Extracting the QCD dipole cross-section

A dissertation submitted to The University of Manchester for the degree of
Master of Science
in the Faculty of Engineering and Physical Sciences

2006

Yu Wei

School of Physics and Astronomy

Contents

List of Figures

List of Tables

Abstract

THE UNIVERSITY OF MANCHESTER

ABSTRACT OF DISSERTATION submitted by Yu Wei for the Degree of Master of Science and entitled Extracting the QCD dipole cross-section. Month and Year of submission: September 2006.

By reviewing important experimental results and theoretical concepts leading to parton saturation, we introduce the colour dipole model as the key object that connects theories containing saturation dynamics with experiment. In particular, the colour dipole model can be used to assess the presence of saturation in the data. By summarizing the results of these assessments, we identify the important question: "to what extent are saturation dynamics present in the data?" We propose the extraction, or unfolding, of the dipole cross-section, without parameterization, as an unbiased way to approach the answer to this question. Then, we identify that an *ill-posed linear Fredholm integral equation* needs to be solved, in order to perform the extraction. To apply numerical methods, the integral equation is discretized and a parameterization of F_2 structure function data is performed. The problem is then successfully solved using Tikhonov regularization with a linear inequality constraint.

Declaration

No portion of this work referred to in this dissertation has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

Acknowledgements

I would like to thank my supervisor Professor Jeff Forshaw for offering me such a interesting project to work with, without his guidance and constant support, I would not have finished this dissertation. Moreover, I really appreciated the talk with him about, "what is most important to be a successful theoretical physicist?" from which I hope to have gained some of the most important guidance for the rest of my academic career.

I would also like to thank my parents, for giving me this precious opportunity to come to United Kingdom, and for respecting my choice to become a physicist. And my girlfriend Rita, for her support and understanding of my enthusiasm towards physics.

I am grateful to Jorge Silva and his father Carlos Silva, for their encouragement when I needed it most.

I would like to thank Dr. Qilong Ming of New York State University for the discussion about the incorporation of the linear inequality constraint into the Tikhonov regularization.

I am also grateful to my friends Tim Coughlin, Elliot Woods, Roger Goldsbrough and Amy for helping me proof read this dissertation.

Chapter 1

Theory and Background

1.1 Introduction

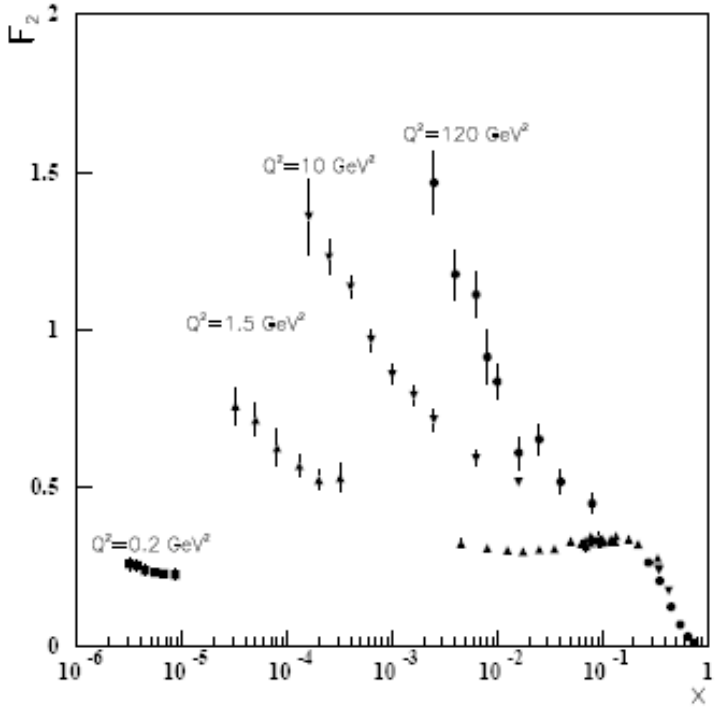

Figure 1.1: The strong rise of F_2 with decreasing x. The HERA data are from the 1994 runs and the data for $Q^2 = 1.5GeV^2$ and $x > 10^{-3}$(triangles) are from fixed target experiments. Figure taken from [1].

The advent of the HERA collider at DESY enabled the study of deep inelastic scattering (DIS) at small values of Bjorken x (below 0.01), where the $q\bar{q}$ sea and

gluon dynamics are dominant. The main observed effect is a strong rise of the proton structure function F_2 in the limit $x \to 0$, the rise becoming more rapid as Q^2 increases (see figure 1.1). Interpreted with the linear evolution equations of Quantum Chromodynamics, this reflects the increase of the parton densities (sea quarks and gluons). However, as we will illustrate later, the parton rise is so strong that for sufficiently small x the computed cross section would violate unitarity, indicating that important physical effects are neglected in the approximation leading to the linear evolution equations.

It has been found [2, 3] that this rise can be tamed by taking into account the non-linear saturation dynamics which are present at high gluon density. That is, while the cross-section for gluon splitting $g \to gg$ increases with decreasing x, the cross-section of gluon self-absorption $gg \to g$ also increases, eventually the rise of the gluon density will be tamed when these two processes balance each other. The question asked by the physicists in the early 90s was "how can we assess whether these non-linear dynamics exist in the data or not?"

Mueller [4,5] answered this question by showing that the gluon structure function can be viewed as the interaction of a colour dipole with the target. This enables us to connect the gluon density function with the dipole cross-section. Thus the taming of the growth of the gluon density as x decreases can now be directly viewed as the taming of the growth of the dipole cross-section. Based on this, phenomenological dipole models [6–8] were constructed and used to assess the presence of saturation in the data.

However, the results [9–12] of these assessments, using various dipole models, are still indefinite as to the role of non-linear effects in the current data. Data from some processes seem to prefer the dipole models containing saturation dynamics, however data from other processes cannot distinguish between the dipole models containing saturation dynamics and those which do not. In this project we propose a new way to assess the presence of saturation in the current data, that is to extract the dipole cross-section without parameterization.

This dissertation is divided into three parts. In part I, we are going to present an overview of the theoretical background of this project. Starting by introducing the

kinematics of DIS and the Quark Parton Model (QPM), by considering its higher order corrections, we write down the DGLAP linear evolution equations that predict the strong rise of the gluon density $xg(x, Q^2)$ as x decreases. We also present the prediction of a strong rise of the gluon density with decreasing x by the BFKL equation. Following this, we will show that the strong rise can be tamed by incorporating saturation dynamics. To assess the presence of these nonlinear effects, we introduce the colour dipole model that can be used to compare the theory containing saturation dynamics with experiment. Following this, we will describe in detail the DIS process in the colour dipole framework, and derive the $\gamma^* P$ cross-section as a function of dipole cross-section. We will present three recent colour dipole models whose predictions are compared with the data to assess the presence of saturation dynamics. We will summarise the results of theses assessments and motivate the extraction of the dipole cross-section without parameterization as an unbiased way to approach the important question "to what extent are saturation dynamics present in the data?" Finally we will identify the problem (a *linear fredholm integration equation*) we need to solve in order to perform this extraction.

In part II, we are going to discuss the preparations needed to solve this integration equation numerically. We will first discretize this integration equation, and then parameterize the structure function data in order to 'produce' enough data to perform the extraction.

In part III, we will first demonstrate a naive attempt to unfold the dipole cross-section. By analyzing its results, using singular value decomposition methods (SVD), we identify our problem as an *ill-posed problem,* which needs to be solved by regularization. We then present four regularization schemes and a brief review of four papers in the field of high energy physics concerning unfolding methods. From this we conclude that the Tikhonov method, which incorporate a "smoothness" constraints on the extracted solution, is the most suitable regularization method for our problem. Following this, we will discuss how accurately the regularized solution can approximate the true solution, and how the regularization parameter effects this accuracy. Following a discussion of two methods for choosing regularization parameters, we will extract the dipole cross-section using Tikhonov methods with several

different choices of regularization parameter. By analyzing these results, we propose to incorporate another constraint, which requires the extracted dipole cross-section to be a monotonic function of increasing dipole size r. The dipole cross-section is then successfully extracted with both constraints. Finally we outline the most important future developments that can be made based on this research.

1.2 Deep Inelastic Scattering

Since Rutherford uncovered the structure of the atom using the scattering of α particles off gold nuclei, scattering experiments have been and continue to be, one of the most powerful tools for studying the fundamental structure of matter and the theories that describe them. In this section, we are going to give a brief overview of a particular type of scattering experiment, called deep inelastic scattering, which ultimately aims to uncover the structure of the proton.

1.2.1 The kinematics of DIS

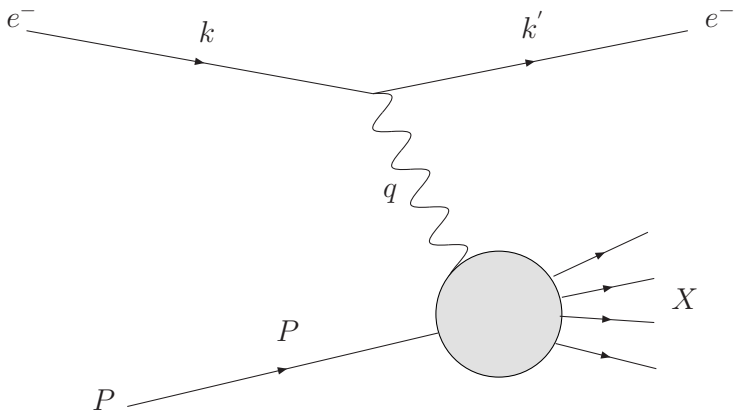

Figure 1.2: Deep inelastic scattering $e^-(k) + proton(P) \rightarrow e^-(k') + X$

The simplest Feynman diagram for the electron-proton DIS process is depicted in Fig.1.2, which shows an electron with four momentum k scattering off a proton with four momentum P and mass M, by exchanging a virtual photon γ^* with four momentum q. This process is characterized by the following kinematic variables:

- The virtuality of photon: $Q^2 = -q^2 = -t = -(k - k')^2$.

- The Bjorken variable: $x = Q^2/2p \cdot q$.

- The centre of mass (c.o.m) energy squared: $s = (P + k)^2$.

- The energy transferred between the lepton and proton systems: $y = P \cdot q/P \cdot k$.

- The recoil mass M_X^2: $W^2 = (P + q)^2 = M^2 + \frac{1-x}{x}Q^2$.

The word "deep" means high resolution, i.e. $Q^2 \gg M^2$, and "inelastic " means $M_X^2 \neq M^2$, i.e. $x < 1$. The cross-section of the above DIS process is given by [13]:

$$\sigma^{ep\to eX} = \sum_X \frac{1}{4ME} \int d\Phi \frac{1}{4} \sum_{spin} |M|^2_{ep\to eX} \tag{1.1}$$

where $\frac{1}{4ME}$ is the flux factor, $\int d\Phi$ is the phase space factor, and $|M|^2_{ep\to eX}$ is the transition amplitude. The spin averaged differential cross-section can be written (neglecting Z° boson exchange) as:

$$\frac{d^2\sigma}{dxdQ^2} = \frac{4\pi\alpha^2}{Q^4}\left[y^2 F_1(x, Q^2) + (\frac{1-y}{x} - \frac{xy^2M^2}{Q^2})F_2(x, Q^2)\right] \tag{1.2}$$

where

$$y = \frac{Q^2}{x(s - M^2)}$$

α is the electromagnetic coupling constant. F_1 and F_2 are the proton structure functions which are predicted to depend on Lorentz invariants x and Q^2 and contain all hadronic information necessary to describe the DIS cross-section.

1.2.2 Bjorken Scaling and the Quark Parton Model (QPM)

In 1969, Bjorken predicted that the structure function will behave like $F_2(x, Q^2) \to \tilde{F}_2(x)$ as $Q^2 \to \infty$ [14], and this was subsequently observed to be at least approximately correct in experiments. Later, Feynman explained Bjorken scaling based on his parton model [15]. The essence of the idea is that the photon scatters elastically off a point-like constituent of the proton. Because the scattering centre is a point, which has no dimensionful scale, then F_2 cannot depend on dimensionful Q^2. In addition, the parton model gives an explicit model for the structure function:

$$F_2(x) = \sum_i e_i^2 x f_i(x) \tag{1.3}$$

where the sum is over partons with charge e_i and $f_i(x)$ is the number density of electrically charged partons of type 'i' in the proton. Thus, deep inelastic scattering can be viewed as a photon scattering incoherently off individual partons, as shown in Fig.1.3. Identify the partons with quarks, the double differential cross-section can be written as [16]

$$\frac{d^2\sigma}{dxdQ^2} = \frac{2\pi\alpha^2}{xQ^4}[1 + (1-y)^2] \sum_i e_i^2 x q_i(x) \tag{1.4}$$

where distribution function $q_i(x)$, in the Bjorken frame [17, 18] (or infinite momentum frame), is the probability that the struck quark i carries a fraction x of the proton's momentum, P. These discoveries formed the foundations of the Quark-Parton Model (QPM). Under this framework, our kinematic variables Q^2 and x gain the following physical interpretations:

- The photon's virtuality Q^2 is now identified as the transverse resolution of the photon or transverse size of measured partons. In the Bjorken frame, $Q^2 \sim q_T^2$

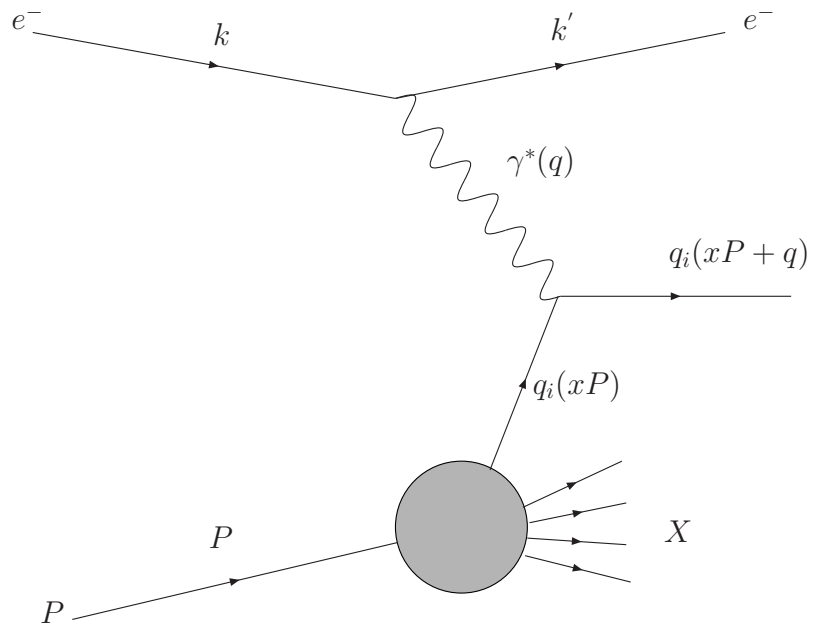

Figure 1.3: Deep inelastic scattering as the incoherent sum of point-like elastic scattering of partons.

(the transverse momentum of γ^*), and γ^* is considered to be absorbed over a transverse distance $\triangle x_\perp \sim 1/Q$.

- The Bjorken variable x is now identified as the longitudinal momentum fraction of the struck parton.

We discuss these variables in the Bjorken frame, because it is in this frame that deep inelastic scattering is most clearly visualized in a space-time picture, where the virtual photon has zero energy.

1.3 Higher order corrections

Subsequent experimental data showed that Bjorken scaling is not exact. The failure is particular dramatic at small x, as can be seen from Fig.1.1. To explain this effect, higher order QCD corrections need to be considered. That is, we need to take into account processes like $q \to qg$ and $g \to q\bar{q}$, which can contribute to DIS as shown in Fig.1.4.

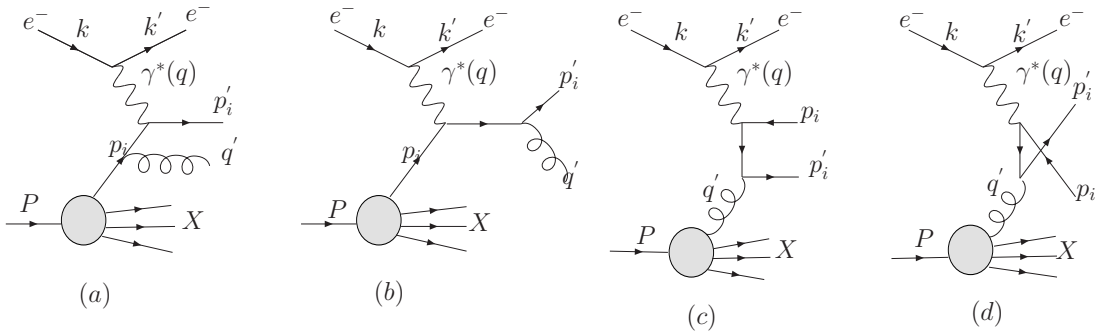

Figure 1.4: The higher order real corrections to the DIS process, where (a) and (b) represents the contribution from the $q \to qg$ process, (c) and (d) represents the contribution from $g \to q\bar{q}$ process

We will briefly discuss the DGLAP equations below, following the approach in [16]. To study these processes and identify the source of the scaling violation, we define the fractional momentum of the parton to be $p_i = \xi P$ and $z = Q^2/(2p_i \cdot q) = x/\xi$ which is the analogue of Bjorken x.

Now let us first consider the $q \to qg$ process. We notice that the γ^*-parton level diagram for processes (a) and (b), is actually the QCD Compton process (QCDC): $\gamma^* q \to gq$ as shown in Fig.1.5. Its hard scattering cross-section can be approximated as:

$$\hat{\sigma}_{QCDC} \simeq \hat{\sigma}_0 e_i^2 \frac{\alpha_s}{2\pi} \frac{4}{3} [\frac{1+z^2}{1-z}] \int_0^{p_t^2(max)} \frac{dp_t^2}{p_t^2} \tag{1.5}$$

where

$$\hat{\sigma}_0 = \frac{W^2 - m^2}{8\pi^2 \alpha_e},$$

W is the $\gamma^* p$ centre of mass energy, m is the proton mass. And

$$p_{t(max)}^2 = \frac{Q^2(1-z)}{4z}$$

is the gluon transverse momentum in the parton centre of mass frame. We regularize the $p_t^2(max)$ integral by introducing the cut-off energy, k^2, at the lower limit, and it gives

$$\int_{k^2}^{p_t^2(max)} \frac{dp_t^2}{p_t^2} = \ln(\frac{Q^2}{k^2}) + \ln(\frac{1-z}{4z}). \tag{1.6}$$

From this result, we identify the origin of scaling violations, i.e. gluon radiation introduces non-zero p_t and integrating over the dp_t^2/p_t^2 will introduce large logs, $\ln(Q^2/k^2)$. We define the function

$$P_{qq}(z) = \frac{4}{3} [\frac{1+z^2}{1-z}] \tag{1.7}$$

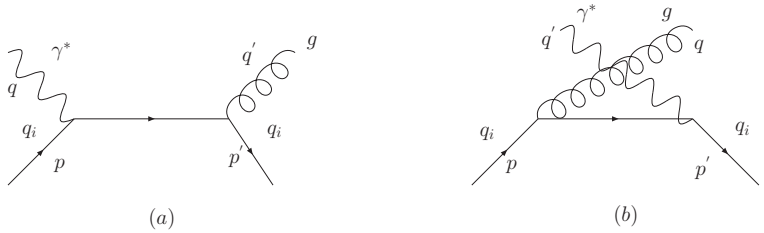

Figure 1.5: The QCD Compton process, (a)s-channel, (b)u-channel

as the unregularized probability distribution for $q \to q(z)g(1-z)$ splitting (where z is the fraction of the initial quark momentum carried by the final quark). Then the cross-section for QCDC is given by:

$$\hat{\sigma}_{QCDC}(z) = \hat{\sigma}_0 e_i^2 \frac{\alpha_s}{2\pi}\big[P_{qq}(z)\ln\big(\frac{Q^2}{k^2}\big) + C(z)\big] \tag{1.8}$$

where $C(z)$ are the terms left over in addition to the leading $\alpha_s \ln(Q^2/k^2)$ term.

