – as it is the case in our study – that most of the correlation and inflation of uncertainty comes indeed from the functional dependences of the model.

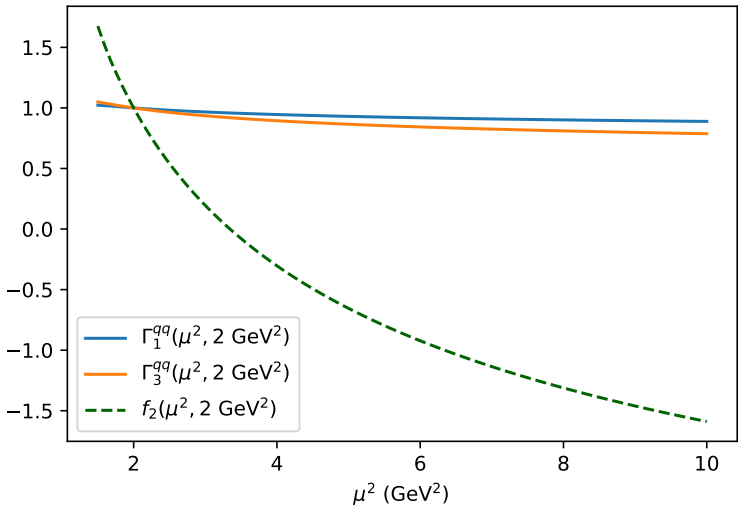

Figure 5.6 – Values of $\Gamma_1^{qq}(\mu^2, 2\ \text{GeV}^2)$, $\Gamma_3^{qq}(\mu^2, 2\ \text{GeV}^2)$ and $f_2(\mu^2, 2\ \text{GeV}^2)$ corresponding to $\epsilon = 0.041$. We show an interval in μ^2 equal to [1.5 GeV^2, 10 GeV^2].

By replacing f_2 by its value (5.62) and comparing with Eq. (5.59), one obtains that

$$\alpha = d_1^{uds}(\mu_0^2) + (1-\epsilon) d_3^{uds}(\mu_0^2)\,, \tag{5.64}$$

$$\beta = \epsilon d_3^{uds}(\mu_0^2)\,. \tag{5.65}$$

One may see therefore the motivation for the peculiar choice of f_2. The fact that Γ_1^{qq} and Γ_3^{qq} only differ by ϵ triggers an inflation of the uncertainty of fitted parameters by $1/\epsilon$ and causes the fact that $d_1^{uds} \approx -d_3^{uds}$. We are specifically trying to counter this effect, by finding a value of ϵ so as to obtain a robust extraction of α and β, that is precisely $\text{cov}(\alpha, \beta) = 0$. Since

$$\text{cov}(\alpha, \beta) = \epsilon\, \text{cov}(d_1^{uds}(\mu_0^2), d_3^{uds}(\mu_0^2)) + \epsilon(1-\epsilon)(\Delta d_3^{uds})^2\,, \tag{5.66}$$

solving $\text{cov}(\alpha, \beta) = 0$ gives either $\epsilon = 0$ or

$$\epsilon = 1 + \frac{\text{cov}(d_1^{uds}(\mu_0^2), d_3^{uds}(\mu_0^2))}{(\Delta d_3^{uds})^2} = 1 + \text{corr}(d_1^{uds}(\mu_0^2), d_3^{uds}(\mu_0^2)) \frac{\Delta d_1^{uds}}{\Delta d_3^{uds}}\,. \tag{5.67}$$

The numerical application with the values of our phenomenological fit produces $\epsilon = 0.041$. The other solution $\epsilon = 0$ corresponds to f_2 infinite. It is a trivial solution of no interest to us since robustness of the extraction of (α, β) is only obtained at the cost that $\beta = 0$ systematically. It corresponds actually to a fit with just the first parameter. On the contrary, $\epsilon = 0.041$ allows us to transform Γ_3^{qq} into a new function f_2 shown on Fig. 5.6 which is now clearly different from Γ_1^{qq}, such that the joint fit of f_2 and Γ_1^{qq} on experimental data produces uncorrelated uncertainties, contrary to the almost fully correlated results obtained when using the initial Γ_3^{qq}. The values of the standard deviations of α and β, noted $\Delta\alpha$ and $\Delta\beta$, contain in a reliable way information on the actual uncertainty of the subtraction constant, unlike $(\Delta d_1^{uds}, \Delta d_3^{uds})$ which contains mostly information about the close proximity of Γ_1^{qq} and Γ_3^{qq}. We can use the equivalent of Eq. (5.60) where we have replaced Γ_3^{qq} by f_2, yielding

$$(\Delta\alpha)^2 (\Gamma_1^{qq}(Q^2, \mu_0^2))^2 + (\Delta\beta)^2 f_2^2(Q^2, \mu_0^2) = (3/8)^2 \sigma_{C_H}^2(Q^2)\,, \tag{5.68}$$

and this time, we have indeed the bound

$$\Delta\beta \leq \frac{3}{8} \inf \frac{\sigma_{C_H}(Q^2)}{|f_2(Q^2,\mu_0^2)|} \,. \tag{5.69}$$

This bound on the parameter β which has no physical meaning is converted into a bound on $d_3^{uds}(\mu_0^2) = \beta/\epsilon$ (5.65) by

$$\Delta d_3^{uds} = \frac{\Delta\beta}{|\epsilon|} \,. \tag{5.70}$$

Noting that using the definition of f_2 (5.62)

$$\frac{1}{\epsilon} = \frac{f_2(Q^2,\mu_0^2) - \Gamma_1^{qq}(Q^2,\mu_0^2)}{\Gamma_3^{qq}(Q^2,\mu_0^2) - \Gamma_1^{qq}(Q^2,\mu_0^2)} \,, \tag{5.71}$$

we obtain

$$\Delta d_3^{uds} = \Delta\beta \left| \frac{f_2(Q^2,\mu_0^2)}{\Gamma_1^{qq}(Q^2,\mu_0^2)} - 1 \right| \left| 1 - \frac{\Gamma_3^{qq}(Q^2,\mu_0^2)}{\Gamma_1^{qq}(Q^2,\mu_0^2)} \right|^{-1} \,, \tag{5.72}$$

$$\leq \frac{3}{8} \left(\inf \frac{\sigma_{C_H}(Q^2)}{|f_2(Q^2,\mu_0^2)|} \right) \left| \frac{f_2(Q^2,\mu_0^2)}{\Gamma_1^{qq}(Q^2,\mu_0^2)} - 1 \right| \left| 1 - \frac{\Gamma_3^{qq}(Q^2,\mu_0^2)}{\Gamma_1^{qq}(Q^2,\mu_0^2)} \right|^{-1} \,. \tag{5.73}$$

This expression can be divided in two parts : the first two factors depend on the actual experimental uncertainty σ_{C_H} and the set of "ideal" functions (Γ_1^{qq}, f_2), whereas the final factor measures the proximity of Γ_1^{qq} and Γ_3^{qq} and produces the observed correlations. The structure of the first two factors with f_2 both on the numerator and denominator and knowing that $\Gamma_1^{qq} \approx 1$, guarantees that it will remain essentially of the order σ_{C_H} on the entire available range in Q^2. Indeed,

— if $|f_2(Q^2,\mu_0^2)| \ll \Gamma_1^{qq}(Q^2,\mu_0^2)$, since $f_2(\mu_0^2,\mu_0^2) = 1$,

$$\left(\inf \frac{\sigma_{C_H}(Q^2)}{|f_2(Q^2,\mu_0^2)|} \right) \left| \frac{f_2(Q^2,\mu_0^2)}{\Gamma_1^{qq}(Q^2,\mu_0^2)} - 1 \right| \lesssim \sigma_{C_H}(\mu_0^2) \,, \tag{5.74}$$

— if on the contrary, $|f_2(Q^2,\mu_0^2)| \gg \Gamma_1^{qq}(Q^2,\mu_0^2)$,

$$\left(\inf \frac{\sigma_{C_H}(Q^2)}{|f_2(Q^2,\mu_0^2)|} \right) \left| \frac{f_2(Q^2,\mu_0^2)}{\Gamma_1^{qq}(Q^2,\mu_0^2)} - 1 \right| \lesssim \frac{\sigma_{C_H}(Q^2)}{|f_2(Q^2,\mu_0^2)|} \left| \frac{f_2(Q^2,\mu_0^2)}{\Gamma_1^{qq}(Q^2,\mu_0^2)} \right| \approx \sigma_{C_H}(Q^2) \,, \tag{5.75}$$

— finally, if $|f_2(Q^2,\mu_0^2)|$ is of the order of $\Gamma_1^{qq}(Q^2,\mu_0^2)$, the result is of the order of $\sigma_{C_H}(Q^2)$. It is precisely the case in our study.

Therefore, we can build an estimate of the uncertainty on $d_3^{uds}(\mu_0^2)$ when fitted on experimental data spanning over a range of $Q^2 \in [\mu_0^2, \mu^2]$ by introducing a factor K of the order of 1 and writing

$$\Delta d_3^{uds} \approx K \sigma_{C_H} \left| 1 - \frac{\Gamma_3^{qq}(Q^2,\mu_0^2)}{\Gamma_1^{qq}(Q^2,\mu_0^2)} \right|^{-1} \,. \tag{5.76}$$

σ_{C_H} is a typical value of the experimental uncertainty of the subtraction constant on the experimental range in Q^2. As discussed previously, if the final term verifies

$$\left| 1 - \frac{\Gamma_3^{qq}(Q^2,\mu_0^2)}{\Gamma_1^{qq}(Q^2,\mu_0^2)} \right|^{-1} \gg 1 \,, \tag{5.77}$$

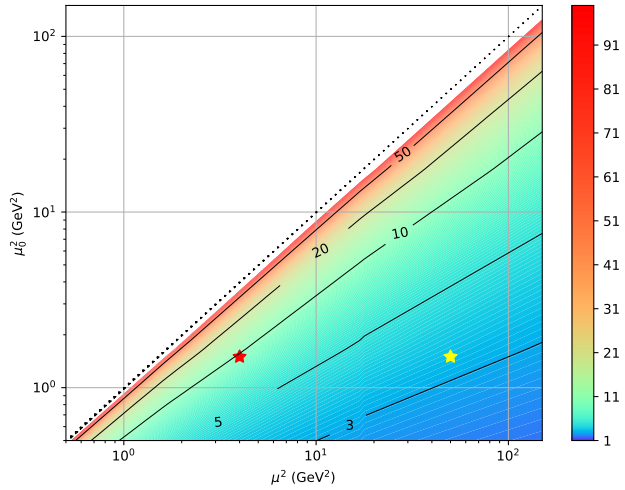

Figure 5.7 – Estimate of the uncertainty on $d_3^{uds}(\mu_0^2)/\sigma_{C_H}$ given by $0.75/(1 - \Gamma_3^{qq}(Q^2,\mu_0^2)/\Gamma_1^{qq}(Q^2,\mu_0^2))$ for experimental data spanning a range $Q^2 \in [\mu_0^2, \mu^2]$. The red star shows the result for $\mu_0^2 = 1.5$ GeV^2 and $\mu^2 = 4$ GeV^2, which is relevant for the world DVCS dataset included in the current neural network study. The estimate gives an uncertainty on d_3^{uds}/σ_{C_H} of 9.5, so a total uncertainty on d_3^{uds} of 28, which is consistent with our phenomenological fit. The yellow star shows an hypothetical EIC configuration, with $\mu^2 = 50$ GeV^2. The uncertainty on d_3^{uds}/σ_{C_H} is expected to drop at 3.4.

we can deduce than uncertainty on d_1^{uds} in the joint fit case is mostly driven by the close proximity of Γ_1^{qq} and Γ_3^{qq} rather than by the precise characteristics of the fitted subtraction constant.

We choose $K = 3/4$ for numerical applications, motivated by the fact that in our extraction shown on Fig. 5.6, a typical order of magnitude of f_2 is -1. Therefore, $3/8|f_2/\Gamma_1^{qq} - 1| \approx 3/8 * 2 = 3/4$. Then the estimate gives an uncertainty on $d_3^{uds}(\mu_0^2)$ of 28 for $\mu_0^2 = 1.5$ GeV^2, $\mu^2 = 4$ GeV^2 and $\sigma_{C_H} = 3$ (see the second panel of Fig. 5.2 at $t = 0$). The values of μ_0^2 and μ^2 are chosen since the vast majority of DVCS measurements included in the neural network fit take values in that range for Q^2. A few measurements go as high as $Q^2 = 10$ GeV^2, but they do not represent a strong statistical constraint to the fit. The result is largely compatible with our own experimental extraction $d_3^{uds}(2\text{ GeV}^2) = -11 \pm 26$. We show on Fig. 5.7 the expected uncertainty on $d_3^{uds}(\mu_0^2)$ normalised by σ_{C_H} depending on the range of evolution $[\mu_0^2, \mu^2]$ where experimental data are available.

Therefore, without any increase in the precision of subtraction constant data, but an increase in the available range in Q^2 up to 50 GeV^2 as could be obtained in the EIC, this quick estimate predicts a reduction of uncertainty on d_3^{uds} from 28 to 10, so about a factor 3. As the uncertainty remains three times larger than σ_{C_H}, it is foreseeable that d_1 and d_3 will remain mostly driven by the correlation between the two, but a significant improvement has been achieved compared to the current situation.

Conclusion

Under the assumptions of simplified evolution – no heavy flavour and negligible radiative gluons contributions –, we showed that it is possible to approximate the uncertainty

on d_3^{uds} by the estimate

$$K\sigma_{C_H}\left|1-\frac{\Gamma_3^{qq}(Q^2,\mu_0^2)}{\Gamma_1^{qq}(Q^2,\mu_0^2)}\right|^{-1}. \tag{5.78}$$

σ_{C_H} is the typical experimental uncertainty of the subtraction constant data on the probed range in Q^2, and $K = 3/4$ gives excellent results on actual data. This estimate allows us to simultaneously gauge the impact of increasing the range in Q^2 of available data, and the impact of adding new functional forms to the estimation of the d_n^{uds} terms for a given range in energy. The results for EIC kinematics look promising, with an expected reduction of uncertainty on (d_1^{uds}, d_3^{uds}) by a factor 3 when only considering the increase in range of Q^2, without even taking into account that experimental data themselves will increase in precision.

5.2.3 . Shadow D-terms

We have only considered the deconvolution problem in terms of similarity between functional forms of evolution operators so far. This approach is practical in the context of GFF extraction from the DVCS subtraction constant since we were only interested in a quantity, d_1, whose LO evolution could be very conveniently written thanks to the Gegenbauer expansion. However, higher order evolution of the D-term, as well as the evolution of full GPDs in the context of the general deconvolution problem are much more difficult to represent. Although conformal moments of GPDs allow a relatively similar representation of the effect of LO evolution compared to what we have described here, the reconstruction of the full GPD from its conformal moments is a rather technical procedure which requires an analytical form for the moments. We will therefore approach the full deconvolution problem of GPDs with a more intuitive and powerful tool : shadow distributions. Let us present briefly the application of this concept to the D-term extraction, before presenting more systematically shadow GPDs in the next chapter.

> **Shadow D-term** A shadow D-term at N^{k}LO and (t,μ_F^2) is a set of functions $D_0^a(t,\mu^2)$ for $a = u,d,s,g,...$ following the same evolution equations that traditional D-terms such that the subtraction constant computed from D_0^a at order N^{k}LO and (t,μ_F^2) noted $C_{H,k}[D_0^a](t,\mu_F^2)$ is exactly vanishing.

Since the LO subtraction constant writes (5.26)

$$C_{H,0}[D^q,D^g](t,\mu_F^2) = 4\sum_q e_q^2 \sum_{\text{odd } n} d_n^q(t,\mu_F^2)\,, \tag{5.79}$$

defining a LO shadow D-term amounts to finding a set of $d_n^q(t,\mu_F^2)$ such that

$$\sum_q e_q^2 \sum_{\text{odd } n} d_n^q(t,\mu_F^2) = 0\,. \tag{5.80}$$

The simplest solution is to choose all light flavours equal and $d_1^{uds}(t,\mu_F^2) = -d_3^{uds}(t,\mu_F^2)$ with $d_n^{uds} = 0$ if $n \geq 5$. Therefore, the simplest LO shadow D-term is defined as

$$D_0^{uds}(z,t,\mu_F^2) = (1-z^2)d_1^{uds}(t,\mu_F^2)\left(C_1^{(3/2)}(z) - C_3^{(3/2)}(z)\right), \tag{5.81}$$

$$= \frac{7z(1-z^2)(3-5z^2)}{2}\, d_1^{uds}(t,\mu_F^2)\,. \tag{5.82}$$

But we could as well define more complicated LO shadow D-terms by choosing $d_1^{uds}(t,\mu_F^2) = d_3^{uds}(t,\mu_F^2)$ and $d_5^{uds}(t,\mu_F^2) = -2\,d_1^{uds}(t,\mu_F^2)$ for instance, which would yield

$$D_0^{uds}(z,t,\mu_F^2) = (1-z^2)d_1^{uds}(t,\mu_F^2)\left(C_1^{(3/2)}(z) + C_3^{(3/2)}(z) - 2C_5^{(3/2)}(z)\right), \tag{5.83}$$

$$= -\frac{z(1-z^2)(123-700z^2+693z^4)}{4}\,d_1^{uds}(t,\mu_F^2)\,. \tag{5.84}$$

Any linear combination of the two previous functions is still a LO shadow D-term, since its contribution to the subtraction constant remains 0. Therefore, we can build a vector space of z-dependent functions which bring no contributions to the subtraction constant at (t,μ_F^2), and only contributions of order $\alpha_s(\mu^2)$ for $\mu^2 \neq \mu_F^2$.

This vector space provides a practical way to assess the uncertainty of a deconvolution procedure. For instance, the extraction of the D-term from subtraction constant data can be represented by $D^a_{best} + \mathcal{D}^a$ where D^a_{best} is the best fit value obtained for a linear model [8] optimised on the central value of experimental data, and $\mathcal{D}^a$ is an element of the allowed functional space spanned by the model such that $|C_{H,k}[\mathcal{D}^a](t,\mu^2)|$ is of the order of the experimental uncertainty for all experimental kinematics (t,μ^2). The diversity of all possible $\mathcal{D}^a$ gives an account of the uncertainty of the extraction. Provided that the functional space spanned by the model intersects the vector space of shadow distributions, the latter form a natural way to determine some of these $\mathcal{D}^a$ and obtain a measure of the uncertainty of extraction.

For instance, let us go back again to our fit with free d_1^{uds} and d_3^{uds}. As we have just discussed before, there is a vector space of dimension one of shadow D-terms which can be obtained with such parameterisation, corresponding to $d_1^{uds}(t,\mu_F^2) = -d_3^{uds}(t,\mu_F^2)$. Let's assume the experimental uncertainty is of the order of σ_{C_H} on the range $[\mu_0^2,\mu^2]$. A measure of the uncertainty on the D-term can be obtained by choosing $\mu_F^2 = \mu_0^2$ and exploring the maximal value of the shadow D-terms such that $|C_{H,k}[\mathcal{D}_0^a](t,\mu^2)| \leq \sigma_{C_H}$. Keeping all our previous assumptions on the simplified evolution behaviour,

$$|C_{H,k}[\mathcal{D}_0^a](t,\mu^2)| = \frac{8}{3}\,|d_3^{uds}(t,\mu_0^2)|\,\left|\Gamma_1^{qq}(Q^2,\mu_0^2) - \Gamma_3^{qq}(Q^2,\mu_0^2)\right|\,, \tag{5.85}$$

and enforcing that $|C_{H,k}[\mathcal{D}_0^a](t,\mu^2)| \leq \sigma_{C_H}$ gives

$$|d_3^{uds}(t,\mu_0^2)| \leq \frac{3}{8}\frac{\sigma_{C_H}}{\Gamma_1^{qq}(Q^2,\mu_0^2)}\left(1 - \frac{\Gamma_3^{qq}(Q^2,\mu_0^2)}{\Gamma_1^{qq}(Q^2,\mu_0^2)}\right)^{-1}\,. \tag{5.86}$$

We retrieve an estimate with exactly the same general behaviour that the one we derived in the previous section [9]. Note that even if we have no precise knowledge of the Γ_n^{qq}

8. What is important is that $D^a_{best} + \mathcal{D}^a$ belongs to the functional space spanned by the model – otherwise it is of course not a relevant object to determine the result of the deconvolution procedure with the model under consideration. This condition is fulfilled immediately if the model is linear and D^a_{best} and $\mathcal{D}^a$ each belong to its functional space, which is exactly the case we have faced so far.

9. What is then exactly the difference between this procedure and the previous estimate of Eq. (5.78) ? The latter is more general, since it does not make reference to the evolution of any specific function, but is based on the scale dependence of evolution operators themselves. For instance, all the derivations of the previous section remain true even if there are no shadow D-terms in the functional space spanned by the chosen model for the D-term. The new procedure estimates the uncertainty specifically in functional spaces where shadow D-terms can be found by measuring the growth of their contributions to the convolution. Although less general, the new method is also far more manageable, especially when evolution is only performed numerically.

operators, but simply a numerical procedure that evolves directly the shadow D-term as a function of z, we are still able to determine the maximal size of shadow D-terms which obey $|C_{H,k}[\mathcal{D}_0^a](t,\mu^2)| \leq \sigma_{C_H}$.

5.2.4 . Enriching the Ansatz

Free gluon contribution

Considering the results of allowing the simultaneous fit of d_1^{uds} and d_3^{uds}, it does not seem useful to increase the number of values of n allowed. However, other hypotheses of our Ansatz can be relaxed. From now on, we return to only allowing the $n = 1$ contribution. Let us fit d_1^g as a free parameter. We obtain at $\mu_0^2 = 2$ GeV^2

Parameter	Value
$d_1^{uds}(\mu_0^2)$	-0.7 ± 1.2
$d_1^g(\mu_0^2)$	51 ± 111
$d_1^c(\mu_0^2)$	0.2 ± 0.4

Compared to the fit of d_1^{uds} as sole free parameter where we obtained $d_1^{uds}(\mu_0^2) = -0.5 \pm 1.2$, the result has barely changed. On the contrary, the gluon uncertainty explodes, from -0.6 ± 1.6 to 51 ± 111. To explain this behaviour, we notice that the contribution of gluons to the quark terms through evolution is particularly small : the explicit solution of D-term evolution derived in Eq. (5.17) shows that the contribution of $d_1^g(\mu_0^2)$ to $d_1^{uds}(\mu^2)$ is suppressed by a factor

$$\Gamma_1^{qg}(\mu^2,\mu_0^2) = \frac{\Gamma_1^-(\mu^2,\mu_0^2) - \Gamma_1^+(\mu^2,\mu_0^2)}{a_1^- - a_1^+}\,. \tag{5.87}$$

Numerically, for the typical range from $\mu_0^2 = 1.5$ GeV^2 to $\mu^2 = 4$ GeV^2, this factor is only of the order of 0.015, while for the same evolution range, $\Gamma_1^{qq} \approx 0.92$, $\Gamma_1^{gq} \approx 0.31$ and $\Gamma_1^{gg} \approx 0.94$. With d_1^g and d_1^{uds} as free parameters and neglecting heavy flavours, our model of the subtraction constant reads

$$C_{H,0}[D^q, D^g](Q^2) = \frac{8}{3}\left(\Gamma_1^{qq}(Q^2,\mu_0^2) d_1^{uds}(\mu_0^2) + \Gamma_1^{qg}(Q^2,\mu_0^2) d_1^g(\mu_0^2)\right)\left(1 - \frac{t}{\Lambda^2}\right)^{-\alpha}\,. \tag{5.88}$$

Since the factor in front of $d_1^g(\mu_0^2)$ is at least 60 times smaller than the one in front of $d_1^{uds}(\mu_0^2)$, it is not a surprise that the resulting distribution of d_1^g is overall much wider than that of d_1^{uds} when d_1^g is a free parameter. On the contrary, for radiatively generated gluon contributions at threshold 0.1 GeV^2, that is when $d_1^g(0.1\text{ GeV}^2) = 0$ is imposed verify

$$\begin{aligned} d_1^g(2\text{ GeV}^2) &= \Gamma_1^{gq}(2\text{ GeV}^2, 0.1\text{ GeV}^2)\, d_1^{uds}(0.1\text{ GeV}^2)\,, &(5.89)\\ &= \frac{\Gamma_1^{gq}(2\text{ GeV}^2, 0.1\text{ GeV}^2)}{\Gamma_1^{qq}(2\text{ GeV}^2, 0.1\text{ GeV}^2)}\, d_1^{uds}(2\text{ GeV}^2)\,, &(5.90)\\ &\approx 1.4\, d_1^{uds}(2\text{ GeV}^2)\,. &(5.91) \end{aligned}$$

Therefore radiatively generated d_1^g is just slightly larger than d_1^{uds}, and the gluon indirect contribution to the LO subtraction constant is more than $60/1.4 \approx 40$ times smaller than the direct quark one. There is therefore virtually no sensitivity to gluon with radiative

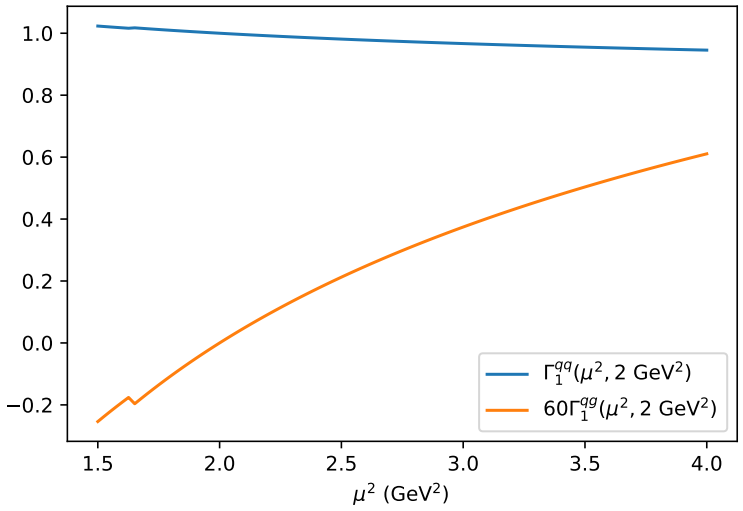

Figure 5.8 – Values of $\Gamma_1^{qq}(\mu^2, 2\ \text{GeV}^2)$ and $60\Gamma^{qg}(\mu^2, 2\ \text{GeV}^2)$ over the relevant experimental range [1.5 GeV^2, 4 GeV^2] to our study. The dent on the curve of Γ_1^{qg} represents the threshold of charm activation. Although we neglect the d_1^c term in the numerical applications, we take into account the change of number of active flavours in the definition of α_s. It brings however only very little difference over all.

generation in a LO subtraction constant study, explaining the remarkable independence of d_1^{uds} on the gluon generation threshold. We had previously observed this fact on Fig. 5.4, and used it as a tool for the simplification of evolution with radiative gluon generation.

Although this explains why free gluon contributions are considerably larger than free quark contributions in a LO analysis, it does not justify why the two are essentially uncorrelated. Fig. 5.8 shows the values of Γ_1^{qq} and $60\Gamma_1^{qg}$ over the relevant experimental range. The two functions exhibit significantly different shapes. This is a strong argument for the decorrelation of $d_1^{uds}(2\ \text{GeV}^2)$ and $d_1^g(2\ \text{GeV}^2)$ following the discussion of Section 5.2.2 [10]

t-dependent Ansatz

In all the previous fits, the parameters α and Λ of the assumed factorised t-dependence have remained fixed at values $\alpha = 3$ and $\Lambda = 0.8$ GeV. Allowing one or the other to be

10. Formally, as we remarked on a previous note, correlations can also emerge from the poor information content of replicas of the subtraction constant themselves. Borrowing the notations of Section 4.1.1, the covariance matrix of the best fit parameters in a linear fit is obtained as $(C^T\Omega^{-1}C)^{-1}$ where C is the matrix of the functional dependences of the model on the probed kinematics, and Ω the covariance matrix of the data. Large correlations in the fitted parameters can emerge from the poor conditioning of either C – which corresponds precisely to the idea which we have already explored that the functional dependences in the model are too similar – or Ω. For instance, if the fitted dataset has been generated from a linear model with r parameters, then Ω is of rank r. It is easy to show that as a consequence, the matrix of fitted parameters is at most of rank r too. This means that even if a totally different linear model with n parameters is fitted on the dataset, perfect correlations between parameters will emerge as soon as $n > r$. However, considering that we perform our fit on non-linear replicas of a neural network with great flexibility, and that we only try to fit a handful of parameters at a time, the issue of insufficient information in the fitted data does not emerge. It is then sufficient to observe that the functional forms of the model are strongly different.

fitted jointly with d_1^{uds} produces however very inconsistent results. The fit is now no longer linear in its parameters, which exposes us to the risk of unexpected behaviour as for instance the separation of replicas into clusters that we highlighted in Section 4.1.1. Overall, no significant trend emerges from the fit. This does not come completely as a surprise, since we know that the experimental coverage in t is one of the weak points of the available data. Let us note however that the limited coverage in t is partly a consequence of the imposed kinematic cut $Q^2 > 1.5\ \mathrm{GeV}^2$ we used to select fitted DVCS experiments. This cut is meant to protect the dataset from higher twist corrections, which are suppressed by a factor t/Q^2. Taking into account higher twist corrections could in principle allow us to incorporate more data and extend the range of t.

From a phenomenological point of view, the t-dependence is crucial since it drives the energy and pressure profiles computed from the GFF. We remind that the pressure anisotropy $s(r)$ can be defined in the Breit frame (see Section 1.4.2, Eq. (1.33)) by

$$s_a(r) = -\frac{4M}{r^2}\int \frac{\mathrm{d}^3\boldsymbol{\Delta}}{(2\pi)^3}\, e^{-i\boldsymbol{\Delta}\cdot\boldsymbol{r}}\, \frac{t^{-1/2}}{M^2}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\left[t^{5/2}\, C^a(t)\right]\,, \tag{5.92}$$

where $t = -||\boldsymbol{\Delta}||^2$. Since $C^a(t) = d_1^a(t)/5$, fixing the values of α and Λ corresponds to entirely fixing the r profile of the pressure anisotropy, and relegating the result of the fit of d_1^a to a simple normalisation factor.

Since the integral depends only on the norm of $\boldsymbol{r}$, we are free to choose any representative for the vector. Let us consider $\boldsymbol{r}$ and $\boldsymbol{\Delta}$ with respective spherical coordinates $(r, 0, 0)$ and (Δ, θ, ϕ) with the radius-colatitude-longitude convention. Then $\boldsymbol{\Delta}\cdot\boldsymbol{r} = r\Delta\cos\theta$, and the differential element $\mathrm{d}^3\boldsymbol{\Delta}$ gives $\Delta^2\sin\theta\mathrm{d}\Delta\mathrm{d}\theta\mathrm{d}\phi$, so

$$s_a(r) = -\frac{4}{5Mr^2(2\pi)^3}\int_0^{2\pi}\mathrm{d}\phi\int_0^{\pi}\mathrm{d}\theta\int_0^{\infty}\mathrm{d}\Delta\,\Delta^2\sin\theta\, e^{-i\Delta r\cos\theta}\, t^{-1/2}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(t^{5/2}d_1^a(t)\right)\,. \tag{5.93}$$

Using

$$\int_0^{\pi}\mathrm{d}\theta\ \sin\theta\, e^{-i\Delta r\cos\theta} = \frac{2\sin(\Delta r)}{\Delta r}\,, \tag{5.94}$$

gives

$$s_a(r) = -\frac{2}{5Mr^3\pi^2}\int_0^{\infty}\mathrm{d}\Delta\,\Delta\sin(\Delta r)\, t^{-1/2}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(t^{5/2}d_1^a(t)\right)\,. \tag{5.95}$$

Assuming

$$d_1^a(t) = d_1^a\left(1-\frac{t}{\Lambda^2}\right)^{-\alpha}\,, \tag{5.96}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(t^{5/2}d_1^a(t)\right) = t^{1/2}\frac{d_1^a(t)}{4(\Lambda^2-t)^2}\left[15\Lambda^4 + t(20\alpha-30)\Lambda^2 + t^2(4\alpha^2-16\alpha+15)\right]\,. \tag{5.97}$$

A closed form solution can be found for $\alpha = 3$:

$$s_a(r) = -\frac{d_1^a}{160M\pi}\Lambda^6 r e^{-\Lambda r}\,. \tag{5.98}$$

For other values, it does not seem that a simple analytical solution exists, but the integral can be evaluated numerically easily. We give on Fig. 5.9 a representation of the pressure profile obtained for various combinations of α and Λ. We observe that even moderate

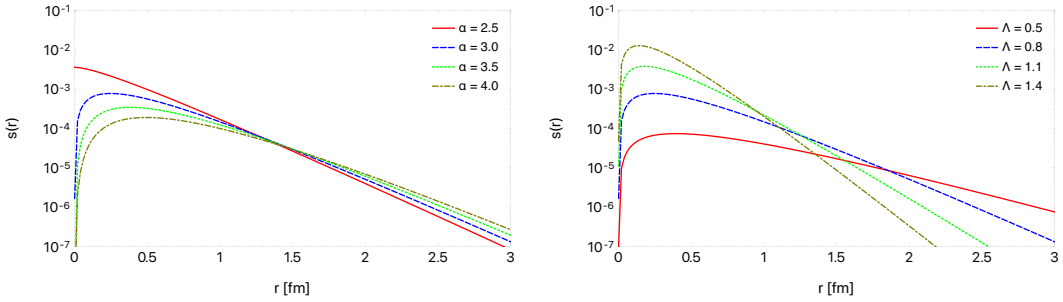

FIGURE 5.9 – Profile of pressure anisotropies $s_a(r)$ obtained for various values of the parameters of the t dependent Ansatz. (left) for $\Lambda = 0.8$ GeV, and various values of α. (right) for $\alpha = 3$ and various values of Λ. Figure taken from Ref. [4].

variations of these two parameters produce drastic impact on the value of the maximum of the curve and its tail behaviour. Considering the weak sensitivity on these parameters achieved with experimental data, a serious caution should be exerted when trying to interpret any physical value emerging from such analyses.

5.3 . Effect of next-to-leading order

5.3.1 . Subtraction constant at NLO

So far, we have only used the LO expression of the subtraction constant (5.2) which is expressed at order 0 in α_s. However, the use of LO evolution equations means that we have already included implicit corrections of order 1 in α_s in our fits. Indeed, the $\Gamma_n^{\pm}$ and Γ_n^0 functions defined in Eq. (5.7)-(5.8) which drive the evolution of the Gegenbauer coefficients of the D-term satisfy

$$\Gamma(\mu^2, \mu_0^2) = \left(\frac{\alpha_s(\mu^2)}{\alpha_s(\mu_0^2)}\right)^{2\gamma/\beta_0}, \tag{5.99}$$

$$= 1 + \frac{2\gamma}{\beta_0}\left[\frac{\alpha_s(\mu^2) - \alpha_s(\mu_0^2)}{\alpha_s(\mu_0^2)}\right] + \dots, \tag{5.100}$$

$$= 1 + \frac{2\gamma}{\beta_0}\frac{1}{\alpha_s(\mu_0^2)} \log\left(\frac{\mu^2}{\mu_0^2}\right) \frac{\partial \alpha_s}{\partial \log \mu^2} + \dots, \tag{5.101}$$

and using the LO solution of the renormalisation group equations (2.53)

$$\frac{\partial \alpha_s}{\partial \log \mu^2} = -\frac{\beta_0}{4\pi}\alpha_s^2, \tag{5.102}$$

so

$$\Gamma(\mu^2, \mu_0^2) = 1 + \gamma \frac{\alpha_s(\mu^2)}{2\pi} \log\left(\frac{\mu_0^2}{\mu^2}\right) + \mathcal{O}(\alpha_s^2). \tag{5.103}$$

Since scale dependence of the order of $\mathcal{O}(\alpha_s)$ has been included in our previous analysis thanks to the effect of evolution, it is interesting for completeness to also include corrections of the same order in the DVCS coefficient function. Following the relations of Eq. (3.27), the definition of the subtraction constant itself will be changed to include

corrections of order $\mathcal{O}(\alpha_s)$. We provide in Appendix A.3 an example of computation of the contribution of the NLO DVCS coefficient functions to the subtraction constant. We present here the resulting relation between the D-term and the subtraction constant at NLO

$$C^q_{H,1}[D^q](t,Q^2) = 2e_q^2 \int_{-1}^{1} \mathrm{d}z \, \frac{D^q(z,t,Q^2)}{1-z} + \alpha_s(Q^2)\frac{e_q^2 C_F}{2\pi} \int_{-1}^{1} \mathrm{d}z \, \frac{D^q(z,t,Q^2)}{1-z} \left[9 + (\log 2)^2 + \log(1-z)\left(-3 + \frac{6}{1+z} - \log(1-z) \right) \right], \tag{5.104}$$

$$C^g_{H,1}[D^g](t,Q^2) = \alpha_s(Q^2)\frac{e_q^2 T_F}{2\pi} \int_{-1}^{1} \mathrm{d}z \, \frac{D^g(z,t,Q^2)}{(1-z)^2} \log\left(\frac{1+z}{2}\right) \left[2 - \frac{3+z}{1+z} \right], \tag{5.105}$$

where the assumption that $Q^2 = \mu^2$ eliminates the $\log(Q^2/\mu^2)$ contributions. Contrary to the LO case where a simple closed form could be found in terms of Gegenbauer coefficients, we are forced to resort to approximate values. Truncation to $n = 3$ gives :

$$C_{H,1}[D^q, D^g](t,Q^2) \approx \sum_q 4e_q^2 \Bigg[\left(1 + 0.060\alpha_s(Q^2)\frac{C_F}{2\pi} \right) d_1^q(t,Q^2) + \left(1 + 3.33\alpha_s(Q^2)\frac{C_F}{2\pi} \right) d_3^q(t,Q^2) \Bigg] + \alpha_s(Q^2) n_e \frac{T_F}{2\pi} \Big[-9.56\, d_1^g(t,Q^2) - 11.1\, d_3^g(t,Q^2) \Big]. \tag{5.106}$$

Notice that contrary to the LO case, we now benefit from a direct sensitivity to gluon contributions in the subtraction constant.

5.3.2 . Results of the neural network fit

As in the LO case, we start by a fit with the simplest Ansatz with only d_1^{uds} as free parameter. The results shown here represent the first fit of the full NLO DVCS subtraction constant. The numerical values presented here will soon be submitted for publication. We recall the results of the LO + LO evolution fit in the first column, and then the result of the full NLO fit in the second at $\mu_0^2 = 2$ GeV^2 :

Parameter	LO + LO evolution fit	Full NLO fit
$d_1^{uds}(\mu_0^2)$	-0.5 ± 1.2	-0.5 ± 1.4
$d_1^c(\mu_0^2)$	-0.002 ± 0.005	-0.002 ± 0.006
$d_1^g(\mu_0^2)$	-0.6 ± 1.6	-0.7 ± 1.9

Addition of the NLO contributions makes very little difference compared to the LO results with LO evolution. With our treatment of heavy quarks, $\alpha_s(2\ \mathrm{GeV}^2) = 0.31$. Using n_e the sum of the squared charges of active quarks (5.28), the coefficient of $d_1^{uds}(2\ \mathrm{GeV}^2)$ changes from $4n_e$ at LO to $4.02n_e$ at NLO, and the coefficient of $d_1^g(2\ \mathrm{GeV}^2)$ from 0 to $-0.24n_e$. Although we now have a direct sensitivity to d_1^g, it is suppressed by a factor 20 approximatively compared to d_1^{uds}. Since gluons are radiatively generated with the same evolution as before, we have demonstrated in Eq. (5.91) that their ratio to the quark distribution is systematically of the order of 1.4 for a gluon threshold of 0.1 GeV^2. Therefore, radiative gluon sensitivity is still reduced by a factor of the order of 10 compared to quark sensitivity at NLO. This is nonetheless a significant improvement compared to the minimal factor of the order of 40 we identified at LO.

Allowing d_1^{uds} and d_3^{uds} as free parameters yields

Parameter	LO + LO evolution fit	Full NLO fit
$d_1^{uds}(\mu_0^2)$	11 ± 25	11 ± 23
$d_3^{uds}(\mu_0^2)$	-11 ± 26	-7.7 ± 17
$d_1^{g}(\mu_0^2)$	15 ± 34	15 ± 32
$d_3^{g}(\mu_0^2)$	-1.8 ± 3.9	-1.2 ± 2.6

We have already mentioned that the coefficient of $d_1^{uds}(2\ \text{GeV}^2)$ barely changes from $4n_e$ at LO to $4.02n_e$ at NLO. The coefficient of $d_3^{uds}(2\ \text{GeV}^2)$ on the other hand changes from $4n_e$ to $4.9n_e$, that is an increase of the order of 20%. We expect therefore d_3^{uds} to decrease by about 20% in the NLO study compared to the LO one. The slight variation of the Q^2 dependence introduced by the NLO $\alpha_s(Q^2)$ term mostly affects d_3^{uds}, and not d_1^{uds} because of their different respective factors. We show on the left panel of Fig. 5.10 the modified μ^2 dependence of the operators defined by

$$\widetilde{\Gamma}_1^{qq}(\mu^2, \mu_0^2) \propto \left(1 + 0.060\alpha_s(\mu^2)\frac{C_F}{2\pi}\right)\Gamma_1^{qq}(\mu^2, \mu_0^2)\,, \tag{5.107}$$

$$\widetilde{\Gamma}_3^{qq}(\mu^2, \mu_0^2) \propto \left(1 + 3.33\alpha_s(\mu^2)\frac{C_F}{2\pi}\right)\Gamma_3^{qq}(\mu^2, \mu_0^2)\,, \tag{5.108}$$

where we have included the μ^2 dependence introduced by the explicit dependence on α_s. Eq. (5.108) uses the $\propto$ symbol because we have rescaled each operator so that $\widetilde{\Gamma}_{1,3}^{qq}(\mu_0^2, \mu_0^2) = 1$ to eliminate the absolute difference size effect of 20% in favor of d_3^{uds}. As can be seen, the new μ^2 dependence has only slightly changed. It is possible to define an estimate of the uncertainty of the extraction of d_3^{uds} in exactly the same fashion as in Eq. (5.78), except we replace Γ_n^{qq} by the normalised $\widetilde{\Gamma}_n^{qq}$. The difference of the result, shown on the right panel of Fig. 5.10, is almost uniquely driven by the overall rescaling of d_3^{uds}. For $\mu_0^2 = 1.5\ \text{GeV}^2$, $\mu^2 = 4\ \text{GeV}^2$ and $\sigma_{C_H} = 3$, we obtain an estimation of the uncertainty of d_3^{uds} of 17 in the NLO case, which is what was found when actually performing the fit, compared to a prediction of 28 in the LO case.

The functional dependence of d_1^{uds} and d_3^{uds} remains largely similar, and the decorrelation effect is minute. As we did in the LO case, we can introduce NLO shadow D-terms. The situation is made slightly more complicated by the direct gluon sensitivity. We have to choose $d_1^{uds}(\mu_0^2)$ and $d_3^{uds}(\mu_0^2)$ such that the quark and gluon contributions cancel out each other in Eq. (5.106). This means that our shadow D-term will depend explicitly on the gluon radiative threshold. The NLO shadow D-term is obtained thanks to

$$\begin{aligned}\left[4\left(1 + 0.060\alpha_s(\mu_0^2)\frac{C_F}{2\pi}\right) - 9.56\alpha_s(\mu_0^2)\frac{T_F}{2\pi}\frac{\Gamma_1^{gq}(\mu_0^2, 0.1\ \text{GeV}^2)}{\Gamma_1^{qq}(\mu_0^2, 0.1\ \text{GeV}^2)}\right] d_1^{uds}(t, \mu_0^2) \\ = -\left[4\left(1 + 3.33\alpha_s(\mu_0^2)\frac{C_F}{2\pi}\right) - 11.1\alpha_s(\mu_0^2)\frac{T_F}{2\pi}\frac{\Gamma_1^{gq}(\mu_0^2, 0.1\ \text{GeV}^2)}{\Gamma_1^{qq}(\mu_0^2, 0.1\ \text{GeV}^2)}\right] d_3^{uds}(t, \mu_0^2)\,,\end{aligned} \tag{5.109}$$

where we have replaced $d_{1,3}^{g}$ by their radiatively generated values as in Eq. (5.91). This gives at $\mu_0^2 = 2\ \text{GeV}^2$

$$d_1^{uds}(t, \mu_0^2) = -1.4 d_3^{uds}(t, \mu_0^2)\,, \tag{5.110}$$

which is slightly different from the LO shadow D-term $d_1^{uds}(t, \mu_0^2) = -d_3^{uds}(t, \mu_0^2)$. Again, the shadow D-term is consistent with the obtained results, since $\sigma_{d_1} \approx 1.4\sigma_{d_3}$.

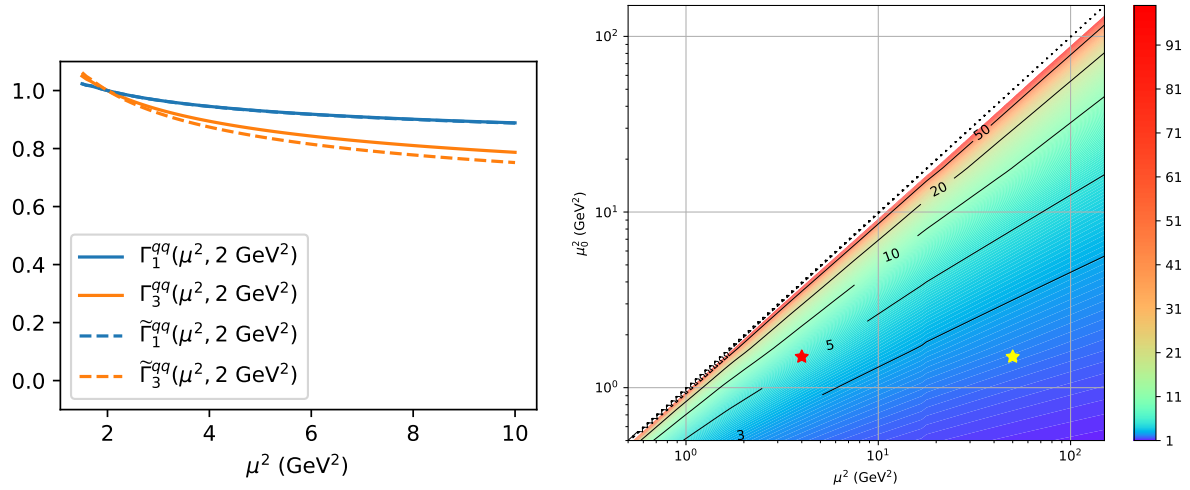

Figure 5.10 – (left) Values of $\widetilde{\Gamma}_1^{qq}(\mu^2, 2\ \text{GeV}^2)$ and $\widetilde{\Gamma}_3^{qq}(\mu^2, 2\ \text{GeV}^2)$ as defined in Eq. (5.108) to take into account the slight modification introduced by NLO contributions, compared to the original $\Gamma_1^{qq}(\mu^2, 2\ \text{GeV}^2)$ and $\Gamma_3^{qq}(\mu^2, 2\ \text{GeV}^2)$. – (right) estimate of the uncertainty of d_3^{uds}/σ_{C_H} built with the NLO operator.

Finally, allowing d_1^{uds} and d_1^g as free parameters gives

Parameter	LO + LO evolution fit	Full NLO fit
$d_1^{uds}(\mu_0^2)$	-0.7 ± 1.2	0.4 ± 2.8
$d_1^g(\mu_0^2)$	51 ± 111	5.3 ± 19
$d_1^c(\mu_0^2)$	0.2 ± 0.4	0.02 ± 0.06

The most striking effect is the reduction by a factor 6 of the uncertainty on gluons. It is coherent with the significant increase in sensitivity for gluons we have mentioned earlier, going from at least 40 times less than quarks thanks to indirect contribution via evolution to only about 10 times less thanks to the direct introduction of d_1^g in the subtraction constant. Contrary to the situation at LO, adding gluon to the NLO fit changes the value of d_1^{uds}, multiplying the uncertainty on d_1^{uds} by a factor 2. It is a sign that d_1^{uds} and d_1^g are now slightly correlated, compared to the situation at LO. However, the joint fit of quark and gluon contributions remains the most robust extension of our model we can perform on actual experimental data.

Conclusion

Extracting either the GFF $C(t)$ – or the more difficult full D-term – from the knowledge of the subtraction constant is a deconvolution problem whose theoretical solution is obtained thanks to evolution equations. However, due to the tame effect of evolution, and the large similarity of anomalous dimensions for Gegenbauer coefficients at large n, the amplification of uncertainty of the problem is particularly large on the current experimental range.

First, on the physical side, we corroborate the findings of Ref. [202] by showing that the subtraction constant derived from a neural network fit of most of the world DVCS data is compatible with 0. Therefore, a significant effort to collect more precise data is necessary for the study of hadron matter properties to produce statistically significant

results in analyses that are not strongly model dependent. We notice that most of the uncertainty comes from the poor knowledge of the real part of the CFF, which could be constrained by charge dependent observables as we demonstrated in Section 4.4.

Then, the analytic solution of the D-term evolution allows us to understand to a good quantitative extent the results of all our fits. Among notable results, we have shown that radiative gluon contributions are so weak that they amount in practice to no sensitivity to gluons, either at LO (sensitivity suppressed by a factor at least 40) or NLO (sensitivity suppressed by a factor 10). The fit of d_1^{uds} is quite resilient to the addition of free gluon contributions either at LO or NLO. On the contrary, it behaves very poorly when confronted with d_3^{uds}. We have developed an estimate of the uncertainty on d_3^{uds} based on the similarity of the fitted functional dependence on μ^2 between $d_1^{uds}(\mu^2)$ and $d_3^{uds}(\mu^2)$. This estimate allows notably to gauge the impact of the lever arm in Q^2 which could be achieved at an EIC.

We have presented many important concepts that allow us to better understand the challenges of deconvolution on a simpler problem of yet significant physical relevance compared to the full extraction of GPDs from CFFs we will tackle in the next chapter. Notably, we have briefly introduced shadow D-terms as an alternative source of knowledge on the uncertainty of the D-term extraction from subtraction constant data. This intuitive and flexible tool will represent most of the forthcoming discussion about the general deconvolution problem.

6 - The full DVCS deconvolution problem

After presenting in the previous chapter the issue of deconvoluting the D-term and the gravitational form factor $C^a(t)$ from the experimental knowledge of the DVCS subtraction constant, it is time to devote interest to the extraction of the rest of the information on GPDs contained in DVCS data. We remind from the presentation of DVCS in Chapter 3 that photon lepto-production – the actual experimental channel in which DVCS is measured – provides sensitivity to GPDs in a leading twist analysis thanks to EFFs and CFFs. EFFs dependence arises from Bethe-Heitler contributions to the photon lepto-production cross-section, which only bring constraints on the non-singlet (C-odd) quark GPDs through (see Eq. (1.18))

$$2F_1^q(t) = \int_{-1}^{1} \mathrm{d}x\, H^{q(-)}(x,\xi,t,\mu^2)\,, \tag{6.1}$$

$$2F_2^q(t) = \int_{-1}^{1} \mathrm{d}x\, E^{q(-)}(x,\xi,t,\mu^2)\,, \tag{6.2}$$

where the factor 2 arises from using explicitly the non-singlet component of the GPD. With no dependence on Q^2 or ξ in these form factors, the potential for deconvolution is inexistent. On the other hand, DVCS is only sensitive to C-even contributions and brings rich constraints on the singlet quark and gluon GPDs thanks to the factorisation of CFFs (see Section 3.1, (3.10)) :

$$2\mathcal{H}(\xi,t,Q^2) = \sum_a T^a\left(\frac{x}{\xi},\frac{Q^2}{\mu^2},\alpha_s(\mu^2)\right)\otimes H^{a(+)}(x,\xi,\mu^2)\,. \tag{6.3}$$

For simplicity, we will present all the results with GPD H and the associated CFF $\mathcal{H}$. Formulas for E, $\widetilde{H}$ and $\widetilde{E}$ can readily be derived.

The DVCS deconvolution problem consists in recovering $H^{a(+)}$ from the experimental knowledge of $\mathcal{H}$. As in the case of the D-term, the flavour separation possibilities are limited. A simpler version of the problem consists in trying to recover parton-by-parton $H^{a(+)}$ from $\mathcal{H}^a = T^a \otimes H^{a(+)}$ assuming that flavour separation has been achieved by other means. These could include a theoretical model of the relation between flavours – the simplest being that all light flavours are equal and heavy flavours are purely radiatively generated as we did in the D-term study in Chapter 5, with either free or radiative gluon contributions – or some experimental input. The latter can for instance stem from changing the target on which DVCS is performed from a proton to a neutron [203], or using meson production which is sensitive to the flavour content of GPDs. We will focus on the simpler parton-by-parton deconvolution in the following. The status of the literature on the feasibility of the deconvolution procedure is contrasted. Already in the early days of GPDs, Ref. [217] argued that evolution provided the theoretical means to perform this deconvolution, in principle without the need to resort to a parametric fit. Aside from the detailed status of Ref. [35], most reviews of GPDs either briefly mention the absence of a generic deconvolution procedure [21, 67, 116] or asserted that a parametric form had to be chosen to fit to experimental data [218, 171, 219, 220]. The quantitative impact of

evolution on typical EIC evolution range from 2 GeV^2 to 50 GeV^2 in a LO analysis has been questioned in Ref. [221].

Our objective in this chapter is to determine an order of magnitude of what we will call the **conditioning of the deconvolution problem** at NLO, that is the factor of amplification applied to experimental CFF uncertainty in the extraction of GPDs in the worst case [1]. As far as we know, no systematic attempt at characterising this issue at NLO has been performed before. We know that this conditioning depends on the modelling assumptions imposed on $H^{a(+)}$ as we have demonstrated in Section 5.2 in the case of the subtraction constant, where a joint fit of d_1^{uds} and d_1^g gave a reasonably robust result, whereas the joint fit of d_1^{uds} and d_3^{uds} was drown in the uncertainty emerging from shadow D-terms. It is therefore crucial in our discussion of the full deconvolution problem to determine clearly what modelling assumptions we want to enforce on $H^{a(+)}$, all the more that this object fulfills theoretical properties, like polynomiality of Mellin moments and positivity.

