UNIVERSITÀ DEGLI STUDI DI MILANO

FACOLTÀ DI SCIENZE E TECNOLOGIE

Laurea Magistrale in Fisica

Optimization of parton density uncertainties

Relatore: Prof. Stefano Forte

Correlatore: Prof. Stefano Carrazza

Luca Talon
Matricola n° 884722
A.A. 2017/2018

Contents

Introduction

The determination of high precision theoretical predictions of the proton-proton collisions is a fundamental task in modern particle physics phenomenology. Since the proton structure is described in terms of Parton Distribution Functions (PDFs), it is of crucial importance to obtain a precise assessment of PDFs and their uncertainties. The PDFs represent at the lowest order in perturbation theory the probability to find a given constituent of the proton carrying a given momentum fraction and can not be determined from first principles by the current theory that describes the proton constituent interactions, namely Quantum Chromodynamics (QCD).

The determination of PDFs thus follows the same procedure of other QCD parameters, i.e. fitting appropriate experimental data given by the standard χ^2 minimization. The main difficulty in parton density fits consists in the determination of a function rather than a single parameter from a finite sample of experimental measurements. Moreover, experimental uncertainties introduce fluctuations in the functional space of PDFs and therefore an accurate description of the functional probability distribution of PDFs is mandatory in order to understand parton density contribution in theoretical predictions.

The modern approaches to the determination of parton densities rely on a parametrization that allows us to reduce the problem of fitting a functional probability distribution into a finite-dimensional problem in the space of parameters. The description of the parameter probability distribution follows two main approaches, namely the Hessian representation and the Monte Carlo representation: the former assumes that the χ^2 near the minimum is a quadratic function of the parameters which follow a multivariate gaussian distribution. The one-sigma contour in the parameter space is then provided by the textbook parameter-fitting criterion $\Delta\chi^2 = 1$. However this criterion does not provide a reliable estimation of the PDF uncertainties for existing Hessian sets which adopt instead the criterion $\Delta\chi^2 = t^2$, where t is called tolerance and typically $t \simeq 5$. In the Monte Carlo representation the parameter probability distribution is provided by a Monte Carlo replica sample and the uncertainties are calculated with statistical estimators.

The aim of this work is to study the problem of the introduction of a toler-

ance for the Monte Carlo PDFs provided by NNPDF collaboration that adopts a methodology based on the neural networks. Furthermore the main reason of the need of a tolerance for the existing Hessian sets consists in the dataset incompatibility, i.e. the results obtained from different datasets are not compatible within their uncertainties produced with the parameter-fitting criterion $\Delta\chi^2 = 1$. The tolerance thus implies an inflation of the PDF uncertainties in order to accommodate all the contributions from each dataset.

However the Monte Carlo sets provided by the NNPDF collaboration do not exhibit any tension between the fit results with different datasets and therefore it is important to understand if the tolerance concerns only the existing Hessian PDFs or it represents a deeper property of PDF determination. Since the Monte Carlo uncertainties do not rely on a parameter-fitting criterion upon the $\Delta\chi^2$, a Hessian conversion of the Monte Carlo PDF sets is required in order to assess the need of a tolerance.

This elaborate is organized as follows: In the first chapter a brief review of how PDFs were historically introduced is given along with a description of the main property of PDFs. In the second chapter we introduce the experimental data adopted in the PDF fits and then we discuss in detail the two main representations of the PDF uncertainties. In the third chapter we introduce the problem of the tolerance for a Monte Carlo set and we present the strategies adopted for the calculation of the results.

Chapter 1

From deep inelastic scattering to parton distributions

In this section an overview of how PDFs arise from deep inelastic scattering is presented.

We first briefly review the fundamental aspects of Quantum Chromodynamics (QCD), particularly those related to the determination of PDFs. Afterwards we approach the deep inelastic scattering in the naive parton model and then we will introduce QCD-improved corrections which lead to PDF evolution equation.

1.1 Fundamentals of QCD

QCD is a gauge field theory which is invariant under local transformation of the gauge group $SU(3)$. The gauge bosons of the theory are called gluons and are massless particles with spin 1. The fermions are called quarks and are massive particles with fractional electric charge (either $2/3$ or $-1/3$ for the quarks, and the opposite sign for the antiquarks). There are three families of quarks each containing a pair of quarks with their corresponding antiparticles and these six types of quarks are generically called flavours. The Lagrangian density which is invariant under local $SU(3)$ transformations is:

$$\mathcal{L} = \sum_{flavours} \overline{\psi}_a (i\gamma_\mu D^\mu - m)_{ab} \psi_b - \frac{1}{4} \operatorname{Tr} \left[G_{\mu\nu} G^{\mu\nu} \right], \tag{1.1}$$

where ψ_a are the quark fields, D_μ is the covariant derivative

$$D_\mu = \partial_\mu + i g_0 T_a A^a_\mu, \tag{1.2}$$

where A^a_μ is the gluon field and T_a are the eight generators of $SU(3)$ in the fundamental representation. $G^a_{\mu\nu}$ is the field strength tensor which can be defined in

terms of the gluon field and the $SU(3)$ structure constants f^{abc} in the following way:

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g_0 f^{abc} A^b_\mu A^c_\nu. \tag{1.3}$$

The parameter g_0 is the bare coupling of the theory and the bare strong coupling is defined by $\alpha_S^0 = \frac{g_0^2}{4\pi}$.

The amplitude of a QCD process can be computed perturbatively in α_S^0 using the Feynman rules that can be deduced from the Lagrangian density in Eq. (1.1). However, higher order terms in α_S^0 typically suffer from divergences that can be cured by redefining some bare quantities such as α_S^0 in order to remove divergent terms. This is the basic idea of a more general procedure called renormalization that requires the introduction of an unphysical energy scale μ in order to isolate the divergences involved in the definition of the new 'renormalized' quantities. Since physical observables do not depend on the particular choice of μ, it can be shown that the renormalized strong coupling constant α_S can be altered so that it absorbs the unphysical dependence of μ. This dependence is determined by the renormalization group equation (Callan-Symanzik):

$$\mu^2 \frac{d}{d\mu^2} \alpha_S(\mu^2) = \beta(\alpha_S(\mu^2)), \tag{1.4}$$

where the β function admits a perturbative expansion in α_S:

$$\beta(\alpha_S) = -\alpha_S(\beta_0 + \beta_1 \alpha_S + \beta_2 \alpha_S^2). \tag{1.5}$$

Eq. (1.4) implies that α_S runs, i.e. the strength of the coupling constant depends on the energy scale of the process in which it enters. At leading order, namely the lowest order in α_S, the running is determined by the β_0 coefficient which is:

$$\beta_0 = \frac{33 - 2n_f}{12\pi}, \tag{1.6}$$

where n_f is the number of flavours that are light at the scale μ^2, i.e. their squared masses are lower than μ^2 . Since $n_f < 17$ at any scale in QCD, the β function is negative, implying that the strength of the coupling increases as the scale of the interaction decreases. This property of QCD is known as asymptotic freedom [1][2] and implies that quarks and gluons can be treated as free particles when $\mu^2 \to \infty$. The solution of Eq. (1.4) at leading order can be written in terms of the value of α_S at some arbitrary fixed scale Q_0^2 in the following way:

$$\alpha_S(\mu^2) = \alpha_S(Q_0^2) \left(1 - \beta_0 \alpha_S(Q_0^2) \log \frac{\mu^2}{Q_0^2}\right), \tag{1.7}$$

thus it is enough to measure the coupling constant at one scale and then it can be computed to any other scale using Eq. (1.7). Equivalently, the solution of Callan-Symanzik equation at leading order can be expressed in terms of the parameter Λ, known as the Landau pole of QCD:

$$\alpha_S(\mu^2) = \frac{1}{\beta_0 \log \frac{\mu^2}{\Lambda}}. \tag{1.8}$$

The value of Λ is not given by the theory and it can be shown that Λ is of the same order of the squared mass of the proton. Since Λ represents the energy scale at which the coupling constant diverges (note that leading order approximation becomes inadequate as a consequence), it can be used as a rough estimate of the scale in which perturbation theory breaks down. This regime is known as long distance physics while the opposite regime where the energy scale is greater than Λ is called short distance physics.

1.2 Deep inelastic scattering

Parton distribution functions were introduced by Feynman in 1969 in the effort to understand the scattering behavior of hadronic states in terms of parton model and successfully describe the deep inelastic scattering measurements.

In this process, a high energy lepton with four-momentum k^μ probes a proton with four-momentum P^μ with the exchange of a gauge boson. For simplicity we only consider the charged current interaction mediated by a virtual photon. If we label the outgoing lepton with k'^μ, the four-momentum carried by the photon is:

$$q^\mu = k^\mu - k'^\mu, \; q^2 = -Q^2 < 0. \tag{1.9}$$

In the inelastic regime Q^2 is large and the proton fragments into an arbitrary hadronic state X as illustrated at tree level in Fig. 1.1.

In the proton rest frame we can define the following invariant kinematic variables:

$$\begin{aligned} M^2 &= P^2, \\ \nu &= P \cdot q = M(E - E'), \\ y &= \frac{q \cdot P}{k \cdot P} = 1 - \frac{E'}{E}, \\ x &= \frac{Q^2}{2P \cdot q} = \frac{Q^2}{2\nu}, \end{aligned} \tag{1.10}$$

where M is the mass of the proton, E and E' are the energies of the ingoing and outgoing lepton respectively, ν is the energy transfer and y is known as inelasticity

and it ranges from 0 (elastic scattering) to 1 (maximum energy transfer). The variable x is known as the Bjorken parameter and it will play a fundamental role in the parton model.

Despite of the large number of kinematic variables, the squared amplitude can be described using only two of them: the 8 unknown kinematic variables (hadronic and leptonic outgoing four-momentum) are constrained by four equations from energy-momentum conservation and on-shell condition for k'^{μ} which reduce the degrees of freedom by 5. Furthermore, the process is invariant under rotation around the lepton direction and therefore the squared amplitude will not depend on the azimuthal angle so that the actual number of degrees of freedom is decreased to two.

We shall consider henceforth x and Q^2 as independent kinematic variables. Since QED is a gauge theory, the propagator of the photon depends on the particular gauge choice used to remove unphysical components of the photon field A^{μ}. If we consider the covariant gauge that fixes $\partial_{\mu}A^{\mu} = 0$, the photon propagator assumes the following form:

$$D_{\mu\nu}(q) = -\frac{i}{q^2}\left(g_{\mu\nu} - (1-\xi)\frac{q_{\mu}q_{\nu}}{q^2}\right), \tag{1.11}$$

where ξ is a finite constant which fixes the specific gauge choice. In the Feynman gauge ξ is set to 1 and the photon propagator becomes:

$$D_{\mu\nu}(q) = -\frac{ig_{\mu\nu}}{q^2}. \tag{1.12}$$

Neglecting spin labels, the amplitude of this process in the Feynman gauge is thus

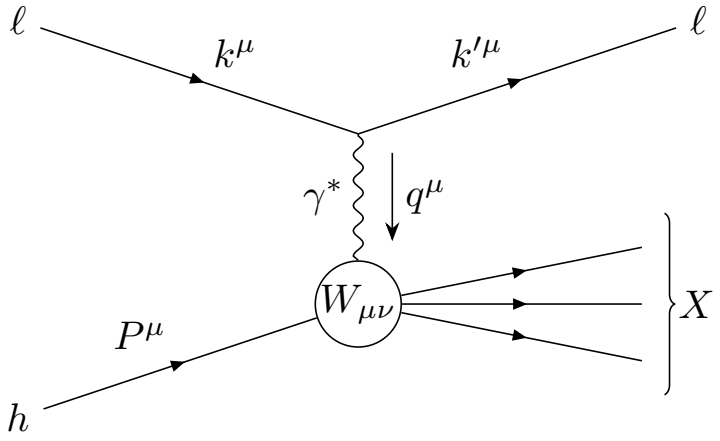

Figure 1.1: *Deep inelastic scattering of a charged lepton with a target proton.*

given by:

$$i\mathcal{M} = \bar{u}(k')(-ie\gamma^\mu)u(k)\left(-\frac{ig_{\mu\nu}}{q^2}\right)\langle X|\mathcal{J}_h^\nu|P\rangle\,, \tag{1.13}$$

where $\mathcal{J}_h^\nu$ is the hadronic current which can not be explicitly computed due to our ignorance of the wavefuctions for the hadronic states $|X\rangle$ and $|P\rangle$.

To isolate the problem, we are able to factorize the inclusive unpolarized squared amplitude into an hadronic ($W_{\mu\nu}$) and leptonic ($L_{\mu\nu}$) part in the following way:

$$\begin{aligned}|i\overline{\mathcal{M}}|^2 =& \frac{e^2}{Q^4}\left[\frac{1}{2}\sum_{pol}\bar{u}(k')\gamma_\mu u(k)\bar{u}(k)\gamma_\nu u(k')\right]\left[\frac{1}{2}\sum_X\langle P|\mathcal{J}_h^{\dagger\mu}|X\rangle\langle X|\mathcal{J}_h^\nu|P\rangle\right]\\ =& \frac{e^2}{Q^4}L_{\mu\nu}W^{\mu\nu}\end{aligned} \tag{1.14}$$

with:

$$W^{\mu\nu} = \frac{1}{2}\sum_X\langle P|\mathcal{J}_h^{\dagger\mu}|X\rangle\langle X|\mathcal{J}_h^\nu|P\rangle = \frac{1}{2}\langle P|\mathcal{J}_h^{\dagger\mu}\mathcal{J}_h^\nu|P\rangle\,. \tag{1.15}$$

Since the term $\langle X|\mathcal{J}_h^\mu|P\rangle$ is the amplitude for the interaction of an incoming proton with a virtual photon that gives a final state X, the hadronic tensor therefore represents the inclusive squared amplitude for the process $\gamma^* + P \to X$.
Instead the leptonic tensor can be straightforwardly calculated in the following way:

$$\begin{aligned}L^{\mu\nu} =& \frac{1}{2}\sum_{pol}\bar{u}(k')\gamma^\mu u(k)\bar{u}(k)\gamma^\nu u(k') = \frac{1}{2}\,\mathrm{Tr}\left[\not{k}'\gamma^\mu\not{k}\gamma^\nu\right]\\ =& 2\left(k'^\mu k^\nu + k'^\nu k^\mu - g^{\mu\nu}k'\cdot k\right),\end{aligned} \tag{1.16}$$

where we have neglected the lepton mass. The hadronic tensor $W^{\mu\nu}$ can not be calculated from first principles in perturbation theory but we can retrieve information about its structure by requiring invariance under parity transformation and the conservation of the hadronic current:

$$q^\mu W_{\mu\nu} = q^\nu W_{\mu\nu} = 0. \tag{1.17}$$

Hence the tensor may be parametrised without loss of generality in terms of the QED structure functions F_1 and F_2:

$$\begin{aligned}W^{\mu\nu} =& -\left(g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2}\right)F_1\left(x,Q^2\right) +\\ &+\left(P^\mu - q^\mu\frac{P\cdot q}{q^2}\right)\left(P^\nu - q^\nu\frac{P\cdot q}{q^2}\right)\frac{1}{\nu}F_2\left(x,Q^2\right).\end{aligned} \tag{1.18}$$

If parity-violating interactions were taken into account, a further function F_3 would arise from the hadronic tensor parametrization. We may now compute the squared amplitude in Eq. (1.14), the flux factor and then the cross section in terms of the structure functions. The differential cross section of DIS mediated by a virtual photon is given by:

$$\frac{d\sigma_{DIS}}{dxdQ^2} = \frac{16\pi M^2 E^2}{Q^4}\Big[\frac{1}{2}\left(1+(1-y)^2\right)xF_1\left(x,Q^2\right) + (1-y)(F_2(x,Q^2) - 2xF_1(x,Q^2)) - \left(\frac{M}{2E}\right)xyF_2(x,Q^2)\Big]. \tag{1.19}$$

It is worth noting that the structure functions are related to DIS experimental measurements through Eq. (1.19) and therefore they represent physical observables.