To include the contribution from the QCDC process in F_2, we need to write the $\gamma^* P$ cross-section using the factorization theorem in QCD, which states that the cross-section for DIS may be written as the convolution of two terms: a calculable hard scattering cross-section and a non-perturbative parton density. Hence, the $\gamma^* P$ cross-section can be factorized as the convolution of parton distribution function (pdf) $q_i(\xi)$, with γ^*-parton scattering cross-section $\hat{\sigma}$:

$$\sigma(x, Q^2) = \sum_i \int_0^1 dz \int_0^1 d\xi q_i(\xi)\delta(x - z\xi)\hat{\sigma}(x, z) \tag{1.9}$$

Using the relation between structure functions and cross-sections for transverse and longitudinal virtual photon scattering

$$\begin{aligned} \sigma_T &= \frac{4\pi^2\alpha}{s} 2F_1 \\ \sigma_L &= \frac{4\pi^2\alpha}{s}\big[\frac{F_2}{x} - 2F_1\big] \end{aligned} \tag{1.10}$$

the expression for F_2 is obtained:

$$F_2 = \frac{s}{4\pi^2\alpha}(\sigma_T + \sigma_L)x = \frac{\sigma}{\sigma_0}x \quad \text{where} \quad \sigma_0 = \frac{4\pi^2\alpha}{s}. \tag{1.11}$$

The γ^*-parton level cross-section is given by

$$\hat{\sigma}_T + \hat{\sigma}_L = e_i^2\hat{\sigma}_0\delta(1-z) + 0. \tag{1.12}$$

Substitute (1.11) and (1.12) into (1.9), the structure function F_2 in terms of the convolution of pdf with γ^*-parton hard scattering cross-section $\hat{\sigma}$ is obtained:

$$F_2(x,Q^2) = x\int_0^1 dz \int_0^1 d\xi q_i(\xi)\delta(x-z\xi)e_i^2\delta(1-z). \tag{1.13}$$

Now, to include the contribution from the QCDC process, its cross-section $\hat{\sigma}_{QCDC}(z)$ is added to the zeroth order γ^*-parton level process to give:

$$\begin{aligned} F_2(x,Q^2) &= x\int_0^1 dz \int_0^1 d\xi q_i(\xi)\delta(x-z\xi)[e_i^2\delta(1-z) + \hat{\sigma}_{QCDC}(z)/\hat{\sigma}_0] \\ &= e_i^2\int_x^1 \frac{d\xi}{\xi} q_i(\xi)\big[\delta(1-\frac{x}{\xi}) + \frac{\alpha_s}{2\pi}P_{qq}\big(\frac{x}{\xi}\big)\ln\big(\frac{Q^2}{k^2}\big)\big]. \end{aligned} \tag{1.14}$$

including only the leading $\alpha_s \ln(Q^2/k^2)$ term. As can be seen, F_2 now depends on the arbitrary cut-off energy k^2. To remove this scale dependence, we introduce a new scale μ^2 such that $\mu^2 \gg k^2$. If μ^2 is chosen to be the collinear factorization scale, then the collinear singularities (soft non-perturbative physics) will be absorbed into the parton density, by rewriting the normalized parton density $q_i(x,\mu^2)$ as

$$q_i(x,\mu^2) = q_i^0(x) + \frac{\alpha_s}{2\pi}\int_x^1 \frac{d\xi}{\xi} q_i^0(\xi)P_{qq}\big(\frac{x}{\xi}\big)\ln\big(\frac{\mu^2}{k^2}\big). \tag{1.15}$$

q_i is now factorization scale dependent, and the structure function can be written as

$$F_2(x,Q^2) = xe_i^2\int_x^1 \frac{d\xi}{\xi} q_i(\xi,\mu^2)\big[\delta(1-\frac{x}{\xi}) + \frac{\alpha_s}{2\pi}P_{qq}\big(\frac{x}{\xi}\big)\ln\big(\frac{Q^2}{\mu^2}\big) + \frac{\alpha_s}{2\pi}C\big(\frac{x}{\xi}\big)\big]. \tag{1.16}$$

The renormalized parton density $q_i(x,\mu^2)$ is fundamentally non-perturbative and cannot be predicted from first principles, but its evolution with $\ln\mu^2$ can be calculated

$$\frac{\partial q_i(x,\mu^2)}{\partial \ln\mu^2} = \frac{\alpha_s}{2\pi}\int_x^1 \frac{d\xi}{\xi} q_i(\xi,\mu^2)P_{qq}(x/\xi). \tag{1.17}$$

As can be seen there is an apparent singularity in the unregularized splitting function (1.7), as $z \to 1$, which is associated with the emission of soft gluons. This will be canceled by its virtual correction at $\mathcal{O}(\alpha_s)$ [19], and the singularity is regularized by the '+ prescription':

$$\frac{1}{1-z} \to \frac{1}{(1-z)_+} \quad \text{where} \quad \int_0^1 dz \frac{f(z)}{1-z)_+} = \int_0^1 dz \frac{f(z)-f(1)}{1-z}. \tag{1.18}$$

From which, the regularized splitting function can be written as

$$P_{qq}(z) = \frac{4}{3}\left[\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z)\right]. \tag{1.19}$$

Similarly to above, the contribution to F_2 from the diagrams (c) and (d) in Fig.1.9 which underlies the processes $g \to q\bar{q}$ can also be calculated:

$$F_2^g(x,Q^2) = x\int_x^1 \frac{d\xi}{\xi} e_i^2 g^0(\xi)\frac{\alpha_s}{2\pi} P_{qg}\left(\frac{x}{\xi}\right)\ln\left(\frac{Q^2}{k^2}\right) \tag{1.20}$$

where

$$P_{qg}(z) = \frac{1}{2}[z^2 + (1-z)^2] \tag{1.21}$$

is the $g \to q\bar{q}$ splitting function. Taking into account this process (1.17) becomes

$$\frac{\partial q_i(x,\mu^2)}{\partial \ln \mu^2} = \frac{\alpha_s}{2\pi}\int_x^1 \frac{d\xi}{\xi}[q_i(\xi,\mu^2)P_{qq}(x/\xi) + g(\xi,\mu^2)P_{qg}(x/\xi)]. \tag{1.22}$$

By taking into account two further processes, $q \to g(z)q(1-z)$ and $g \to g(z)g(1-z)$ corresponding to the (regularized) splitting functions P_{gq} and P_{gg}, which give rise to the evolution of the gluon density function $g(x,\mu^2)$, together with the virtual correction in α_s, the DGLAP linear evolution equations [20–22] are obtained:

$$\frac{\partial \mathbf{q}(x,Q^2)}{\partial \ln Q^2} = \frac{\alpha_s(Q^2)}{2\pi}\sum_j \int_x^1 \frac{d\xi}{\xi}\mathbf{P}(x/\xi,\alpha_s(Q^2))\mathbf{q}(\xi,Q^2) \tag{1.23}$$

where $\mathbf{q}$ is vector $\mathbf{q} = \mathbf{q}(q_i, g)$, and $\mathbf{P}$ is a 2×2 matrix with the splitting function $P_{q_iq_j}$,P_{q_ig} in the first row and P_{gq_j},P_{gg} in the second row. From the DGLAP equation, the x and Q^2 dependent structure function F_2 is now generally predicted to be

$$\frac{F_2(x,Q^2)}{x} = \int_x^1 \frac{d\xi}{\xi}\left[\sum_i e_i^2 q_i(\xi,Q^2)\, C_q\left(\frac{x}{\xi},\alpha_s\right) + \bar{e}^2 g(\xi,q^2)\, C_g\left(\frac{x}{\xi},\alpha_s\right)\right] \tag{1.24}$$

where $\bar{e}^2 = \sum_i e_i^2$ and the sum is over all active quark flavors. C_q and C_g are called the coefficient functions (they contain the partonic scattering cross section), and they are "factorization scheme dependent" beyond lowest order. To the order we compute here $C_q = \delta(1 - x/\xi)$ and $C_g = 0$.

1.3.1 Prediction of the gluon density from DGLAP linear evolution

By keeping only the gluon splitting term P_{qg} and using the double leading log approximation (DLLA) where all terms $\sim [\alpha_s \ln(Q^2) \ln(1/x)]^n$ are summed, with a non-singular input, DGLAP predicts the gluon density $xg(x, Q^2)$ at low x to be [23]:

$$xg(x, Q^2) \sim \exp(\sqrt{ln(1/x)}) \tag{1.25}$$

which corresponds to the rise of $xg(x, Q^2)$ as $e^{\sqrt{ln(1/x)}}$ as x decreases.

1.3.2 Prediction of $\sigma^{\gamma^* P}$ from BFKL linear evolution equation

The BFKL equation, which provides an alternative to DGLAP for summing the contributions of multiple gluon emissions at small x. BFKL sums gluon ladders, taking into account all leading contributions $[\bar{\alpha}_s \ln(1/x)]^n$, with $\bar{\alpha}_s = \alpha_s N_c/\pi$. At lowest order it predicts the DIS cross-section to be [24]:

$$\sigma \sim \frac{x^{-\omega_0}}{\sqrt{\ln 1/x}} \tag{1.26}$$

where

$$\omega_0 = 4\bar{\alpha}_s \ln 2. \tag{1.27}$$

Like DGLAP, this result raises a serious problem with decreasing x, as the solution features a rapid rise with x which will eventually violate the Froissart unitarity bound: $\sigma \leq \ln^2 s$.

1.4 The role of non-linear effects

As already discussed at the beginning of the chapter, an important physical effect was neglected in the approximation which led to the linear evolution equations: the saturation dynamics is non-linear in nature. This should be important at the high parton densities which occur at low Bjorken x. Moreover these effects could tame the growth of gluon density from the $g \to gg$ splitting process, by taking into account its inverse $gg \to g$ self-absorption process. Qualitatively, the gluon-gluon recombination or self-absorption cross-section is of order $\alpha_s(Q^2)/Q^2$ at a transverse scale Q^2. In the proton infinite momentum frame, $xg(x, Q^2)$ is the number of gluons in the proton wave function which are localized within an area $\triangle x_\perp \sim 1/Q$ in the transverse plane. Therefore the saturation dynamics might be expected when the total effective gluon-gluon recombination cross-section in the proton approaches the size of the proton disk:

$$\frac{\alpha_s(Q^2)}{Q^2} xg(x, Q^2) \sim \pi R^2. \tag{1.28}$$

At this scale gluon fusion balances gluon splitting to create an equilibrium-like system of partons with a definite value for the average transverse momentum, which is called the "saturation scale" $Q_s(x)$. This x dependent scale will decrease as x decreases.

Gribov, Levin and Ryskin (1983) [2] and later Mueller and Qiu (1986) [3] extended the low-x DLLA (Double Leading Logarithmic Approximation) by including saturation dynamics in the equation which predicts the gluon density:

$$\frac{\partial^2 xg(x, Q^2)}{\partial \ln Q^2 \partial \ln(1/x)} = \frac{3\alpha_s}{\pi} xg(x, Q^2) - \frac{81\alpha_s^2}{16Q^2R^2} (xg(x, Q^2))^2 \tag{1.29}$$

where the second term takes into account the nonlinear effect from gluon recombination which will damp the growth of xg at low x.

In the early 1990s, A. Mueller proved two important results [4, 5], which placed the colour dipole model at centre stage for connecting theories containing saturation dynamics with experiment. Firstly, he showed that the gluon structure function can be viewed as the interaction of a colour dipole with the target, as shown in Fig.1.6.

Specifically, the interaction cross-section (called the dipole cross-section), can be written as [16]

$$\sigma_{dipole}(x, r^2) = \frac{8\pi\alpha_s}{3} \int \frac{d^2k}{k^4} F_g(x, k)(1 - e^{ik.r}) \tag{1.30}$$

where k is transverse momenta of gluons and $F_g(x, k)$ is the unintegrated gluon density function. The gluon density function is given by [16]:

$$xg(x, Q^2) = \frac{1}{\pi} \int_0^{Q^2} \frac{d^2k}{k^2} \Theta(Q^2 - k^2) F_g(x, k), \tag{1.31}$$

As can be seen the dipole cross-section $\sigma(x, r^2)$ and the gluon density $xg(x, Q^2)$ are bridged by the unintegrated gluon density $F_g(x, k)$. Thus, the damping of the growth of the gluon density function with decreasing x will be directly reflected as the damping of the growth of the dipole cross-section.

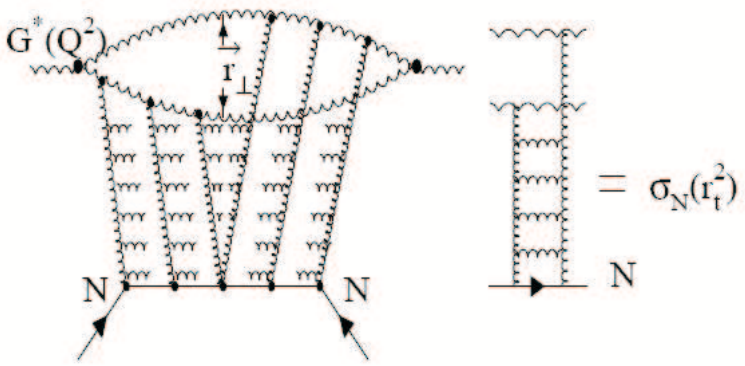

Figure 1.6: Interaction of a colour dipole with the target, where $G^*(Q^2)$ is a gluon probe, $r_\perp$ is the size of the colour dipole, N is the target hadron and $\sigma_N(r_t^2)$ is the dipole cross-section. Figure taken from [25].

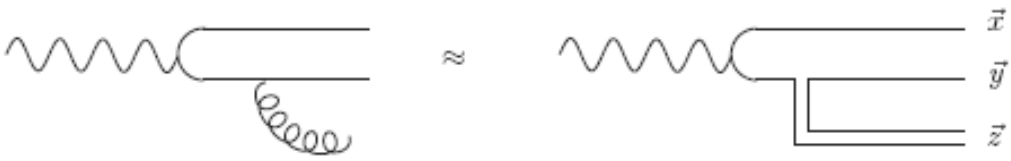

Figure 1.7: The BFKL evolution can be expressed as dipole evolution via gluon emission, where $\vec{x}$ and $\vec{y}$ are the transverse position of the original dipole and $\vec{z}$ is the transverse position of the emitted dipole. Figure taken from [26].

Secondly, Mueller proved that BFKL evolution can be rewritten as a decay of one dipole into two dipoles for large N_c, as shown in Fig.1.7. The amount of evolution depends on the centre-of-mass energy or rapidity, $Y = log(s/s_0)$ of the system. From this the deep inelastic scattering process can be considered as a classic dipole branching process, plus the process of dipole-proton scattering, as shown in Fig.1.8. Thus, the dipole cross-section satisfies the evolution equation [26]

$$\frac{\partial\sigma(\vec{x},\vec{y},Y)}{\partial Y} = \frac{N_c\alpha_s}{\pi}\int\frac{d^2\vec{z}}{2\pi}\frac{(\vec{x}-\vec{y})^2}{(\vec{x}-\vec{z})^2(\vec{y}-\vec{z})^2}\times[\sigma(\vec{x},\vec{z},Y)+\sigma(\vec{y},\vec{z},Y)-\sigma(\vec{x},\vec{y},Y)] \tag{1.32}$$

where the first two terms express the process of original dipole branching into two dipoles, one of which may then scatter off the proton, while the second term expresses the fact that the original dipole is destroyed after emission. As can be seen, this linear evolution equation results in an exponential growth of the dipole cross-section with rapidity:

$$\sigma \approx exp\{[4(N_c\alpha_s/\pi)log2]Y\} \tag{1.33}$$

as is the case for gluon density $xg(x,Q^2)$ predicted by BFKL. As already stated equation (1.32) ignores non-linear effects resulting from the fact that both newly

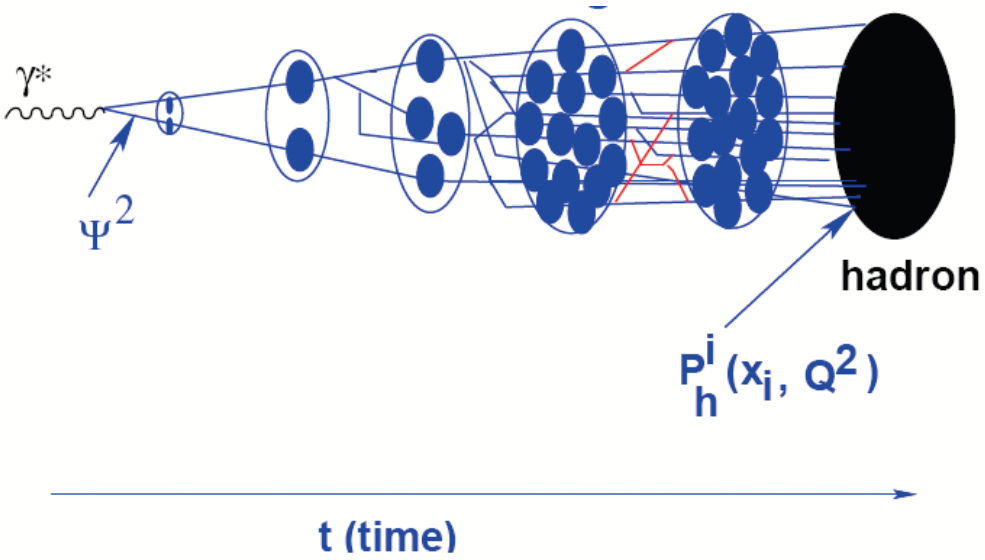

Figure 1.8: Deep inelastic scattering described as a dipole branching process and the process of dipole-proton scattering. Where Ψ^2 is the photon wave function and $P_h^i(x_i,Q^2)$ is the hadron wavefunction. Figure taken from [25].

created dipoles can scatter off the proton. Taking this process into account and assuming that the two dipoles scatter independently off the proton, our dipole cross-section can now be rewritten as

$$\frac{\partial\sigma(\vec{x},\vec{y},Y)}{\partial Y} = \frac{N_c\alpha_s}{\pi}\int\frac{d^2\vec{z}}{2\pi}\frac{(\vec{x}-\vec{y})^2}{(\vec{x}-\vec{z})^2(\vec{y}-\vec{z})^2}\times \\ [\sigma(\vec{x},\vec{z},Y)+\sigma(\vec{y},\vec{z},Y)-\sigma(\vec{x},\vec{y},Y)-\frac{1}{2}\sigma(\vec{x},\vec{z},Y)\sigma(\vec{y},\vec{z},Y)] \tag{1.34}$$

where the non-linear term $\frac{1}{2}\sigma(\vec{x},\vec{z},Y)\sigma(\vec{y},\vec{z},Y)$ should tame the growth of the dipole cross-section. This is the Balitsky-Kovchegov (BK) equation [27, 28].

Thus, based on the above important concepts, phenomenological colour dipole models containing or not containing saturation dynamics, can be constructed to assess the presence of saturation in the experimental data.

1.5 The general framework of the colour dipole picture

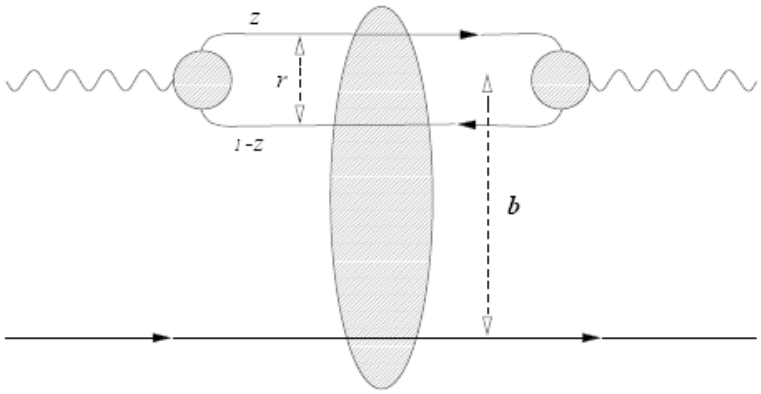

Figure 1.9: The colour dipole model for $\gamma^* + p \rightarrow \gamma^* + p$. Figure taken from [9].

Having discussed how non-linear effects can tame the growth of the gluon density and how the damping of the growth of the gluon density can be transformed into a damping of the growth of the dipole cross-section, we shall now describe the process of deep inelastic scattering in the colour dipole framework and derive the $\gamma^* P$ reaction cross-section as a function of dipole cross-section. The damping of

the growth of the dipole cross-section will then tame the growth of the $\gamma^* P$ total cross-section predicted by the DGLAP and BFKL equations.

The essential idea of the colour dipole model is that photon-hadron interactions at high energies can be described by the following mechanism (illustrated in Figure 1.9): In the rest frame of the hadron, the incoming virtual photon undergoes a fluctuation into virtual partonic or hadronic states a "long" time ($\sim 1/(mx)$) before the interaction with the hadron, and these then scatter coherently from the proton in a time which is short in comparison to $\sim 1/(mx)$.

In the following, the hadronic fluctuation of photon $\gamma^* \to q\bar{q}$ is first discussed, and then the dipole cross-section ($\sigma^{q\bar{q}+P\to X}$) is introduced using the Optical Theorem. Finally, the total deep inelastic scattering cross-section ($\sigma^{\gamma*+P\to X}$) is written as a function of the dipole cross-section.

1.5.1 The hadronic fluctuation of photon: $\gamma^* \to q\bar{q}$

The hadronic and partonic states generated by fluctuations of a photon are assumed to be dominated by colour singlet quark-antiquark pairs, which are characterized by their transverse size r and the fraction z of the light-cone momentum of the pair carried by the quark. Such states are called "colour dipoles" and are assumed to be eigenstates of diffraction, that is, they scatter without change of r and z in the diffractive limit. Before describing these states we shall first review the general expansion of hadronic states.

In the light-cone quantization, a hadronic state $|\Psi\rangle$ can be expanded in the form

$$|\Psi\rangle = \sum_n \Psi_n |n\rangle \tag{1.35}$$

where Ψ_n are the light-cone wavefunctions of the hadronic state and each colour singlet partonic state $|n\rangle$ is characterized by the transverse momenta $\vec{k}_{\perp i}$ and helicities λ_i of the partons, together with their longitudinal momenta k_i^+. As the diffractive eigenstates are those where the partons have fixed impact parameters, it will be convenient to proceed our discussion in a mixed configuration-momentum space description, where the transverse momenta $\vec{k}_{\perp i}$ are transformed into their conjugate

transverse position vectors $\vec{b}_i$ and the light-cone energies k_i^+ are transformed into the fractional light-cone energies x_i:

$$\vec{k}_{\perp i} \sim \vec{b}_i \qquad k_i^+ \sim x_i \tag{1.36}$$

where

$$x_i = \frac{k_i^+}{P^+} \qquad P^+ = \sum_i k_i^+ \qquad \sum_i x_i = 1 \tag{1.37}$$

and

$$P_\perp = \sum_i \vec{k}_{\perp i} \tag{1.38}$$

where P^+ is the light-cone energy and $P_\perp$ is the transverse momentum of the hadron state. Thus, the expansion (1.35) can be written as

$$|\Psi\rangle = \sum_i \Psi(x_i, \vec{b}_i, \lambda_i) \, |x_i, \vec{b}_i, \lambda_i\rangle \, . \tag{1.39}$$

Thus, the strongly interacting fluctuations of the photon can be decomposed into a superposition of the Fock states in the quark-gluon basis:

$$|\gamma^*\rangle_H = \sum \Psi_{q\bar{q}} \, |q\bar{q}\rangle + \sum \Psi_{q\bar{q}g} \, |q\bar{q}g\rangle + \text{ Higher Fock States.} \tag{1.40}$$

As discussed previously, the $q\bar{q}$ states dominate in the diffractive process. Denoting the transverse positions of the quark and antiquark by vectors $\vec{b}_1$ and $\vec{b}_2$ respectively, the new impact vector $\vec{b}$ can be defined as

$$\vec{b} = (\vec{b}_1 + \vec{b}_2)/2, \qquad b = |\vec{b}| \tag{1.41}$$

and the relative transverse position vector

$$\vec{r} = \vec{b}_1 - \vec{b}_2, \qquad r = |\vec{r}| \tag{1.42}$$

is introduced, where r is the size or transverse size of the dipole. From equations (1.39), (1.41) and (1.42), the expansion (1.40) becomes

$$|\gamma^*, \lambda\rangle_H = \sum_{h\bar{h}} \int dz d^2\vec{r}\, \Psi^{\gamma,\lambda}_{h,\bar{h}}(\vec{r}, z)\, |z, \vec{b}_i, h, \bar{h}\rangle + \text{ Higher Fock States} \tag{1.43}$$

where λ is the photon helicity and $h, \bar{h}$ are the helicities of the quark and antiquark respectively.