Our approach will be based on shadow GPDs that we introduced in Refs. [6, 7]. As we have explored in the case of the subtraction constant, finding shadow distributions – that is essentially distributions with vanishing imprint on CFFs at some scale μ_F^2 – within the parameter space spanned by the model of GPD and studying their growth with μ^2 provides an estimation of the difficulty of the deconvolution problem on a experimental range $[\mu_F^2, \mu^2]$. Evolution of shadow distributions determines the conditioning of the extraction, but their actual functional dependence is also a very interesting input. Our study of shadow GPDs will be the basis for the modelling efforts we conduct in the next chapter to pave the way for data-driven extractions of GPDs from experimental data.

Neighbour concepts to shadow GPDs have been envisioned in the literature, although no systematic search comparable to what we perform here has been attempted as far as we are aware, and particularly not a NLO. Ref. [222] discusses "zero modes" of the LO subtraction constant, which are exactly LO shadow D-terms at a fixed scale, although the influence of scale dependence was not explicitely evaluated. Ref. [221] discussed the possible weakness of LO evolution of GPDs even on EIC kinematics by considering three different models of GPDs with quite similar CFFs at scale $\mu_0^2 = 2$ GeV^2, and observing that they remain barely distinguishable at $\mu^2 = 50$ GeV^2. As we will see, the difference between the models used in this study can be considered as an approximate LO shadow GPD.

6.1 . Shadow generalised parton distributions

Let us start by formally defining the notion of shadow GPD. Due to the number of theoretical and experimental constraints that GPDs obey, our definition is more constrai-

1. We borrow the terminology from the conditioning of a matrix, defined as the ratio of its largest eigenvalue over its smallest one. The conditioning measures the worst-case amplification of error through the inversion of the matrix. In practice, let A be the matrix of interest, and X_1 and X_ϵ two eigenvectors of norm 1 associated respectively to the largest and smallest eigenvalues λ_1 and λ_ϵ. Let us assume that their ratio $\lambda_1/\lambda_\epsilon$, which we defined as the conditioning of A, equals $1/\epsilon$. Then $A(X_1 + X_\epsilon) = \lambda_1 X_1 + \lambda_\epsilon X_\epsilon = \lambda_1[X_1 + \mathcal{O}(\epsilon)]$. Therefore, as the conditioning increases, ϵ decreases, and the contribution of X_ϵ may be considered as noise. However, provided A^{-1} exists, $A^{-1}[\lambda_1 X_1 + \lambda_\epsilon X_\epsilon] = X_1 + X_\epsilon$. Applying A^{-1} to what we considered as negligible noise just before now yields contributions of size 1 : the error has been amplified by the inversion by a factor $1/\epsilon$, which is the conditioning of the matrix.

ning than that for shadow D-terms. What we call "actual GPD" in this definition is any GPD which satisfies the properties that we have reviewed in Chapter 2 while giving a satisfying account of experimental data.

Shadow GPDs A shadow GPD at N^{k}LO and μ_F^2 is a set of functions $H_0^{a(+)}(x,\xi,\mu^2)$ for $a = u, d, s, g, ...$ which verify

- $H_0^{a(+)}(x,\xi,\mu^2)$ has the same parity in x and ξ that an actual singlet GPD of the same type (even in ξ, odd in x for quarks and even in x for gluons),
- it vanishes at least as fast as an actual GPD at end point $x = 1$,
- it respects the polynomiality of Mellin moments or equivalently admits a double distribution (DD) representation – which shall be called shadow DD,
- it follows the same evolution equations as actual GPDs,
- its convolution with the N^{k}LO DVCS coefficient function is exactly vanishing at μ_F^2 :

$$T^a\left(\frac{x}{\xi}, \frac{Q^2}{\mu_F^2}, \alpha_s(\mu_F^2)\right) \otimes H_0^{a(+)}(x,\xi,\mu_F^2) = 0\,, \tag{6.4}$$

- its forward limit is exactly vanishing

$$H_0^{a(+)}(x,0,\mu_F^2) = 0\,. \tag{6.5}$$

The interest of shadow GPDs is similar to that of shadow D-terms detailed in Section 5.2.3. More precisely, the uncertainty on a linear GPD extraction can be represented by $H_{best}^{a(+)} + H^{a(+)}$, where $H_{best}^{a(+)}$ is the best fit value obtained for the linear [2] model of the GPD and $H^{a(+)}$ is an element of the allowed functional space spanned by the model such that $|T^a \otimes H^{a(+)}(\mu^2)|$ is of the order or less than the experimental uncertainty over the probed range in μ^2. Since shadow GPDs are intended to bring particularly small contributions to the CFF convolution, they are ideal candidates to obtain maximal such functions $H^{a(+)}$.

Since the forward limit of GPDs at $t = 0$ is related to ordinary parton distribution functions (PDFs), we have added the requirement of vanishing forward limit in the definition of the shadow GPD. This is particularly useful for H, since its forward limit is the unpolarised PDF which is very well constrained. Therefore, we may consider that the experimental uncertainty on the forward limit is virtually 0, and for our shadow GPDs to remain below, they have to vanish in this limit. $\widetilde{H}$ gives polarised PDFs in its forward limit, which are also quite well constrained by a wealth of experimental data. On the opposite, the forward limit of E and $\widetilde{E}$ are poorly known, so the requirement can be lifted then. It is also the case if we are explicitly interested in a value of $t \neq 0$. Let us notice that, for our general purpose of evaluating the conditioning of the deconvolution problem, putting too much constraints on our shadow GPDs will result in an underestimation of their potential maximal effect, so will give an inferior bound on the worse conditioning which is still useful. Since the forward limit of GPDs follows its own evolution equation DGLAP,

2. Same comment as for shadow D-terms. In order to quantify the uncertainty of the deconvolution given a modelling assumption, shadow GPDs are only useful if $H_{best}^{a(+)} + H^{a(+)}$ belongs to the functional space spanned by the model, which is easier to fulfill in case of a linear model encompassing independently $H_{best}^{a(+)}$ and $H^{a(+)}$.

vanishing forward limit at scale μ_F^2 implies the same result at any scale, regardless of the perturbative order under consideration.

All other conditions are intended to guarantee that the sum of a shadow GPD and the best fit GPD is still a GPD. One of the important properties of actual GPDs which we did not list is positivity. Shadow GPDs themselves violate positivity inequalities maximally. We remind that the inequalities bound the value of combinations of GPDs by their forward limit. Since shadow GPDs have vanishing forward limits, satisfying positivity inequalities would condemn them to remain equally 0. But as shadow GPDs are meant to be added to actual good GPD candidates, their positivity bound must be derived from the forward limit of the good GPD candidate. We will deal with this question in the next chapter where we build models based on shadow GPDs, but for now we will not concern ourselves with positivity bounds.

We note that shadow GPDs verify a fundamental property : they can be written simply as the difference of two actual singlet GPDs with the same N^{k}LO CFF at μ_F^2 and the same forward limit. This would immediately fulfill all conditions enumerated in our previous definition, and the opposite is true. We will use this fact to our advantage in the modelling part, but it is less useful to explicitly demonstrate the existence of such objects.

An important consequence of the fact that shadow GPDs admit a double distribution representation is that they can be decomposed in the Polyakov-Weiss scheme into a DD $f^{a(+)}(\beta,\alpha,\mu^2)$ and a D-term. It is easy to verify thanks to the dispersive formalism that the D-term of a shadow GPD is a shadow D-term as defined in Section 5.2.3. Indeed, the real and imaginary parts of the N^{k}LO CFF associated to the shadow GPD are 0 at scale μ_F^2 by definition, so Eq. (3.25) gives that the subtraction constant itself is 0. As we have already studied the conditioning of the deconvolution problem for the subtraction constant, we can focus here solely on the shadow DD $f^{a(+)}(\beta,\alpha,\mu^2)$. In the absence of a D-term, the dispersion relation shows that $\mathrm{Im}\, T^a \otimes H_0^{a(+)}(\mu_F^2) = 0 \Longrightarrow \mathrm{Re}\, T^a \otimes H_0^{a(+)}(\mu_F^2) = 0$, so the cancelling the CFF of a shadow DD $f^{a(+)}(\beta,\alpha,\mu^2)$ simply amounts to solving

$$\mathrm{Im}\, T^a\left(\frac{x}{\xi},\frac{Q^2}{\mu_F^2},\alpha_s(\mu_F^2)\right)\otimes\int_\Omega \mathrm{d}\alpha\mathrm{d}\beta\,\delta(x-\beta-\alpha\xi)f^{a(+)}(\beta,\alpha,\mu_F^2)=0 \tag{6.6}$$

that is for quarks,

$$\int_\Omega \frac{\mathrm{d}\alpha\mathrm{d}\beta}{\xi}\,\mathrm{Im}\, T^q\left(\alpha+\frac{\beta}{\xi},\frac{Q^2}{\mu_F^2},\alpha_s(\mu_F^2)\right) f^{q(+)}(\beta,\alpha,\mu_F^2)=0\,. \tag{6.7}$$

We will now demonstrate explicitly that there exist shadow DDs for some very lax modelling assumptions at LO and NLO.

6.2 . Proof of the existence of LO and NLO shadow DDs at a given scale

In order to demonstrate the existence of shadow DDs, we first have to determine a general functional space in which we allow our DDs to live. The determination of this space is crucial : there may exist shadow DDs, but they may only live in such exotic spaces that they never come close to any reasonable parameterisation, and are therefore completely irrelevant from a phenomenological point of view. Since we are not interested in such exotic constructions, we choose a very reasonable space : polynomials in the variable α

and β. The freedom of our model is determined by the maximal order N which we allow for our polynomials. Thanks to the Stone-Weierstrass theorem, any continuous function on a compact separated space can be approximated by polynomials provided N is large enough. They represent therefore a simple but quite generic representation for DDs.

It is true that phenomenological models of GPDs would require a prohibitively large number of terms, because polynomials have a hard time producing the observed power divergence of singlet PDFs at small x. But we are trying to model here a shadow GPD, that is the difference of two singlet GPDs with a vanishing forward limit, and it is very reasonable to expect them to cancel out their small x divergence. Another pitfall is that high order polynomials can have a tendency to oscillate quickly. What kind of acceptable oscillating behaviour for DDs is allowed is not a settled question as far as we are aware. Widely accepted phenomenological GPDs exhibit noticeable irregularities in the (x, ξ) plane. For instance, the presence of a D-term translates into a contribution of the form $\mathrm{sgn}(\xi) D(x/\xi)$ in the GPD, which produces frequently significant oscillations and a discontinuity in the small $|\xi| > |x|$ region. Ref. [179] demonstrates that the covariant extension of a completely well-behaved GPD defined by saturation of its positivity bound in the region $|x| \geq |\xi|$ can produce significant oscillations in the region $|x| \leq |\xi|$. Nonetheless, we shall check that even N of the order of 20 or more does not produce obviously unacceptable oscillating behaviour with our shadow GPDs.

With our polynomial assumption, a quark shadow DD writes for $\beta > 0$

$$f^{q(+)}(\beta, \alpha) = \sum_{\substack{m,n=0 \\ m \text{ even}}}^{m+n \leq N} c_{m,n} \alpha^m \beta^n \,, \tag{6.8}$$

with implicit (t, μ_F^2) parameters which we will omit unless necessary.

For $f^{q(+)}$ to be the DD of a singlet GPD, it must be odd in β. We can either guarantee this by defining $f^{q(+)}(\beta, \alpha) = -f^{q(+)}(|\beta|, \alpha)$ for $\beta < 0$ or by only considering n odd in Eq. (6.8) as we did in Ref. [6]. The first option means that $f^{q(+)}$ is not actually a true polynomial over the whole rhombus, but only separately for $\beta > 0$ and $\beta < 0$. Due to the second option, we will systematically consider that N is odd in the following. Then the definition of $f^{q(+)}$ requires $(N+1)(N+3)/4$ coefficients $c_{m,n}$.

The two prescriptions for negative β are useful, the first one spanning a large functional space and serving as a stepping stone for the second one as we will see in the following. Let us note already an interesting consequence of the first prescription. If $f^{q(+)}(0^+, \alpha) \neq 0$, then the DD is not continuous on the $\beta = 0$ line. The consequence of this discontinuity is quite similar to that of introduced by a D-term – which is also a discontinuity on the $\beta = 0$ line – from the point of view of the regularity of the obtained GPD. In the region $|x| \geq |\xi|$, since Radon integration lines never cross the $\beta = 0$ line inside the rhombus, the GPD will be perfectly well-behaved. Troubles appear when going to the $(x, \xi) = (0, 0)$ limit with $|\xi| > |x|$. Assume that $x, \xi > 0$ and $x = \lambda\xi$ for a fixed value of $\lambda < 1$. Then the Radon transform gives

$$H^{q(+)}(\lambda\xi, \xi) = \int_{(\lambda\xi - 1)/(1+\xi)}^{(1+\lambda\xi)/(1+\xi)} \mathrm{d}\alpha \, f^{q(+)}((\lambda - \alpha)\xi, \alpha) \,. \tag{6.9}$$

Using that $|\lambda| < 1$ and the parity of $f^{q(+)}$ in β,

$$H^{q(+)}(\lambda\xi,\xi) = \int_{(\lambda\xi-1)/(1+\xi)}^{\lambda} \mathrm{d}\alpha\, f^{q(+)}((\lambda-\alpha)\xi,\alpha) - \int_{\lambda}^{(1+\lambda\xi)/(1+\xi)} \mathrm{d}\alpha\, f^{q(+)}((\alpha-\lambda)\xi,\alpha)\,. \tag{6.10}$$

Then taking the limit $\xi \to 0$ gives

$$H^{q(+)}(\lambda\xi,\xi) \to \int_{-1}^{\lambda} \mathrm{d}\alpha, f^{q(+)}(0^+,\alpha) - \int_{\lambda}^{1} \mathrm{d}\alpha\, f^{q(+)}(0^+,\alpha)\,. \tag{6.11}$$

Then

$$\frac{\partial}{\partial\lambda} \lim_{\xi\to 0} H^{q(+)}(\lambda\xi,\xi) = 2f^{q(+)}(0^+,\lambda)\,. \tag{6.12}$$

One sees therefore that unless $f^{q(+)}(0^+,\alpha)$ is identically vanishing, the limit when $(x,\xi) \to (0,0)$ of $H^{q(+)}(x,\xi)$ is not well defined in the sector $|\xi| > |x|$, and depends on the "angle of approach" defined by λ. The situation is similar for the D-term, which produces a term $\mathrm{sgn}(\xi)D^q(x/\xi)$ in the GPD, whose limit when $(x,\xi) \to (0,0)$ depends also on the angle of approach unless the D-term is 0.

6.2.1 . Vanishing forward limit

The first condition that we choose to fulfill to build our shadow GPD is the cancellation of the forward limit of the associated GPD. With the form of Eq. (6.8), the associated forward limit reads thanks to the Radon transform

$$H^{q(+)}(x,0) = \int_{x-1}^{1-x} \mathrm{d}\alpha\, f^{q(+)}(x,\alpha)\,, \tag{6.13}$$

$$= 2(1-x)\sum_{v=0}^{N} x^v \left[\sum_{\substack{m=0\\ m\text{ even}}}^{N-1} \sum_{k=0}^{\min(v,m)} \frac{c_{m,v-k}}{m+1}\binom{m}{k}(-1)^k\right]\,. \tag{6.14}$$

Cancelling the forward limit therefore amounts to solving the system on the coefficients $c_{m,n}$ such that the bracketed term of Eq. (6.14) equals 0 for each value of $v \in \{0,...,N\}$. This is a set of $N+1$ linear independent equations [3] for $(N+1)(N+3)/4$ variables $c_{m,n}$. The explicit solution of the system can be found relatively easily [4] for $\beta > 0$ as

$$f^{q(+)}(\beta,\alpha) = \sum_{\substack{m=2\\ m\text{ even}}}^{N-1} P_m(\beta)((m+1)\alpha^m - (1-\beta)^m)\,, \tag{6.15}$$

where $P_m(\beta)$ is an arbitrary polynomial in β of order $N-m$. Notice that we have not imposed any parity condition in β on our result. On the other hand, characterising true

3. Independence can be demonstrated by observing that each equation involves a coefficient $c_{m,n}$ which is used in no other. Precisely, $c_{0,v}$ is only involved in the equation associated to x^v.

4. The easiest demonstration of this result consists in starting from the solution. First we verify that Eq. (6.15) yields indeed systematically a vanishing forward limit. Second we check that the dimension of the space of solutions that we have exhibited corresponds exactly to the expected dimension, which we know from the number of variables and independent equations in the system contained in Eq. (6.14). We will not detail such kind of straightforward but numerically cumbersome verifications in the following.

β-odd polynomial DDs with vanishing forward limit is significantly more difficult. We could solve directly the system made only of $c_{m,n}$ with n odd at a fixed value of N. This would yield a first solution where the kernel of the matrix is non trivial at order $N = 5$. The β-odd polynomial DDs with vanishing forward limit of lowest order are proportional to

$$f_5^{q(+)}(\beta, \alpha) = \beta(5\alpha^4 - 6\alpha^2\beta^2 - 6\alpha^2 + \beta^4 - 2\beta^2 + 1)\,. \tag{6.16}$$

However, each new order N will then bring new functions, with a more and more complicated shape. This is where it becomes interesting to have solved the system in the more general and simpler case of Eq. (6.15). We know that $f_5^{q(+)}$ is a particular case of that general solution and we can rewrite it as

$$P_2(\beta) = -2\beta(1 + \beta^2)\,, \quad \text{and} \quad P_4(\beta) = \beta\,. \tag{6.17}$$

This is the simplest combination of polynomials $P_m(\beta)$ which satisfies that each $P_m(\beta)$ is an odd function and $\sum_m P_m(\beta)(1-\beta)^m$ is also odd. Sparing further technical details, having translated the system no longer in terms of the $c_{m,n}$ coefficients but in terms of the $P_m(\beta)$ polynomials makes it easier to find the general solution for a truly polynomial singlet DD which cancels its forward limit as

$$\begin{aligned} f^{q(+)}(\beta, \alpha) = &\sum_{\substack{m=4 \\ m \text{ even}}}^{N-1} \beta Q_m(\beta)((m+1)\alpha^m - (1-\beta)^m) + \\ &+ \frac{1}{4}\left(3\alpha^2 - (1-\beta)^2\right) \sum_{\substack{m=4 \\ m \text{ even}}}^{N-1} Q_m(\beta)\left((1-\beta)^m - (1+\beta)^m\right)\,, \end{aligned} \tag{6.18}$$

where $Q_m(\beta)$ are now arbitrary *even* polynomials in β of order $N - m - 1$. The simplest solution $f_5^{q(+)}$ can be written in this new representation as

$$Q_4(\beta) = 1\,, \tag{6.19}$$

which is quite a simplification compared to the form of Eq. (6.16). There are $(N-1)(N-3)/8$ independent β-odd solutions which can be generated for order N odd, corresponding for instance to taking $P_m(\beta)$ successively as all allowed monomials. The 10 different GPDs obtained for $N = 11$ are represented on Fig. 6.1. Each monomial produces a quite similar overall shape, but with a different scale in x which allows us to build by linear combination a very flexible behaviour away from $\xi = 0$.

> **Remark** The generality of Eqs. (6.15) and (6.18) which involve arbitrary polynomials allows one to consider changing $P_m(\beta)$ and $Q_m(\beta)$ into arbitrary functions. Then of course the resulting DD will no longer be a polynomial, but it will inherit the property of vanishing forward limit. This is another argument why the polynomial representation is pertinent : it can serve as a practical calculation platform to derive more general results not specifically linked to the polynomial assumption. In practice, there is a very general and simple way to build a DD with a vanishing associated forward limit. We can first set the DD to just any

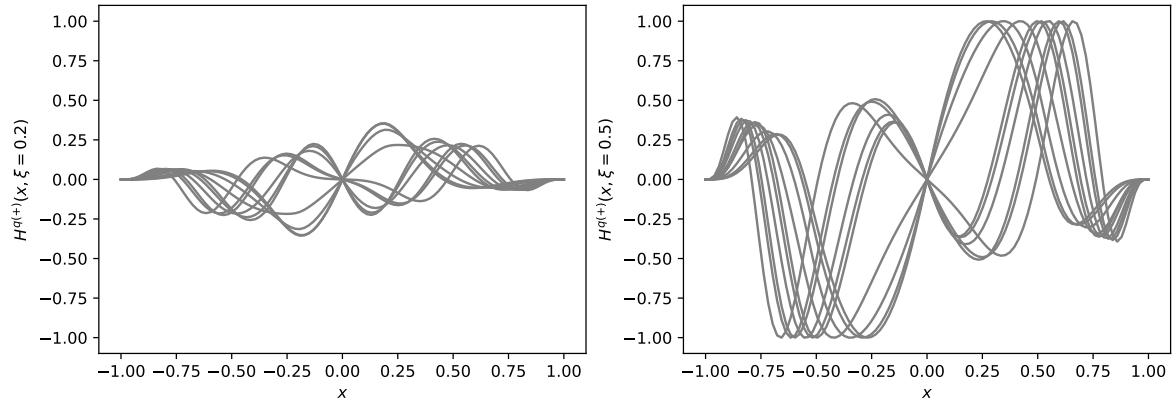

Figure 6.1 – The GPDs on the lines $\xi = 0.2$ (left) and $\xi = 0.5$ (right) associated to 10 independent β-odd polynomial DDs of order $n \leq 11$ with vanishing forward limit $\xi = 0$. The GPDs are scaled to reach maximal value of ± 1 on their $\xi = 0.5$ line. Each line corresponds to a monomial $Q_m(\beta)$ in Eq. (6.18). The GPDs decrease in size as ξ becomes closer to 0, which is of course a consequence of the fact that all these GPDs cancel exactly when $\xi = 0$.

value, and correct its α-independent part so as to cancel the forward limit. For $\beta > 0$, this method writes

$$f^{q(+)}(\beta,\alpha) = \widetilde{f}(\beta,\alpha) - \frac{1}{2(1-\beta)} \int_{\beta-1}^{1-\beta} \mathrm{d}\gamma\, \widetilde{f}(\beta,\gamma)\,, \tag{6.20}$$

where $\widetilde{f}(\beta,\alpha)$ is any function with correct support property. The formulation of Eq. (6.20) is more general and might seem more powerful than the one we have presented in Eq. (6.15). The latter is however far better suited for our purpose because it provides a rigorous computation scheme to add more and more constraints on the DD. For instance, we will later want to cancel CFFs, which we see *a priori* no easy way to do with the formulation of Eq. (6.20). On the contrary, the polynomial parameterisation will allow us to translate this constraint into another system on the $c_{m,n}$ coefficients which we can actually solve. We have already shown with Eq. (6.18) how we can build upon the solution of Eq. (6.15) to enforce more constraints on the DD.

6.2.2 . Radon transform with polynomial DDs

Before going further and cancelling the contribution of shadow GPDs to CFFs, it is necessary both to be able to express the Radon transform of a polynomial DD, and to recover a polynomial DD from the inverse Radon transform of a GPD. As we have seen in Section 3.4.2, computing the inverse Radon transform is an important issue for modelling based on the covariant extension programme. Our explicit analytical formulation of the Radon transform for polynomial DDs will therefore provide an additional angle of attack to this issue, in parallel with finite element methods.

The associated GPD computed thanks to the Radon transform of our polynomial DD

is the following rational fraction for $x > |\xi|$:

$$H^{q(+)}(x,\xi) = \int_{(x-1)/(1+\xi)}^{(1-x)/(1-\xi)} \mathrm{d}\alpha\, f^{q(+)}(x-\alpha\xi,\alpha)\,, \tag{6.21}$$

$$= \sum_{u=1,v=0}^{N+1} \left[\frac{1}{(1+\xi)^u} + \frac{1}{(1-\xi)^u}\right] q_{u,v} x^v\,, \tag{6.22}$$

where uniqueness of the partial fraction decomposition guarantees that coefficients $q_{u,v}$ are unambiguously determined. The linear relation between $c_{m,n}$ and $q_{u,v}$ derived from the Radon transform reads

$$q_{u,v} = \sum_{\substack{m=0 \\ m \text{ even}}}^{\min(u,N)-1} \sum_{n=\max(u,v)-m-1}^{N} R^{m,n}_{u,v} c_{m,n}\,, \tag{6.23}$$

where

$$R^{m,n}_{u,v} = \sum_{j=0}^{n} \frac{(-1)^{u+v+j+n+1}}{m+j+1} \binom{n}{j} \binom{j}{m-u+j+1} \binom{m+j+1}{v-n+j}\,. \tag{6.24}$$

A demonstration of this result is provided in Appendix A.4. From Eq. (6.22), we deduce that a polynomial DD of order N odd gives rise to $(N+1)(N+2)$ $q_{u,v}$ coefficients. Note that the linear relation of Eq. (6.23) guarantees that the apparent divergence of Eq. (6.22) for $|\xi| \to 1$ is only fictitious, since the simultaneous limit $x \to 1$ cancels out the issue [5].

Knowing explicitly the image of polynomial DDs through the Radon transform brings a new light to the covariant extension procedure presented in Section 3.4.2. We remind that the covariant extension consists in practice in building a DD which reproduces the values of a GPD in the region $|x| \geq |\xi|$, and then using this DD to extend the GPD to the region $|x| \leq |\xi|$. Ref. [179] elaborating on results of Ref. [77] demonstrated how to perform practically this inversion of the Radon transform in the region $|x| \geq |\xi|$ thanks to a finite elements method. Having discretised the (β,α) plane of DDs into hundreds of nodes, the authors build a relation between the values of the DD at those nodes and values of the GPD on a large number of $x_i \geq \xi_i \geq 0$ kinematics under the form of a linear system. The inversion of the system allows us to recover the values of the DD at the nodes. Our previous calculation provides an alternative. We can parameterise our DD by a polynomial of order N, and build likewise a relation between the coefficients $c_{m,n}$ and the values of the GPD of interest. Our linear model is given by

$$H^q(x,\xi) = \sum_{\substack{m,n=0 \\ m \text{ even}}}^{m+n\leq N} c_{m,n} R_{m,n}(x,\xi)\,, \tag{6.25}$$

where

$$R_{m,n}(x,\xi) = \sum_{u=m+1,v=0}^{N+1} \left[\frac{1}{(1+\xi)^u} + \frac{1}{(1-\xi)^u}\right] x^v R^{m,n}_{u,v}\,. \tag{6.26}$$

5. One can notice that GPDs computed in the perturbative limit $x \to 1$ (see Section 2.1, Eq. (2.9)) have a form which is quite reminiscent of the ξ pole structure of Eq. (6.22). With our formulation, we can easily makes powers of $(1-\xi^2)$ appear at the denominator of H.

Using the machinery of the linear fit introduced in Section 4.1.1, and considering we are fitting the GPD on $x_i \geq \xi_i \geq 0$ kinematics with $1 \leq i \leq M$ and no uncertainty, we introduce the matrix

$$C = \begin{pmatrix} R_{0,0}(x_1,\xi_1) & R_{0,1}(x_1,\xi_1) & \cdots & R_{N-1,1}(x_1,\xi_1) \\ & & \ddots & \\ R_{0,0}(x_M,\xi_M) & R_{0,1}(x_M,\xi_M) & \cdots & R_{N-1,1}(x_M,\xi_M) \end{pmatrix} . \tag{6.27}$$

C can be entirely computed analytically once the grid of values (x_i, ξ_i) is fixed. We use a rectangular grid of 210 points regularly [6] spaced to span the entire (x, ξ) plane where $1 > x \geq \xi \geq 0$. Then if Y is the vector of values of the GPD at the probed kinematics, the coefficients $c_{m,n}$ fitting best the GPD are obtained as $(C^T C)^{-1} C^T Y$. The extension of the GPD in the region $|x| < |\xi|$ can then easily be computed from the knowledge of the DD. To compare our results with Ref. [179], we work in the Pobylitsa DD scheme and adopt one of the pion GPDs discussed in the paper, namely for $x > |\xi|$ and $t = 0$:

$$H^u(x,\xi) = \sqrt{u\left(\frac{x+\xi}{1+\xi}\right) u\left(\frac{x-\xi}{1-\xi}\right)} , \tag{6.28}$$

where the pion PDF is given by

$$u(x) = 213.32\, x^2(1-x)^2 \left[1 + 2.29\, x(1-x) - 2.93\sqrt{x(1-x)}\right] . \tag{6.29}$$

This definition of the GPD saturates the pion positivity property. The specific shape of the PDF is intended to give an account of dynamical chiral symmetry breaking at low scale [223]. In addition, the soft pion theorem [39, 224] allows one to fix a part of the ambiguity of the DD on the $\beta = 0$ line – which is unconstrained by measurements solely in the region $|x| \geq |\xi|$ – by requiring that the reconstructed GPD at $\xi = 1$ is an even function of x. We follow the same prescription which consists in subtracting to the reconstructed GPD the following D-term for $x > 0$

$$D^u(x) = \frac{1}{2}\Big[H^u(x,1) - H^u(-x,1)\Big] . \tag{6.30}$$

The result of our fit with the addition of the aforementioned D-term, is shown on Fig. 6.2 for three different orders N of the polynomial DD : $N \in \{5, 7, 9\}$. We only fit the DD for $\beta > 0$, and keep $f^u(\beta, \alpha) = 0$ for $\beta < 0$ to satisfy the requirement that $H^u(x, \xi) = 0$ for $x < -|\xi|$ (no antiquarks for this valence GPD). As can be seen on the first three panels, as N increases, the agreement of the reconstructed GPD (dotted lines) with the fitted GPD (solid lines) becomes excellent. On the lower left panel, we show in addition to the best fit value an account of the uncertainty of the extraction. What is shown precisely are replicas generated using the covariance matrix of the fit $(C^T C)^{-1}$, where we assumed that each of the fitted values of the GPD varies independently by a normal distribution of absolute standard deviation 0.0023. This value is selected because 68% of the absolute distances between the best fit and the targeted GPD on the fitted

6. The pion valence PDF we will consider has no divergence at small x, so a linear grid is more adapted than a logarithmic one. The choice of the grid is absolutely transparent in the formalism.

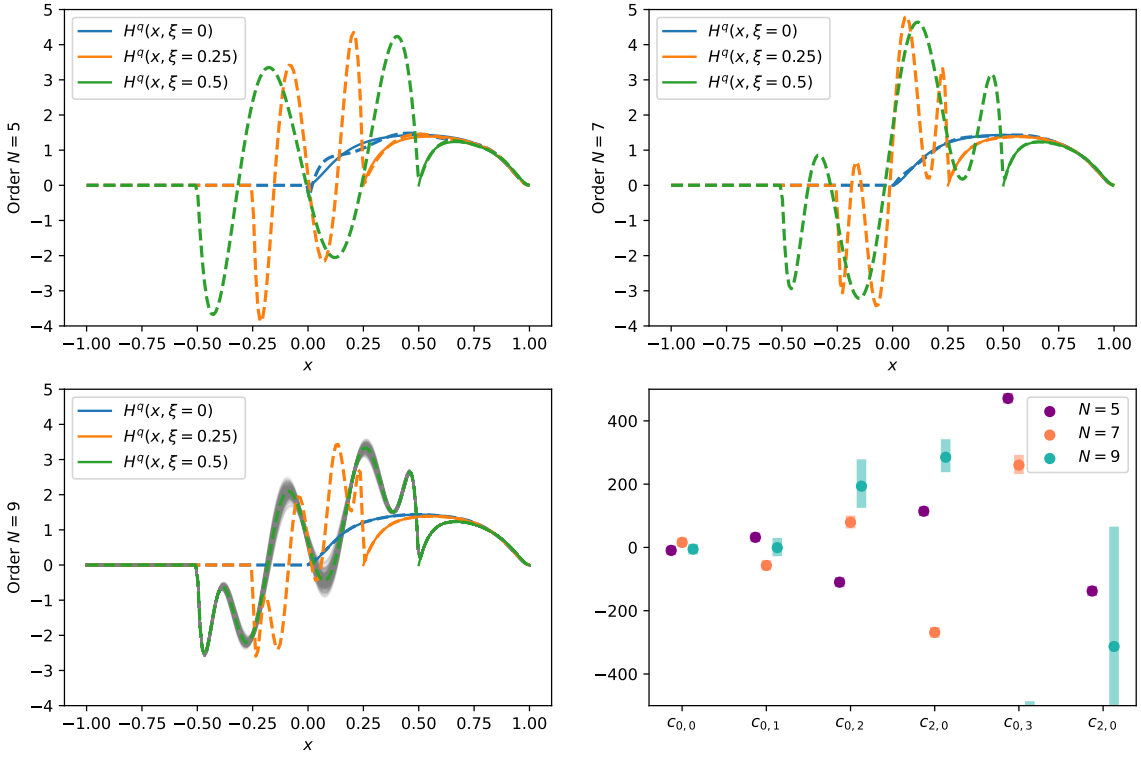

Figure 6.2 – (first three panels) Fit of the GPD defined by Eq. (6.28) in the region $|x| \geq |\xi|$ (solid lines) by the image through the Radon transform of a polynomial DD (dotted lines) of degree $N = 5$ (upper left), $N = 7$ (upper right) and $N = 9$ (lower left). The image of the Radon transform is also computed in the region $|x| \leq |\xi|$ to demonstrate the possibility of covariant extension. The fit is performed on 210 (x, ξ) points. $N = 5$ corresponds to 12 free parameters, $N = 7$ to 20 and $N = 9$ to 30. The χ^2 divided by degrees of freedom (number of points minus number of free parameters) is 3.5e-3 for $N = 5$, 1.8e-4 for $N = 7$ and 1.2e-5 for $N = 9$. In the case $N = 9$, we show in addition replicas demonstrating the uncertainty of our fit when the 210 fitted points are smeared by uncorrelated uncertainty of absolute size 0.0023. – (lower right) Central value and uncertainty on a few parameters $c_{m,n}$ for $N = 5$, $N = 7$ and $N = 9$. The dot represents the best fit value, and the band the uncertainty if the fitted GPD is smeared as in the previous panel.

points are inferior to it. It is so minute that it is impossible to see in the region $|x| \geq |\xi|$ where the fit is performed. However, it translates into a visible uncertainty in the region $|x| \leq |\xi|$, which can reach a standard deviation of about 0.3. That represents therefore an increase of error by a factor 100, which we will consider more precisely below. The last panel of Fig. 6.2 shows the uncertainty on a few different $c_{m,n}$ coefficients of lowest order depending on the total order of the fit. When increasing order N, the values of the coefficients change a lot and χ^2 still decreases significantly, signifying that higher orders in N are likely to bring changes to the extracted DD. This is visible overall by the change of extension in the region $|x| \leq |\xi|$ produced at each order.

When comparing with the result of Ref. [179] (Fig. 5), noticeable differences are visible in the region $|x| \leq |\xi|$, some of them being significantly outside both uncertainty evaluations. A first answer is that our polynomial fit will keep evolving as N increases, until it probably comes closer to the finite elements extraction. Indeed, at $N = 9$, we still use only 30 free parameters compared to the 427 nodes used in Ref. [179]. Furthermore, the uncertainty band represents only aleatoric uncertainty, obtained by varying the inputs but not the model used for fitting. They bring therefore only a limited picture of the real uncertainty of the fit. However, the agreement in the region $|x| \geq |\xi|$ is already excellent at $N = 9$. This is a striking example of the ill-posed character of the inversion of the Radon transform [225], or equivalently of its very large conditioning. Precisely, the issue is that there exist particularly small eigenvalues of the linear Radon transform operator, such that the image of associated eigenvectors is small in size. Therefore, when applying the inverse Radon transform to the image of these eigenvectors, a tiny feature of the fitted data gets considerably amplified. The stability of the inverse Radon transform will depend on the ability of the fitted functional form to not incorporate these eigenvectors [7]. This explains why, as the space of polynomial increases with N, the sensitivity of the fitted coefficients to noise on the fitted data increases significantly. We are facing once again the dilemma of bad conditioning. Do we want to relax strong model assumptions which bias our result, but at the possible cost of letting poorly constrainable eigenvectors sneak into our probed functional space ?

Finally let us note that the polynomial form is at the same time an advantage, since it allows generic calculations with solely additions and multiplications, but a drawback as well. Already at $N = 9$, to obtain a reliable value of the fit, we had to resort to exact fraction calculations and inversion of matrices thanks to Gauss-Jordan elimination. With the lowest values of x and ξ for fitted kinematics of the order of 0.05 and the largest of the order of 1, raising to power 9 forces to consider in the same matrix values spanning 10^{-12} to 1. Fortunately, if the (x, ξ) points on which the fit is performed are taken as fractions, C can always be computed exactly, so it is not a problem as long as the polynomial order remains small enough – so that the size of the numerator and denominator of the fractions under consideration do not become prohibitively large as well.

7. Maybe we should call them shadow eigenvectors ? :) In practice, shadow distributions as we have defined them are elements of the kernel of a number of linear operators (notably the linear operators which compute the forward limit and the CFFs from the GPD) at their definition scale, so formally eigenvectors associated exactly to the eigenvalue 0 at that scale. Evolution whose effect we will study in Section 6.3 prevents shadow distributions to remain exactly eigenvectors of the mentioned operators at other scales, but the analogy we draw here further illustrates how shadow distributions are a tool to quantify the conditioning of a linear inversion / deconvolution problem.

Analytic inversion of the Radon transform

The formal demonstration of the results of this section can be found in Appendix A.4. There are $(N+1)(N+2)$ coefficients $q_{u,v}$ for only $(N+1)(N+3)/4$ coefficients $c_{m,n}$ for a polynomial DD of order N. Therefore, the $q_{u,v}$ form a family with too much freedom to parameterise the image of the Radon transform of polynomials DDs. Some functions which can be written under the form of Eq. (6.22) for $x > |\xi|$ do not stem from the Radon transform a polynomial DD. Therefore, assuming a GPD derives indeed from a polynomial DD, the coefficients $c_{m,n}$ can be retrieved from only a subset of $q_{u,v}$. A systematic way to do so is to select

$$\begin{aligned} &q_{1,1}, q_{1,2}, ..., q_{1,N+1}, \\ &\quad q_{3,3}, q_{3,4}, ..., q_{3,N+1}, \\ &\quad\quad ...., q_{N,N}, q_{N,N+1} \end{aligned} \tag{6.31}$$

Then ordering of the coefficients $c_{m,n}$ by lexicographic order of $(m+n, m)$ and the selected subset of $q_{u,v}$ by lexicographic order of (v, u), Eq. (6.23) can be recasted in a matrix formulation – for instance for $N = 3$:

$$\begin{pmatrix} q_{1,1} \\ q_{1,2} \\ q_{1,3} \\ q_{3,3} \\ q_{1,4} \\ q_{3,4} \end{pmatrix} = \begin{pmatrix} -1 & & & & & \\ & -1/2 & & & & \\ & & -1/3 & & & \\ & & -1/3 & -1/3 & & \\ & & & & -1/4 & \\ & & & & -1/4 & -1/12 \end{pmatrix} \begin{pmatrix} c_{0,0} \\ c_{0,1} \\ c_{0,2} \\ c_{2,0} \\ c_{0,3} \\ c_{2,1} \end{pmatrix} . \tag{6.32}$$

The above matrix is block-diagonal, and grows in an orderly fashion when N increases : an increase in N results simply in the addition of a new block at the end of the diagonal. It is an excellent and non-trivial property of the chosen subset of the $q_{u,v}$ coefficients, which allows us to compute explicitly its inverse. We find for instance for $N = 3$

$$\begin{pmatrix} c_{0,0} \\ c_{0,1} \\ c_{0,2} \\ c_{2,0} \\ c_{0,3} \\ c_{2,1} \end{pmatrix} = \begin{pmatrix} -1 & & & & & \\ & -2 & & & & \\ & & -3 & & & \\ & & 3 & -3 & & \\ & & & & -4 & \\ & & & & 12 & -12 \end{pmatrix} \begin{pmatrix} q_{1,1} \\ q_{1,2} \\ q_{1,3} \\ q_{3,3} \\ q_{1,4} \\ q_{3,4} \end{pmatrix} . \tag{6.33}$$

The general solution of the inversion is found as

$$c_{m,n} = -\binom{n+m}{m} (n+m+1) \sum_{\substack{k=0 \\ k \text{ even}}}^{m} \binom{m}{k} E_{m-k} q_{k+1,n+m+1} , \tag{6.34}$$

where the E_k are known as Euler numbers, defined by

$$\frac{2}{\exp(x) + \exp(-x)} = \sum_{k=0}^{\infty} \frac{E_k}{k!} x^k . \tag{6.35}$$

We adhere to the convention of alternating signs. In particular, E_m is null for m odd and $E_0 = 1$, $E_2 = -1$, $E_4 = 5$, $E_6 = -61$, ..., $E_m \approx 2m!(2/\pi)^{m+1}(-1)^{m/2}$. Eq. (6.34) has practical consequences, since it provides a way to determine whether or not a GPD which writes under the form of Eq. (6.22) for $x > |\xi|$ derives indeed from a polynomial DD. After obtaining a candidate polynomial DD with Eq. (6.34), computing its image through the Radon transform and verifying whether or not the initial GPD is recovered will prove if the GPD derives from a polynomial DD. Notice however that deriving from a polynomial DD is DD representation dependent.

6.2.3 . LO shadow GPDs

To progress on our construction of shadow GPDs, we now wish to cancel the imaginary part of LO CFF. We remind that (3.15)

$$\mathrm{Im}\, T_0^q\left(\frac{x}{\xi}, \frac{Q^2}{\mu_F^2}, \alpha_s(\mu_F^2)\right) \otimes H^q(x,\xi,\mu^2) = \pi e_q^2 H^{q(+)}(\xi,\xi,t,\mu^2)\,. \tag{6.36}$$

Using the result of Eq. (6.22),

$$H^{q(+)}(\xi,\xi) = \sum_{u=1,v=0}^{N+1} \left[\frac{1}{(1+\xi)^u} + \frac{1}{(1-\xi)^u}\right] q_{u,v}\xi^v\,. \tag{6.37}$$

$$\tag{6.38}$$

Since

$$\xi^v = (\pm 1)^v(-1+1\pm\xi)^v = (\pm 1)^v \sum_{w=0}^{v} \binom{v}{w} (-1)^{w+v}(1\pm\xi)^w\,, \tag{6.39}$$

so

$$\frac{\xi^v}{(1\pm\xi)^u} = (\mp 1)^v \sum_{w=-u}^{v-u} \binom{v}{w+u} (-1)^{w+u}(1\pm\xi)^w\,, \tag{6.40}$$

and we obtain the partial fractions expansion of $H^{q(+)}(\xi,\xi)$ as

$$H^{q(+)}(\xi,\xi) = \sum_{w=-N-1}^{N} \left[(1-\xi)^w + (-1)^v(1+\xi)^w\right] \left[\sum_{u=1,v=0}^{N+1} (-1)^{w+u} q_{u,v} \binom{v}{w+u}\right]\,. \tag{6.41}$$

As we mentioned before, there exist however GPDs which can write under the form of Eq. (6.22) and do not actually derive from a polynomial DD. When only spanning the $q_{u,v}$ coefficients that can actually result from the Radon transform of a polynomial DD,

Eq. (6.41) takes the considerably friendlier form [8]

$$H^{q(+)}(\xi,\xi) \overset{\text{pol. DD}}{=} \sum_{w=1}^{N+1} \frac{1}{(1+\xi)^w} \left[\sum_{u=w,v=0}^{N+1} (-1)^{w+u+v} q_{u,v} \binom{v}{u-w} \right] . \tag{6.43}$$

With the same reasoning as for the forward limit in Eq. (6.14), the cancellation of the LO CFF amounts to solving a system on the coefficients $q_{u,v}$ such that the bracketed term in Eq. (6.43) equals 0 for each value of $w \in \{1, ..., N+1\}$. As the $q_{u,v}$ are themselves expressed as a linear function of the $c_{m,n}$ thanks to the action of the Radon transform which we have derived in the previous section, we can produce a fully matricial formulation of all the systems to solve [9].

Matricial form of the system giving LO shadow GPDs

To write the matricial form of the systems, it is useful to introduce a multi-index notation

$$\underline{m} = (m,n)\,, \tag{6.44}$$
$$\underline{u} = (u,v)\,. \tag{6.45}$$

Then we can denote by R the matrix of the Radon transform relating the $c_{\underline{m}}$ to the $q_{\underline{u}}$, whose general term $R_{\underline{u},\underline{m}}$ was given in Eq. (6.24). It is a matrix with $(N+1)(N+2)$ lines and $(N+1)(N+3)/4$ columns which yields

$$\left(q_{\underline{u}}\right) = \left(R_{\underline{u},\underline{m}}\right) \cdot \left(c_{\underline{m}}\right) . \tag{6.46}$$

The cancellation of the LO CFF can be written from Eq. (6.43) as

$$\left(C_{w,\underline{u}}\right) \cdot \left(q_{\underline{u}}\right) = \left(C_{w,\underline{u}}\right) \cdot \left(R_{\underline{u},\underline{m}}\right) \cdot \left(c_{\underline{m}}\right) = \left(0\right) , \tag{6.47}$$

where C is a matrix of $N+1$ lines and $(N+1)(N+2)$ columns of general term

$$C_{w,\underline{u}} = (-1)^{w+u+v} \binom{v}{u-w} . \tag{6.48}$$

8. This simplification is best understood by writing directly the Radon transform of the polynomial DD for $x = \xi$ in Eq. (6.21) :

$$H^{q(+)}(\xi,\xi) = \int_{(\xi-1)/(1+\xi)}^{1} \mathrm{d}\alpha\, f^{q(+)}(\xi(1-\alpha),\alpha) = \sum_{m,n} c_{m,n}\xi^n \int_{(\xi-1)/(1+\xi)}^{1} \mathrm{d}\alpha\, \alpha^m (1-\alpha)^n . \tag{6.42}$$

It is already clear from this expression that the pole structure at $\xi = 1$ of Eq. (6.41) has to disappear as only a pole structure at $\xi = -1$ will emerge from Eq. (6.42). The fact that the polynomial part of Eq. (6.41) (so for $w \geq 0$ in that expression) disappears as well is found by a cumbersome but straightforward partial fractions expansion of Eq. (6.42).

9. Computing fully the partial fractions expansion of Eq. (6.42) would yield directly a system on the $c_{m,n}$ coefficients to cancel the LO CFF. Instead, we have chosen an hybrid approach, where we derive a system on the $q_{u,v}$ but need to perform manually simplifications like the one from Eq. (6.41) to (6.43) due to the fact that not all $q_{u,v}$ arise from polynomial DDs. Then we build a system on the $c_{m,n}$ thanks to the Radon matrix. The reason for this two step procedure is that the system expressed at the level of the $q_{u,v}$ has a much nicer overall formulation, and it becomes increasingly helpful as we incorporate more constraints on the DD.

Owing to the relation derived straightforwardly from Eq. (6.22) :

$$H^{q(+)}(x,0) = \sum_{w=0}^{N+1} x^w \sum_{u=1,v=0}^{N+1} 2q_{u,w}\,, \tag{6.49}$$

one can likewise reinterpret the system we solved to cancel the forward limit in Eq. (6.14) as

$$\big(Q_{w,\underline{u}}\big)\cdot\big(R_{\underline{u},\underline{m}}\big)\cdot\big(c_{\underline{m}}\big) = \big(0\big)\ , \tag{6.50}$$

where Q is a matrix of $N+1$ lines and $(N+1)(N+2)$ columns of general term [10]

$$Q_{w,\underline{u}} = 2\delta_{wv}\,. \tag{6.51}$$

This formulation is therefore far more compact and practical to implement numerically compared to the sole expressions of Eqs. (6.14) and (6.43).

Cancelling only the LO CFF

Finding shadow LO DDs now requires to solve both systems (6.47) and (6.50) simultaneously. Since we have already derived the general solution of Eq. (6.50) given in Eqs. (6.15) or (6.18) depending on how the extension to $\beta<0$ is done, we can simply focus first on the cancellation of the LO CFF (6.47). Characterising the general solution space becomes increasingly difficult, and the upcoming results are found by solving the exact value of the system (6.47) for large values of N and uncovering their remarkable regularity. Although no demonstration will be given, these results are thoroughly checked up to a very large order in N. If we do not impose direct parity in β, the solution space is given by

$$f^{q(+)}(\beta,\alpha) = \sum_{\substack{m=2\\ m\ \text{even}}}^{N-1}\ \sum_{k=0}^{N-m} \lambda_{m,k}\beta^k\left[\alpha^m - \sum_{j=0}^{m}\beta^j\binom{m}{j}(-1)^j\widetilde{P}_{m-j}(k+j)\right]\,, \tag{6.52}$$

where the $\lambda_{m,k}$ are a series of arbitrary real coefficients, and the rational fractions $\widetilde{P}_j(x)$ are defined through the recursion relation

$$(x+j)(x+j+1)\widetilde{P}_j(x) = x(x+1) + j(j-1)\widetilde{P}_{j-2}(x)\,, \tag{6.53}$$

with initialisation [11] $\widetilde{P}_0(x)=1$ and $\widetilde{P}_1(x)=x/(x+2)$. There is a similarity with the solution for the forward limit of Eq. (6.15), which would correspond almost to the same form, but where $\widetilde{P}_j(x)=1$ for all j, up to a factor $m+1$ in front of α^m. We will come back on that aspect later. In particular, the space of polynomial DDs cancelling their forward limit and the one cancelling their LO CFF have the same dimension. However, the fact that the bracketed expression in Eq. (6.52) depends on k prevents the convenient

10. The interest of the hybrid approach of expressing the conditions in terms of $q_{u,v}$ and not $c_{m,n}$ mentioned in the previous footnote might appear more clearly here when comparing with the system in terms of the $c_{m,n}$ expressed in Eq. (6.14).

11. Initialisation would almost seem unnecessary, considering that putting $j=0$ and $j=1$ in Eq. (6.53) gives immediately $x(x+1)\widetilde{P}_0(x) = x(x+1)+0\times\widetilde{P}_{-2}(x)$ and $(x+1)(x+2)\widetilde{P}_1(x) = x(x+1)+0\times\widetilde{P}_{-1}(x)$.

formulation obtained for the forward limit. To illustrate the actual complexity of the expression hidden behind the $\widetilde{P}$ functions, the first two terms of the development in m of Eq. (6.52) read

$$f^{q(+)}(\beta,\alpha) = \sum_{k=0}^{N-2} \lambda_{2,k}\beta^k \left[\alpha^2 - \beta^2 + 2\beta\frac{k+1}{k+3} - \frac{k^2+k+2}{(k+2)(k+3)}\right] + \sum_{k=0}^{N-4} \lambda_{4,k}\beta^k \Bigg[\alpha^4 - \beta^4 + 4\beta^3\frac{k+3}{k+5} - 6\beta^2\frac{k^2+5k+8}{(k+4)(k+5)} + 4\beta\frac{k^3+6k^2+17k+12}{(k+3)(k+4)(k+5)} - \frac{k^4+6k^3+23k^2+18k+24}{(k+2)(k+3)(k+4)(k+5)}\Bigg] + \dots . \tag{6.54}$$

Based on this first result, the solution space for explicitly β-odd polynomials is obtained as

$$f^{q(+)}(\beta,\alpha) = \sum_{\substack{m=4\\ m \text{ even}}}^{N-1} \sum_{\substack{k=1\\ k \text{ odd}}}^{N-m} \lambda_{m,k}\beta^k \left[\alpha^m - \sum_{j=0}^{m} \beta^j \binom{m}{j} (-1)^j \widetilde{P}_{m-j}(k+j)\right] - \frac{1}{2}\sum_{\substack{j=1\\ j \text{ odd}}}^{N-2} \beta^j \times \left[\alpha^2 - \beta^2 + 2\beta\frac{j+1}{j+3} - \frac{j^2+j+2}{(j+2)(j+3)}\right] \sum_{\substack{m=4\\ m \text{ even}}}^{N-1} \sum_{\substack{k=1\\ k \text{ odd}}}^{N-m} \lambda_{m,k} \binom{m}{j+1-k} \widetilde{Q}_{m+k-j}(m+k)\,, \tag{6.55}$$

where the $\widetilde{Q}_j(x)$ are closely related to the rational fractions $\widetilde{P}_j(x)$, since they are also rational fractions defined as :

$$x(x+1)\widetilde{Q}_j(x) = (x-j+2)(x-j+3) + (j-1)(j-2)\widetilde{Q}_{j-2}(x-2)\,, \tag{6.56}$$

with initialisation $\widetilde{Q}_1(x) = (x+2)/x$ and $\widetilde{Q}_2(x) = 1$. The simplest polynomial solution with β-odd can be represented as

$$\lambda_{4,1} = 1\,, \tag{6.57}$$

corresponding to

$$f^{q(+)}(\beta,\alpha) = \beta\left(\alpha^4 - 2\alpha^2\beta^2 - \frac{6}{5}\alpha^2 + \beta^4 - \frac{2}{3}\beta^2 + \frac{1}{5}\right). \tag{6.58}$$

As for the vanishing of the forward limit, the solution space for β-odd solutions is of dimension $(N-1)(N-3)/8$. We show the 10 solutions corresponding to $N = 11$ on Fig. 6.3.