To analyse the hadronic tensor is now convenient to introduce the Sudakov decomposition of a generic four-vector k in terms of two light-like vectors, namely p and n, along with a space-like two-dimensional transverse vector:

$$k^\mu = ap^\mu + bn^\mu + k_T^\mu, \tag{1.20}$$

with

$$\begin{gathered} p^2 = n^2 = k_T \cdot p = k_T \cdot n = 0, \\ p \cdot n = 1. \end{gathered} \tag{1.21}$$

Therefore, we can write the initial state four-momenta as follows:

$$\begin{aligned} P^\mu &= p^\mu + \frac{M^2}{2}n^\mu, \\ q^\mu &= \nu n^\mu + q_T^\mu. \end{aligned} \tag{1.22}$$

Given the Sudakov decomposition, the structure functions can be projected out of the hadronic tensor by:

$$\begin{aligned} F_2 &= \nu n^\mu n^\nu W_{\mu\nu}, \\ F_L &:= F_2 - 2xF_1 = \frac{Q^4}{\nu^3}p^\mu p^\nu W_{\mu\nu}. \end{aligned} \tag{1.23}$$

The term F_L which appears in the second equation of Eq. (1.23) is known as the longitudinal structure function.

So far, few assumptions have been made to describe the tensor $W_{\mu\nu}$: we successfully manage to encode all the proton QED structure in two dimensionless scalar

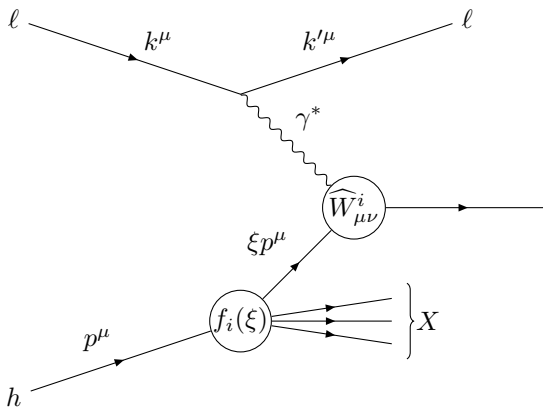

Figure 1.2: *Parton model picture for DIS.*

functions $F_1(x, Q^2)$ and $F_2(x, Q^2)$ by requiring charge conservation and Lorentz covariance.

In the limit where $Q^2, \nu \to +\infty$, known as the Bjorken limit, the structure functions were observed to obey an approximate scaling law, i.e. they depend only on the dimensionless variable x:

$$F_i\left(x, Q^2\right) \to F_i(x)\,. \tag{1.24}$$

Bjorken scaling implies that the photon scatters off a pointlike particle, since otherwise the dimensionless structure functions would depend on Q^2 only through Q^2/Q_0^2, where $1/Q_0^2$ refers to some length scale of the interacting particle.

1.3 Parton model

The first formulation of the parton model, also known as 'naive' parton model, states that for a sufficiently hard interaction the virtual photon scatters off a single point-like parton inside the proton (Fig. 1.2) and we can treat the partons as approximately free particles. A generic quantity which depends on the hadronic state can be computed as the sum of partonic contributions calculable in perturbation theory weighted by a parton distribution function which encodes the probability of the parton to carry a fraction ξ of proton total momentum. PDFs can not be calculated from first principles as they depend on non-perturbative internal structure of the proton. The 'naive' parton model and its main features can be summarized as follows:

- When Q^2 is high enough that the binding energy of the proton can be neglected, the photon becomes sensitive to proton constituents and interacts only with one parton.

- Parton carries a fraction ξ of the proton total momentum p, namely ξp, and its intrinsic momentum due to parton dynamics inside the proton is neglected.

- Each parton is associated with a probability density $f_i(\xi)$, where i represents the type of the parton. The PDFs are fitted from data and thus can be considered 'functional parameters' of the theory.

- The hadronic tensor is assumed to follow:

$$W_{\mu\nu}\left(x,Q^2\right) = \sum_i^{partons} f_i \otimes \widehat{W}^i_{\mu\nu}\left(x,Q^2\right), \tag{1.25}$$

  where $\widehat{W}^i_{\mu\nu}$ refers to the parton-virtual photon scattering and we have introduced the multiplicative convolution:

$$f \otimes g\,(x) = \int_0^1 \frac{d\xi}{\xi} f\,(\xi)\, g\left(\frac{x}{\xi}\right). \tag{1.26}$$

The parton level tensor must obey the same symmetries as the hadron level tensor $W^{\mu\nu}$. Therefore, the form of $\widehat{W}^{\mu\nu}$ is:

$$\widehat{W}^{\mu\nu}_i = -\left(g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2}\right)\widehat{F}^i_1\left(\frac{x}{\xi},Q^2\right) + \left(p^\mu - q^\mu\frac{p\cdot q}{q^2}\right)\left(p^\nu - q^\nu\frac{p\cdot q}{q^2}\right)\frac{\xi^2}{\nu}\widehat{F}^i_2\left(\frac{x}{\xi},Q^2\right), \tag{1.27}$$

where $\widehat{F}^i_1$ and $\widehat{F}^i_2$ are the parton level structure functions and are related to hadronic ones as follows:

$$\begin{aligned} F_J(x,Q^2) &= \sum_i^{partons} \int_0^1 \frac{d\xi}{\xi} f_i\,(\xi)\,\widehat{F}^i_J\left(\frac{x}{\xi},Q^2\right),\ J=1,L \\ F_2(x,Q^2) &= \sum_i^{partons} \int_0^1 \xi d\xi f_i\,(\xi)\,\widehat{F}^i_2\left(\frac{x}{\xi},Q^2\right). \end{aligned} \tag{1.28}$$

The crucial advantage of parton model is that the partonic tensor and structure functions are now computable in perturbation theory by calculating the squared amplitude for the subprocess $q_i(\xi p) + \gamma^*(q) \to q_i(l)$ since the partonic version of $W^{\mu\nu}$ can be represented in the following way:

$$\widehat{W}^{\mu\nu}_i = \frac{1}{2}\sum_{pol}\langle q_i|\mathcal{J}^{\dagger\mu}_h|q_i\rangle\,\langle q_i|\mathcal{J}^{\nu}_h|q_i\rangle\,. \tag{1.29}$$

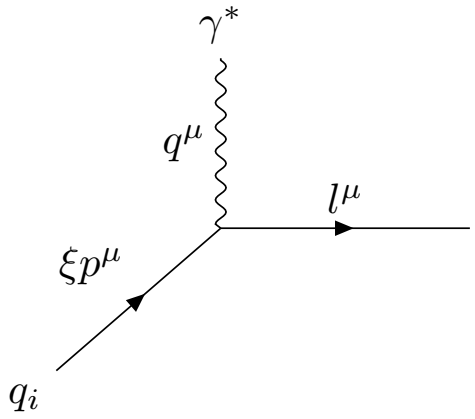

Figure 1.3: *Leading order Feynman's diagram for* $q_i(\xi p) + \gamma^*(q) \to q_i(l)$.

At leading order, also known as Born level, the amplitude is:

$$\begin{aligned} i\mathcal{M}_i^\mu &= -ie_{q_i}\bar{u}\left(l\right)\gamma^\mu u\left(\xi p\right), \\ \widehat{W}_{\mu\nu}^i &= \frac{1}{2}\sum_{pol}|\mathcal{M}_i|_{\mu\nu}^2. \end{aligned} \tag{1.30}$$

The unpolarized squared modulus of Eq. (1.30) can be found analogously to leptonic tensor in Eq. (1.16) (neglecting parton masses). Ultimately, including the phase space of the final parton and using parton level projectors in Eq. (1.23), we find in the C.M. frame:

$$\begin{aligned} \widehat{F}_2^i &= \nu\frac{n^\mu n^\nu}{\xi^2}\widehat{W}_{\mu\nu}^i = 4\nu e_{q_i}^2\delta(l^2), \\ \widehat{F}_L^i &= \frac{Q^4}{\xi\nu^3}p^\mu p^\nu\widehat{W}_{\mu\nu}^i = 0, \\ \widehat{F}_1^i &= \frac{\xi^2}{2x}\left(\widehat{F}_2^i - F_L^i\right) = 2\nu e_{q_i}^2\frac{\xi^2}{x}\delta(l^2). \end{aligned} \tag{1.31}$$

We can rewrite the delta in terms of ξp and q using four-momentum conservation:

$$\delta(l^2) = \delta((\xi p + q)^2) = \delta(2\xi p\cdot q - Q^2) = \frac{1}{2\nu}\delta\left(\xi - x\right). \tag{1.32}$$

We have found that the parton model prediction at leading order explains Bjorken scaling since the structure functions do not depend on Q^2; furthermore, the scaling variable x actually describes the momentum fraction carried by the interacting parton. The full QED structure functions of the proton can be evaluated inserting

the parton level structure functions in Eq. (1.28):

$$\begin{aligned} F_1\left(x,Q^2\right) &= \int_0^1 d\xi \sum_i f_i(\xi) e_{q_i}^2 \frac{\xi}{x}\delta\left(\xi - x\right) = \sum_i e_{q_i}^2 f_i(x), \\ F_2\left(x,Q^2\right) &= 2\int_0^1 \xi d\xi \sum_i f_i(\xi) e_{q_i}^2 \delta\left(\xi - x\right) = 2x\sum_i e_{q_i}^2 f_i(x). \end{aligned} \tag{1.33}$$

We may notice from Eq. (1.33) that $F_2 = 2xF_1$ and since $F_L = F_2 - 2xF_1$ we find $F_L = 0$ which is known as Callan-Gross relation. Furthermore Eq. (1.33) implies that parton distribution functions are tightly related to proton structure functions which can be measured in deep inelastic scattering experiments.

1.4 Higher order corrections

The naive parton model was able to provide a good phenomenological description of early DIS measurements. Its success also provided great support for QCD as the correct description of the strong interaction. As matter of fact, the Bjorken scaling introduced substantial constraints upon the theory governing the internal dynamics of the proton. The asymptotic freedom of QCD allows for a consistent description of Bjorken-scaling, where the constituents of the hadron can be viewed as independent, non-interacting point like particles at high values of the resolution parameter Q^2 . The partons in Feynman's model were therefore quickly associated with the quarks and gluons of QCD and a theoretical proof of 'naive' parton model assumptions was given by Product Operators Expansion formalism.

We shall now investigate perturbative QCD corrections to structure functions beyond leading order. As we shall see below, next-to-leading order (NLO) corrections in α_S lead to divergences which can be regularized, factorized and re-absorbed in the definition of PDFs that acquire a dependence of the hard scale Q^2. This procedure gives rise to a structure function correction proportional to $\alpha_S \log Q^2$ that breaks Bjorken scaling.

The NLO amplitude involves the Feynman diagrams in Fig. 1.4 which are given

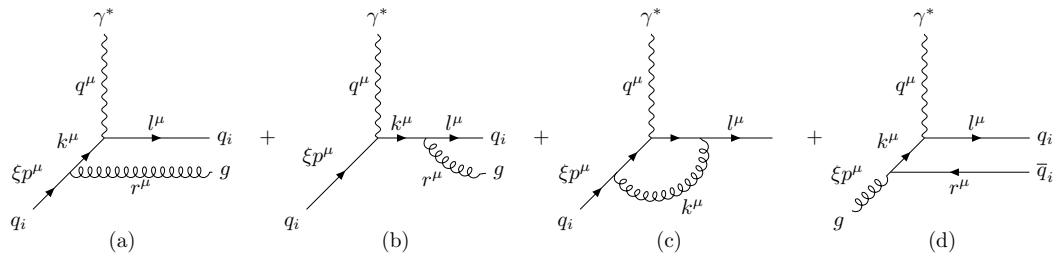

Figure 1.4: *NLO Feynman's diagrams for real gluon emission in initial (a) and final (b) state, virtual vertex correction (c) and initial gluon splitting (c).*

by:

$$
\begin{aligned}
i\mathcal{M}^{\mu}_{a,i} &= \overline{u}(l)(-ie_{q_i}\gamma^{\mu})\frac{i\not{k}}{k^2}(-ig_S\gamma^{\alpha}T^a_{lm})u(\xi p)\varepsilon^*_{\alpha}(r), \\
i\mathcal{M}^{\mu}_{b,i} &= \overline{u}(l)(-ig_S\gamma^{\alpha}T^a_{lm})\frac{i\not{k}}{k^2}(-ie_{q_i}\gamma^{\mu})u(\xi p)\varepsilon^*_{\alpha}(r), \\
i\mathcal{M}^{\mu}_{c,i} &= \overline{u}(l)\Bigg[\int\frac{dk^4}{(2\pi)^4}(-ig_S\gamma^{\alpha}T^a_{lm})\frac{i(\xi\not{p}+\not{q}-\not{k})}{(\xi p+q-k)^2}(-ie_{q_i}\gamma^{\mu}) \\
&\qquad \frac{i(\xi\not{p}-\not{k})}{(\xi p-k)^2}(-ig_S\gamma^{\beta}T^b_{ml})\frac{-ig_{\alpha\beta}}{k^2}\Bigg]u(\xi p), \\
i\mathcal{M}^{\mu}_{d,i} &= \overline{u}(l)(-ie_{q_i}\gamma^{\mu})\frac{i\not{k}}{k^2}(-ig_S\gamma^{\alpha}T^a_{lm})v(r)\varepsilon_{\alpha}(\xi p),
\end{aligned}
\tag{1.34}
$$

where T^a_{ml} are the Gell-Mann matrices and g_S is the strong coupling constant with $\alpha_S = \frac{g_S^2}{4\pi}$.