1.5.2 The dipole cross-section $\sigma^{q\bar{q}+P\to X}$

Using the Optical Theorem, the dipole cross-section, which is the total cross-section for scattering a dipole of size r from a proton ($q\bar{q} + P \to X$), is expressed in terms of the imaginary part of the forward elastic scattering ($q\bar{q} + P \to q\bar{q} + P$) amplitude:

$$\begin{aligned} \sigma_{dipole}(s, r, z) &= \frac{\text{Im } A(s, \vec{r}, z, \vec{P}_\perp)(s, t = 0)}{s} \\ &= \frac{\text{Im } \int d^2\vec{b} e^{i\vec{P}_\perp \cdot \vec{b}} \langle z, \vec{b}_i, h, \bar{h}|\hat{T}|z, \vec{b}_i, h, \bar{h}\rangle \,|_{P_\perp=0}}{s} \end{aligned} \tag{1.44}$$

where A is the amplitude for elastically scattering a dipole from the proton, $\vec{P}_\perp$ is the transverse momentum of the outgoing proton, s is the center of mass energy and t is the energy transferred to the proton. Because, in the diffractive limit, the dipole states are eigenstates of the scattering operator:

$$\hat{T}\, |z, \vec{b}_i, h, \bar{h}\rangle = \tau(\vec{b}, s; z, \vec{r})\, |z, \vec{b}_i, h, \bar{h}\rangle\,, \tag{1.45}$$

the dipole cross section can be written as

$$\sigma_{dipole}(s, r, z) = \int d^2\vec{b} \frac{\text{Im } \tau(\vec{b}, s; z, \vec{r})}{s}. \tag{1.46}$$

1.5.3 The reaction cross-section $\sigma^{\gamma^*+P\to X}$

Again, using the Optical Theorem, the $\gamma^* P$ reaction cross-section ($\gamma^* + P \to X$) is expressed in terms of the imaginary part of the forward elastic scattering ($\gamma^* + P \to \gamma^* + P$) amplitude:

$$\sigma_{L,T} = s^{-1}\text{Im } \langle \gamma^*, \lambda|\hat{T}|\gamma^*, \lambda\rangle\, (s, t = 0). \tag{1.47}$$

Inserting the dipole states (1.43) and using expression (1.46) (for convenience we denote σ_{dipole} as simply σ) gives the $\gamma^* P$ cross-section:

$$\sigma_{L,T}^{\gamma^* p \to X} = \int dz d^2\vec{r} |\Psi_\gamma^{L,T}(r,z)|^2 \sigma(s,r,z). \tag{1.48}$$

Because the reaction cross-section $\sigma_{L,T}(\gamma^* P \to X)$is connected with the F_2 structure function by

$$F_2(x,Q^2) = \frac{Q^2}{4\pi^2\alpha_{em}}(\sigma_L + \sigma_T), \tag{1.49}$$

F_2 can be expressed directly in terms of dipole cross-section σ:

$$F_2(x,Q^2) = \frac{4\pi^2\alpha_{em}}{Q^2} \int dz d^2\vec{r} \left[|\Psi_\gamma^L(r,z)|^2 + |\Psi_\gamma^T(r,z)|^2 \right] \sigma(s,r,z). \tag{1.50}$$

1.5.4 The Photon wavefunction

For small r, the light-cone photon wavefunctions are given by the tree level QED expressions [26]:

$$|\Psi_\gamma^L(r,z)|^2 = \frac{6}{\pi^2}\alpha_{em} \sum_{f=1}^{n_f} e_f^2 Q^2 z^2 (1-z)^2 K_0^2(\epsilon r) \tag{1.51}$$

$$|\Psi_\gamma^T(r,z)|^2 = \frac{3}{2\pi^2}\alpha_{em} \sum_{f=1}^{n_f} e_f^2 \{ [z^2 + (1-z)^2] \epsilon^2 K_1^2(\epsilon r) + m_f^2 K_0^2(\epsilon r) \} \tag{1.52}$$

where

$$\epsilon = \sqrt{z(1-z)Q^2 + m_f^2} \tag{1.53}$$

and the sum is over all n_f quark flavors f and $K_0(x)$ and $K_1(x) = -\partial_x K_0(x)$ are modified Bessel functions [29] with the asymptotic behaviors:

$$K_0(x) \approx (\frac{\pi}{2x})^{1/2} e^{-x} \approx K_1(x), \qquad x \to \infty, \tag{1.54}$$

$$K_0(x) \approx -\log x \quad K_1(x) \approx x^{-1}, \qquad x \to 0. \tag{1.55}$$

The following qualitative features are immediately observed from the above expressions:

1. At large Q^2 and fixed z, it can be seen from (1.54), that the wavefunctions decrease exponentially with increasing r, i.e.

$$\Psi(r,z) \sim \exp\{-rQ[z(1-z)]^{\frac{1}{2}}\}$$

provided $Q^2 z(1-z) \gg m_f^2$. Thus, only small dipoles can contribute unless z is close to its end-point values 0 or 1. As can be seen from equations (1.51) and (1.52)

- $|\Psi_\gamma^L(r,z)|$ is suppressed as $z \to 0$ *or* 1.
- $|\Psi_\gamma^T(r,z)|$ is not suppressed as $z \to 0$ *or* 1.

This is why longitudinal photon process are more inherently perturbative.

2. For small Q^2, particularly when $Q^2 = 0$ and $\epsilon = m_f$, the wavefunctions fall off as

$$\Psi(r,z) \sim \exp(-m_f\, r).$$

Hence, large dipoles ($r \approx m_f^{-1}$) with light quarks can typically contribute. However, when $r > 1$ fm, strong forces between quarks start to play an important role, leading to confinement. In this region, it is perhaps more appropriate to use the Generalized Vector Dominance (GVD) models [30].

1.6 The color dipole models

Different phenomenological colour dipole models are characterized by different forms for the dipole cross-sections. In the following, we will first predict the behavior of the dipole cross-section for small dipoles using perturbative QCD. This, together with the observation from above that large dipoles tend to have soft hadronic behavior, forms the general guidance for constructing dipole models. Then, we will discuss three particular colour dipole models.

Properties of the dipole cross-section for small dipoles

At low Bjorken x, the P_{qq} term in the DGLAP evolution equation derived in section 1.3 is negligible in comparison to the gluon splitting term P_{qg}, thus the evolution of $F_2(x, Q^2)$ can be approximated to be:

$$\frac{\partial F_2(x,Q^2)}{\partial \ln Q^2} \approx \sum_q e_q^2 \frac{\alpha_s}{2\pi} \int_x^1 dz G(x/z, Q^2) P_{qg}(z) \tag{1.56}$$

where e_q is the electric charge of a quark q (in units of the electron charge) and $G(x, Q^2)$ is the gluon momentum density. Substituting the prediction of $F_2(x, Q^2)$ from (1.50), the DGLAP prediction of the dipole cross-section is given by [26]:

$$\sigma(s, r) \approx \frac{\pi^2 \alpha_s}{3} r^2 G(x, A^2/r^2) \tag{1.57}$$

where the constant $A \approx 3$ is obtained from numerical studies. From this result, we observe that the dipole cross-section behaves roughly like $\sigma(s, r) \to r^2$ as $r \to 0$. Thus, at small separations the $q\bar{q}$ pair appears to have a much reduced cross-section with partons inside proton. This behavior of the dipole cross-section is called "color transparency".

1.6.1 Golec-Biernat-Wüsthoff model

The GW dipole model [7, 31, 32] has the following parametrization form for the dipole cross-section:

$$\sigma = \sigma_0 \left\{ 1 - \exp\left[-\frac{r}{2R_0(x)} \right] \right\} \tag{1.58}$$

with

$$R_0(x) = \frac{1}{Q_0} \left(\frac{x_{mod}}{x_0} \right)^{\lambda/2} \qquad x_{mod} = x\big(1 + \frac{4m_f^2}{Q^2}\big) \tag{1.59}$$

where m_f is the quark mass with $m_f = 0.14$ GeV for the light quarks (u,d,s) and $m_f = 1.5$ GeV for the charm quark, $Q_0 = 1$ GeV. Using a purely perturbative photon wavefunction, the 3 free parameters x_0, σ_0 and λ are determined by fitting to F_2 structure function data, to be $\sigma_0 = 29.12$mb, $\lambda = 0.277$ and $x_0 = 0.41 \times 10^{-4}$. As

can be seen from (1.58), $\sigma \to r^2$ when $r \to 0$, implying this model does incorporate "colour transparency".

Qualitatively, saturation is realized from the scaling of the $q\bar{q}$ transverse separation r in the dipole cross-section by the saturation radius $R_0(x)$. Thus, at fixed r, while $r < R_0(x)$, σ will rise rapidly as x decreases until $R_0(x)$ becomes smaller than r (caused by the decreasing of x), whereupon σ will tend to have a much softer dependence on x.

1.6.2 The CGC model

The CGC model from Iancu, Itakura and Munier [8] aims to embody the theory of Colour Glass Condensate [33, 34]. The model has the following form:

$$\sigma = 2\pi R^2 N_0 \left(\frac{rQ_s}{2}\right)^{2[\gamma_s + \frac{\log(2/rQ_s)}{k\lambda\log(1/x)}]} \qquad for \qquad rQ_s \leq 2 \tag{1.60}$$

$$= 2\pi R^2 \{1 - \exp[-a\log^2(brQ_s)]\} \qquad for \qquad rQ_s > 2 \tag{1.61}$$

where the saturation scale is $Q_s \equiv (x_0/x)^{\lambda/2}$ GeV. The coefficients a and b are uniquely determined by ensuring continuity of the cross-section and its first derivative at $rQ_s = 2$. For $rQ_s < 2$ the solution matches that of the leading order BFKL equation and fixes γ_s to be 0.63 and k to be 9.9. The coefficient N_0 is strongly correlated to the definition of the saturation scale and it was found that the quality of fit to F_2 data is only weakly dependent upon N_0. Thus, for a fixed value of N_0, the parameters: x_0, λ and R need to be fixed by a fit to the data.

1.6.3 FS04 Model

This model [10, 35] was based on the hypothesis of two pomerons, which had been explored in [36]. The cross-section has the form

$$\sigma(s, r) = \sigma_{soft}(s, r) + \sigma_{hard}(s, r) \tag{1.62}$$

which takes into account the contribution from both 'soft' and 'hard' pomerons. The FS04 model actually contains two sub models: the 'FS04 Regge model' which

does not contain the idea of saturation and 'FS04 saturation model' which does. The dipole cross-section has the following form:

$$\sigma = A_H r^2 x_m^{-\lambda_H} \qquad for \qquad r < r_0 \tag{1.63}$$

$$= A_S x_m^{-\lambda_S} \qquad for \qquad r > r_1 \tag{1.64}$$

where

$$x_m = \frac{Q^2}{Q^2 + W^2}\left(1 + \frac{4m^2}{Q^2}\right). \tag{1.65}$$

m is the quark mass, which is taken to be fixed at 1.4 GeV for charm quark dipoles and is a parameter to be fitted for light quark dipoles. A_H, A_S, λ_S, λ_H and r_1 are also parameters to be fitted. The dipole cross-section σ in the intermediate region $r_0 \leq r \leq r_1$ is determined by interpolating linearly between the two forms of (1.63) and (1.64). The difference between FS04 Regge model and FS04 saturation model comes from the mechanism by which the boundary parameter r_0 is determined. If r_0 is a constant, then the parametrization contains no saturation, which is the case of FS04 Regge model. Alternatively, in the FS04 saturation model r_0, is determined by requiring that the hard component of the cross section is some fixed fraction of the soft component:

$$\sigma(s, r_0)/\sigma(s, r_1) = f \tag{1.66}$$

where f is fixed and r_0 is a parameter to be fitted. Thus, as x decreases, r_0 will decrease. At a fixed value of r, the rise of the dipole cross-section will begin to be be damped as soon as r_0 becomes smaller than r. This will give rise to saturation effects. The predictions of the CGC, GW and FS04 Regge models are shown in Fig.1.10.

From the above discussion, the characteristics of saturation embodied in colour dipole models can be summarized as follows: The strong rise with energy of the dipole cross-section, which holds for small dipoles, pertains only for $r < r_s(x)$, where $r_s(x)$ is the saturation radius that decreases monotonically as x decreases.

1.7 Confront the data

As we have discussed, being able to assess the presence of saturation dynamics in experimental data is the key to verifying that the incorporation of saturation dynamics into the theories is the right direction to take for reconciling QCD theory with experiment. Moreover, it is also important to test the validity of the color dipole models. Thus, the predictions of the above dipole models are compared with data and the results are summarized below.

Forshaw and Shaw *et al* [10–12] have compared the predictions from the various colour dipole models with the data from deep inelastic scattering (DIS), deeply virtual Compton scattering (DVCS) and diffractive deep inelastic scattering (DDIS) processes. They found that the colour dipole models are indeed able to explain the data very well. However, the results on the presence of saturation effects in the data are inconclusive. The F_2 data suggest the presence of saturation dynamics, while the data from DVCS and DDIS processes are unable to distinguish between models which include or exclude saturation.

The key question is: can we, without any parameterization, extract the dipole cross-section from the data and assess directly the extent to which saturation is

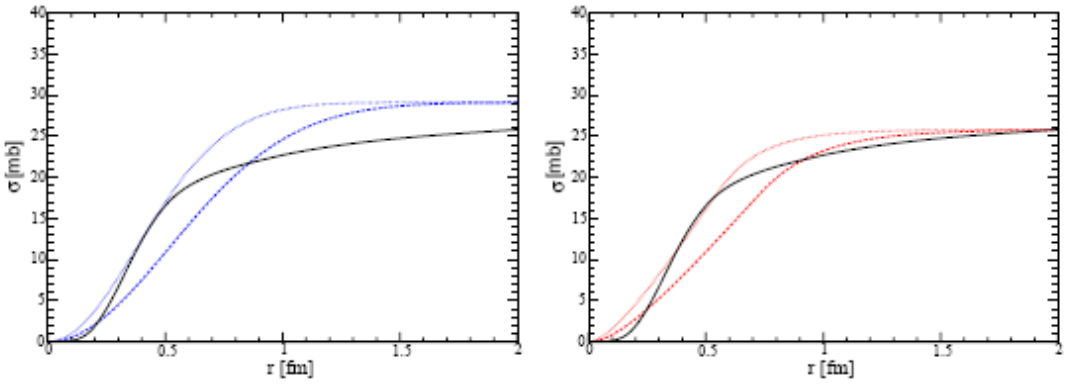

Figure 1.10: The GW model dipole cross-section (left) and the CGC dipole cross-section (right) is shown at $Q^2 = 2\ GeV^2$ (dotted line) and $Q^2 = 20\ GeV^2$ (dashed line). The Q^2-independent FS04 Regge model dipole cross-section (solid line) at the same energy is shown on both graphs. Figure taken from [26].

present in the data?

The second part of this thesis is devoted to analyzing the nature of this problem, exploring the methods for extracting the dipole cross-section and identifying the best way of performing this task.

1.8 Identify the problem

All the physical information needed for extracting, or unfolding, the dipole cross-section σ from the structure function F_2 is contained in equation (1.50), with the photon light-cone wavefunction $|\Psi_\gamma^{L,T}(r,z)|^2$ well determined. More precisely, because $|\Psi_\gamma^{L,T}(r,z)|^2$ depends explicitly on Q^2, as observed in expressions (1.51) and (1.52), the dipole cross-section is assumed to be independent of z, thus the equation (1.50) can be written as

$$F_2(Q^2,x) = \frac{4\pi^2\alpha_{em}}{Q^2}\int dz d^2\vec{r}\left[|\Psi_\gamma^L(r,z)|^2 + |\Psi_\gamma^T(r,z)|^2\right]\sigma(r,x). \tag{1.67}$$

If we we substitute $d^2r = 2\pi r dr$ and define

$$G(Q^2,r) = \int_0^1 dz\,\frac{8\pi^3\alpha_{em}}{Q^2}\,r\left[|\Psi_\gamma^L(r,z)|^2 + |\Psi_\gamma^T(r,z)|^2\right] \tag{1.68}$$

then

$$F_2(x,Q^2) = \int d\vec{r}\,G(Q^2,r)\,\sigma(r,x). \tag{1.69}$$

Notice that because the integration kernel $G(Q^2,r)$ doesn't depend on x, x can be considered as an external variable. Hence, the integration equation can at some value of x be written as:

$$F_2(Q^2) = \int d\vec{r}\,G(Q^2,r)\,\sigma(r). \tag{1.70}$$

We identify this equation as a special class of mathematical equation called a *linear Fredholm integral equation of the first kind* with the general form:

$$g(y) = \int dy\, K(x, y)\, f(y). \tag{1.71}$$

Thus, by performing the above simplification, our physical problem of unfolding σ from F_2 has been transformed into a well-defined mathematical problem.

Chapter 2

The Preparation

According to [37] and [38], to solve the linear Fredholm integral equation numerically, it needs to be discretized i.e. the integration (1.70) needs to be approximated by matrix multiplication. At some particular value of Bjorken x we write:

$$F_2(Q^2) = \int dr\, G(Q^2, r)\, \sigma(r) \quad \sim \quad b = A \times x$$

where A is a matrix, b and x (this is different from Bjorken x) are column vectors. The accuracy of this approximation depends on the number of intervals the integration range has been divided into, which subsequently depends on the number of columns of the matrix A. However, to be able to solve the matrix equation, the number of rows need to be larger than the number of columns (the number of equations need to be more than the number of variables). Thus, how accurately the integration can be approximated in our case is ultimately determined by the number of rows in vector b, i.e. how many data points are available.

The structure function F_2 is usually measured at a quite limited number (typically less than 10) of Q^2 values for each value of Bjorken x. However to approximate the integration accurately, hundreds of data points will be needed for each value of x. To be able to "provide" these data, a parameterization is done for structure function data in the range of $0.045 \leq Q^2 \leq 45\, GeV^2$.

In section 2.1 the discretization of the integration equation (1.70) will be discussed and then the parameterization of the F_2 structure function will be presented

in section 2.2.

2.1 Discretization of the integration equation

The integration equation (1.70) is discretized using the quadrature method [39]. In this method, a quadrature rule with integration nodes (abscissas) $t_1, \ldots, t_n$ and corresponding weights $\omega_1, \ldots, \omega_n$ is used to approximate an integral as

$$\int_0^1 \phi(t)dt \approx \sum_{j=1}^{n} \omega_j \phi(t_j). \tag{2.1}$$

To apply this method to our equation:

$$F_2(Q^2) = \int dr\, G(Q^2, r)\, \sigma(r)$$

the range of integration from r_1 to r_n is divided into n integration nodes with logarithmic spacing between them, i.e.

$$r_i = \exp[\ln r_1 + r_0 \times (i-1)] \quad \text{where} \quad r_0 = (\ln r_n - \ln r_1)/(n-1) \tag{2.2}$$

where r_i is the i^{th} integration node. The range of variable Q^2 is also divided into n nodes from Q_1^2 to Q_n^2 with logarithmic spacing between them, i.e.

$$Q_i^2 = \exp[\ln Q_1^2 + Q_0^2 \times (i-1)] \quad \text{where} \quad Q_0^2 = [\ln Q_n^2 - \ln Q_1^2]/(n-1). \tag{2.3}$$

The integration weight ω_i is the spacing between the r_i^{th} and r_{i+1}^{th} integration nodes:

$$\omega_i = dr_{i\to i+1} = r_{i+1} - r_i = \exp[\ln r_1 + r_0 \times i] - \exp[\ln r_1 + r_0 \times (i-1)]. \tag{2.4}$$

Thus using the quadrature method, our integration equation becomes

$$F_2(Q_i^2) = \sum_{j=1}^{n-1} G(Q_i^2, r_j)\omega_j \cdot \sigma(r_j). \tag{2.5}$$

This equation can now be written in matrix form:

$$b = Ax \tag{2.6}$$

where

$$(A)_{ij} = G(Q_i^2, r_j)\omega_j \tag{2.7}$$

$$x = (\sigma(r_1), \sigma(r_2), \ldots, \sigma(r_n)) \tag{2.8}$$

$$b = (F_2(Q_1^2), F_2(Q_2^2), \ldots, F_2(Q_n^2)) \tag{2.9}$$

2.1.1 Computing matrix A

The $n \times n$ matrix A in equation (2.6) is

$$\begin{aligned}(A)_{ij} = G(Q_i^2, r_j) &= \int_0^1 dz \, \frac{8\pi^3 \alpha_{em}}{Q_i^2} \, r_j |\Psi(Q_i^2, r_j)|^2 \\ &= \int_0^1 dz \, \frac{8\pi^3 \alpha_{em}}{Q_i^2} \, r_j \{|\Psi_\gamma^L(r_j, z)|^2 + |\Psi_\gamma^T(r_j, z)|^2\}\end{aligned} \tag{2.10}$$

with

$$|\Psi_\gamma^L(r, z)|^2 = \frac{6}{\pi^2}\alpha_{em} \sum_{f=1}^{n_f} e_f^2 Q^2 z^2 (1-z)^2 K_0^2(\epsilon r)$$

$$|\Psi_\gamma^T(r, z)|^2 = \frac{3}{2\pi^2}\alpha_{em} \sum_{f=1}^{n_f} e_f^2 \{[z^2 + (1-z)^2]\epsilon^2 K_1^2(\epsilon r) + m_f^2 K_0^2(\epsilon r)\}$$

$(A)_{ij}$ was computed using Matlab [40], with the following choice of parameters:

- $r_1 = 10^{-6}$, $r_n = 200$ in natural units, with $n = 1000$.
- $Q_1^2 = 0.28\, GeV^2$, $Q_n^2 = 45\, GeV^2$, with $n = 1000$.
- The quark electric charges e_f are $2/3$ for up and charm quark, $-1/3$ for down and strange quark.
- The quark masses m_f are summed over for u,d,s,c quarks with $m_f = 0.14\, GeV$ for the light quarks (u,d,s), and $m_f = 1.4\, GeV$ for the charm quark.

To show that the discretization is successful, the matrix equation (2.6) was used to predict the structure function F_2, using the FS04 saturation model [10] at $Q^2 = 4.5$, 8.5, 22 and 45 GeV^2 respectively. The result is shown in Fig.2.1.