A similar generalisation compared to the one we exposed in Section 6.2.1 is also possible, following the idea that solutions can be found by setting the DD to just any value, and correcting its α-independent part so as to cancel the LO CFF. It is done in the following way, which we will use in Chapter 7 for modelling purposes for $\beta > 0$:

$$f^{q(+)}(\beta,\alpha) = \widetilde{f}(\beta,\alpha) - \frac{\mathrm{d}}{\mathrm{d}\beta}\left(\frac{\beta}{2-\beta}\int_{\beta-1}^{1} \mathrm{d}\gamma\, \widetilde{f}\left(\frac{\beta(1-\gamma)}{2-\beta},\gamma\right)\right) \tag{6.59}$$

The fact that this DD has systematically a vanishing LO CFF comes from, for $\xi > 0$,

$$H(\xi,\xi) = \int_{(\xi-1)/(1+\xi)}^{1} \mathrm{d}\alpha\, f(\xi-\alpha\xi,\alpha) = \frac{1}{\xi}\int_{0}^{2\xi/(1+\xi)} \mathrm{d}\beta\, f\left(\beta, 1-\frac{\beta}{\xi}\right), \tag{6.60}$$

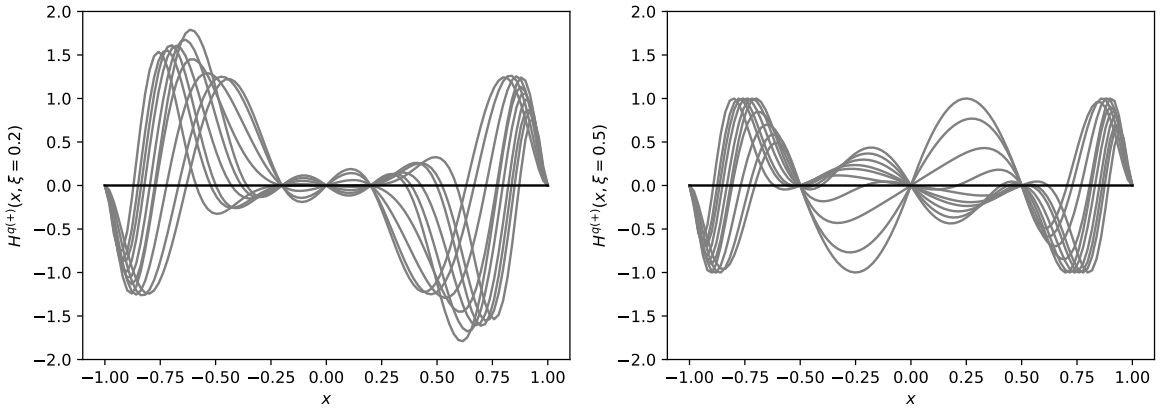

Figure 6.3 – The GPDs on the lines $\xi = 0.2$ (left) and $\xi = 0.5$ (right) associated to 10 independent β-odd polynomial DDs of order $n \leq 11$ with vanishing LO CFF. The GPDs are scaled to reach maximal value of ± 1 on their $\xi = 0.5$ line. Each line corresponds to a unique non-zero $\lambda_{m,k}$ in Eq. (6.55). The GPDs vanish at $x = 0$ by parity, and $x = \pm\xi$.

and

$$\frac{1}{\xi}\int_0^{2\xi/(1+\xi)} \mathrm{d}\beta\, \frac{\mathrm{d}}{\mathrm{d}\beta}\left(\frac{\beta}{2-\beta}\int_{\beta-1}^{1} \mathrm{d}\gamma\, \widetilde{f}\left(\frac{\beta(1-\gamma)}{2-\beta}, \gamma\right)\right)$$

$$= \frac{1}{\xi}\left(\frac{2\xi/(1+\xi)}{2-2\xi/(1+\xi)}\int_{2\xi/(1+\xi)-1}^{1} \mathrm{d}\gamma\, \widetilde{f}\left(\frac{2\xi/(1+\xi)}{2-2\xi/(1+\xi)}(1-\gamma), \gamma\right)\right) \tag{6.61}$$

$$= \int_{(\xi-1)/(1+\xi)}^{1} \mathrm{d}\gamma\, \widetilde{f}\left(\xi - \gamma\xi, \gamma\right) , \tag{6.62}$$

which is exactly the diagonal generated by the DD $\widetilde{f}$. However, just like for the similar formula obtained for the forward limit in Eq. (6.20), it is very impractical to use this formulation to try to impose more constraints on the DD, like parity, simultaneous cancellation of the forward limit, and later higher order considerations.

Building a complete LO shadow GPD

Having determined the solution space of DDs cancelling only their forward limit on the one hand, and only their LO CFF on the other, we need to compute the intersection of the two spaces to construct a LO shadow GPD. We remind that it corresponds to solving the linear system of the joint conditions Eqs. (6.47) and (6.50). The solution space is further reduced, but owing to the general argument that we still only have $\mathcal{O}(N)$ independent equations for $\mathcal{O}(N^2)$ coefficients $c_{m,n}$, we know we will find an infinite vector space of solutions. Due to the difficulties of determining the general solution space, we will now only propose an infinite family of β-odd solutions which cancel both their forward limit

and LO CFF defined by, for $N \geq 9$

$$f^{q(+)}(\beta,\alpha) = \beta^{N-8}\Bigg[\alpha^8 - \frac{28}{9}\alpha^6\left(\frac{N^2-3N+20}{(N+1)N} + \beta^2\right) + \frac{10}{3}\alpha^4\left(\frac{N^2-7N+40}{(N+1)N} + \frac{2(N^2-3N+44)}{3(N+1)N}\beta^2 + \beta^4\right) - \frac{4}{3}\alpha^2\left(\frac{N^2-11N+60}{(N+1)N} - \frac{N-8}{N}\beta^2 - \frac{N^2-3N-28}{(N+1)N}\beta^4 + \beta^6\right) + \frac{1}{9}(1-\beta^2)^2\left(\frac{N^2-15N+80}{(N+1)N} - \frac{2(N-8)}{N}\beta^2 + \beta^4\right)\Bigg]. \tag{6.63}$$

These represent only a small subset of the total solution space of dimension $(N-5)(N-7)/8$, as testified by its very limited exploration of the dependence in α.

Having cancelled the forward limit and LO CFF of a DD with correct parity, we merely have to verify a final criterion when referring to our definition of a shadow GPD in Section 6.1, that is the behaviour at $x = 1$. It is however made unnecessary by the observation that GPDs generated by Eq. (6.63) are already factorised by $(1-x)^3$. We observe that as we keep adding more constraints to our shadow GPDs, they get even factorised by larger powers of $(1-x)$, making this constraint essentially void.

The existence of infinitely many GPDs with a null LO CFF at a given scale was already known and expressed in different forms – see for instance, Ref. [226]. However, to the best of our knowledge the vanishing of the LO CFF has never been studied before in conjunction with the vanishing of the forward limit of the GPD. Moreover, the advantage of the analysis we have developed here will become obvious when extended at NLO. The general strategy of reducing the problem to linear set of $\mathcal{O}(N)$ equations although we have at our disposal $\mathcal{O}(N^2)$ coefficients will allow us to demonstrate the infinite number of NLO shadow GPDs, as well as to explicit some of them.

6.2.4 . NLO shadow GPDs

Considering the inflation of the size of formulas as we implemented further constraints, one can legitimately fear about the results once we actually consider NLO CFFs, which are true integrals, and not point-wise information on the GPD as the forward limit and LO CFF were. Actually, as long as one does not intend to provide a full analytic description of shadow GPDs, and sticks to the minimal programme of demonstrating their existence and exhibiting a few realisations, things are not that bad.

At NLO, the DVCS coefficient function is composed of two parts in addition to the LO part (see for instance Ref. [149]) :

$$T_1^a\left(\frac{x}{\xi}, \frac{Q^2}{\mu^2}, \alpha_s(\mu^2)\right) = T_0^a\left(\frac{x}{\xi}\right) + \alpha_s(\mu^2)\log\left(\frac{Q^2}{\mu^2}\right)T_{1,0}^a\left(\frac{x}{\xi}\right) + \alpha_s(\mu^2)T_{1,1}^a\left(\frac{x}{\xi}\right). \tag{6.64}$$

We devote a significant attention to the role of these terms in Section 6.3. For now, let us take them as granted, and focus first on $T_{1,0}^q$, which is also often called collinear term. Explicit calculations demonstrated in Appendix A.3 yield (A.48)

$$\mathrm{Im}\, T_{1,0}^q\left(\frac{x}{\xi}\right) \otimes H^q(x,\xi,\mu^2) = \frac{e_q^2 C_F}{2}\left(\left[\frac{3}{2} + \log\left(\frac{1-\xi}{2\xi}\right)\right]H^{q(+)}(\xi,\xi) + \int_\xi^1 \mathrm{d}x\, \frac{H^{q(+)}(x,\xi) - H^{q(+)}(\xi,\xi)}{x-\xi}\right). \tag{6.65}$$

For a β-odd polynomial DD, the integral can be written as

$$\int_\xi^1 \mathrm{d}x\, \frac{H^{q(+)}(x,\xi) - H^{q(+)}(\xi,\xi)}{x-\xi} = \sum_{w=1}^{N+1} \frac{1}{(1+\xi)^w} \sum_{u=w,v=0}^{N+1} D_{w,\underline{u}} q_{\underline{u}}\,, \tag{6.66}$$

where a new matrix of constraints is introduced [12]

$$D_{w,\underline{u}} = (-1)^{u+v+w} \sum_{k=1}^{v} \frac{(-1)^k}{k} \binom{v-k}{u-w} - \frac{1}{k} \binom{v}{u-w}\,. \tag{6.67}$$

It is remarkable that the polynomial DD representation allows us to decompose $\operatorname{Im} T^q_{1,0} \otimes H^{q(+)}$ into a rational fraction (6.66) with poles of finite order at $\xi = -1$ and a logarithmic singularity $\log\left(\frac{1-\xi}{2\xi}\right)$ factorised by the previously studied term $H^{q(+)}(\xi,\xi)$. It means that we can build on the solution found at previous order to immediately cancel the logarithmic term, and only need to preoccupy ourselves with the linear system formed by $(D_{w,\underline{u}}) \cdot (q_{\underline{u}})$, which brings at most $N+1$ new constraints. This behaviour is also observed when studying the real part of the LO CFF, which involves likewise a separation into logarithmic singularities factorised by $H^{q(+)}(\xi,\xi)$ and a finite number of poles – which are exactly cancelled by the vanishing of the imaginary part by virtue of the dispersion relation.

The situation is also encountered for the term $T^q_{1,1}$ – the 1-loop term, by noting that if $H^{q(+)}(\xi,\xi) = 0$,

$$\begin{aligned}\operatorname{Im} T^q_{1,1}\left(\frac{x}{\xi}\right) \otimes H^q(x,\xi,\mu^2) = \log\left(\frac{1-\xi}{2\xi}\right) \operatorname{Im} T^q_{1,0}\left(\frac{x}{\xi}\right) \otimes H^q(\xi,t,\mu^2) \\ + \frac{e_q^2 C_F}{4} \sum_{w=1}^{N-1} \frac{1}{(1+\xi)^w} \sum_{u=w,v=0}^{N+1} E_{w,\underline{u}} q_{\underline{u}}\,,\end{aligned} \tag{6.68}$$

where

$$\begin{aligned}E_{w,\underline{u}} = (-1)^{u+w} \sum_{k=1}^{v} \frac{3}{k} \binom{v-k}{u-w} - \frac{3(-1)^k}{k} \binom{v}{u-w} \\ - \frac{2(-1)^v}{k} \sum_{j=1}^{k} \frac{(-1)^j}{j} \binom{v-j}{u-w} - \frac{1}{j} \binom{v}{u-w}\,.\end{aligned} \tag{6.69}$$

Again, we can (and must) build on the previous cancellation of the collinear term to cancel the 1-loop term. The latter then brings at most $N-1$ new linear constraints.

We also note that a GPD resulting from a DD $f^{q(+)}$ may exhibit a discontinuity at $(x,\xi) = (1,1)$. The reasoning is very similar to the one we led about the discontinuity at $(x,\xi) = (0,0)$, and the reason is essentially the same. While the discontinuity at $(x,\xi) = (0,0)$ was a consequence of a discontinuity of the DD on the $\beta = 0$ line, the

12. This formula is obtained in a similar fashion as Eq. (6.43), that is first by inserting the expression of the GPD $H^{q(+)}(x,\xi)$ (6.22) in the left hand side (l.h.s.) of Eq. (6.66), then performing the partial fractions decomposition, and finally simplifying the expression by considering only the $q_{u,v}$ coefficients which can actually result from the Radon transform of a polynomial DD.

discontinuity at $(x,\xi) = (1,1)$ is a consequence of the fact that the DD does not go smoothly to 0 on the edges of the rhombus. For instance, for any $\lambda \geq 1$,

$$\lim_{\varepsilon\to 0} H^{q(+)}\left(1-\frac{\varepsilon}{\lambda}, 1-\varepsilon\right) = \int_0^{1/\lambda} \mathrm{d}\alpha\, f^{q(+)}\left(1-\alpha,\alpha\right) , \tag{6.70}$$

and the limit at $(x,\xi) = (1,1)$ may depend on the actual path to $(1,1)$ in the (x,ξ)-plane, unless $f^{q(+)}(1-\alpha,\alpha) = 0$. This adds another set of equations on the DD with

$$f^{q(+)}(1-\alpha,\alpha) = \sum_{\substack{m,n=0\\ m \text{ even}}}^{m+n\leq N} c_{m,n}\alpha^m(1-\alpha)^n = \sum_{w=0}^{N}\alpha^w \sum_{\substack{m=w,n=0\\ m \text{ even}}}^{m+n\leq N} B_{w,\underline{m}}\, c_{\underline{m}} , \tag{6.71}$$

where

$$B_{w,\underline{m}} = (-1)^w \binom{n}{w-m} . \tag{6.72}$$

Notice that working in the Pobylitsa representation of DDs spares us from the trouble of implementing this condition, since the $(1-x)$ factor introduced by this alternative representation guarantees better regularity in the limit $x \to 1$.

Adding all these constraints means that the order N at which a non zero solution can be found increases. When asking only the cancellation of the LO CFF, collinear term and 1-loop term, first solutions arise at $N = 17$. With the further requirements that the forward limit vanishes as well as the DD on the edges of the rhombus, the first solutions are found for $N = 25$. One is displayed in its full complexity in Appendix A.5, and depicted as the dotted brown curve in Fig. 6.4 where it is added to the popular phenomenological GK model. Another NLO shadow GPD of order 27 is also depicted as the dashed orange curve. Therefore, at scale μ_F^2 where the shadow NLO GPDs are defined, the three models have exactly the same forward limit and full NLO CFF.

In spite of the large polynomial order, the shadow GPDs do not oscillate in a way *a priori* excluded on physical grounds. Additionally, they exhibit very different shapes in both $|x| > |\xi|$ and $|x| < |\xi|$ regions, illustrating already on this very small subset of the solution space that NLO shadow GPDs can produce a large range of different effects.

We foresee that the argument of the relative increase of the number of constraints and of free parameters can be extended *mutatis mutandis* to guarantee the existence of shadow gluon GPDs and more generally of shadow GPDs at any finite order in pQCD. Having not searched for non-trivial shadow gluon GPDs, we will work with the solution $(H_0^{q(+)}, H_0^{g(+)} = 0)$ where $H_0^{q(+)}$ is a non-trivial shadow quark GPD to study the effect of evolution. 0 is obviously a shadow GPD, so the pair will exhibit all required properties.

6.3 . The effect of evolution on shadow GPDs

The size of shadow GPDs that can remain invisible in DVCS data is inversely proportional to the maximal contribution they bring to CFFs on the probed Q^2 range. By design, this contribution is exactly 0 at the definition scale of the shadow distribution μ_F^2. If no evolution effects are taken into account, the problem is then formally not invertible, and infinite shadow GPDs will plague the extraction [13]. If however shadow GPDs bring

13. Neglecting the effect of positivity bounds which we will consider in the next chapter.

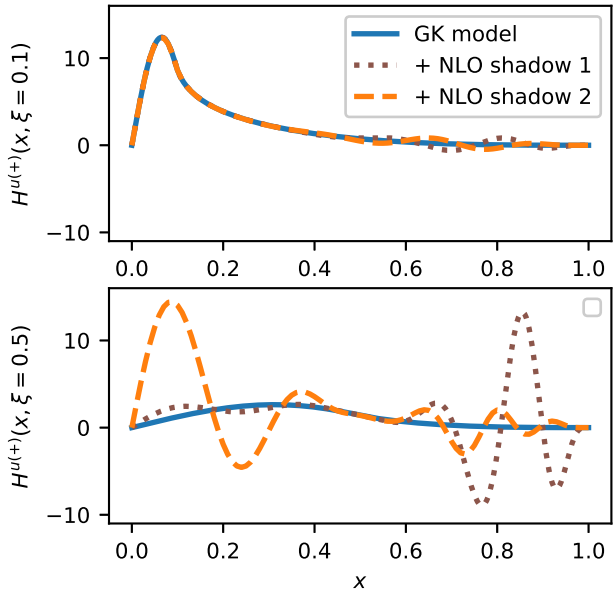

Figure 6.4 – $H^{u(+)}$ as a function of x for $\xi = 0.1$ and 0.5, $t = -0.1$ GeV^2 and $\mu_F^2 = 1$ GeV^2. Solid blue : GK model. Dashed orange and dotted brown : GK model with the addition of two different NLO shadow GPDs. In all cases one obtains exactly the same NLO CFF and forward limit at scale μ_F^2. Figure taken from Ref. [6].

sizable contributions to the CFF as soon as Q^2 strays from μ_F^2, then they are relatively easy to discriminate and will not present a serious issue to the deconvolution.

By design, our construction of shadow GPDs cancels all terms of order $\mathcal{O}(\alpha_s)$ in the DVCS coefficient function. We would therefore expect any contribution of the shadow GPDs to the CFF to be of order α_s^2. However, considering that we lead this study with EIC kinematics in mind, that is very roughly speaking a range in Q^2 of the order of 1 to 100 GeV^2, the size of the logarithmic corrections generated by the rather large evolution range might spoil this $\mathcal{O}(\alpha_s^2)$ behaviour, and generate larger than expected contributions to the CFF.

Let us note that by neutralising the collinear term $T_{1,0}^q$ with the system (6.67), we have already cancelled a part of the logarithmic corrections. Before investigating numerically the effect of the resummation of the full leading logarithmic corrections, we would like to present in a bit more precise way their origin, and clarify the interplay between fixed order truncation and leading logarithm expansion. This will allow us to understand more clearly what is the theoretical origin of the residual effect of evolution on shadow GPDs once all terms of order $\mathcal{O}(\alpha_s)$ have been cancelled. Although the results shown in the next section are globally already known and belong to the common knowledge of the behaviour of the perturbative expansion, they have not been frequently spelled out in detail, and particularly not in the context of GPD phenomenology – probably mainly due to the fact that currently probed DVCS experimental range in Q^2 is limited enough that evolution effects at high order were deemed secondary until now. Similar aspects to those discussed below are found for instance in Ref. [227] in the case of DIS, although presented in less detail than we do now.

6.3.1 . Combining evolution equations and the DVCS coefficient function

The link between truncation of the renormalisation group equation (RGE) at fixed order in α_s and the leading logarithm (LL) expansion is probably the most clearly seen when solving the RGE of α_s itself, which we presented already in Eq. (2.53). We remind that

$$\frac{\partial \alpha_s(\mu^2)}{\partial \log \mu^2} = \beta(\alpha_s(\mu^2)) = -b_0 \alpha_s^2(\mu^2) - b_1 \alpha_s^3(\mu^2) - \dots , \tag{6.73}$$

The general solution can be written under the form of a Taylor expansion [14]

$$\alpha_s(Q^2) = \sum_{j=0}^{\infty} \frac{1}{j!} \log^j \left(\frac{Q^2}{\mu^2} \right) \frac{\partial^j \alpha_s(\mu^2)}{\partial \log^j(\mu^2)} , . \tag{6.78}$$

Differentiating again Eq. (6.73) yields

$$\frac{\partial^2 \alpha_s(\mu^2)}{\partial \log^2 \mu^2} = \left[-2b_0 \alpha_s(\mu^2) - 3b_1 \alpha_s^2(\mu^2) - \dots \right] \frac{\partial \alpha_s(\mu^2)}{\partial \log \mu^2} , \tag{6.79}$$

$$= 2b_0^2 \alpha_s^3(\mu^2) + 5 b_0 b_1 \alpha_s^4(\mu^2) + \dots \tag{6.80}$$

A trivial recursion gives immediately that the term of lowest order in the expansion of $\partial^j \alpha_s(\mu^2)/\partial \log^j \mu^2$ is $j!\, b_0^j \alpha_s^{j+1}(\mu^2)$. It is therefore possible to introduce $\eta_{j,l}$ coefficients depending on $b_0, b_1, \dots$ such that

$$\alpha_s(Q^2) = \alpha_s(\mu^2) + \sum_{j=2}^{\infty} \alpha_s^j(\mu^2) \left[\log^{j-1} \left(\frac{Q^2}{\mu^2} \right) (-b_0)^{j-1} + \sum_{l=2}^{j-1} \log^{j-l} \left(\frac{Q^2}{\mu^2} \right) \eta_{j,l} \right] . \tag{6.81}$$

One observes therefore that **the mere form of the RGE gives rise to an expansion of the solution both in terms of powers of α_s and logarithmic corrections**. Let us introduce the notation

$$L = \log \left(\frac{Q^2}{\mu^2} \right) . \tag{6.82}$$

14. The mixture of dependences in μ^2 and $\log(\mu^2)$ might appear displeasing. Either one can elude the question by considering that $\alpha_s(\mu^2)$ is a short-hand notation for $\alpha_s(\log(\mu^2))$, in which case Eq. (6.78) is just a standard Taylor expansion in the variable $\log(\mu^2)$, or one can notice the following implication of the fact that f admits a series expansion in terms of powers of g :

$$f(x) = \sum_{k=0}^{\infty} a_k g^k(x) \implies f(x) = \sum_{j=0}^{\infty} \frac{(g(x) - g(x_0))^j}{j!} \frac{d^j f(x_0)}{dg^j(x)} , \tag{6.74}$$

which corresponds exactly to the result we are interested in with $f = \alpha_s$, $x = Q^2$, $x_0 = \mu^2$ and $g = \log$. Indeed

$$\frac{\mathrm{d}^j f(x_0)}{\mathrm{d}g^j(x)} = \sum_{k=j}^{\infty} a_k \frac{k!}{(k-j)!} g^{k-j}(x_0) , \tag{6.75}$$

and

$$\sum_{j=0}^{\infty} \frac{(g(x) - g(x_0))^j}{j!} \frac{d^j f(x_0)}{dg(x)^j} = \sum_{j=0}^{\infty} \frac{(g(x) - g(x_0))^j}{j!} \sum_{k=j}^{\infty} a_k \frac{k!}{(k-j)!} g^{k-j}(x_0) , \tag{6.76}$$

$$= \sum_{k=0}^{\infty} a_k \sum_{j=0}^{k} \binom{k}{j} (g(x) - g(x_0))^j g^{k-j}(x_0) = \sum_{k=0}^{\infty} a_k g^k(x) = f(x) . \tag{6.77}$$

.

The LL terms in Eq. (6.81) are therefore $\alpha_s^j(\mu^2)(-Lb_0)^{j-1}$ for $j \geq 1$. Their resummation to infinity while ignoring sub-leading logarithms yields

$$\alpha_s(Q^2) \overset{LL}{=} \frac{\alpha_s(\mu^2)}{1+\alpha_s(\mu^2)Lb_0}\,. \tag{6.83}$$

On the other hand, solving directly the RGE truncated at LO $\partial\alpha_s(\mu^2)/\partial\log(\mu^2) = -b_0\alpha_s^2(\mu^2)$ gives

$$\alpha_s(Q^2) \overset{LO}{=} \frac{1}{b_0\log(Q^2/\Lambda^2)}\,, \tag{6.84}$$

for some integration constant Λ, which corresponds to the Landau pole of divergence of the perturbative strong coupling at low scale. It is easy to verify that Eqs. (6.83) and (6.84) are exactly equal, so **resumming the LL expansion of α_s is equivalent to exactly solving its LO evolution equation**.

General solution of GPD evolution

GPDs follow a RGE whose general form is close to the one followed by α_s itself. Therefore, the general features of the previous paragraph still apply. The singlet GPD RGE can be written as

$$\frac{\partial H^{a(+)}(\mu^2)}{\partial\log(\mu^2)} = \sum_b \left[\alpha_s(\mu^2)K^{ab,(0)} + \alpha_s^2(\mu^2)K^{ab,(1)} + ...\right] \otimes H^{b(+)}(\mu^2)\,, \tag{6.85}$$

where a and b stand for the various parton types, and $K^{ab,(j)}$ are the splitting kernels computed at N^{j}LO. The $\otimes$ notation serves as a compact way to represent the integral depicted fully in Eq. (2.52) as in the case of the coefficient function. There again, the general solution is found as

$$H^{a(+)}(Q^2) = \sum_{j=0}^{\infty} \frac{L^j}{j!}\frac{\partial^j H^{a(+)}(\mu^2)}{\partial\log^j(\mu^2)}\,. \tag{6.86}$$

The differentiation of Eq. (6.85) gives a slightly more tedious expression compared to Eq. (6.80)

$$\begin{aligned}\frac{\partial^2 H^{a(+)}(\mu^2)}{\partial\log^2(\mu^2)} &= \sum_b \frac{\partial\alpha_s(\mu^2)}{\partial\log(\mu^2)}\left(K^{ab,(0)} + 2\alpha_s(\mu^2)K^{ab,(1)} + ...\right)\otimes H^{b(+)}(\mu^2) + \Big(\alpha_s(\mu^2)K^{ab,(0)} \\ &\quad + \alpha_s^2(\mu^2)K^{ab,(1)} + ...\Big)\otimes\sum_c\left(\alpha_s(\mu^2)K^{bc,(0)} + \alpha_s^2(\mu^2)K^{bc,(1)} + ...\right)\otimes H^{c(+)}(\mu^2)\,, \qquad (6.87)\\ &= \sum_b\left[-\alpha_s^2(\mu^2)b_0K^{ab,(0)} + \sum_c\alpha_s^2(\mu^2)K^{ac,(0)}\otimes K^{cb,(0)}\right]\otimes H^{b(+)}(\mu^2) + \mathcal{O}(\alpha^3(\mu^2))\,. \qquad (6.88)\end{aligned}$$

Higher order derivatives are of order α_s^3 at least, so we deduce

$$H^{a(+)}(Q^2) = \sum_b \Bigg[\delta_{ab} + \alpha_s(\mu^2) L K^{ab,(0)} + \frac{1}{2}\alpha_s^2(\mu^2) L^2 \left(-b_0 K^{ab,(0)} + \sum_c K_{ac}^{(0)} \otimes K_{cb}^{(0)}\right) + \alpha_s^2(\mu^2) L K^{ab,(1)}\Bigg] \otimes H^{b(+)}(\mu^2) + \mathcal{O}(\alpha_s^3)\,, \tag{6.89}$$

where δ_{ab} is the Kronecker symbol. In general, a recursion on higher order derivatives of $H^{a(+)}$ allows us to conclude that the GPD evolution operator defined by

$$H^{a(+)}(Q^2) = \sum_b \Gamma^{ab,(k)}(Q^2, \mu^2) \otimes H^{b(+)}(\mu^2) \tag{6.90}$$

can be expressed at order α_s^k thanks to the introduction of the $\gamma_{j,l}^{ab}$ coefficients such that

$$\Gamma^{ab,(k)}(Q^2, \mu^2) = \delta_{ab} + \sum_{j=1}^{k} \alpha_s^j(\mu^2) \sum_{l=0}^{j-1} L^{j-l} \gamma_{j,l}^{ab}\,. \tag{6.91}$$

Note that apart from the initial δ_{ab} term, all terms come with a non-zero power of log. Once again, the double expansion in terms of powers of α_s and logarithmic corrections come from the fact that the RGE takes the generic form $\partial f(\mu^2)/\partial \log(\mu^2) =$ polynomial in $\alpha_s(\mu^2) \otimes f(\mu^2)$.

DVCS coefficient function and the requirements of consistency of evolution

The DVCS coefficient function shows a very similar form with also a double expansion in α_s and L. When expanded at order α_s^k, it writes in general as

$$T_k^a\left(\omega, \frac{Q^2}{\mu^2}, \alpha_s(\mu^2)\right) = \sum_{j=0}^{k} \alpha_s^j(\mu^2) \sum_{l=0}^{j} L^{j-l} T_{j,l}^a(\omega)\,, \tag{6.92}$$

$$= T_{0,0}^a(\omega) + \alpha_s(\mu^2) \log\left(\frac{Q^2}{\mu^2}\right) T_{1,0}^a(\omega) + \alpha_s(\mu^2) T_{1,1}^a(\omega) + \dots\,, \tag{6.93}$$

where $T_{j,l}^a(\omega)$ is a complex-valued distribution, and we have highlighted the beginning of the DVCS coefficient function expansion which we have precisely cancelled for our NLO shadow GPDs. The LL terms read $\alpha_s^j(\mu^2) L^j T_{j,0}^a(\omega)$. Contrary to the case of the evolution operator, there are also terms with no logarithm power, namely $\alpha_s^j(\mu^2) T_{j,j}^a(\omega)$. We will call $T_{0,0}^a(\omega)$ the tree-level term, and $T_{j,j}^a(\omega)$ for $j \geq 1$ the j-loop term. They are computed from j-loop perturbative diagrams of the hard scattering of the DVCS process.

We will now demonstrate that the independence of CFFs on the factorisation scale – up to sub-leading contributions of order $\mathcal{O}(\alpha_s^{k+1})$ if the CFFs are computed at order α_s^k – implies that the terms with non-zero logarithm power in the expansion of the DVCS coefficient function (6.92) are entirely determined by the interplay of the j-loop terms $T_{j,j}^a(\omega)$ and the coefficients $\gamma_{j,l}^{ab}$ of the expansion of the evolution operator (6.91). The requirement of independence of CFFs on the factorisation scale can be translated into the

following equation

$$T_k^a\left(\omega,\frac{Q^2}{\mu^2},\alpha_s(\mu^2)\right)\otimes H^{a(+)}(\mu^2)=T_k^a\left(\omega,1,\alpha_s(Q^2)\right)\otimes H^{a(+)}(Q^2)+\mathcal{O}(\alpha_s^{k+1})\,, \tag{6.94}$$

$$=\sum_b T_k^a\left(\omega,1,\alpha_s(Q^2)\right)\otimes\Gamma^{ab,(k)}(Q^2,\mu^2)\otimes H^{b(+)}(\mu^2)+\mathcal{O}(\alpha_s^{k+1})\,. \tag{6.95}$$

Expansion at order α_s^2 To get an intuition of what is happening, let us first work at fixed order α_s^2. We will show that the adequacy of the left hand side (l.h.s.) and right hand side (r.h.s.) of Eq. (6.95) brings mutual constraints between the evolution equation and DVCS coefficient function.

The r.h.s. of Eq. (6.95) involves the expression

$$\sum_b T_k^a\left(\omega,1,\alpha_s(Q^2)\right)\otimes\Gamma^{ab,(k)}(Q^2,\mu^2)\otimes H^{b(+)}(\mu^2)$$

$$=\sum_b\left[T_{0,0}^a+\alpha_s(Q^2)T_{1,1}^a+\alpha_s^2(Q^2)T_{2,2}^a\right]\otimes\Bigg[\delta_{ab}+\alpha_s(\mu^2)L\gamma_{1,0}^{ab}+\alpha_s^2(\mu^2)L^2\gamma_{2,0}^{ab}$$

$$+\alpha_s^2(\mu^2)L\gamma_{2,1}^{ab}\Bigg]\otimes H^{b(+)}(\mu^2)+\mathcal{O}(\alpha_s^3) \tag{6.96}$$

where we have used the expressions (6.92) and (6.91). Then using that $\alpha_s(Q^2)=\alpha_s(\mu^2)-\alpha_s^2(\mu^2)Lb_0+\mathcal{O}(\alpha_s^3)$ and developing the expression,

$$\sum_b T_k^a\left(\omega,1,\alpha_s(Q^2)\right)\otimes\Gamma^{ab,(k)}(Q^2,\mu^2)\otimes H^{b(+)}(\mu^2)$$

$$=\sum_b\Bigg[T_{0,0}^a\otimes\delta_{ab}+\alpha_s(\mu^2)LT_{0,0}^a\otimes\gamma_{1,0}^{ab}+\alpha_s(\mu^2)T_{1,1}^a\otimes\delta_{ab}+\alpha_s^2(\mu^2)L^2T_{0,0}^a\otimes\gamma_{2,0}^{ab}$$

$$+\alpha_s^2L(-b_0T_{1,1}^a\otimes\delta_{ab}+T_{1,1}^a\otimes\gamma_{1,0}^{ab}+T_{0,0}\otimes\gamma_{2,1})+\alpha_s^2(\mu^2)T_{2,2}^a\otimes\delta_{ab}\Bigg]\otimes H^{b(+)}(\mu^2)+\mathcal{O}(\alpha_s^3)\,. \tag{6.97}$$

On the other hand, the l.h.s of Eq. (6.95) gives

$$T_k^a\left(\omega,\frac{Q^2}{\mu^2},\alpha_s(\mu^2)\right)\otimes H^{a(+)}(\mu^2)=\sum_b\Bigg[T_{0,0}^a+\alpha_s(\mu^2)\Big(LT_{1,0}^a+T_{1,1}^a\Big)$$

$$+\alpha_s^2(\mu^2)\Big(L^2T_{2,0}^a+LT_{2,1}^a+T_{2,2}^a\Big)\Bigg]\otimes\delta_{ab}\otimes H^{b(+)}(\mu^2)\,. \tag{6.98}$$

Therefore, identification of Eqs. (6.97) and (6.98) leads to the following relations at order α_s^2

$$\begin{cases}T_{1,0}^a\otimes\delta_{ab}=T_{0,0}^a\otimes\gamma_{1,0}^{ab}\\ T_{2,0}^a\otimes\delta_{ab}=T_{0,0}^a\otimes\gamma_{2,0}^{ab}\\ T_{2,1}^a\otimes\delta_{ab}=T_{1,1}^a\otimes\gamma_{1,0}^{ab}-b_0T_{1,1}^a\otimes\delta_{ab}+T_{0,0}^a\otimes\gamma_{2,1}^{ab}\,.\end{cases} \tag{6.99}$$

One sees on this example that $\mathrm{N}^j\mathrm{LL}$ terms $T_{k,j}^a$ are exactly obtained from convolutions of $T_{l,l}^a$ – that is what we called l-loop terms before – where $0\leq l\leq j$ with evolution operators.

General result For simplicity, we will only derive explicitly the constraints induced by Eq. (6.95) on the LL and NLL terms of the DVCS coefficient function. The method can be applied straightforwardly to any order in the leading logarithms expansion. The r.h.s. of Eq.(6.95) involves the expression

$$\begin{aligned}&\sum_b T_k^a\left(\omega,1,\alpha_s(Q^2)\right)\otimes\Gamma^{ab,(k)}(Q^2,\mu^2)\otimes H^{b(+)}(\mu^2)\\&=\sum_b\left[\sum_{m=0}^k\alpha_s^m(Q^2)T^a_{m,m}\right]\otimes\left[\delta_{ab}+\sum_{j=1}^k\alpha_s^j(\mu^2)\sum_{l=0}^{j-1}L^{j-l}\gamma^{ab}_{j,l}\right]\otimes H^{b(+)}(\mu^2)+\mathcal{O}(\alpha_s^{k+1})\,,\end{aligned}\tag{6.100}$$

using Eqs. (6.92) and (6.91). We first need to compute the LL expansion of $\alpha_s^m(Q^2)$ involved in the first term of Eq. (6.100), which is obtained from the LL terms of Eq. (6.81) as

$$\alpha_s^m(Q^2)\overset{LL}{=}\sum_{n=m}^\infty\alpha_s^n(\mu^2)L^{n-m}(-b_0)^{n-m}\binom{n-1}{m-1}\,.\tag{6.101}$$

Therefore, the LL term of $\alpha_s^m(Q^2)$ is of order $\alpha_s^n(\mu^2)L^{n-m}$ and has a "deficit" of m powers of logarithms compared to the LL term in the DVCS coefficient function, which is of order $\alpha_s^n(\mu^2)L^n$. Only $m=0$ in the first term of Eq. (6.100) contributes therefore to the LL term in the DVCS coefficient function. It must be associated to $l=0$ in the second term of Eq. (6.100) to indeed yield a LL term.

NLL terms are obtained either with $m=0$ and $l=1$, or $m=1$ and $l=0$. By applying these counting rules in Eq. (6.100), we obtain the NLL expansion

$$\begin{aligned}&\sum_b T_k^a\left(\omega,1,\alpha_s(Q^2)\right)\otimes\Gamma^{ab,(k)}(Q^2,\mu^2)\otimes H^{b(+)}(\mu^2)\\&\overset{NLL}{=}\sum_b\left[T^a_{0,0}\otimes\left(\delta_{ab}+\sum_{j=1}^k\alpha_s^j(\mu^2)L^j\gamma^{ab}_{j,0}\right)+T^a_{0,0}\otimes\left(\sum_{j=2}^k\alpha_s^j(\mu^2)L^{j-1}\gamma^{ab}_{j,1}\right)\right.\\&\left.+\sum_{n=1}^k\alpha_s^n(\mu^2)L^{n-1}(-b_0)^{n-1}T^a_{1,1}\otimes\left(\delta_{ab}+\sum_{j=1}^k\alpha_s^j(\mu^2)L^j\gamma^{ab}_{j,0}\right)\right]\otimes H^{b(+)}(\mu^2)+\mathcal{O}(\alpha_s^{k+1})\,.\end{aligned}\tag{6.102}$$

Mind that the final sum of the second line starts at $j=2$ due to the fact that $\gamma^{ab}_{1,1}$ is undefined with the formula of Eq. (6.91). The sum can be started at $j=1$ by conveniently defining $\gamma^{ab}_{1,1}=0$. A slight reshuffling of the third line of Eq. (6.103) yields

$$\begin{aligned}\sum_b T_k^a\left(\omega,1,\alpha_s(Q^2)\right)\otimes\Gamma^{ab,(k)}(Q^2,\mu^2)\otimes H^{b(+)}(\mu^2)&\overset{NLL}{=}\sum_b\left[T^a_{0,0}\otimes\left(\delta_{ab}+\sum_{j=1}^k\alpha_s^j(\mu^2)L^j\gamma^{ab}_{j,0}\right)\right.\\&+T^a_{0,0}\otimes\left(\sum_{j=2}^k\alpha_s^j(\mu^2)L^{j-1}\gamma^{ab}_{j,1}\right)+\sum_{n=1}^k\alpha_s^n(\mu^2)L^{n-1}(-b_0)^{n-1}T^a_{1,1}\otimes\delta_{ab}\\&\left.+\sum_{p=2}^k\alpha_s^p(\mu^2)L^{p-1}\sum_{j=1}^{p-1}(-b_0)^{p-j-1}T^a_{1,1}\otimes\gamma^{ab}_{j,0}\right]\otimes H^{b(+)}(\mu^2)+\mathcal{O}(\alpha_s^{k+1})\,.\end{aligned}\tag{6.103}$$

By identifying with the l.h.s. of Eq. (6.95), we obtain the following conditions

$$\begin{cases} T^a_{j,0} \otimes \delta_{ab} = T^a_{0,0} \otimes \gamma^{ab}_{j,0} \text{ for } j \geq 1\,, \\ T^a_{j,1} \otimes \delta_{ab} = \sum_{m=1}^{j-1} (-b_0)^{j-m-1} T^a_{1,1} \otimes \gamma^{ab}_{m,0} + (-b_0)^{j-1} T^a_{1,1} \otimes \delta_{ab} + T^a_{0,0} \otimes \gamma^{ab}_{j,1} \text{ for } j \geq 2\,. \end{cases} \tag{6.104}$$

The result for $j \leq 2$ is in agreement with the previously found results of Eq. (6.99). This system demonstrates that **the LL DVCS coefficient $T^a_{j,0}(\omega)$ is exactly the tree-level term $T^a_{0,0}(\omega)$ convoluted with the LL GPD evolution.** Resumming the entire NLL expansion of the DVCS coefficient function thanks to Eq. (6.104) yields

$$\begin{aligned} \sum_{j=1}^{\infty} \alpha_s^j(\mu^2) L^{j-1} T^a_{j,1} \otimes \delta_{ab} &= \alpha_s(\mu^2) T^a_{1,1} \otimes \delta_{ab} + \sum_{j=2}^{\infty} \alpha_s^j(\mu^2) L^{j-1} \sum_{m=1}^{j-1} (-b_0)^{j-m-1} T^a_{1,1} \otimes \gamma^{ab}_{m,0} \\ &\quad + \sum_{j=2}^{\infty} \alpha_s^j(\mu^2) L^{j-1} (-b_0)^{j-1} T^a_{1,1} \otimes \delta_{ab} + T^a_{0,0} \otimes \sum_{j=2}^{\infty} \alpha_s^j(\mu^2) L^{j-1} \gamma^{ab}_{j,1} \end{aligned} \tag{6.105}$$

$$\begin{aligned} &= \alpha_s(\mu^2) T^a_{1,1} \otimes \delta_{ab} + \frac{\alpha_s(\mu^2)}{1 + \alpha_s(\mu^2) L b_0} \sum_{m=1}^{\infty} \alpha_s^m(\mu^2) L^m T^a_{1,1} \otimes \gamma^{ab}_{m,0} \\ &\quad - \frac{\alpha_s^2(\mu^2) L b_0}{1 + \alpha_s(\mu^2) L b_0} T^a_{1,1} \otimes \delta_{ab} + T^a_{0,0} \otimes \sum_{j=2}^{\infty} \alpha_s^j(\mu^2) L^{j-1} \gamma^{ab}_{j,1} \end{aligned} \tag{6.106}$$

$$= \frac{\alpha_s(\mu^2)}{1 + \alpha_s(\mu^2) L b_0} T^a_{1,1} \otimes \left(\delta_{ab} + \sum_{m=1}^{\infty} \alpha_s^m L^m \gamma^{ab}_{m,0} \right) + T^a_{0,0} \otimes \sum_{j=2}^{\infty} \alpha_s^j(\mu^2) L^{j-1} \gamma^{ab}_{j,1}\,. \tag{6.107}$$

The final term is the tree-level term convoluted with the NLL GPD evolution, whereas the first term is the 1-loop term convoluted with the LL GPD evolution. The factor in front of $T^a_{1,1}$, which we have already met in Eq. (6.83), is actually the LL expansion of $\alpha_s(Q^2)$.

Remark We have expressed the results for the LL and NLL DVCS coefficient functions in Eq. (6.104) in terms of the $\gamma^{ab}_{k,j}$ elements of the evolution operator expansion defined in Eq. (6.91). However, this is a redundant source of information, which can be reduced to the sole $K^{ab,(i)}$ introduced in Eq. (6.85). Eq. (6.89) shows for instance that we have already derived

$$\begin{cases} \gamma^{ab}_{1,0} = K^{ab,(0)} \\ \gamma^{ab}_{2,0} = 1/2(-b_0 K^{ab,(0)} + \sum_c K^{ac,(0)} \otimes K^{cb,(0)}) \\ \gamma^{ab}_{2,1} = K^{ab,(1)}\,. \end{cases} \tag{6.108}$$

The study of the LL expansion of the evolution operator Γ^{ab} demonstrates that it is exactly the solution of the LO evolution, involving only the kernel $K^{ab,(0)}$. Similar results can be obtained for the N^{k}LL expansion, which only involves kernels up to $K^{ab,(k)}$.

Conclusion

We understand therefore that the convolution of the tree-level coefficient function $T^a_{0,0}$ and the solution of the LO evolution gives the full LL expansion of the DVCS coefficient function, by convoluting $T^a_{0,0}$ with combinations of $K^{ab,(0)}$. The full NLL expansion is obtained from the additional knowledge of $T^a_{1,1}$ and $K^{ab,(1)}$ and so on so forth.

When only considering a narrow range μ^2 around Q^2, $|L| \ll 1$ and the size of logarithmic contributions is small. Then fixed order perturbation can be envisioned [15] and we could satisfy ourselves from having cancelled all terms of order $\mathcal{O}(\alpha_s)$ in the DVCS coefficient function by construction of the shadow GPDs. Formally, within a fixed order α_s truncation, we would then have shown the deconvolution problem to be non invertible.

However, as evolution range gets significant, development in terms of leading logarithms becomes more relevant than fixed order due to the increase of L. At LL, we are led to consider an infinite resummation of terms $\alpha_s^j L^j$ and the presence of the logarithms spoils the expected $\mathcal{O}(\alpha_s^2)$ behaviour of the CFF. Having cancelled the contributions of $T^a_{0,0}$ and $T^a_{1,0}$ only in the LL expansion, our shadow GPDs will now leave an imprint on CFF. The evaluation of the numerical size of this contribution is the object of the next section. Let us note that the limit of validity of fixed order perturbation at α_s^k can be probed numerically by observing to what extend the residuals are well approximated by $A\alpha_s^{k+1}(\mu^2)$ as we will explore in the following section.

6.3.2 . Practical effect of evolution on a NLO shadow GPD

The software APFEL++ [228, 229, 1], interfaced with the PARTONS framework, provides the solution of the LO GPD evolution equation thanks to a fourth order Runge-Kutta method. The entire LL expansion of the evolution operator is therefore resummed. As we have demonstrated in the previous section, combining the tree-level DVCS coefficient function $T^a_{0,0}$ with the LL GPD evolution equation amounts to resumming the entire LL expansion of the DVCS coefficient function. Combining the 1-loop coefficient function $T^a_{1,1}$ with LL evolution gives additionally the resummation of a part of the NLL expansion of the DVCS coefficient function. The combination of the tree-level DVCS coefficient with NLL GPD evolution is missing to obtain the complete NLL expansion of the DVCS coefficient function. However, no NLL GPD evolution code in momentum space is available at the time of writing, so we will work only with a partial account of NLL contributions.

We compute the convolution of the full NLO DVCS coefficient function at $\mu^2 = Q^2$ (which makes the collinear term vanish) with the LL evolution of the shadow GPD from its definition scale μ_F^2. Thus, using Eq. (6.91) and the fact that we take the gluon shadow GPD as 0 at μ_F^2, the CFF convolution reads

$$\mathcal{H}(Q^2) = \left[T^q_{0,0} + \alpha_s(Q^2)T^q_{1,1}\right] \otimes \left[1 + \sum_{j=1}^{\infty} \alpha_s^j(\mu_F^2) \log^j\left(\frac{Q^2}{\mu_F^2}\right) \gamma^{qq}_{j,0}\right] \otimes H_0^{q(+)}(\mu_F^2)\,, \quad (6.109)$$

$$= T^q_{0,0} \otimes H_0^{q(+)}(\mu_F^2) + \alpha_s(\mu_F^2) \log\left(\frac{Q^2}{\mu_F^2}\right) T^q_{0,0} \otimes \gamma^{qq}_{1,0} \otimes H_0^{q(+)}(\mu_F^2)$$
$$+ \alpha_s(Q^2) T^q_{1,1} \otimes H_0^{q(+)}(\mu_F^2) + \mathcal{O}(\alpha_s^2)\,. \quad (6.110)$$

15. provided the terms of the higher order are numerically small and leaving aside any question related to the asymptotic nature of the perturbative series.

Since we have demonstrated that (6.99)

$$T^q_{0,0} \otimes \gamma^{qq}_{1,0} = T^q_{1,0}\,, \tag{6.111}$$

we find back the expected result that the NLO shadow GPD cancels exactly all terms of order α_s^0 and α_s^1 even under evolution, and

$$\mathcal{H}(Q^2) = \mathcal{O}(\alpha_s^2)\,. \tag{6.112}$$

We want to verify numerically the predicted behaviour of Eq. (6.112), observe the effect of logarithmic corrections which is expected to break down this property, and measure the overall size of the term $\mathcal{O}(\alpha_s^2)$ on an evolution range of 1 to 100 GeV^2. To verify Eq. (6.112), APFEL++ provides the possibility to arbitrarily set $\alpha_s(\mu_0^2)$ at some reference scale μ_0^2. By choosing particularly small values of α_s, we suppress the size of LL contributions $\alpha_s^j L^j$ where j is large (L is fixed by our choice of evolution range), so we expect to observe the prescribed quadratic behaviour. If however α_s becomes relatively large, we expect the effect of all the logarithms and larger powers of α_s to come into play and spoil the quadratic behaviour.

We show the result of evolution to $Q^2 = 100$ GeV^2 from an initial definition scale of the shadow GPD of $\mu_F^2 = 1$ GeV^2 on Fig. 6.5. We have chosen the dotted brown NLO shadow GPD depicted in Fig. 6.4 – we have of course subtracted the phenomenological model to only keep the shadow component. The continuous blue line on Fig. 6.5 represents the fit of the total (quarks + radiatively generated gluons) imaginary part of the CFF by a quadratic function of α_s. The agreement is very satisfactory as long as $\alpha_s \leq 0.15$, after which it starts to break as expected. A phenomenologically relevant value of $\alpha_s(100$ GeV$^2)$ in the $\overline{MS}$ scheme is of the order of 0.18.

The most striking feature of this graph is the extremely small absolute value of the contribution of the NLO shadow GPD to the CFF even on a considerable evolution range from 1 to 100 GeV^2. At the phenomenologically relevant value of $\alpha_s(100$ GeV$^2) \approx 0.15$, the imaginary part of the CFF is only of order 10^{-5} for $\xi = 0.1$ although the shadow GPD itself is of order 1. Provided the functional space chosen to extract GPDs can accommodate flexible enough objects to include NLO shadow GPDs like the one we are discussing here, we see therefore that a possible factor of amplification of uncertainty when going from the CFF to the GPD could reach 10^5. The conditioning of the extraction is very poor even on a very large evolution range corresponding to the future EIC.

We notice finally that the fact that contributions of quarks and gluons are of opposite sign in Fig. 6.5 is the cause of an additional suppression of the effect of evolution by a factor 5 or more. This fact is already well-known, see for instance Refs. [124, 149]. The very recent publication in Ref. [126] of NNLO DVCS coefficient functions observes a similar phenomenon.

6.3.3 . Consequences of the deconvolution problem for the phenomenology of GPDs

The developments on shadow GPDs presented in this chapter demonstrate that there exist spaces with extremely poor conditioning of the extraction of GPDs from DVCS data in an NLO framework. This result is true even over kinematic domains relevant for the experiments at JLab, CERN, EIC or EIcC. The form of the GPDs living in these poorly constrained spaces has been carefully selected on general theoretical principles to make

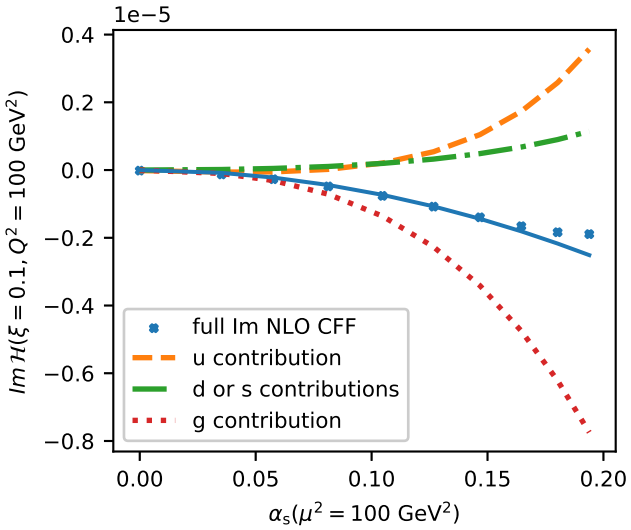

Figure 6.5 – Imaginary part of the NLO CFF $\mathcal{H}(\xi = 0.1, Q^2 = 100\ \mathrm{GeV}^2)$ evaluated with the NLO shadow GPD 1 shown in Fig. 6.4 and $H^{d(+)} = H^{s(+)} = H^{g(+)} = 0$ at $\mu_F^2 = 1\ \mathrm{GeV}^2$. The blue dots correspond to computations made with different values of $\alpha_s(Q^2 = 100\ \mathrm{GeV}^2)$ and the solid blue line results from a quadratic fit to the first seven points. The dashed orange, dash-dotted green and dotted brown lines indicate u, d or s, and g contributions to the CFF, respectively. Figure taken from Ref. [6].

sure these objects remain relevant for the usual phenomenology of GPDs. This will be demonstrated by the prominent role they will play in different modelling proposals in the next chapter.

Even if lattice QCD (see Section 3.2) might in the long term provide an evaluation of the functional shape of GPDs, experimental data – and particularly DVCS data – are, and will remain for years, the main source of quantitative knowledge about GPDs.

Our results on DVCS can straightforwardly be extended to TCS and LO DVMP due to the similarity of the coefficient functions involved in these processes. The evaluation of the first few Mellin moments from lattice QCD is not expected to qualitatively modify the answer we have provided here to the deconvolution problem. In terms of fits to experimental data, only a multichannel analysis beyond leading order, over large kinematic domains accessible in collider experiments, and within a complete framework such as PARTONS or GeParD may provide the needed leverage to quantitatively constrain GPDs. Double DVCS or processes where the active parton emits several particles provide a richer kinematic dependence of the structure functions which could radically change the considerations on the deconvolution problem. Notably, DDVCS offers a direct access to GPDs at $x \neq \xi$ at LO and seems a natural candidate to make the deconvolution well-defined. However, the smaller cross-section and more demanding experimental analysis reduce for now the possibilities coming from these channels.

In parallel to joint experimental analyses at higher order and lattice inputs in x-space, it is also desirable to constrain GPDs with additional physical principles. In particular, implementation of positivity constraints is a main aspect of the reduction of the size of shadow distributions which we will discuss in the following chapter.