When computing the squared amplitude we must take into account only terms of $O(\alpha_S)$ and thus we may consider only the squared modulus of real emissions and twice the real part of their interference along with the interference term of virtual vertex correction with Born amplitude (as shown in Fig. 1.5). As it regards the initial state gluon diagrams, the squared amplitude is given by Fig. 1.6. Each of these terms contains integrals which diverge in both the ultraviolet (UV) and infrared (IR) regions. UV divergences are dealt with following QCD renormalization procedure. IR divergences for real emission occur in two regions of the gluon phase-space, namely the soft limit where gluon momentum approaches zero and the collinear limit where the gluon transverse momentum becomes small.

When virtual correction divergences are taken into account, QCD infrared and collinear safety theorem ensures the cancellation of real and virtual singularities if the process is not sensitive to collinear and soft emission. However the collinear divergences present in the real emission diagrams from the initial state partons are not subject to the same cancellations as they modify the momenta at the

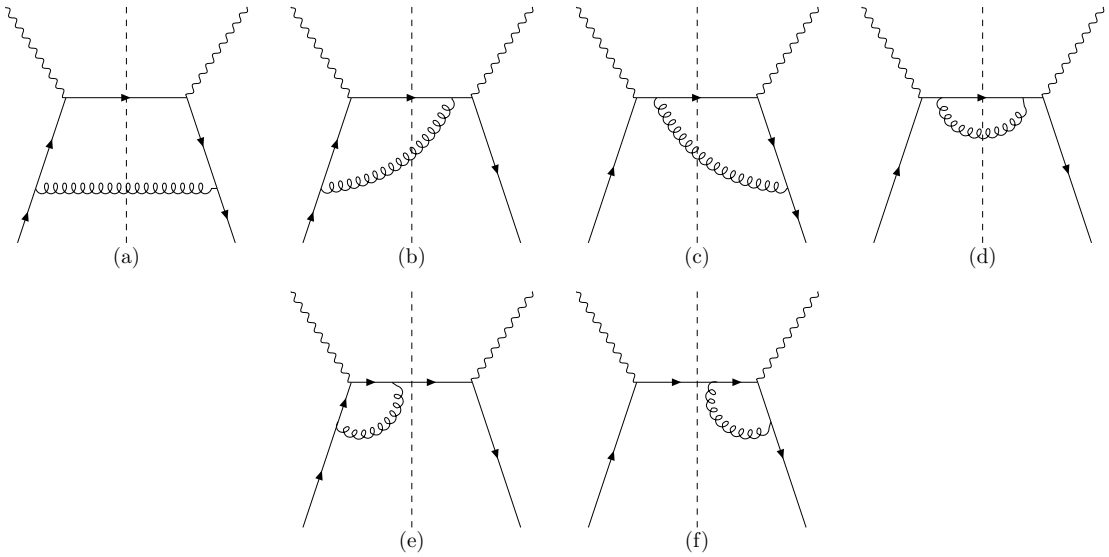

Figure 1.5: *Contributions to the squared amplitude at NLO: the first row describes real emission terms and the second represents the interference between vertex correction and Born amplitude. Divergences in (b), (c), (d) and soft singularities in (a) cancel out with (e) and (f) while collinear singularity in (a) is not involved in this cancelation.*

interaction vertex.

We may treat this infinities in a similar way as the renormalization cures the UV divergences: by factorizing and reabsorbing them into some bare quantity. In this case, we may assume PDFs as bare functional parameters of the theory which are not related to any physical observables and they are redefined in such a way they correspond to finite measurable quantities. We can explicitly show this procedure by calculating the parton structure function at $O(\alpha_S)$:

$$\widehat{F}_2^i\left(\xi, Q^2\right) = 2e_{q_i}^2\,\delta\left(\xi - x\right) + \frac{\alpha_S}{2\pi}\sum_j\left(P_{ij}\left(\frac{x}{\xi}\right)\log\left(\frac{Q^2}{\kappa^2}\right) + H_{ij}(\xi)\right) + O\left(\alpha_S^2\right) \tag{1.35}$$

where i and j refer to interacting and initial state partonic species respectively, $H_{ij}(\xi)$ contains all the finite contributions and $P_{ij}(\xi)$ are the Altarelli-Parisi splitting functions. We have introduced the unphysical cutoff κ^2 to regulate infrared divergences that arise in the limit $\kappa^2 \to 0$. After convoluting the NLO parton level structure functions with the bare PDFs denoted by $f_i^0(\xi)$, we obtain the full

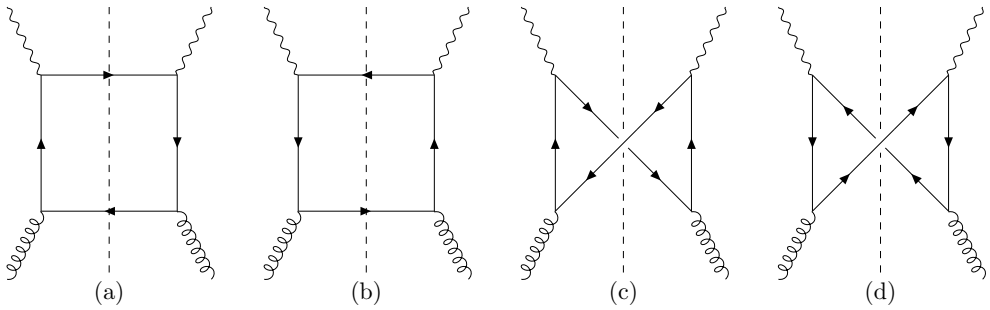

Figure 1.6: *Contribution to squared- amplitude at NLO with initial state gluon. Each term contains both soft and collinear singularities.*

structure function:

$$F_2\left(x,Q^2\right)=\sum_i xe_{q_i}^2\Big[f_i^0(x)+ \\ +\frac{\alpha_S}{2\pi}\int_0^1\frac{d\xi}{\xi}\left(P_{ij}\left(\frac{x}{\xi}\right)\log\left(\frac{Q^2}{\kappa^2}\right)+H_{ij}(\xi)\right)f_j^0(\xi)\Big]+O(\alpha_S^2). \tag{1.36}$$

This expression for structure function still suffers from an IR divergence when $\kappa^2\to 0$. We may attempt to factorize the divergent terms introducing the factorization scale μ_F in order to separate long and short distance physics. The IR singularities can then be reabsorbed into the bare PDF which already describes the strong coupled dynamics of the hadron in the following way:

$$f_i\left(\xi,\frac{Q^2}{\mu_F^2}\right)=f_i^0\left(\xi\right)+\frac{\alpha_S}{2\pi}\int_0^1\frac{d\xi}{\xi}\sum_j\Delta_{ij}^{(1)}\left(\frac{x}{\xi},\frac{\mu_F^2}{\kappa^2}\right)f_j^0\left(\xi\right)+O\left(\alpha_S^2\right), \tag{1.37}$$

where the counter terms $\Delta_{ij}^{(n)}$ are made up by a regular part $\Delta_{r,ij}^{(n)}$ and a singular part $\Delta_{s,ij}^{(n)}$. The singular part is uniquely specified by having to remove the collinear divergence present in parton level structure functions. As regards Eq. (1.36), the singularity may be subtracted by setting:

$$\Delta_{s,ij}^{(n)}\left(\frac{x}{\xi},\frac{\mu_F^2}{\kappa^2}\right)=P_{ij}\left(\frac{x}{\xi}\right)\log\left(\frac{\mu_F^2}{\kappa^2}\right), \tag{1.38}$$

where Altarelli-Parisi splitting functions at NLO are:

$$P_{q\bar{q}}(x)=C_F\left[\frac{1+x^2}{(1-x)_+}+\frac{3}{2}\delta(1-x)\right]+O\left(\alpha_S^2\right), \tag{1.39}$$

$$P_{qg}(x)=T_r\left[x^2+(1-x)^2\right]+O\left(\alpha_S^2\right). \tag{1.40}$$

where T_r and C_F are respectively $\frac{1}{2}$ and $\frac{4}{3}$ and the *plus distribution* in the first equation in Eq. (1.39) is defined so that its integral with a sufficiently smooth function $g(x)$ is:

$$\begin{aligned} &\int_0^1 dx \frac{g(x)}{(1-x)_+} = \int_0^1 dx \frac{g(x)-g(1)}{(1-x)}, \\ &\frac{1}{(1-x)_+} = \frac{1}{(1-x)},\ 0 \leq x \leq 1. \end{aligned} \tag{1.41}$$

Unlike the singular part, the regular part $\Delta_{r,ij}^{(n)}$ is not constrained by any requirement and in principle it can be set arbitrarily. The aforementioned procedure and the choice of a specific regular part are known as factorization scheme. For example, we may include all the finite terms in the regular part of the counter term so that $\Delta_{r,ij}^{(1)} = H_{ij}$. This particular choice is called DIS factorization scheme and allows us to write the hadronic structure function in a particularly simple way:

$$F_2\left(x, Q^2\right) = 2 \int_0^1 \xi d\xi \sum_i f_i^{\mathrm{DIS}}\left(\xi, \frac{Q^2}{\mu_F^2}\right) e_{q_i}^2. \tag{1.42}$$

However the DIS factorization scheme depends on the process where it is defined and thus it does not allow a practical definition of PDFs across multiple experiments. We may fulfill this request by introducing the *modified minimum subtraction* ($\overline{\mathrm{MS}}$) scheme where the regular term contains only process independent contributions which at NLO are given by $\Delta_{r,ij}^{(1)} = \log 4\pi - \gamma_E$. In $\overline{\mathrm{MS}}$ scheme the PDFs assume the following form:

$$f_i^{\overline{\mathrm{MS}}}\left(\xi, \frac{Q^2}{\mu_F^2}\right) = f_i^0\left(\xi\right) + \frac{\alpha_S}{2\pi} \sum_j \left[P_{ij}(x) \log \frac{Q^2}{\kappa^2} + \log 4\pi - \gamma_E\right] \otimes f_j^0\left(x\right) + O\left(\alpha_S^2\right). \tag{1.43}$$

As a consequence of factorization procedure, the new physically observable PDFs acquire an explicit dependence on the hard scale Q^2 as shown in Eq. (1.37).
In this factorization scheme the structure function F_2 becomes:

$$F_2\left(x, Q^2\right) = \sum_i x e_{q_i}^2 \left[f_i^{\overline{\mathrm{MS}}}\left(x, \frac{Q^2}{\mu_F^2}\right) + \frac{\alpha_S}{2\pi} \int_x^1 \frac{d\xi}{\xi} f_i^{\overline{\mathrm{MS}}}\left(\frac{x}{\xi}, \frac{Q^2}{\mu_F^2}\right) \widetilde{H}_{ij}(\xi)\right] + O(\alpha_S^2), \tag{1.44}$$

where $\widetilde{H}_{ij}(\xi)$ represents the finite term remaining after factorization. We may notice that the NLO structure function in Eq. (1.44) involves logarithms of the hard scale Q^2 which break Bjorken scaling.
If higher order are taken into account, the general structure function in a process

independent factorization scheme (such as $\overline{\text{MS}}$) can be written as:

$$F(x,Q^2) = \sum_i \int_x^1 \frac{d\xi}{\xi} C_i\left(\frac{x}{\xi}, \frac{Q^2}{\mu_F^2}, \alpha_S(Q^2)\right) f_i\left(\xi, \frac{Q^2}{\mu_F^2}\right), \tag{1.45}$$

where C_i are finite functions which can be computed perturbatively and represent the Wilson's coefficients if we perform Operators Product Expansion (OPE) of the hadronic tensor. Eq. (1.45) states that the structure functions can be computed convoluting process dependent coefficients C_i with the same set of PDFs, both calculated at a fixed perturbative order in α_S. This condition holds only if the factorization of IR divergences is process independent, i.e. the singular part of the counterterm in Eq. (1.37) is a universal function which has no sensibility to vertex dynamics. This fundamental property of factorization procedure is due to *Universal Collinear factorization theorem* [3] which is valid for a large number of processes.

1.5 PDF Evolution Equation

The factorization procedure introduces an arbitrary scale μ_F in order to separate the IR region from short distance physics. The scale μ_F is commonly set on the same order of magnitude of the hard scale Q^2 and has no physical meaning since the structure functions do not depend on factorization scheme. As we may notice in Eq. (1.45), the right hand side exhibits an explicit dependence on μ_F while the left hand side does not depend on μ_F. The requirement of the independence of structure function from the factorization scale can be expressed as follows:

$$\mu_F^2 \frac{d}{d\mu_F^2} F\left(x,Q^2\right) = 0. \tag{1.46}$$

This relation leads to a renormalization group equation for parton distributions and Wilson's coefficients in terms of Altarelli Parisi splitting functions P_{ij} :

$$\mu_F^2 \frac{d}{d\mu_F^2} f_i\left(y, \frac{Q^2}{\mu_F^2}\right) = \sum_j \int_y^1 \frac{dz}{z} P_{ij}\left(\frac{y}{z}, \alpha_S(\mu_F^2)\right) f_j\left(z, \frac{Q^2}{\mu_F^2}\right), \tag{1.47}$$

$$\mu_F^2 \frac{d}{d\mu_F^2} C_i\left(y, \frac{Q^2}{\mu_F^2}, \alpha_S(\mu_F^2)\right) = -\sum_k \int_y^1 \frac{dz}{z} P_{ik}\left(\frac{y}{z}, \alpha_S(\mu_F^2)\right) C_k\left(z, \frac{Q^2}{\mu_F^2}, \alpha_S(\mu_F^2)\right). \tag{1.48}$$

The above equations are known as Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations and the results can be proven with OPE formalism and hold order by order in perturbation theory. The right hand side of DGLAP equations may be

interpreted as an anomalous dimension for PDFs and Wilson's coefficients which differ only by a minus sign. This feature ensures that a modification of factorization scale μ_F for the PDFs cancels out with the same scale variation for the Wilson's coefficients leaving the structure function invariant under μ_F reparametrization. We may thus define the parton densities at a fixed energy scale and measure them by fitting experimental data and afterwards evolve PDFs at the new energy scale by solving DGLAP equation in Eq. (1.47).
The general form of parton density evolution equations consists in a system of coupled integro-differential equations with Altarelli-Parisi splitting functions as kernel elements. However, the rank of the evolution matrix P_{ij} is not maximal since the charge conjugation invariance and the flavour symmetry (QCD is flavour blind) reduce the number of independent splitting functions according to the following constrains:

$$\begin{aligned} P_{q_i q_j} &= P_{\bar{q}_j \bar{q}_j}, \\ P_{q_i \bar{q}_j} &= P_{\bar{q}_i q_j}, \\ P_{q_i g} &= P_{\bar{q}_i g}, \\ P_{g q_i} &= P_{g \bar{q}_i}. \end{aligned} \tag{1.49}$$