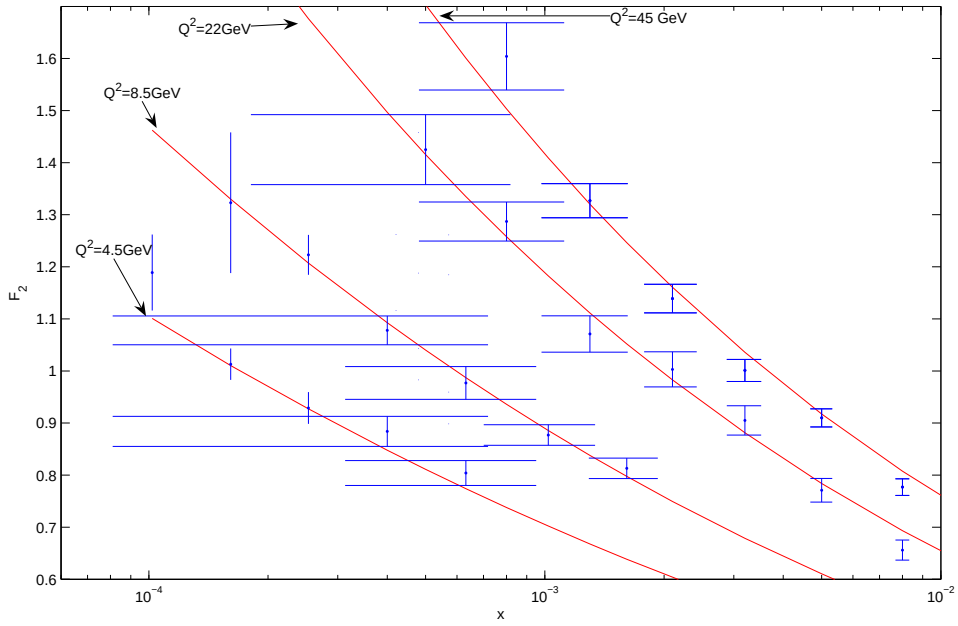

Figure 2.1: Comparison of the structure function (red curves) calculated using the matrix equation (2.6) and the FS04 saturation model, with structure function data from Zeus [41]. The data does not have x error bars, they are the artifacts from Matlab's plotting function.

2.2 Parameterization of structure function data

Our first goal is to find a parametrization for the structure function data, so that its value and associated error can be predicted at any required values of x and Q^2 within the desired range. This is achieved by minimizing the χ^2 of the parameterization with respect to a set of parameters being fitted. The χ^2 is defined as

$$\chi^2(\alpha) = \sum_{i=1,j=1}^{n} \frac{(F_2^D(x_i, Q_j^2) - F_2^P(\alpha, x_i, Q_j^2))^2}{e_{ij}^2} \tag{2.11}$$

where $F_2^D(x_i, Q_j^2)$ is the structure function data, e_{ij}^2 is its corresponding error. α is the vector of free parameters being fitted and $F_2^P(\alpha, x_i, Q_j^2)$ is the parameterized function that predicts the value of structure function. $F_2^P(\alpha, x_i, Q_j^2)$ was rather arbitrarily chosen to have the following form:

$$F_2^P(\alpha, x_i, Q_j^2) = \left(\frac{Q^2}{Q^2+\mu^2}\right)^{\gamma} \cdot x^{\lambda_0+\lambda_1 \ln Q^2+\lambda_2 \ln^2 Q^2} \cdot (1-x)^{\beta_1 \ln Q^2+\beta_2 \ln^2 Q^2} \cdot (1 + a\,x^{0.2} + b\,x^{0.4} + c\,x^{0.6} + d\,x^{0.8}) \tag{2.12}$$

where γ, μ, λ_0, λ_1, λ_2, β_1, β_2, a, b, c, d are the free parameters to be fitted. The minimization is carried out using the program 'Minuit' [42] from the CERN program library [43]. In the following, we will briefly discuss how Minuit determines the desired free parameters.

Minuit works such that, by feeding it the structure function data and χ^2 as a function of the free parameters, it minimizes χ^2 by varying these free parameters according to the instructed method. When the minimization process has converged, it returns the values of the variable parameters at the minimum of χ^2, and their estimated errors.

There are four minimization methods available in Minuit, corresponding to four different minimizers. The one called 'Migrad', which is recommended by the authors of Minuit and capable of providing reliable information about parameter errors is used. The other three minimizers are, according to the authors, either "kept in the program mainly for sentimental reasons" or "[give] no reliable information about parameter errors".

The (one standard deviation) parameter errors in Migrad are defined as the change needed in a parameter to change the χ^2 value by one, while keeping all other $N-1$ parameters fixed. They are computed by calculating the error matrix (or covariance matrix) [44], which is the inverse of the matrix of second derivatives of the χ^2 function (2.11), and then multiplied by one if one standard deviation error is required, and multiplied by two, if two standard deviation error is required etc. The square roots of the diagonal elements of this matrix are the calculated parameter errors, therefore they take into account all the parameter correlations

Parameters	Value	Error	Parameters	Value	Error
γ	2.873×10^{-1}	1.761×10^{-2}	β_1	-1.533	3.039×10^{-1}
μ	2.334	3.191×10^{-1}	a	-7.124	3.523×10^{-2}
λ_0	5.043×10^{-2}	1.872×10^{-3}	b	3.391×10^{1}	1.239×10^{-1}
λ_1	4.029×10^{-2}	8.484×10^{-4}	c	-7.997×10^{1}	3.560×10^{-1}
λ_2	-2.647×10^{-4}	1.881×10^{-4}	d	7.521×10^{1}	8.488×10^{-1}
β_1	1.287×10^{1}	1.287			

Table 2.1: The variable parameters obtained at the minimum of the χ^2 value.

The minimization technique currently implemented in Migrad is a stable variation of the Davidon-Fletcher-Powell variable-metric algorithm [45]. It uses its current estimate of the covariance matrix to determine the current search direction, thus the algorithm converges to the correct error matrix as it converges to the function minimum. A good approximation to the gradient vector at the current best point is also needed.

The F_2 structure function data from Zeus [41] with Q^2 from $0.045\,GeV^2$ to $45\,GeV^2$ and 180 data points are used in the parameterization. The minimization process is found to converge, with $\chi^2 \approx 0.625$ per experimental point. The obtained parameter values with errors are shown in table 2.1.

The prediction of $F_2(x, Q^2)$ from this parameterization is compared with Zeus F_2 data in Fig.2.2.

2.2.1 Computing the 'pseudodata' using the parameterization

The vector b in (2.6) corresponding to our pesudodata $F_2(Q^2)$ is subsequently computed using the parameterization $F_2^P(\alpha, x, Q^2)$. For the remainder of this thesis we work at only one value of $x = 0.0004$ and focus upon the unfolding procedure at that x. Of course future work would also consider unfolding at any x. We consider Q^2 in the range $0.28\,GeV^2 < Q^2 < 45\,GeV^2$ (with log spacing).

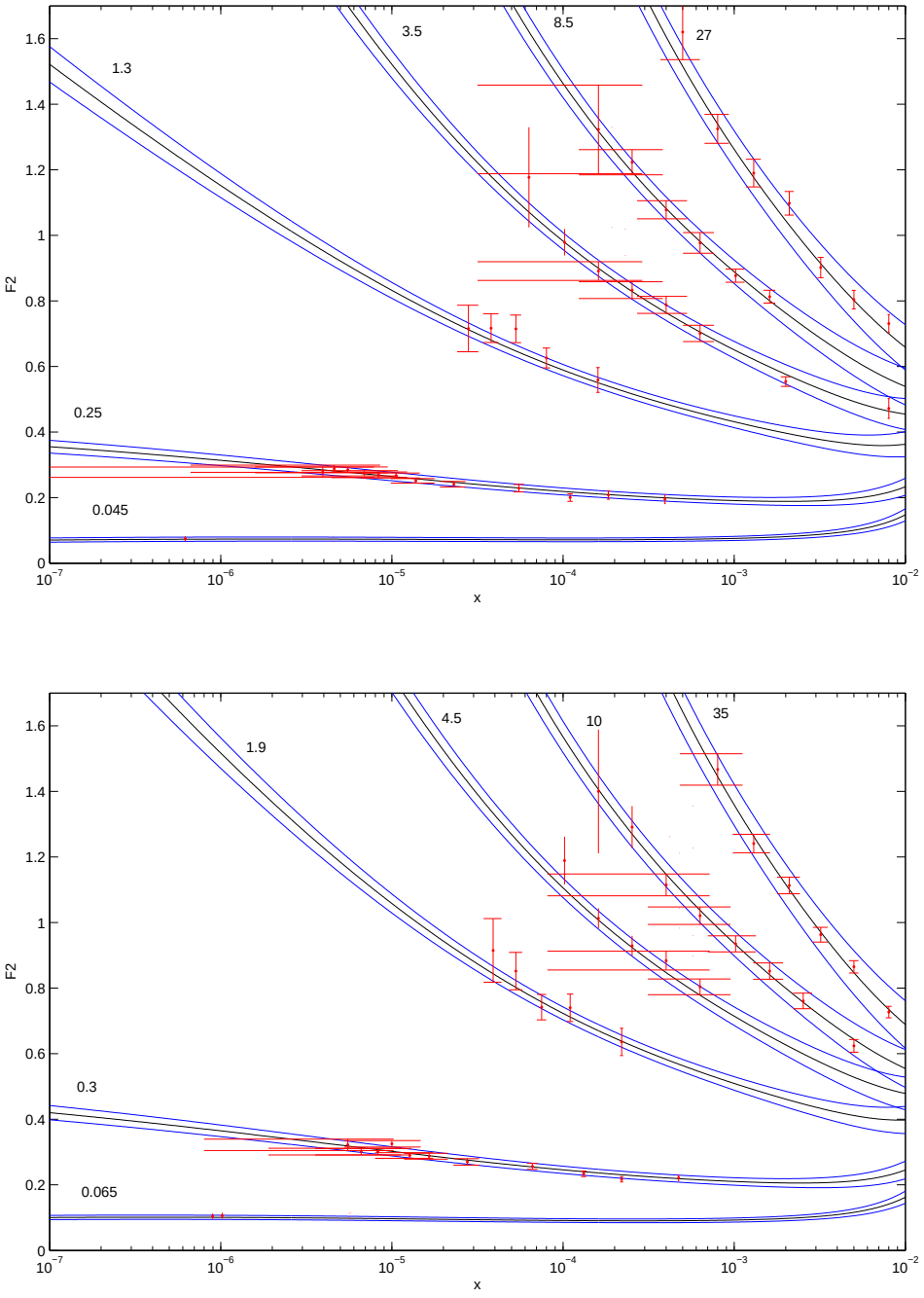

Figure 2.2: The parameterization of the structure function data, the black curves are the predicted structure function F_2. The blue curves are their error bands. The numbers on the graph are the values of Q^2. Again the x error bars here are the artifacts from Matlab's plotting function.

Chapter 3

The Unfolding

3.1 Understanding the nature of the problem

After the above preparation, matrix A which 'contains' the photon light-cone wave-function, and vector b, which 'contains' the pesudodata $F_2(Q^2)$ are generated. Now our problem of solving the integral equation (1.70) had been converted to a problem of solving the matrix equation (2.6):

$$b = Ax.$$

The most immediate method for solving this equation is by inverting the matrix A, however this leads to a meaningless solution, as will be demonstrated in section 3.1.1. By analyzing the problems of this direct inversion method using a powerful tool in numerical linear algebra called the singular value decomposition in section 3.1.2, the origins of these problems are identified. Finally, the general strategies for treating these problems are discussed.

The units of all the variables discussed below are natural units.

3.1.1 The problem with direct unfolding by inverting A

As mentioned above x can naively be computed by inverting the matrix A, i.e.:

$$x = A^{-1}b. \tag{3.1}$$

The result is shown in Fig.3.1 and Fig.3.2. To compare, the GW [32] dipole model

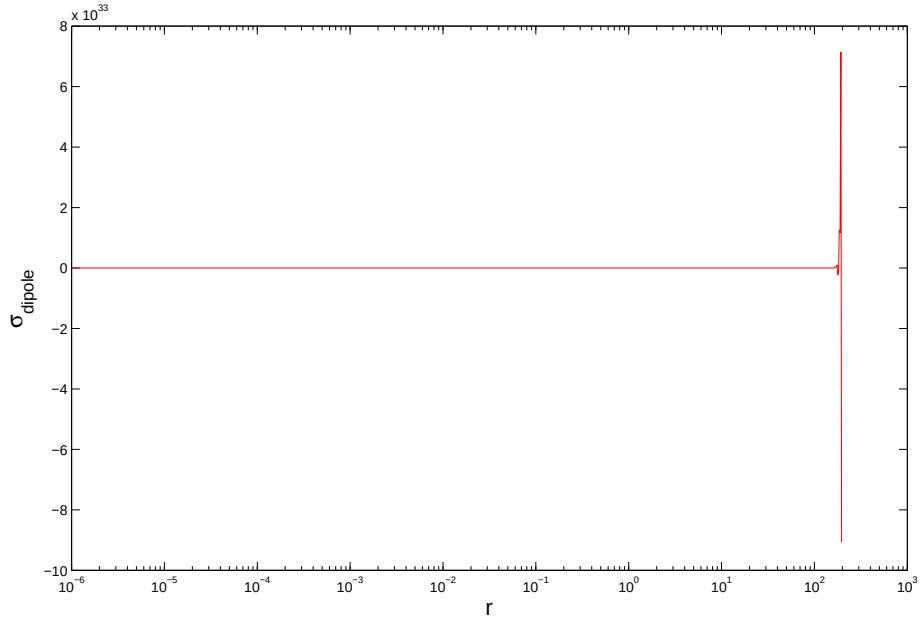

Figure 3.1: The dipole cross-section σ obtained from directly inverting the matrix A, as a function of dipole size r.

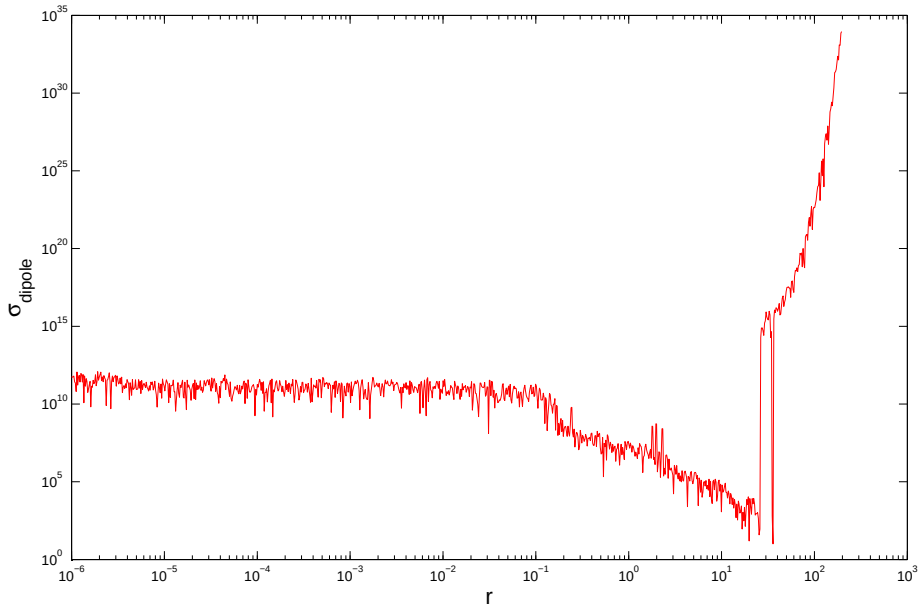

Figure 3.2: The absolute value of obtained dipole cross-section with y-axis in logarithmic scale as a function of dipole size r.

is shown in Fig.3.3. As can be seen, the details of the dipole cross-section curve

are dominated by its behavior at $r \approx 200$ with amplitude $\approx \pm 8 \times 10^{33}$. Thus, Fig.3.2 with y-axis in logarithmic scale is presented. As can be seen, the value of the extracted dipole cross section (denoted σ) oscillates wildly over two orders of magnitude up to $r \approx 20$, and then raised dramatically when r approaches the cut-off value. The extracted dipole cross-section from this method is distinctly different from the GW model and the solution is apparently meaningless.

To gain insight about the above behavior, the condition number [46] of matrix A which measures the solution's sensitivity to perturbations, is computed:

$$cond(A) \equiv ||A||_2 \, ||A'||_2 = \frac{s_1}{s_n} \tag{3.2}$$

where A' is the pseudoinverse of A, s_1 and s_n are the largest and smallest singular values of A (their meaning will be explained in the following sections). In our case, $cond(A) = 2.565 \times 10^{43}$, this shows the matrix A is apparently ill-conditioned, and the solution will be extremely sensitive to perturbations. This means that a very small perturbation in the structure function data b can cause very large variation in the extracted dipole cross-section x. With such sensitivity to perturbations the solu-

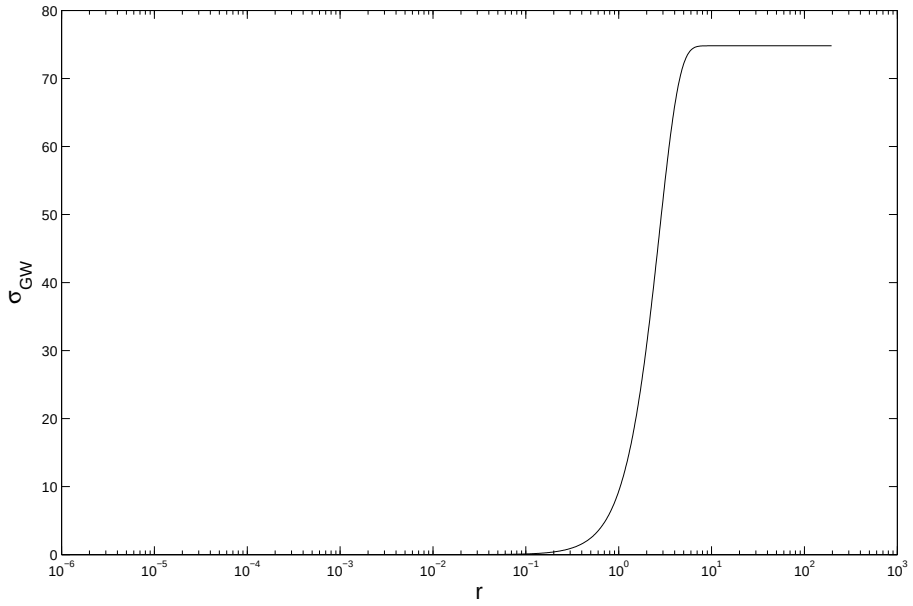

Figure 3.3: The dipole cross-section from the GW model at Bjorken $x = 0.0004$, as function of dipole size r.

tion obtained above can be completely dominated by errors from the measurements of b.

To understand the origin of this large condition number, and how this give rise to the oscillatory and large norm characteristic of our solution, the problem is analyzed using the singular value decomposition (SVD) method.

3.1.2 The origin of the problem

The SVD [47] of a $n \times n$ matrix A is defined as

$$A = U\Sigma V^T = \sum_{i=1}^{n} u_i s_i v_i^T \tag{3.3}$$

where

$$U^T U = V^T V = I_n$$

and I_n is the identity matrix. u_i and v_i are the column vectors of the $n \times n$ orthogonal matrices U and V. $u_i s_i v_i^T$ is a matrix generated by the outer product of a column vector u_i and row vector v_i^T. Thus, matrix A is decomposed as a sum of n matrices. The diagonal matrix Σ has nonnegative diagonal elements appearing in nonincreasing order such that:

$$s_1 \geq s_2 \geq \cdots \geq s_n \geq 0.$$

The elements s_i are called the singular values of A while the vectors u_i and v_i are called the left and right singular vectors of A respectively. Geometrically, the SVD of A provides two sets of orthonormal basis vectors (i.e. the columns of U and V) such that the matrix A becomes diagonal when transformed into these two bases. It is related to the eigenvalue decomposition of $A^T A$ and AA^T in such way that:

$$A^T A = V\Sigma^2 V^T \quad and \quad AA^T = U\Sigma^2 U^T. \tag{3.4}$$

Now we are able to write down the explicit expression for the pseudoinverse of A as

$$A' = \sum_{i=1}^{\mathrm{A}} v_i s_i^{-1} u_i^T.$$

To analyze our problem, the matrix A was decomposed using SVD and the singular values s_i as a function of a index i are plotted in Fig.3.4. The right singular vectors v_i at various value of index i are plotted in Fig.3.5. The following properties

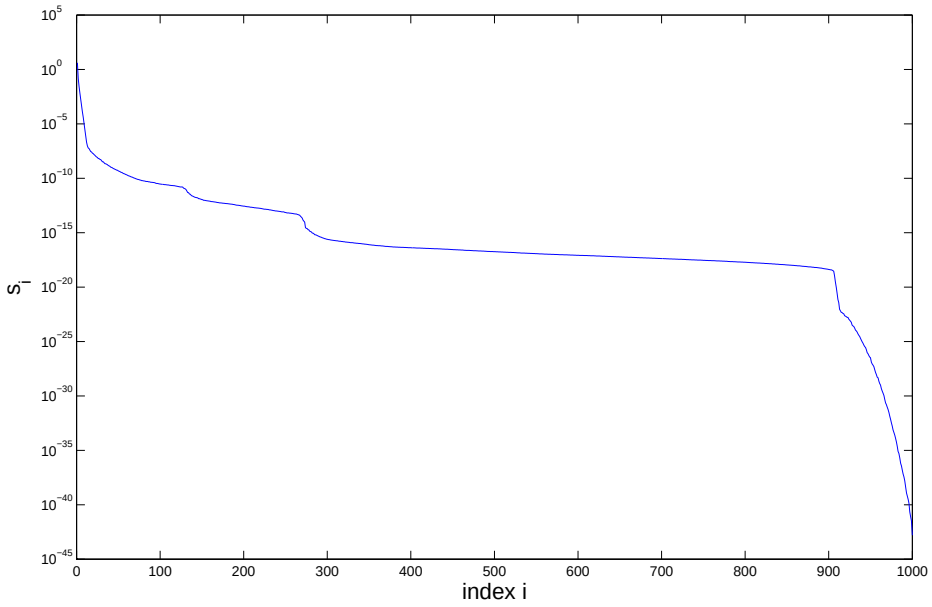

Figure 3.4: The singular value decreases to zero as i increases

of the results are observed:

- The singular values s_i decay rapidly to zero as i increases.
- The right singular vectors v_i tend to be more oscillatory as the index i increases, thus as s_i decreases. (This characteristic is possessed by the right singular vector u_i as well)

We identify problems with the above properties as a special class of problem in mathematics called *a discrete ill-posed problem*. Some of the most important literature for treating this problem are [48], [49] and [46]. To interpret these results using SVD, we notice that following relations can derived from Eq.(3.3):

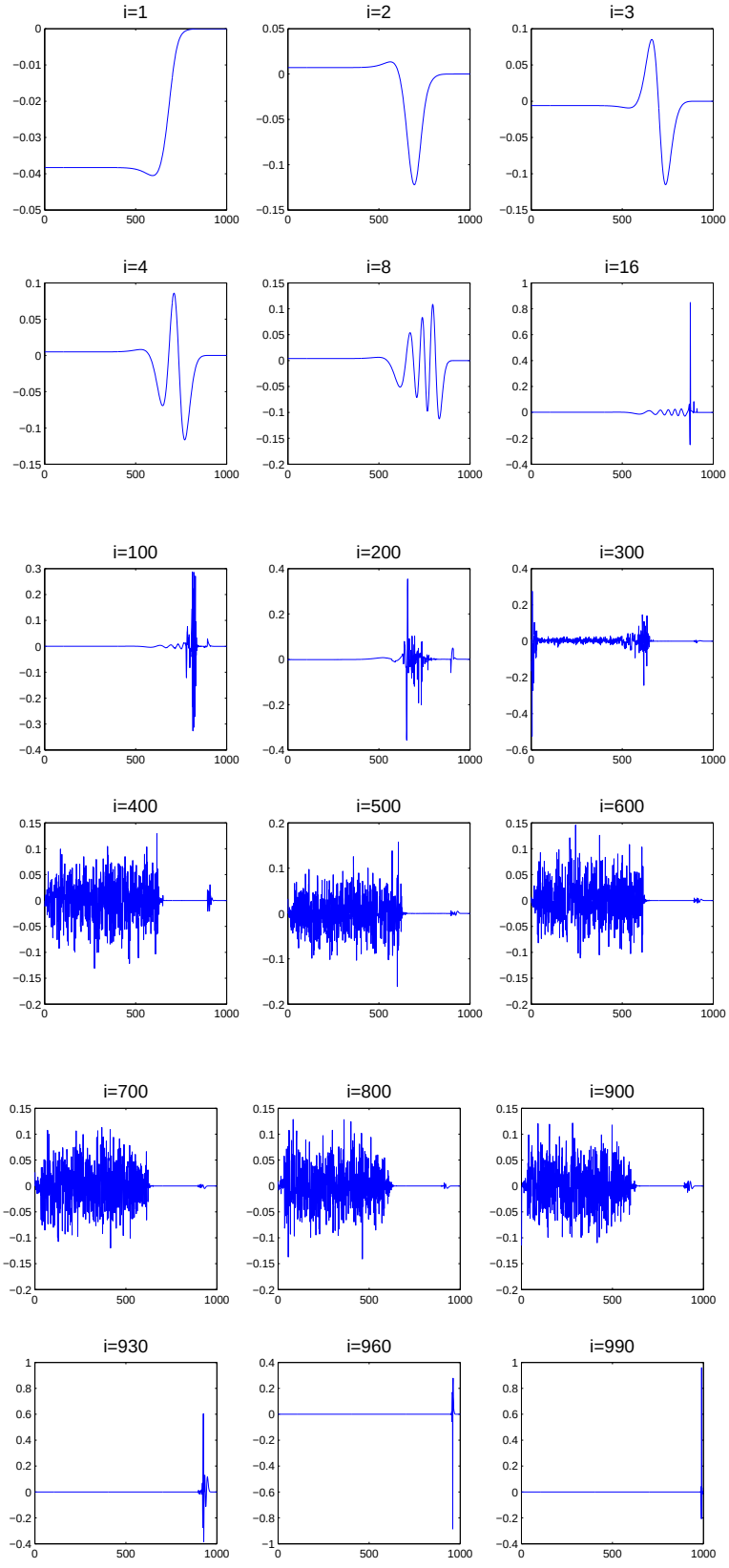

Figure 3.5: The frequency of oscillation of the singular vectors increases as index i increases.

$$Av_i = s_i u_i, \qquad ||Av_i||_2 = s_i \tag{3.5}$$

$$A^T u_i = s_i v_i, \qquad ||A^T u_i||_2 = s_i \tag{3.6}$$

where $i = 1, \ldots, n$, and $||Av_i||_2$ denotes the two-norm of the vector Av_i. The two-norm of a vector x is defined as:

$$||\, x||_2 = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{\frac{1}{2}} = \sqrt{x^* x}. \tag{3.7}$$

It can be seen from these relations that very small singular values s_i, as observed in Fig.3.4, imply that there exists certain linear combinations of the columns of A, characterized by the elements of the right singular vector v_i, such that $||Av_i||_2 = s_i$ is very close to zero. The same holds for u_i and rows of A. Thus the fact that many small singular values ($\leq 10^{-15}$) are observed for our matrix A implies many columns of A are almost linearly dependent (actually they are, in a numerical sense [48]). Thus our problem is underdetermined, and the solution is not unique. Actually it is more appropriate to state that it is the linear dependencies in columns of A that are manifested as the existence of many small singular values when subject to SVD.

The effect of small singular values, s_i, and the origin of extreme sensitivity to perturbations

We have noticed that as s_i decreases, the singular vectors u_i and v_i shown in Fig.3.5 become more and more oscillatory. Consider a mapping Ay of an arbitrary vector y. Using SVD

$$y = \sum_{i=1}^{n} (v_i^T y) v_i \quad Ay = \sum_{i=1}^{n} s_i (v_i^T y) u_i. \tag{3.8}$$

The first equation shows the vector y is expressed in the basis of v_i. As discussed above, when i increases, the singular value s_i decreases and the elements of the singular vector v_i have more and more frequent sign changes. Thus, from the second equation, we observe the projection of y onto the higher frequency components are more damped in Ay, due to multiplication of smaller s_i, than the lower-frequency

components. The inverse problem of computing y from $Ay = c$, has the opposite effect, i.e. it amplifies the high-frequency oscillations in c. Specifically, in our problem, the direct unfolding by inversion can be written as

$$\begin{aligned} x = A^{-1}b &= V\Sigma^{\dagger}U^{T}b \\ &= \sum_{i=1}^{n} v_i s_i^{-1} u_i^T\, b \\ &= \sum_{i=1}^{n} \frac{u_i^T b}{s_i} v_i. \end{aligned} \tag{3.9}$$

Hence, it is the division by the small singular values s_i that amplifies the high-frequency components of b. Consequently, a tiny error in b can be amplified dramatically by these divisions of small singular values. Thus it is the linear dependencies in the columns of A that are manifested as small singular values, ultimately causing extreme sensitivity to perturbations.

To summarise, multiplication by a matrix A in an ill posed problem generally has a "smoothing" effect, and the inverse generally has a"sharpening" effect.

The origin of the oscillatory and large norm characteristic of our solution

In the ideal case, the coefficient $\frac{u_i^T b}{s_i}$ in our solution (3.9) should decay gradually to zero, as a physical dipole cross-section should not be oscillatory, i.e. the coefficient $\frac{u_i^T b}{s_i}$ for high frequency components should be small. In fact, the dipole cross-section is assumed to be a monotonic function of the dipole size r. Therefore, as the singular values s_i decay gradually to zero, we would expect that the Fourier coefficient $|u_i^T b|$ decays to zero faster than the singular value s_i. However as shown in Fig.3.6, it levels off at about $i \gtrapprox 100$. This is because $b \neq b^{exact}$ but $b = b^{exact} + e$, where e is the perturbation which takes into account the errors from the measurement of the F_2 structure function, the discretization process, and the rounding errors of computation. Thus $|U_i^T b|$ is actually

$$u_i^T b = u_i^T b^{exact} + u_i^T e \tag{3.10}$$

the first term $u_i^T b^{exact}$ characterizes the physical solution and should decay gradually to zero faster than the singular value s_i, however the random error in the second term will not. Consequently the Fourier coefficient $u_i^T b$ will decay until the second term $u_i^T e$ becomes dominant compared to $u_i^T b^{exact}$.

Therefore, if we assume the perturbation e is unbiased and uncorrelated[1], then its covariance matrix is given by

$$\sigma_0^2 I_n$$

where σ_0 is the standard deviation of the error. The expected value of $||e||_2$ satisfies $\varepsilon(||e||_2^2) = n\sigma_0^2$ [48]. The expected value of the Fourier coefficient of e can be written as

$$\varepsilon(||u_i^T e||_2) = \sigma_0, \qquad i = 1, \ldots, n. \tag{3.11}$$

Thus the coefficient $|u_i^T b|$ levels off at $\approx \sigma_0$. Assuming this happens at $i = i_b$, then $|u_i^T b|$ will be completely dominated by $|u_i^T e|$ for $i > i_b$. The error level, σ_0,

[1]This will not strictly be the case for F_2, but that doesn't matter as we are only making a qualitative argument here

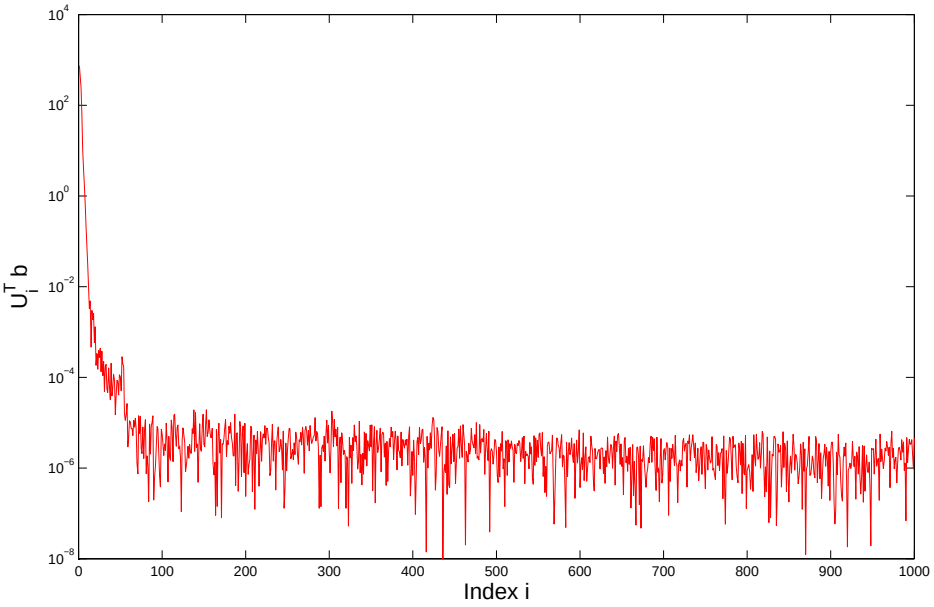

Figure 3.6: The coefficient $U_i^T b$ decreases as i increases. However, it settles at about 10^{-6} after $i \gtrapprox 100$

also determines how much information about the underlying behavior (b^{exact}) can be extracted from the given behavior of b. Mathematically, only those SVD components of the solution for which the error in $|u_i^T b|$ does not dominate can be recovered, i.e. the components of $|u_i^T b|$ for $i \leq i_b$.

3.1.3 Key strategies for solving the problem

To summarize, the main difficulty faced in our problem is that the underdetermined nature of A, manifest as the existence of many small singular values of matrix A, caused our solution x to be extremely sensitive to perturbations in b. An important consequence is that the matrix equation $b = Ax$ should not be solved exactly, as the solution x will be dominated by the errors in b. Furthermore, to remove the underdetermined nature and stabilize the solution, appropriate extra information about the desired solution has to be incorporated into the system. This is the essence of regularization. Thus, instead of solving

$$b = Ax$$

exactly, we solve this equation in a least squares sense

$$b \cong Ax.$$

This is equivalent to minimizing the two-norm of the vector $Ax - b$ defined as

$$||Ax - b||_2$$

which is sometimes called the residual norm of the solution. To regularize this problem, in addition to solving

$$\min \{||Ax - b||_2\} \tag{3.12}$$

we also want to constraint the behavior of x (e.g. smoothness, size). If we denote the regularized solution as x_{reg} and define $\Xi(x_{reg}) = ||Ax_{reg} - b||_2$, there are four general schemes for regularization [48]:

- Minimize $\Xi(x_{reg})$ subject to the constraint that x_{reg} belongs to a specific subset, $x_{reg} \in \mathbf{S}_{x_{reg}}$. e.g. Minimize $\Xi(x_{reg})$, subject to $x_{reg} \geq 0$.

- Minimize $\Xi(x_{reg})$ subject to the constraint that a measure, $\omega(x_{reg})$, of the "size" of x_{reg} is less than some specified upper bound δ, i.e. $\omega(x_{reg}) \leq \delta$.

- Minimize $\omega(x_{reg})$ subject to the constraint $\Xi(x_{reg}) \leq \alpha$.

- Minimize a linear combination of $\Xi^2(x_{reg})$ and $\omega^2(x_{reg})$:

$$\min\{\Xi^2(x_{reg}) + \lambda^2\omega^2(x_{reg})\} \tag{3.13}$$

where α, δ and λ are regularization parameters, and the function ω is often referred to as the 'discrete smoothing norm'. The last scheme is called Tikhonov regularization, when the $\omega(x_{reg})$ has the explicit form:

$$\omega(x_{reg}) = ||L\,x||_2 \tag{3.14}$$

this often measures the smoothness of the solution, the matrix L will be defined shortly. The requirement that the regularized solution be smooth can be translated into the requirement that the total squared derivative (of some order) be small. For example, the total squared first derivative can be calculated as

$$D_1 = \sum_{i=1}^{n-1}(x_i - x_{i+1})^2 \tag{3.15}$$

and in the case of second derivative:

$$D_2 = \sum_{i=2}^{n-1}[(x_{i+1} - x_i) - (x_i - x_{i-1})]^2. \tag{3.16}$$

Thus the matrix L is typically a matrix approximation of the derivative operator of x, e.g. a 2^{nd} derivative operator can be approximated by the following matrix:

$$L = \begin{pmatrix} -1+\xi & 1 & 0 & 0 & \dots & & & \\ 1 & -2+\xi & 1 & 0 & 0 & \dots & & \\ 0 & 1 & -2+\xi & 1 & 0 & 0 & \dots & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \\ & & & & \dots & 1 & -2+\xi & 1 \\ & & & & & \dots & 1 & -1+\xi \end{pmatrix}$$

where L is a banded matrix with full rank. To avoid degeneracy (i.e. every column and every row sums up to zero) and enable inversion of L, a number ξ, is added to every diagonal component. ξ is chosen such that it is large enough to enable the inversion, but small enough so that L should be a good approximation of the second derivative operator. Its value is chosen to be 0.0001 as recommended in [50]. Matrix L can also be an identity matrix I_n, in which case, $||L\,x||_2$ measures the two-norm or the "size" of x.

From the above discussion, we observe that the fourth scheme, the Tikhonov regularization, is the most suitable scheme for our problem. Because the dipole cross-section is indeed a smooth function with respect to varying dipole size r, thus a smoothness constraint incorporated on the regularized solution using Tikhonov method will be an appropriate physical constraint.

3.2 Regularization methods

In the following, Tikhonov regularization will be discussed in detail, and a brief survey of unfolding methods used in high energy physics will be presented in section 3.2.2

3.2.1 Tikhonov Regularization

Tikhonov regularization [51] as briefly discussed above is a type of direct (as opposed to iterative) regularization method. The key idea is to impose a smoothness constraint on the desired solution in the form of the discrete smoothing norm $||Lx||_2$, in order to damp the oscillatory behavior of the solution. It has the general form:

$$\min\{||Ax - b||_2^2 + \lambda^2 ||Lx||_2\} \tag{3.17}$$

where λ is the regularization parameter that controls how important is the smoothness of the regularized solution x_{reg}, relative to satisfying the least squares equation $Ax \cong b$. How to choose λ is a very important issue that is going to be discussed in later sections.

To solve (3.17) numerically, it needs to be first transformed into the standard form.

Transformation to standard form

If the matrix L in the discrete smoothing norm $\omega(x) = ||Lx||_2$ is the identity matrix I_n, then the regularization problem is said to be in standard form. The transformation is done by first writing the Tikhonov regularization in the following form:

$$\min\left\{\left|\left|\begin{pmatrix} A \\ \lambda L \end{pmatrix} x - \begin{pmatrix} b \\ 0 \end{pmatrix}\right|\right|_2^2\right\}$$

hence

$$\min\left\{\left|\left|\begin{pmatrix} AL^{-1} \\ \lambda \end{pmatrix} Lx - \begin{pmatrix} b \\ 0 \end{pmatrix}\right|\right|_2^2\right\}$$

finally

$$\min\left\{\left|\left|\begin{pmatrix} \hat{A} \\ \lambda \end{pmatrix} \hat{x} - \begin{pmatrix} b \\ 0 \end{pmatrix}\right|\right|_2^2\right\}$$

where

$$x = L^{-1}\hat{x} \qquad A = \hat{A}L \tag{3.18}$$

thus Tikhonov regularization in standard form is obtained:

$$\min\{||\hat{A}\hat{x} - b||_2^2 + \lambda^2 ||\hat{x}||_2^2\} \tag{3.19}$$

From the above, our problem has been transformed into the problem of solving for $\hat{x}$, which is the vector of second derivative of x. Thus our problem will be divided into two steps, first find the regularized the solution $\hat{x}$ and then transform it back to x.

Solving Tikhonov regularization using SVD of A

If we apply SVD to $\hat{A}$, the matrix equation $\hat{A}\hat{x} = b$ becomes

$$\begin{aligned} USV^T\hat{x} &= b \\ SV^T\hat{x} &= U^Tb. \end{aligned} \tag{3.20}$$

Hence minimizing the residual norm

$$||\hat{A}\hat{x} - b||_2^2$$

is equivalent to minimizing the norm

$$||SV^T\hat{x} - U^Tb||_2^2.$$

Thus Tikhonov regularization can be written as:

$$\min\{||SV^T\hat{x} - U^Tb||_2^2 + \lambda^2||\hat{x}||_2^2\}. \tag{3.21}$$

Writing the square of 2-norms in the form of an inner product gives

$$\min\{(SV^T\hat{x} - U^Tb)^T(SV^T\hat{x} - U^Tb) + \lambda^2\hat{x}^T\hat{x}\}. \tag{3.22}$$

The $\hat{x}$ in the second term can be written in the column basis v_i of V as

$$\hat{x} = \sum_{i=1}^{n}(v_i^T\hat{x})v_i \tag{3.23}$$

and because S is a diagonal matrix, (3.22) becomes

$$\min\{\sum_{i=1}^{n}(s_iv_i^T\hat{x} - u_i^Tb)^2 + \lambda^2[(v_i^T\hat{x})v_i]^2\}. \tag{3.24}$$

Expanding the first and second terms gives

$$\min\{\sum_{i=1}^{n} s_i^2(u_i^T b)^2 + s_i^2(v_i^T \hat{x})^2 - 2s_i(v_i^T \hat{x})(u_i^T b) + \lambda^2(v_i^T \hat{x})^2(v_i^T v_i)\}. \tag{3.25}$$

Notice that $v_i^T v_i = 1$ as V is an orthogonal matrix. To find the solution $\hat{x}$ that minimizes the function in the bracket, we set the first derivative of the function with respect to $v_i^T x$ equal to 0 which gives

$$\sum_{i=1}^{n} 2s_i^2(v_i^T \hat{x}) - 2s_i(u_i^T b) + 2\lambda^2(v_i^T \hat{x}) = 0. \tag{3.26}$$

Combining the $v_i^T \hat{x}$ terms gives

$$\sum_{i=1}^{n}(v_i^T \hat{x})(s_i^2 + \lambda^2) = \sum_{i=1}^{n} s_i(u_i^T b) \tag{3.27}$$

hence

$$\sum_{i=1}^{n} v_i^T \hat{x} = \sum_{i=1}^{n} \frac{s_i}{s_i^2 + \lambda^2}\, u_i^T b. \tag{3.28}$$

Finally

$$\begin{aligned} \hat{x} &= \sum_{i=1}^{n} \frac{s_i}{s_i^2 + \lambda^2} u_i^T b\, v_i \\ &= \sum_{i=1}^{n} \frac{s_i^2}{s_i^2 + \lambda^2} \frac{u_i^T b}{s_i}\, v_i \\ &= \sum_{i=1}^{n} f_i\, \frac{u_i^T b}{s_i}\, v_i \end{aligned} \tag{3.29}$$

where

$$x_{reg} = L^{-1}\hat{x} \tag{3.30}$$

is the solution to Tikhonov regularization. The f_i are called the filter factors and will be discussed shortly. The regularized solution can also be written in matrix form:

$$x_{reg} = L^{-1} V \Theta \Sigma^{\dagger} U^T\, b \tag{3.31}$$

where Θ is a $n \times n$ diagonal matrix: $\Theta = F = diag(f_i)$.