7 - New models of generalised parton distributions

We have demonstrated in Chapters 5 and 6 that the extraction of GPDs from DVCS is ambiguous in practice due to the presence of shadow terms, that give essentially negligible contributions to CFFs over the range of virtuality Q^2 accessible to current and envisioned facilities, and are therefore unconstrained by experimental data. However, this does certainly not mean that nothing of value can be said about GPDs from the DVCS process. First, the small ξ limit benefits from the excellent knowledge of PDFs. By construction, shadow terms tend to zero as ξ gets smaller, which gives rise to modelling opportunities which we will explore in more detail at the end of this chapter. Furthermore, even if a considerable uncertainty exists due to the shadow terms outside the diagonal and small ξ limit, dealing with ill-defined, badly conditioned, or simply non-invertible problems is an issue arising in every field of science without becoming insurmountable impediments.

A remarkable example of a way around this kind of issues is given by missing data inpainting, that is the recreation of a full set of data from only a portion of it. Even in cases where the missing data is completely lost and unknown, statistical tools provide reconstruction procedures with a very high level of satisfaction (see for instance a work on data inpainting increasing considerably the robustness of the recovery of gravitational wave signals in Ref. [230]). The key element is to identify a good systematic way to perform the extrapolation, either through the selection of a specific basis deemed relevant to the studied case – for instance selecting a number of modes in a principal component analysis, or a basis of relevant wavelets for image reconstruction, or by constructing a satisfactory model architecture as in the case of neural network modelling.

In this chapter, we explore several models we have developed in order to give a reasonable account of the uncertainty associated with GPDs extracted from DVCS data. We present first a simple analytical model that reproduces exactly LO CFFs, which can be used in association with polynomial shadow DDs found in the previous chapter. We give afterwards a presentation of a grounding work on the neural network modelling of double distributions, which we have published in Ref. [8]. We end up the chapter by discussing specific aspects of the modelling of GPDs in the small ξ region. This region attracts a lot of interest for reasons we will detail, and evades the issue of shadow GPDs in a way that is instructive to study.

7.1 . A simple analytical model

Having determined exactly the solution space of polynomial DDs that either cancel their forward limit – Eqs. (6.15) and (6.18) – or their LO CFF – Eqs. (6.52) and (6.55) –, and provided an infinite family of polynomial DDs which cancel both – Eq. (6.63) –, as well as the building blocks to extract many more in case the need would be felt, we have at our disposal an extensive characterisation of the solution space of the homogeneous system "vanishing forward limit, vanishing LO CFF". Since it is a linear system, we know that the complete solution to the problem "known forward limit, known LO CFF" is obtained by adding a particular solution to the previously determined homogeneous solution. We present now a systematic way to find such a particular solution.

7.1.1 . Description of the model

Let us consider that at a determined value of t and μ_F^2, we know a forward limit $q^{(+)}(x)$ and a CFF $\mathcal{H}^q(\xi)$. We look for a model of GPD which reproduces exactly these two limits at LO. Thanks to our earlier observation of Eq. (6.60), the DD defined for $\beta > 0$ by

$$f_1^{q(+)}(\beta,\alpha) = \frac{1}{\pi e_q^2}\frac{\mathrm{d}}{\mathrm{d}\beta}\left(\frac{\beta}{2-\beta}\mathrm{Im}\,\mathcal{H}^q\left(\frac{\beta}{2-\beta}\right)\right), \tag{7.1}$$

and extended by parity to $\beta < 0$ produces exactly the expected imaginary part of the LO CFF. We note $H_1^{q(+)}$ the Radon transform of $f_1^{q(+)}$, and further define for $\beta > 0$

$$f_2^{q(+)}(\beta,\alpha) = \frac{q^{(+)}(\beta) - H_1^{q(+)}(\beta,0)}{2(1-\beta)\beta^2}, \tag{7.2}$$

with again a parity extension to $\beta < 0$. The form of $f_2^{q(+)}$ is chosen to enforce the expected forward limit while correcting the unwanted one produced by $f_1^{q(+)}$. We define $H_2^{q(+)}$ as the Radon transform associated to $f_2^{q(+)}$. Then

$$H_3^{q(+)}(x,\xi) = H_1^{q(+)}(x,\xi) + (x^2-\xi^2)H_2^{q(+)}(x,\xi) \tag{7.3}$$

has both the correct imaginary part of the LO CFF since $(x^2-\xi^2)$ neutralises the contributions of $H_2^{q(+)}$, and the correct forward limit :

$$H_3^{q(+)}(x,0) = H_1^{q(+)}(x,0) + x^2\int_{x-1}^{1-x}\mathrm{d}\alpha\, f_2^{q(+)}(x,\alpha), \tag{7.4}$$

$$= H_1^{q(+)}(x,0) + x^2\frac{q^{(+)}(x) - H_1^{q(+)}(x,0)}{2(1-x)x^2}2(1-x), \tag{7.5}$$

$$= q^{(+)}(x). \tag{7.6}$$

Let us show in addition that $H_3^{q(+)}$ verifies the important property of polynomiality of the Mellin moments. $H_1^{q(+)}$ and $H_2^{q(+)}$ satisfy polynomiality since they write as Radon transform of double distributions. To show that $H_3^{q(+)}$ satisfies it as well, we just need to observe that since $H_2^{q(+)}$ is a singlet GPD with no D-term, its Mellin moments for n even are 0, and for n odd,

$$\int_{-1}^{1}\mathrm{d}x\, x^n H_2^{q(+)}(x,\xi) = \sum_{k=0\ \text{even}}^{n-1}\xi^k A_{n,k}. \tag{7.7}$$

Multiplying $H_2^{q(+)}$ by $(x^2-\xi^2)$ simply adds a term ξ^{n+1} :

$$\int_{-1}^{1}\mathrm{d}x\, x^n(x^2-\xi^2)H_2^{q(+)}(x,\xi) = \int_{-1}^{1}\mathrm{d}x\, x^{n+2}H_2^{q(+)}(x,\xi) - \xi^2\int_{-1}^{1}\mathrm{d}x\, x^n H_2^{q(+)}(x,\xi), \tag{7.8}$$

$$= A_{n+2,0} + \sum_{\substack{k=2\\ \text{even}}}^{n+1}(A_{n+2,k} - A_{n,k-2})\xi^k. \tag{7.9}$$

Therefore, $H_3^{q(+)}$ satisfies polynomiality, but with a D-term generated by the interplay of $H_2^{q(+)}$ and the factor $(x^2-\xi^2)$. Technically, multiplying a DD $f^{q(+)}$ by a factor $(x^2-\xi^2)$

may be considered as a new DD scheme, in the same way that $(1-x)f_P^{q(+)}$ was presented as the Pobylitsa representation in Eq. (2.39). Our new scheme is however only adapted to singlet GPDs with vanishing diagonals [1].

Having fulfilled polynomiality, desired forward limit and imaginary part of the LO CFF, we still need to concern ourselves with its real part. The LO dispersion relation reads

$$\mathrm{Re}\,\mathcal{H}^q(\xi) = 2e_q^2 \int_{-1}^{1} \mathrm{d}z\, \frac{D^q(z)}{1-z} + \frac{1}{\pi}\mathrm{p.v.}\int_0^1 \mathrm{d}\xi'\, \mathrm{Im}\,\mathcal{H}^q(\xi') \left(\frac{1}{\xi-\xi'} - \frac{1}{\xi+\xi'}\right) . \tag{7.10}$$

Therefore, obtaining the correct real part of the LO CFF simply amounts to obtaining the correct subtraction constant, which we note C_H. We have introduced a spurious D-term – and therefore a spurious subtraction constant – by multiplying a GPD by $(x^2-\xi^2)$ in Eq. (7.3). This spurious subtraction constant, noted C_{H_3} can be identified by the use of the dispersion relation (7.10) on $H_3^{q(+)}$. We can correct it by adding a D-term D_3^q to our model such that

$$2e_q^2 \int_{-1}^{1} \mathrm{d}z\, \frac{D_3^q(z)}{1-z} = C_H - C_{H_3} . \tag{7.11}$$

One particularly easy solution is provided by

$$D_3^q(z) = \frac{3}{4e_q^2}(C_H - C_{H_3}) z(1-z^2) , \tag{7.12}$$

which yields the final model

$$H^{q(+)}(x,\xi,t) = H_3^{q(+)}(x,\xi,t) + \frac{3}{4e_q^2}\,(C_H - C_{H_3})\,\frac{x}{|\xi|} R\left(1-\frac{x^2}{\xi^2}\right) , \tag{7.13}$$

where $R(x) = (x+|x|)/2$ is the ramp function.

Summary of the model

To produce a singlet GPD model satisfying the polynomiality property and having the exact forward limit $q^{(+)}(x)$ and LO CFF $\mathcal{H}^q(\xi)$ at some fixed t and μ_F^2 values :

1. Compute the CFF subtraction constant C_H.
2. Compute the Radon transform $H_1^{q(+)}(x,\xi)$ of the DD defined for $\beta > 0$ – and expanded for $\beta < 0$ so that it is β-odd – by

$$f_1^{q(+)}(\beta,\alpha) = \frac{1}{\pi e_q^2}\frac{\mathrm{d}}{\mathrm{d}\beta}\left[\frac{\beta}{2-\beta}\mathrm{Im}\,\mathcal{H}^q\left(\frac{\beta}{2-\beta}\right)\right] .$$

3. Compute the Radon transform $H_2^{q(+)}(x,\xi)$ of the DD defined for $\beta > 0$ –

1. It is technically possible to represent singlet GPDs with a non zero limit for $x \to \xi$ in this scheme (the diagonal $x = \xi$ itself would then be undefined except using a continuous extension). We will even face a GPD with divergent behaviour when $x \to \xi$ when applying our model to the proton case in the following. It necessitates however a singular DD to overcome the factor $(x^2-\xi^2)$ and it is therefore ill advised to use this representation as a general DD scheme.

and expanded for $\beta < 0$ so that it is β-odd – by

$$f_2^{q(+)}(\beta,\alpha) = \frac{q^{(+)}(\beta) - H_1^{q(+)}(\beta,0)}{2(1-\beta)\beta^2}\,.$$

4. Compute the CFF subtraction constant C_{H_3} of

$$H_3^{q(+)}(x,\xi) = H_1^{q(+)}(x,\xi) + (x^2-\xi^2)H_2^{q(+)}(x,\xi)\,.$$

5. The GPD model is given by

$$H_3^{q(+)}(x,\xi) + \frac{3}{4e_q^2}\left(C_H - C_{H_3}\right)\frac{x}{|\xi|}R\left(1-\frac{x^2}{\xi^2}\right)\,,$$

where $R(x) = (x+|x|)/2$ is the ramp function.

7.1.2 . Practical use

We will show the results of the model on two practical cases, one for pions and one for protons.

Pion model

For the pion model, we use successively the two PDFs discussed in Ref. [179], namely

$$u_1(x) = \Theta(x)\left(30x^2(1-x)^2\right)\,, \tag{7.14}$$

$$u_2(x) = \Theta(x)\left(213.32\,x^2(1-x)^2\left[1+2.29\,x(1-x)-2.93\sqrt{x(1-x)}\right]\right)\,, \tag{7.15}$$

where $\Theta(x)$ is the Heaviside step function. Because the PDFs are vanishing for $x<0$, the associated singlet PDFs read

$$u_1^{(+)}(x) = \mathrm{sgn}(x)\left(30x^2(1-|x|)^2\right)\,, \tag{7.16}$$

$$u_2^{(+)}(x) = \mathrm{sgn}(x)\left(213.32\,x^2(1-|x|)^2\left[1+2.29\,|x|(1-|x|)-2.93\sqrt{|x|(1-|x|)}\right]\right)\,. \tag{7.17}$$

Positivity constraint for the pion on the singlet GPD takes the simple form for $x > |\xi|$

$$|H^{u(+)}(x,\xi)| \leq \sqrt{u^{(+)}\left(\frac{x+\xi}{1+\xi}\right)u^{(+)}\left(\frac{x-\xi}{1-\xi}\right)}\,. \tag{7.18}$$

As a consequence,

$$\lim_{\substack{x\to\xi\\x>\xi}} H^{u(+)}(x,\xi) \leq \sqrt{u^{(+)}\left(\frac{2\xi}{1+\xi}\right)u^{(+)}(0)}\,. \tag{7.19}$$

Considering $u^{(+)}(0) = 0$ for both considered pion PDFs, we conclude therefore that the imaginary part of the LO CFF for our model has to be 0. We have no way *a priori* to fix a sensible value for the real part of the LO CFF or equivalently the expected subtraction

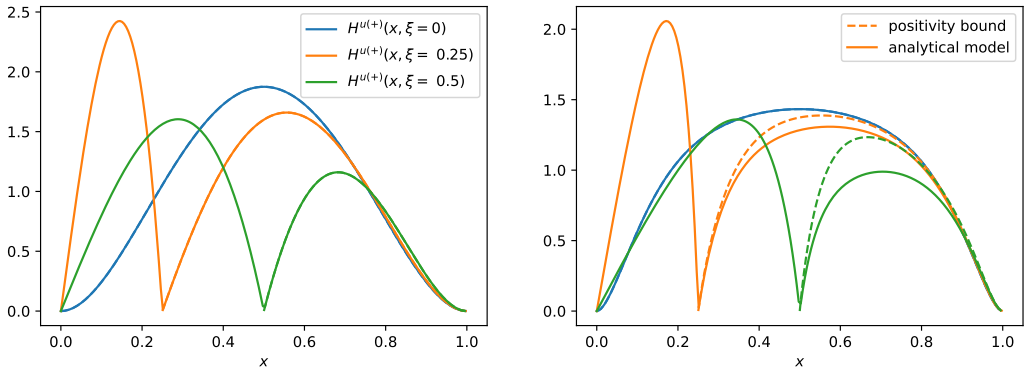

Figure 7.1 – Singlet analytical models that reproduce exactly a given PDF and imaginary part of a LO CFF, and that vanish at $\xi = 1$ to satisfy the requirements of the soft pion theorem. We use two pion PDFs $u_1^{(+)}$ (Eq. (7.16) - left) and $u_2^{u(+)}$ (Eq. (7.17) - right). Due to positivity constraints, the imaginary part of the LO CFF is constrained to be 0. The models are shown in solid lines for $\xi = 0$ (blue), $\xi = 0.25$ (orange) and $\xi = 0.5$ (green). On the right panel, the dotted lines represent the positivity bound of Eq. (7.18). It is exactly equal to our analytical model on the left panel.

constant in the absence of actual experimental data. We can however replace steps 1 and 5 of our modelling procedure – which are intended to adjust the model to the expected real part of the LO CFF by defining a custom D-term contribution – by another procedure intended to satisfy the soft pion theorem [39, 224] already mentioned in Section 6.2.2. We remind that the latter requires that the singlet quark GPD vanishes on the $\xi = 1$ line. Therefore, step 1 is suppressed, and step 5 replaced by

$$H_3^{u(+)}(x,\xi) - \mathrm{sgn}(\xi)\Theta\left(1-\frac{x^2}{\xi^2}\right) H_3^{u(+)}\left(\frac{x}{\xi},1\right) . \tag{7.20}$$

Starting directly at step 2, since $\mathrm{Im}\,\mathcal{H}^u(\xi) = 0$, we find that $f_1^{q(+)}(\beta,\alpha) = 0$, and following the rest of the procedure with modified step 5 brings the results of Fig. 7.1. We reproduce exactly the correct PDFs and the vanishing imaginary part of the LO CFF in both cases. We compare the model with the positivity bound defined for $x > |\xi|$ by the r.h.s. of Eq. (7.18). Surprisingly, for the PDF choice $u_1^{(+)}$ (left panel), our analytical model is exactly equal to the positivity bound. For $u_2^{(+)}$ (right panel), our analytical model remains below the positivity bound. In both cases, the behaviour looks excellent. Therefore, this model could represent an alternative to the procedure of Ref. [231, 179] or the polynomial fit we presented in Section 6.2.2 consisting in finding a DD which approximately saturates the positivity bound of Eq. (7.18) to extend the GPD to the ERBL region. **We have here a model that directly satisfies polynomiality, is defined on both the DGLAP and ERBL regions, reproduces exactly a given PDF, imaginary part of the LO CFF and allows further constraints on the D-term – either to reproduce the real part of the LO CFF or satisfy the soft-pion theorem for instance – and seems to co-exist well with the constraint of positivity.** Note that even if positivity was not obtained by the analytical model for some requirements of PDF and LO CFF, we could add to the model a LO shadow GPD to try and obtain the correct positivity property. LO shadow GPDs would not spoil the

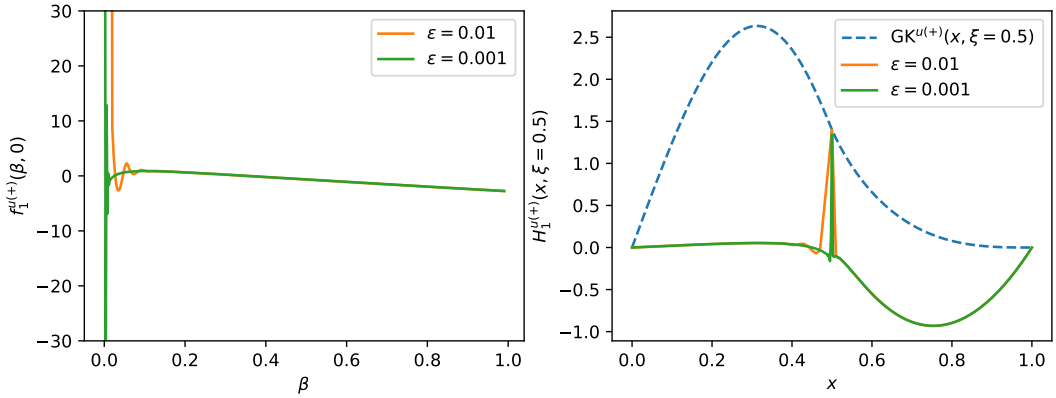

Figure 7.2 – (left) DD $f_1^{u(+)}(\beta, \alpha)$ intended to reproduce the imaginary part of the u quark contribution to the LO CFF in GK model for $\xi \in [\epsilon, 1]$, and $\epsilon = 0.01$ or 0.001. As ϵ decreases and we reproduce correctly the imaginary part of the LO CFF on a wider interval, the DD tends to a continuous function plus a singularity at $\beta = 0$. – (right) Radon transform $H_1^{u(+)}(x, \xi = 0.5)$ of the previous DDs (orange and green), with the value of the singlet u quark GK model in dotted blue. One can observe that we reproduce indeed with this exotic model the correct diagonal of the GPD, but the rest of it looks nothing like a traditional GPD.

exact agreement of the forward limit and LO CFF by definition, and produce a wide range of behaviours away from the regions $x = \xi$ and $x = 0$ that can help correct a lack of positivity. First attempts in this direction have been very promising, justifying the use of shadow GPDs in the neural network modelling of Section 7.2.

Proton model

We can apply our model to try to reproduce the phenomenological GK model. We extract the singlet PDF and LO CFF of the u quark GK model, and use them as seeds to fabricate our own. The subtraction constant in the GK model is 0 owing to the absence of a modelling of the D-term. However, already at the second step of our procedure, a notable difficulty appears. Due to the divergence of the imaginary part of the LO CFF at small ξ, the derivative required to compute $f_1^{u(+)}(\beta, \alpha)$ is difficult to estimate at small β. To understand what is happening, it is useful to only try to reproduce the imaginary part of the LO CFF starting from $\xi > \epsilon$. We demonstrate on Fig. 7.2 the effect of choosing $\epsilon = 0.01$ (orange curves) and $\epsilon = 0.001$ (green curves). The DD $f_1^{u(+)}(\beta, \alpha)$ on the left panel seems to be overall very well-behaved, except in the limit where β is close to 0 where it oscillates very strongly. As ϵ decreases, the range in β of oscillations decreases quickly, but the their amplitude increases on the other hand. In the limit where $\epsilon \to 0$ and we reproduce the imaginary part of the LO CFF on the entire interval $]0, 1]$, the DD becomes a continuous background with an superimposed distribution at $\beta = 0$. This is visible on the Radon transform $H_1^{q(+)}(x, \xi)$ shown on the right panel, where as $\epsilon \to 0$, the GPD becomes a continuous background, plus a finite height but infinitely narrow peak at $x = \xi$. The second step of our model produces therefore an "exotic" solution to the problem of getting the correct imaginary part of a diverging LO CFF while remaining polynomial. Let us stress that the distribution at the $\beta = 0$ line does not create a D-term, but still belong to the DD f in the Polyakov-Weiss representation.

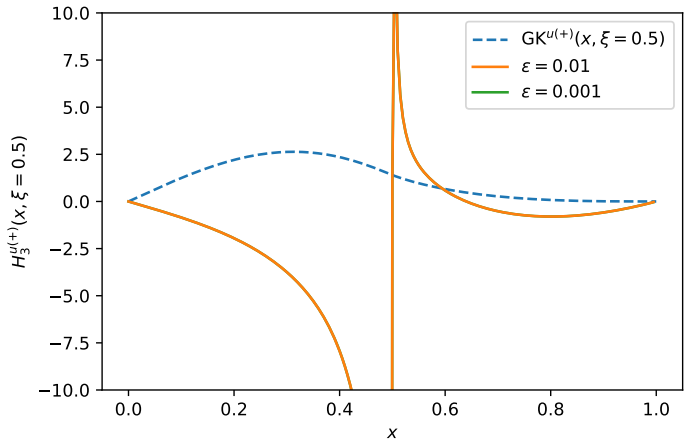

Figure 7.3 – $H_3^{u(+)}(x, \xi = 0.5)$ of our model which reproduces the correct forward limit and diagonal of the GK model, but with a divergence when $x \to \xi$ with $x \neq \xi$. The variation of ϵ makes no observable difference in the model.

One may wonder if following the further steps of our modelling procedure allows us to somehow regularise this exotic GPD. It is however not the case, the third step with $f_2^{u(+)}$ bringing even more issues. Indeed, $f_2^{u(+)}$ is singular at $\beta = 0$ because of the factor $1/\beta^2$ in its definition. When $\xi \neq 0$, the factor $(x^2 - \xi^2) = (x + \xi)(x - \xi)$ by which we multiply the Radon transform of f_2 can tame only a divergence of order $1/(x - \xi)$ on the diagonal, but not $1/(x - \xi)^2$. For the pion model, we had the chance that $u^{(+)}(\beta) - H_1^{u(+)}(\beta, 0)$ was vanishing in the limit $\beta \to 0$, providing the necessary regularisation. However, for the GK model, the ill-defined behaviour of $H_1^{u(+)}(\beta, 0)$ at small β and the divergence of $u^{(+)}(\beta)$ do not offer this relief. We end up therefore with an even more peculiar GPD : it is polynomial, has the correct diagonal and forward limit, but diverges when x gets close to ξ from either superior or inferior values, while being finite and equal to the expected diagonal at $x = \xi$. This effect can be observed on Fig. 7.3.

As a conclusion, we have proposed a model which gives excellent results for GPDs whose PDF remains finite when $x \to 0$, reproducing correctly the forward limit and LO CFF while respecting polynomiality and showing a good compatibility with positivity inequalities. However, in the case of diverging PDFs at small x, the results are quite exotic, although they still produce the correct forward limit and LO CFF while satisfying polynomiality. It is possible to envision an hybrid modelling where the divergent part of the GPD is modelled first, for instance thanks to an RDDA-like approach, and the resulting GPD now cleared from divergences is modelled thanks to our procedure with the addition of shadow LO GPDs. We want however to explore the possibility of dealing directly with the phenomenological small x divergence.

7.2 . Neural network modelling of the double distribution

Our objective of being able to exactly reproduce a forward limit and LO CFF while respecting positivity and polynomiality has only been partially fulfilled in the previous section owing to the bad behaviour of our model when facing the phenomenological Regge behaviour of PDFs at small x. We consider nonetheless that our objective is important

since it encompasses directly some of the most stringent experimental constraints on GPDs. It is known that models based on the RDDA (see Section 3.4.1) for instance offer only a limited flexibility : although they reproduce exactly a given forward limit by design, their ability to accommodate the LO CFF is entirely contained in just one free parameter and is therefore quite limited (see for instance Ref. [166]).

Having used neural network models of CFFs precisely for their flexibility and the lesser model dependence they introduce compared to physically motivated parameterisations, it is only natural to try to satisfy our objective of reproducing the forward limit and LO CFF with a neural network model of the double distribution. Neural network modelling has now become common practice in the field of PDFs, see for instance Refs. [188, 190, 191]. However, it is not completely straightforward to apply the PDF strategy to the GPD case, owing to the higher dimensionality of GPDs and extra non-trivial theoretical constraints. We will notably illustrate how to deal with the simplified positivity inequality for $x > |\xi|$ inspired from Eq. (2.44) [2] :

$$|H^q(x,\xi,t)| \leq \sqrt{q\left(\frac{x+\xi}{1+\xi}\right) q\left(\frac{x-\xi}{1-\xi}\right) \frac{1}{1-\xi^2}}\,. \tag{7.21}$$

This section is largely based on the published work of Ref. [8].

7.2.1 . Description of the model

Our neural network model for singlet DDs consists of three parts

$$f^{q(+)}(\beta,\alpha) = (1-x^2) f_C^{q(+)}(\beta,\alpha) + (x^2-\xi^2) f_S^{q(+)}(\beta,\alpha) + \xi f_D^{q(+)}(\beta,\alpha)\,. \tag{7.22}$$

The first term, $(1-x^2) f_C^{q(+)}(\beta,\alpha)$, is inspired from the RDDA, except it will use a much more flexible profile function to be able to adapt to different LO CFFs. The second term, $(x^2-\xi^2) f_S^{q(+)}(\beta,\alpha)$, is directly dictated by our study of shadow GPDs. Its contribution vanishes by design for $x = \xi$, and we will make sure it vanishes as well for $\xi = 0$. The inclusion of this term is important for a proper estimation of model uncertainties when GPDs are constrained by a sparse set of data. We have demonstrated previously that adding factors x^2 or ξ^2 to singlet DDs did not violate the polynomiality property, but simply resulted in the production of a D-term contribution. The third term, $\xi f_D^{q(+)}(\beta,\alpha)$, only contributes to the D-term. The inclusion of this term gives an extra flexibility to the model required to reproduce all x-moments of GPDs and correct the possibly unwanted D-term contributions of the two previous terms. Let us detail more precisely the exact modelling of each term.

First term

As in the RDDA, we factorise the well-known forward limit $q^{(+)}(x)$ with a profile function $h_C(\beta,\alpha)$, except the profile function is produced by a neural network. We have added a factor $(1-x^2)$ in front of $f_C^{q(+)}$ in Eq. (7.22) which helps to satisfy positivity in the large x region where the inequalities are particularly stringent. It translates into the

2. This simplified positivity inequality overlooks the role of the GPD E which we have not specifically modelled so far.

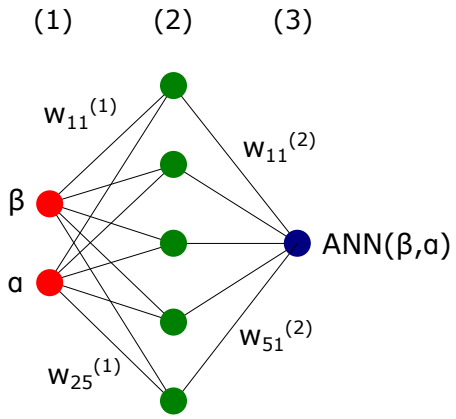

Figure 7.4 – Architecture of the neural network used to represent the profile function $h_C(\beta, \alpha)$. Figure taken from Ref. [8].

factoring of a term $1/(1-\beta^2)$ in the expression of $f_C^{q(+)}$, that is

$$f_C^{q(+)} = q^{(+)}(\beta) h_C(\beta, \alpha) \frac{1}{1-\beta^2} \,. \tag{7.23}$$

To ensure the correct behaviour of the GPD, the profile function $h_C(\beta, \alpha)$ must be even in the α and β variables, and normalised as follows

$$\int_{\beta-1}^{1-\beta} \mathrm{d}\alpha \, h_C(\beta, \alpha) = 1 \,, \tag{7.24}$$

so that the proper reduction to $q^{(+)}$ is obtained for the Radon transform of this term when $\xi = 0$. In addition, to increase the regularity of the term and facilitate the implementation of positivity constraints, we require that the profile function vanishes on the edges of the rhombus, that is

$$h_C(\beta, \alpha) = 0 \text{ for } |\alpha| + |\beta| = 1 \,. \tag{7.25}$$

To satisfy all these requirements, we propose to model the profile function by

$$h_C(\beta, \alpha) = \frac{\mathrm{ANN}_C(|\beta|, \alpha)}{\int_{|\beta|-1}^{1-|\beta|} \mathrm{ANN}_C(|\beta|, \alpha)} \,, \tag{7.26}$$

where $\mathrm{ANN}_C(\beta, \alpha)$ is the output of a neural network whose architecture is detailed in Fig. 7.4. It is a feed-forward network with a single hidden layer, taking $|\beta|$ and α as inputs. Denoting by φ the activation function of the hidden layer neurons, the output of the k-th hidden neuron is

$$o_k^{(2)} = \varphi\left(b_k + w_{1,k}^{(1)}|\beta| + w_{2,k}^{(1)} \frac{\alpha}{1-|\beta|}\right) - \varphi\left(b_k + w_{1,k}^{(1)}|\beta| + w_{2,k}^{(1)}\right) + \left[w_{2,k}^{(1)} \to -w_{2,k}^{(1)}\right] \,, \tag{7.27}$$

where b_k is the bias of the neuron, and $w_{1,k}^{(1)}$, $w_{2,k}^{(1)}$ are the weights connecting it to the input layer. They are free parameters to be fitted on data. The activation function is chosen as

the sigmoid $\varphi(x) = 1/(1+\exp(-x))$. One can see therefore that the neurons do not act simply as the composition of the activation function and the mere weighted sum of its inputs.

The reason for the form of Eq. (7.27) is to satisfy by design most of the requirements of our profile function. First, the term $\left[w_{2,k}^{(1)} \to -w_{2,k}^{(1)}\right]$ guarantees that the output is even in α, while using directly $|\beta|$ gives the correct parity in β. Then rescaling α by $1-|\beta|$ serves two purposes. First it is easy to see in Eq. (7.27) that the output is 0 if $\alpha = 1-|\beta|$, so the profile function vanishes on the edges of the rhombus. Then it provides a desirable transformation from variable α which spans $[|\beta|-1, 1-|\beta|]$ to $\alpha/(1-|\beta|)$ which spans $[-1,1]$. This standardisation of variables produces a better behaviour in terms of convergence of the neural network.

The output of the network itself is then obtained this time as a simple weighted sum of the outputs of the hidden layers $o_k^{(2)}$:

$$\mathrm{ANN}_C(|\beta|,\alpha) = \sum_k w_{k,1}^{(2)} o_k^{(2)} \,. \tag{7.28}$$

This amounts to using the linear activation function in the output layer, without any biases. The parameters $w_{k,1}^{(2)}$ are free. It is possible to compute explicitly the normalisation factor in Eq. (7.26) as

$$\begin{aligned}\int_{|\beta|-1}^{1-|\beta|} \mathrm{ANN}_C(|\beta|,\alpha) = 2\sum_k \frac{w_{k,1}^{(2)}}{w_{2,k}^{(1)}}(|\beta|-1)\Bigg[&\log\left(\cosh\left(b_k - w_{2,k}^{(1)} + |\beta| w_{1,k}^{(1)}\right)\right) \\ &- \log\left(\cosh\left(b_k + w_{2,k}^{(1)} + |\beta| w_{1,k}^{(1)}\right)\right) + w_{2,k}^{(1)} \tanh\left(b_k - w_{2,k}^{(1)} + |\beta| w_{1,k}^{(1)}\right) \\ &+ w_{2,k}^{(1)} \tanh\left(b_k + w_{2,k}^{(1)} + |\beta| w_{1,k}^{(1)}\right)\Bigg] \,. \end{aligned} \tag{7.29}$$

Second term

We have extracted shadow GPDs in Chapter 6 thanks to a polynomial parameterisation of DDs which yielded many remarkable properties. It is an excellent framework to demonstrate the existence of these objects and perform their explicit extraction at NLO in view of the quantification of the conditioning of the deconvolution problem. However, since no constraint has been imposed on the large β region of polynomial DDs, shadow GPDs extracted in this manner are not particularly small at large x, as can be witnessed for the brown curve on Fig. 6.4 for instance. Since positivity constraints are so stringent in the large x region, it means that contributions to the uncertainty stemming from this NLO shadow GPD will be extremely suppressed. The orange curve on Fig. 6.4, extracted at a slightly higher polynomial order, is better behaved in the large x region, so will contribute more to the uncertainty while taking into account positivity, but we find it preferable to use an alternative modelling of shadow GPDs.

We remind that shadow GPDs can be either defined explicitly through the criteria enumerated in Section 6.1, or more simply be defined as difference of actual GPDs with the same forward limit and (N)LO CFF. We will take advantage of that second definition

to produce the term $f_S^{q(+)}(\beta,\alpha)$ in a similar way to $f_C^{q(+)}(\beta,\alpha)$:

$$f_S^{q(+)}(\beta,\alpha) = q^{(+)}(\beta) h_S(\beta,\alpha)\,, \tag{7.30}$$

where $h_S(\beta,\alpha)$ is a new profile function whose normalisation is not 1 as for the previous term, but 0 to cancel the forward limit :

$$\int_{\beta-1}^{1-\beta} \mathrm{d}\alpha\, h_S(\beta,\alpha) = 0\,. \tag{7.31}$$

All other properties of parity and cancellation on the edges of the rhombus are preserved. Thanks to the $(x^2-\xi^2)$ factor in Eq. (7.22), the imaginary part of the LO CFF is equally cancelled. It is not the case of the real part, due to the D-term induced by this factor, but we will take care of this in the final term. One may wonder why we introduced a reference to $q^{(+)}$ in Eq. (7.30) although we are explicitly building a term whose forward limit vanishes. It is because we are interested in the large x behaviour induced by factoring the phenomenologically relevant PDF, which will prove very useful to enforce positivity.

A model of the profile function $h_S(\beta,\alpha)$ can be proposed as

$$h_S(\beta,\alpha) = N_S \left(\frac{\mathrm{ANN}_S^{(1)}(|\beta|,\alpha)}{\int_{|\beta|-1}^{1-|\beta|} \mathrm{ANN}_S^{(1)}(|\beta|,\alpha)} - \frac{\mathrm{ANN}_S^{(2)}(|\beta|,\alpha)}{\int_{|\beta|-1}^{1-|\beta|} \mathrm{ANN}_S^{(2)}(|\beta|,\alpha)} \right)\,. \tag{7.32}$$

In practice, we are subtracting two GPDs having the same forward limit, obtained thanks to two neural networks $\mathrm{ANN}_S^{(1,2)}$. For simplicity, we will fix $\mathrm{ANN}_S^{(2)} = \mathrm{ANN}_C$, and choose for $\mathrm{ANN}_S^{(1)}$ a similar architecture to ANN_C, but with free parameters unrelated to that of the latter. N_S is a normalisation which can be in principle arbitrary without producing any forward limit or contribution to the imaginary part of the LO CFF. In practice, it will of course be limited by positivity inequalities, and to give the most complete account of uncertainty related to shadow GPDs, we will systematically try to maximise this normalisation without violating positivity.

Third term

We model $f_D^{q(+)}(\beta,\alpha)$ as a traditional D-term contribution

$$f_D^{q(+)}(\beta,\alpha) = \delta(\beta) D(\alpha)\,, \tag{7.33}$$

where D is represented by its expansion in terms of Gegenbauer polynomials truncated at order N

$$D(\alpha) = (1-\alpha^2) \sum_{i=1 \text{ odd}}^{N} d_i C_i^{(3/2)}(\alpha)\,. \tag{7.34}$$

The coefficients d_i are free parameters.

7.2.2 . Practical use

Ability to reproduce the GK model

To assess the potential of the model presented above, we generate pseudo-data from the GK model and first verify our ability to reproduce them. Considering the GK model

is based on the RDDA and our model is also built on factoring the forward limit with a profile function, what we are actually mainly probing through this test is the ability of the neural network to reproduce Radyushkin's choice of profile function. Precisely, we sample the GK model on $I = 400$ points probing logarithmically the domain $x, \xi \in [10^{-4}, 0.95]$ for $t = 0$ and $Q^2 = 4$ GeV^2, and train the ANN to minimise the root mean squared relative error (RMSE)

$$\mathrm{RMSE} = \sum_{i=1}^{I} \sqrt{\frac{1}{I}\left(\frac{\mathrm{GK}^{u(+)}(x_i,\xi_i) - H^{u(+)}(x_i,\xi_i)}{\mathrm{GK}^{u(+)}(x_i,\xi_i)}\right)}, \tag{7.35}$$

where $H^{u(+)}$ denotes the Radon transform of our neural network model of DD. We use 5 neurons in the hidden layers of ANN_C and $\mathrm{ANN}_S^{(1)}$, and an expansion of the D-term in $f_D^{q(+)}$ of 5 elements, that is up to $i = 9$ in Eq. (7.34). The scaling parameter of the $f_S^{q(+)}(\beta,\alpha)$ term is maintained fixed. Training is performed thanks to a genetic algorithm, with a drop-out regularisation to prevent overfitting (see Section 4.1.2 where the method is presented). We remind that drop-out amounts to randomly de-activating neurons – here one neuron in each network is de-activated at any time – during training while keeping in memory their previously trained parameters. The output of other neurons is scaled to compensate for the loss of signal intensity. The fraction of de-activated neurons has been validated by confronting the model to independent samples and observing the absence of over-fitting effects. This method allows us to introduce an element of stochasticity in the fit which prevents excessive sensitivity to minute features of the fitted data, and amounts in practice to training in parallel several different architectures. We re-activate all neurons for testing after the training, so the output of the full neural network may be considered as a superposition of the output of all architectures trained in parallel. We additionally perform the minimisation 100 times starting from different random initialisations of the neural networks, producing therefore what we will call 100 replicas. This allows us to reduce issues with imperfect convergence to a local minimum, and give an account of the uncertainty generated by the stochastic contribution of the drop-out method.

Finally, we apply outlier suppression thanks to a modified 3σ exclusion procedure : once training of the 100 replicas is achieved, we evaluate each on 1000 random points in the (x,ξ) space and discard replicas which are more than 3σ away from the mean of the distribution formed by all replicas in more than 10% of the probed kinematics.

The result is shown on Fig. 7.5. The fit of a single replica takes about 1.9h on 40 computing cores and ends with an RMSE of 0.016 on average. As can be assessed on the plot, the agreement is overall very satisfactory. Increasing the number of neurons in the hidden layers, which yields an even better agreement with the fitted data, is therefore not deemed necessary. A comparison of Mellin moments of the GK model and the distribution of Mellin moments of our model is shown on Fig. 7.6. We remind that the Mellin moments of a singlet GPD are exactly 0 for n even, and satisfy for n odd

$$\int_{-1}^{1} \mathrm{d}x\, x^n H^{q(+)}(x,\xi) = \sum_{k=0 \text{ even}}^{n+1} \xi^k A_{n,k}\,. \tag{7.36}$$

The abscissa for each Mellin coefficient $A_{n,k}$ on the plot of Fig. 7.6 is designed to include 0 to allow a visual measurement of the relative uncertainty of the extraction – except for

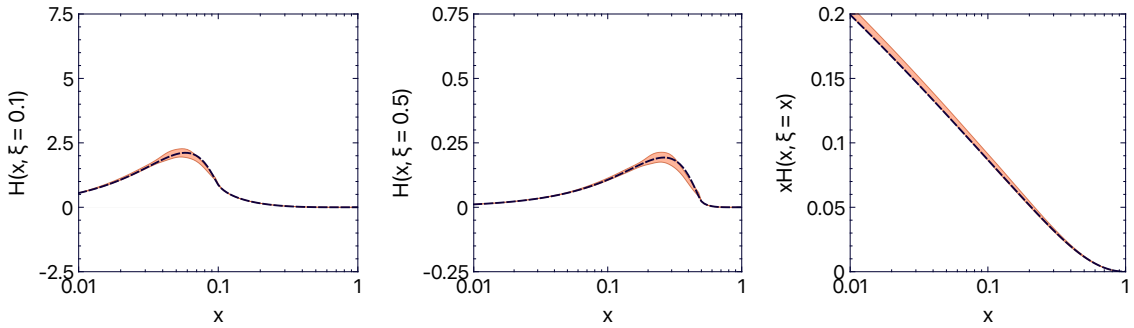

Figure 7.5 – Fit of the GK model (black dashed line) with our neural network DD model on 400 points covering a large range in (x, ξ). The orange band represents the 1σ interval computed from 100 minimisations starting from random initialisations of the network trained with drop-out method. We compared the GK model with our own for $\xi = 0.1$ (left), $\xi = 0.5$ (center) and $\xi = x$ (right). Figure taken from Ref. [8].

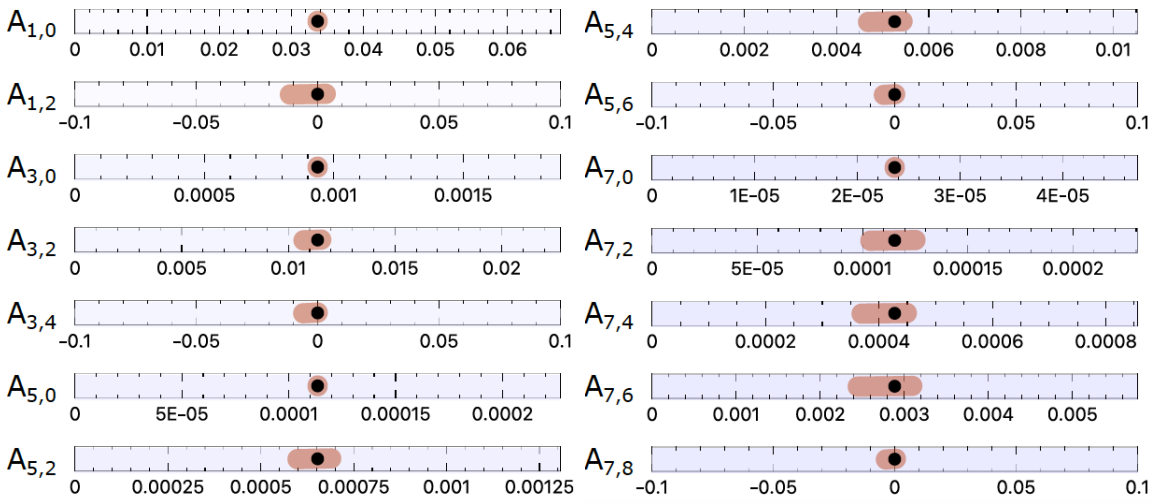

Figure 7.6 – Comparison of the Mellin coefficients defined in Eq. (7.36) for the GK model and our fitted model. Figure taken from Ref. [8].

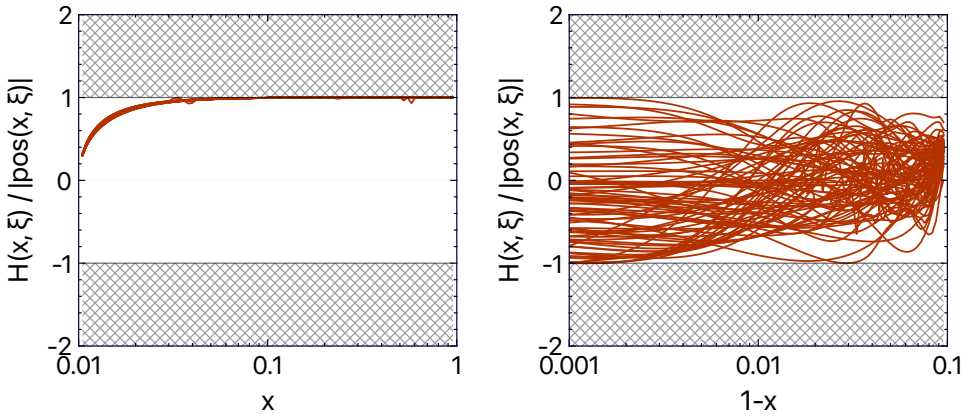

Figure 7.7 – 100 replicas (solid red lines) of the fit of our neural network DD model on the PDF and diagonal of the GK model PDF while enforcing the positivity condition of Eq. (7.21). We show the result for $\xi = 0.01$ (left) and $\xi = 0.9$ (right). The grey hatched bands are exclusion zones prohibited by the positivity condition. Figure taken from Ref. [8].

coefficients $A_{n,n+1}$ which are associated to the D-term and vanish in the GK model due to the absence of a D-term. We observe that moments $A_{n,0}$, which correspond to moments of the forward limit, are particularly well reconstructed. This is not a surprise since our model is factorised by the forward limit and reproduces it exactly. Overall this test of reproducibility of the GK model is very conclusive, and the fact that GK is right in the uncertainty bands shows that the neural network has the ability to reproduce GK without significant bias.

Uncertainty associated to only knowing the forward limit and diagonal of the GPD

We would like to lead now a more ambitious test. Assuming we only know the PDF and imaginary LO CFF – that is the diagonal of the GPD – and enforce positivity constraints, what kind of uncertainty on the extraction of the GPD does our model produce ? Since as we mentioned before, we always reproduce exactly any desired forward limit, we just have to fit our model on the diagonal of the GK model. We use the same loss function as above (7.35), except we now apply it only to $I = 200$ points logarithmically spaced with $x = \xi \in [10^{-4}, 0.95]$. To enforce positivity, we use the following procedure during training with the genetic algorithm. For each candidate considered during training, we check by how much we can scale $f_S^{q(+)}$ by changing the N_S parameter to saturate the simplified positivity inequality of Eq. (7.21). We check this on 10 000 points covering the $x > |\xi|$ domain. If we are not able to scale $f_S^{q(+)}$ to respect the constraint on all points, we simply discard the candidate in the training. If we can, we choose the highest allowed value of $|N_S|$. This method to enforce positivity requires a substantial computing power, but gives satisfactory results. Note that $f_C^{q(+)}$ and $f_S^{q(+)}$ may not fulfill positivity when taken separately, but fulfill it in sum after the minimisation. This indicates that inclusion of $f_S^{q(+)}$ gives extra-freedom to satisfy this important theoretical requirement.

We demonstrate the result of our procedure to guarantee the satisfaction of positivity condition (7.21) in Fig. 7.7. As one can witness, the replicas do not enter the exclusion

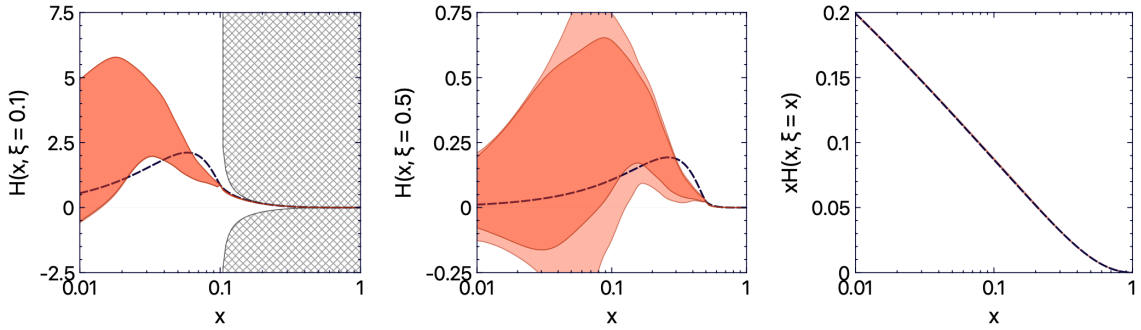

Figure 7.8 – Fit of the diagonal of the GK model (black dashed line) with our neural network DD model on 200 points covering a large range in $x = \xi$. The deep orange band represents the 1σ interval computed from 100 replicas of $f_C^{q(+)}$ alone, the light orange one of the sum $f_C^{q(+)} + f_S^{q(+)}$. We compared the GK model with our own for $\xi = 0.1$ (left), $\xi = 0.5$ (center) and $\xi = x$ (right). The hatched bands represent the excluded zone by the positivity condition of Eq. (7.21). Figure taken from Ref. [8].

band prohibited by the positivity condition. At small ξ, the variance of the replicas is very reduced, due to the strong constraining power of the imposed forward limit. On the contrary, at large ξ, the replicas fill the entire space authorised by positivity constraints.

The general result of the fit of the diagonal with positivity enforced is shown on Fig. 7.8. On the right panel, the excellent agreement of our model with the diagonal of the GK model is visible. The RMSE is on average 0.0075. This time, computation for a single replica takes however about 26h on 40 computing cores. For $\xi = 0.1$ and $\xi = 0.5$, we show the uncertainty bands associated to 100 replicas for the contribution of $f_C^{q(+)}$ alone (deep orange) and the sum $f_C^{q(+)} + f_S^{q(+)}$ (light orange). Considering fitting the model only on its diagonal brings no constraint on the D-term, we do not show the unconstrained contribution of $f_D^{q(+)}$. Two main observations can be made. First, the shadow term $f_S^{q(+)}$ brings no visible contribution to the extraction at $\xi = 0.1$, but a noticeable one at $\xi = 0.5$. Very practically, this is due to the fact that shadow contributions vanish for $\xi = 0$. A deeper understanding of the small ξ behaviour of GPDs will be the subject of the final section of this document. Second, there is a systematic deviation between our model and the target GK for x slightly smaller than ξ. The cause of this discrepancy is not that our model would somehow fail to reproduce the GK one, since we know well from our first test that we are able to do it. The actual reason is that the u singlet GPD H in the GK model does not satisfy the simplified positivity condition of Eq. (7.21). Indeed, a more complete version of the positivity condition would notably imply GPD E, which is out of the scope of this exploratory study.

Conclusion

The two tests we have shown demonstrate that our neural network modelling of DDs has the ability to reproduce the general features of a phenomenologically relevant model, while offering at the same time a far greater flexibility. It provides an interesting possibility to quantify the uncertainty associated to the joint knowledge of PDFs and CFFs, while allowing to implement positivity conditions in a framework guaranteeing the polynomiality property. One can envision straightforwardly the addition of further constraints, like the

knowledge of some Mellin moments which could be jointly fitted together with the CFF knowledge thanks to an adapted loss function. A complete working example of this model based on experimental data is still to be done. Among aspects which remain to be dealt with are a deeper reflection on inclusion of t dependence, the need to take into account other types of GPDs and managing the complex link between GPDs and experimental data.

We believe this model can be generalised to include these effects. As an increased control of systematic effects is highly desirable in front of the precision era of GPD physics, we hope that our model may conveniently be used in current and future analyses of GPDs, or in connection for instance to lattice QCD, either in x-space or through Mellin moments.

7.3 . Specificities of the small ξ region

As we have already noticed, the proximity of the well-known forward limit at small ξ, at least for the GPDs H and $\widetilde{H}$ at $t = 0$, restricts a lot the relevence of shadow GPDs in that region. The only way to introduce shadow GPDs would be by means of abrupt functions which jump suddenly from 0 at $\xi = 0$ to a significant value at still negligible values of ξ – objects that we found too exotic to be considered so far. The assumed regularity of GPDs in the small ξ limit has given rise to interesting modelling proposals, already since the early days of GPD studies more than two decades ago. The subject has triggered a renewed phenomenological interest thanks to the perspective of using exclusive measurements at very small values of ξ to constrain gluon PDFs at equally small longitudinal momentum fraction x. Gaining a precise knowledge of gluon PDFs in this region is crucial for the search of signs of gluon saturation [232], calculations of event rates in ultra-high energy neutrino astrophysics [233], and the characterization of the gluonic system that defines the initial state of heavy-ion collisions [232].

The uncertainty on gluon PDFs for $x \sim 10^{-4}$ or less is still quite large, mainly because experimental access is scarcer than for larger values of x. Experimental sensitivity to the very small ξ and x region can be achieved through different measurements in collider facilities which provide the largest collision energies $\sqrt{s}$. At the LHC, inclusive particle production in particular of charm and beauty quarks has been proposed to constrain gluon PDFs in the proton for instance in Refs. [234, 235, 236]. Although a noticeable reduction of uncertainty is produced by inclusion of these measurements, the impact is limited by missing higher order corrections indicated by large scale uncertainties, hadronization uncertainties and other effects not accounted for in perturbative QCD. In addition to these observables, exclusive hard photoproduction processes which can also be measured at colliders are interesting to consider since they should be less or, at least, differently affected by phenomena not accounted for in state-of-the-art pQCD calculations.

Of particular interest is exclusive heavy vector meson production (HVMP) of J/ψ or Υ mesons, where the hard scale is provided by the mass of the vector meson m_V in the final state, and typically probed values of ξ [3] are of the order of $e^{-y} m_V/\sqrt{s}$ where y is the pseudorapidity. HERA [237, 238, 239, 240, 241, 242, 243], LHC [244, 245, 246, 247, 248, 249] or a future LHeC allow measurements at small values of ξ of the order of 10^{-4} down to 10^{-6}. The future EIC will provide precise HVMP data with ξ of the order of 10^{-3} to

3. As it is customary for experimentalists to talk in terms of Bjorken's variable x_B, we remind that in the kinematic region of interest here, $x_B \approx 2\xi$ (see Eq. (3.5)).

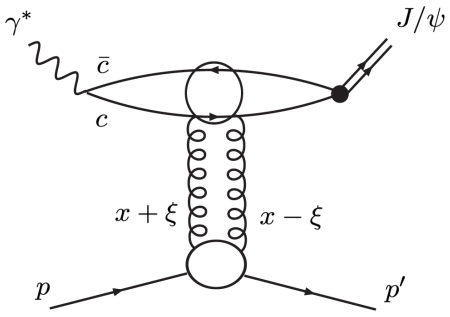

Figure 7.9 – Photoproduction of heavy vector meson J/ψ, adapted from [251].

10^{-4}.