It is now convenient to choose an appropriate basis in the space of PDFs in order to make the matrix evolution P_{ij} as diagonal as possible. We first define:

$$q_i^{\pm} = q_i \pm \bar{q}_i. \tag{1.50}$$

Because of the baryon number conservation, the combinations q_i^- (also called *valences*) are preserved from evolution and thus decouple. Similarly, the following *triplets* combinations decouple from the evolution of the remaining parton densities:

$$\begin{aligned} T_3 &= u^+ + d^+, \\ T_8 &= u^+ + d^+ - 2s^+, \\ T_{15} &= u^+ + d^+ + s^+ - 3c^+, \\ T_{24} &= u^+ + d^+ + s^+ + c^+ - 4b^+, \\ T_{35} &= u^+ + d^+ + s^+ + c^+ + b^+ - 5t^+. \end{aligned} \tag{1.51}$$

Instead, the *singlet* combination given by

$$\Sigma = \sum_i q_i^+ \tag{1.52}$$

couples with the gluon PDF.
Therefore, the *non-singlets*, namely the *valences* and the *triplets*, evolve according

to:

$$\mu_F^2 \frac{d}{d\mu_F^2} f_i^{NS}\left(x, \mu_F^2\right) = \int_x^1 \frac{dz}{z} P_i^{NS}\left(\frac{x}{z}, \alpha_S\right) f_i^{NS}\left(z, \mu_F^2\right), \tag{1.53}$$

where P_i^{NS} are obtained inserting Eq. (1.50) and Eq. (1.51) into DGLAP equation Eq. (1.47). The *singlet* combination couples with the gluon PDF so that their evolution is given by the following system:

$$\mu_F^2 \frac{d}{d\mu_F^2} \begin{pmatrix} g\left(x, \mu_F^2\right) \\ \Sigma\left(x, \mu_F^2\right) \end{pmatrix} = \int_x^1 \frac{dz}{z} \begin{pmatrix} P_{gg} & P_{g\Sigma} \\ P_{\Sigma g} & P_{\Sigma\Sigma} \end{pmatrix} \begin{pmatrix} g\left(z, \mu_F^2\right) \\ \Sigma\left(z, \mu_F^2\right) \end{pmatrix}. \tag{1.54}$$

This particular choice of the basis in the space of PDFs allows us to reduce the $1 + 2n_f$ coupled equations in Eq. (1.47) to $2n_f - 1$ independent equations (n_f for *valences* and $n_f - 1$ for *triplets*) along with an irreducible system of two equations for *singlet* and gluon PDFs. The solutions of Eq. (1.53) and Eq. (1.54) are uniquely determined once the initial conditions are given, i.e. the PDF explicit functional form at some initial scale Q_0^2. However there is no analytical solution and the determination of the PDF evolution typically follows one of two numerical procedures; the most direct consists in solving DGLAP equations iteratively in the x-space and it is implemented in codes such as HOPPET, QCDNUM and APFEL. The alternative procedure relies on Mellin transformation which is give by:

$$f(n) = \int_0^1 x^{n-1} f(x) dx, \; n \in \mathbb{C}. \tag{1.55}$$

The main advantage is that multiplicative convolution in the x-space is reduced to a product in the Mellin-space:

$$\begin{aligned} \int_0^1 dx\; x^{n-1} \int_x^1 \frac{dy}{y} h(y) g(x/y) &= \int_0^1 dx\; x^{n-1} \left[\int_0^1 dy \int_0^1 dz\; h(y) g(z) \delta(x - yz) \right] \\ &= \int_0^1 dy\; y^{n-1} h(y) \int_0^1 dz\; z^{n-1} g(z) \\ &= h(n) \cdot g(n). \end{aligned} \tag{1.56}$$

Therefore the solution of DGLAP equations becomes trivial in the Mellin-space and the trade-off is then recovering the PDFs in x-space which usually requires a numerical inversion procedure. The QCD-PEGASUS implements this approach.

1.6 Heavy quarks

So far we have being assuming the approximation that the partons involved in PDF evolution equations are massless. Since the masses of the three lightest quarks u,

d and s are far below Λ, this approximation is completely reasonable for them. A more accurate treatment of the remaining flavours must be taken into account, especially when the energy scale approaches the quark masses. In particular we are interested in studying the case where the characteristic scale of the interaction is smaller or bigger than quarks masses and how the transition between these two regimes takes place in parton distribution evolution.

The effects of heavy quarks are encoded in several *flavour number schemes* which can be summarized in two limiting cases, depending on the relation between the quark mass m_q and the energy scale Q^2 where the PDFs are defined:

- $m_q^2 \gtrsim Q^2$: the quark can be considered as a completely final state and it does not participate in the evolution equation since there is no energy to produce it. Moreover all the mass effects must be included in the calculation of squared amplitude of the final state.

- $m_q^2 \ll Q^2$: the quark is treated as another massless parton so that its PDF is perturbatively generated by DGLAP equation.

The first limit is well realized in the fixed flavour number scheme (FFNS) which is based on the assumption that heavy quarks are treated as purely final state particles and the only partons of the theory are the n_l lightest quarks and the gluon. When a single heavy quark with mass m_h is introduced, the structure function in this scheme becomes:

$$F(n_l, Q^2, m_h) = \sum_i^{n_l} C_i \left(n_l, \frac{Q^2}{m_h}, \frac{\mu^2}{m_h}, \frac{Q^2}{\mu^2} \right) \otimes f_i(n_l, \mu^2). \tag{1.57}$$

Notice that the sum runs only on the light flavours since there is no heavy quark PDF and the Wilson's coefficients acquire an explicit dependence on heavy quark mass. Eq. (1.57) is accurate when the energy scale is near the mass threshold and below but when Q^2 increases this scheme may suffer from large logarithms of $\frac{Q^2}{m_h^2}$ that are not treated in factorization procedure and threat perturbative series convergence.

Instead the zero-mass variable flavour number scheme (ZM-VFNS) is accurate in the second limit since it introduces a heavy quark PDF. In this scheme a generic structure function is simply:

$$F(n_l + 1, Q^2) = \sum_i^{n_l+1} C_i \left(n_l, \frac{Q^2}{\mu^2} \right) \otimes f_i(n_l + 1, \mu^2). \tag{1.58}$$

ZM-VFNS takes into account mass threshold by setting the heavy quark PDF to zero when Q^2 is below m_q^2. Moreover, the heavy quark PDF evolves with

the DGLAP equations for scales greater than the heavy quark mass. While this method is accurate in the regime where FFNS fails, its treatment of the heavy quarks in terms of massless parton completely ignores the massive contribution to the Wilson's coefficients and its reliability reduces in the large Q^2 region where powers of $(\frac{Q^2}{m_q^2})$ become significant.

Each of these schemes succeeds where the other breaks down. The General Mass Variable Flavour Number Scheme (GM-VFNS) attempts to unify the advantages of both ZM-VFNS and FFNS so that the effects of the heavy quarks are accounted for all scales. The basic idea consists in switching from a FFNS with n_f flavours to a FFNS with $n_f + 1$ flavours when Q^2 matches the mass threshold. The relation between PDFs at scales above and below the quark mass m_h can be established in terms of a $(n_f + 1, n_f)$ transition matrix:

$$f_i(n_f + 1, \mu \to m_h^+) = \sum_j^{n_f} A_{ij}\left(n_f, \frac{Q^2}{m_h^2}\right) \otimes f_j\left(n_l, \mu \to m_h^-\right). \tag{1.59}$$

The matrix A can be computed perturbatively and is known at NNLO. Requiring the continuity of the theoretical expression of the structure functions at mass threshold, we find:

$$\begin{aligned} F(x, Q^2) &= \sum_k C_k^-(n_f, m_h) \otimes f_k^-(n_l) = \sum_j C_j^+(n_f + 1, m_h) \otimes f_j^+(n_l + 1) = \\ &= \sum_{jk} C_j^-(n_f + 1, m_h) \otimes A_{jk}\left(n_f, \frac{Q^2}{m_h^2}\right) \otimes f_k^-(n_l). \end{aligned} \tag{1.60}$$

Since the PDFs are continuous functions, the Wilson's coefficients must obey:

$$C_k^-(n_f, m_h) = \sum_j C_j^-(n_f + 1, m_h) \otimes A_{jk}\left(n_f, \frac{Q^2}{m_h^2}\right), \tag{1.61}$$

where in the previous equations the superindexes + and - refer to the direction of the limit and the dependence on μ is suppressed. This guideline is refined in various GM-VFNS such as ACO, TR and FONLL.

1.7 General properties of parton distribution functions

The GM-VFNS allows us to generate perturbatively the PDFs of heavy quarks when the energy scale Q^2 exceeds the mass threshold. Therefore the number of independent PDFs is constrained by the initial scale Q_0^2 where they are measured as

the quark distributions with mass above the initial scale arise from DGLAP equation. Usually parton distributions are determined at some scale $m_s^2 < Q_0^2 < m_c^2$ in order to minimize the independent distributions while avoiding non-perturbative effects. The seven PDFs of gluon, quarks u, d, s and their antiparticles represent the building block of the parton model applications since every theoretical prediction involving hadronic states relies ultimately on them.

Although the parton densities describe the non perturbative dynamics of hadronic constituents and therefore they can not be computed explicitly, some general statements may shed light on their x and hard scale dependence. The most important constrains are the parton distribution sum rules which fix the relative normalization of PDFs. Firstly, the momentum sum rule states that the total fraction of proton momentum carried by each parton must sum up to one, namely:

$$\int_0^1 dx \left[x\Sigma(x,Q^2) + xg(x,Q^2)\right] = 1, \tag{1.62}$$

where Σ is the singlet combination defined above. Moreover the valence sum rules ensure the conservation of the quantum numbers that characterize the hadron; for the proton we have:

$$\begin{aligned}
\int_0^1 dx\ u^-(x,Q^2) &= \int_0^1 dx[u(x,Q^2) - \bar{u}(x,Q^2)] = 2, \\
\int_0^1 dx\ d^-(x,Q^2) &= \int_0^1 dx[d(x,Q^2) - \bar{d}(x,Q^2)] = 1, \\
\int_0^1 dx\ s^-(x,Q^2) &= \int_0^1 dx[s(x,Q^2) - \bar{s}(x,Q^2)] = 0.
\end{aligned} \tag{1.63}$$

Additional constrains can be inferred about the PDF structure from momentum and valence sum rules. Eq. (1.62) requires that the first momentum of singlet and gluon distributions must be integrable while the integrability of the distributions themselves is not needed. Furthermore these parton densities must vanish in the limit where $x \to 1$ in order to mitigate large contribution to the integral. Moreover the valence sum rules in Eq. (1.63) enforce the integrability of the light quark PDFs over the whole x range. The information obtained from momentum and valence sum rules constrain the small and large x behavior of the singlet and valence PDFs and thus we may isolate these dominant contributions in the following way:

$$\begin{aligned}
q^-(x,Q^2) &= N_V\ x^{\alpha_V}(1-x)^{\beta_V} r_V(x), \\
\Sigma(x,Q^2) &= N_\Sigma\ x^{\alpha_\Sigma}(1-x)^{\beta_\Sigma} r_\Sigma(x),
\end{aligned} \tag{1.64}$$

where the coefficients α and β control the small and large x regions respectively. These coefficients along with normalization N must implement the aforementioned

considerations inferred from the sum rules. The remaining term $r(x)$ describes the functional form of the parton densities between the two x-limits and represents the main object of research in the PDF determination.

Chapter 2

Parton density representation

A precise description of the proton structure is a crucial task in modern particle physics as the parton densities connect the partonic dynamics to the physical observables. The predictive power of a theory is therefore tightly correlated to an accurate assessment of PDFs and their uncertainties. Since hadrons consist in strong coupled QCD bound states, the perturbative approach to PDF determination is doomed to fail. Although a great deal of effort and progress has been made in understanding PDFs through non perturbative methods such as Lattice QCD [4], results are far from the accuracy requirements of practical applications at hadron colliders.

The determination of PDFs thus follows the same procedure of other QCD parameters, i.e. fitting appropriate experimental data. The main difficulty in parton density fits consists in the determination of a function rather than a single parameter from a finite sample of experimental measurements. Moreover, the experimental uncertainties introduce fluctuations in the functional space of PDFs and therefore an accurate description of the functional probability distribution of PDFs is mandatory in order to understand the parton density contribution in theoretical predictions.

2.1 Experimental data

All the different approaches to the parton density determination are based on the selection of the datasets used in the fitting procedure. The experimental data represent thus the backbone of PDF analysis and are selected in order to provide the maximum sensibility to the parton densities. Furthermore, the measured points entering the fit must be described by accurate theoretical predictions in order to minimize the PDF uncertainties. For this purpose experimental cuts are introduce to remove e.g. non perturbative kinematic regions where non perturbative

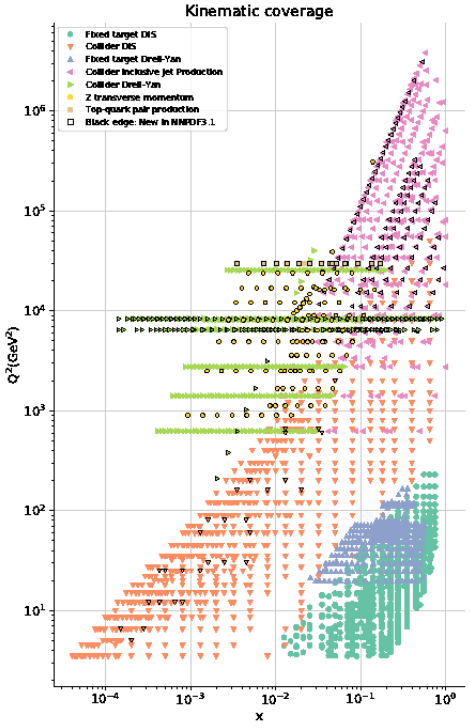

Figure 2.1: *Kinematic coverage adopted by the NNPDF collaboration [5].*

corrections become relevant.