Filter factors

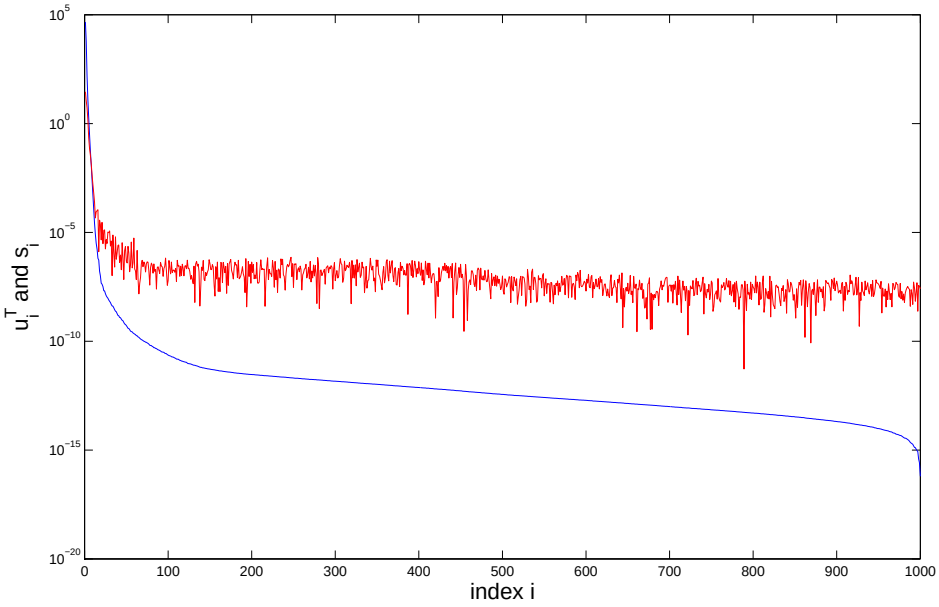

Figure 3.7: The Fourier coefficients $|U_i^T b|$ (red solid curve) and the singular values s_i (blue solid curve) as a function of index i. The y-axis is in logarithmic scale

From the above, the regularized solution can be explicitly written as

$$x_{reg} = L^{-1}\hat{x} = L^{-1}\sum_{i=1}^{n} f_i \frac{u_i^T b}{s_i}. \tag{3.32}$$

The Fourier coefficients $|u_i^T b|$ and the singular values s_i in the solution of $\hat{x}$, are plotted in Fig.3.7, which shows the characteristics of an ill-posed problem. The coefficients $|u_i^T b|$ of the solution are completely dominated by errors after a certain index $i > i_b$, as discussed in section 3.1.2. Thus the regularization of a discrete ill-posed problem is actually a matter of filtering out or damping the erroneous SVD components and extract the significant ones. Notice that the filter factors satisfy

$$f_i = \frac{s_i^2}{s_i^2 + \lambda^2} \quad \begin{cases} f_i \approx 1 & for \quad \lambda < s_i \\ f_i \ll 1 & for \quad \lambda > s_i \end{cases} \tag{3.33}$$

which is just designed to effectively filter, or damp, the SVD components for which $s_i < \lambda$.

3.2.2 A survey of unfolding methods used in high energy physics

The following four unfolding papers [50, 52–54], in the field of high energy physics are studied, the result will be summarized below.

In the discussion of the first two papers, the problem to be solved has the identical form to our problem:

$$y = Ax$$

where y is the measured distribution in the experiment, A represents the detector's response matrix and x is the true distribution which we want to extract.

Höcker and Kartvelishvili [50], took the following form of the expression to be minimized:

$$(Ax - b)^T(Ax - b) + \tau \cdot (Cx)^T Cx \tag{3.34}$$

which is very similar to our expression (3.22). Here, the matrix C is the second derivative matrix (3.1.3). The term $(Cx)^T Cx$, measures the squared sum of the second derivative quantifying the smoothness of the extracted true distribution. The solution is given by

$$x^\tau = C^{-1} \sum_{i=1}^{n} \frac{s_i^2}{s_i^2 + \tau} \frac{u_i^T b}{s_i} v_i \tag{3.35}$$

which has exactly the same form as (3.29), except that in our case $\hat{x}$ was converted to x in a later step by multiplying L^{-1}. They proposed to choose the regularization parameter as $\tau = s_{i_b}$, that is, the singular value at which the coefficient $|u_i^T b|$ becomes dominated by error. This will effectively damp the singular components for which $i > i_b$. As can be see, the method used in this paper is Tikhonov regularization with exactly the same form as we derived ealier.

In Blobel's approach [53], the expression

$$F(x) = -log\, L(x, y, A) + \tau \cdot C_B(x) \tag{3.36}$$

is minimized. Here τ is the regularization parameter, $-log\, L(x, y, A)$ is the negative logarithm of the total likelihood function with respect to the elements x_j of vector

x. If the measured data y obeys the Poisson distribution, $-log\ L(x, y, A)$ takes the following form:

$$-log\ L(x, y, A) = \sum_{i=1}^{n} (\hat{y}_i - y \cdot \ln \hat{y}_i) \tag{3.37}$$

where

$$\hat{y}_i = \sum_{j=1}^{m} a_{ij} x_j. \tag{3.38}$$

$\hat{y}$ is the vector of expected value of the measured distribution and a_{ij} is the matrix element of A. The function $C_B(x)$ measures the smoothness of the vector x and is the sum of the squared second derivative of x, which has the following form

$$C_B(x) = x^T C_B x. \tag{3.39}$$

Together with equation (3.34), we identify that matrix C_B is actually the matrix $C^T C$ where C is the matrix (3.1.3). The solution to (3.36) is given by

$$(x_j)_{reg} = \left(\frac{1}{1 + \tau \cdot S_j} \right) (x_j)_{unreg} \tag{3.40}$$

where S_j is the eigenvalue of the transformed[2] matrix C, S_j has the characteristic that it decays gradually to zero as j increases, like the singular values. $(x_j)_{unreg}$ is the solution without regularization, i.e. the solution obtained by purely minimizing the negative logarithm of the total likelihood function. The factors between the regularized and unregularized solutions are called 'fixing factors' and they act like filter factors. The regularization parameter is chosen as follows:

$$n_{df} = \sum_{j=1}^{m} \left(\frac{1}{1 + \tau \cdot S_{jj}} \right) \tag{3.41}$$

is the sum of all fixing factors called the 'effective degrees of freedom'. Blobel argues n_{df} should be equal to, or larger than, the number of significant terms, by which he means the terms that are not dominated by errors, i.e. $n_{df} \geq i_b$. Thus the contribution from terms with large errors will be damped by the fixing factors. Thus, again Tikhonov's method is used, where $-log\ L(x, y, A)$ is similar to our residual

[2] please see [53] for details of the transformations

norm $||Ax - b||_2$, and $C_B(x) = x^T C_B x$ represents exactly our discrete smoothing norm.

In Cowan's paper [54], the problem to be solved, takes the following form:

$$\nu = R\mu$$

where ν is the measured distribution in the experiment, R represents the detector's response matrix and μ is the true distribution. The expression to be minimized is

$$\Phi(\mu) = \alpha \ln L(\mu) + S(\mu) \tag{3.42}$$

where α is the regularization parameter, L is the log likelihood function as discussed previously, and $S(\mu)$ which measures the smoothness of the vector μ, is the sum of the squared second derivative of μ, this is similar to the function $C_B(x)$ in Blobel's paper. Cowan presented two methods for choosing the regularization parameter, one of them based on the estimation of the mean squared error averaged over the elements of the measured data, the other based on the calculation of the bias of the maximum likelihood estimator. As can be seen Cowan also used the Tikhonov's method. Furthermore, an iterative method proposed by D.Agostini [55] was also discussed in Cowan's paper.

In Barlow's paper [52], Tikhonov regularization is once again advocated. However, there is no advice on choosing the regularization parameter. Rather, another iterative method, the maximum Entropy method, is discussed in detail. This method is also discussed in [56]. An algorithm for implementing this method is available at [57].

As can be seen, all four papers have promoted Tikhonov regularization, however they each advocate different strategies for choosing the regularization parameter.

3.3 How regularization affects the solution

From the above brief review, we observe that Tikhonov regularization is a very popular method, and in section 3.2.1, we have derived its regularized solution, and

discussed the properties of filter factors. In the following, we will derive the error of the regularized solution and subsequently present its covariance matrix. After that, the most important graphical tool for analyzing regularization: the L-curve, will be discussed.

3.3.1 The Resolution matrix

The resolution matrix is an important tool, that characterizes the error from regularization. To derive it, we define the matrix of regularized inverse, $A^{\#}$, and use it to write the regularized solution as:

$$x_{reg} = A^{\#} b. \tag{3.43}$$

Thus from (3.31), the matrix $A^{\#}$ can be written as

$$A^{\#} = L^{-1} V F \Sigma^{\dagger} U^{T}. \tag{3.44}$$

Now the $n \times n$ resolution matrix ν can be defined as

$$\nu = A^{\#} A = L^{-1} V F V^{T} L. \tag{3.45}$$

Thus ν is symmetric. It quantifies the smoothing of the exact solution by the particular regularization method in the following way If

$$b = b^{exact} + e = A x^{exact} + e \tag{3.46}$$

then

$$x_{reg} = A^{\#} b = \nu x^{exact} + A^{\#} e \tag{3.47}$$

thus

$$x^{exact} - x_{reg} = (I_n - \nu) x^{exact} - A^{\#} e \tag{3.48}$$

where the term νx^{exact} in (3.47) is the regularized or "smoothed" version of the exact solution x^{exact}, while the second term $A^{\#} e$ is the contribution from the perturbation in b to x_{reg}. Thus $(I_n - \nu) x^{exact}$ is the pure regularization error in x_{reg}, which is characterized by the deviation of ν from I_n. In essence, ν describes how well

the exact solution x^{exact}, in the noise-free case, is approximated by the regularized solution $A^{\#}b^{exact}$.

Hansen [58] studied the upper bound of the pure regularization error of Tikhonov regularization. He proved mathematically that, for problems (e.g. our dipole cross-section) whose Fourier coefficient $|u_i^T b|$, in the absence of perturbation error, decay faster than the singular values s_i, there exist an appropriate regularization parameter λ, whose corresponding solution is guaranteed to have approximately the same properties as the exact solution.

The covariance matrix of the solution

In the statistical framework [59], if CC^T $(C = \sigma_0 I_m)$) is the covariance matrix for measurements, then the covariance matrix for x_{reg} is given by $Cov(x_{reg}) = A^{\#}CC^T(A^{\#})^T$ [48]:

$$Cov(x_{reg}) = \sigma_0^2 L^{-1} V F^2 (\Sigma^{\dagger})^2 V^T L. \tag{3.49}$$

This expression shows that the error in the regularized solution depends explicitly on filter factors in a way that more damping, or filtering (small f_i), corresponds to decreasing of the error in the regularized solution.

3.3.2 L-curve analysis

The most important graphical tool for analysis of regularization of discrete ill-posed problems is the so-called L-curve which is a plot of the discrete smoothing norm $\omega(x_{reg})$, i.e. $||Lx_{reg}||_2$ of the regularized solution versus the corresponding residual norm $||Ax_{reg} - b||_2$ for all valid regularization parameters (e.g. λ).

A typical L-curve is plotted on a log-log scale in Fig.3.8. As can be seen there is a distinct corner separating the vertical and horizontal parts, the residual norm $||Ax - b||$ and the smoothing norm $||Lx||$ are monotonic functions of the regularization parameter λ, such that as λ increases, $||Ax - b||$ always increases and $||Lx||$ always decreases. Thus, the L-curve displays the 'trade off' between minimizing these two quantities. This demonstrates the questions at the heart of regularization, i.e. how

much regularization does our problem need? And ultimately, how accurately can a regularized solution approximate the exact solution? The first question concerning how to choose the regularization parameter λ will be discussed in detail in section 3.4. Here we will present a qualitative analysis of the second question: How accurately x_{reg} approximates the exact solution. This can be quantified by using (3.47) and (3.48). We obtain:

$$x^{exact} - x_{reg} = (x^{exact} - A^{\#}b^{exact}) - A^{\#}e \tag{3.50}$$

$$= \sum_{i=1}^{n}(1 - f_i)\frac{u_i^T b^{exact}}{s_i}v_i - \sum_{i=1}^{n} f_i \frac{u_i^T e}{s_i}v_i \tag{3.51}$$

$$= \sum_{i=1}^{n}\left(1 - \frac{s_i^2}{s_i^2 + \lambda^2}\right)\frac{u_i^T b^{exact}}{s_i}v_i - \sum_{i=1}^{n}\frac{s_i^2}{s_i^2 + \lambda^2}\frac{u_i^T e}{s_i}v_i. \tag{3.52}$$

As can be seen, the regularization parameter λ is the key factor affecting the accuracy of the regularized solution. Qualitatively, in the limiting case when λ is very small, which corresponds to very little regularization, filter factors $f_i \approx 1$ for most SVD components, and the error $x^{exact} - x_{reg}$ is dominated by the perturbation error $A^{\#}e$. This situation is called 'undersmooth', and normally corresponds to the part of the

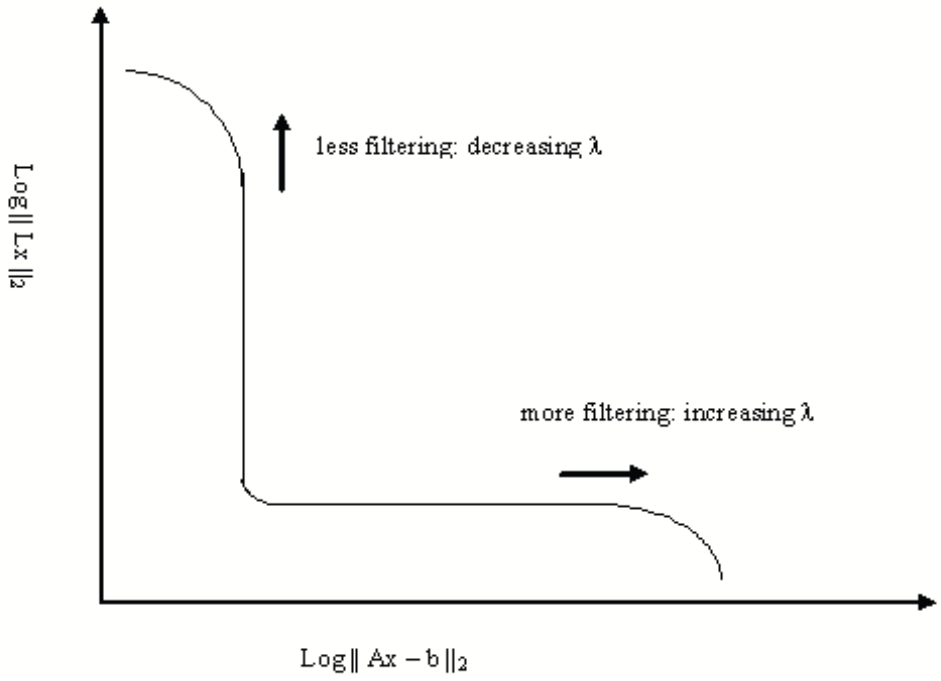

Figure 3.8: A typical L-curve in log-log scale

L-curve to the left of the corner. When λ is very large, which corresponds to a large amount of regularization and $f_i \ll 1$ for most SVD components, in which case the error $x^{exact} - x_{reg}$ is dominated by the regularization error $x^{exact} - A^{\#}b^{exact}$. This is called 'oversmooth', and corresponds to the part of the L-curve to the right of the corner. The L-curve method for choosing the regularization parameter will be discussed further in section 3.4.3.

3.4 Regularization parameters

As discussed above, choosing the most appropriate regularization parameter λ is crucial for obtaining an accurate regularized solution. From a mathematical point of view, different methods of choosing the regularization parameter λ can be analyzed by studying how fast the regularized solutions converge to the "true" solution as the error in the measurement b tends to zero, the methods that produce faster convergence being desired. This approach will not be presented here, however we refer the reader to the literature [60]. Alternatively, we will present two practical approaches, one called the discrepancy principle, which chooses λ based on the estimation of the error norm $||e||_2$ and the other based on the L-curve analysis presented in the previous section. Before we start discussing these parameter choice methods, we want to demonstrate the characteristic of the solution for various regularization parameters λ.

3.4.1 The regularization parameter dependence of the solution

To show how the extracted dipole cross-section and its predicted structure function change with regularization parameter we present two figures, each showing the extracted dipole cross-section with a particular range of regularization parameter λ. We also show their corresponding two tables with the specific values of λ and the χ^2 per degrees of freedom for their predicted F_2 structure function. Note that for small enough values of λ the χ^2 per point can be very small. This is mainly because

λ	χ^2/dof	λ	χ^2/dof	λ	χ^2/dof
2×10^{-5}	1.459×10^{-7}	3.2×10^{-4}	3.912×10^{-6}	5.12×10^{-3}	1.28×10^{-4}
4×10^{-5}	3.174×10^{-7}	6.4×10^{-4}	9.324×10^{-6}	1.024×10^{-2}	3.36×10^{-4}
8×10^{-5}	7.265×10^{-7}	1.28×10^{-3}	2.119×10^{-5}	2.048×10^{-2}	8.403×10^{-4}
1.6×10^{-4}	1.733×10^{-6}	2.56×10^{-3}	5.255×10^{-5}	4.096×10^{-2}	1.953×10^{-3}

Table 3.1: The value of regularization parameter λ and the corresponding χ^2 per degree of freedom for Fig.3.9.

the data are not real data but pseudo-data. A more realistic and important development of this analysis would be to move the central value of the pseudo-data by a random amount in a manner consistent with the known error. This would not only stop the χ^2 from becoming too small, it would also provide an important test of the robustness of our unfolding. In the next few sections we will discuss suggested ways to fix λ.

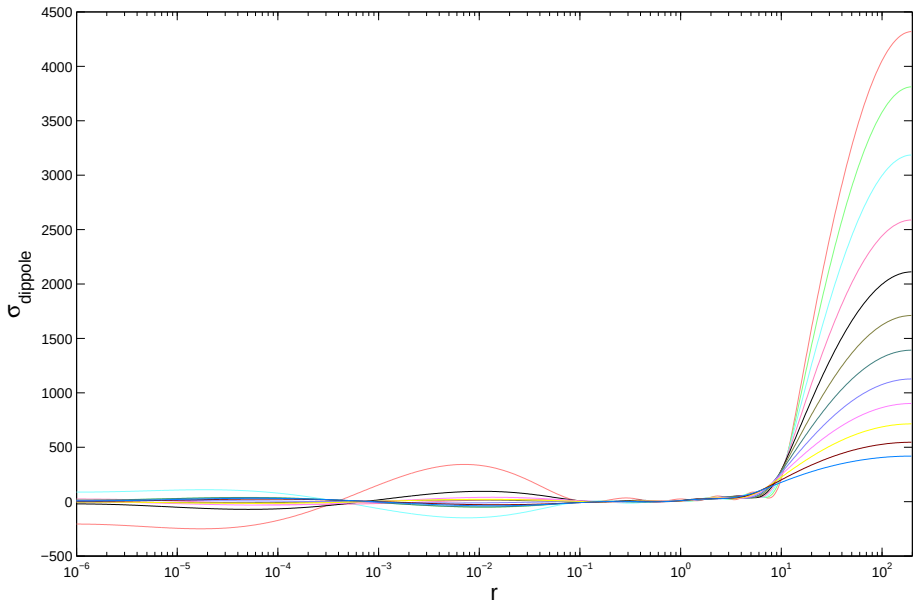

Figure 3.9: The extracted dipole cross-section for "relatively small" values of regularization parameters λ, their specific value is shown in table 3.1. The maximum value of the dipole cross-section decreases as λ increases.

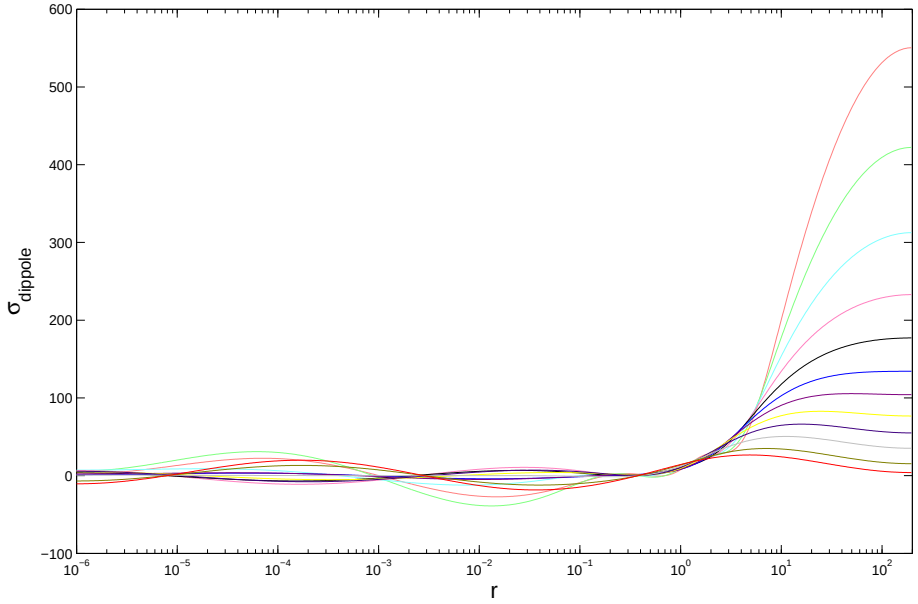

Figure 3.10: The extracted dipole cross-section for "relatively large" values of regularization parameters λ, their specific value is shown in table 3.2. The maximum value of the dipole cross-section decreases as λ increases.

λ	χ^2/dof	λ	χ^2/dof	λ	χ^2/dof
2×10^{-2}	8.168×10^{-4}	3.2×10^{-1}	1.831×10^{-2}	5.12	5.609×10^{-1}
4×10^{-2}	1.895×10^{-3}	6.4×10^{-1}	3.683×10^{-2}	1.024×10^{1}	1.927
8×10^{-2}	4.489×10^{-3}	1.28	7.980×10^{-2}	2.048×10^{1}	5.228
1.6×10^{-1}	8.978×10^{-3}	2.56	2.066×10^{-1}	4.096×10^{1}	9.236

Table 3.2: The value of regularization parameter λ and the corresponding χ^2 per degree of freedom for Fig.3.10

3.4.2 The discrepancy principle

As discussed in section 3.3.2, the residual norm $||Ax - b||$ and the smoothing norm $||Lx||$ are monotonic functions of the regularization parameter λ. Thus we can fix λ by fixing these residual norms, and this is the idea behind the discrepancy principle. In 1966, Morozov [61] proposed that when a good estimate of $||e||_2$ is available, and the relation $Ax^{exact} = b^{exact}$ is exact, then the regularization parameter λ can be chosen such that the residual norm is equal to an upper bound δ_e for $||e||_2$:

$$||Ax_\lambda - b||_2 = \delta_e \qquad \text{where} \qquad ||e||_2 \leq \delta_e. \tag{3.53}$$

However, studies [62] and [63], show that this method often leads an oversmooth solution. To counter this, a compensated discrepancy principle [64] is introduced, which compensates for the fact that although $\sigma_0\sqrt{n}$ (where n is number of entries) is a valid estimate for $||e||_2$ neither $\sigma_0\sqrt{n}$ nor $||e||_2$ may be a valid estimate of the residual norm $||Ax_\lambda - b||$. It takes the following form:

$$\begin{aligned} ||Ax_{reg} - b|| &= (||e||_2^2 - \sigma_0^2 \text{Trace}(AA^\#))^{\frac{1}{2}} \\ &= (\sigma_0^2 m - \sigma_0^2 \text{trace}(AA^\#))^{\frac{1}{2}} \end{aligned} \tag{3.54}$$

$A^\#$ is defined in equation (3.44).