The leading order two-gluon exchange in the t-channel depicted on Fig. 7.9 is the dominant contribution to the HVMP cross-section. Because there is a transfer of four-momentum between initial and final hadron states p and p', the description of this process in the framework of collinear factorization does not involve usual PDFs, but GPDs. As we will see shortly, simply equating GPDs in the small ξ region with their forward limit is a tempting possibility, which comes however with a number of issues. A more sophisticated proposal to relate GPDs to PDFs in the small ξ region, which we will refer to in the following as Shuvaev's proposal for concision since it is based on the Shuvaev transform, was proposed in Refs. [110, 250] and was applied to LHC and HERA data for instance in Ref. [251] or more recently to LHCb data for instance in Ref. [252]. The recent development of the software APFEL++ for the LO evolution of GPDs in momentum-space [1] allows us to reinterpret Shuvaev's proposal, and suggest an alternative procedure where we find the evaluation of the theoretical uncertainty associated to linking GPDs to PDFs easier to perform.

We will start by exposing some properties of the evolution operators of GPDs and PDFs in the very small x and ξ limits. Then we revisit Shuvaev's proposal to reconstruct the small ξ dependence of GPDs from PDFs and reinterpret its validity in the context of the properties of evolution operators discussed in the previous section. After that, we introduce an alternative proposal, and discuss the uncertainty associated with it.

7.3.1 . Properties of evolution operators in the small x and ξ limit

The problem of reconstructing GPDs from PDFs in the small ξ limit might appear quite trivial. We know well that PDFs correspond to the forward limit of GPDs when $t \to 0$, and therefore $\xi \to 0$. For simplicity, we will express all results with the GPD H, but similar results can be obtained for E, $\widetilde{H}$ and $\widetilde{E}$. We will also systematically consider $\xi \geq 0$, parity in ξ allowing to derive all result for $\xi < 0$. By proposing a factorised t-Ansatz as we have done for instance for the study of the GFF $C^a(t)$ in Chapter 4, we could think about suggesting

$$H^a(x, \xi, t, \mu^2) = x^{p_a} a(x, \mu^2) A(t)\,, \tag{7.37}$$

in the limit where ξ is small, where $p_a = 1$ if $a = g$ and 0 otherwise [4]. Surely the factorised t-dependence is a serious assumption *per se* [5] but there is a bigger theoretical issue. It is possible to write Eq. (7.37) at some scale μ_0^2, but as we will see, the form of the equation is not preserved by evolution. Then this would lead us to wonder what scale μ_0^2 would be suitable, and what level of uncertainty we can expect from this approximation.

Let us notice that, provided we accept the assumption of factorised t-dependence, the suggestion (7.37), which is exactly true if $\xi = 0$, is still very satisfactory if ξ is not strictly zero, but $x \gg \xi$. Indeed, in that limit, GPD evolution tends towards the DGLAP evolution of PDFs, and Eq. (7.37) becomes valid for all scales. The situation is more difficult when ξ is not negligible compared to x, due to the relation x, ξ and μ^2 enjoy via evolution equations. This case deserves a particular attention since the convolutions of GPDs that enter the description of exclusive HVMP exhibit a strong sensitivity to the region $x \approx \xi$ – as in the case of DVCS. Likewise, HVMP is only sensitive to singlet GPDs.

Thanks to the discussion of Section 6.3.1 explaining the link between the LL resummation and the LO splitting kernels $K^{ab,(0)}$ introduced in Eq. (2.52), we define the LO evolution of GPDs with full LL resummation as

$$\frac{H^a(x,\xi,t,\mu^2)}{x^{p_a}} = \sum_b \int_0^1 \frac{\mathrm{d}z}{x}\, \Gamma^{ab,(0)}\left(\frac{z}{\xi}, \frac{\xi}{x}; \mu^2, \mu_0^2\right) \frac{H^b(z,\xi,t,\mu_0^2)}{z^{p_b}}\,. \tag{7.38}$$

In light of the definition of Eq. (7.38), the usual LO DGLAP evolution equation of PDFs can be written as

$$a(x,\mu^2) = \sum_b \int_0^1 \frac{\mathrm{d}z}{x}\, \Gamma^{ab,(0)}_{DGLAP}\left(\frac{z}{x}; \mu^2, \mu_0^2\right) b(z,\mu_0^2)\,, \tag{7.39}$$

where the DGLAP evolution operator $\Gamma^{ab,(0)}_{DGLAP}$ is linked to the general one $\Gamma^{ab,(0)}$ through

$$\Gamma^{ab,(0)}_{DGLAP}(\alpha; \mu^2, \mu_0^2) = \lim_{r\to 0} \Gamma^{ab,(0)}\left(\frac{\alpha}{r}, r; \mu^2, \mu_0^2\right)\,. \tag{7.40}$$

One can consider the $\Gamma^{ab,(0)}$ operators as **weights** of the GPD at initial scale μ_0^2 to produce GPDs at final scale μ^2. We will call z the longitudinal momentum fraction at initial scale μ_0^2, and x the same quantity at final scale μ^2. $\Gamma^{ab,(0)}$ operators provide therefore a measure of the importance of various z regions of the GPD at initial scale to the GPD at x at final scale.

The properties of evolution at small values of ξ have for instance been studied in Refs. [253, 254, 250, 256], but to the best of our knowledge, studies of this kind have been performed thanks to models of GPDs and DDs, often with the assumption of power-law behaviour at small x, and have therefore a lesser generality than our discussion which takes place directly at the level of the $\Gamma^{ab,(0)}$ operators. The evolution code APFEL++ allows indeed the explicit extraction of these objects.

4. $H^q(x,\xi,t,\mu^2) = q(x+\xi,\mu^2)A(t)$ has been likewise proposed [253, 254] but suffers in the end from the same issues as Eq. (7.37) in addition to requiring more care to implement the correct parity of the result in ξ.

5. Let us note for instance that the large x behaviour described in Eq. (2.9) is independent of t, and therefore incompatible with a factorised t dependence, and that lattice results seem to confirm it [255].

For numerical applications, we will use the typical hard scales encountered in HVMP, given by half the mass of the vector meson as argued in Ref. [257]. Indeed, Ref. [258] demonstrates the large sensitivity of the NLO HVMP cross-section to the precise choice of factorisation scale, largely due to the $\log(1/\xi)$ enhancement of the NLO corrections. Ref. [257] argues that fixing the factorisation scale at half the mass of the vector meson allows a resummation of the $\log(1/\xi)$ contributions and a stabilisation of the perturbative expansion. We will therefore use the final scales $\mu_c = m_{J/\psi}/2 = 1.5$ GeV and $\mu_b = m_\Upsilon/2 = 4.7$ GeV. Since we only perform LO evolution, we need an initial scale to start evolution which is high enough to produce consistent results. We choose as initial scale $\mu_0 = 1$ GeV and plot on Fig. 7.10 evolution operators $\Gamma^{ab,(0)}(y, \xi/x; \mu^2, \mu_0^2)$ evaluated for $\mu^2 = \mu_c^2$ and μ_b^2. We have named the ratio $y = z/\xi$ because this variable will play a predominant role in the discussion. Mind that only values of y in the range $[0, 1/\xi]$ should be considered as a consequence of the fact that z varies between 0 and 1. The plot allows therefore to observe evolution operators for ξ as small as 10^{-6}. Let us comment on some important features.

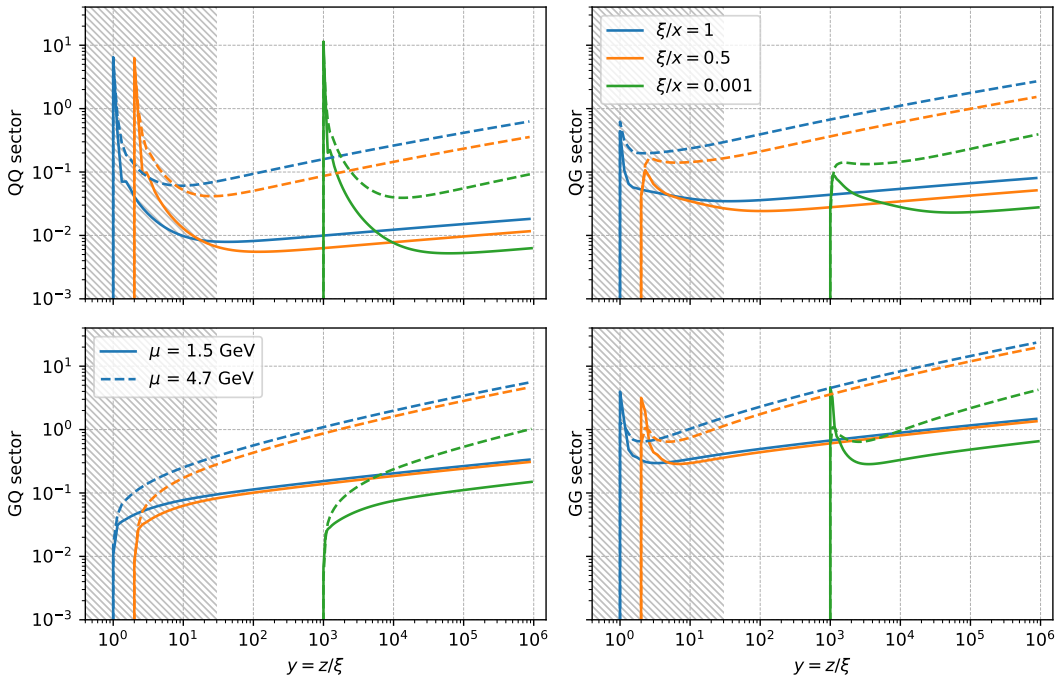

Figure 7.10 – Value of $\Gamma^{ab,(0)}(y, \xi/x; \mu^2, \mu_0^2 = 1\ \text{GeV}^2)$ where $y = z/\xi$ in the four sectors $ab = qq$ (upper left), $ab = qg$ (upper right), $ab = gq$ (lower left) and $ab = gg$ (lower left). The continuous lines stand for $\mu = m_{J/\psi}/2 = 1.5$ GeV and the dotted lines for $\mu = m_\Upsilon/2 = 4.7$ GeV. The colours refer to various values of the ratio ξ/x : $\xi = x$ (blue), $\xi = 0.5x$ (orange) and $\xi = 0.001x$ (green). The light grey region is $y < 30$ which serves as our criterion to discriminate between the small and large y regions.

— **Feature 1** : We have only shown curves for $x \geq \xi$ since it is the region where we have the best hope of directly relating the GPD to the PDF. In the $0 < x < \xi$ region, contributions of the D-term, which evolve with their own independent ERBL equation, are completely independent from the PDF. If $x \geq \xi$, $H^a(x, \xi, t, \mu)$ only depends of values of $H^b(z, \xi, t, \mu_0)$ such that $z \geq x$, or $y \geq x/\xi$, and the weights

$\Gamma^{ab,(0)}(y, \xi/x)$ are null if $y < x/\xi$ as can be assessed on Fig. 7.10.

— **Feature 2** : The qq and gg sectors exhibit a strong peak at $z = x$, or $y = x/\xi$, all the more that μ^2 is close to μ_0^2. It is easily understandable since in the limit of no evolution where $\mu^2 = \mu_0^2$, $\Gamma^{ab,(0)}(y, \xi/x; \mu^2, \mu_0^2) = \delta_{ab}\delta(1 - y\,\xi/x)$ where δ_{ab} is the Kronecker delta [6]. On the contrary, as μ^2 increases, the contribution from the region $y > x/\xi$ becomes increasingly important, especially for gluons.

The large y region, where $z \gg \xi$, corresponds to a kinematic region where at initial scale μ_0^2, the asymmetry between the four-momenta of incoming and outgoing hadron states $z + \xi$ and $z - \xi$ is negligible, and the GPD can safely be replaced with its forward limit. We choose $y > 30$ as criterion to discriminate between small and large y, which is enough to ensure both $z \gg \xi$, and to clear away from the peak region of $\Gamma^{ab,(0)}(y, \xi/x = 1; \mu^2, \mu_0^2)$ which will be of specific interest to this study. Conversely, it would be in general unwise to replace the GPD with the PDF in the small y region (in grey on the figure), where they may differ significantly.

— **Feature 3** : As it is our general goal to extract the largest possible information on GPDs at small ξ from the PDF, we are led to wonder what are the criteria so that the large y region plays the most important possible role in evolution. We notice that weights $\Gamma^{ab,(0)}$ increase in the large y region if :

1. the evolution range characterised by μ^2 increases as previously mentioned. This is demonstrated graphically on Fig. 7.10 by the difference between the solid ($\mu = 1.5$ GeV) and dotted ($\mu = 4.7$ GeV) lines.
2. the ratio x/ξ gets larger, since $\Gamma^{ab,(0)}$ only takes non-zero values for $y \geq x/\xi$ if $x \geq \xi$. It corresponds roughly to pushing the curves of Fig. 7.10 to the right, outside of the grey region.
3. at a fixed value of the ratio x/ξ, the value of ξ decreases. Indeed, for $x \geq \xi$, relevant values of y range from x/ξ to $1/\xi$, so it corresponds to increasing the range of y to the right on Fig. 7.10. Note however that the actual value of the evolved GPD depends not only on the weights $\Gamma^{ab,(0)}$, but also on the value of the GPD at initial scale. Whether the contributions of the large y region actually increase when both ξ and x decrease will depend on whether the increase of $\Gamma^{ab,(0)}$ at large y exceeds the increase of the initial GPD at small x. This depends on the chosen shape of the GPD at initial scale, and we will provide a quantitative answer to the dominance of the large y region in Section 7.3.3.

— **Feature 4** : A final remarkable feature which is only partly visible on Fig. 7.10 is the similarity of shape. We present in Fig. 7.11 the same plot where we multiply y by ξ/x so the peaks are systematically aligned at $z/x = 1$. The curves in the gg and gq sectors seem almost independent of ξ/x as soon as $z/x > 5$. They actually differ by an almost constant factor of 5% between $\xi = x$ and $\xi = 0.5x$ in the large z/x

6. We observe that there is also a noticeable peak in the qg sector which increases fastly when x gets closer to ξ. This tends to imply that gluons close to the diagonal would produce more quarks than those away from the diagonal although we have not investigated precisely the reason for this phenomenon.

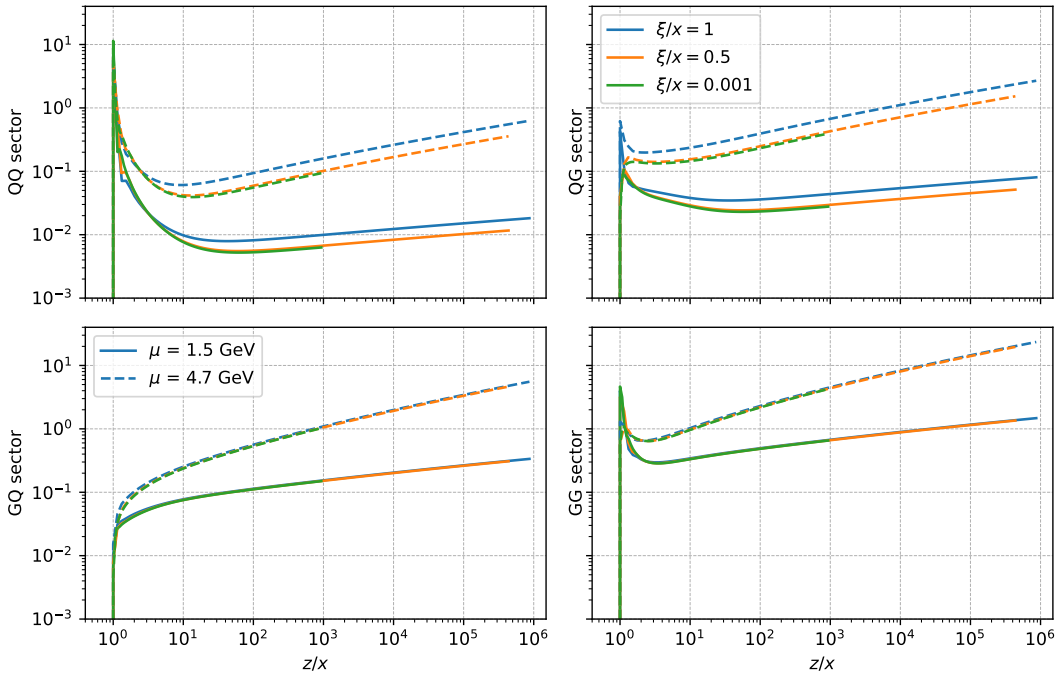

Figure 7.11 – Same as Fig. 7.10 except we now represent $\Gamma^{ab,(0)}$ as a function of z/x instead of $y = z/\xi$.

domain at $\mu = 4.7$ GeV. The difference is much starker in the qq and qg sectors, with an almost constant factor of 50% of difference for the same parameters. Significant differences are furthermore observed in the region where z is of the order of x.

Practically, the last feature we indentified means that provided the region $z \gg x$ at initial scale μ_0^2 is dominant in the evolution, which is the case if the GPD does not increase too quickly at small x, the change brought by evolution to the diagonal $x = \xi$ of the quark GPD is 50% larger than the change brought to the PDF. This fact was previously highlighted for instance in Refs. [253, 254, 250, 259], where it is shown analytically that for a power-law behaviour of the forward limit at small x and scale μ_0^2, then for $\mu^2 > \mu_0^2$ and $x \geq \xi$,

$$\lim_{(x,\xi)\to(0,0)} \frac{H^a(x,\xi,\mu^2)}{H^a(x,0,\mu^2)} = h\left(\frac{\xi}{x}, \mu^2, \mu_0^2\right) . \tag{7.41}$$

$h(\xi/x, \mu^2)$ tends to 1 when $\xi/x \to 0$ since the GPD tends towards its forward limit in that case, and the evolution equation of the GPD tends towards the traditional PDF DGLAP evolution equation. However, $h(\xi/x, \mu^2)$ is larger than 1 at finite $\xi/x > 0$, all the more that μ^2 is larger than μ_0^2 and ξ/x close to 1. Therefore, the diagonal of the GPD receives larger contributions from evolution compared to the forward limit. Ref. [260] finds for instance a ratio of the diagonal to the forward limit of gluon GPDs of 1.6 for $\mu_0^2 = 0.6$ GeV^2 to $\mu^2 = 100$ GeV^2 for a Regge-inspired Ansatz. By observing that all $\Gamma^{ab,(0)}$ operators are proportional to one another in the large z/x regime, we obtain similarly that

$$\lim_{(x,\xi)\to(0,0)} \frac{\Gamma^{ab,(0)}(z/\xi, \xi/x; \mu^2, \mu_0^2)}{\Gamma^{ab,(0)}(z/x, 0; \mu^2, \mu_0^2)} = \widetilde{h}\left(\frac{\xi}{x}, \mu^2, \mu_0^2\right) . \tag{7.42}$$

This observation does not rely on an analytical demonstration with assumption of power-law behaviour of the forward limit at some scale – which Ref. [256] shows to be only approximatively preserved by evolution – but on an observation of the properties of the weights $\Gamma^{ab,(0)}$ with the assumption that the large z region is dominating evolution.

This demonstrates that the naive proposal of Eq. (7.37) cannot remain valid under LO evolution if x is of the order of ξ. Assuming Eq. (7.37) is valid at $\mu_0 = 1$ GeV, already at $\mu = 4.7$ GeV, the l.h.s. of the relation may be about 50% bigger than the r.h.s. for small values of $x = \xi$. It is therefore clear that any attempt at linking HVMP data, which is particularly sensitive to the $x \approx \xi$ region, to PDFs must pay attention to the entanglement of x and ξ dependence via evolution equations. We will discuss in the following section Shuvaev's proposal regarding this link, and illustrate it in view of the $\Gamma^{ab,(0)}$ operators.

7.3.2 . Revisiting the role of evolution in Shuvaev's proposal

The Shuvaev transform [110, 250] allows one to relate the representation of GPDs in momentum space as we have mostly used them so far to the representation of GPDs in the space of conformal moments defined in Section 2.4, Eqs (2.54) and (2.55). We remind that conformal moments provide a convenient tool to study evolution as they do not mix under LO evolution. We also noticed in Section 2.4 that the anomalous dimensions γ_n (5.9) governing the general evolution of GPDs are precisely the same that those governing the DGLAP evolution of PDFs. This property will be at the basis of Shuvaev's proposal to reconstruct small ξ GPDs through the effect of evolution.

The Shuvaev operator, which we denote $\mathcal{S}^a(x,\xi;n)$, allows us to relate the two representations of GPDs through

$$H^a(x,\xi,t,\mu^2) = \mathcal{S}^a(x,\xi,n) * \mathcal{O}^a_n(\xi,t,\mu^2)\,, \tag{7.43}$$

where we use the $*$ notation to indicate the action of the Shuvaev operator on an analytical continuation of the conformal moments. The transform provides therefore a resummation of the infinite formal expansion of Eq. (3.42).

Reconstructing the ξ dependence of GPDs can equivalently take the form of defining the ξ dependence of conformal moments. Shuvaev's proposal consists in approximating in the limit $\xi \ll 1$ the conformal moments by their value for $\xi = 0$, *i.e.* (2.56)

$$\mathcal{O}^a_n(\xi,t,\mu^2) \simeq \mathcal{O}^a_n(\xi=0,t,\mu^2) = \int_{-1}^{1} \mathrm{d}x\, x^{n-p_a} H^a(x,\xi=0,t,\mu^2)\,, \tag{7.44}$$

which are exactly the Mellin moments of the PDF if $t = 0$. This proposal greatly simplifies the expression of Eq. (7.43), since it can be turned into the following form

$$H^a(x,\xi,t,\mu^2) \approx \mathcal{S}'^a\left(\frac{\xi}{x},\frac{\mathbf{x'}}{\xi}\right) * a(\mathbf{x'},\mu^2)A(t)\,, \tag{7.45}$$

where we assumed a factorised t-dependence and $\mathcal{S}'^a$ is now the composition of the Shuvaev operator $\mathcal{S}^a$ and the simple Mellin transform (see Appendix A.6 for details). We use boldface character **x'** so that no ambiguity arises on the actual integration variable subtended by the symbol $*$. The practical consequences on GPDs of the assumption of ξ independence of conformal moments has been discussed in Ref. [165]. The proposal has been extended in Ref. [111] which provides a technical fix so that the reconstructed GPD

does not expand outside the correct support $x \in [-1, 1]$. Since both Shuvaev's proposal and its extension of Ref. [111] take the general form

$$\mathcal{O}_n^a(\xi, t, \mu^2) = \mathcal{F}(\xi)\mathcal{O}_n^a(\xi = 0, t, \mu^2)\,, \tag{7.46}$$

with $\mathcal{F}(\xi) = 1$ in Shuvaev's case and $\mathcal{F}(\xi) = 2(\xi/2)^{n+1}T_{n+1}(1/\xi)$ where $T_n(x)$ are Chebychev polynomials of the first kind in Ref. [111], these reconstructions of ξ dependence are more compatible with evolution than the naive (7.37), in the sense that for the non-singlet component, reconstructing first the ξ dependence from the PDF at scale μ_0^2 and evolving the obtained GPD to μ^2, or evolving first the PDF from μ_0^2 to μ^2 and then reconstructing the ξ dependence produce the same result. It can be seen by the fact that the ratio $\mathcal{O}_n^q(\xi, t, \mu^2)/\mathcal{O}_n^q(\xi = 0, t, \mu^2)$ for n even is independent of μ^2 thanks to Eq. (2.57) for LO evolution.

It has been however argued in Ref. [109] that the operations of truncating the ξ expansion of conformal moments - in this case to its first term - and performing their analytical continuation to apply the Shuvaev transform are non-commutative, so that in some cases, the reconstructed GPD would not present the correct forward limit. The authors also describe this procedure as a model among others to reconstruct the small ξ dependence of GPDs not without its own theoretical uncertainty. It is precisely our objective to better understand the domain of validity of this approximation, and affect a quantitative uncertainty to it.

A way to understand Shuvaev's proposal relies on the following remarkable property linking the DGLAP LO evolution operators $\Gamma_0^{ab,(0)}$ introduced in Eq. (7.39) to the general LO evolution operators defined in Eq. (7.38) : provided $y = z/\xi \gg 1$, and any $\mu^2 > \mu_0^2$,

$$\frac{1}{x}\Gamma^{ab,(0)}\left(y, \frac{\xi}{x}; \mu^2, \mu_0^2\right) \approx \frac{1}{x^{p_a}}\mathcal{S}'^a\left(\frac{\xi}{x}, \frac{\mathbf{x'}}{\xi}\right) * \frac{1}{\mathbf{x'}}\Gamma_{DGLAP}^{ab,(0)}\left(\frac{z}{\mathbf{x'}}; \mu^2, \mu_0^2\right)\,. \tag{7.47}$$

Fig. 7.12 shows the excellent quality of this approximation in the qq and gg sectors for $\xi/x = 1$, where we expect the largest deviations between the DGLAP evolution operators and the GPD ones. The agreement in the large y region is of the percent level, and we believe that this discrepancy for $y > 10^3$ is dominated by the precision of the grid we used in the extraction of evolution operators from APFEL++, and not the actual inacurracy of Eq. (7.47). However, in the small y region (in grey on the figure), the agreement of Eq. (7.47) is poor. It is obvious from the details of Appendix A.6 since for $\xi/x = 1$, the r.h.s. of Eq. (7.47) gives non vanishing contributions down to $y = 1/2$, while the l.h.s. vanishes as soon as $y < 1$.

We briefly discussed in Section 7.3.1 general criteria so that most of the weights $\Gamma^{ab,(0)}$ reside in the large y region. Provided there exists an initial scale μ_0^2 such that the large y region completely dominates the evolution to μ^2 for a given PDF, Shuvaev's proposal (7.45) gives

$$\begin{aligned}
&\mathcal{S}'^a\left(\frac{\xi}{x}, \frac{\mathbf{x'}}{\xi}\right) * a(\mathbf{x'}, \mu^2)A(t) \\
&\qquad = \mathcal{S}'^a\left(\frac{\xi}{x}, \frac{\mathbf{x'}}{\xi}\right) * \sum_b \int_0^1 \frac{\mathrm{d}z}{\mathbf{x'}}\,\Gamma_{DGLAP}^{ab,(0)}\left(\frac{z}{\mathbf{x'}}; \mu^2, \mu_0^2\right)\, b(z, \mu_0^2)A(t)\,, &(7.48)\\
&\qquad \approx x^{p_a}\sum_b \int_0^1 \frac{\mathrm{d}z}{x}\,\Gamma^{ab,(0)}\left(y, \frac{\xi}{x}; \mu^2, \mu_0^2\right)\, b(z, \mu_0^2)A(t)\,, &(7.49)
\end{aligned}$$

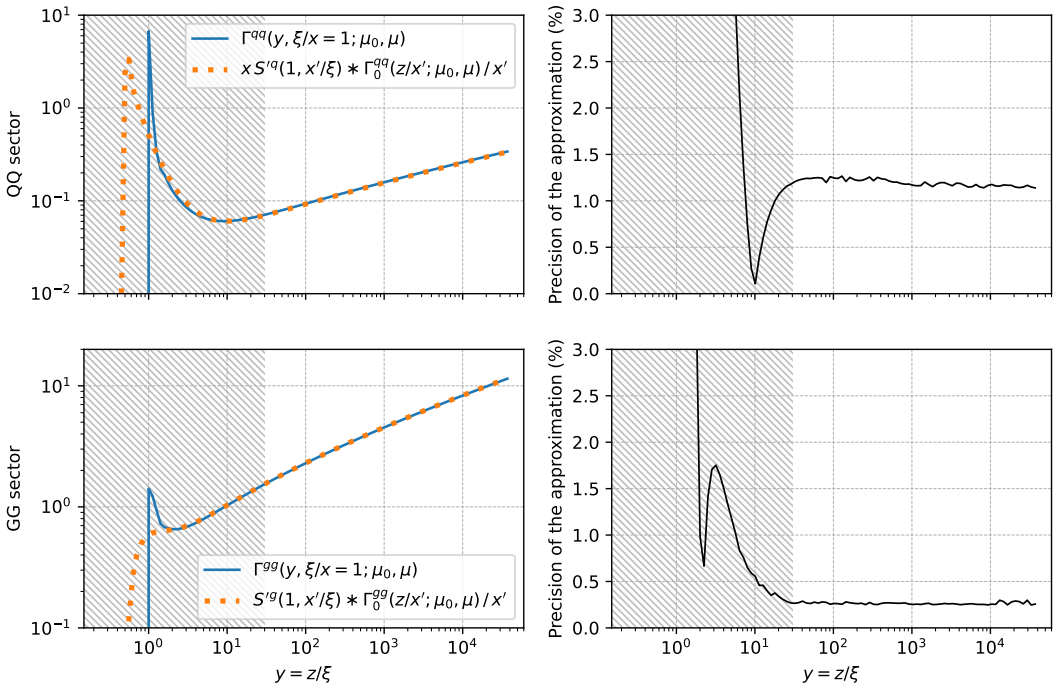

Figure 7.12 – Test of the approximation of LO evolution operators $\Gamma^{ab,(0)}(y, \xi/x; \mu^2, \mu_0^2)$ by Eq. (7.47), as a function of $y = z/\xi$ in the case $\xi/x = 1$, $\mu_0 = 1$ GeV, $\mu = 4.7$ GeV. Test in the qq sector (upper left) and gg sector (lower left). On the right, depiction of the precision of the approximation, reaching $\sim 1\%$ in the qq sector and $\sim 0.3\%$ in the gg sector in the large y region. The light grey region is $y < 30$ which serves as our criterion to discriminate between the small and large y region.

where we used Eq. (7.39) in the first line, and applied the approximation of Eq. (7.47) in the second. Note that by virtue of our assumption that the large y region dominates evolution, the integral of Eq. (7.49) actually takes non negligible values only for $y \geq 30$, that is $z \geq 30\xi$. Then in this region, we can also safely assume $b(z, \mu_0)A(t) \approx z^{-p_b} H^b(z, \xi, t, \mu_0)$, and using Eq. (7.38)

$$\mathcal{S}'^{a}\left(\frac{\xi}{x}, \frac{\mathbf{x'}}{\xi}\right) * a(\mathbf{x'}, \mu^2)A(t) \approx x^{p_a} \sum_b \int_0^1 \frac{\mathrm{d}z}{x} \Gamma^{ab,(0)}\left(y, \frac{\xi}{x}; \mu^2, \mu_0^2\right) \frac{H^b(z, \xi, t, \mu_0^2)}{z^{p_b}}, \quad (7.50)$$

$$= H^a(x, \xi, t, \mu^2)\,. \quad (7.51)$$

This demonstration of the validity of Shuvaev's proposal introduces explicitly of a lower scale μ_0^2 such that the large y region dominates evolution. As the ξ dependence of the GPD at initial scale μ_0^2 is assumed negligible in the large y region, **the Shuvaev transform relies on the idea that the ξ dependence generated by evolution dominates compared to the ξ dependence of the GPD at initial scale μ_0^2**. However any reference to μ_0^2 disappears in the final formulation of the proposal because the relation between general LO and DGLAP evolution operators of Eq. (7.47) holds independently of μ^2. Yet, the disappearance of scale μ_0^2 deprives us of a handle on the systematic uncertainty created by the procedure.

We notice that some generic arguments used to argue of the numerical accuracy of Shuvaev's method have been disputed. For instance, Ref. [261] argues that the ξ depen-

dence of conformal moments only comes from moments with $n > 1$, and can therefore be neglected as all these moments tend to 0 as μ^2 increases. Ref. [35] notes that first, this argument implies that the method is also valid even if ξ is not small although we have built an intuition that it is an important element of the discussion. Second, the ξ independent part of higher order conformal moments tends to 0 at the same speed than the ξ dependent part. There is therefore no reason to neglect only the ξ dependent part of higher order conformal moments, but neglecting them altogether amounts to considering only the asymptotic form of GPDs, which is clearly unsuited from a phenomenological point of view.

Another argument found in Ref. [250] is that the ξ dependent part of conformal moments starts at ξ^2 due to the polynomiality property of GPDs, so that uncertainty caused by neglecting this ξ dependence is of order $\mathcal{O}(\xi^2)$. However, as Ref. [35] notes again, this time the argument seems independent on μ^2, so the actual numerical size of corrections hidden in the $\mathcal{O}(\xi^2)$ must be clarified.

We wish therefore to produce an estimate of the systematic uncertainty which takes into account both ξ and the values of μ^2 and μ_0^2. One could wonder what μ_0^2 actually means as it seems completely arbitrary when trying to relate the GPD to the PDF at μ^2. Intuitively, the best choice of μ_0^2 to maximise the validity of an approach *à la* Shuvaev is always the lowest possible value, since as evolution range increases, so does the dominance of the large y region as we have seen in Section 7.3.1. But owing to the LO formalism developed here, we would still like μ_0^2 to remain large enough. We will therefore fix $\mu_0^2 = 1$ GeV^2, and will derive our estimate of the systematic uncertainty of the reconstruction of the small ξ dependence depending on the value of μ^2.

In addition to the desire of producing an uncertainty estimation, we are also motivated to present an alternative proposal by numerical difficulties in the implementation of Shuvaev's method. As the Shuvaev transform presents itself under the form of a "quite poorly convergent double integral" [262] (see Appendix A.6 for details), it is sometimes not directly used, but simplified under the assumption of power-law behaviour at small x. It is argued in Ref. [262] that this simplification leads sometimes to inaccuracies.

7.3.3 . An alternative proposal

We remind that our demonstration of the validity of Shuvaev's proposal based on evolution operators in the previous section relied crucially on the possibility to make two approximations at a low scale μ_0^2 : first approximating the general LO evolution operators from the DGLAP ones via Eq. (7.47), and then approximating the GPD by the PDF itself. Considering that APFEL++ provides a straightforward way to perform evolution at non-vanishing ξ, the need for the first approximation does not seem particularly crucial. Furthermore, the first approximation is rooted in properties of the LO evolution operator, which we do not expect to hold at higher orders. It seems therefore to us that only keeping the second approximation is simpler, and closely equivalent to Shuvaev's one provided the large y region actually dominates evolution from μ_0^2 to μ^2. This is exactly the naive intuition we had in Eq. (7.37) and which we discarded because it was not scale independent. As we have seen, the Shuvaev transform essentially provides a sophisticated way to transform this assumption into a (partly) scale independent one, but we feel that keeping the scale dependence is on the contrary a great way to quantify the uncertainty in the procedure.

Our proposal is therefore very simply Eq. (7.37) at some scale μ_0^2

$$H^a(x,\xi,t,\mu_0^2) = x^{p_a} a(x,\mu_0^2) A(t)\,, \tag{7.52}$$

and performing the actual GPD evolution to μ^2 afterwards. To the sole purpose of making the difference with Shuvaev's proposal obvious, we express it in the space of conformal moments for the quark GPD and n even, where evolution of conformal moments is very simply summarised by Eq. (2.57) :

— Shuvaev's proposal :

$$\mathcal{O}_n^q(\xi,t,\mu^2) = \int_{-1}^{1} \mathrm{d}x\, x^n q(x,\mu^2) A(t)\,. \tag{7.53}$$

— our proposal with LO evolution :

$$\mathcal{O}_n^q(\xi,t,\mu^2) = \eta(\xi,\mu_0^2) \int_{-1}^{1} \mathrm{d}x\, x^n q(x,\mu^2) A(t)\,, \tag{7.54}$$

where

$$\eta(\xi,\mu_0^2) = \frac{\Gamma(n+1)\Gamma(3/2)}{2^n\Gamma(n+3/2)} \frac{\xi^n \int_{-1}^{1} \mathrm{d}x\, C_n^{(3/2)}(x/\xi) q(x,\mu_0^2)}{\int_{-1}^{1} \mathrm{d}x\, x^n q(x,\mu_0^2)}\,. \tag{7.55}$$

Notice that the dependence on ξ and μ_0^2 of conformal moments with our proposal is entirely dictated by the profile of the PDF at scale μ_0^2.

As we have previously discussed, HVMP data constrains mostly the gluon GPD $H^g(x,\xi,t,\mu^2)$ with μ of the order of half the mass of the heavy meson, and x in the vicinity of ξ, both being very small for events we are interested in. Shuvaev's proposal directly translates this constraint into one on the gluon PDF at the same μ^2 and similarly small x. The situation is very schematically described on the upper left panel of Fig. 7.13. This figure is merely intended for illustration purposes and makes no claim of correct orders of magnitude in the respective uncertainties. We assume that the value of $H^g(x=\xi=10^{-5},\mu^2)$ is extracted with some uncertainty from a LO HVMP analysis, and transformed into a point with uncertainty – represented by the purple dot in the upper left panel of Fig. 7.13 – on the gluon PDF thanks to Shuvaev's proposal. No systematic uncertainty is associated to this transformation. It is then possible to perform a reweighting of the gluon PDF with respect to this new information in a similar way to what we presented in Section 4.4, or a complete refit. The new information is independent of our previous knowledge of the gluon PDF.

The situation is more involved with our proposal which intends to provide a systematic uncertainty on the transformation from GPD to PDF. As our procedure depends on the profile of the gluon PDF at μ_0^2 (7.55), systematic uncertainty will depend on the current knowledge of the gluon PDF.

First, we compute approximate replicas of the GPD $H^g(x=\xi=10^{-5},\mu^2)$ thanks to the ξ dependent GPD evolution operator applied to replicas of our current knowledge of the gluon PDF at initial scale μ_0^2. The latter is represented schematically in the upper right panel of Fig. 7.13, while the constructed approximate GPD is shown on the lower left panel. Our approximate replicas of $H^g(x=\xi=10^{-5},\mu^2)$ are then reweighted against

the HVMP experimental data, giving rise to the blue band. For the reweighting, we add to the experimental uncertainty a systematic one (yellow band) designed to give an account of the approximation we have made by building the GPD from the ξ-dependent evolution of the PDF. The derivation of an estimate of this uncertainty will be detailed below. In practice, the more the large y region dominates evolution from μ_0^2 to μ^2 at the values of x and ξ selected, the lower the systematic uncertainty. With this reweighting performed, it is then possible to measure its impact on the PDF at μ^2 as sketched out in the lower right panel by evolving the reweighted replicas of the gluon PDF from initial scale μ_0^2 to μ^2 [7].

The explicit reference to μ_0^2 provides a tool to estimate the uncertainty associated to the reconstruction of the small ξ dependence. For instance, if μ_0^2 is taken as μ^2 – so no evolution is considered in practice – then our proposal leads to assuming that the GPD is exactly equal to the PDF. We know that this approximation can be wrong by about 50% as we previously discussed at the end of Section 7.3.1. We expect however that taking lower values of μ_0^2 will yield a better approximation. We build an estimate of the systematic uncertainty of the reconstruction of the ξ dependence by measuring the proportion of the contribution of the large y region at initial scale – where we consider that we can safely conflate GPD and PDF at μ_0^2 – to the evolved GPD at scale μ^2. We know that our assumption that the GPD is equal to the t-factorised PDF at initial scale does introduce a significant bias in the small y region, but if the latter is bringing only insignificant contributions to the evolution of the GPD, the bias is considerably reduced and the procedure is sound. However, as we have already mentioned in Section 7.3.1, the actual dominance of the large y region is not just a matter of the properties of the $\Gamma^{ab,(0)}$ operators, but also of the growth of the GPD at small x.

For the sake of illustration of our method, we use the gluon PDF of the MMHT 2014 LO PDF datasets [263], which is precisely defined at $\mu_0^2 = 1$ GeV^2. We propose an estimate of systematic uncertainty of the ξ reconstruction as a relative factor of $1/R$, where

$$R = \frac{\displaystyle\int_{30\xi}^{1} \frac{\mathrm{d}z}{x}\, \Gamma^{gg,(0)}\left(y, \frac{\xi}{x}; \mu^2, \mu_0^2\right)\, g(z, \mu_0^2)}{\displaystyle\int_{0}^{1} \frac{\mathrm{d}z}{x}\, \Gamma^{gg,(0)}\left(y, \frac{\xi}{x}; \mu^2, \mu_0^2\right)\, g(z, \mu_0^2)}\,, \tag{7.56}$$

which measures precisely the proportion of large $y = z/\xi$ contribution to the evolved gluon GPD, where large y is defined as $y > 30$. Fig. 7.14 shows the results obtained for the evolved diagonal $x = \xi$, and $x = 2\xi$. This plot allows us to recover the general features we identified in Section 7.3.1. First, as either μ^2 or the ratio x/ξ increase, so does the contribution of the large y region. Second, as ξ decreases for a fixed value of the ratio x/ξ, two trends are noticeable. When ξ is still rather large, the decrease of ξ triggers a very quick increase of the contribution of the large y region. A first part of this increase is trivial, since the large y region is defined by $z > 30\xi$, and $z < 1$. Therefore, as long as $\xi > 1/30$, there is no contribution from the large y region at all, because ξ is simply too

7. Another way to see the difference between our proposal and Shuvaev's is to notice that, for each replica of the gluon PDF at scale μ_0^2, the relation between the GPD reconstructed at ξ and μ^2 and the PDF at μ^2 is slightly different, as it depends on the profile of the gluon PDF at initial scale. On the contrary, Shuvaev's proposal produces systematically the same relation on the GPD and PDF at scale μ^2.

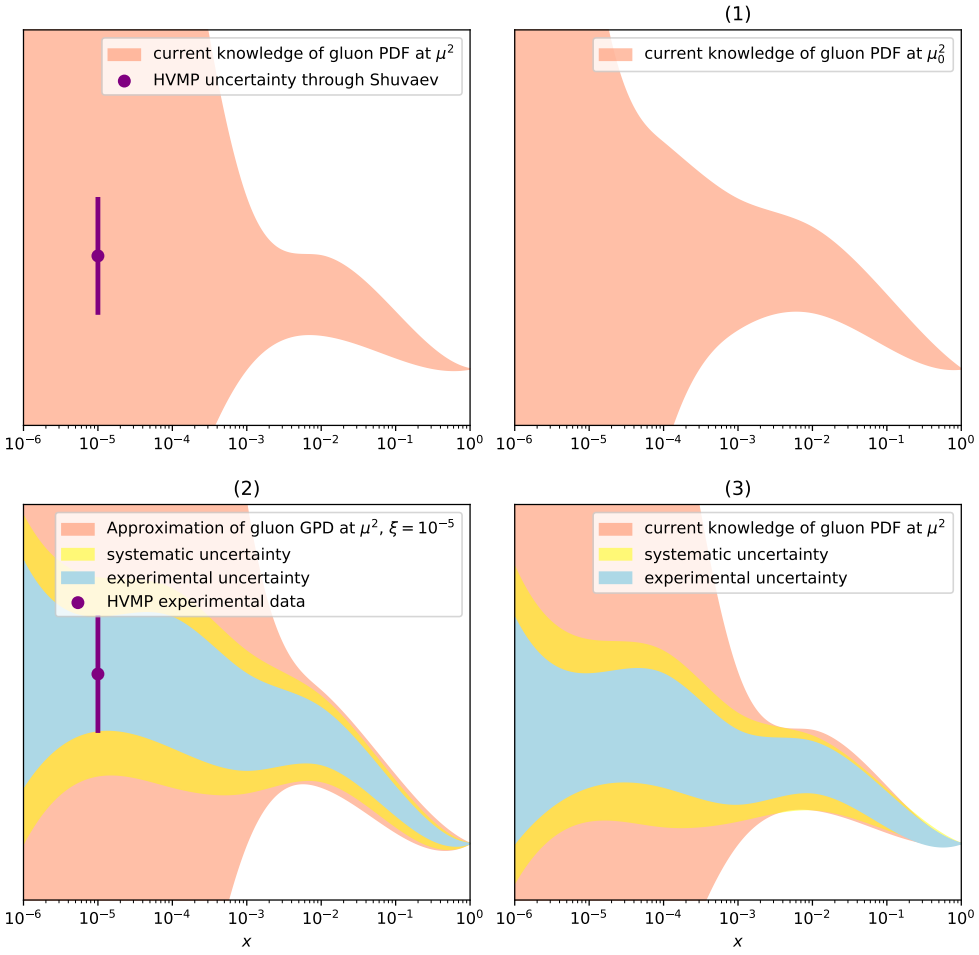

Figure 7.13 – Schematic depiction of the difference between Shuvaev's proposal (upper left panel) and our proposal (the next three panels). Shuvaev's proposal transforms directly uncertainty on the GPD derived from HVMP experimental uncertainty into uncertainty on the PDF. No account of the systematic uncertainty introduced by the transformation is given. On the contrary, we are able to introduce a systematic uncertainty created by the assumption that the GPD is equal to the PDF at a low scale μ_0^2, and add it to the experimental uncertainty to reweight our approximation of the GPD, and therefore the PDF at μ^2.

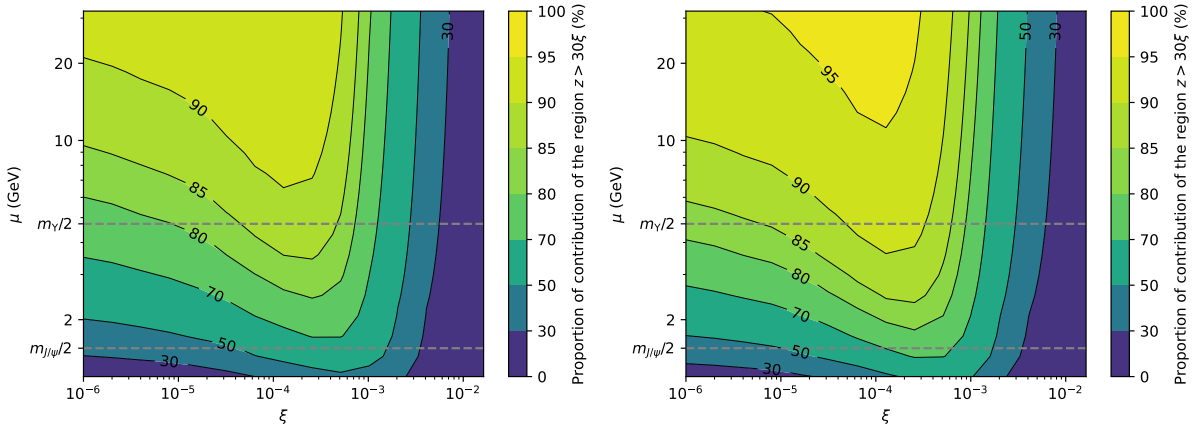

Figure 7.14 – The MMHT 2014 LO gluon PDF $g(x)$ is evolved from $\mu_0 = 1$ GeV to μ with the GPD LO evolution operator at $\xi/x = 1$ (left) and $\xi/x = 0.5$ (right) for various values of ξ and μ. The proportion of contribution of the large $y = z/\xi$ region to the final result is shown.

large for this region to exist. As soon as ξ becomes smaller than $1/30$, the large y region becomes quickly dominant. However, when ξ keeps decreasing below roughly 3×10^{-4}, the contribution of the large y region starts to decrease slowly. The cause of this behavior is an interplay between the increase of the operator weights $\Gamma^{gg,(0)}$ in the large y region and the steep increase of the gluon PDF itself at small y. The power law behavior of the gluon PDF at small y erodes the domination of the large y region in evolution.

The value of μ^2 plays a crucial role in the dominance of the large y region. Whereas at the low value of $\mu = m_{J/\psi}/2 = 1.5$ GeV, the large y region represent just about 50% of the contributions to the evolved GPD in the case $x = \xi$, it quickly increases to about 80% at $\mu = m_\Upsilon/2$. Let us notice however that our estimator of uncertainty is very conservative. It roughly amounts to assuming that the GPD would differ from the PDF by about 100% in the entire region $z \in [x, 30\xi]$ at initial scale μ_0^2. It is interesting to observe for instance what is actually this difference in a popular phenomenological model like GK. We show the difference between the GPD at $\xi = 3 \times 10^{-4}$ and $\xi = 0$ for the u singlet quark on Fig. 7.15. As expected, when nearing the diagonal $z = \xi = 3 \times 10^{-4}$ on the orange curve, the GPD is larger than its forward limit. However, it is not uniformly larger by 100% on the full small y region (grey band), but only very locally in the vicinity of the diagonal and not by more than 25% (about 2600 vs 2100 for $z = \xi$ on the right panel of Fig. 7.15). The difference between the GPD and the PDF is even smaller for the gluon distribution. The actual uncertainty of the reconstruction procedure would therefore be considerably reduced compared to our estimator of Fig. 7.14. As we have previously discussed, models based on the RDDA lack flexibility and only explore a limited range of plausible phenomenological models. We may expect therefore that they produce a lower bound of uncertainty, whereas our conservative estimate is a large upper bound.

7.3.4 . Conclusion

Our proposal (7.52) can be used without any change with any perturbative order of GPD evolution. With beyond LO evolution, the general statement at the basis of Shuvaev's proposal – that is conformal moments of GPDs are independent of ξ and equal to Mellin moments of the PDF – is no longer valid regardless of the scale. Therefore,

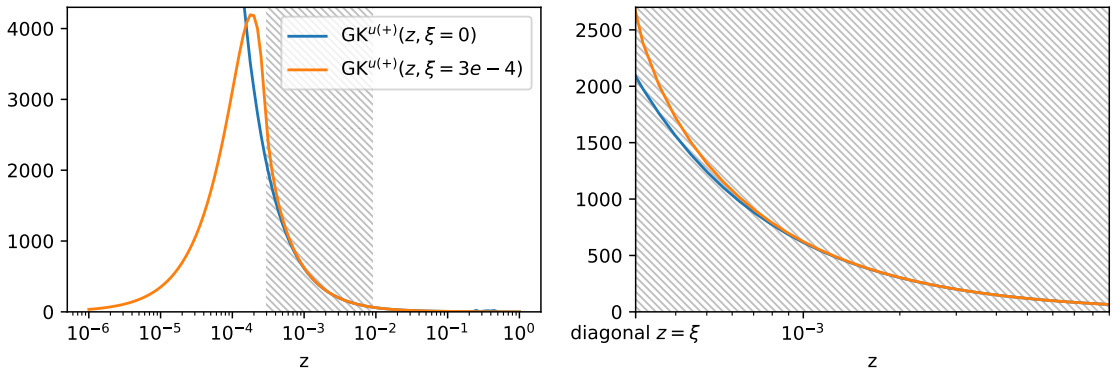

Figure 7.15 – u singlet GPD in the GK model at $\mu_0^2 = 1$ GeV^2 and $t = -0.1$ GeV^2 for $\xi = 0$ (blue curve) and $\xi = 3 \times 10^{-4}$ (orange curve). The small y region, corresponding to $z \in [\xi, 30\xi]$ is represented by the grey area. It is zoomed in on the right panel, where the discrepancy between the PDF and the diagonal can be observed on the left most value of z.

an explicit reference to the arbitrary scale were this statement is chosen to be true has to be made, losing the main advantage of Shuvaev's method compared to ours. Using higher order evolution with our proposal brings also the advantage that it is possible to lower the value of μ_0^2, and therefore decrease the uncertainty of ξ reconstruction. Indeed, although one can always lower the value of the arbitrary value of μ_0^2 to benefit from a larger evolution range, it comes at the expense of reducing the stability of the perturbative expansion and increasing the size of radiative corrections. A trade-off must therefore be found, and lower values of μ_0^2 can be used jointly with higher order evolution. Higher order evolution is also intrinsically important regardless of the question of the lowest achievable value of μ_0^2 since usual techniques to quantify missing higher order corrections, like varying the scale or introducing a custom coupling α_s will not enlighten us on the ξ dependence induced by higher orders evolution.

Conclusion

We have offered a presentation of some aspects of the extraction of GPDs from experimental data, focusing mostly on the quantification of the potential of the actual extraction of GPDs from DVCS.

We have first demonstrated the physical interest of collecting more DVCS observables in varied kinematic regions by showing the potential constraining power of a beam charge asymmetry dataset on the real part of the CFF $\mathcal{H}$. In this impact study related to the installation of a positron beam at JLab, we have detailed the generation of preliminary pseudo-data and their use to reweight a flexible model of the current knowledge of CFFs. The potential for a considerable improvement of our knowledge of $\mathrm{Re}\,\mathcal{H}$ has been demonstrated. As the uncertainty on $\mathrm{Re}\,\mathcal{H}$ is a major cause of the large uncertainty on the DVCS subtraction constant – which we find to be compatible with 0 with a flexible modelling of CFFs fitted on most of the world DVCS dataset –, the improvement of the experimental knowledge of CFFs should be viewed as a priority for a data-driven extraction of proton mechanical properties with lesser model bias. Our study also shows how the Bayesian reweighting technique, which had not been used yet in the context of GPD studies as far as we are aware, provides a useful tool to scan the space of kinematics and observables and select those with the largest expected constraining power.

However, experimental uncertainty is only a part of the overall issue of the extraction of mechanical properties in the current state of experimental data. Thanks to a detailed study of the numerical properties of the D-term evolution, we have been able to measure quantitatively the impact of the range in Q^2 on the robustness of the GFF extraction, and therefore the ability to reduce model dependence in the analysis. We have built an estimate of the impact of EIC kinematics on this uncertainty based on general principles, and shown that a reduction of uncertainty by a factor 3 could be anticipated by the sole effect of increasing the range in Q^2 without even assuming an increased precision of measurements. Furthermore, we have presented results for fits of the GFF at NLO accuracy for the first time. The most striking feature is the increased gluon sensitivity achieved by the direct presence of a gluon contribution to the subtraction constant. We have devoted a significant attention to the discussion of the difference of results between radiative gluon generation from a threshold where the gluon contribution is fixed to 0, and allowing gluon contributions to be freely fitted. We hope that the detailed statistical analysis we have presented will prove useful to interpret the results of fits performed on future datasets whose coverage in Q^2 is expected to be much larger, and effects of evolution will therefore need to be taken into account all the more crucially.

Our presentation of the extraction of the GFF has been largely made under the prism of the deconvolution problem. Our estimate of the impact of the EIC kinematics provides a first tool to understand the numerical stability of the deconvolution procedure. We have also introduced a new formulation in terms of shadow D-terms, whose similarity with our estimate we have demonstrated. These objects are specifically designed to cancel their contribution to the DVCS subtraction constant at a given scale, where they are mathematically impossible to extract from experimental data. The growth of their contribution to the subtraction constant with increasing scale provides a quantitative tool to measure the amplification factor of experimental uncertainty in D-term extractions.