In the following sections the most important processes in the determination of PDFs are presented, focusing on how the experimental data affect parton density combinations.

2.1.1 Fixed target and collider DIS

Deep inelastic scattering data represent the wider and more important dataset for PDF analysis. The electron-proton scattering data from HERA and SLAC explore the medium and large x region and improve statistical and systematic uncertainties at medium and high Q^2. This dataset provides the most reliable probe of the proton electromagnetic structure functions while fixed-target DIS experiments introduce important constrains at high x. Moreover, the neutral current DIS measurements shed light on $q_i + \overline{q}_i$ combination and the Z mediated processes give information about the PDF flavour separation through F_3 structure function. In addiction to proton measurements, the data obtained from scattering off deuterium target provide constrains on light quark PDFs, namely $u - d$ and u/d combinations if the isospin symmetry is assumed.

Although the gluon contribution at DIS appears at NLO, the scaling violations

of structure functions and the wide range of energy scale covered by the data ensure an indirect but robust determination of gluon PDFs. DIS has been the historical benchmark for QCD and parton model predictions and it represents one of the best understood scattering process since his theoretical predictions are known at second order in α_S^2 with full heavy quark mass contributions. Moreover, the clean ep or μp environment provides datasets unaffected by nuclear correction even if low energy data may suffer from non perturbative corrections. Appropriate kinematic cuts avoid this problem.

2.1.2 Neutrino DIS

The scattering processes involving neutrino beams probe the electro-weak behavior of the proton and allow the measurement of F_2^ν and F_3^ν structure functions whose form in terms of partonic densities at leading is given by (suppressing CKM matrix element):

$$\begin{aligned} F_2^\nu(x) &= x\left(u^+(x) + d^+(x) + 2s(x) + 2\overline{c}(x)\right), \\ F_2^{\overline{\nu}}(x) &= x\left(u^+(x) + d^+(x) + 2\overline{s}(x) + 2c(x)\right), \end{aligned} \tag{2.1}$$

and for F_3 we have:

$$\begin{aligned} F_3^\nu(x) &= x\left(u^-(x) + d^-(x) + 2s(x) - 2\overline{c}(x)\right), \\ F_3^{\overline{\nu}}(x) &= x\left(u^-(x) + d^-(x) - 2\overline{s}(x) + 2c(x)\right). \end{aligned} \tag{2.2}$$

The fit of these distributions allows us to obtain a better description of valence quark combination $q - \overline{q}$. However the lack of an accurate description of large nuclear corrections could lead to a misleading assessment of the uncertainties.

A further contribution from neutrino physics concerns semi-inclusive dimuon production process $pN \to \mu\mu X$ which gives a direct handle on the strange distribution $s(x)$ which is Cabibbo favoured. These experimental measurements have been provided by NuTeV/CCFR collaboration and introduce important constrains upon strange quark distribution.

2.1.3 Drell-Yan and boson production

The measurements of electroweak boson production and Drell-Yan cross sections have became really accurate in the LHC era and represent the most important dataset after DIS results. In particular, Drell-Yan phenomenology views the production of a pair of leptons from the scattering of two initial state hadrons. Since the neutrinos pair measurements in scattering processes represent an extremely prohibitive task, the interest in Drell-Yan experiments focuses on the lepton-lepton

and lepton-neutrino pairs originating from γ^*/Z and W mediated interactions respectively as shown in Fig. 2.2. The relevant kinematic variables in neutral current process are the invariant mass of the lepton pair which depends on their energy and momentum by

$$M_{ll}^2 = (E_1 + E_2)^2 + (\boldsymbol{p}_1 + \boldsymbol{p}_2)^2, \tag{2.3}$$

and the intermediate boson rapidity, given in the detector frame by

$$y = \frac{1}{2} \log \frac{E + p_L}{E - p_L}, \tag{2.4}$$

with E and p_L the energy and the longitudinal momentum of the intermediate boson respectively.
These kinematic variables are directly associated to the Bjorken scaling parameter by the following relation:

$$x_{\pm} = \frac{M_{ll}}{\sqrt{s}} e^{\pm y}, \tag{2.5}$$

where s is the centre of mass energy squared and the $\pm$ denotes the parton direction with respect to the beam frame. Therefore high rapidity measurements provide precious information about the PDF behavior in both high and low x regions. When charged current interactions are taken into account, the neutrino in the final state can not be detected and the resolution of W rapidity deteriorates. This problem can be avoided by expressing data in terms of the pseudorapidity of the detected lepton, namely:

$$\eta = -\log \tan \theta, \tag{2.6}$$

where θ is the angle between the lepton direction and the beam axis.
Another relevant contribution to the PDF determination is the lepton asymmetry in W-mediated Drell-Yan and it is defined by:

$$\mathcal{A}_W^l = \frac{d\sigma_{l^+}/d\eta_l - d\sigma_{l^-}/d\eta_l}{d\sigma_{l^+}/d\eta_l + d\sigma_{l^-}/d\eta_l} \tag{2.7}$$

and $d\sigma_{l^\pm}/d\eta_l$ refers to the differential cross section for $W^\pm \to l^\pm \nu_l$.
Generally Drell-Yan measurements with fixed target (hydrogen or deuterium) provide a relatively clean probe of u/d PDF combination but poorly understood nuclear corrections from deuterium scattering could affect theoretical predictions.

On the contrary, collider experiments provide the theoretical cleanest environment for the Drell-Yan process since the high energy scale suppresses non perturbative effects. Data from $p\overline{p}$ interactions at Tevatron include neutral current cross sections, asymmetry measurements, Z rapidity distribution and lepton asymmetry in Eq. (2.7). These datasets provide important information on u/d ratio and

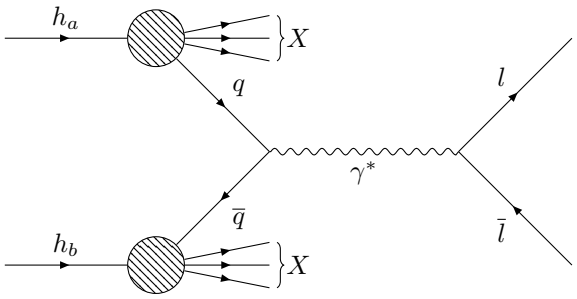

Figure 2.2: *Drell-Yan process mediated by a virtual photon.*

quark valence distributions. Moreover, less inclusive processes such as W production in association with a charm jet give relevant constrains upon the strange PDF because of the large strange-charm CKM matrix element.

The measurements of vector boson production and Drell-Yan cross sections from LHC experiment represent another conspicuous contribution to PDF fit. ATLAS and CMS provide precise measurements of W and Z distribution in both rapidity and transverse momentum along with lepton charge asymmetry while the LHCb sensibility to vector boson production in the very forward region probes high rapidity regime.

2.1.4 Jet production data

So far the constrains upon the gluon PDF have been produced in a quiet indirect way from scaling violations in DIS experiments. The request of an explicit determination of gluon parton density is met by jet production measurements which provide important information in the large x region. Jets are narrow cone of particles produced by the hadronization of quark and gluon radiation and are experimentally reconstructed from single particle data via appropriate clustering algorithms. These clustering algorithms are required to provide a good description of the jet structure and must satisfy QCD infrared and collinear safety theorem. More recent experiments typically utilise sequential-combination algorithms such as the Cambridge-Aachen [6], k_T [7] or anti-k_T [8] algorithms.

The cross section for the inclusive jet and dijet (i.e. the emission of a pair of jets) data in hadron-hadron collisions are known at NLO and only approximate NNLO results are available. The inclusive jet and dijet cross section measurements are available from CDF, D0, ATLAS and CMS experiments.

2.2 Treatment of multiplicative uncertainties

In general the task of the fitting procedure is to provide the best estimation for an unknown parameter of a given theory by requiring the maximum likelihood between the theoretical predictions of a specific observable and its experimental measurements. Moreover the fitting process must provide an accurate assessment of how the parameter probability distribution depends on the experimental uncertainties.

Since the parton densities can be viewed as QCD functional parameters, the problem of their determination consists in providing the best estimation of a functional probability density in the space of PDFs from a finite number of experimental data. In general this task can be achieved by introducing an appropriate error function χ^2 that measures the fit quality and commonly is defined by:

$$\chi^2 = \sum_{i,j}^{N_{dat}} (t_i - m_i)(\mathrm{Cov}^{-1})_{ij}(t_j - m_j), \tag{2.8}$$

where t are the theoretical predictions which depend on the PDFs, m are the experimental data points and Cov^{-1} is the inverse of the experimental covariance matrix. Although Eq. (2.8) represents the most common definition of χ^2, other definition can be adopted as we shall see later. The full experimental uncertainty information is contained in the covariance matrix that is characterized by the sum of three different contributions:

$$\mathrm{Cov}_{i,j} = \sigma_i^{unc}\sigma_j^{unc} + \sum_k^{N_{add}} \sigma_{ik}^{add}\sigma_{kj}^{add} + \sum_k^{N_{mul}} (\sigma_{ik}^{mul}\sigma_{kj}^{mul}) m_i m_j, \tag{2.9}$$

where σ_i^{unc} is the uncorrelated uncertainty for the data point m_i, σ_{ik}^{add} is the correlated addictive systematic error between m_i and m_j and σ_{ik}^{mul} is the correlated multiplicative systematic error. Typically the main sources of multiplicative uncertainties are the normalization of the data points. This method of constructing the covariance matrix is unambiguously defined by the experimental results but becomes unreliable for use directly in fitting procedure. Indeed, the multiplicative uncertainties are proportional to the value of the data points and therefore smaller data points are assigned a smaller uncertainty than bigger data points. It can be shown that these uncertainties introduce a bias [9] in the fitting procedure when combining datasets from independent experiments since the theoretical predictions determined via χ^2 minimization are systematically shifted lower than the true values. In particular this effect worsens as the number of the points that share the same multiplicative error increases.

A typical method to avoid this bias consists in including the normalizations in the fitting procedure as unknown parameters with an additional penalty terms to

the χ^2 to avoid large deviations of the normalizations n_k in terms of its experimental uncertainty s_k. Eq. (2.8) thus becomes:

$$\chi^2(t,n) = \sum_k^{N_{mul}} \sum_{i,j}^{N_{dat}} (t_i/n_k - m_i)(\mathrm{Cov}^{-1})_{ij}(t_j/n_k - m_j) + \sum_k^{N_{mul}} \frac{(1-n_k)^2}{s_k^2}. \quad (2.10)$$

This procedure mitigates the effects of multiplicative uncertainties but still suffers from the bias, especially when combining several normalizations from different experiments.

A full unbiased description of the experimental covariance matrix is available using the t_0 prescription introduced by the NNPDF collaboration [10]. The basic idea of this method consists in multiplying the normalization uncertainties for a fixed value t_0 rather than using the experimental data points m. The value of t_0 is determined before the fit and can be tuned to be consistent with the final result t. Typically self-consistence is achieved using an iterative procedure that adopts the results of a previous fit as the new input t_0 for the subsequent fit and generally convergence is very rapid. Assuming t_0 prescription, the χ^2 becomes:

$$\chi^2 = \sum_{i,j}^{N_{dat}} (t_i - m_i)(\mathrm{Cov}_{t_0}^{-1})_{ij}(t_j - m_j), \quad (2.11)$$

where the covariance matrix is given by:

$$(\mathrm{Cov}_{t_0})_{i,j} = \sigma_i^{unc}\sigma_j^{unc} + \sum_k^{N_{add}} \sigma_{ik}^{add}\sigma_{kj}^{add} + \sum_k^{N_{mul}} (\sigma_{ik}^{mul}\sigma_{kj}^{mul})t_{0,i}t_{0,j}. \quad (2.12)$$

2.3 Representation of PDF uncertainties

Fitting a function is a procedure that requires an infinite number of degrees of freedom and since the experimental data are always a finite number this problem is theoretically under-constrained. Nevertheless we can cope with this issue by projecting the infinite-dimensional space of parton densities onto a finite-dimensional subspace. In fact PDFs are supposed to be smooth function of the variable x and therefore they may be represented with a finite accuracy in terms of an appropriate basis in a finite-dimensional subspace. For this reason it is possible to express the parton densities in terms of a parametrization with a finite number of parameters. The problem is then reduced to find the optimal parametrization that provides the best description of parton densities without introducing potential biases.

Once the parametrization is fixed, the fitting procedure is well defined and typically depends on how the PDFs and their uncertainties are represented. Indeed

parton distributions are delivered for practical applications as a set of member functions which contains the best fit result, also called central value, along with an error set that allows the propagation of uncertainties and is determined by following two main strategies: the Hessian approach and the Monte Carlo approach. This two strategies are presented in the following sections along with the choice of the parametrization adopted in each case.