3.4.3 The L-curve method

As discussed in section 3.3.2, the regularization parameter λ controls the relative contribution to x_{reg} from the measurement error and the regularization error. The key idea behind the L-curve method is that a regularization parameter should be chosen so that the errors from these two sources are well balanced. Hansen [65] found that this optimal regularization parameter is not far from the the regularization parameter that corresponds to the L-curve's corner. Thus a good approximation to the optimal regularization parameter can be obtained by computing the corner of the L-curve. Mathematically, it is defined as the point on the curve [48]:

$$(\zeta(\lambda), \mu(\lambda)) = (\log||Ax_{reg} - b||, \log \omega(x_{reg})) \tag{3.55}$$

that has maximum curvature, where the curvature k is defined as

$$k(\lambda) = \frac{\zeta' \mu'' - \zeta'' \mu'}{[(\zeta')^2 + (\mu\prime)^2)]^{\frac{3}{2}}}. \tag{3.56}$$

Thus the regularization λ is chosen to maximize the curvature of the L-curve.

In addition to the methods discussed above, there are other important methods such as the 'generalized cross-validation method' [66], and the 'quasi-optimality criterion' [67], due to limited time we do not discuss these here.

From these discussions, we can see that one advantage of Tikhonov regularization is that many other regularization schemes can be realized by choosing the regularization parameter with different philosophies. For example, the second regularization scheme is equivalent to choose the largest λ such that $||Lx||_2 \leq \delta$, and the third scheme can be realized as choosing the smallest λ, such that $\Xi(x_{reg}) \leq \alpha$.

3.4.4 Choosing λ: Blobel's method

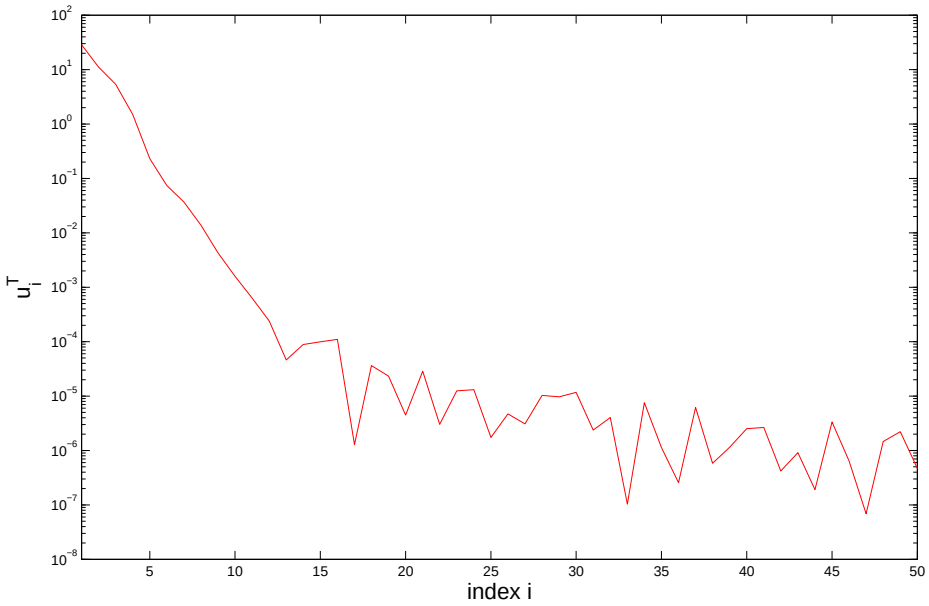

Figure 3.11: The Fourier coefficient $|u_i^T b|$ is plotted against the index i

Blobel [53], proposed a method for choosing λ based on a good estimation of the number of significant terms. He explicitly used the plot of $|u_i^T b|$ against index i as his guidance. It is stated that "it is clearly seen on the plot, as the value of i where the behavior d_i ($|u_i^T b|$) changes from exponentially falling to a constant". To be able to observe this behavior clearly, figure 3.7 is plotted for index i from 1 to 50 in Fig.3.11. As can be seen, the Fourier coefficient stops falling exponentially for $i > 13$, which corresponds to following value of λ

$$\lambda = s_{13} \approx 8.375 \times 10^{-6}.$$

The result computed by matlab based on this estimation is shown in Fig.3.12. As can be seen, the solution still possess a large norm and oscillating characteristic, which corresponds to a typical undersmooth solution, and implies that the regularization parameter is too small.

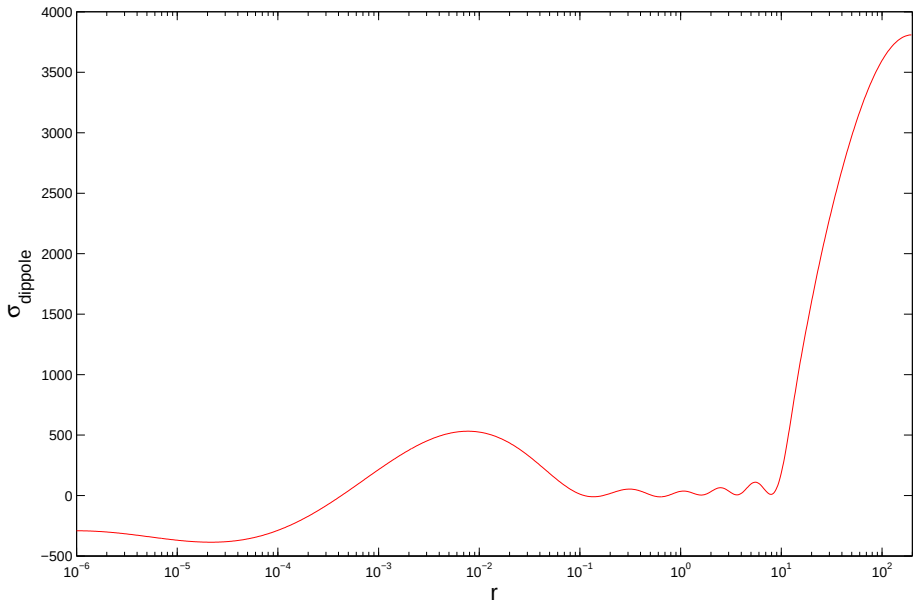

Figure 3.12: The extracted dipole cross-section (red curve) based on Blobel's parameter choice method at Bjorken $x = 0.0004$.

3.4.5 Choosing λ: The discrepancy principle

Based on the estimation of the error norm $||e||_2 = 1.863$, the regularization parameter is chosen to satisfy equation (3.53):

$$||Ax_\lambda - b||_2 = \delta_e \qquad \text{where} \qquad ||e||_2 \leq \delta_e.$$

This equation is identical to a least squares problem with a quadratic constraint. Algorithms available for solving this are [57, 68, 69]. The matlab function 'discrep' in [57] is used, which computes the regularization parameter λ using the singular value decomposed matrices U, Σ and V of A together with the value of error norm $||e||_2$ and structure function data F_2 in vector b. λ is computed to be

$$\lambda \approx 22.41.$$

The extracted dipole cross-section is computed using Matlab and is shown in Fig.3.13, also shown is the GW dipole model which has a χ^2/dof of ≈ 1.5. As can be seen the

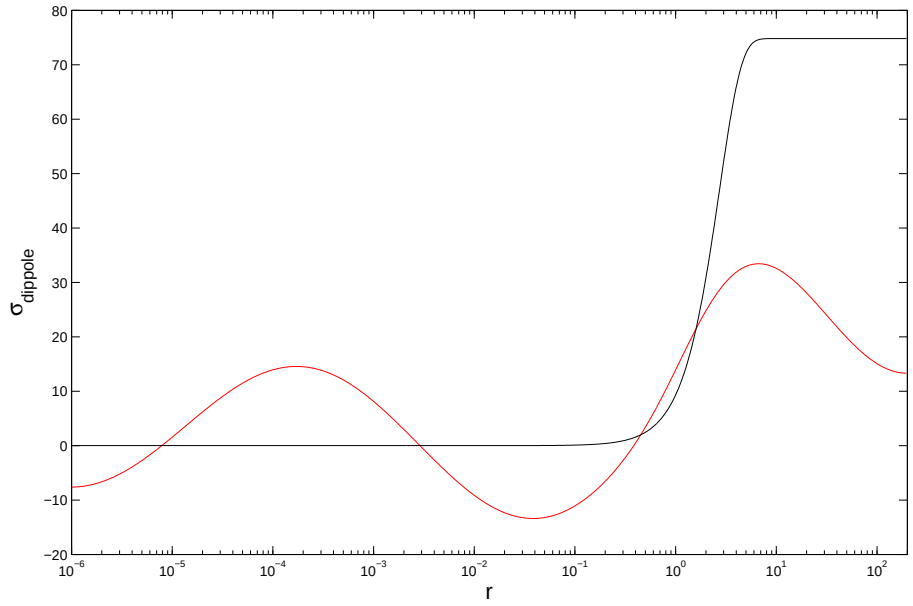

Figure 3.13: The extracted dipole cross-section (red solid curve) based on the discrepancy principle at Bjorken $x = 0.0004$. The GW model (black solid curve), is shown for comparison at the same Bjorken x.

solution still possess oscillatory behavior, but with much smaller frequency and has much smaller norm than the result obtained from Blobel's method. However, there are significant negative components in the solution.

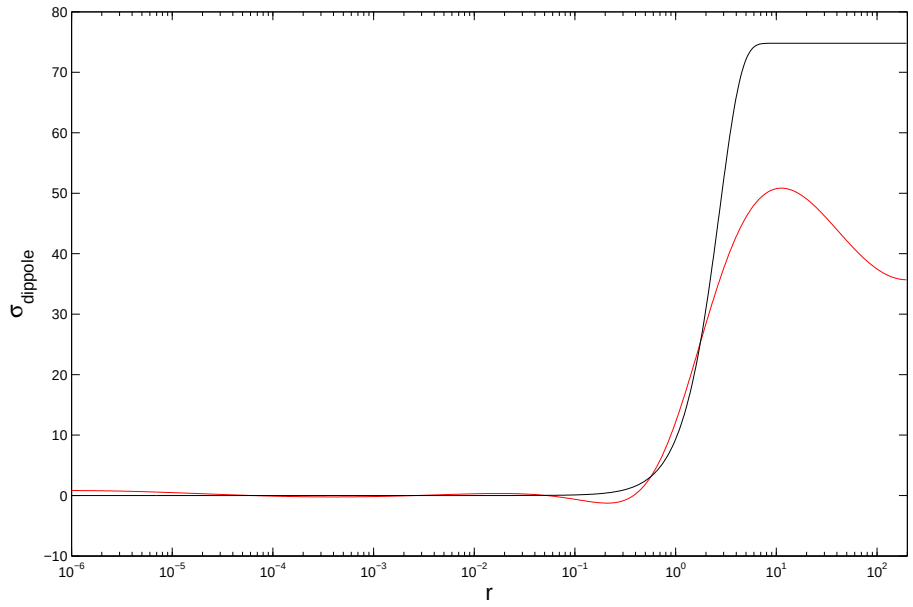

Figure 3.14: The extracted dipole cross-section (red solid curve) based on the compensated discrepancy principle at Bjorken $x = 0.0004$. The GW model (black solid curve), is shown for comparison at the same Bjorken x.

Using the compensated discrepancy principle where the regularization parameter is chosen to satisfy equation (3.54):

$$||Ax_{reg} - b|| = (||e||_2^2 - \sigma_0^2 \text{trace}(AA^{\#}))^{\frac{1}{2}}$$

the regularization parameter is computed using matlab function 'discrep' in [57] to be

$$\lambda \approx 10.074.$$

The result of regularization based on this estimation is shown in Fig.3.14. As can be seen, the solution is a lot more stable than previous solutions, with only mild oscillatory behavior, and the smallest negative components of all obtained solutions.

3.4.6 Choosing λ: The L-curve method

The regularization parameter that minimizes the curvature of the L-curve can be computed using the matlab function 'lcorner' in [57], which computes the regularization parameter λ using the singular value decomposed matrices U, Σ and V of A together with the structure function data F_2 in b. However, Fig.3.15 illustrates that the point of maximum curvature is not clearly defined for our problem and we therefore do not use the L-curve method. Vogel has done a comprehensive analysis on the possible cause of failure for L-curve method in [70].

3.4.7 Choosing λ: The physical choice

As can be seen, all the solutions obtained above, regardless of parameter choice methods, yield negative components in the solution, this is partly due to the oscillatory nature of the solution. Based on the fact that the dipole cross-section must be positive, a method can be devised for choosing the regularization parameter λ

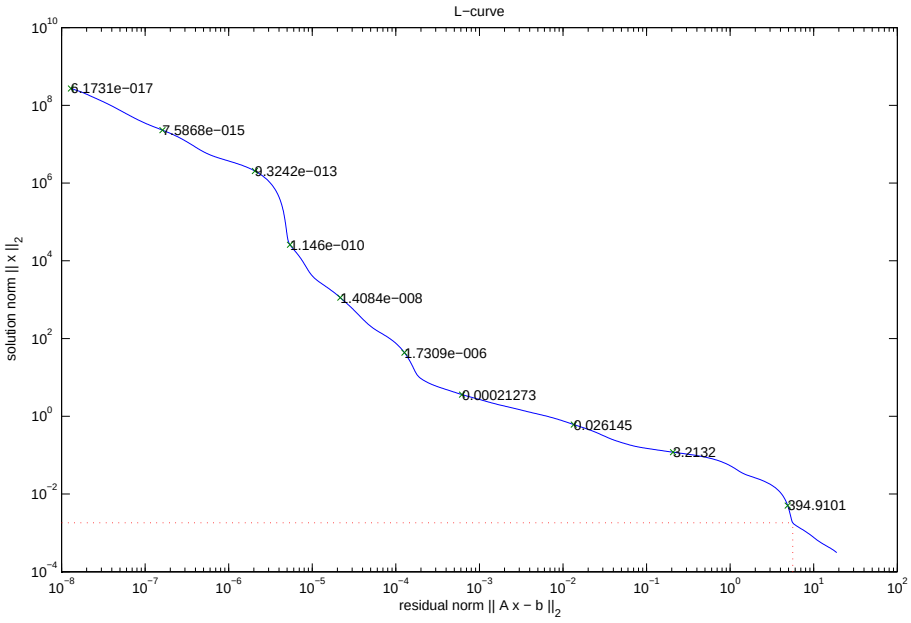

Figure 3.15: The L-curve is plotted on a log-log scale, formally the corner of this L-curve is at the intersection of the two red lines, and the numbers on the curve are the values of the regularization parameters λ.

such that it minimizes the sum of the negative components in the solution:

$$\min = \sum_{j=1}^{l} -x_j \quad \text{where} \quad x_j < 0 \tag{3.57}$$

A vector of $100,000$ entries was computed with each entry containing the sum of the negative components of x corresponding to $\lambda = 0.001, 0.002, \ldots, 99.999, 100$ respectively. We only computed the $\lambda > 0.001$ and $\lambda < 100$, because for $\lambda < 10.1$, the sum of negative components was observed to increase as λ becomes smaller, and for $\lambda > 10.1$, the sum of negative components was observed to increase as λ becomes larger. The regularization parameter λ which corresponds to smallest sum of negative component is then obtained to be

$$\lambda = 10.1.$$

The extracted dipole cross-section corresponding to such a choice of parameter is shown in Fig.3.16. As can be seen the solution is similar to that obtained from the compensated discrepancy principle.

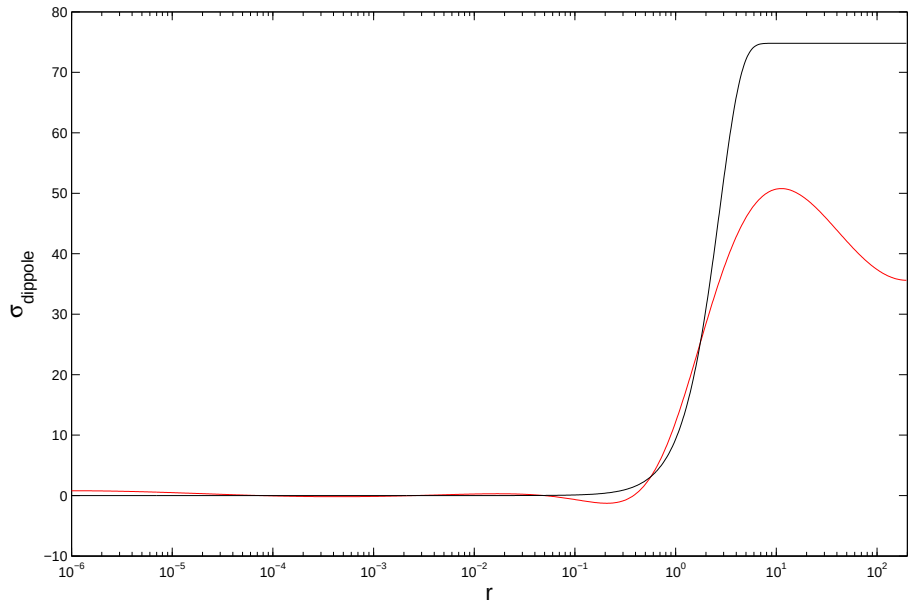

Figure 3.16: The extracted dipole cross-section (red solid curve) based on the physical constraint for choosing λ, at the Bjorken $x = 0.0004$. The GW model (black solid curve), is shown for comparison at the same Bjorken x.

3.4.8 Discussion

As can be seen from the above, the solutions are dramatically different for different regularization parameters λ, and solutions obtained from the compensated discrepancy principle and the physical choice are closest to a physical dipole cross-section.

There are clearly two undesirable features in our regularized solutions. Firstly, all solutions obtained, regardless of parameter choice method, have negative components whereas the physical dipole cross-section should be strictly positive. Secondly, the solution usually possesses a mild oscillatory pattern, whereas the physical dipole cross section should probably be a monotonic function with increasing dipole size r. Moreover, it is likely that the first problem of negativity of the solution is caused by its oscillatory behavior. Thus to extract the dipole cross-section more accurately, we propose to incorporate another constraint on the solution, namely that the extracted dipole cross-section should be a monotonic function of dipole size r. Consequently, our problem now becomes to solve

$$\min\{||Ax - b||_2^2 + \lambda^2 ||Lx||_2\} \tag{3.58}$$

subject to

$$Gx \geq 0 \tag{3.59}$$

where the $(n-1) \times n$ matrix G, is the matrix approximation to the first derivative operator:

$$G = \begin{pmatrix} -1+\xi & 1 & 0 & 0 & \dots & & & \\ 0 & -1+\xi & 1 & 0 & \dots & & & \\ 0 & 0 & -1+\xi & 1 & \dots & & & \\ \vdots & \vdots & \vdots & \vdots & \ddots & & & \\ & & & & \dots & -1+\xi & 1 & 0 \\ & & & & & \dots & -1+\xi & 1 \end{pmatrix}$$

G is a banded matrix with full row rank. And the number ξ (with typical value $\xi \approx 0.0001$) is added to the diagonal elements, for the same reason as in the case of L. We identify this problem as belonging to a special class of problems in mathematics

called *solving (regularized) least squares problems with linear inequality constraints*. Two of the most important literatures for treating this problem are [71] and [72].

3.5 Incorporating the "monotonicity" constraint

3.5.1 Method

We have surveyed the methods [73–76] for solving our new problem, and found that [74] only considered the case where the matrix A in equation (3.58) is of full rank. The algorithm described in [73] is not completely general in the sense that it assumed the matrix G in equation (3.59) to have full rank. The method described in [76] though is a general iterative method that can be applied to situations where A and G are both rank-deficient, however it requires a prior input of a "guess" of the extracted dipole cross-section. Thus, we choose to use the method by Haskell and Hanson [75], which specializes in solving ill-conditioned problems (where A and G can be rank deficient), and which does not require any prior input of the solution. It solves the linear least squares problem with both equality and inequality constraints (LSEI) of the following form:

$$Ex = f, \tag{3.60}$$

$$\tilde{A}x \cong \tilde{b}, \tag{3.61}$$

$$Gx \geq h. \tag{3.62}$$

In our case, $E = 0$, $f = 0$, as the linear equality constraint is not needed. Equation (3.61) contains our regularized linear least squares system, with

$$\tilde{A} = \begin{pmatrix} A \\ \lambda L \end{pmatrix}, \quad \tilde{b} = \begin{pmatrix} b \\ 0 \end{pmatrix} \tag{3.63}$$

and equation (3.62) corresponds to our equation (3.59), where $h = 0$ and G is the matrix described in the last section. It is by this equation that the extracted dipole cross-section is constrained to increase monotonically as the dipole size, r, increases. Thus our problem takes the following form:

$$\tilde{A}x \cong \tilde{b}, \tag{3.64}$$

$$Gx \geq h. \tag{3.65}$$

The idea behind Haskell and Hanson's method is that the LSEI problem can be transformed into another class of problem called the 'Non-Negativity constrained Linear least Squares problem with Equality constraints (NNLSE)' of the following form:

$$Ex = f, \tag{3.66}$$

$$Ax \cong b, \tag{3.67}$$

$$x_i \geq 0, \quad i = l+1, \ldots, n, \quad 0 \leq l \leq n. \tag{3.68}$$

Hanson and Haskell illustrated three different approaches to this transformation in their paper. Here one of the methods, which uses slack variables, is illustrated. An n dimensional vector, w, of nonnegative variables is introduced into the inequality constraints of (3.62), so that they become equality constraints, $Gx - w = h$, because requiring

$$Gx > h$$

is equivalent to requiring

$$Gx - w = h \quad \text{for} \quad w_i \geq 0, \quad i = 1, \ldots, n.$$

Thus our problem can now be written as

$$\tilde{A}x \cong \tilde{b}, \tag{3.69}$$

$$Gx - w = 0, \tag{3.70}$$

$$w \geq 0. \tag{3.71}$$

Subsequently this problem can be transformed into the problem of NNLSE:

$$\bar{E}\bar{x} = \bar{f}, \tag{3.72}$$

$$\bar{A}\bar{x} \cong \tilde{b}, \tag{3.73}$$

$$w \geq 0 \tag{3.74}$$

with the following choices of matrices

$$\bar{E} \equiv [G : -I], \tag{3.75}$$

$$\bar{f} \equiv [0], \tag{3.76}$$

$$\bar{A} \equiv [\tilde{A} : 0], \tag{3.77}$$

$$\bar{x} \equiv \begin{bmatrix} x \\ w \end{bmatrix}. \tag{3.78}$$

In $[G : -I]$ and $[\tilde{A} : 0]$, the symbol ":" means to augment the two matrices horizontally. I is an $(n-1) \times n$ identity matrix[3], $\tilde{A}$ is given by equation (3.63). The problem NNLSE can now be treated by solving the differentially weighted least squares problem [77]:

$$\min\{||\bar{E}\bar{x} - \bar{f}||_2^2 + \epsilon||\bar{A}\bar{x} - \tilde{b}||_2^2\} \quad \text{subject to} \quad \bar{x} \geq 0. \tag{3.79}$$

The least squares equations are each weighted by a small parameter ϵ, relative to the equality constraint.