We have generalised this approach to the full deconvolution problem which consists in recovering a complete GPD from the knowledge of CFFs. As far as we know, no systematic attempt at a quantification of the numerical stability of the deconvolution problem at next-to-leading order has been led before. We dedicated a lot of attention to the details of the derivation of NLO shadow GPDs, demonstrating the numerous advantages of the polynomial parameterisation of DDs and the remarkable nested structure which appears when trying to cancel CFFs built from these objects. We have provided detailed formulas to allow the reproduction and extension of this work, and explicit examples of shadow GPDs. The precise study of their behaviour under evolution has led to a quantification of the amplification of the experimental uncertainty on CFFs when recovering GPDs by a factor 10^5 over an evolution range from 1 GeV^2 to 100 GeV^2. Note that this result has been properly demonstrated only if the chosen GPD modelling is flexible enough to include our shadow GPDs. Models with a physical parameterisation with few parameters are less subject to this amplification of uncertainty, although at the cost of a much larger modelling bias. We have discussed the practical consequences of this severe uncertainty issue and argued for the necessity of a multi-channel approach of the extraction of GPDs intended to mitigate the effect of these shadow GPDs.

In addition to being a tool to quantify the amplification of experimental errors during the deconvolution procedure, shadow GPDs give also interesting modelling insights due to the peculiar properties they obey by definition. Having a vanishing forward limit and CFF at the desired order, they can be added to GPD models without bringing any change to DVCS and DIS experimental inputs on the GPD. They are therefore a convenient tool to ensure that a satisfactory model which violates nonetheless positivity constraints can be corrected with minimal changes. We have proposed a model precisely intended to reproduce these experimental inputs at LO at will, so that the affine space made of the addition of this model and the vector space of shadow distributions provides a fully generic analytical solution to the question of extracting a GPD from DIS and LO DVCS. The behaviour of our model is excellent for non-diverging PDFs and CFFs at low x, but rather peculiar in the presence of phenomenological Regge-inspired divergences. We resorted therefore to a neural network modelling of double distributions where we put to use our understanding of shadow GPDs to add flexibility to the model. In practice, we have built a shadow scheme of DDs by adding a factor $(x^2 - \xi^2)$ in front of the singlet GPD and cancelling its forward limit thanks to appropriate normalisations. We have demonstrated that our modelling strategy is able to reproduce a phenomenological divergence at small x, that it can be adapted to fulfill positivity constraints and produces an interesting account of uncertainty related to the sole knowledge of forward limit and CFFs.

Noting that shadow GPDs bring only little contribution to the small ξ region due to their vanishing forward limit and good regularity properties, we have explored the possibility of providing a realistic uncertainty associated with the reconstruction of the small ξ region from the knowledge of the PDF. This matter is of significant importance to constrain gluon PDFs at small x, and has been the subject of several modelling proposals. Among them, the popular Shuvaev method provides an interesting reconstruction based on properties of evolution, which unfortunately does not provide easily the means for the evaluation of the systematic uncertainty of the reconstruction procedure. It is understood from the traditional justification of the method that its validity relies on the small value of ξ and the comparatively large effect of ξ dependence generated by evolution. The recent

publication of a LO GPD evolution code in momentum space gave us an interesting opportunity to propose a closely related but alternative method to reconstruct the ξ dependence. Our proposal allows the computation of an explicit bound of the systematic uncertainty of the reconstruction. The study of the small ξ dependence is important even in a larger context than the determination of small x gluon PDFs, as many physical applications of GPDs are performed at zero skewness – like hadron tomography in the impact parameter representation – although most experimental constraints are obtained at non vanishing skewness. We understand therefore how the uncertainty on experimental extraction characterised by shadow GPDs is translated into an uncertainty in the zero skewness limit.

Among the extensions of this work which could be considered, the most important one in our opinion is to apply the formalism of shadow GPDs to other hard processes like DVMP beyond leading order [8], in order to get a better idea of the potential of a joint analysis at this stage. As the possibilities of analytical derivation of shadow distributions are probably reaching their limits with the complexity of the expressions we had to deal with for NLO DVCS, switching to numerical approximate methods seems necessary. Our neural network model of double distributions including LO shadow contributions seems the ideal basis to start this study. First, the model would have to be improved to include the already known NLO shadow GPDs. This could be performed in an approximate way by not requiring that shadow terms brings an exactly vanishing contribution to the NLO CFF as we did so far, but only a very small contribution compared to the typical size of the PDF for instance. Then an assessment of the impact of a NLO DVMP analysis would yield enlightening conclusions on the potential of tri-dimensional experimental processes [9] on the extraction of DVCS.

Our quantification of the numerical stability of the deconvolution procedure invites to collect experimental measurements for processes with a richer kinematic structure than DVCS and DVMP. We are notably thinking about DDVCS or processes where the active parton emits several particles. The larger number of kinematic dependences of the structure functions of these processes provides a handle that radically changes the considerations on the deconvolution problem. However, the smaller cross-section and more demanding experimental analysis reduce for now the possibilities coming from these channels. On the other hand, lattice QCD brings promises of extractions of x-dependent parton distributions which could change radically our knowledge of hadron structure provided systematic uncertainties are well controlled. Experiment would still have its importance in the case of a lattice breakthrough, since it would be used as a reference to demonstrate precisely the consistence of lattice results.

With the perspective of a significant improvement on the experimental access to GPDs with the JLab 12 GeV upgrade and the projects of future facilities like the EIC, GPDs are on the verge of entering a precision era where the level of modelling bias which is currently considered as a necessity for the sake of producing results with "reasonable" error bars will no longer remain acceptable. We hope that this work has paved the way for this change of paradigm, and will provide useful insights on the behaviour of phenomenological

8. The current formalism is adapted to LO DVMP as the amplitudes are closely related to the DVCS ones.

9. The structure functions of both DVCS and DVMP depend on the three kinematic variables (ξ, t, Q^2).

extractions when the flexibility of the modelling is increased.

A - Appendix

A.1 . Demonstration of the effect of adding a new point to a linear fit

We remind from the derivation in Section 4.1.1 that for a linear fit with generalised least squares loss function, the uncertainty on the extracted parameters $A^{(0)}$ is given by the covariance matrix $\Sigma = (C^T \Omega^{-1} C)^{-1}$ where Ω is the covariance matrix of the vector of experimental measurements F, and C the matrix of probed kinematic configurations. The mean of $A^{(0)}$ is given by $\Sigma C^T \Omega^{-1} F$ (4.6). Assume we add to the current fit a new uncorrelated datapoint evaluated on the kinematic configuration $X^{(1)}$ with central value f_1 and normally distributed experimental uncertainty σ_1. F, Ω and C are upgraded to F', C' and Ω' in the form of block-matrices :

$$F' = \left(\begin{array}{c} F \\ \hline f_1 \end{array} \right), \quad C' = \left(\begin{array}{c} C \\ \hline X^{(1)T} \end{array} \right), \quad \Omega' = \left(\begin{array}{c|c} \Omega & 0 \\ \hline 0 & \sigma_1^2 \end{array} \right), \tag{A.1}$$

and

$$C'^T \Omega'^{-1} C' = \Sigma^{-1} + \frac{1}{\sigma_1^2} X^{(1)} X^{(1)T}. \tag{A.2}$$

Hence the covariance of the extracted parameters $\Sigma' = (C'^T \Omega'^{-1} C')^{-1}$ after addition of the new uncorrelated point can be deduced from the former Σ via

$$\Sigma' = \Sigma \left[\mathrm{Id}_n - \frac{1}{\sigma_1^2} X^{(1)} X^{(1)T} \Sigma + \frac{1}{\sigma_1^4} \left(X^{(1)} X^{(1)T} \Sigma \right)^2 - \cdots \right]. \tag{A.3}$$

where Id_n is the identity matrix of size n. When evaluated at a given kinematic configuration X, the model $f_{A^{(0)}}(X) = X^T A^{(0)}$ follows a simple normal distribution with variance

$$(\Delta f_{A^{(0)}}(X))^2 = X^T \Sigma X. \tag{A.4}$$

Therefore

$$X^{(1)T} \Sigma X^{(1)} = (\Delta f_{A^{(0)}}(X^{(1)}))^2, \tag{A.5}$$

and

$$\Sigma' = \Sigma - \frac{\Sigma X^{(1)} X^{(1)T} \Sigma}{\sigma_1^2 + (\Delta f_{A^{(0)}}(X^{(1)}))^2}, \tag{A.6}$$

where we have used the fact that (A.5) is a scalar. For the new parameters $A'^{(0)}$ obtained after inclusion the additional data point, the uncertainty at kinematic configuration X is given by

$$(\Delta f_{A^{(1)}}(X))^2 = X^T \Sigma' X = (\Delta f_{A^{(0)}}(X))^2 - \frac{(X^T \Sigma X^{(1)})^2}{\sigma_1^2 + (\Delta f_{A^{(0)}}(X^{(1)}))^2}. \tag{A.7}$$

Notice that the uncertainty is systematically reduced or equal when adding a new data point, with equality for X orthogonal to $\Sigma X^{(1)}$ for the usual Euclidian scalar product. For

$X = X^{(1)}$, that is at the very location where the new point is added, Eq. (A.7) simplifies to

$$(\Delta f_{A^{(1)}}(X^{(1)}))^2 = (\Delta f_{A^{(0)}}(X^{(1)}))^2 - \frac{(\Delta f_{A^{(0)}}(X^{(1)}))^4}{\sigma_1^2 + (\Delta f_{A^{(0)}}(X^{(1)}))^2}\,, \tag{A.8}$$

$$= \left(\frac{1}{\sigma_1^2} + \frac{1}{(\Delta f_{A^{(0)}}(X^{(1)}))^2}\right)^{-1}\,. \tag{A.9}$$

A similar result can be obtained concerning the mean of the fit, which is given by

$$\langle f_{A^{(0)}}(X)\rangle = X^T \Sigma C^T \Omega^{-1} F\,. \tag{A.10}$$

Then

$$\langle f_{A^{(1)}}(X^{(1)})\rangle = X^{(1)T}\Sigma' C'^T \Omega'^{-1} F'\,, \tag{A.11}$$

$$= X^{(1)T}\Sigma' \left[C^T\Omega^{-1}F + \frac{f_1}{\sigma_1^2}\,X^{(1)}\right]\,, \tag{A.12}$$

$$= X^{(1)T}\Sigma' C^T\Omega^{-1}F + \frac{(\Delta f_{A^{(1)}}(X^{(1)}))^2}{\sigma_1^2}\,f_1\,. \tag{A.13}$$

The first term can be expressed thanks to Eq. (A.6) :

$$X^{(1)T}\Sigma' C^T\Omega^{-1}F = \langle f_{A^{(0)}}(X^{(1)})\rangle \left[1 + \frac{(\Delta f_{A^{(0)}}(X^{(1)}))^2}{\sigma_1^2}\right]^{-1}\,, \tag{A.14}$$

and finally, introducing

$$\lambda = \left[1 + \frac{(\Delta f_{A^{(0)}}(X^{(1)}))^2}{\sigma_1^2}\right]^{-1}\,, \tag{A.15}$$

we obtain the expected result

$$\begin{cases} \langle f_{A^{(0)\prime}}(X^{(1)})\rangle = \lambda\langle f_{A^{(0)}}(X^{(1)})\rangle + (1-\lambda) f_1\,, \\ (\Delta f_{A^{(0)\prime}}(X^{(1)}))^2 = \lambda(\Delta f_{A^{(0)}}(X^{(1)}))^2\,. \end{cases} \tag{A.16}$$

A.2 . Impact of having to very close functional forms in a linear fit

As we have presented in Section 5.1.2, the deconvolution problem of the D-term should admit a unique solution thanks to the different μ^2 behaviour of the Γ_n^{qq} functions. Although linearly independent, these functions take actually close numerical values, especially if the lever arm in μ^2 is small. Let us try to formalise briefly the impact of enriching a linear model by adding a new functional dependence which is very close numerically to the existing one. Consider the following situation. A fit with uncertainties has been obtained with the following model on some experimental data :

$$\alpha f_1(\mu^2) + \beta f_2(\mu^2)\,, \tag{A.17}$$

where α and β are the free parameters and f_1 and f_2 are known and linearly independent. Let us assume the best fit values (α, β) follow a bivariate normal distribution with variance

$((\Delta\alpha)^2, (\Delta\beta)^2)$ and covariance $\mathrm{cov}(\alpha, \beta)$. Assume now that we replace f_2 by another function

$$f_\epsilon(\mu^2) = (1-\epsilon) f_1(\mu^2) + \epsilon f_2(\mu^2)\,. \tag{A.18}$$

Therefore, as ϵ tends to 0, f_ϵ becomes increasingly close to f_1 while remaining linearly independent. If we fit again the original dataset by replacing f_2 with f_ϵ, the new best fit parameters are noted $(\alpha_\epsilon, \beta_\epsilon)$, and verify

$$\alpha_\epsilon f_1(\mu^2) + \beta_\epsilon f_\epsilon(\mu^2) = \alpha f_1(\mu^2) + \beta f_2(\mu^2)\,. \tag{A.19}$$

Since

$$\alpha_\epsilon f_1(\mu^2) + \beta_\epsilon f_\epsilon(\mu^2) = \left[\alpha_\epsilon + \beta_\epsilon(1-\epsilon)\right] f_1(\mu^2) + \epsilon\beta_\epsilon f_2(\mu^2)\,, \tag{A.20}$$

we obtain readily

$$\alpha_\epsilon = \alpha - \beta\frac{1-\epsilon}{\epsilon}\,, \qquad \beta_\epsilon = \frac{\beta}{\epsilon}\,, \tag{A.21}$$

and

$$(\Delta\alpha_\epsilon)^2 = (\Delta\alpha)^2 + \left(\frac{1-\epsilon}{\epsilon}\right)^2 (\Delta\beta)^2 - 2\frac{1-\epsilon}{\epsilon}\mathrm{cov}(\alpha,\beta)\,, \tag{A.22}$$

$$(\Delta\beta_\epsilon)^2 = \frac{(\Delta\beta)^2}{\epsilon^2}\,, \tag{A.23}$$

$$\mathrm{cov}(\alpha_\epsilon, \beta_\epsilon) = \frac{\mathrm{cov}(\alpha,\beta)}{\epsilon} - \frac{1-\epsilon}{\epsilon^2}(\Delta\beta)^2\,. \tag{A.24}$$

In particular, the linear correlation coefficient reads

$$\mathrm{corr}(\alpha_\epsilon, \beta_\epsilon) = \frac{\mathrm{cov}(\alpha_\epsilon, \beta_\epsilon)}{(\Delta\alpha_\epsilon)(\Delta\beta_\epsilon)}\,, \tag{A.25}$$

$$= -1 + \frac{\epsilon^2}{2}\left[\frac{(\Delta\alpha)^2(\Delta\beta)^2 - \mathrm{cov}(\alpha,\beta)^2}{(\Delta\beta)^4}\right] + \mathcal{O}(\epsilon^3)\,. \tag{A.26}$$

Therefore, as f_1 and f_ϵ become increasingly similar, the distributions of best fit parameters $(\alpha_\epsilon, \beta_\epsilon)$ see their standard deviations grow as $1/\epsilon$, and their correlation coefficient converge quickly to -1. If $|f_1 - f_\epsilon| \ll f_1$, we can propose additionally the following approximation for the model :

$$\alpha_\epsilon f_1(\mu^2) + \beta_\epsilon f_\epsilon(\mu^2) \approx (\alpha_\epsilon + \beta_\epsilon)\, f_1(\mu^2) = (\alpha+\beta)\, f_1(\mu^2)\,. \tag{A.27}$$

Therefore, the fit with only the functional dependence $f_1(\mu^2)$ gives as best fit approximately $\alpha + \beta$, while it gives exactly $\alpha + \beta - \beta/\epsilon$ (A.21) when fitted jointly with $f_\epsilon(\mu^2)$.

A.3 . DVCS coefficient function at next-to-leading order

With the notations of Eq. (3.10), the DVCS coefficient functions at NLO read

$$T_1^q\left(\omega, \frac{Q^2}{\mu^2},\ \alpha_s(\mu^2)\right) = T_0^q\left(\omega\right) + \alpha_s(\mu^2)\log\left(\frac{Q^2}{\mu^2}\right) T_{1,0}^q\left(\omega\right) + \alpha_s(\mu^2) T_{1,1}^q\left(\omega\right)\,, \tag{A.28}$$

$$T_1^g\left(\omega, \frac{Q^2}{\mu^2},\ \alpha_s(\mu^2)\right) = \alpha_s(\mu^2)\log\left(\frac{Q^2}{\mu^2}\right) T_{1,0}^g\left(\omega\right) + \alpha_s(\mu^2) T_{1,1}^g\left(\omega\right)\,. \tag{A.29}$$

We refer to the expressions found in Ref. [149]. As we have mentioned in Section 3.1.1, Ref. [149] uses the traditional $i\epsilon$ prescription which we find quite inconvenient both for the numerical and exact computation of the DVCS convolution (3.10). In particular, in Chapter 5 where we compute the NLO contributions to the subtraction constant and in Chapter 6 where we cancel exactly the NLO CFF of a GPD, we need to use the exact limit $\epsilon \to 0^+$ of the $i\epsilon$ prescription. In both cases, only the imaginary part of the coefficient function will be of interest to us. At LO, the Sokhotski-Plemelj formula provides immediately the solution as (3.13)

$$\lim_{\epsilon\to 0^+} \mathrm{Im}\, T_0^q\left(\omega, \epsilon\right) = e_q^2\pi[\delta(1-\omega) - (\omega \to -\omega)]\,. \tag{A.30}$$

To express higher order terms, it is useful to introduce the plus-prescription which we define as

$$\left(\frac{1}{\omega-1}\right)_+ = \frac{1}{\omega-1} + \delta(\omega-1)\,\mathrm{v.p.}\int_0^\omega \frac{\mathrm{d}z}{z-1}\,. \tag{A.31}$$

The action of the plus-prescription on a test function f yields

$$\int_1^x \mathrm{d}\omega\, \left(\frac{1}{\omega-1}\right)_+ f(\omega) = \int_1^x \mathrm{d}\omega\, \left(\frac{1}{\omega-1}\right) f(\omega) + f(1)\,\mathrm{v.p.}\int_0^1 \frac{\mathrm{d}\omega}{\omega-1}\,, \tag{A.32}$$

$$= \int_1^x \mathrm{d}\omega\, \frac{f(\omega)-f(1)}{\omega-1} + f(1)\,\mathrm{v.p.}\int_0^x \frac{\mathrm{d}\omega}{\omega-1}\,, \tag{A.33}$$

$$= \int_1^x \mathrm{d}\omega\, \frac{f(\omega)-f(1)}{\omega-1} + f(1)\log|x-1|\,. \tag{A.34}$$

Ref. [149] gives that the so-called quark collinear term of the DVCS coefficient function, which we denoted by $T_{1,0}^q$ for reasons which will become clear in Section 6.3.1, writes

$$\frac{1}{\xi} T_{1,0}^q\left(\frac{x}{\xi}, \epsilon\right) = \frac{e_q^2 C_F}{4\pi} \frac{1}{x+\xi-i\epsilon}\left[-3 - 2\log\left(\frac{x+\xi}{2\xi} - i\epsilon\right)\right] - (x \to -x)\,, \tag{A.35}$$

where $C_F = 4/3$. Let us demonstrate that the explicit limit $\epsilon \to 0^+$ yields

$$\lim_{\epsilon\to 0^+} \mathrm{Im}\, T_{1,0}^q\left(\omega, \epsilon\right) = \frac{e_q^2 C_F}{2}\left(\Theta(\omega-1)\left(\frac{1}{\omega-1}\right)_+ + \delta(\omega-1)\left[\frac{3}{2} - \log(2)\right]\right) - (\omega \to -\omega)\,. \tag{A.36}$$

Demonstration Let us assume that $\xi > 0$. Thanks to parity properties, the DVCS convolution can be written for $x \geq 0$ with the explicit introduction of $H^{q(+)}$:

$$\int_{-1}^{1} \frac{\mathrm{d}x}{\xi} T^q\left(\frac{x}{\xi}\right) H^q(x,\xi) = \int_{0}^{1} \frac{\mathrm{d}x}{\xi} T^q\left(\frac{x}{\xi}\right) H^{q(+)}(x,\xi)\,. \tag{A.37}$$

We observe in Eq. (A.35) that no singularity emerges from the term which we have made explicit, due to the fact that $x+\xi > 0$ for $x \geq 0$ and $\xi > 0$. On the contrary, the term symbolically denoted as $-(x \to -x)$ contains singularities for $x = \xi$, which generate an imaginary contribution. Using that

$$\lim_{\epsilon \to 0^+} \frac{1}{1-\omega-i\epsilon} = \text{p.v.}\frac{1}{1-\omega} + i\pi\delta(1-\omega)\,, \tag{A.38}$$

$$\lim_{\epsilon \to 0^+} \log\left(\frac{1-\omega}{2} - i\epsilon\right) = \log\left|\frac{\omega-1}{2}\right| - i\pi\Theta(\omega-1)\,, \tag{A.39}$$

we obtain that the imaginary part of the coefficient function reads

$$\lim_{\epsilon \to 0^+} \mathrm{Im}\, T^q_{1,0}(\omega,\epsilon) = \frac{e_q^2 C_F}{2}\left[\frac{\Theta(\omega-1)}{\omega-1} + \delta(\omega-1)\left(\frac{3}{2} + \log\left|\frac{\omega-1}{2}\right|\right)\right]\,, \tag{A.40}$$

where we have suppressed the principal value due to the Heaviside distribution. Noting that the plus-prescription (A.31) can also be written as

$$\left(\frac{1}{\omega-1}\right)_+ = \frac{1}{\omega-1} + \delta(\omega-1)\log|\omega-1|\,, \tag{A.41}$$

we observe that Eq. (A.40) may be rewritten as

$$\lim_{\epsilon \to 0^+} \mathrm{Im}\, T^q_{1,0}(\omega,\epsilon) = \frac{e_q^2 C_F}{2}\left[\Theta(\omega-1)\left(\frac{1}{\omega-1}\right)_+ + \delta(\omega-1)\left(\frac{3}{2} - \log(2)\right)\right]\,. \tag{A.42}$$

When coming back to the traditional expression of the convolution with integration for ξ from -1 to 1 and the full GPD (and not just its singlet component), it is necessary to complete Eq. (A.42) with $-(\omega \to -\omega)$, producing the expected result. □

Having expressed , $\mathrm{Im}\, T^q_{1,0}(\omega)$ as a distribution without reference to the $i\epsilon$ term, we can compute the exact value of its convolution with a GPD for $\xi > 0$ as

$$\int_{-1}^{1} \frac{\mathrm{d}x}{\xi} \mathrm{Im}\, T^a_{1,0}\left(\frac{x}{\xi}\right) H^q(x,\xi) = \int_{0}^{1} \frac{\mathrm{d}x}{\xi} \mathrm{Im}\, T^a_{1,0}\left(\frac{x}{\xi}\right) H^{q(+)}(x,\xi)\,, \tag{A.43}$$

$$= \int_0^{1/\xi} \mathrm{d}\omega\, \mathrm{Im}\, T^a_{1,0}(\omega)\, H^{q(+)}(\omega\xi,\xi)\,, \tag{A.44}$$

$$= \frac{e_q^2 C_F}{2}\left[H^{q(+)}(\xi,\xi)\left(\frac{3}{2} - \log(2)\right) + \int_1^{1/\xi} \mathrm{d}\omega \left(\frac{1}{\omega-1}\right)_+ H^{q(+)}(w\xi,\xi)\right]\,. \tag{A.45}$$

Eq. (A.34) gives

$$\int_1^{1/\xi} \mathrm{d}\omega \left(\frac{1}{\omega-1}\right)_+ H^{q(+)}(w\xi,\xi)$$
$$= \int_1^{1/\xi} \mathrm{d}\omega \, \frac{H^{q(+)}(\omega\xi,\xi) - H^{q(+)}(\xi,\xi)}{\omega-1} + H^{q(+)}(\xi,\xi) \log\left|\frac{1}{\xi}-1\right| , \tag{A.46}$$
$$= \int_\xi^1 \mathrm{d}x \, \frac{H^{q(+)}(x,\xi) - H^{q(+)}(\xi,\xi)}{x-\xi} + H^{q(+)}(\xi,\xi) \log\left(\frac{1-\xi}{\xi}\right) , \tag{A.47}$$

hence the final result used in Eq. (6.65)

$$\mathrm{Im}\, T^q_{1,0}\left(\frac{x}{\xi}\right) \otimes H^q(x,\xi) =$$
$$\frac{e_q^2 C_F}{2}\left(\left[\frac{3}{2} + \log\left(\frac{1-\xi}{2\xi}\right)\right] H^{q(+)}(\xi,\xi) + \int_\xi^1 \mathrm{d}x \, \frac{H^{q(+)}(x,\xi) - H^{q(+)}(\xi,\xi)}{x-\xi}\right) . \tag{A.48}$$

Contribution to the subtraction constant

The form of Eq. (A.36) also allows us to evaluate the contribution of the quark collinear term to the subtraction constant according to Eq. (3.27) :

$$\frac{2}{\pi}\int_1^\infty \mathrm{d}\omega \, \mathrm{Im}\, T^q_{1,0}(\omega) \int_{-1}^1 \mathrm{d}z \, \frac{D^q(z)}{\omega - z} = \frac{e_q^2 C_F}{\pi}\left[\left(\frac{3}{2} - \log(2)\right)\int_{-1}^1 \mathrm{d}z \, \frac{D^q(z)}{1-z}\right.$$
$$\left. + \lim_{x\to\infty} \int_1^x \frac{\mathrm{d}\omega}{\omega-1}\int_{-1}^1 \mathrm{d}z \, D^q(z)\left(\frac{1}{\omega-z} - \frac{1}{1-z}\right) + \log(x-1)\int_{-1}^1 \mathrm{d}z \, \frac{D^q(z)}{1-z}\right] . \tag{A.49}$$

Using that

$$\frac{1}{\omega-1}\left(\frac{1}{\omega-z} - \frac{1}{1-z}\right) = -\frac{1}{(1-z)(\omega-z)} , \tag{A.50}$$

we obtain

$$\frac{2}{\pi}\int_1^\infty \mathrm{d}\omega \, \mathrm{Im}\, T^q_{1,0}(\omega) \int_{-1}^1 \mathrm{d}z \, \frac{D^q(z)}{\omega - z} = \frac{e_q^2 C_F}{\pi}\int_{-1}^1 \mathrm{d}z \, \frac{D^q(z)}{1-z}\left[\frac{3}{2} - \log(2)\right.$$
$$\left. + \lim_{x\to\infty} \log(x-1) - \int_1^x \frac{\mathrm{d}\omega}{\omega - z}\right] . \tag{A.51}$$

Then

$$\lim_{x\to\infty} \log(x-1) - \int_1^x \frac{\mathrm{d}\omega}{\omega-z} = \lim_{x\to\infty} \log\left(\frac{(z-1)(x-1)}{z-x}\right) = \log(1-z) , \tag{A.52}$$

so we obtain finally the contribution of the quark collinear term to the subtraction constant as

$$\frac{2}{\pi}\int_1^\infty \mathrm{d}\omega \, \mathrm{Im}\, T^q_{1,0}(\omega) \int_{-1}^1 \mathrm{d}z \, \frac{D^q(z)}{\omega - z} = \frac{e_q^2 C_F}{\pi}\int_{-1}^1 \mathrm{d}z \, \frac{D^q(z)}{1-z}\left[\frac{3}{2} + \log\left(\frac{1-z}{2}\right)\right] \tag{A.53}$$

The derivation is therefore rather straightforward although a bit cumbersome, and we give directly the results for the 1-loop quark and gluon terms in Eqs. (5.104) and (5.105).

A.4 . Radon transform in the context of the polynomial parameterisation of DDs

$$\text{If } \; f^{q(+)}(\beta,\alpha) = \sum_{\substack{m,n=0\\ m \text{ even}}}^{m+n\leq N} c_{m,n}\alpha^m\beta^n\,, \tag{A.54}$$

then, for $x > |\xi|$,

$$H^{q(+)}(x,\xi) = \sum_{u=1,v=0}^{N+1}\left[\frac{1}{(1+\xi)^u} + \frac{1}{(1-\xi)^u}\right] q_{u,v}x^v\,, \tag{A.55}$$

where

$$q_{u,v} = \sum_{\substack{m=0\\ m \text{ even}}}^{\min(u,N)-1}\;\sum_{n=\max(u,v)-m-1}^{N} c_{m,n}\sum_{j=0}^{n}\frac{(-1)^{u+v+j+n+1}}{m+j+1}\binom{n}{j}\binom{j}{m-u+j+1}\binom{m+j+1}{v-n+j}\,. \tag{A.56}$$

Demonstration In the region $x > |\xi|$,

$$H^{q(+)}(x,\xi) = \sum_{\substack{m,n=0\\ m \text{ even}}}^{m+n\leq N} c_{m,n}\int_{(x-1)/(1+\xi)}^{(1-x)/(1-\xi)} d\alpha\;\alpha^m(x-\alpha\xi)^n\,, \tag{A.57}$$

$$= \sum_{\substack{m,n=0\\ m \text{ even}}}^{m+n\leq N} c_{m,n}\sum_{j=0}^{n}\binom{n}{j}(-\xi)^j x^{n-j}\int_{(x-1)/(1+\xi)}^{(1-x)/(1-\xi)} d\alpha\;\alpha^{m+j}\,, \tag{A.58}$$

$$= \sum_{\substack{m,n=0\\ m \text{ even}}}^{m+n\leq N} c_{m,n}\sum_{j=0}^{n}\binom{n}{j}(-\xi)^j x^{n-j}\frac{(1-x)^{m+j+1}}{m+j+1}\left(\frac{1}{(1-\xi)^{m+j+1}} + \frac{(-1)^j}{(1+\xi)^{m+j+1}}\right). \tag{A.59}$$

Developing the term $(1-x)^{m+j+1}$ and then performing the partial fractions expansion thanks notably to Eq. (6.40) yields

$$H^{q(+)}(x,\xi) = \sum_{\substack{m,n=0\\ m \text{ even}}}^{m+n\leq N} c_{m,n}\sum_{j=0}^{n}\sum_{k=0}^{m+j+1}\sum_{l=0}^{j}\binom{n}{j}\binom{j}{l}\binom{m+j+1}{k} \times \frac{(-1)^{j+l+k}}{m+j+1}x^{n-j+k}\left(\frac{1}{(1-\xi)^{m+j+1-l}} + \frac{1}{(1+\xi)^{m+j+1-l}}\right). \tag{A.60}$$

A straight-forward variable change $v = n-j+k$ and $u = m+j+1-l$ gives the expected result, and the refinement of the summation bounds for m and n comes from a simple study of the non vanishing binomial coefficients in Eq. (A.60). □

Inversion of the Radon transform matrix

As described in the main text, we extract a subset of $q_{u,v}$ such that the matrix relating them to the $c_{m,n}$ is invertible by selecting

$$\begin{aligned} &q_{1,1}, q_{1,2}, ..., q_{1,N+1}, \\ &\quad q_{3,3}, q_{3,4}, ..., q_{3,N+1}, \\ &\quad\quad ...., q_{N,N}, q_{N,N+1} \end{aligned} \tag{A.61}$$

Ordering the selected $q_{u,v}$ by alphabetic order of (v, u) and the $c_{m,n}$ by alphabetic order of $(m + n, m)$, we obtain a triangular, block-diagonal, invertible matrix of the Radon transform.

Demonstration Thanks to eq. (A.56), we know that the coefficient $q_{u,v}$ is calculated using $c_{m,n}$ coefficients such that

$$m - u + j + 1 \leq j \implies m \leq u - 1\,, \tag{A.62}$$

$$v - n + j \leq m + j + 1 \implies v - m - 1 \leq n \implies n \geq v - u\,. \tag{A.63}$$

If we select a series of coefficients $q_{u,v}$ by either increasing u or decreasing v at u kept fixed, the last selected coefficient $q_{u,v}$ involves the $c_{u-1,v-u}$ coefficient which was never used to calculate any previous $q_{u',v'}$. This selection requires that u is odd since $u - 1$ has to be even, and that $v \geq u$. It is the choice made in (A.61). With this selection and ordering, the matrix relating the $q_{u,v}$ to the $c_{m,n}$ is therefore triangular inferior. We can check additionally that the coefficient of $c_{u-1,v-u}$ in the calculation of $q_{u,v}$ is equal to

$$\sum_{j=0}^{v-u} \frac{(-1)^{j+1}}{u+j} \binom{v-u}{j} = -\frac{(v-u)!(u-1)!}{v!} = -\frac{1}{u}\binom{v}{u}^{-1}\,, \tag{A.64}$$

which in particular is never 0. Therefore, using the selection of (A.61) will give an injective matrix. It is then enough to notice that the matrix is square for all N to show that it is invertible. We obtain for instance at $N = 3$

$$\begin{pmatrix} q_{1,4} \\ q_{1,3} \\ q_{1,2} \\ q_{1,1} \\ q_{3,4} \\ q_{3,3} \end{pmatrix} = \begin{pmatrix} -1/4 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1/3 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ -1/4 & 0 & 0 & 0 & -1/12 & 0 \\ 0 & -1/3 & 0 & 0 & 0 & -1/3 \end{pmatrix} \begin{pmatrix} c_{0,3} \\ c_{0,2} \\ c_{0,1} \\ c_{0,0} \\ c_{2,1} \\ c_{2,0} \end{pmatrix}\,. \tag{A.65}$$

Although this demonstrates the reason behind the selection of $q_{u,v}$ and that the matrix is invertible, it is not a good ordering of the elements, since when N increases, new lines are inserted in the middle of the matrix and not simply at the end. The next step is to observe that for our selection of $q_{u,v}$, the contribution of $c_{m,n}$ is 0 if $m+n > v-1$. This is a non-trivial observation which we will justify later. Therefore, since $m+n \geq v-1$ (A.63), only terms where $m+n = v-1$ are

non-zero, and the suggested ordering in the main text transforms the matrix into a block-diagonal one with excellent properties with respect to the growth of N. Finally, note that the calculated terms of the diagonal are also the eigenvalues of the matrix. This gives a hint at the conditioning of the Radon transform. However, due to the fact that we are working in the quite peculiar bases of the coefficients $c_{m,n}$ and a subset of the coefficients $q_{u,v}$, it is not clear how to properly interpret this hint practically. □

For instance, at order $N = 3$ with the proper ordering,

$$\begin{pmatrix} q_{1,1} \\ q_{1,2} \\ q_{1,3} \\ q_{3,3} \\ q_{1,4} \\ q_{3,4} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1/3 & 0 & 0 & 0 \\ 0 & 0 & -1/3 & -1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1/4 & 0 \\ 0 & 0 & 0 & 0 & -1/4 & -1/12 \end{pmatrix} \begin{pmatrix} c_{0,0} \\ c_{0,1} \\ c_{0,2} \\ c_{2,0} \\ c_{0,3} \\ c_{2,1} \end{pmatrix} . \tag{A.66}$$

The structure of the block corresponding to $m+n = N$ even in the block-diagonal matrix is given by

$$\begin{pmatrix} q_{1,N+1} \\ q_{3,N+1} \\ \vdots \\ q_{N+1,N+1} \end{pmatrix} = \begin{pmatrix} A_1(1,N+1) & 0 & \cdots & 0 \\ A_3(3,N+1) & A_1(3,N+1) & & 0 \\ \vdots & & \ddots & \\ A_{N+1}(N+1,N+1) & A_{N-1}(N+1,N+1) & \cdots & A_1(N+1,N+1) \end{pmatrix} \begin{pmatrix} c_{0,N} \\ c_{2,N-2} \\ \vdots \\ c_{N,0} \end{pmatrix} \tag{A.67}$$

where

$$A_k(u,N+1) = -\frac{(N+k-u)!(u-1)!}{(k-1)!(N+1)!} . \tag{A.68}$$

The only change brought by choosing N odd would be that the last line stops at $q_{N,N+1}$ and not $q_{N+1,N+1}$.

Demonstration We start from the expression of eq. (A.56). As explained before, only the coefficients such that $m \leq u-1$ and $m+n = v-1$ are used to compute $q_{u,v}$ if said $q_{u,v}$ belongs to the selection (A.61). This means we can rewrite (A.56)

$$q_{u,v} = \sum_{m=0 \text{ even}}^{u-1} c_{m,v-m-1} \sum_{j=0}^{v-m-1} \frac{(-1)^{j+1}}{m+j+1} \binom{v-m-1}{j} \binom{j}{m-u+j+1} \binom{m+j+1}{m+j+1} , \tag{A.69}$$

$$= \sum_{m=0 \text{ even}}^{u-1} c_{m,v-m-1} \frac{(v-m-1)!}{(u-m-1)!} \sum_{j=0}^{v-m-1} \frac{(-1)^{j+1}}{m+j+1} \frac{1}{(v-m-1-j)!(m-u+j+1)!} , \tag{A.70}$$

and obtain that

$$q_{u,v} = -\sum_{m=0 \text{ even}}^{u-1} \frac{(v-m-1)!}{(u-m-1)!} \frac{(u-1)!}{v!} c_{m,v-m-1} = \sum_{m=0 \text{ even}}^{u-1} A_{u-m}(u,v)\, c_{m,v-m-1} , \tag{A.71}$$

which is a much simpler formulation than Eq. (A.56), but only valid for u odd and $v \geq u$ since it relies crucially on the choice of (A.61). The fact that coefficients with $m+n > v-1$ are systematically 0, which we have not demonstrated so far, is obtained by a similar calculation. One can even use directly the current result, since these terms would correspond to $A_k(u, N+1)$ with negative values of k. Eq. (A.68) can still be used where $(k-1)!$ is replaced by Euler's $\Gamma(k)$, which is infinite for negative integer values – giving indeed that the coefficients are 0. □

We are now ready to verify that the inverse of the Radon transform matrix is obtained thanks to

$$c_{m,n} = -\binom{n+m}{m}(n+m+1)\sum_{k=0\,\text{even}}^{m}\binom{m}{k}E_{m-k}\;q_{k+1,n+m+1}\,, \tag{A.72}$$

where the E_{2i} are Euler numbers. We adhere to the convention of alternating signs (sequence A028296 in the OEIS – https://oeis.org/search?q=A028296) : $E_0 = 1$, $E_2 = -1$, $E_4 = 5$, $E_6 = -61$, ...

Demonstration We verify that the combination of our formulas (A.71) and (A.72) gives back the identity :

$$q_{u,v} = -\sum_{m=0\,\text{even}}^{u-1}\frac{(v-m-1)!}{(u-m-1)!}\frac{(u-1)!}{v!}\;c_{m,v-m-1}\,, \tag{A.73}$$

$$= \sum_{m=0\,\text{even}}^{u-1}\frac{(v-m-1)!}{(u-m-1)!}\frac{(u-1)!}{v!}\frac{(v-1)!}{(v-m-1)!m!}\;v\sum_{k=0\,\text{even}}^{m}\binom{m}{k}E_{m-k}\;q_{k+1,v}\,, \tag{A.74}$$

$$= \sum_{m=0\,\text{even}}^{u-1}\binom{u-1}{m}\sum_{k=0\,\text{even}}^{m}\binom{m}{k}E_{m-k}\;q_{k+1,v}\,, \tag{A.75}$$

$$= \sum_{k=0\,\text{even}}^{u-1}q_{k+1,v}\sum_{m=k\,\text{even}}^{u-1}\binom{u-1}{m}\binom{m}{k}E_{m-k}\,, \tag{A.76}$$

$$= \sum_{k=0\,\text{even}}^{u-1}q_{k+1,v}\binom{u-1}{k}\sum_{m=0\,\text{even}}^{u-k-1}\binom{u-1-k}{m}E_m\,. \tag{A.77}$$

$$\tag{A.78}$$

To conclude, it is enough to use the remarkable, yet rarely mentioned property that

$$\sum_{m=0\,\text{even}}^{u}\binom{u}{m}E_m = 1 \text{ if } u=0, \text{ and } 0 \text{ for } u \text{ even} > 0\,. \tag{A.79}$$

(see for instance `https://mathworld.wolfram.com/EulerNumber.html` where it is symbolically noted $(E-i)^u = 0$ where E^m is to be interpreted as $|E_m|$ due to our sign convention). □

A.5 . An example of NLO shadow GPD

Here is the exact expression of one of the three polynomial DDs of lowest order $N = 25$ which cancel analytically their forward limit, LO, collinear and 1- loop CFFs and vanish on the edges of the rhombus. The coefficients of the monomials are fractions with about 25 figures in their numerator and denominator.

The numerical value of the associated GPD rounded up to 20 decimals is available in the PARTONS framework [144] as the module GPDBDMMS21. Note that even with such a precise rounding, numerical errors can appear for $\xi \geq 0.8$ in the region $|x| > |\xi|$. If sampling at very large values of ξ is required, we recommend to use the exact formula below with a fraction module.

The code of the PARTONS framework is open source and can be found online at https://drf-gitlab.cea.fr/partons/core/partons on version 3 of the GPL (GPLv3).

Value of the DD

$f^{q(+)}(\beta,\alpha) =$ - 902951624458552763993/7949224070189526888655776 $\alpha^0\beta^1$ +16451533817059854624472208117797274868580182160 $\alpha^0\beta^3$ - 150154394071077855923032649741356729842296218592 $\alpha^0\beta^5$ +769580331186256476958944 $\alpha^0\beta^7$ - 622099668901285366186234162355945497371603643 $\alpha^0\beta^9$ +378876730074805795892507/662435339182460574054648 $\alpha^0\beta^{11}$ - 126065012408676306437886472604 $\alpha^0\beta^{13}$ +5166507853713552899224471104058898637434290091080 $\alpha^0\beta^{15}$ - 211243440125717184838063512649741356729842296218592 $\alpha^0\beta^{17}$ +3578983979552941705255986739746120350947634443278888 $\alpha^0\beta^{19}$ - 5644458778059943411870494/883247118909947432072864 $\alpha^0\beta^{21}$ +4257963054831799340828089/1656088347956151435136620 $\alpha^0\beta^{23}$ - 8871941542190270307968751987306017547381722163944 $\alpha^0\beta^{25}$ +5370264770844120458307/4416235594549737160364322 $\alpha^2\beta^1$ - 25332177063896702834235/4416235594549737160364322 $\alpha^2\beta^3$ +342324919768286206311522808441739780757175683121 $\alpha^2\beta^5$ - 868412204315398253945855520294493187171450455422 $\alpha^2\beta^7$ +9511643086800499945411092208117797274868580182162 $\alpha^2\beta^9$ - 5622610481573927305388687/662435339182460574054648 $\alpha^2\beta^{11}$ +829252629367280752867485/1104058898637434290091082 $\alpha^2\beta^{13}$ +153552787808340786492937911040588986374342900910802 $\alpha^2\beta^{15}$ - 735394660909338889880260151324870678364921148109296 $\alpha^2\beta^{17}$ +341514340153127157185258374416235594549737160364322 $\alpha^2\beta^{19}$ - 5658457905176554327251771/1104058898637434290091082 $\alpha^2\beta^{21}$ +443597077109513515398437533121766959123028702732424 $\alpha^2\beta^{23}$ - 213132741228535919841225/883247118909947432072864 $\alpha^4\beta^1$ +778916879993220033913751104058898637434290091082 $\alpha^4\beta^3$ - 6093099596344738990353551104058898637434290091082 $\alpha^4\beta^5$ +129536794349729230160177651104058898637434290091082 $\alpha^4\beta^7$ - 2205165815055358060644518513248706783649211481092960 $\alpha^4\beta^9$ +848816622290825400692585/1104058898637434290091082 $\alpha^4\beta^{11}$ +2146239115711889506366275/2208117797274868580182160 $\alpha^4\beta^{13}$ +948995215014585960533798533121766959123028702732424 $\alpha^4\beta^{15}$ - 1418651433041252583889017058832471189099474320728640 $\alpha^4\beta^{17}$ +607762826111919118893125027601472465935857252277 $\alpha^4\beta^{19}$ - 2217985385547567576992187522081177797274868580182160 $\alpha^4\beta^{21}$ +673028602010250304175875/3312176695912302870273242 $\alpha^6\beta^1$ - 2924764390261684821156391104058898637434290091082 $\alpha^6\beta^3$ +363746206801820942803694911040588986374342900910802 $\alpha^6\beta^5$ - 4965346936133528256678709/3312176695912302870273242 $\alpha^6\beta^7$ +250355122310797643251580311040588986374342900910802 $\alpha^6\beta^9$ +185182936903681558611902311040588986374342900910802 $\alpha^6\beta^{11}$ +365407620652463764559599111040588986374342900910802 $\alpha^6\beta^{13}$ +749145728647330227939961911040588986374342900910802 $\alpha^6\beta^{15}$ - 884490285330684650085362760147246593585785725227 $\alpha^6\beta^{17}$ +3105179539766594607789062582804417397807571756831 $\alpha^6\beta^{19}$ - 4136704165765632413399985/4416235594549737160364322 $\alpha^8\beta^1$ - 1404535763174073640435712208117797274868580182160 $\alpha^8\beta^3$ - 5304100246073677603631171144162355945497371603643 $\alpha^8\beta^5$ - 1279039388130044691525061511040588986374342900910802 $\alpha^8\beta^7$ - 7310656858730326172409746744162355945497371603643 $\alpha^8\beta^9$ - 4490202550472259046779947522081177797274868580182160 $\alpha^8\beta^{11}$ - 8838041651437840971261293744162355945497371603643 $\alpha^8\beta^{13}$ - 8724255239384306834681727/5520294493187171450455422 $\alpha^8\beta^{15}$ - 9315538619299783823367187511040588986374342900910802 $\alpha^8\beta^{17}$ +5860376206888412998007681/2208117797274868580182160 $\alpha^{10}\beta^1$ +3347564527946820174796102311040588986374342900910802 $\alpha^{10}\beta^3$ +330255915744555853066478587/1104058898637434290091082 $\alpha^{10}\beta^5$ +1816111521589272643794440031331217665991230287027324 $\alpha^{10}\beta^7$ +17765915587545490542589182122081177797274868580182160 $\alpha^{10}\beta^9$ +2217004357375462576805350392208117797274868580182160 $\alpha^{10}\beta^{11}$ +3195412329265643150989287272760147246593585725225770 $\alpha^{10}\beta^{13}$ +6831394987486508137135937555202944931871714504554 $\alpha^{10}\beta^{15}$ - 6482789265452565167313456713247870678364921148109296 $\alpha^{12}\beta^1$ - 1019152265613712949299990511040588986374342900910802 $\alpha^{12}\beta^3$ - 1110876464157759481154225372208117797274868580182160 $\alpha^{12}\beta^5$ - 4605952000027962345645929543202304729290907003564 $\alpha^{12}\beta^7$ - 6852346283547674813301586374416235594549737160364322 $\alpha^{12}\beta^9$ - 4689012936098504900327104327601472465935857252277 $\alpha^{12}\beta^{11}$ - 4036733401696572990125781253312176695912302870273242 $\alpha^{12}\beta^{13}$ +6630001426758260012167725/1104058898637434290091082 $\alpha^{14}\beta^1$ +1509182674425984939534892511040588986374342900910802 $\alpha^{14}\beta^3$ +5921708616715109122570491511040588986374342900910802 $\alpha^{14}\beta^5$ +1169762976905723554626365551104058898637434290091082 $\alpha^{14}\beta^7$ +3482554489662818348169367527601472465935857252277 $\alpha^{14}\beta^9$ +2217985385547567576992187527601472465935857252277 $\alpha^{14}\beta^{11}$ - 4310966330912662373405258832471189099474320728642 $\alpha^{16}\beta^1$ - 4945304719603865543197622944162355945497371603643 $\alpha^{16}\beta^3$ - 2924639540888481054147184378832471189099474320728642 $\alpha^{16}\beta^5$ - 5588304616620016928064137711040588986374342900910802 $\alpha^{16}\beta^7$ - 7541150310861729761773437522081177797274868580182160 $\alpha^{16}\beta^9$ +1005122053400872829170087153974612035094763444327888 $\alpha^{18}\beta^1$ +2152419387959174662566175744162355945497371603643 $\alpha^{18}\beta^3$ +3324384120434767286579520733121766959123028702732424 $\alpha^{18}\beta^5$ +8428344465080756792570312599365300877369086108197244 $\alpha^{18}\beta^7$ - 6691392647034077116454397/883247118909947432072864 $\alpha^{20}\beta^1$ - 1001266706229323003191134276011586087665544 $\alpha^{20}\beta^3$ - 6210359079533189215578125/662435339182460574054648 $\alpha^{20}\beta^5$ + $\alpha^{22}\beta^1$

Value of the GPD for $x > |\xi|$

The nullity of the forward limit and LO CFF are made very explicit respectively by the ξ^2 and $(x^2-\xi^2)$ factors, whereas $(1-x)^5$ guarantees a very smooth behaviour at large x.