2.3.1 Hessian representation

In the Hessian approach the parton densities are described by introducing a fixed parametrization based on theoretical constrains such as the sum rules. The PDFs at some fixed scale Q_0^2 are thus determined by a set of parameters and the probability density in the space of PDFs is replaced by a probability density in the parameter space which is assumed to follow a multivariate gaussian distribution. For example, one possible parametrization of parton densities is

$$xf(x, Q_0^2) = a_0 x^{a_1}(1-x)^{a_2} \exp[a_3 x + a_4 x^2 + a_5\sqrt{x} + a_6 x^{a_7}] \tag{2.13}$$

with different parameter sets $\vec{a} = (a_1, a_2, ...)$ for each flavour. The functional form in Eq. (2.13) was used in the early fits while modern Hessian approaches adopt more advanced functional form. The parameters $\vec{a}$ are then fixed by minimizing an appropriate error function like χ^2 which becomes a functional in the parameter space given by:

$$\chi^2(\vec{a}) = \sum_{i,j}^{N_{dat}} (t_i(\vec{a}) - m_i)(\mathrm{Cov}^{-1})_{ij}(t_j(\vec{a}) - m_j), \tag{2.14}$$

where the theoretical predictions $t(\vec{a})$ depend on the parameters $\vec{a}$ through PDF evolution. The best fit set of parameters $\vec{a}_0$ is defined so that $\chi^2(\vec{a}_0)$ corresponds to the absolute minimum of χ^2 and under the assumption that the minimum is unique and that the error function is quadratic around the minimum, a small deviation from $\vec{a}_0$ induces an increment of the χ^2 given by

$$\Delta\chi^2(\vec{a}) = \chi^2(\vec{a}) - \chi^2(\vec{a}_0) = (\vec{a} - \vec{a}_0)H(\vec{a} - \vec{a}_0), \tag{2.15}$$

where H is the Hessian matrix of the error function χ^2 (with an extra factor 1/2) evaluated at the minimum,

$$H_{ij} = \frac{1}{2} \left. \frac{\partial^2 \chi^2(\vec{a})}{\partial a_i \partial a_j} \right|_{\vec{a}=\vec{a}_0}. \tag{2.16}$$

It can be shown that the inverse of H represents the covariance matrix C that describes the multivariate gaussian distribution of the parameters that is given by:

$$p(\vec{a}) = \frac{(\det(C))^{-\frac{1}{2}}}{\sqrt{(2\pi)^N}} \exp\left[\frac{1}{2}\sum_{ij}^{N}(a-a_0)_i C_{ij}^{-1}(a-a_0)_j\right], \tag{2.17}$$

where N is the number of parameters. If we now define the shift in the parameter space as $\vec{\delta} = \vec{a} - \vec{a}_0$ then Eq. (2.15) reduces to:

$$\Delta\chi^2 = \vec{\delta} H \vec{\delta}. \tag{2.18}$$

Since H is a real symmetric matrix, it can be diagonalized in terms of a complete orthonormal basis of eigenvectors $\{\vec{v}_i\}$ with $i = 1, ..., N_{eig}$. It is now convenient to rescale each eigenvector $\vec{v}_i$ by the square root of its eigenvalue λ_i and we may define the new rescaled basis as $\vec{e}_i = \vec{v}_i/\sqrt{\lambda_i}$. The decomposition of $\vec{\delta}$ in the new basis can be expressed in the following way:

$$\vec{\delta} = \sum_i z_i \vec{e}_i, \tag{2.19}$$

with $z_i = \vec{\delta} \cdot \vec{e}_i$. Inserting this decomposition in Eq. (2.18), we find:

$$\Delta\chi^2 = \sum_i z_i^2, \tag{2.20}$$

which describes a hypersphere with radius $\sqrt{\Delta\chi^2}$ in the parameter space and the one-σ contour (i.e. the 68% confidence level) around the central value is defined by the condition $\Delta\chi^2 = 1$, known as parameter-fitting criterion. Furthermore, the choice of the basis $\{\vec{e}\}$ allows also the diagonalization of the matrix C since $C = H^{-1}$ and thus the multivariate gaussian distribution in Eq. (2.17) reduces to a product of independent univariate gaussian distributions for each eigenvector.

Assuming linear error propagation, the Hessian set of PDFs is composed by a central value that corresponds to the best fit parameters $\vec{a}_0$ while the error set corresponds to a shift in the parameter space from $\vec{a}_0$ along the direction of each rescaled eigenvector $\vec{e}_i$ given by:

$$\vec{a}^{(i)} = \vec{a}_0 + t\vec{e}_i, \tag{2.21}$$

where $t = \sqrt{\Delta\chi^2}$ and is set to one for one-sigma deviation.

Any quantity $\mathcal{O}$ which depends on PDFs is also a function of the parameters in the Hessian method and its best estimation is therefore $\mathcal{O}(\vec{a}_0)$ while the uncertainty

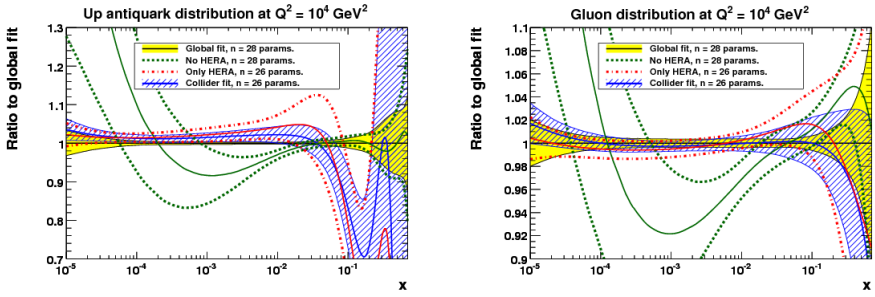

Figure 2.3: *Effect on the up and gluon PDFs of fitting subsets of MSTW 2008 global data from Ref. [14].*

induced on $\mathcal{O}$ by the Hessian error set can be calculated in the following way assuming again linear error propagation:

$$\sigma_{\mathcal{O}} = \sqrt{\sum_{i}^{N} \left(\mathcal{O}(\vec{a}^{(i)}) - \mathcal{O}(\vec{a}_0)\right)^2}. \tag{2.22}$$

The relation in Eq. (2.22) holds also when the observable $\mathcal{O}$ is the PDFs themselves and allows the determination of the PDF standard deviation σ and hence the one-σ band, namely the symmetric interval centred in the best fit value with amplitude σ.

However, in practical applications the parameter-fitting criterion $\Delta\chi^2 = 1$ is not adequate since the one-sigma band does not provide a realistic estimation of the parton density fluctuations. In particular the results for the global fit, i.e. the fit performed with the maximum number of available data points, and the results fitted from a dataset of a single experiment are generally not compatible within their error bands produced with $\Delta\chi^2 = 1$ criterion, as shown in Fig. 2.3. This tension between the results from different datasets emerges also when the minimum chi squared for the global fit χ^2_{tot} is compared to the minimum chi squared χ^2_i of the single dataset: it can be shown [13] that often the difference between χ^2_i and χ^2_{tot} can not be explained in terms of statistical fluctuations.

This suggests that a correct assessment of global fit uncertainties must take into account the compatibility with each dataset. This requirement is met by introducing the concept of tolerance, namely the PDF error bands for the global fit are produced by tuning the terms t in Eq. (2.21) known as tolerance in order to provide a good agreement with the results of each dataset. Since $t^2 = \Delta\chi$, the tolerance coincides with the radius of the hypersphere defined in Eq. (2.20) that represents the one-sigma contour in the space of parameters. The optimal value for the tolerance therefore corresponds to the minimum radius that accommodates

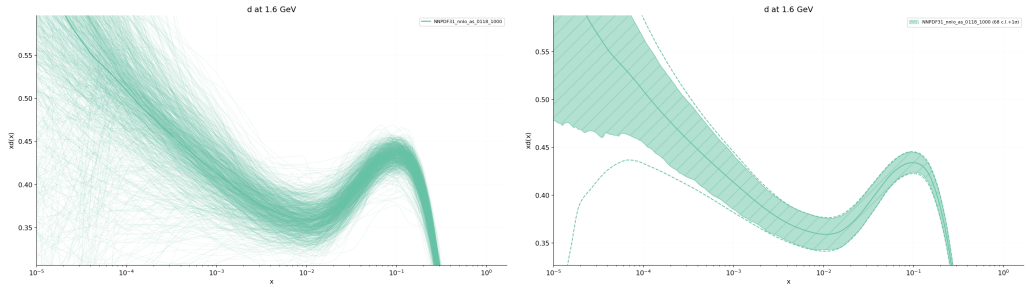

Figure 2.4: *Down quark PDF from global fit at* $Q^2 = 1.65$ GeV^2 *produced by NNPDF3.1 collaboration: the left plot shows all the 1000 replicas and the central values while the right plot shows the one-sigma band around the central value (solid line) and the 68% band (dashed line).*

all the minima $\vec{a}_{0,i}$ of each experiment within their one-sigma contours produced with the parameter-fitting criterion $\Delta\chi^2 = 1$ which is described by an hypersphere with unitary radius centred in $\vec{a}_{0,i}$. This procedure may be refined by introducing a different tolerance t_k for each eigenvector instead of a global tolerance for all eigenvectors. This approach is called 'dynamical' tolerance and typically t_k ranges in $2 < t_k < 5$ while the global tolerance method requires $t \simeq 10$. The dynamical tolerance thus mitigates the large variation from $\Delta\chi^2 = 1$ of the global tolerance method.

Although the tolerance was introduced to cope with the incompatibility of the datasets, other explanations for the need of a tolerance have been studied and can be summarized in two main categories: dataset incompatibilities and parametrization bias. The dataset incompatibilities are related to the uncertainties of the quantities entering the fit and include the discrepancies between experimental datasets, non gaussian deviations and theoretical uncertainties due to missing higher order perturbative corrections or QCD parameter uncertainties (such as α_S). An assessment of dataset incompatibilities can be performed studying the variation of $\Delta\chi^2$ when a new experiment is added to the fit or analyzing the error propagation of pseudodata produced from 'a priori' known distribution. Instead the parametrization bias depends on how the choice of a particular parametrization affects the uncertainty propagation through the fitting procedure and its effects can be determined by adopting a more flexible or constraining parametrization. Both dataset incompatibilities and parametrization bias have been studied [14] and generally they contribute evenly to the determination of tolerance.

2.3.2 Monte Carlo representation

In the Monte Carlo approach the probability distribution in the parameter space is given by a Monte Carlo sample, i.e. a list of sets of parameters that describes the unknown underlying distribution. Each set of parameters is associated to a function in the PDF space called Monte Carlo replica. The error set in this representation is thus constituted by N_{rep} Monte Carlo replicas and the best fit is given by averaging over all the replicas. Any quantity $\mathcal{O}$ that depends on PDFs, including PDFs themselves, assumes a different value $\mathcal{O}_k$ for each replica and the dependence of $\mathcal{O}$ from the PDF fluctuations can be evaluated by computing statistical estimators over the set of values $\{\mathcal{O}_k\}$. For example, the central value is given by the sample mean defined by

$$\langle \mathcal{O} \rangle = \frac{1}{N_{rep}} \sum_{k}^{N_{rep}} \mathcal{O}_k, \tag{2.23}$$

while the uncertainty due to PDF fluctuations is the standard deviation of $\{\mathcal{O}_k\}$:

$$\sigma_{\mathcal{O}} = \sqrt{\frac{1}{N_{rep}-1} \sum_{k}^{N_{rep}} (\mathcal{O}_k - \langle \mathcal{O} \rangle)^2} \quad . \tag{2.24}$$

Unlike Hessian approach, Monte Carlo parametrization does not assume *a priori* distribution and therefore is sensitive to non gaussian effects. Moreover, a practical way to assess non gaussianity consists in comparing the one standard deviation band around the central PDF and the 68% confidence level band which contains the central 68% of the replica sample: since a gaussian distribution requires these two bands to coincide, a difference between them is a symptom of a non gaussian behavior. However in most cases the one-sigma band does not differ too much from the 68% band with the exception of large and small x regions as shown in Fig. 2.4.

Unlike the Hessian representation, the Monte Carlo representation does not require the introduction of the tolerance since there are no contrasts between global and data subset results. The Monte Carlo approach is adopted by the NNPDF [11] collaboration with a parametrization based on neural network [12]. Neural networks allow an extremely flexible parametrization and are 'unbiased' since they are able to fit a very wide class of functions with a finite number of parameters without adjusting the functional form according to the problem.

2.4 Hessian conversion of a Monte Carlo set

The way PDF uncertainties are determined in the Hessian approach deeply differs from the Monte Carlo representation: the former is founded on the knowledge of

the χ^2 in the vicinity of the minimum and requires the introduction of a tolerance to provide reliable predictions while the latter requires the calculation of statistical estimators and does not rely on a tolerance. Nevertheless, both these approaches must provide in principle the same description of the structure of the proton and therefore it may be possible to pass from a representation to the other without (too much) loss of information.

The Monte Carlo representation of a Hessian set can be carried out with ease by drawing a Monte Carlo sample of parameters from the gaussian distribution provided by the Hessian representation. The list of replicas in the space of parameters is then uniquely associated to a Monte Carlo list of replicas in the space of PDFs through the Hessian parametrization and thus the new Monte Carlo distribution is by construction gaussian.

Instead the Hessian conversion of a Monte Carlo set requires a more careful treatment. First the Monte Carlo representation does not require the assumption of gaussianity which on the contrary represents the fundamental hypothesis of the Hessian approach. Therefore, a Hessian representation of a Monte Carlo set can only make sense if the Monte Carlo distribution is gaussian. Whereas deviation from gaussianity may be important in specific kinematic regions, typically when PDF uncertainties are dominated by theoretical constraints due to limited experimental points, the assumption of gaussianity provides a good approximation, especially when PDF fluctuations are driven by a wide number of experimental data which follow a gaussian distribution.

Once gaussianity is provided,the basic idea of the Hessian conversion [15] is to construct a covariance matrix in the space of Monte Carlo replicas that allows a gaussian representation of PDF distribution. The eigenvectors of the covariance matrix are then represented as a linear combination of replicas. We shall assume that the prior Monte Carlo set contains N_{rep} replicas $\{f_\alpha^{(k)}\}, k = 1, .., N_{rep}$ where $\alpha = 1, .., N_{pdf}$ represents the type of PDF with $N_{pdf} = 2N_q + 1$ and N_q is the number of quarks. We may define a discrete covariance matrix by introducing a sample of N_x points in the x-space for each PDF flavour. This sampling requires only to be fine grained enough that the differences between neighboring points are no-negligible. The $(N_x N_{pdf} \times N_x N_{pdf})$ covariance matrix is thus defined by the corresponding statistical estimator in the space of PDFs:

$$\mathrm{cov}_{lm} = \frac{N_{rep}}{N_{rep}-1}\left(\langle f_\alpha^{(k)}(x_i)\ f_\beta^{(k)}(x_j)\rangle_{rep} - \langle f_\alpha^{(k)}(x_i)\rangle_{rep}\langle f_\beta^{(k)}(x_j)\rangle_{rep}\right), \qquad (2.25)$$

where $l = N_x(\alpha - 1) + i$ and $m = N_x(\beta - 1) + j$ run over all $N_x N_{pdf}$ points and $\langle\ \cdot\ \rangle_{rep}$ is the average over all replicas. An analogous definition of Eq. (2.25) can be provided introducing the rectangular $(N_x N_{pdf} \times N_{rep})$ matrix:

$$X_{lk} = f_\alpha^{(k)}(x_i) - f_\alpha^{(0)}(x_i), \qquad (2.26)$$

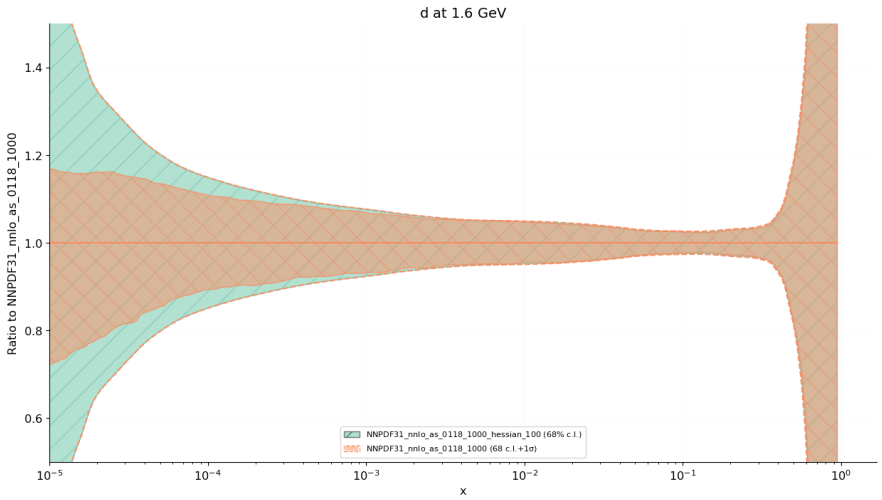

Figure 2.5: *Comparison between the Monte Carlo uncertainties (red) and the Hessian conversion uncertainties with 100 eigenvalues (blue) of the down PDF normalized to the central PDF.*