3.5.2 Result

Firstly, the regularization parameter λ in $\bar{A}$ from $\tilde{A}$ needs to be chosen (see Eq.(3.63)). According to Eldén [78], it is more difficult to choose the regularization parameter λ when there is a linear inequality constraint, as the mapping of b into x is now nonlinear. For simplicity, we followed Hanson and Phillips' approach [79], i.e. we choose the regularization parameter λ without the linear inequality constraint first,

[3]All entries of the last column are set to equal to zero

and then use this λ for the linear constraint case. Thus, we use the regularization parameter

$$\lambda \approx 10.074$$

obtained in last section from the compensated discrepancy method. A more sophisticated review for choosing the regularization parameter in the presence of a linear inequality constraint is available in [80].

The equation (3.79) is solved using the Fortran subroutine DWNNLS [81] (also available from the Slactec program library [82]), written by Haskell and Hanson, which computes the solution $\bar{x}$, using $\bar{A}$, $\bar{E}$, $\bar{f}$ and $\tilde{b}$. It chooses the parameter ϵ to be

$$\epsilon = \sqrt{\frac{10^{-4}\eta}{\gamma}} \tag{3.80}$$

where γ is

$$\gamma = \left|\left|\begin{pmatrix} \bar{E} \\ \bar{A} \end{pmatrix}\right|\right|$$

$||\cdot||$ is the subordinate matrix norm of l_∞ vector norm [47]. The subordinate matrix norm of l_∞ vector norm for a $n \times n$ matrix P is defined as:

$$||P|| = \max_{\substack{1 \le i \le n, \\ 1 \le j \le n}} (A_{ij}) \tag{3.81}$$

and η is the machine relative arithmetic precision of the computer performing the minimization.

The extracted dipole cross-section with both "smoothness" and "monotonicity" constraints is shown in Fig.3.17. As can be seen, the extracted dipole cross-section is now strictly positive, and "almost" monotonic (though it can't be seen clearly from the graph, the value of dipole cross-section drops from 54.941 to 54.885 as the dipole size r increases from about 40 to 200). The resulting prediction for the F_2 structure function is shown in Fig.3.18.

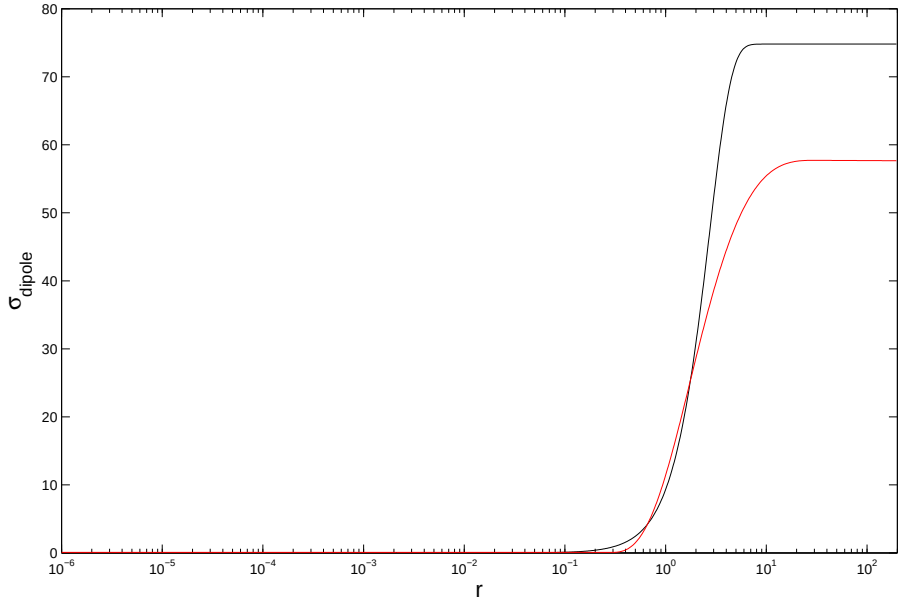

Figure 3.17: The extracted dipole cross-section (red solid curve) at Bjorken $x = 0.0004$, with "smoothness" and "monotonicity" constraint. The GW model (black solid curve), is shown for comparison at the same Bjorken x.

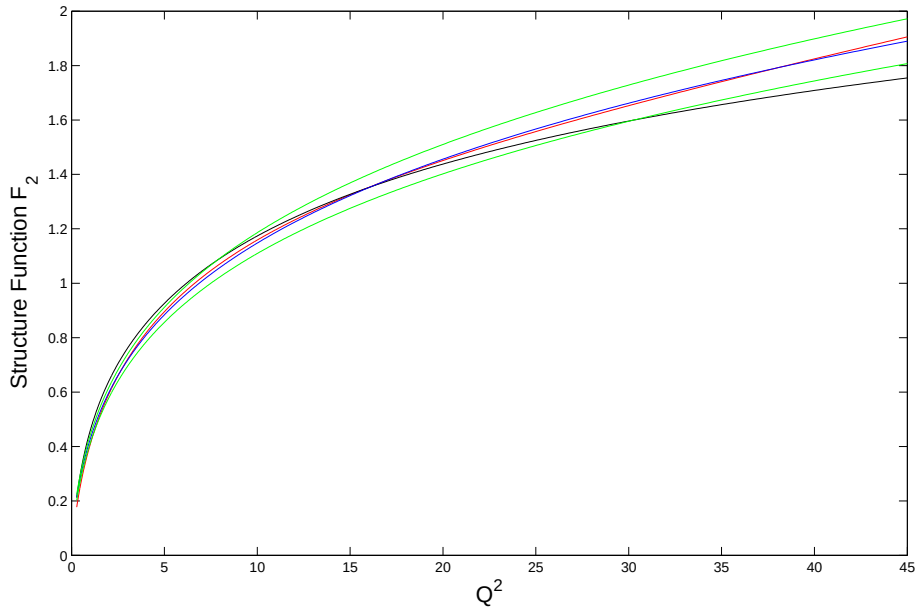

Figure 3.18: The predicted F_2 structure function from the extracted dipole cross section (red solid curve) and the F_2 pseudodata (blue solid curve) with error (green solid curve). The prediction from the GW model is also shown (black solid curve).

To show how the extracted dipole cross-section and its predicted structure function change with regularization parameter in the linearly constrained case, we again present a plot of extracted dipole cross-section with a range of regularization parameter λ. We also show its corresponding table with the specific values of λ and the χ^2 per degrees of freedom for their predicted F_2 structure function.

Note that the χ^2 values we quote ought not to be taken too seriously sine they related to the pseudo-data rather than the real data and we have not been particularly careful to use the pseudo-data only in the regions where there is also actual data. Of course this is not a problem for this thesis, since its main purpose is to establish the feasibility of unfolding a dipole cross-section. Future work is needed to make an unfolding with reliable errors.

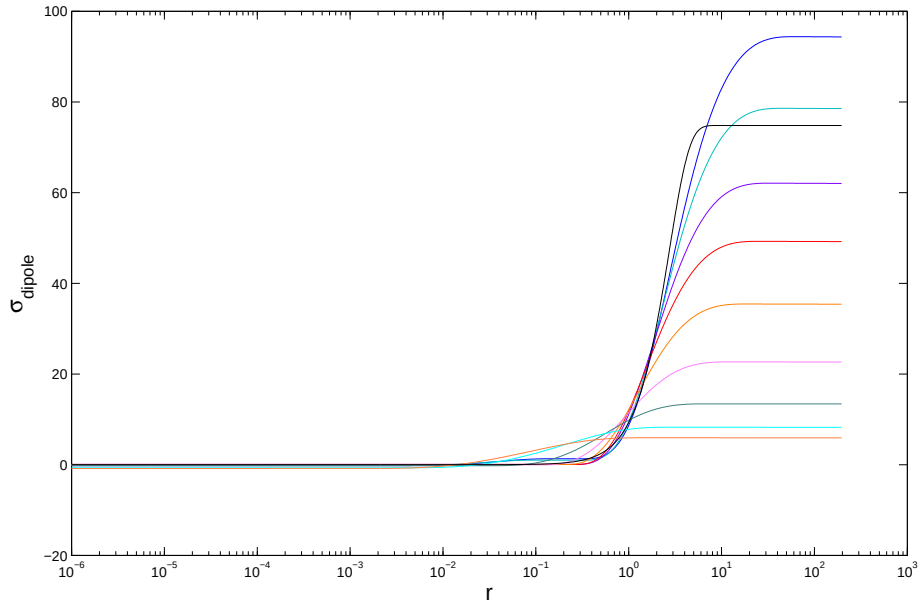

Figure 3.19: The extracted dipole cross-section with both 'monotonicity' and 'smoothness' constraint for a range of regularization parameters λ. The maximum value of the dipole cross-section decreases as λ increases.

3.5.3 Discussion

By imposing the "smoothness" and "monotonicity" constraints, we have successfully extracted a physical dipole cross-section and its prediction of the structure function

λ	χ^2/dof	λ	χ^2/dof	λ	χ^2/dof
2	0.14	16	3.252	128	48.34
4	0.373	32	9.871	256	76.7
8	1.172	64	24.68	512	113.4

Table 3.3: The value of regularization parameter λ and their corresponding χ^2 per degree of freedom for Fig.3.19

F_2 agrees with the data.

Possible developments of of this project include studying the choice of the regularization parameter with a linear inequality constraint more accurately, e.g. by studying the methods available in [80]. Also, the algorithm could be further studied to determine a covariance matrix of the solution. An alternative iterative algorithm LSSOL [76] which is strongly recommended by Björck [71] might also be applied to our problem.

The immediate next step is to go on to study the whole F_2 data set and extract the dipole cross-section at different values of x. Our ultimate goal is to assess if the extracted dipole cross-section $\sigma(r, x)$ requires saturation as claimed in the studies [10–12].

Chapter 4

Conclusion

In the first part of this thesis, we reviewed the important experimental results and theoretical concepts that lead to parton saturation at low Bjorken x. We showed how the dipole cross-section can be an important tool for describing the deep inelastic scattering process and studying saturation effects at low x. Then, we briefly reviewed three recent dipole models, and summarized the results of the comparison of their predictions with data. Subsequently, we identified the nature of the difficulty in answering the question: to what extent are saturation dynamics presented in the data? We discussed the possible extraction of the dipole cross-section without a prior parameterization as an unbiased way to answer this question. Subsequently, we identified the integration equation that needs to be solved, or unfolded, in order to perform this task.

In the second part, we discussed the discretization of this integration equation, and produced a parameterization of the structure function data, in order to "produce" enough 'pseudo' data, to perform the unfolding.

In the third part, we firstly illustrated a failed attempt of direct unfolding, by analyzing the results using singular value decomposition (SVD) we identified our problem as a *discrete ill-posed problem*. Subsequently, we discussed the general strategies for solving this problem. Moreover, four papers in high energy physics discussing unfolding are reviewed and Tikhonov regularization is identified as the the most appropriate method for solving our problem. We also studied the effect of

regularization and discussed the question: how accurate can a regularized solution approximate the exact solution. We concluded that the key is to choose the most appropriate regularization parameter λ. Finally, we unfolded the dipole cross-section by requiring that it is both "smooth" and "monotonic" with respect to varying dipole size r. The result shows. A physical dipole cross-section was successfully extracted and its prediction agrees with data.

We summarize here the most important steps for future work. Firstly, as the regularization parameter λ holds the key to the accuracy of the extracted dipole cross-section, it is crucial to study thoroughly all parameter choice methods available, in order to choose the best method or devise an appropriate method for our problem. Also how to choose the regularization parameter in the case of linear inequality constraints needs to be attacked. Secondly, the Haskell and Hanson [75] algorithm could be improved, so that the "monotonicity" constraint can be more strictly incorporated. Finally, to compute the error of the regularized solution accurately, the error from discretization and rounding errors needs to be included. Moreover the Haskell and Hanson [75] algorithm could be further modified to compute the covariance matrix of the regularized solution taking into account all the above errors.

Once the dipole cross-section is accurately unfolded, we will be able to directly assess the extent to which the saturation characteristics are presented in our extracted dipole cross-section, and subsequently be able to make a more definite statement about the extent to which the saturation dynamics are present in the data. The same method can also be applied to unfold the dipole cross-section using the diffractive deep inelastic scattering data [83–85].

References

[1] H. Abramowicz and A. Caldwell, Rev. Mod. Phys. **71**, 1275 (1999), hep-ex/9903037.

[2] L. V. Gribov, E. M. Levin, and M. G. Ryskin, Phys. Rept. **100**, 1 (1983).

[3] A. H. Mueller and J.-w. Qiu, Nucl. Phys. **B268**, 427 (1986).

[4] A. H. Mueller, Nucl. Phys. **B335**, 115 (1990).

[5] A. H. Mueller, Nucl. Phys. **B415**, 373 (1994).

[6] J. R. Forshaw, G. Kerley, and G. Shaw, Phys. Rev. **D60**, 074012 (1999), hep-ph/9903341.

[7] K. Golec-Biernat and M. Wusthoff, Phys. Rev. **D59**, 014017 (1999), hep-ph/9807513.

[8] E. Iancu, K. Itakura, and S. Munier, Phys. Lett. **B590**, 199 (2004), hep-ph/0310338.

[9] J. R. Forshaw and G. Shaw, JHEP **12**, 052 (2004), hep-ph/0411337.

[10] J. R. Forshaw and G. Shaw, JHEP **12**, 052 (2004), hep-ph/0411337.

[11] J. R. Forshaw and G. Shaw, Nucl. Phys. Proc. Suppl. **146**, 206 (2005).

[12] J. R. Forshaw, R. Sandapen, and G. Shaw, (2006), hep-ph/0608161.

[13] T. Feldmann, DIS: Theoretical introduction, at Workshop on DIS, Graduiertenkolleg Basel, 2004.

[14] J. D. Bjorken, Phys. Rev. **179**, 1547 (1969).

[15] R. P. Feynman, Phys. Rev. Lett. **23**, 1415 (1969).

[16] R. Devenish and A. Cooper-Sarkar, *Deep Inelastic Scattering* (Oxford University Press, Oxford, New York, 2004).

[17] A. Harindranath, (1996), hep-ph/9612244.

[18] R. Venugopalan, (1998), nucl-th/9808023.

[19] J. Silva, "QCD Corrections to the Process $e^+e^- \to q\overline{q}$", Master's thesis, School of Physics and Astronomy, University of Manchester, Manchester, UK, 2006.

[20] V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972).

[21] G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977).

[22] Y. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).

[23] R. D. Ball and S. Forte, Phys. Lett. **B335**, 77 (1994), hep-ph/9405320.

[24] J. R. Forshaw and D. A. Ross, *Quantum Chromodynamics and the Pomeron* (Cambridge University Press, Cambridge, UK, 1997).

[25] E. Levin, (2001), hep-ph/0105205.

[26] J. R. Forshaw and G. Shaw, (Cambridge University Press, Cambridge, UK, 2007), chap. Diffraction and Colour Dipoles.

[27] Y. V. Kovchegov, Phys. Rev. **D60**, 034008 (1999), hep-ph/9901281.

[28] I. Balitsky, Nucl. Phys. **B463**, 99 (1996), hep-ph/9509348.

[29] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970).

[30] A. Donnachie and G. Shaw*Generalized Vector Dominace in Electromagnetic Interactions of Hadrons* Vol. 2 (Plenum Press, 1978).

[31] K. Golec-Biernat and M. Wusthoff, Phys. Rev. **D60**, 114023 (1999), hep-ph/9903358.

[32] K. Golec-Biernat, Acta Phys. Polon. **B33**, 2771 (2002), hep-ph/0207188.

[33] E. Iancu, A. Leonidov, and L. McLerran, (2002), hep-ph/0202270.

[34] L. D. McLerran, Lect. Notes Phys. **583**, 291 (2002), hep-ph/0104285.

[35] J. R. Forshaw, G. Kerley, and G. Shaw, Phys. Rev. **D60**, 074012 (1999), hep-ph/9903341.

[36] A. Donnachie and P. V. Landshoff, Phys. Lett. **B437**, 408 (1998), hep-ph/9806344.

[37] M. G. Wing and J. D. Zahrt, *A Primer on integral equations of the first kind: The problem of deconvolution and unfolding* (Port City, Baltimore, Maryland, 1991).

[38] C. W. Groetsch, *Inverse Problems in the Mathematical Sciences* (Vieweg-Sohn, Wiesbaden, 1993).

[39] L. Delves and J. L. Mohamed, *Computational Methods for Integral Equations* (Cambridge University Press, Cambridge, UK, 1985).

[40] Matlab 7.1 (service pack 3), The MathWorks Inc, 1984-2005.

[41] ZEUS, S. Chekanov *et al.*, Eur. Phys. J. **C21**, 443 (2001), hep-ex/0105090.

[42] F.James, Minuit: Function minimizaiton and error analysis, Version.94.1. CERN, Geneva, Swizerland, 1998.

[43] Cern program library, http://cernlib.web.cern.ch/cernlib/.

[44] W. Eadie and D. D. etc., *Statistical Methods in Experimental Physics*, 1st ed. (Elsevier Noth-Holland, New York, 1977).

[45] J. J. More, Derivation and theory for the davidon-fletcher-powell and broyden-fletcher-goldfarb-shanno methods, in *Bulletin of the Operations Research Society of America*, 1975.

[46] L. N. Trefethen and D. Bau, *Numerical Linear Algebra* (Siam, Philadephia, 1997).

[47] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3^{rd} ed. (The Johns Hopkins University Press, Baltimore, 1996).

[48] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems* (Siam, Philadelphia, 1997).

[49] A. Bukhgeim, *Introduction to the Theory of Inverse Problems* (VSP BV, Netherlands, 2000).

[50] A. Hocker and V. Kartvelishvili, Nucl. Instrum. Meth. **A372**, 469 (1996), hep-ph/9509307.

[51] A. Tikhonov, Sov.Math **5**, 1034 (1963).

[52] R. Barlow, SLUO lectures on statistics and numerical methods in hep, lecture 9: Unfolding, http://www.hep.man.ac.uk/u/roger/.

[53] V. Blobel, (2002), hep-ex/0208022.

[54] G. Cowan, Prepared for Conference on Advanced Statistical Techniques in Particle Physics, Durham, England, 18-22 Mar 2002.

[55] G. D'Agostini, Nucl. Instrum. Meth. **A362**, 487 (1995).

[56] R. Narayan and R. Nityananda, Ann. Rev. Astron. Astrophys. **24**, 127 (1986).

[57] P. C. Hansen, Regularization tools: A matlab package for analysis and solution of discrete ill-posed problems, Numer. Algorithms, 1992.

[58] P. C. Hansen, BIT **30**, 658 (1990).

[59] D. Marquardt, Technometrics **12**, 591 (1970).

[60] H. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems* (Dordrecht, Netherlands, 1996).

[61] V. Morozov, Soviet Math. **7**, 414 (1966).

[62] B. Hofmann, *Regularization for Applied Inverse and Ill-Posed Problems* (Teuber, Germany, 1986).

[63] G. Wahba, Spline models for observational data, in *CBMS-NSF Regional Conferences Series in Applied Mathematics* Vol. 59, Philadelphia, 1990, SIAM.

[64] A. Davies and M. Hassan, *Optimality in the regularization of ill-posed inverse problems* (Academic Press, London, UK, 1987), .

[65] P. Hansen, SIAM review **34**, 561 (1992).

[66] M. H. G.H. Golub and G. Wahba, Technometrics **21**, 215 (1979).

[67] M. Hanke and T. Raus, SIAM J. Sci. Comput. **17**, 956 (1996).

[68] T. Chan and J. O. etc., BIT **32**, 481 (1992).

[69] V. Morozov, *Methods for Solving Incorrectly Posed Problems* (Springer-Verlag, New York, 1984).

[70] C. R. Vogel, Inverse Problems **12**, 535 (1996).

[71] A. Björck, *Numcerical Methods for Least Squares Problems* (SIAM, Philadelphia, 1996).

[72] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*, 2nd ed. (SIAM, Philadelphia, 1995).

[73] R. Mifflin, Mathematical Programming **16**, 141 (1979).

[74] K. Schittkowski and J. Stoer, Numerische Mathematick **31**, 431 (1979).

[75] K. H. Haskell and R. J. Hanson, Mathematical Programming **21**, 98 (1981).

[76] P. Gill and S. etc., User's guide for lssol(version 1.0): a fortran package for constrained linear least-squares and convex quadratic programming, Report SOL, Standford University, CA, USA, 1986.

[77] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*, 1st ed. (SIAM, Philadelphia, 1974).

[78] A. Björck and L. Eldén, Dept. of Mathematis, Linköping University Report No. LITH-R-33-1979, 1979 (unpublished).

[79] R. Hanson, Comm. ACM **15**, 883 (1972).

[80] G.Wahba, Dept. of Statistics, University of Wisconsin Report No. 595, 1980 (unpublished).

[81] K. H. Haskell and R. J. Hanson, ACM Transactions on Mathmatical software (1982).

[82] Slatec common mathematical library, http://www.netlib.org/slatec/.

[83] H1, A.Aktas et al, (2006), hep-ex/0606003.

[84] ZEUS, S. Chekanov *et al.*, Eur. Phys. J. **C38**, 43 (2004), hep-ex/0408009.

[85] ZEUS, S. Chekanov *et al.*, Nucl. Phys. **B713**, 3 (2005), hep-ex/0501060.