$H^{q(+)}(x,\xi)$ = (65536 (1 - x)^5 (x^2 - ξ^2) ξ^2 (-83637711531927981 15x - 418188555765963990575 x^2 - 2291958055186829401725 x^3 - 11125239429806435084025 x^4 + 913222467358957809358284 x^5 + 46546458810486617307691 80 x^6 + 23806967947745329028806 6140 x^7 - 29627608879356577216781 89860 x^8 + 65698703668100841508783 05780 x^9 + 18761012723785934692103 395128 x^{10} - 76928400823343726746081 064520 x^{11} + 41017836020009001958211 348760 x^{12} + 76834572510678115789139 906520 x^{13} - 22420022896664405471246 022060 x^{14} - 66864462549508597243006 970172 x^{15} - 15512628910724367358088 291100 x^{16} + 23726225441438503939679 841780 x^{17} + 16767969514739610882060 937500 x^{18} + 33535939029479221764121 87500 x^{19} + 19364775039847258768 20$x\xi^2$ + 96823875199236293841 00 $x^2\xi^2$ - 11481314695808908530826 80 $x^3\xi^2$ - 58181164480638433004862 00 $x^4\xi^2$ + 12192384117487031237581 6872 $x^5\xi^2$ - 44783613359519589710267 64120 $x^6\xi^2$ + 41569429075455721504118 181624 $x^7\xi^2$ - 65873134592387896080837 321450 $x^8\xi^2$ - 48954563055788594220981 1607250 $x^9\xi^2$ + 20047040127781831350293 54484054 $x^{10}\xi^2$ - 22345348562634265924902 73718850 $x^{11}\xi^2$ - 30787013167180182704695 6893912 $x^{12}\xi^2$ + 12614278367070327586220 49674280 $x^{13}\xi^2$ + 43080239118009489911 6978662500 $x^{14}\xi^2$ - 36158943379255526362444 9291116 $x^{15}\xi^2$ - 27045363208960472006094 9014760 $x^{16}\xi^2$ - 43359225928487593047 670802952 $x^{17}\xi^2$ + 67071878058958443528243 75000 $x^{18}\xi^2$ + 13414375611791688705648 75000 $x^{19}\xi^2$ + 33668139410495638498 9659$x\xi^4$ + 16834069705247819249932 95 $x^2\xi^4$ - 41355875799181768444651 1355 $x^3\xi^4$ + 78447825943227161306867 6865 $x^4\xi^4$ - 33833380732598545306708 111584 $x^5\xi^4$ - 86920937387611069554046 719120 $x^6\xi^4$ + 46889018506328928904029 8761392 $x^7\xi^4$ + 22377286106141284751585 36274810 $x^8\xi^4$ - 13923931065860427209313 8614309 10 $x^9\xi^4$ + 23653486346969934917789 786747702 $x^{10}\xi^4$ - 10763700454260477933616 617268674 $x^{11}\xi^4$ - 63096797269571616693 75775894156 $x^{12}\xi^4$ + 25223198741560002025 13219862340 $x^{13}\xi^4$ + 22272140810385597205040 91473100 $x^{14}\xi^4$ + 28806301233350446950337 2891612 $x^{15}\xi^4$ - 97840707258698883729 294814380 $x^{16}\xi^4$ - 18971946980104590581 163462876 $x^{17}\xi^4$ + 37262154477199135293 4687500 $x^{18}\xi^4$ + 74524308954398270586 937500 $x^{19}\xi^4$ + 10200525243684812551454 8332$x\xi^6$ - 84325769585404075230 9458340 $x^2\xi^6$ - 18783135855456536941 745093160 $x^3\xi^6$ + 16422813710063331056408 7429420 $x^4\xi^6$ + 22183855741924236616 7978256396 $x^5\xi^6$ - 43135898422718764555 15907072700 $x^6\xi^6$ + 64478176331213941784754 93189684 $x^7\xi^6$ + 24140786122579579681685 842862640 $x^8\xi^6$ - 81227147644572001593190 607666000 $x^9\xi^6$ + 75735173003934541979411 445995792 $x^{10}\xi^6$ - 77825422947289147545 16730052912 $x^{11}\xi^6$ - 13299786278241568754 818657921132 $x^{12}\xi^6$ - 14863591665806686601571 969885980 $x^{13}\xi^6$ + 72409046983870994985 7175005740 $x^{14}\xi^6$ + 13406571016366396688 14090937532 $x^{15}\xi^6$ - 67202398775487632233400 39760 $x^{16}\xi^6$ - 13440479755097526446680 07952 $x^{17}\xi^6$ - 41023462979179785791 1150000 ξ^8 + 96491580346592470096302 40014$x\xi^8$ - 34056348347179890223302 830670 $x^2\xi^8$ - 19515667532728044661788 8738970 $x^3\xi^8$ - 24847107426709255943484 8457510 $x^4\xi^8$ + 13083103872462977884215 019554492 $x^5\xi^8$ - 49071201503618589151810 860329640 $x^6\xi^8$ + 44876009242554338181103 408360920 $x^7\xi^8$ + 72961294997833345788792 874766340 $x^8\xi^8$ - 15852705462834937851963 7491849020 $x^9\xi^8$ + 74024545904866816892142 596859824 $x^{10}\xi^8$ + 87329701278457185324 81110708496 $x^{11}\xi^8$ - 37252701723726271984154 39773660 $x^{12}\xi^8$ - 66540633638063877648 1155793580 $x^{13}\xi^8$ + 49779811308679164501 207600720 $x^{14}\xi^8$ + 99559622617358329002 41520144 $x^{15}\xi^8$ - 63222841977319114801 29714786 ξ^{10} + 23811068688981025617088 355790$x\xi^{10}$ + 67434947639572369660085 5203390 $x^2\xi^{10}$ - 40307841208886244536004 15212370 $x^3\xi^{10}$ - 37195755433941144402087 61411320 $x^4\xi^{10}$ + 69618417243516856573764 653634936 $x^5\xi^{10}$ - 15982445464923346542864 1309651380 $x^6\xi^{10}$ + 91948799033337891281434 484285820 $x^7\xi^{10}$ + 91893298890700035896165 670144570 $x^8\xi^{10}$ - 10766905313956394750150 8486017950 $x^9\xi^{10}$ + 16394590072795823456831 785110650 $x^{10}\xi^{10}$ + 29186634837490492750181 87393490 $x^{11}\xi^{10}$ - 22515908175632213521760 6017900 $x^{12}\xi^{10}$ - 45031816351264427043521 203580 $x^{13}\xi^{10}$ + 22412583894051240034914 095002 ξ^{12} - 10273785530809082489585 3978360$x\xi^{12}$ + 90276635585731382795740 84318884 $x^2\xi^{12}$ - 25541305658609626188640 420868200 $x^3\xi^{12}$ - 14386558308741781612899 98412650 $x^4\xi^{12}$ + 12132120010750090691630 0415372708 $x^5\xi^{12}$ - 18129908033149165389306 7232486820 $x^6\xi^{12}$ + 53265622454080043369740 96968759 6 $x^7\xi^{12}$ + 46032948972668967621414 613291950 $x^8\xi^{12}$ - 18916968152364568311797 563289850 $x^9\xi^{12}$ + 78616731411537887583430 1294850 $x^{10}\xi^{12}$ + 15723346282307577516686 0258970 $x^{11}\xi^{12}$ + 36078749865031134056986 5791056 ξ^{14} - 55966664633721524751600 52745128$x\xi^{14}$ + 27787270863557584809501 810705912 $x^2\xi^{14}$ - 50678573641634010525222 989395120 $x^3\xi^{14}$ + 45040703891953369915779 64577320 $x^4\xi^{14}$ + 82110525390855025726060 079248344 $x^5\xi^{14}$ - 69353831846842296024416 345356632 $x^6\xi^{14}$ + 44851175938554147499719 37902408 $x^7\xi^{14}$ + 50098228817296335022521 09096000 $x^8\xi^{14}$ - 67697403219595718740188 4924800 $x^9\xi^{14}$ + 97196832145991484635153 6213416 ξ^{16} - 94299431736753464634526 61694799$x\xi^{16}$ + 30259316417155987363625 807014449 $x^2\xi^{16}$ - 34791985535830528468266 741444325 $x^3\xi^{16}$ + 18727157183705027772269 7958315 $x^4\xi^{16}$ + 19955470635025218070583 559080964 $x^5\xi^{16}$ - 59169588799667922241 42796119536 $x^6\xi^{16}$ - 16582936267379480337015 715584 $x^7\xi^{16}$ + 61725684137569260582 805395000 $x^8\xi^{16}$ + 96212114733680516304 5785576402 ξ^{18} - 61075282423361845266134 83834738$x\xi^{18}$ + 12164497352096521345934 596394910 $x^2\xi^{18}$ - 73516881659869621586 51585230390 $x^3\xi^{18}$ - 10342203391970081392370 36005940 $x^4\xi^{18}$ + 85472587030955343845526 3228588 $x^5\xi^{18}$ - 56921890611922093007 847101232 $x^6\xi^{18}$ + 38607277104697484636960 6881810 ξ^{20} - 14912737669049308949223 16575295$x\xi^{20}$ + 15070724023309553911 62157111185 $x^2\xi^{20}$ - 20947374307114636563021 5101705 $x^3\xi^{20}$ - 21693382885093885956940 224275 $x^4\xi^{20}$ + 59070702889243936616 432386680 ξ^{22} - 10679057351156706572428 0857180$x\xi^{22}$ + 16682728710185112558340 817460 $x^2\xi^{22}$ + 25864284926781680457923 04420 ξ^{24}))/(27709612688932234336443 554842455 (1 - ξ^2)^{17})

A.6 . Formulas for Shuvaev's proposal

The operator S'^a resulting from the composition of the Shuvaev operator S^a and the simple Mellin transform acts on a PDF $a(x)$ in the following way [250]

$$\mathcal{S}'^a\left(\frac{\xi}{x},\frac{\mathbf{x'}}{\xi}\right)*a(\mathbf{x'}) = \int_{-1}^{1}\mathrm{d}x'\left[\frac{2}{\pi}\,\mathrm{Im}\int_0^1\mathrm{d}s\left(\frac{4s(1-s)}{(1+\frac{\xi}{x}(1-2s))^{1+p_a}}\sqrt{1-\frac{4s\frac{x'}{\xi}(1-s)}{\frac{x}{\xi}+1-2s}}\right)^{-1}\right]\frac{\mathrm{d}}{\mathrm{d}x'}\left(\frac{a(x')}{|x'|}\right), \tag{A.80}$$

where $p_a = 1$ for gluons ($a = g$) and 0 otherwise.

Support of integration

Let us determine exactly the values of x' and s such that the imaginary part in Eq. (A.80) is non zero. We consider the case where $x > \xi \geq 0$. Then the only source of imaginary part comes from

$$1-\frac{4sx'(1-s)}{x+\xi(1-2s)} \leq 0 \implies x' \geq \frac{x+\xi(1-2s)}{4s(1-s)}\,. \tag{A.81}$$

In particular, $x > \xi$ and $0 \leq s \leq 1$ imply that $x' > 0$. It is also necessary that

$$x+\xi(1-2s)-4sx'(1-s) \leq 0\,, \tag{A.82}$$

which yields the constraint

$$s_1 \leq s \leq s_2 \quad \text{with} \quad s_{1,2} = \frac{1}{2}+\frac{\xi}{4x'} \mp \frac{\sqrt{4x'^2+\xi^2-4xx'}}{4x'}\,, \tag{A.83}$$

where we used that $x' > 0$. Furthermore, the fact that $x > \xi$ implies that

$$\xi - 2x' < -\sqrt{4x'^2+\xi^2-4xx'}\,, \tag{A.84}$$

as seen by squaring the inequality. It gives that $s_2 < 1$. The fact that $x + \xi > 0$ gives similarly that $s_1 > 0$. Therefore, we never need to integrate down to $s = 0$ or up to $s = 1$ where non-integrable singularities would appear.

Finally, the imaginary part is not null only if $4x'^2 + \xi^2 - 4xx' > 0$ in Eq. (A.83), that is

$$\text{either} \quad x' < \frac{x}{2}-\frac{\sqrt{x^2-\xi^2}}{2} \quad \text{or} \quad x' > \frac{x}{2}+\frac{\sqrt{x^2-\xi^2}}{2}\,. \tag{A.85}$$

The first possibility implies

$$2x'-\xi < x-\xi-\sqrt{x^2-\xi^2} < x-\xi-\sqrt{(x-\xi)^2} = 0\,, \tag{A.86}$$

which is ruled out by $\xi - 2x' < -\sqrt{4x'^2+\xi^2-4xx'} < 0$ (A.84). To summarise, for $x > \xi \geq 0$,

$$\mathcal{S}'^a\left(\frac{\xi}{x},\frac{\mathbf{x'}}{\xi}\right)*a(\mathbf{x'}) = \int_{x/2+\sqrt{x^2-\xi^2}/2}^1 \mathrm{d}x'\,C^a\left(\frac{\xi}{x},\frac{x'}{\xi}\right)\,\frac{\mathrm{d}}{\mathrm{d}x'}\left(\frac{a(x')}{x'}\right), \tag{A.87}$$

where

$$C^a\left(\frac{\xi}{x},\frac{x'}{\xi}\right)=\frac{2}{\pi}\,\mathrm{Im}\int_{s_1}^{s_2}\mathrm{d}s\left(\frac{4s(1-s)}{(1+\frac{\xi}{x}(1-2s))^{1+p_a}}\sqrt{1-\frac{4s\frac{x'}{\xi}(1-s)}{\frac{x}{\xi}+1-2s}}\right)^{-1}. \tag{A.88}$$

The fact that the lower bound on the x' integration in Eq. (A.87) is $x/2+\sqrt{x^2-\xi^2}/2$, which is strictly less than x if $\xi>0$, is crucial to understand why **the relation between the DGLAP and general evolution operators expressed in Eq.** (7.47) **cannot be exact**.

Using the notations of Eq. (7.47), the DGLAP operator $1/x'\Gamma_0^{ab}(z/x';\mu^2,\mu_0^2)$ is non zero for $z\geq x'$. Since values of $x'\in[x/2+\sqrt{x^2-\xi^2}/2,x]$ give non vanishing contributions to the integral defining the operator S'^a (A.87), it means that DGLAP operators with $z\geq x/2+\sqrt{x^2-\xi^2}/2$ contribute to the approximation of the general evolution operator at $x>\xi$ in Eq. (7.47). We know however that the latter should only depend on values $z\geq x$. The approximation is therefore bound to behave poorly in the low z (or low y) region, which is precisely observed on Fig. 7.12.

We can check nonetheless that in the case where $\xi=0$ – where the lower bound on the x' integration in Eq. (A.87) is x – the operator S'^a reconstructs correctly the forward limit.

Case $\xi\to 0$

If the reconstruction procedure is sound, it should reproduce the PDF in the limit $\xi\to 0$ with our gluon GPD convention, that is $xS'^a(\xi/x,\mathbf{x'}/\xi)*a(\mathbf{x'})\to a(x)$. We verify this analytically.

$$\lim_{\xi\to 0}C^a\left(\frac{\xi}{x},\frac{x'}{\xi}\right)=\frac{2\sqrt{x}}{\pi}\,\mathrm{Im}\int_{s_1}^{s_2}\frac{\mathrm{d}s}{4s(1-s)\sqrt{x-4sx'(1-s)}}, \tag{A.89}$$

$$\text{where}\quad s_{1,2}=\frac{1}{2}\left(1\mp\sqrt{1-\frac{x}{x'}}\right). \tag{A.90}$$

The analytical computation of the integral in Eq. (A.89) gives $C^a(\xi/x,x'/\xi)\to -1$ when $\xi\to 0$. Hence

$$\lim_{\xi\to 0}xS'^a\left(\frac{\xi}{x},\frac{\mathbf{x'}}{\xi}\right)*a(\mathbf{x'})=-x\int_x^1\mathrm{d}x'\,\frac{\mathrm{d}}{\mathrm{d}x'}\left(\frac{a(x')}{x'}\right)=a(x), \tag{A.91}$$

if we take into account that $f(1)=0$ in the last equality.

Case $x=\xi$

We have so far only concerned ourselves with $x>\xi$ since it is clear that the integral is finite in that case as we never get to integrate up to the problematic boundaries $s=0$ and 1. The result can be extended without any change for $x=\xi$. It is sufficient to notice that the integral remains convergent although we are integrating up to $s=1$ where a potential divergence might occur.

$$C^a\left(1,\frac{x'}{x}\right)=\frac{2}{\pi}\,\mathrm{Im}\int_{x/2x'}^1\mathrm{d}s\left(\frac{4s(1-s)}{(2(1-s))^{1+p_a}}\sqrt{1-\frac{2sx'}{x}}\right)^{-1}. \tag{A.92}$$

Therefore, even though we integrate up to $s = 1$ if $x = \xi$, no divergence appears thanks to the cancellation of terms $(1 - s)$ in Eq. (A.92).

B - Résumé détaillé en Français

Les distributions de partons généralisées

Les théorèmes de factorisation constituent l'un des rares moyens pratiques pour caractériser la structure des hadrons à partir de données expérimentales. Ils permettent de décomposer les amplitudes de diffusion en termes de distributions de partons d'une part, qui sont des objets non-perturbatifs contenant des informations universelles sur la structure de la cible hadronique, et d'autre part en termes de fonctions coefficients dépendant du processus expérimental, calculables en théorie des perturbations. Les fonctions de distribution de partons habituelles (PDF), qui entrent notamment dans la description de la diffusion profondément inélastique (DIS), reçoivent une interprétation probabiliste en temps que densités de nombre de partons (quarks et gluons) en fonction de la fraction d'impulsion longitudinale x qu'ils portent dans un hadron en mouvement rapide. Les processus exclusifs – où toutes les particules impliquées dans l'état final de l'interaction sont mesurées – comme la diffusion Compton profondément virtuelle (DVCS) ou la production de mésons profondément virtuelle (DVMP) motivent l'introduction de distributions de partons de dimension supérieure, appelées distributions de partons généralisées (GPD). Contrairement aux PDF, les GPD peuvent être utilisées pour décrire les amplitudes de diffusion où un transfert d'impulsion vers la cible hadronique se produit. Par conséquent, outre la fraction d'impulsion longitudinale x, elles dépendent également du transfert d'impulsion longitudinal mesuré par ξ, et du transfert d'impulsion total mesuré par t. Les PDF constituent une limite des GPD quand aucune impulsion n'est transférée à la cible hadronique ($\xi = t = 0$). Les GPD généralisent également les facteurs de forme élastiques (EFF) qui sont obtenus en intégrant les GPD sur x.

En plus des trois variables x, ξ et t, et de manière similaire aux PDF, les GPD dépendent d'une échelle de renormalisation μ^2 qui découle de la nécessité d'éliminer la divergence UV des opérateurs de champ qui définissent formellement les GPDs. La dépendance des distributions de partons en fonction de l'échelle de renormalisation peut être exprimée en théorie des perturbations grâce à des équations intégro-différentielles connues sous le nom d'équations d'évolution.

En tant que généralisations des PDF, les GPD possèdent également une interprétation probabiliste. Elles encodent la corrélation entre la fraction d'impulsion longitudinale x et la position transversale dans le hadron par rapport au barycentre de l'impulsion longitudinale. Cette cartographie de la distribution radiale de l'impulsion longitudinale est connue sous le nom de tomographie hadronique et constitue l'une des principales motivations du programme d'étude des GPD. Une autre motivation physique cruciale est la possibilité remarquable d'exprimer les éléments de matrice du tenseur énergie-impulsion en termes de facteurs de forme gravitationnels (GFF) dérivés des GPD. Les GFF permettent à leur tour de définir des distributions d'énergie et de pression à l'intérieur de la matière hadronique. Ce sujet est au centre d'un intérêt théorique et expérimental considérable ces dernières années, et plusieurs extractions phénoménologiques basées sur divers ensembles de données DVCS disponibles avec différentes hypothèses de modélisation ont été publiées.

Le problème de déconvolution du DVCS

Bien que les GPD soient impliquées dans plusieurs processus expérimentaux différents, le DVCS a reçu la plus grande attention théorique et expérimentale ces dernières années. En effet, il présente à la fois l'avantage d'une section efficace significative – si nécessaire en tenant compte de son interférence avec le processus de Bethe-Heitler – et d'une description théorique relativement propre – comparée par exemple au DVMP qui implique une autre fonction non-perturbative appelée amplitude de distribution (DA). Un théorème de factorisation démontre que l'amplitude du DVCS peut être paramétrée par des facteurs de forme Compton (CFF) $\mathcal{F}$ calculés à partir des GPD F de la manière générique suivante

$$\mathcal{F}(\xi,t,Q^2) = \sum_a \int_{-1}^{1} \frac{\mathrm{d}x}{\xi} T^a\left(\frac{x}{\xi}, \frac{Q^2}{\mu^2}, \alpha_s(\mu^2)\right) \frac{F^a(x,\xi,t,\mu^2)}{x^{p_a}}\,, \tag{B.1}$$

où la somme s'étend sur les types de partons (avec $a = q$ pour les quarks et g pour les gluons), T^a est la fonction coefficient du DVCS calculable en théorie des perturbations, $p_a = 0$ si $a = q$ et 1 si $a = g$, α_s est le couplage fort et Q^2 est la virtualité du photon médiant l'interaction entre le lepton entrant et la cible hadronique. La question de savoir si la valeur de F^a peut être extraite de la connaissance expérimentale de $\mathcal{F}$ est appelée "problème de déconvolution" du DVCS et constitue un élément central de ce document. La question doit être affinée, car par exemple la parité de T^q implique déjà que seule la partie x-impaire de la GPD H^q, connue sous le nom de GPD singlet de quark et notée par

$$H^{q(+)}(x,\xi,t,\mu^2) = H^q(x,\xi,t,\mu^2) - H^q(-x,\xi,t,\mu^2)\,, \tag{B.2}$$

contribue à l'intégrale de l'équation (B.1). De plus, la séparation des contributions des différents types de partons est notoirement difficile à réaliser. Le problème de déconvolution est d'autant plus difficile que les GPD doivent suivre un certain nombre de contraintes théoriques. Notamment, leurs moments de Mellin en x doivent être polynomiaux en ξ en raison de la covariance de Lorentz :

$$\int_{-1}^{1} \mathrm{d}x\, x^n H^q(x,\xi,t,\mu^2) = \sum_{\substack{k=0 \\ \text{even}}}^{n+1} H^q_{n,k}(t,\mu^2)\xi^k\,. \tag{B.3}$$

Il est en fait équivalent pour H^q de satisfaire l'équation (B.3) et de s'écrire sous la forme d'une transformée intégrale connue sous le nom de transformée de Radon

$$H^q(x,\xi,t,\mu^2) = \int \mathrm{d}\alpha \mathrm{d}\beta\, \delta(x-\beta-\alpha\xi)\Big[f^q(\beta,\alpha,t,\mu^2) + \xi\delta(\beta)D^q(\alpha,t,\mu^2)\Big]\,, \tag{B.4}$$

où le support de $f^q(\beta,\alpha)$ est restreint à $|\alpha| + |\beta| \leq 1$ et celui de $D^q(\alpha)$ à $\alpha \in [-1,1]$. f^q est appelée une double distribution (DD) et D^q la contribution de q au D-terme. Bien que la dépendance en x et ξ des GPD soit intriquée par la condition de l'équation (B.3), la dépendance en α et β des DD n'est pas contrainte par cette relation, sauf pour la propriété de support.

Une étude des propriétés analytiques des amplitudes de diffusion avec quelques hypothèses sur le comportement de $\mathrm{Im}\,\mathcal{H}(\xi,t,Q^2)$ pour $\xi \to 0$ permet de dériver des relations de dispersion qui relient les parties réelles et imaginaires des CFF :

$$\mathrm{Re}\,\mathcal{H}(\xi,t,Q^2) = \frac{1}{\pi}\mathrm{p.v.}\int_0^1 \mathrm{d}\xi'\,\mathrm{Im}\,\mathcal{H}(\xi',t,Q^2)\left(\frac{1}{\xi-\xi'} - \frac{1}{\xi+\xi'}\right) + C_H(t,Q^2)\,, \tag{B.5}$$

où C_H, indépendant de ξ, est appelé constante de soustraction, et est lié au D-terme par

$$C_H(t,Q^2) = \frac{2}{\pi}\int_1^\infty \sum_a \frac{\mathrm{d}\omega}{\omega^{p_a}} \operatorname{Im} T^a\left(\omega, \frac{Q^2}{\mu^2}, \alpha_s(\mu^2)\right) \int_{-1}^1 \mathrm{d}z\, \frac{D^a(z,t,\mu^2)}{\omega - z}\,. \tag{B.6}$$

En fin de compte, l'information sur les GPD contenue dans les données expérimentales DVCS est exactement constituée de la partie imaginaire de l'équation (B.1) qui contraint uniquement la partie β-impaire de la DD – désignée par $f^{q(+)}$ –, et de l'équation (B.6) qui contraint uniquement le D-terme.

À l'ordre dominant (LO) en α_s, la partie imaginaire de l'équation (B.1) se lit comme suit

$$\operatorname{Im}\mathcal{H}(\xi,t,Q^2) \overset{LO}{=} \pi \sum_q e_q^2 H^{q(+)}(\xi,\xi,t,\mu^2)\,, \tag{B.7}$$

$$= \pi \sum_q e_q^2 \int_{(\xi-1)/(1+\xi)}^1 \mathrm{d}\alpha\, f^{q(+)}(\xi(1-\alpha),\alpha,t,\mu^2)\,, \tag{B.8}$$

et l'équation (B.6) donne

$$C^q(t,Q^2) \overset{LO}{=} 2\sum_q e_q^2 \int_{-1}^1 \mathrm{d}z\, \frac{D^q(z,t,\mu^2)}{1-z}\,, \tag{B.9}$$

où il est habituel de choisir μ^2 proche de Q^2. Dans ce qui suit, nous supposerons que $\mu^2 = Q^2$, sauf indication contraire explicite.

Puisque les deux équations (B.8) et (B.9) impliquent des intégrales de la DD et du D-terme où une variable est intégrée, il semble au premier abord improbable que même une connaissance parfaite des données expérimentales DVCS sur une large gamme cinématique soit suffisante pour retrouver effectivement la partie β-impaire de la DD et le D-terme. Cependant, la connaissance de la dépendance de ces deux objets selon μ^2 grâce aux équations d'évolution fournit l'argument théorique qui démontre la possibilité d'effectuer cette extraction.

Quantifier l'incertitude du problème de déconvolution pour la constante de soustraction

L'évolution LO du D-terme est exprimée de manière pratique grâce à une expansion en termes de polynômes de Gegenbauer :

$$D^q(z,t,\mu^2) = (1-z^2)\sum_{\text{odd } n}^\infty d_n^q(t,\mu^2) C_n^{(3/2)}(z)\,, \tag{B.10}$$

$$D^g(z,t,\mu^2) = \frac{3}{2}(1-z^2)^2 \sum_{\text{odd } n}^\infty d_n^g(t,\mu^2) C_{n-1}^{(5/2)}(z)\,. \tag{B.11}$$

Ensuite, l'équation (B.9) donne

$$C_H(t,Q^2) \overset{LO}{=} 4\sum_q e_q^2 \sum_{\text{odd } n}^\infty d_n^q(t,\mu^2)\,. \tag{B.12}$$

L'intérêt de cette représentation est qu'il est possible d'écrire les équations d'évolution LO sous la forme suivante

$$d_n^q(t,\mu^2) \overset{LO}{=} \Gamma_n^{qq}(\mu^2,\mu_0^2) d_n^q(t,\mu_0^2) + \Gamma_n^{qg}(\mu^2,\mu_0^2) d_n^g(t,\mu_0^2)\,, \tag{B.13}$$

$$d_n^g(t,\mu^2) \overset{LO}{=} \Gamma_n^{gq}(\mu^2,\mu_0^2) d_n^q(t,\mu_0^2) + \Gamma_n^{gg}(\mu^2,\mu_0^2) d_n^g(t,\mu_0^2)\,. \tag{B.14}$$

Les opérateurs d'évolution Γ_n^{ab} dictent le comportement de d_n^a en tant que fonctions de μ^2. Si les d_n^a forment une famille de fonctions de μ^2 linéairement indépendantes, alors la mesure des données expérimentales DVCS sur une gamme de valeurs variées en Q^2 devrait donner en principe le pouvoir de discriminer chaque d_n^q et de reconstruire ainsi le terme $D^q(\alpha,t,\mu^2)$ complet à partir de la simple connaissance de $C_H(t,Q^2)$.

L'évaluation de la possibilité pratique d'effectuer une telle extraction, à la fois en ce qui concerne l'ensemble actuel de données DVCS et l'impact attendu du collisionneur électron-ion (EIC), est l'une des contributions originales de ce document. Une extraction du D-terme indépendante de modèle nécessiterait d'obtenir d_n^a pour toutes les valeurs de $n \in \{1,3,5,...\}$ directement à partir des données expérimentales. Par souci de simplification, nous ne permettons d'abord que l'extraction de d_1, puis de d_1 et d_3 conjointement. En étudiant l'effet de l'ajout d'un nouveau paramètre libre dans la représentation du D-terme, nous pouvons quantifier la possibilité d'une extraction moins biaisée. Nous montrons que sous un certain nombre d'hypothèses de modélisation qui sont couramment utilisées dans les extractions phénoménologiques du D-terme, il est possible d'estimer l'incertitude sur d_3^q par

$$\Delta d_3^q \approx \frac{3}{4}\sigma \left(1 - \frac{\Gamma_3^{qq}(\mu^2_{\max},\mu^2_{\min})}{\Gamma_1^{qq}(\mu^2_{\max},\mu^2_{\min})}\right)^{-1}\,, \tag{B.15}$$

où $[\mu^2_{\min},\mu^2_{\max}]$ représente la plage disponible en Q^2 où les données précises de la constante de soustraction DVCS sont collectées, et σ leur incertitude typique. L'équation (B.15) met en évidence que l'incertitude liée à l'extraction du D-terme provient de deux sources principales : l'incertitude sur les données expérimentales σ, et l'incertitude due au bras de levier limité en Q^2 ici quantifiée par la similarité de l'évolution des différents d_n sur la plage disponible en Q^2. C'est en effet le fait que chaque d_n se comporte différemment sous l'effet de l'évolution qui fournit le levier théorique permettant de résoudre le problème de déconvolution.

Nous démontrons comment réinterpréter l'estimation de l'incertitude de la procédure de déconvolution (B.15) en introduisant ce que nous avons appelé des D-termes "fantômes". Ces objets n'apportent exactement aucune contribution à la constante de soustraction $C_H(t,Q^2)$ à l'ordre perturbatif choisi pour $Q^2 = \mu_0^2$, et leur contribution reste inférieure à la valeur de σ sur la plage disponible en Q^2. Nous montrons que la taille maximale des D-termes fantômes qui appartiennent à l'espace de modélisation autorisé pour le D-terme fournit une quantification similaire de l'incertitude de la procédure de déconvolution par rapport à (B.15).

Phénoménologie du DVCS

Nous montrons que notre estimation est capable de rendre compte de la phénoménologie actuelle des DVCS. Nous présentons d'abord en détail les outils de modélisation par réseaux de neurones que nous utilisons pour rendre compte de la propagation de l'incertitude dans les analyses complexes de données expérimentales. Nous discutons des stratégies

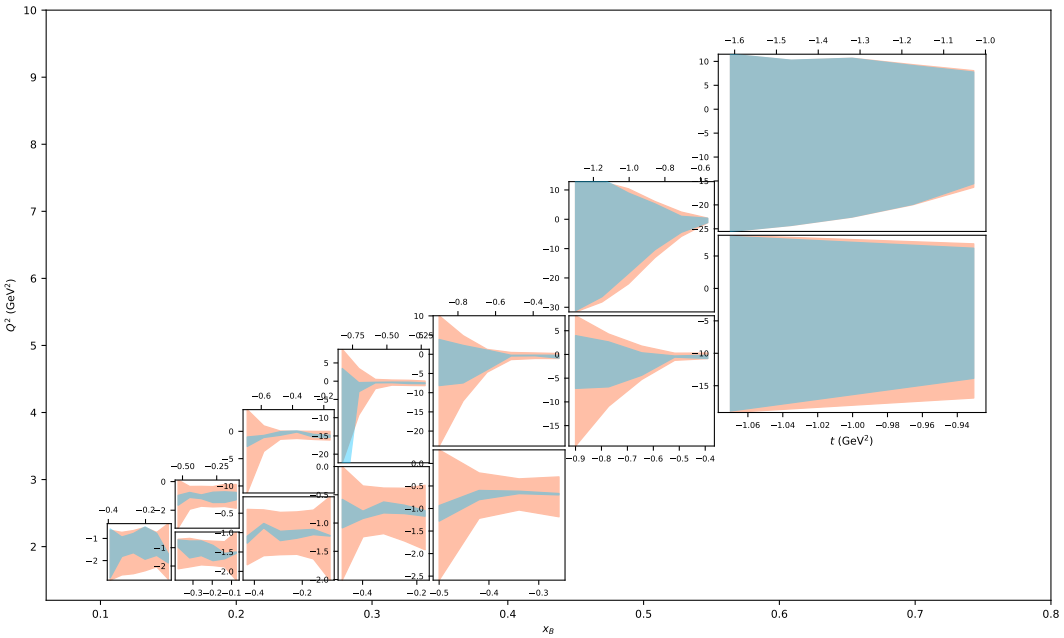

Figure B.1 – Régions de confiance à 68% pour $\mathrm{Re}\,\mathcal{H}$ avec l'ensemble actuel de données DVCS (bande orange) et l'incertitude réduite attendue grâce au programme positrons à JLab (bande bleue). Pour chacun des treize bins de (x_B, Q^2) $[\xi \approx x_B/(2 - x_B)]$, le sous-graphe montre les résultats sur les bins en t.

d'entraînement et de validation des réseaux de neurones et soulignons l'impact de la suppression des valeurs aberrantes. Nous comparons cette technique qui nécessite beaucoup de calculs avec des outils statistiques traditionnels tels que la matrice hessienne, et nous montrons dans les deux cas comment les fits peuvent prendre en compte de nouveaux points de données sans nécessiter un nouveau calcul fastidieux.

Nous utilisons notamment l'un de ces outils, appelé repondération bayésienne, pour étudier l'impact d'une future installation expérimentale potentielle sur l'incertitude expérimentale σ dans l'équation (B.15). Nous montrons comment la prise de données avec un faisceau de positrons au Thomas Jefferson National Accelerator Facility (JLab) pourrait permettre une réduction significative de l'incertitude sur $\mathrm{Re}\,\mathcal{H}$, qui joue un rôle crucial dans l'incertitude de la constante de soustraction, grâce à la mesure des asymétries de charge du faisceau. L'évaluation de la réduction attendue de l'incertitude sur $\mathrm{Re}\,\mathcal{H}$ selon les hypothèses détaillées dans le document est présentée sur la Fig. B.1.

Pour étudier l'incertitude de la procédure de déconvolution liée à l'effet de levier en Q^2 dans l'équation (B.15), nous menons une analyse numérique détaillée des opérateurs d'évolution Γ_n^{qq}. Cette étude permet de comprendre plusieurs caractéristiques intrigantes de notre extraction phénoménologique du D-terme : nous sommes notamment en mesure d'expliquer l'augmentation d'un facteur 20 de l'incertitude lors du fit conjoint de d_1 et d_3 par rapport à d_1 seul dans notre extraction phénoménologique du D-terme avec l'ensemble actuel de données DVCS, mais aussi le manque de sensibilité aux contributions des gluons lorsqu'elles sont générées radiativement à partir de celles des quark. Nous sommes en outre en mesure de prédire qu'avec une couverture étendue en Q^2 fournie par EIC, une réduction de l'incertitude d'un facteur $\sim$3 sur le fit conjoint de d_1 et d_3 peut être attendue du seul effet de la plage en Q^2, sans tenir compte d'un effet de précision accrue des mesures σ.

Nous présentons également les premiers résultats d'une étude phénoménologique de la constante de soustraction menée à l'ordre sous-dominant (NLO) sur l'ensemble actuel de données DVCS, et nous concluons que la plupart des caractéristiques identifiées à l'ordre dominant restent pertinentes.

Le problème de déconvolution pour les doubles distributions

Après avoir discuté en détail de l'incertitude de la procédure de déconvolution permettant de récupérer le D-terme à partir de la constante de soustraction, nous abordons la question de l'extraction de la DD de la partie imaginaire du CFF. Nous utilisons pour cela une extension de notre concept de D-termes fantômes, à savoir les DD fantômes. Ce sont des DD dont la contribution au CFF à l'ordre perturbatif choisi est exactement 0 à l'échelle $Q^2 = \mu_0^2$. Comme dans le cas des D-termes fantômes, la taille de la contribution au CFF des DDs fantômes évoluées sur la plage expérimentale en Q^2 disponible fournit une mesure de l'incertitude de la déconvolution.

Nous démontrons en détail l'existence de DD fantômes à NLO en utilisant des DD polynomiales en α et β. Nous montrons comment ce choix permet de construire une série de systèmes linéaires avec une forme commode, donnons explicitement les formules requises pour la solution itérative de ces systèmes, et fournissons des solutions explicites. Nous discutons également de certains avantages et inconvénients de l'utilisation de DD polynomiales pour la phénoménologie des GPD, et notamment le programme d'extension covariante.

Pour clarifier l'effet de l'évolution LO sur les DD fantômes, nous consacrons une présentation détaillée à la différence entre l'expansion de la fonction coefficient du DVCS en termes de logarithmes dominants d'une part et de puissances de α_s d'autre part. Bien que nous travaillions à NLO, et donc que les DDs fantômes annulent strictement les termes d'ordre α_s^0 et α_s^1 de la fonction coefficient du DVCS, la resommation de tous les termes d'ordre $\alpha_s^k \log^k(Q^2/\mu^2)$ pour $k \geq 2$ (expansion en logarithmes dominants de la fonction coefficient du DVCS) donne une contribution non nulle au CFF sous évolution LO. Nous fournissons donc le premier résultat à l'ordre α_s^1 avec évolution LO montrant que même sur un intervalle de $[1, 100]$ GeV^2, les trois GPD présentés sur la Fig. B.2 sont impossibles à distinguer avec les seules données DVCS, car leurs CFF sont indiscernables dans l'incertitude expérimentale typique. En pratique, les DD fantômes représentent des archétypes de modes particulièrement difficiles à contraindre dans le cadre d'une extraction non biaisée des GPD : à mesure que la flexibilité du modèle GPD augmente, et que de telles GPD fantômes entrent dans l'espace de modélisation choisi, l'incertitude de déconvolution explose.

Nouveaux modèles de distributions de partons généralisées

Dans l'optique d'une phénoménologie des GPD moins biaisée, les modèles avec réseaux de neurones, qui sont déjà utilisés pour représenter les CFF extraits de données expérimentales, apparaissent comme une possibilité naturelle. Nous démontrons qu'il est possible de satisfaire de nombreuses contraintes théoriques sur les GPD en utilisant une représentation des DD par réseau de neurones qui satisfait par conception la polynomialité des moments de Mellin (voir l'équation (B.3)), les symétries discrètes attendues, et reproduit en même temps la limite des PDF et un CFF LO donnés. Un exemple de résultat d'un tel

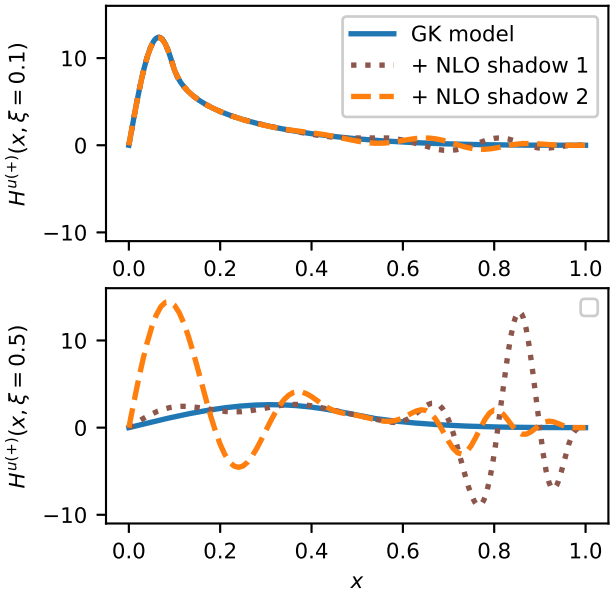

Figure B.2 – Trois GPD représentées en fonction de x pour $\xi = 0,1$ et $\xi = 0,5$ dont le CFF NLO à $\xi = 0,1$ diffère de moins de $\sim 10^{-5}$ bien que les GPD elles-mêmes diffèrent de $\sim$1.

réseau de neurones appliqué à un modèle phénoménologique populaire est présenté dans la Fig. B.3. Nous avons conçu une représentation spécifique des DD pour rendre compte de l'incertitude liée aux DD fantômes. Nous avons également accordé une attention particulière aux inégalités de positivité, qui apportent des contraintes fortes sur la forme des distributions dans la région des grands x. Nous démontrons notamment comment la procédure d'apprentissage permet l'implémentation de ces contraintes et soulignons leur impact sur l'incertitude du fit.

Nous observons que les GPDs fantômes n'apportent pas d'incertitude significative sur la région des petits ξ en raison d'une hypothèse implicite de régularité des DDs. Dans la limite des petits ξ et x, nous nous concentrons sur une proposition de modélisation populaire pour les GPDs basée sur la transformée de Shuvaev, qui permet d'approximer les GPDs à petit ξ entièrement à partir de leur PDF. Cette possibilité est particulièrement séduisante pour contraindre les PDFs de gluons à très petit x à partir de la mesure de la production de mésons vecteurs lourds dans les collisions ultra-périphériques. Nous montrons comment il est possible de réinterpréter cette proposition de modélisation grâce à un nouveau code d'évolution nommé APFEL++. Comme la validité de la procédure repose de manière cruciale sur l'idée que la région des grands x de la PDF à une certaine échelle initiale basse μ_0^2 contrôle la région des petits x et ξ à une grande échelle μ^2, nous proposons une modification de la méthode pour introduire une quantification de son incertitude systématique. Notre quantification de l'incertitude consiste en une mesure de la domination effective de la région des grands x à une certaine échelle initiale sur l'évolution de la GPD. Elle permet d'établir certains critères sur la taille de x et ξ, et sur la taille de l'échelle d'énergie impliquée dans le processus pour que la reconstruction de la GPD à partir de sa PDF apparaisse comme une procédure maîtrisée.

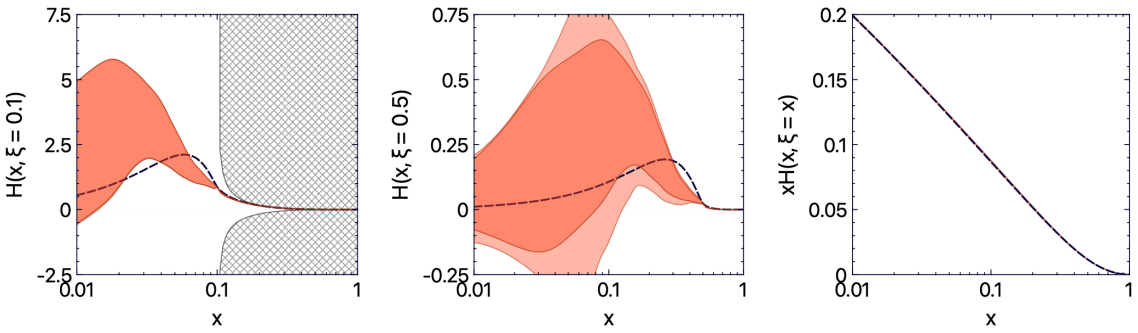

Figure B.3 – Démonstration de l'incertitude pour un modèle avec réseau de neurones entraîné à reproduire la PDF et le CFF LO d'un modèle phénoménologique (Goloskokov-Kroll) tout en respectant la polynomialité des moments de Mellin et une contrainte de positivité simplifiée. Comparaison entre le réseau de neurones (bande orange) et le modèle Goloskokov-Kroll (ligne noire) pour $\xi = 0,1$ (gauche), $\xi = 0,5$ (centre) et $\xi = x$ (droite). Les bandes hachurées représentent la zone exclue par la condition de positivité.

Bibliography

[1] V. Bertone et al. "Revisiting evolution equations for generalised parton distributions". In : (juin 2022). arXiv : 2206.01412 [hep-ph].

[2] H. Dutrieux et al. "Impact of a positron beam at JLab on an unbiased determination of DVCS Compton form factors". In : *Eur. Phys. J. A* 57.8 (2021), p. 250. DOI : 10.1140/epja/s10050-021-00560-2. arXiv : 2105.09245 [hep-ph].

[3] A. Accardi et al. "An experimental program with high duty-cycle polarized and unpolarized positron beams at Jefferson Lab". In : *Eur. Phys. J. A* 57.8 (2021), p. 261. DOI : 10.1140/epja/s10050-021-00564-y. arXiv : 2007.15081 [nucl-ex].

[4] H. Dutrieux et al. "Phenomenological assessment of proton mechanical properties from deeply virtual Compton scattering". In : *Eur. Phys. J. C* 81.4 (2021), p. 300. DOI : 10.1140/epjc/s10052-021-09069-w. arXiv : 2101.03855 [hep-ph].

[5] H. Dutrieux et al. "Proton internal pressure from deeply virtual Compton scattering on collider kinematics". In : (2022, in preparation).

[6] V. Bertone et al. "Deconvolution problem of deeply virtual Compton scattering". In : *Phys. Rev. D* 103.11 (2021), p. 114019. DOI : 10.1103/PhysRevD.103.114019. arXiv : 2104.03836 [hep-ph].

[7] V. Bertone et al. "Shadow generalized parton distributions : a practical approach to the deconvolution problem of DVCS". In : *SciPost Phys. Proc.* (8 2022), p. 107. DOI : 10.21468/SciPostPhysProc.8.107. URL : https://scipost.org/10.21468/SciPostPhysProc.8.107.

[8] H. Dutrieux et al. "Artificial neural network modelling of generalised parton distributions". In : *Eur. Phys. J. C* 82.3 (2022), p. 252. DOI : 10.1140/epjc/s10052-022-10211-5. arXiv : 2112.10528 [hep-ph].

[9] H. Dutrieux, M. Winn et V. Bertone. "When exclusive meets inclusive at low Bjorken-x_B : how to use exclusive measurements to constrain PDFs based on evolution equations". In : (2022, in preparation).

[10] D. Müller et al. "Wave functions, evolution equations and evolution kernels from light ray operators of QCD". In : *Fortsch. Phys.* 42 (1994), p. 101-141. DOI : 10.1002/prop.2190420202. arXiv : hep-ph/9812448.

[11] A. V. Radyushkin. "Scaling limit of deeply virtual Compton scattering". In : *Phys. Lett. B* 380 (1996), p. 417-425. DOI : 10.1016/0370-2693(96)00528-X. arXiv : hep-ph/9604317.

[12] X.-D. Ji. "Deeply virtual Compton scattering". In : *Phys. Rev.* D55 (1997), p. 7114-7125. DOI : 10.1103/PhysRevD.55.7114. arXiv : hep-ph/9609381 [hep-ph].

[13] A. Radyushkin. "Asymmetric gluon distributions and hard diffractive electroproduction". In : *Phys. Lett. B* 385 (1996), p. 333-342. DOI : 10.1016/0370-2693(96)00844-1. arXiv : hep-ph/9605431.

[14] J. C. Collins, L. Frankfurt et M. Strikman. "Factorization for hard exclusive electroproduction of mesons in QCD". In : *Phys. Rev. D* 56 (1997), p. 2982-3006. DOI : 10.1103/PhysRevD.56.2982. arXiv : hep-ph/9611433.

[15] A. V. Radyushkin. "Nonforward parton distributions". In : *Phys. Rev. D* 56 (1997), p. 5524-5557. DOI : 10.1103/PhysRevD.56.5524. arXiv : hep-ph/9704207.

[16] J. C. Collins et A. Freund. "Proof of factorization for deeply virtual Compton scattering in QCD". In : *Phys. Rev. D* 59 (1999), p. 074009. DOI : 10.1103/PhysRevD.59.074009. arXiv : hep-ph/9801262.

[17] X.-D. Ji et J. Osborne. "One loop corrections and all order factorization in deeply virtual Compton scattering". In : *Phys. Rev. D* 58 (1998), p. 094018. DOI : 10.1103/PhysRevD.58.094018. arXiv : hep-ph/9801260.

[18] J. C. Collins, D. E. Soper et G. F. Sterman. "Factorization of Hard Processes in QCD". In : *Adv. Ser. Direct. High Energy Phys.* 5 (1989), p. 1-91. DOI : 10.1142/9789814503266_0001. arXiv : hep-ph/0409313.

[19] C. Mezrag. "An introductory lecture on Generalised Parton Distributions". In : (juill. 2022). arXiv : 2207.13584 [hep-ph].

[20] R. L. Jaffe. "Spin, twist and hadron structure in deep inelastic processes". In : *Ettore Majorana International School of Nucleon Structure : 1st Course : The Spin Structure of the Nucleon*. Jan. 1996, p. 42-129. arXiv : hep-ph/9602236.

[21] A. V. Belitsky et A. V. Radyushkin. "Unraveling hadron structure with generalized parton distributions". In : *Phys. Rept.* 418 (2005), p. 1-387. DOI : 10.1016/j.physrep.2005.06.002. arXiv : hep-ph/0504030.

[22] H. D. Politzer. "Reliable Perturbative Results for Strong Interactions ?" In : *Phys. Rev. Lett.* 30 (1973). Sous la dir. de J. C. Taylor, p. 1346-1349. DOI : 10.1103/PhysRevLett.30.1346.

[23] D. J. Gross et F. Wilczek. "Ultraviolet Behavior of Nonabelian Gauge Theories". In : *Phys. Rev. Lett.* 30 (1973). Sous la dir. de J. C. Taylor, p. 1343-1346. DOI : 10.1103/PhysRevLett.30.1343.

[24] D. J. Gross et F. Wilczek. "Asymptotically Free Gauge Theories - I". In : *Phys. Rev. D* 8 (1973), p. 3633-3652. DOI : 10.1103/PhysRevD.8.3633.

[25] D. Amati, R. Petronzio et G. Veneziano. "Relating Hard QCD Processes Through Universality of Mass Singularities". In : *Nucl. Phys. B* 140 (1978), p. 54-72. DOI : 10.1016/0550-3213(78)90313-9.

[26] E. R. Berger, M. Diehl et B. Pire. "Time - like Compton scattering : Exclusive photoproduction of lepton pairs". In : *Eur. Phys. J. C* 23 (2002), p. 675-689. DOI : 10.1007/s100520200917. arXiv : hep-ph/0110062.

[27] M. Guidal et M. Vanderhaeghen. "Double deeply virtual Compton scattering off the nucleon". In : *Phys. Rev. Lett.* 90 (2003), p. 012001. DOI : 10.1103/PhysRevLett.90.012001. arXiv : hep-ph/0208275.

[28] A. V. Belitsky et D. Mueller. "Exclusive electroproduction of lepton pairs as a probe of nucleon structure". In : *Phys. Rev. Lett.* 90 (2003), p. 022001. DOI : 10.1103/PhysRevLett.90.022001. arXiv : hep-ph/0210313.

[29] B. Pire, L. Szymanowski et J. Wagner. "NLO corrections to timelike, spacelike and double deeply virtual Compton scattering". In : *Phys. Rev. D* 83 (2011), p. 034009. DOI : 10.1103/PhysRevD.83.034009. arXiv : 1101.0555 [hep-ph].

[30] A. Pedrak et al. "Electroproduction of a large invariant mass photon pair". In : *Phys. Rev. D* 101.11 (2020), p. 114027. DOI : 10.1103/PhysRevD.101.114027. arXiv : 2003.03263 [hep-ph].

[31] O. Grocholski et al. "Collinear factorization of diphoton photoproduction at next to leading order". In : *Phys. Rev. D* 104.11 (2021), p. 114006. DOI : 10.1103/PhysRevD.104.114006. arXiv : 2110.00048 [hep-ph].

[32] O. Grocholski et al. "Phenomenology of diphoton photoproduction at next-to-leading order". In : *Phys. Rev. D* 105.9 (2022), p. 094025. DOI : 10.1103/PhysRevD.105.094025. arXiv : 2204.00396 [hep-ph].

[33] R. Boussarie et al. "Exclusive photoproduction of a $\gamma\,\rho$ pair with a large invariant mass". In : *JHEP* 02 (2017). [Erratum : JHEP 10, 029 (2018)], p. 054. DOI : 10.1007/JHEP02(2017)054. arXiv : 1609.03830 [hep-ph].

[34] G. Duplančić et al. "Probing axial quark generalized parton distributions through exclusive photoproduction of a $\gamma\,\pi^{\pm}$ pair with a large invariant mass". In : *JHEP* 11 (2018), p. 179. DOI : 10.1007/JHEP11(2018)179. arXiv : 1809.08104 [hep-ph].

[35] M. Diehl. "Generalized parton distributions". In : *Phys. Rept.* 388 (2003), p. 41-277. DOI : 10.1016/j.physrep.2003.08.002,10.3204/DESY-THESIS-2003-018. arXiv : hep-ph/0307382 [hep-ph].

[36] D. Y. Ivanov et al. "Probing chiral odd GPD's in diffractive electroproduction of two vector mesons". In : *Phys. Lett. B* 550 (2002), p. 65-76. DOI : 10.1016/S0370-2693(02)02856-3. arXiv : hep-ph/0209300.

[37] M. Diehl et T. Gousset. "Time ordering in off diagonal parton distributions". In : *Phys. Lett. B* 428 (1998), p. 359-370. DOI : 10.1016/S0370-2693(98)00439-0. arXiv : hep-ph/9801233.

[38] M. Diehl et al. "Probing partonic structure in gamma* gamma —> pi pi near threshold". In : *Phys. Rev. Lett.* 81 (1998), p. 1782-1785. DOI : 10.1103/PhysRevLett.81.1782. arXiv : hep-ph/9805380.

[39] M. V. Polyakov. "Hard exclusive electroproduction of two pions and their resonances". In : *Nucl. Phys. B* 555 (1999), p. 231. DOI : 10.1016/S0550-3213(99)00314-4. arXiv : hep-ph/9809483.

[40] N. Kivel, L. Mankiewicz et M. V. Polyakov. "NLO corrections and contribution of a tensor gluon operator to the process gamma* gamma —> pi pi". In : *Phys. Lett. B* 467 (1999), p. 263-270. DOI : 10.1016/S0370-2693(99)01148-X. arXiv : hep-ph/9908334.

[41] M. Diehl, T. Gousset et B. Pire. "Exclusive production of pion pairs in gamma* gamma collisions at large Q**2". In : *Phys. Rev. D* 62 (2000), p. 073014. DOI : 10.1103/PhysRevD.62.073014. arXiv : hep-ph/0003233.

[42] S. Meissner et al. "Generalized parton correlation functions for a spin-0 hadron". In : *JHEP* 08 (2008), p. 038. DOI : 10.1088/1126-6708/2008/08/038. arXiv : 0805.3165 [hep-ph].

[43] S. Meissner, A. Metz et M. Schlegel. "Generalized parton correlation functions for a spin-1/2 hadron". In : *JHEP* 08 (2009), p. 056. DOI : 10.1088/1126-6708/2009/08/056. arXiv : 0906.5323 [hep-ph].

[44] N. Chouika. "Generalized Parton Distributions and their covariant extension : towards nucleon tomography". Thèse de doct. IRFU, Saclay, DPHN, 2018.

[45] C. Lorce, B. Pasquini et M. Vanderhaeghen. "Unified framework for generalized and transverse-momentum dependent parton distributions within a 3Q light-cone picture of the nucleon". In : *JHEP* 05 (2011), p. 041. DOI : 10.1007/JHEP05(2011)041. arXiv : 1102.4704 [hep-ph].

[46] A. Accardi et al. "Electron Ion Collider : The Next QCD Frontier". In : *Eur. Phys. J.* A52.9 (2016). Sous la dir. d'A. Deshpande, Z. E. Meziani et J. W. Qiu, p. 268. DOI : 10.1140/epja/i2016-16268-9. arXiv : 1212.1701 [nucl-ex].

[47] R. Abdul Khalek et al. *Science Requirements and Detector Concepts for the Electron-Ion Collider : EIC Yellow Report*. Mars 2021. arXiv : 2103.05419 [physics.ins-det].

[48] X. Chen. "A Plan for Electron Ion Collider in China". In : *PoS* DIS2018 (2018), p. 170. DOI : 10.22323/1.316.0170. arXiv : 1809.00448 [nucl-ex].

[49] D. P. Anderle et al. "Electron-ion collider in China". In : *Front. Phys. (Beijing)* 16.6 (2021), p. 64701. DOI : 10.1007/s11467-021-1062-0. arXiv : 2102.09222 [nucl-ex].

[50] J. L. Abelleira Fernandez et al. "A Large Hadron Electron Collider at CERN : Report on the Physics and Design Concepts for Machine and Detector". In : *J. Phys.* G39 (2012), p. 075001. DOI : 10.1088/0954-3899/39/7/075001. arXiv : 1206.2913 [physics.acc-ph].

[51] M. Burkardt. "Impact parameter dependent parton distributions and off forward parton distributions for zeta —> 0". In : *Phys. Rev. D* 62 (2000). [Erratum : Phys.Rev.D 66, 119903 (2002)], p. 071503. DOI : 10.1103/PhysRevD.62.071503. arXiv : hep-ph/0005108.

[52] M. Diehl. "Generalized parton distributions in impact parameter space". In : *Eur. Phys. J. C* 25 (2002). [Erratum : Eur.Phys.J.C 31, 277–278 (2003)], p. 223-232. DOI : 10.1007/s10052-002-1016-9. arXiv : hep-ph/0205208.

[53] H. Moutarde, P. Sznajder et J. Wagner. "Border and skewness functions from a leading order fit to DVCS data". In : *Eur. Phys. J.* C78.11 (2018), p. 890. DOI : 10.1140/epjc/s10052-018-6359-y. arXiv : 1807.07620 [hep-ph].

[54] J. Ashman et al. "A Measurement of the Spin Asymmetry and Determination of the Structure Function g(1) in Deep Inelastic Muon-Proton Scattering". In : *Phys. Lett. B* 206 (1988). Sous la dir. de V. W. Hughes et C. Cavata, p. 364. DOI : 10.1016/0370-2693(88)91523-7.

[55] K. Goeke et al. "Nucleon form-factors of the energy momentum tensor in the chiral quark-soliton model". In : *Phys. Rev.* D75 (2007), p. 094021. DOI : 10.1103/PhysRevD.75.094021. arXiv : hep-ph/0702030 [hep-ph].

[56] M. V. Polyakov et P. Schweitzer. "Forces inside hadrons : pressure, surface tension, mechanical radius, and all that". In : *Int. J. Mod. Phys.* A33.26 (2018), p. 1830025. DOI : 10.1142/S0217751X18300259. arXiv : 1805.06596 [hep-ph].