and $f^{(0)}$ represents the central value PDF. The covariance matrix can be expressed in terms of X by the following relation:

$$\text{cov}_{lm} = \frac{1}{N_{rep} - 1} XX^T. \tag{2.27}$$

We can diagonalize the covariance matrix using Singular Value Decomposition (SVD) of the matrix X which allows us to factorize X in the following way:

$$X = USV^T, \tag{2.28}$$

where U and V are orthogonal matrices with dimensions $(N_x N_{pdf} \times N_x N_{pdf})$ and $(N_{rep} \times N_{rep})$ respectively and S is a diagonal rectangular matrix with dimensions $(N_x N_{pdf} \times N_{rep})$ whose diagonal elements are called singular values of X and represent the square roots of the eigenvalues of the covariance matrix multiplied by the normalization constant $(N_{rep} - 1)^{-\frac{1}{2}}$. The singular values may be listed in descending order along the diagonal entries of S and it can be shown that the columns of U are the eigenvectors of the covariance matrix. We can define the matrix $Z = US$ which has the property:

$$ZZ^T = US(US)^T = US(V^T V)S^T U^T = (USV^T)(USV^T)^T = XX^T, \tag{2.29}$$

where we used $V^T V = \mathbb{1}$. However Z can be expressed in the following way:

$$Z = US(V^T V) = (USV^T)V = XV \tag{2.30}$$

and thus it provides the sought-for representation of the multigaussian covariance matrix in terms of the original PDF replicas: specifically, V_{kj} is the expansion coefficient of the j-th eigenvector over the k-th replica.

We have thus provided an exact Hessian representation of the covariance matrix in terms of Monte Carlo replicas. However the number of Hessian eigenvalues is equal to $N_{eig}^{(0)} = N_x N_{pdf}$ which is generally large. In practice the smallest eigenvalues will give a modest contribution to the covariance matrix and so we can select a smaller set of $N_{eig} < N_{eig}^{(0)}$ eigenvectors which corresponds to the largest singular values. Due to the ordering of diagonal elements of S, the matrices U and S are therefore replaced by their submatrices u and s with dimensions $(N_x N_{pdf} \times N_{eig})$ and $(N_{eig} \times N_{rep})$. Because s has only N_{eig} non-vanishing diagonal entries, only the $(N_{rep} \times N_{eig})$ submatrix of V contributes. We call this principal submatrix P and replace the matrix V when only the largest N_{eig} eigenvalues are considered. Fig. 2.5 shows a good agreement between the one-sigma bands of the Monte Carlo set and its Hessian conversion with 100 eigenvalues with the exception of small x region where non-gaussian contributions become relevant.

Chapter 3

χ^2 for Hessian converted Monte Carlo set

The Monte Carlo approach allows a definition of parton densities in terms of statistical estimators that ignore the shape of the error function χ^2 in the vicinity of the minimum. On the contrary, the Hessian approach relies on the knowledge of the χ^2 when a parametrization is given; however, for existing Hessian PDF sets the parameter-fitting criterion $\Delta\chi^2 = 1$ does not provide an adequate description of PDF uncertainties and the introduction of the tolerance is needed in order to accommodate the effects of dataset incompatibilities and parametrization bias.

Since the tolerance represents a shift from the minimum of χ^2 along each Hessian eigenvector direction, the need of a tolerance is a requirement that concerns only the Hessian representation of uncertainties and therefore it can not be introduced for Monte Carlo sets unless a Hessian conversion is provided. Indeed, the Hessian representation of a Monte Carlo set of PDFs provides a set of Hessian eigenvectors that describe the one-sigma band of the prior Monte Carlo set. Since each eigenvector is associated to a variation $\Delta\chi^2$ from the minimum of χ^2, the problem of the tolerance for a Monte Carlo set may be treated in terms of its Hessian conversion that can be used to study the χ^2 shape in the proximity of the minimum.

These aspects are investigated in detail in the following sections, focusing on the construction of a reliable Hessian conversion and the description of the χ^2 near the minimum which is directly connected to the problem of the tolerance.

3.1 $\Delta\chi^2$ for Hessian eigenvectors

We first consider the Monte Carlo NNLO global set of PDFs provided by the NNPDF collaboration [5] with 1000 replicas and the Hessian conversion of this

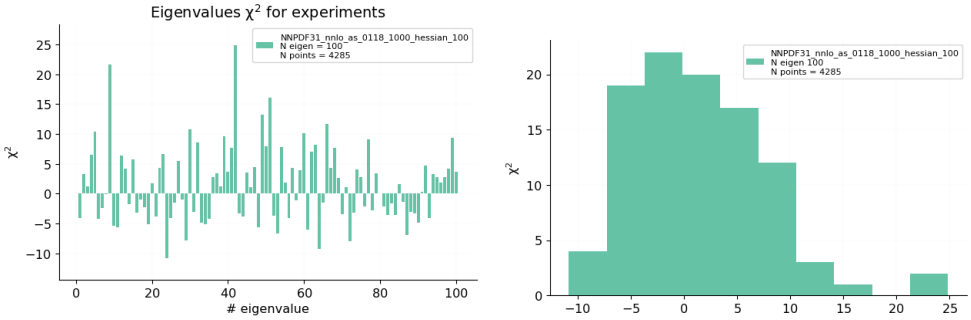

Figure 3.1: $\Delta\chi^2$ *for each eigenvector (left) and the histogram of* $\Delta\chi^2$ *distribution (right) from the Hessian representation with 100 eigenvalues of the prior Monte Carlo NNPDF3.1 NNLO with 1000 replicas.*

prior Monte Carlo set is provided by the code `mc2hessian` [15]. Since the eigenvectors of the Hessian conversion are linear combinations of Monte Carlo replicas and are not determined by any parameter-fitting criterion on the $\Delta\chi^2$, we may thus find the variation $\Delta\chi^2$ induced by each eigenvector, where $\Delta\chi^2$ is the difference between the eigenvector χ^2 and the χ^2 of the central value of the Hessian conversion (which is the same of the prior Monte Carlo set by construction); both these χ^2 are calculated with the definition in Eq. (2.8).

The values of $\Delta\chi^2$ for a Hessian conversion with 100 eigenvalues are presented in Fig. 3.1: not only the $\Delta\chi^2$ for the single eigenvector does not correspond to the parameter-fitting criterion $\Delta\chi^2 = 1$, but also about the 50% of the values of $\Delta\chi^2$ are negative and the $\Delta\chi^2$ ranges in $-10 \lesssim \Delta\chi^2 \lesssim 25$. However, the χ^2 defined in Eq. (2.8) contains the experimental covariance matrix in Eq. (2.9) which adopts the data points for the determination of the multiplicative uncertainties. Since the NNPDF collaboration relies on the t_0 prescription to avoid the bias arising from multiplicative uncertainties, the actual experimental covariance matrix used in the PDF fits is the matrix Cov_{t_0} defined in Eq. (2.12). Therefore the definition of χ^2 must be replaced with Eq. (2.11) which contains the experimental covariance matrix produced with the t_0 prescription. The results obtained from the definition of χ^2 with the t_0 covariance matrix are shown in Fig. 3.2: the range of $\Delta\chi^2$ variations is almost halved and it spans the interval given by $-7 \lesssim \Delta\chi^2 \lesssim 15$ but there is still a large number of negative values of $\Delta\chi^2$. The t_0 prescription therefore mitigates the variation of $\Delta\chi^2$ and henceforth it will be adopted for the computations of the χ^2. It is worth noting that the differences between the experimental covariance matrix and the t_0 covariance matrix are extremely small but produce a considerable effect in the determination of the $\Delta\chi^2$ which also represents a rather small effect compared to the minimum χ^2 which is $\chi^2_{min} \simeq 5070$.

The negative values of $\Delta\chi^2$ however are quite disturbing since they are related

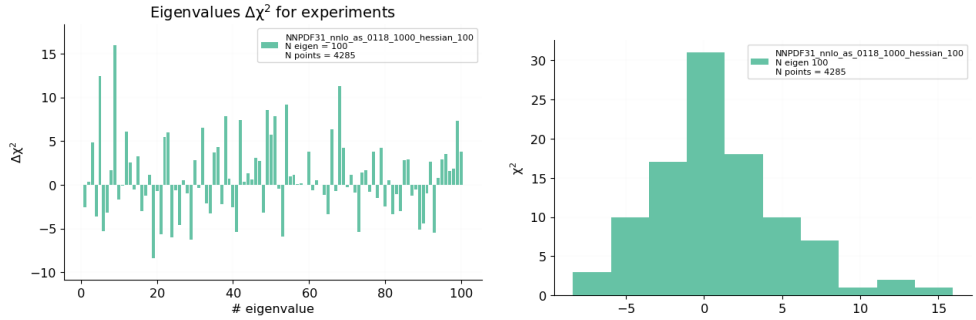

Figure 3.2: *$\Delta\chi^2$ for each eigenvector (left) and the histogram of $\Delta\chi^2$ distribution (right) computed with the χ^2 definition provided by the t_0 prescription. The eigenvectors represent the error set for the Hessian conversion with 100 eigenvalues of the prior Monte Carlo NNPDF3.1 NNLO with 1000 replicas.*

to the directions in the space of PDFs pointed by the corresponding eigenvectors along which the χ^2 decreases over the minimum. This interpretation may suggest an inefficiency in the fitting procedure, i.e. the minimization algorithm does not find the real minimum of the χ^2.

Although the inefficiency could be a possible cause for the negative values of $\Delta\chi^2$, the large variations of $\Delta\chi^2$ can be referred to the fluctuations of the one-sigma contour in the space generated by the basis of the eigenvectors. In particular, this space is completely analogous to the space of the Hessian parameters with dimension equal to the number of eigenvalues adopted in the Hessian conversion. For example, in a ideal case the shape of the χ^2 near the minimum is a multi-dimensional paraboloid in the space of parameters defined by Eq. (2.15) and the one-sigma contour is given by the canonical parameter-fitting criterion $\Delta\chi^2 = 1$. However in the real case the edge of the paraboloid given by $\Delta\chi^2 = 1$ is subjected to large fluctuations that deform the one-sigma contour in the parameter space. Therefore a careful assessment of the possible causes of these fluctuations is needed in order to provide a correct explanation for the negative values of $\Delta\chi^2$.

Since the previous results were produced from a prior Monte Carlo set with 1000 replicas, we may repeat the previous calculations for the $\Delta\chi^2$ of the Hessian eigenvectors with a subset of the prior set with 1000 replicas. The error members of this new set are thus a subset of the 1000 replicas and the new central value is calculated averaging over the new error members. Therefore we may produce 10 batches of Monte Carlo PDFs each containing 100 replicas. We may then compute the $\Delta\chi^2$ for each of these Monte Carlo subsets analogously to the previous case with 1000 replicas. This procedure has the double purpose of testing the $\Delta\chi^2$ dependence from the size of the prior Monte Carlo set and validating the results over a sample of independent batches.

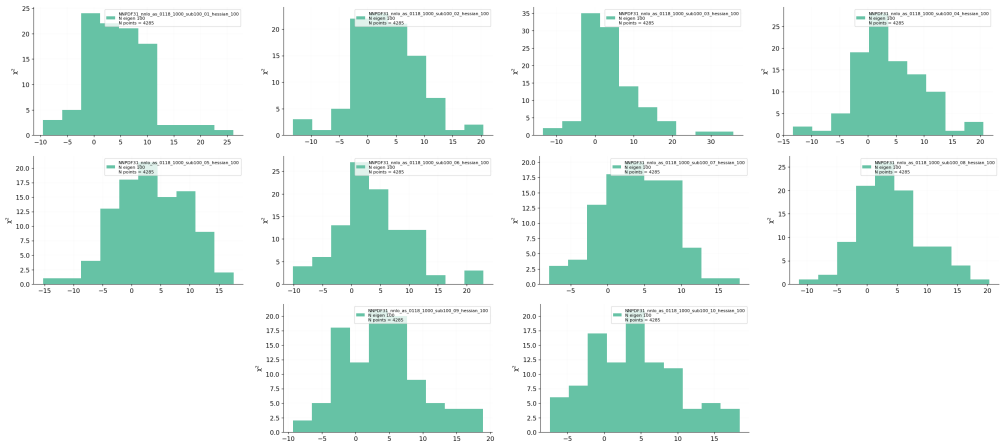

Figure 3.3: *Histograms of $\Delta\chi^2$ distribution for the Hessian conversion with 100 eigenvalues of the 10 Monte Carlo subsets.*

The histograms for the 100 values of $\Delta\chi^2$ for each batch are shown in Fig. 3.3: even if the number of negative $\Delta\chi^2$ is slightly decreased compared to the case with 1000 replicas, the ranges of $\Delta\chi^2$ values span a wider interval than the previous case. Indeed in the most cases the lower extreme of $\Delta\chi^2$ values is $\Delta\chi^2 \simeq -10$ while the upper extreme occurs for $\Delta\chi^2 \simeq 20$. The increase of $\Delta\chi^2$ fluctuations when a smaller number of replicas is considered suggests that a component of $\Delta\chi^2$ fluctuations may be related to finite-size effects. Indeed, the probability distribution given by the Monte Carlo replica sample provides an approximation of the real probability distribution and therefore it is subjected to statistical fluctuations that will vanish in the limit $N_{rep} \to +\infty$.

This hypothesis is corroborated by the comparison between the one-sigma bands of the Hessian converted set with 1000 replicas and 100 replicas: in most cases the one-sigma uncertainties produced from the ten Monte Carlo subsets differ significantly from the uncertainties of the set with 1000 replicas. Fig. 3.4 provides two explicit examples of this discordance. Moreover, Fig. 3.4 shows both the full one-sigma band and the one-sigma band calculated without the contribution of the eigenvectors with negative $\Delta\chi^2$ that we will refer to as negative eigenvectors for simplicity. Since these two bands does not coincide, the negative eigenvectors represent a genuine (and thus non-negligible) contribution to the PDF uncertainties.