[57] M. V. Polyakov. "Generalized parton distributions and strong forces inside nucleons and nuclei". In : *Phys. Lett.* B555 (2003), p. 57-62. DOI : 10.1016/S0370-2693(03)00036-4. arXiv : hep-ph/0210165 [hep-ph].

[58] B. L. G. Bakker, E. Leader et T. L. Trueman. "A Critique of the angular momentum sum rules and a new angular momentum sum rule". In : *Phys. Rev. D* 70 (2004), p. 114001. DOI : 10.1103/PhysRevD.70.114001. arXiv : hep-ph/0406139.

[59] E. Leader et C. Lorcé. "The angular momentum controversy : What's it all about and does it matter ?" In : *Phys. Rept.* 541.3 (2014), p. 163-248. DOI : 10.1016/j.physrep.2014.02.010. arXiv : 1309.4235 [hep-ph].

[60] C. Lorcé. "The light-front gauge-invariant energy-momentum tensor". In : *JHEP* 08 (2015), p. 045. DOI : 10.1007/JHEP08(2015)045. arXiv : 1502.06656 [hep-ph].

[61] E. Leader. "A critical assessment of the angular momentum sum rules". In : *Phys. Lett. B* 720 (2013). [Erratum : Phys.Lett.B 726, 927–927 (2013)], p. 120-124. DOI : 10.1016/j.physletb.2013.01.050. arXiv : 1211.3957 [hep-ph].

[62] K. Tanaka. "Operator relations for gravitational form factors of a spin-0 hadron". In : *Phys. Rev. D* 98.3 (2018), p. 034009. DOI : 10.1103/PhysRevD.98.034009. arXiv : 1806.10591 [hep-ph].

[63] X.-D. Ji. "Gauge-Invariant Decomposition of Nucleon Spin". In : *Phys. Rev. Lett.* 78 (1997), p. 610-613. DOI : 10.1103/PhysRevLett.78.610. arXiv : hep-ph/9603249.

[64] C. Lorcé, H. Moutarde et A. P. Trawiński. "Revisiting the mechanical properties of the nucleon". In : *Eur. Phys. J. C* 79.1 (2019), p. 89. DOI : 10.1140/epjc/s10052-019-6572-3. arXiv : 1810.09837 [hep-ph].

[65] F. Yuan. "Generalized parton distributions at x —> 1". In : *Phys. Rev. D* 69 (2004), p. 051501. DOI : 10.1103/PhysRevD.69.051501. arXiv : hep-ph/0311288.

[66] A. H. Mueller. "Perturbative QCD at High-Energies". In : *Phys. Rept.* 73 (1981), p. 237. DOI : 10.1016/0370-1573(81)90030-2.

[67] S. Boffi et B. Pasquini. "Generalized parton distributions and the structure of the nucleon". In : *Riv. Nuovo Cim.* 30 (2007), p. 387. DOI : 10.1393/ncr/i2007-10025-7. arXiv : 0711.2625 [hep-ph].

[68] W. Gordon. "Der Strom der Diracschen Elektronentheorie". In : *Zeitschrift für Physik* 50 (1928), p. 630-632. DOI : /10.1007/BF01327881.

[69] A. Hertle. "Continuity of the radon transform and its inverse on Euclidean space". In : *Mathematische Zeitschrift* 184 (1983), p. 165-192. DOI : 10.1007/BF01252856.

[70] J. Radon. "Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten". In : *Akad. Wiss.* 69 (1917), p. 262-277.

[71] J. Radon. "On the determination of functions from their integral values along certain manifolds". In : *IEEE Transactions on Medical Imaging* 5.4 (1986), p. 170-176. DOI : 10.1109/TMI.1986.4307775.

[72] F. Hausdorff. "Summationsmethoden und momentfolgen". In : *Mathematische Zeitschrift* 9 (1921), 74—109.

[73] O. Teryaev. "Crossing and radon tomography for generalized parton distributions". In : *Phys.Lett.* B510 (2001), p. 125-132. DOI : 10.1016/S0370-2693(01)00564-0. arXiv : hep-ph/0102303 [hep-ph].

[74] M. V. Polyakov et C. Weiss. "Skewed and double distributions in pion and nucleon". In : *Phys. Rev. D* 60 (1999), p. 114017. DOI : 10.1103/PhysRevD.60.114017. arXiv : hep-ph/9902451.

[75] B. C. Tiburzi. "Double distributions : Loose ends". In : *Phys. Rev. D* 70 (2004), p. 057504. DOI : 10.1103/PhysRevD.70.057504. arXiv : hep-ph/0405211.

[76] P. V. Pobylitsa. "Integral representations for nonperturbative GPDs in terms of perturbative diagrams". In : *Phys. Rev. D* 67 (2003), p. 094012. DOI : 10.1103/PhysRevD.67.094012. arXiv : hep-ph/0210238.

[77] N. Chouika et al. "Covariant Extension of the GPD overlap representation at low Fock states". In : *Eur. Phys. J. C* 77.12 (2017), p. 906. DOI : 10.1140/epjc/s10052-017-5465-6. arXiv : 1711.05108 [hep-ph].

[78] S. J. Brodsky, H.-C. Pauli et S. S. Pinsky. "Quantum chromodynamics and other field theories on the light cone". In : *Phys. Rept.* 301 (1998), p. 299-486. DOI : 10.1016/S0370-1573(97)00089-6. arXiv : hep-ph/9705477.

[79] G. P. Lepage et S. J. Brodsky. "Exclusive Processes in Perturbative Quantum Chromodynamics". In : *Phys. Rev. D* 22 (1980), p. 2157. DOI : 10.1103/PhysRevD.22.2157.

[80] M. Diehl et al. "The overlap representation of skewed quark and gluon distributions". In : *Nucl. Phys. B* 596 (2001). [Erratum : Nucl.Phys.B 605, 647–647 (2001)], p. 33-65. DOI : 10.1016/S0550-3213(00)00684-2. arXiv : hep-ph/0009255.

[81] C. Mezrag, H. Moutarde et J. Rodriguez-Quintero. "From Bethe–Salpeter Wave functions to Generalised Parton Distributions". In : *Few Body Syst.* 57.9 (2016), p. 729-772. DOI : 10.1007/s00601-016-1119-8. arXiv : 1602.07722 [nucl-th].

[82] N. Chouika et al. "A Nakanishi-based model illustrating the covariant extension of the pion GPD overlap representation and its ambiguities". In : *Phys. Lett. B* 780 (2018), p. 287-293. DOI : 10.1016/j.physletb.2018.02.070. arXiv : 1711.11548 [hep-ph].

[83] M. Diehl et al. "Linking parton distributions to form-factors and Compton scattering". In : *Eur. Phys. J. C* 8 (1999), p. 409-434. DOI : 10.1007/s100529901100. arXiv : hep-ph/9811253.

[84] B. Pire, J. Soffer et O. Teryaev. "Positivity constraints for off - forward parton distributions". In : *Eur. Phys. J. C* 8 (1999), p. 103-106. DOI : 10.1007/s100529901063. arXiv : hep-ph/9804284.

[85] A. V. Radyushkin. "Double distributions and evolution equations". In : *Phys. Rev. D* 59 (1999), p. 014030. DOI : 10.1103/PhysRevD.59.014030. arXiv : hep-ph/9805342.

[86] P. V. Pobylitsa. "Inequalities for generalized parton distributions H and E". In : *Phys. Rev. D* 65 (2002), p. 077504. DOI : 10.1103/PhysRevD.65.077504. arXiv : hep-ph/0112322.

[87] P. V. Pobylitsa. "Disentangling positivity constraints for generalized parton distributions". In : *Phys. Rev. D* 65 (2002), p. 114015. DOI : 10.1103/PhysRevD.65.114015. arXiv : hep-ph/0201030.

[88] P. V. Pobylitsa. "Positivity bounds on generalized parton distributions in impact parameter representation". In : *Phys. Rev. D* 66 (2002), p. 094002. DOI : 10.1103/PhysRevD.66.094002. arXiv : hep-ph/0204337.

[89] P. Zyla et al. "Review of Particle Physics". In : *PTEP* 2020.8 (2020). and 2021 update, p. 083C01. DOI : 10.1093/ptep/ptaa104.

[90] Y. L. Dokshitzer. "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics." In : *Sov. Phys. JETP* 46 (1977), p. 641-653.

[91] V. N. Gribov et L. N. Lipatov. "Deep inelastic e p scattering in perturbation theory". In : *Sov. J. Nucl. Phys.* 15 (1972), p. 438-450.

[92] G. Altarelli et G. Parisi. "Asymptotic Freedom in Parton Language". In : *Nucl. Phys. B* 126 (1977), p. 298-318. DOI : 10.1016/0550-3213(77)90384-4.

[93] A. V. Efremov et A. V. Radyushkin. "Factorization and Asymptotical Behavior of Pion Form-Factor in QCD". In : *Phys. Lett. B* 94 (1980), p. 245-250. DOI : 10.1016/0370-2693(80)90869-2.

[94] G. P. Lepage et S. J. Brodsky. "Exclusive Processes in Quantum Chromodynamics : Evolution Equations for Hadronic Wave Functions and the Form-Factors of Mesons". In : *Phys. Lett. B* 87 (1979), p. 359-365. DOI : 10.1016/0370-2693(79)90554-9.

[95] G. Altarelli, S. Forte et G. Ridolfi. "On positivity of parton distributions". In : *Nucl. Phys. B* 534 (1998), p. 277-296. DOI : 10.1016/S0550-3213(98)00661-0. arXiv : hep-ph/9806345.

[96] J. Collins, T. C. Rogers et N. Sato. "Positivity and renormalization of parton densities". In : *Phys. Rev. D* 105.7 (2022), p. 076010. DOI : 10.1103/PhysRevD.105.076010. arXiv : 2111.01170 [hep-ph].

[97] P. V. Pobylitsa. "Virtual Compton scattering in the generalized Bjorken region and positivity bounds on generalized parton distributions". In : *Phys. Rev. D* 70 (2004), p. 034004. DOI : 10.1103/PhysRevD.70.034004. arXiv : hep-ph/0211160.

[98] J. Blumlein, B. Geyer et D. Robaschik. "On the evolution kernels of twist-2 light ray operators for unpolarized and polarized deep inelastic scattering". In : *Phys. Lett. B* 406 (1997), p. 161-170. DOI : 10.1016/S0370-2693(97)00680-1. arXiv : hep-ph/9705264.

[99] A. V. Belitsky et D. Mueller. "Next-to-leading order evolution of twist-2 conformal operators : The Abelian case". In : *Nucl. Phys. B* 527 (1998), p. 207-234. DOI : 10.1016/S0550-3213(98)00310-1. arXiv : hep-ph/9802411.

[100] A. V. Belitsky et D. Mueller. "Exclusive evolution kernels in two loop order : Parity even sector". In : *Phys. Lett. B* 464 (1999), p. 249-256. DOI : 10.1016/S0370-2693(99)01003-5. arXiv : hep-ph/9906409.

[101] A. V. Belitsky, A. Freund et D. Mueller. "Evolution kernels of skewed parton distributions : Method and two loop results". In : *Nucl. Phys. B* 574 (2000), p. 347-406. DOI : 10.1016/S0550-3213(00)00012-2. arXiv : hep-ph/9912379.

[102] V. M. Braun et al. "Two-loop evolution equations for flavor-singlet light-ray operators". In : *JHEP* 02 (2019), p. 191. DOI : 10.1007/JHEP02(2019)191. arXiv : 1901.06172 [hep-ph].

[103] V. M. Braun et al. "Three-loop evolution equation for flavor-nonsinglet operators in off-forward kinematics". In : *JHEP* 06 (2017), p. 037. DOI : 10.1007/JHEP06(2017)037. arXiv : 1703.09532 [hep-ph].

[104] V. M. Braun, G. P. Korchemsky et D. Müller. "The Uses of conformal symmetry in QCD". In : *Prog. Part. Nucl. Phys.* 51 (2003), p. 311-398. DOI : 10.1016/S0146-6410(03)90004-4. arXiv : hep-ph/0306057.

[105] T. Ohrndorf. "Constraints From Conformal Covariance on the Mixing of Operators of Lowest Twist". In : *Nucl. Phys. B* 198 (1982), p. 26-44. DOI : 10.1016/0550-3213(82)90542-9.

[106] M. K. Chase. "THE Q**2 EVOLUTION OF FLAVOR SINGLET WAVE FUNCTIONS IN QCD". In : *Nucl. Phys. B* 174 (1980), p. 109-122. DOI : 10.1016/0550-3213(80)90192-3.

[107] V. N. Baier et A. G. Grozin. "Certain exclusive processes in QCD taking into account two-gluon states". In : *Sov. J. Nucl. Phys.* 35 (1982), p. 596-605.

[108] K. Kumericki, D. Mueller et K. Passek-Kumericki. "Towards a fitting procedure for deeply virtual Compton scattering at next-to-leading order and beyond". In : *Nucl. Phys. B* 794 (2008), p. 244-323. DOI : 10.1016/j.nuclphysb.2007.10.029. arXiv : hep-ph/0703179.

[109] K. Kumerički et D. Mueller. "Deeply virtual Compton scattering at small x_B and the access to the GPD H". In : *Nucl. Phys. B* 841 (2010), p. 1-58. DOI : 10.1016/j.nuclphysb.2010.07.015. arXiv : 0904.0458 [hep-ph].

[110] A. Shuvaev. "Solution of the off forward leading logarithmic evolution equation based on the Gegenbauer moments inversion". In : *Phys. Rev. D* 60 (1999), p. 116005. DOI : 10.1103/PhysRevD.60.116005. arXiv : hep-ph/9902318.

[111] J. D. Noritzsch. "Off forward parton distributions and Shuvaev's transformations". In : *Phys. Rev. D* 62 (2000), p. 054015. DOI : 10.1103/PhysRevD.62.054015. arXiv : hep-ph/0004012.

[112] D. Müller, M. V. Polyakov et K. M. Semenov-Tian-Shansky. "Dual parametrization of generalized parton distributions in two equivalent representations". In : *JHEP* 03 (2015), p. 052. DOI : 10.1007/JHEP03(2015)052. arXiv : 1412.4165 [hep-ph].

[113] A. V. Belitsky et al. "Two loop effects in the evolution of nonforward distributions". In : *Phys. Lett. B* 437 (1998), p. 160-168. DOI : 10.1016/S0370-2693(98)00874-0. arXiv : hep-ph/9806232.

[114] A. V. Belitsky et al. "On the leading logarithmic evolution of the off forward distributions". In : *Phys. Lett. B* 421 (1998), p. 312-318. DOI : 10.1016/S0370-2693(98)00007-0. arXiv : hep-ph/9710427.

[115] A. V. Vinnikov. "Code for prompt numerical computation of the leading order GPD evolution". In : (avr. 2006). arXiv : hep-ph/0604248.

[116] K. Kumericki, S. Liuti et H. Moutarde. "GPD phenomenology and DVCS fitting : Entering the high-precision era". In : *Eur. Phys. J. A* 52.6 (2016), p. 157. doi : 10.1140/epja/i2016-16157-3. arXiv : 1602.02763 [hep-ph].

[117] A. Bacchetta et al. "Single-spin asymmetries : The Trento conventions". In : *Phys. Rev. D* 70 (2004), p. 117504. doi : 10.1103/PhysRevD.70.117504. arXiv : hep-ph/0410050.

[118] D. Boer et al. "Gluons and the quark sea at high energies : Distributions, polarization, tomography". In : (août 2011). arXiv : 1108.1713 [nucl-th].

[119] A. V. Belitsky, D. Mueller et A. Kirchner. "Theory of deeply virtual Compton scattering on the nucleon". In : *Nucl. Phys. B* 629 (2002), p. 323-392. doi : 10.1016/S0550-3213(02)00144-X. arXiv : hep-ph/0112108.

[120] A. V. Belitsky, D. Müller et Y. Ji. "Compton scattering : from deeply virtual to quasi-real". In : *Nucl. Phys. B* 878 (2014), p. 214-268. doi : 10.1016/j.nuclphysb.2013.11.014. arXiv : 1212.6674 [hep-ph].

[121] L. Mankiewicz et al. "NLO corrections to deeply virtual Compton scattering". In : *Phys. Lett. B* 425 (1998). [Erratum : Phys.Lett.B 461, 423–423 (1999)], p. 186-192. doi : 10.1016/S0370-2693(98)00190-7. arXiv : hep-ph/9712251.

[122] A. V. Belitsky et al. "Deeply virtual Compton scattering in next-to-leading order". In : *Phys.Lett.* B474 (2000), p. 163-169. doi : 10.1016/S0370-2693(99)01283-6. arXiv : hep-ph/9908337 [hep-ph].

[123] A. Freund et M. McDermott. "A Detailed next-to-leading order QCD analysis of deeply virtual Compton scattering observables". In : *Eur. Phys. J. C* 23 (2002), p. 651-674. doi : 10.1007/s100520200928. arXiv : hep-ph/0111472.

[124] A. Freund et M. F. McDermott. "A Next-to-leading order QCD analysis of deeply virtual Compton scattering amplitudes". In : *Phys. Rev. D* 65 (2002), p. 074008. doi : 10.1103/PhysRevD.65.074008. arXiv : hep-ph/0106319.

[125] V. M. Braun et al. "Two-loop coefficient function for DVCS : vector contributions". In : *JHEP* 09 (2020), p. 117. doi : 10.1007/JHEP09(2020)117. arXiv : 2007.06348 [hep-ph].

[126] V. M. Braun, Y. Ji et J. Schoenleber. "Deeply-virtual Compton scattering at the next-to-next-to-leading order". In : (juill. 2022). arXiv : 2207.06818 [hep-ph].

[127] N. d'Hose, S. Niccolai et A. Rostomyan. "Experimental overview of Deeply Virtual Compton Scattering". In : *Eur. Phys. J. A* 52.6 (2016), p. 151. doi : 10.1140/epja/i2016-16151-9.

[128] M. Diehl et S. Sapeta. "On the analysis of lepton scattering on longitudinally or transversely polarized protons". In : *Eur. Phys. J. C* 41 (2005), p. 515-533. doi : 10.1140/epjc/s2005-02242-9. arXiv : hep-ph/0503023.

[129] P. Kroll, H. Moutarde et F. Sabatie. "From hard exclusive meson electroproduction to deeply virtual Compton scattering". In : *Eur. Phys. J. C* 73.1 (2013), p. 2278. doi : 10.1140/epjc/s10052-013-2278-0. arXiv : 1210.6975 [hep-ph].

[130] V. M. Braun, A. N. Manashov et B. Pirnay. "Finite-t and target mass corrections to DVCS on a scalar target". In : *Phys. Rev. D* 86 (2012), p. 014003. DOI : 10.1103/PhysRevD.86.014003. arXiv : 1205.3332 [hep-ph].

[131] V. M. Braun, A. N. Manashov et B. Pirnay. "Finite-t and target mass corrections to deeply virtual Compton scattering". In : *Phys. Rev. Lett.* 109 (2012), p. 242001. DOI : 10.1103/PhysRevLett.109.242001. arXiv : 1209.2559 [hep-ph].

[132] V. M. Braun et al. "Deeply Virtual Compton Scattering to the twist-four accuracy : Impact of finite-t and target mass corrections". In : *Phys. Rev. D* 89.7 (2014), p. 074022. DOI : 10.1103/PhysRevD.89.074022. arXiv : 1401.7621 [hep-ph].

[133] M. Defurne et al. "E00-110 experiment at Jefferson Lab Hall A : Deeply virtual Compton scattering off the proton at 6 GeV". In : *Phys. Rev. C* 92.5 (2015), p. 055202. DOI : 10.1103/PhysRevC.92.055202. arXiv : 1504.05453 [nucl-ex].

[134] M. Guidal. "A Fitter code for Deep Virtual Compton Scattering and Generalized Parton Distributions". In : *Eur. Phys. J. A* 37 (2008). [Erratum : Eur.Phys.J.A 40, 119 (2009)], p. 319-332. DOI : 10.1140/epja/i2008-10630-6. arXiv : 0807.2355 [hep-ph].

[135] H. Moutarde. "Extraction of the Compton Form Factor H from DVCS measurements at Jefferson Lab". In : *Phys. Rev. D* 79 (2009), p. 094021. DOI : 10.1103/PhysRevD.79.094021. arXiv : 0904.1648 [hep-ph].

[136] M. Guidal et H. Moutarde. "Generalized Parton Distributions from Deeply Virtual Compton Scattering at HERMES". In : *Eur. Phys. J. A* 42 (2009), p. 71-78. DOI : 10.1140/epja/i2009-10840-4. arXiv : 0905.1220 [hep-ph].

[137] K. Kumerički, D. Müller et M. Murray. "HERMES impact for the access of Compton form factors". In : *Phys. Part. Nucl.* 45.4 (2014), p. 723-755. DOI : 10.1134/S1063779614040108. arXiv : 1301.1230 [hep-ph].

[138] M. Boër et M. Guidal. "Generalized Parton Distributions and Deeply Virtual Compton Scattering". In : *J. Phys. G* 42.3 (2015), p. 034023. DOI : 10.1088/0954-3899/42/3/034023. arXiv : 1412.4651 [hep-ph].

[139] V. D. Burkert, L. Elouadrhiri et F. X. Girod. "The pressure distribution inside the proton". In : *Nature* 557.7705 (2018), p. 396-399. DOI : 10.1038/s41586-018-0060-z.

[140] R. Dupre, M. Guidal et M. Vanderhaeghen. "Tomographic image of the proton". In : *Phys. Rev. D* 95.1 (2017), p. 011501. DOI : 10.1103/PhysRevD.95.011501. arXiv : 1606.07821 [hep-ph].

[141] R. Dupré et al. "Analysis of Deeply Virtual Compton Scattering Data at Jefferson Lab and Proton Tomography". In : *Eur. Phys. J. A* 53.8 (2017), p. 171. DOI : 10.1140/epja/i2017-12356-8. arXiv : 1704.07330 [hep-ph].

[142] K. Kumericki, D. Mueller et A. Schafer. "Neural network generated parametrizations of deeply virtual Compton form factors". In : *JHEP* 07 (2011), p. 073. DOI : 10.1007/JHEP07(2011)073. arXiv : 1106.2808 [hep-ph].

[143] H. Moutarde, P. Sznajder et J. Wagner. "Unbiased determination of DVCS Compton Form Factors". In : *Eur. Phys. J. C* 79.7 (2019), p. 614. DOI : 10.1140/epjc/s10052-019-7117-5. arXiv : 1905.02089 [hep-ph].

[144] B. Berthou et al. "PARTONS : PARtonic Tomography Of Nucleon Software : A computing framework for the phenomenology of Generalized Parton Distributions". In : *Eur. Phys. J. C* 78.6 (2018), p. 478. DOI : 10.1140/epjc/s10052-018-5948-0. arXiv : 1512.06174 [hep-ph].

[145] I. Anikin et O. Teryaev. "Dispersion relations and subtractions in hard exclusive processes". In : *Phys. Rev. D* 76 (2007), p. 056007. DOI : 10.1103/PhysRevD.76.056007. arXiv : 0704.2185 [hep-ph].

[146] M. Diehl et D. Y. Ivanov. "Dispersion representations for hard exclusive processes : beyond the Born approximation". In : *Eur. Phys. J. C* 52 (2007), p. 919-932. DOI : 10.1140/epjc/s10052-007-0401-9. arXiv : 0707.0351 [hep-ph].

[147] J. S. Toll. "Causality and the Dispersion Relation : Logical Foundations". In : *Phys. Rev.* 104 (1956), p. 1760-1770. DOI : 10.1103/PhysRev.104.1760.

[148] G. R. Goldstein et S. Liuti. "The Use of Dispersion Relations in Hard Exclusive Processes and the Partonic Interpretation of Deeply Virtual Compton Scattering". In : *Phys. Rev. D* 80 (2009), p. 071501. DOI : 10.1103/PhysRevD.80.071501. arXiv : 0905.4753 [hep-ph].

[149] H. Moutarde et al. "Timelike and spacelike deeply virtual Compton scattering at next-to-leading order". In : *Phys. Rev.* D87.5 (2013), p. 054029. DOI : 10.1103/PhysRevD.87.054029. arXiv : 1301.3819 [hep-ph].

[150] P. Chatagnon et al. "First Measurement of Timelike Compton Scattering". In : *Phys. Rev. Lett.* 127.26 (2021), p. 262501. DOI : 10.1103/PhysRevLett.127.262501. arXiv : 2108.11746 [hep-ex].

[151] D. Müller et al. "Towards a fitting procedure to deeply virtual meson production - the next-to-leading order case". In : *Nucl. Phys. B* 884 (2014), p. 438-546. DOI : 10.1016/j.nuclphysb.2014.04.012. arXiv : 1310.5394 [hep-ph].

[152] J. P. Lansberg et al. "Single-Transverse-Spin Asymmetries in Exclusive Photoproduction of J/ψ in Ultra-Peripheral Collisions in the Fixed-Target Mode at the LHC and in the Collider Mode at RHIC". In : *Phys. Lett. B* 793 (2019), p. 33-40. DOI : 10.1016/j.physletb.2019.03.061. arXiv : 1812.04553 [hep-ph].

[153] C. Alexandrou et al. "Moments of nucleon generalized parton distributions from lattice QCD simulations at physical pion mass". In : *Phys. Rev. D* 101.3 (2020), p. 034519. DOI : 10.1103/PhysRevD.101.034519. arXiv : 1908.10706 [hep-lat].

[154] K. Cichy. "Overview of lattice calculations of the x-dependence of PDFs, GPDs and TMDs". In : *EPJ Web Conf.* 258 (2022), p. 01005. DOI : 10.1051/epjconf/202225801005. arXiv : 2111.04552 [hep-lat].

[155] X. Ji. "Parton Physics on a Euclidean Lattice". In : *Phys. Rev. Lett.* 110 (2013), p. 262002. DOI : 10.1103/PhysRevLett.110.262002. arXiv : 1305.1539 [hep-ph].

[156] A. V. Radyushkin. "Quasi-parton distribution functions, momentum distributions, and pseudo-parton distribution functions". In : *Phys. Rev. D* 96.3 (2017), p. 034025. DOI : 10.1103/PhysRevD.96.034025. arXiv : 1705.01488 [hep-ph].

[157] M. Constantinou. "The x-dependence of hadronic parton distributions : A review on the progress of lattice QCD". In : *Eur. Phys. J. A* 57.2 (2021), p. 77. DOI : 10.1140/epja/s10050-021-00353-7. arXiv : 2010.02445 [hep-lat].

[158] C. Adloff et al. “Measurement of deeply virtual Compton scattering at HERA”. In : *Phys. Lett. B* 517 (2001), p. 47-58. DOI : 10.1016/S0370-2693(01)00939-X. arXiv : hep-ex/0107005.

[159] S. Chekanov et al. “Measurement of deeply virtual Compton scattering at HERA”. In : *Phys. Lett. B* 573 (2003), p. 46-62. DOI : 10.1016/j.physletb.2003.08.048. arXiv : hep-ex/0305028.

[160] F. D. Aaron et al. “Deeply Virtual Compton Scattering and its Beam Charge Asymmetry in e+- Collisions at HERA”. In : *Phys. Lett. B* 681 (2009), p. 391-399. DOI : 10.1016/j.physletb.2009.10.035. arXiv : 0907.5289 [hep-ex].

[161] A. Airapetian et al. “Measurement of the beam spin azimuthal asymmetry associated with deeply virtual Compton scattering”. In : *Phys. Rev. Lett.* 87 (2001), p. 182001. DOI : 10.1103/PhysRevLett.87.182001. arXiv : hep-ex/0106068.

[162] S. Stepanyan et al. “Observation of exclusive deeply virtual Compton scattering in polarized electron beam asymmetry measurements”. In : *Phys. Rev. Lett.* 87 (2001), p. 182002. DOI : 10.1103/PhysRevLett.87.182002. arXiv : hep-ex/0107043.

[163] H. S. Jo et al. “Cross sections for the exclusive photon electroproduction on the proton and Generalized Parton Distributions”. In : *Phys. Rev. Lett.* 115.21 (2015), p. 212003. DOI : 10.1103/PhysRevLett.115.212003. arXiv : 1504.02009 [hep-ex].

[164] M. Defurne et al. “A glimpse of gluons through deeply virtual compton scattering on the proton”. In : *Nature Commun.* 8.1 (2017), p. 1408. DOI : 10.1038/s41467-017-01819-3. arXiv : 1703.09442 [hep-ex].

[165] I. V. Musatov et A. V. Radyushkin. “Evolution and models for skewed parton distributions”. In : *Phys. Rev. D* 61 (2000), p. 074027. DOI : 10.1103/PhysRevD.61.074027. arXiv : hep-ph/9905376.

[166] C. Mezrag, H. Moutarde et F. Sabatié. “Test of two new parametrizations of the generalized parton distribution H”. In : *Phys. Rev. D* 88.1 (2013), p. 014001. DOI : 10.1103/PhysRevD.88.014001. arXiv : 1304.7645 [hep-ph].

[167] C. Mezrag. “Generalised Parton Distributions : from phenomenological approaches to Dyson-Schwinger equations”. Thèse de doct. Orsay, 2015.

[168] M. Vanderhaeghen, P. A. M. Guichon et M. Guidal. “Hard electroproduction of photons and mesons on the nucleon”. In : *Phys. Rev. Lett.* 80 (1998), p. 5064-5067. DOI : 10.1103/PhysRevLett.80.5064.

[169] M. Vanderhaeghen, P. A. M. Guichon et M. Guidal. “Deeply virtual electroproduction of photons and mesons on the nucleon : Leading order amplitudes and power corrections”. In : *Phys. Rev. D* 60 (1999), p. 094017. DOI : 10.1103/PhysRevD.60.094017. arXiv : hep-ph/9905372.

[170] M. Guidal et al. “Nucleon form-factors from generalized parton distributions”. In : *Phys. Rev. D* 72 (2005), p. 054013. DOI : 10.1103/PhysRevD.72.054013. arXiv : hep-ph/0410251.

[171] K. Goeke, M. V. Polyakov et M. Vanderhaeghen. "Hard exclusive reactions and the structure of hadrons". In : *Prog. Part. Nucl. Phys.* 47 (2001), p. 401-515. DOI : 10.1016/S0146-6410(01)00158-2. arXiv : hep-ph/0106012.

[172] S. V. Goloskokov et P. Kroll. "Vector meson electroproduction at small Bjorken-x and generalized parton distributions". In : *Eur. Phys. J. C* 42 (2005), p. 281-301. DOI : 10.1140/epjc/s2005-02298-5. arXiv : hep-ph/0501242.

[173] S. V. Goloskokov et P. Kroll. "The Role of the quark and gluon GPDs in hard vector-meson electroproduction". In : *Eur. Phys. J. C* 53 (2008), p. 367-384. DOI : 10.1140/epjc/s10052-007-0466-5. arXiv : 0708.3569 [hep-ph].

[174] S. V. Goloskokov et P. Kroll. "An Attempt to understand exclusive pi+ electroproduction". In : *Eur. Phys. J. C* 65 (2010), p. 137-151. DOI : 10.1140/epjc/s10052-009-1178-9. arXiv : 0906.0460 [hep-ph].

[175] K. Kumerički et D. Müller. "Description and interpretation of DVCS measurements". In : *EPJ Web Conf.* 112 (2016). Sous la dir. de C. Marquet, B. Pire et F. Sabatie, p. 01012. DOI : 10.1051/epjconf/201611201012. arXiv : 1512.09014 [hep-ph].

[176] M. V. Polyakov et A. G. Shuvaev. "On "dual" parametrizations of generalized parton distributions". arXiv :hep-ph/0207153. Juill. 2002. arXiv : hep-ph/0207153.

[177] D. Mueller et A. Schafer. "Complex conformal spin partial wave expansion of generalized parton distributions and distribution amplitudes". In : *Nucl. Phys. B* 739 (2006), p. 1-59. DOI : 10.1016/j.nuclphysb.2006.01.019. arXiv : hep-ph/0509204.

[178] J. Boman et E. T. Quinto. "Support theorems for real-analytic Radon transforms". In : *Duke Mathematical Journal* 55.4 (1987), p. 943 -948. DOI : 10.1215/S0012-7094-87-05547-5. URL : https://doi.org/10.1215/S0012-7094-87-05547-5.

[179] J. M. M. Chavez et al. "Pion generalized parton distributions : A path toward phenomenology". In : *Phys. Rev. D* 105.9 (2022), p. 094012. DOI : 10.1103/PhysRevD.105.094012. arXiv : 2110.06052 [hep-ph].

[180] A. D. Kiureghian et O. Ditlevsen. "Aleatory or epistemic ? Does it matter ?" In : *Structural Safety* 31.2 (2009). Risk Acceptance and Risk Communication, p. 105-112. ISSN : 0167-4730. DOI : https://doi.org/10.1016/j.strusafe.2008.06.020. URL : https://www.sciencedirect.com/science/article/pii/S0167473008000556.

[181] A. Aitken. "On least squares and linear combination of observations". In : 55 (1934), p. 42-48. URL : http://scholar.google.com/scholar.bib?q=info:JP6Q2eRpT0QJ:scholar.google.com/&output=citation&hl=en&ct=citation&cd=0.

[182] J. Pumplin et al. "Uncertainties of predictions from parton distribution functions. 2. The Hessian method". In : *Phys. Rev. D* 65 (2001), p. 014013. DOI : 10.1103/PhysRevD.65.014013. arXiv : hep-ph/0101032.

[183] J. Pumplin et al. "New generation of parton distributions with uncertainties from global QCD analysis". In : *JHEP* 07 (2002), p. 012. DOI : 10.1088/1126-6708/2002/07/012. arXiv : hep-ph/0201195.

[184] J. Gao, L. Harland-Lang et J. Rojo. "The Structure of the Proton in the LHC Precision Era". In : *Phys. Rept.* 742 (2018), p. 1-121. DOI : 10.1016/j.physrep.2018.03.002. arXiv : 1709.04922 [hep-ph].

[185] J. Butterworth et al. "PDF4LHC recommendations for LHC Run II". In : *J. Phys. G* 43 (2016), p. 023001. DOI : 10.1088/0954-3899/43/2/023001. arXiv : 1510.03865 [hep-ph].

[186] L. van der Maaten et G. Hinton. "Visualizing Data using t-SNE". In : *Journal of Machine Learning Research* 9.86 (2008), p. 2579-2605. URL : http://jmlr.org/papers/v9/vandermaaten08a.html.

[187] S. Forte et al. "Neural network parametrization of deep inelastic structure functions". In : *JHEP* 05 (2002), p. 062. DOI : 10.1088/1126-6708/2002/05/062. arXiv : hep-ph/0204232.

[188] R. D. Ball et al. "Parton distributions with LHC data". In : *Nucl. Phys. B* 867 (2013), p. 244-289. DOI : 10.1016/j.nuclphysb.2012.10.003. arXiv : 1207.1303 [hep-ph].

[189] R. D. Ball et al. "Parton distributions for the LHC Run II". In : *JHEP* 04 (2015), p. 040. DOI : 10.1007/JHEP04(2015)040. arXiv : 1410.8849 [hep-ph].

[190] E. R. Nocera et al. "A first unbiased global determination of polarized PDFs and their uncertainties". In : *Nucl. Phys. B* 887 (2014), p. 276-308. DOI : 10.1016/j.nuclphysb.2014.08.008. arXiv : 1406.5539 [hep-ph].

[191] L. Del Debbio et al. "Neural-network analysis of Parton Distribution Functions from Ioffe-time pseudodistributions". In : *JHEP* 02 (2021), p. 138. DOI : 10.1007/JHEP02(2021)138. arXiv : 2010.03996 [hep-ph].

[192] F. Rosenblatt. "The Perceptron : A Probabilistic Model for Information Storage and Organization in The Brain". In : *Psychological Review* (1958), p. 65-386.

[193] D. E. Rumelhart, G. E. Hinton et R. J. Williams. "Learning representations by back-propagating errors". In : *Nature* 323.6088 (oct. 1986), p. 533-536. DOI : 10.1038/323533a0.

[194] A. Krizhevsky, I. Sutskever et G. E. Hinton. "ImageNet classification with deep convolutional neural networks". In : *Communications of the ACM* 60 (2012), p. 84 -90.

[195] K. Fukushima. "Neocognitron : A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position". In : *Biological Cybernetics* 36 (2004), p. 193-202.

[196] G. V. Cybenko. "Approximation by superpositions of a sigmoidal function". In : *Mathematics of Control, Signals and Systems* 2 (1989), p. 303-314.

[197] K. Hornik. "Approximation capabilities of multilayer feedforward networks". In : *Neural Networks* 4.2 (1991), p. 251-257. ISSN : 0893-6080. DOI : https://doi.org/10.1016/0893-6080(91)90009-T. URL : https://www.sciencedirect.com/science/article/pii/089360809190009T.

[198] S. Basodi et al. "Gradient amplification : An efficient way to train deep neural networks". In : *Big Data Mining and Analytics* 3.3 (2020), p. 196-207. DOI : 10.26599/BDMA.2020.9020004.

[199] M. Mitchell. *An Introduction to Genetic Algorithms*. Cambridge, MA, USA : MIT Press, 1998. ISBN : 0262631857.

[200] G. E. Hinton et al. *Improving neural networks by preventing co-adaptation of feature detectors*. 2012. DOI : 10.48550/ARXIV.1207.0580. URL : https://arxiv.org/abs/1207.0580.

[201] N. Srivastava et al. "Dropout : A Simple Way to Prevent Neural Networks from Overfitting". In : *Journal of Machine Learning Research* 15.56 (2014), p. 1929-1958.

[202] K. Kumericki. "Measurability of pressure inside the proton". In : *Nature* 570.7759 (2019), E1-E2. DOI : 10.1038/s41586-019-1211-6.

[203] M. Čuić, K. Kumerički et A. Schäfer. "Separation of Quark Flavors Using Deeply Virtual Compton Scattering Data". In : *Phys. Rev. Lett.* 125.23 (2020), p. 232005. DOI : 10.1103/PhysRevLett.125.232005. arXiv : 2007.00029 [hep-ph].

[204] I. F. Ilyas et X. Chu. *Data Cleaning*. ACM, 2019. ISBN : 978-1-4503-7152-0. DOI : 10.1145/3310205.

[205] C. Leys et al. "Detecting outliers : Do not use standard deviation around the mean, use absolute deviation around the median". In : *Journal of Experimental Social Psychology* 49.4 (2013), p. 764-766. ISSN : 0022-1031. DOI : https://doi.org/10.1016/j.jesp.2013.03.013. URL : https://www.sciencedirect.com/science/article/pii/S0022103113000668.

[206] H. Dutrieux. *On the parametrization and extraction of Generalized Parton Distributions*. KTH Dissertation. 2020. Available from : http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-267065.

[207] R. D. Ball et al. "Reweighting and Unweighting of Parton Distributions and the LHC W lepton asymmetry data". In : *Nucl. Phys. B* 855 (2012), p. 608-638. DOI : 10.1016/j.nuclphysb.2011.10.018. arXiv : 1108.1758 [hep-ph].

[208] C. E. Shannon. "A Mathematical Theory of Communication". In : *The Bell System Technical Journal* 27 (), 379-423, 623–656.

[209] A. Stuart. *Kendall's Advanced Theory of Statistics*. Wiley, ISBN 047023380X. 2010.

[210] R. D. Ball et al. "Reweighting NNPDFs : the W lepton asymmetry". In : *Nucl. Phys. B* 849 (2011). [Erratum : Nucl.Phys.B 854, 926–927 (2012), Erratum : Nucl.Phys.B 855, 927–928 (2012)], p. 112-143. DOI : 10.1016/j.nuclphysb.2011.03.017. arXiv : 1012.0836 [hep-ph].

[211] H.-W. Lin et al. "Parton distributions and lattice QCD calculations : a community white paper". In : *Prog. Part. Nucl. Phys.* 100 (2018), p. 107-160. DOI : 10.1016/j.ppnp.2018.01.007. arXiv : 1711.07916 [hep-ph].

[212] A. Afanasev et al. "Physics with Positron Beams at Jefferson Lab 12 GeV". In : (juin 2019). arXiv : 1906.09419 [nucl-ex].

[213] V. D. Burkert et al. “The CLAS12 Spectrometer at Jefferson Laboratory”. In : *Nucl. Instrum. Meth. A* 959 (2020), p. 163419. DOI : 10.1016/j.nima.2020.163419.

[214] V. D. Burkert et al. *Beam Charge Asymmetries for Deeply Virtual Compton Scattering on the Proton at CLAS12*. Proposal to PAC48. 2020.

[215] I. Gabdrakhmanov et O. Teryaev. “QCD motivated subtractions in hard photonic and mesonic reactions”. In : *PoS* Baldin-ISHEPP-XXI (2012), p. 035. DOI : 10.22323/1.173.0035.

[216] M. V. Polyakov et H.-D. Son. “Nucleon gravitational form factors from instantons : forces between quark and gluon subsystems”. In : *JHEP* 09 (2018), p. 156. DOI : 10.1007/JHEP09(2018)156. arXiv : 1808.00155 [hep-ph].

[217] A. Freund. “On the extraction of skewed parton distributions from experiment”. In : *Phys. Lett. B* 472 (2000), p. 412-419. DOI : 10.1016/S0370-2693(99)01332-5. arXiv : hep-ph/9903488.

[218] X.-D. Ji. “Off forward parton distributions”. In : *J. Phys. G* 24 (1998), p. 1181-1205. DOI : 10.1088/0954-3899/24/7/002. arXiv : hep-ph/9807358.

[219] M. Guidal, H. Moutarde et M. Vanderhaeghen. “Generalized Parton Distributions in the valence region from Deeply Virtual Compton Scattering”. In : *Rept. Prog. Phys.* 76 (2013), p. 066202. DOI : 10.1088/0034-4885/76/6/066202. arXiv : 1303.6600 [hep-ph].

[220] D. Mueller. “Generalized Parton Distributions – visions, basics, and realities –”. In : *Few Body Syst.* 55 (2014). Sous la dir. de N. G. Stefanis, p. 317-337. DOI : 10.1007/s00601-014-0894-3. arXiv : 1405.2817 [hep-ph].

[221] K. Kumericki et al. *Accessing GPDs from Experiment — Potential of A High-Luminosity EIC —*. Mai 2011. arXiv : 1105.0899 [hep-ph].

[222] K. Kumericki, D. Mueller et K. Passek-Kumericki. “Sum rules and dualities for generalized parton distributions : Is there a holographic principle ?” In : *Eur. Phys. J. C* 58 (2008), p. 193-215. DOI : 10.1140/epjc/s10052-008-0741-0. arXiv : 0805.0152 [hep-ph].

[223] M. Ding et al. “Symmetry, symmetry breaking, and pion parton distributions”. In : *Phys. Rev. D* 101.5 (2020), p. 054014. DOI : 10.1103/PhysRevD.101.054014. arXiv : 1905.05208 [nucl-th].

[224] C. Mezrag et al. “Sketching the pion’s valence-quark generalised parton distribution”. In : *Phys. Lett. B* 741 (2015), p. 190-196. DOI : 10.1016/j.physletb.2014.12.027. arXiv : 1411.6634 [nucl-th].

[225] F. Natterer. *The Mathematics of Computerized Tomography*. Society for Industrial et Applied Mathematics, 1991. ISBN : 978-0-89871-493-7. DOI : 10.1137/1.9780898719284.

[226] M. V. Polyakov. “Tomography for amplitudes of hard exclusive processes”. In : *Phys. Lett. B* 659 (2008), p. 542-550. DOI : 10.1016/j.physletb.2007.11.012. arXiv : 0707.2509 [hep-ph].

[227] W. L. van Neerven et A. Vogt. "NNLO evolution of deep inelastic structure functions : The Nonsinglet case". In : *Nucl. Phys. B* 568 (2000), p. 263-286. DOI : 10.1016/S0550-3213(99)00668-9. arXiv : hep-ph/9907472.

[228] V. Bertone, S. Carrazza et J. Rojo. "APFEL : A PDF Evolution Library with QED corrections". In : *Comput. Phys. Commun.* 185 (2014), p. 1647-1668. DOI : 10.1016/j.cpc.2014.03.007. arXiv : 1310.1394 [hep-ph].

[229] V. Bertone. "APFEL++ : A new PDF evolution library in C++". In : *PoS* DIS2017 (2018). Sous la dir. d'U. Klein, p. 201. DOI : 10.22323/1.297.0201. arXiv : 1708.00911 [hep-ph].

[230] A. Blelly, J. Bobin et H. Moutarde. "Sparse data inpainting for the recovery of Galactic-binary gravitational wave signals from gapped data". In : *Mon. Not. Roy. Astron. Soc.* 509.4 (2021), p. 5902-5917. DOI : 10.1093/mnras/stab3314. arXiv : 2104.05250 [gr-qc].

[231] J. M. M. Chávez et al. "Accessing the Pion 3D Structure at US and China Electron-Ion Colliders". In : *Phys. Rev. Lett.* 128.20 (2022), p. 202501. DOI : 10.1103/PhysRevLett.128.202501. arXiv : 2110.09462 [hep-ph].

[232] F. Gelis et al. "The Color Glass Condensate". In : *Ann. Rev. Nucl. Part. Sci.* 60 (2010), p. 463-489. DOI : 10.1146/annurev.nucl.010909.083629. arXiv : 1002.0333 [hep-ph].

[233] A. Cooper-Sarkar, P. Mertsch et S. Sarkar. "The high energy neutrino cross-section in the Standard Model and its uncertainty". In : *JHEP* 08 (2011), p. 042. DOI : 10.1007/JHEP08(2011)042. arXiv : 1106.3723 [hep-ph].

[234] R. Gauld et al. "Charm production in the forward region : constraints on the small-x gluon and backgrounds for neutrino astronomy". In : *JHEP* 11 (2015), p. 009. DOI : 10.1007/JHEP11(2015)009. arXiv : 1506.08025 [hep-ph].

[235] R. Gauld et al. "The prompt atmospheric neutrino flux in the light of LHCb". In : *JHEP* 02 (2016), p. 130. DOI : 10.1007/JHEP02(2016)130. arXiv : 1511.06346 [hep-ph].

[236] R. A. Khalek et al. "nNNPDF3.0 : Evidence for a modified partonic structure in heavy nuclei". In : (jan. 2022). arXiv : 2201.12363 [hep-ph].

[237] J. Breitweg et al. "Exclusive electroproduction of ρ^0 and J/ψ mesons at HERA". In : *Eur. Phys. J. C* 6 (1999), p. 603-627. DOI : 10.1007/s100529901051. arXiv : hep-ex/9808020.

[238] S. Chekanov et al. "Exclusive photoproduction of J / ψ mesons at HERA". In : *Eur. Phys. J. C* 24 (2002), p. 345-360. DOI : 10.1007/s10052-002-0953-7. arXiv : hep-ex/0201043.

[239] C. Adloff et al. "Diffractive photoproduction of psi(2S) mesons at HERA". In : *Phys. Lett. B* 541 (2002), p. 251-264. DOI : 10.1016/S0370-2693(02)02275-X. arXiv : hep-ex/0205107.

[240] S. Chekanov et al. "Exclusive electroproduction of J/ψ mesons at HERA". In : *Nuclear Physics B* 695.1-2 (2004), 3–37. ISSN : 0550-3213. DOI : 10.1016/j.nuclphysb.2004.06.034. URL : http://dx.doi.org/10.1016/j.nuclphysb.2004.06.034.

[241] A. Aktas et al. "Elastic J/ψ production at HERA". In : *Eur. Phys. J. C* 46 (2006), p. 585-603. DOI : 10.1140/epjc/s2006-02519-5. arXiv : hep-ex/0510016.

[242] S. Chekanov et al. "Measurement of J/ψ photoproduction at large momentum transfer at HERA". In : *JHEP* 05 (2010), p. 085. DOI : 10.1007/JHEP05(2010)085. arXiv : 0910.1235 [hep-ex].

[243] C. Alexa et al. "Elastic and Proton-Dissociative Photoproduction of J/ψ Mesons at HERA". In : *Eur. Phys. J. C* 73.6 (2013), p. 2466. DOI : 10.1140/epjc/s10052-013-2466-y. arXiv : 1304.5162 [hep-ex].

[244] R Aaij et al. "Exclusive J/ψ and ψ(2S) production in pp collisions at $\sqrt{s} = 7$ TeV". In : *J. Phys. G* 40 (2013), p. 045001. DOI : 10.1088/0954-3899/40/4/045001. arXiv : 1301.7084 [hep-ex].

[245] R. Aaij et al. "Updated measurements of exclusive J/ψ and ψ(2S) production cross-sections in pp collisions at $\sqrt{s} = 7$ TeV". In : *J. Phys. G* 41 (2014), p. 055002. DOI : 10.1088/0954-3899/41/5/055002. arXiv : 1401.3288 [hep-ex].

[246] B. Abelev et al. "Exclusive J/ψ photoproduction off protons in ultra-peripheral p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV". In : *Phys. Rev. Lett.* 113.23 (2014), p. 232504. DOI : 10.1103/PhysRevLett.113.232504. arXiv : 1406.7819 [nucl-ex].

[247] R. Aaij et al. "Measurement of the exclusive Υ production cross-section in pp collisions at $\sqrt{s} = 7$ TeV and 8 TeV". In : *JHEP* 09 (2015), p. 084. DOI : 10.1007/JHEP09(2015)084. arXiv : 1505.08139 [hep-ex].

[248] R. Aaij et al. "Central exclusive production of J/ψ and $\psi(2S)$ mesons in pp collisions at $\sqrt{s} = 13$ TeV". In : *JHEP* 10 (2018), p. 167. DOI : 10.1007/JHEP10(2018)167. arXiv : 1806.04079 [hep-ex].

[249] A. M. Sirunyan et al. "Measurement of exclusive Υ photoproduction from protons in pPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV". In : *Eur. Phys. J. C* 79.3 (2019), p. 277. DOI : 10.1140/epjc/s10052-019-6774-8. arXiv : 1809.11080 [hep-ex].

[250] A. G. Shuvaev et al. "Off diagonal distributions fixed by diagonal partons at small x and xi". In : *Phys. Rev. D* 60 (1999), p. 014015. DOI : 10.1103/PhysRevD.60.014015. arXiv : hep-ph/9902410.

[251] S. P. Jones et al. "Probes of the small x gluon via exclusive J/ψ and Υ production at HERA and the LHC". In : *JHEP* 11 (2013), p. 085. DOI : 10.1007/JHEP11(2013)085. arXiv : 1307.7099 [hep-ph].

[252] C. A. Flett et al. "How to include exclusive J/ψ production data in global PDF analyses". In : *Phys. Rev. D* 101.9 (2020), p. 094011. DOI : 10.1103/PhysRevD.101.094011. arXiv : 1908.08398 [hep-ph].

[253] L. Frankfurt et al. "Nondiagonal parton distribution in the leading logarithmic approximation". In : *Phys. Lett. B* 418 (1998). [Erratum : Phys.Lett.B 429, 414 (1998)], p. 345-354. DOI : 10.1016/S0370-2693(97)01152-0. arXiv : hep-ph/9703449.

[254] A. D. Martin et M. G. Ryskin. "The effect of off diagonal parton distributions in diffractive vector meson electroproduction". In : *Phys. Rev. D* 57 (1998), p. 6692-6700. DOI : 10.1103/PhysRevD.57.6692. arXiv : hep-ph/9711371.

[255] P. Hagler et al. “Helicity dependent and independent generalized parton distributions of the nucleon in lattice QCD”. In : *Eur. Phys. J. A* 24S1 (2005), p. 29-33. doi : 10.1140/epjad/s2005-05-005-y. arXiv : hep-ph/0410017.

[256] M. Diehl et W. Kugler. “Some numerical studies of the evolution of generalized parton distributions”. In : *Phys. Lett. B* 660 (2008), p. 202-211. doi : 10.1016/j.physletb.2007.12.047. arXiv : 0711.2184 [hep-ph].

[257] S. P. Jones et al. “Exclusive J/ψ and Υ photoproduction and the low x gluon”. In : *J. Phys. G* 43.3 (2016), p. 035002. doi : 10.1088/0954-3899/43/3/035002. arXiv : 1507.06942 [hep-ph].

[258] D. Y. Ivanov et al. “Exclusive photoproduction of a heavy vector meson in QCD”. In : *Eur. Phys. J. C* 34.3 (2004). [Erratum : Eur.Phys.J.C 75, 75 (2015)], p. 297-316. doi : 10.1140/epjc/s2004-01712-x. arXiv : hep-ph/0401131.

[259] M. Kirch, A. Manashov et A. Schafer. “Evolution equation for generalized parton distributions”. In : *Phys. Rev. D* 72 (2005), p. 114006. doi : 10.1103/PhysRevD.72.114006. arXiv : hep-ph/0509330.

[260] A. Freund et V. Guzey. “Study of nondiagonal parton distribution models”. In : *Phys. Lett. B* 462 (1999), p. 178-188. doi : 10.1016/S0370-2693(99)00881-3. arXiv : hep-ph/9806267.

[261] K. J. Golec-Biernat, A. D. Martin et M. G. Ryskin. “Diagonal input for the evolution of off diagonal partons”. In : *Phys. Lett. B* 456 (1999), p. 232-239. doi : 10.1016/S0370-2693(99)00504-3. arXiv : hep-ph/9903327.

[262] L. A. Harland-Lang. “Simple form for the low-x generalized parton distributions in the skewed regime”. In : *Phys. Rev. D* 88.3 (2013), p. 034029. doi : 10.1103/PhysRevD.88.034029. arXiv : 1306.6661 [hep-ph].

[263] L. A. Harland-Lang et al. “Parton distributions in the LHC era : MMHT 2014 PDFs”. In : *Eur. Phys. J. C* 75.5 (2015), p. 204. doi : 10.1140/epjc/s10052-015-3397-6. arXiv : 1412.3989 [hep-ph].