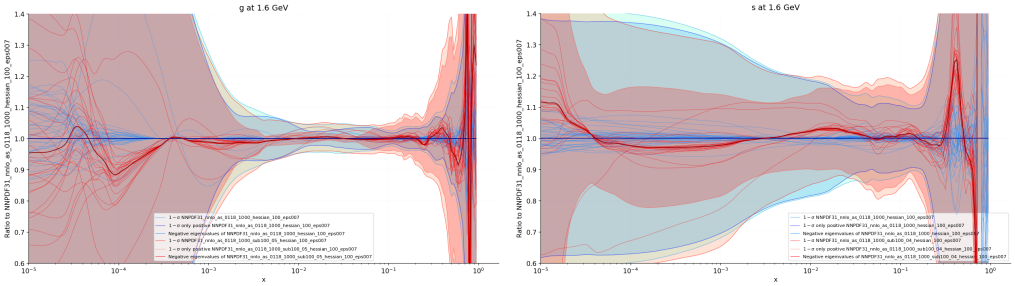

Figure 3.4: *Ratio plot of gluon (left) and strange (right) PDFs given by the Hessian conversion with 100 eigenvalues of the set with 1000 replicas (blue) and the set 100 replicas (red). The inner bands are the one-sigma uncertainties without the contribution of the eigenvectors with negative $\Delta\chi^2$ while the outer bands are the full one-sigma uncertainties. The thin lines are the negative eigenvectors.*

3.2 Non gaussianity

Since the most direct way to get rid of (or at least mitigate) the finite-size effects is to increase the number of replicas, we may attempt to assess the contribution of finite-size effects from $\Delta\chi^2$ fluctuations by producing a Monte Carlo set with the largest number of replicas as possible. Thought so far only global Monte Carlo sets have been considered, we choose to fit only the deep inelastic scattering data in order to provide a large Monte Carlo sample. The reasons of this choice are both practical and theoretical: the deep inelastic scattering provides the most consistent dataset and thus the Monte Carlo fits are extremely faster than the global fits even if DIS dataset contains a great number of data points, namely 3092 experimental points. Furthermore, the consistency of deep inelastic scattering data allows us to remove the potential effects due to dataset incompatibilities.

Fig. 3.5 shows the ratio plots of the light quark and gluon PDFs given by the Monte Carlo set with 3000 replicas produced by the fit at NNLO of deep inelastic scattering datasets. Unlike the global Monte Carlo set with 1000 replicas, these plots manifest a non gaussian behavior across the whole x range; for example the peaks of the 68% band for $x \simeq 0.1$ exhibit a clear deviation from the one-sigma band for the d, $\overline{d}$, u and $\overline{u}$ PDFs.

Since the gaussian assumption is the fundamental hypothesis of the Hessian conversion, a careful treatment of non-gaussian deviations of the Monte Carlo sets must be provided in order to obtain a meaningful Hessian representation. Because the Hessian conversion is possible once the covariance matrix in the space of replicas is provided, we can cope with the problem of non-gaussianity by simply rejecting the points of the x sample for each flavour that do not exhibit a gaussian behavior. For this purpose we may quantify the non gaussianity by introducing

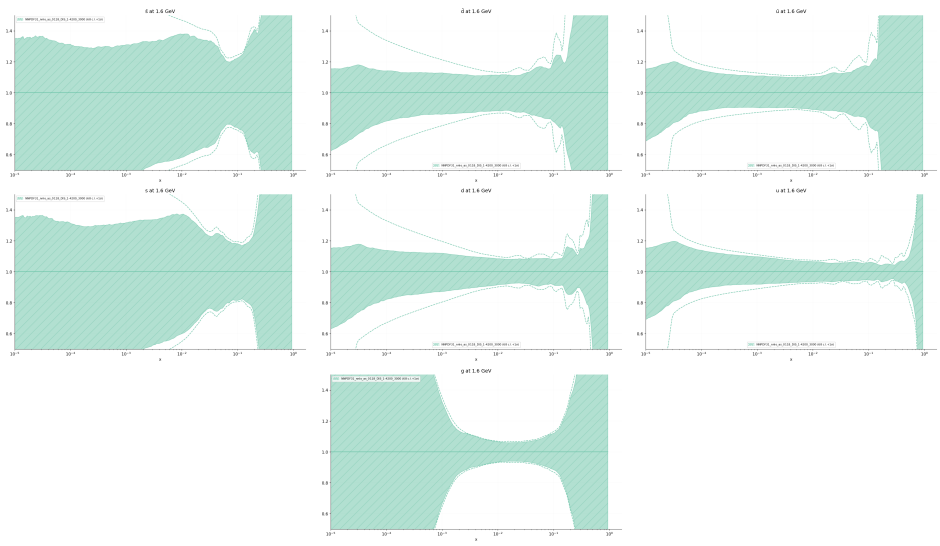

Figure 3.5: *Ratio plots of light quark and gluon PDFs for the DIS only Monte Carlo set with 3000 replicas. The solid lines are the mean value and the one-sigma band while the dashed lines represent the 68% confidence level band.*

the parameter ε which corresponds to the percentage difference between the 68% band ($\sigma_{68\%}$) and the one standard deviation band (σ_{std}) at a given x that thus is defined by:

$$\varepsilon(x) = \frac{|\sigma_{68\%}(x) - \sigma_{std}(x)|}{\sigma_{std}(x)}. \tag{3.1}$$

We may choose a threshold value ε independent of x that can be interpreted as the goodness of gaussian approximation and then discard the x points that give a value of $\varepsilon(x)$ above this threshold.

The ideal Hessian conversion therefore requires a prior Monte Carlo set with $\varepsilon = 0$ along the whole x range but in practice a reliable Hessian conversion requires ε to be small enough so that the gaussian assumption provides a good approximation. Moreover, the optimal value of ε must produce a connected covariance matrix in the space of PDFs, namely the points below ε must provide a connected interval of the x-sample and therefore only the points at large and small x can be removed from the computation of the covariance matrix. We introduce this requirement because the central x region contains a wide number of experimental data whose gaussian uncertainties lead the PDF fluctuations. A large value of ε in this kinematic region is therefore more likely to depend on statistical fluctuations rather than non-gaussianity. Nevertheless we may relax the requirement of a full connected covariance matrix by considering a connected interval in the x range with at most few isolated rejected points.

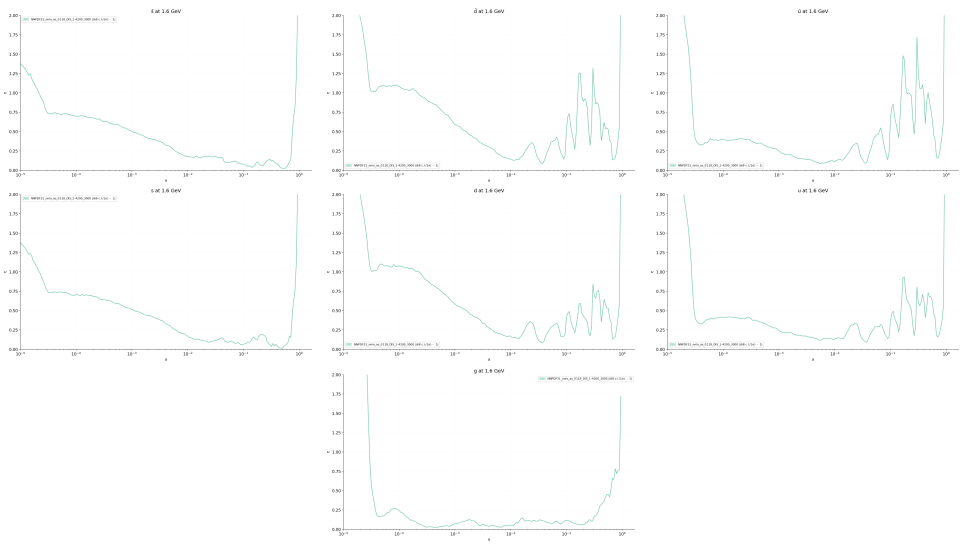

Figure 3.6: *Plots of $\varepsilon(x)$ for the light quark and gluon PDFs from the DIS only Monte Carlo set with 3000 replicas.*

Since the threshold value ε is the same for all flavours, Fig. 3.6 shows that the optimal values of ε for the DIS only Monte Carlo set with 3000 replicas is $\varepsilon = 1.25$ as it manages to accommodate the large peaks at $x \simeq 0.1$ for the $\overline{u}$ and $\overline{d}$ PDFs. Such a large value of ε corresponds to a 68% band that may become twice as large as the one-sigma band and therefore the gaussian assumption provides a poor approximation for this Monte Carlo set. We then attempt to produce a new Monte Carlo set from the previous set with 3000 replicas by introducing more strict replica selection criteria, i.e. removing the outlier replicas with a χ^2 that lies outside the three sigma interval of the χ^2 distribution of all replicas. Even if the ε optimal value for this improved Monte Carlo set decreases to $\varepsilon = 1.0$, we may conclude that such large values of ε for these DIS Monte Carlo sets do not allow a meaningful Hessian conversion.

The experience of the DIS Monte Carlo sets suggests that the global Monte Carlo PDF set may provide a more gaussian behavior since it is determined by a greater number of experimental data. We then repeat the same analysis of non-gaussianity for the global Monte Carlo set and we find that the optimal value for ε corresponds to $\varepsilon = 0.30$ which provides both a reasonable gaussian approximation and a connected x sampling as shown in Fig. 3.7.

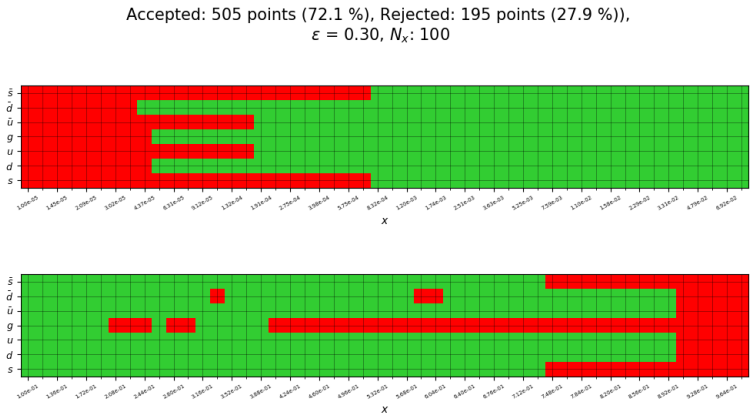

Figure 3.7: *Accepted (green) and rejected (red) x points for the Monte Carlo global set with 1000 replicas with* $\varepsilon = 0.30$ *for the fitted PDF flavours (i.e. the light quarks and antiquarks and the gluon).*

3.3 Hessian conversion with sigma fractions

Since the non-gaussianity of the DIS only Monte Carlo set prevented us to increase the number of replicas we may not remove the finite-size effects. We are then forced to adopt a different strategy that takes into account all the possible contributions to $\Delta\chi^2$ fluctuations that can be summarized as follows:

- **Non-gaussian behavior**: as widely discussed in the previous section, the Hessian conversion requires gaussian uncertainties for the prior Monte Carlo set. We can control non-gaussian deviations within the level of accuracy provided by the parameter ε.

- **Finite-size effects**: these effects are related to the statistical fluctuations of the PDF probability distribution. In principle they can be removed increasing the number of replicas. Alternately, we may assess these effects by analysing the dependence of the $\Delta\chi^2$ fluctuations from the number of replicas N_{rep}.

- **Inefficiency**: since the fitting procedure requires the minimization of the χ^2, a potential inefficiency in the minimization algorithm could lead to the wrong evaluation of the χ^2 minimum which thus introduces a bias when we perform the Hessian conversion.

- **Parabolic deviation**: the assumption of a quadratic χ^2 behavior near the minimum may not provide a good approximation when the actual shape of

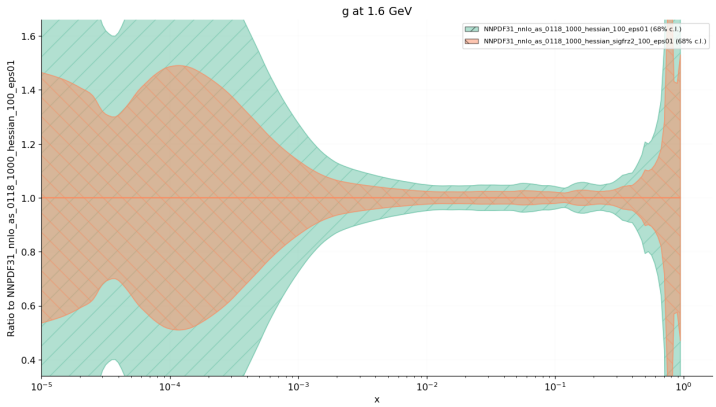

Figure 3.8: *Comparison between the one-sigma (blue) and half-sigma (red) Hessian conversion of the same Monte Carlo set normalized to the central value. The prior Monte Carlo set is the global set with 1000 replicas.*

the χ^2 is more complex. For example, the actual χ^2 could have a valley of equivalent multiple minima or it could have a relative minimum which hides the absolute minimum. Both these examples introduce a deviation from the parabolic assumption which the Hessian conversion relies on.

A possible way to quantify these effects is to produced a Hessian conversion that takes into account different sizes of $\Delta\chi^2$. This procedure allows us to move along the Hessian eigenvectors so that we may explore the shape of the $\Delta\chi^2$ in the vicinity of the minimum.

In general, the construction of a Hessian conversion that takes into account a different size of fluctuations may be carried out by multiplying the PDF covariance matrix by a fixed quantity that corresponds to the size of the sought for representation. For example, a Hessian conversion of a given Monte Carlo sample of an unidimensional variable is determined by the variance of the sample σ^2 and thus the parameter associated to this Hessian conversion follows a gaussian distribution with variance σ^2. If we now want to reduce the amplitude of the parameter fluctuations by a factor 2, we can simply define the variance of the Hessian conversion as $(\frac{\sigma}{2})^2$ instead of σ^2. This trivial example can be generalized to the case of parton densities and therefore we define the Hessian conversion with sigma fraction k as the Hessian set of eigenvectors produced from the Monte Carlo PDF covariance matrix rescaled by a factor k^2. As a consequence of the rescaling of the PDF covariance matrix, the Hessian error band is as well reduced by a factor k as shown in Fig. 3.8.

Since this method allows us to probe the shape of $\Delta\chi^2$, it can be adopted also to study the contribution due to the inefficiency. Indeed in the Hessian representation

the inefficiency of the minimization algorithm implies that the best fit value $\vec{a}_0$ does not correspond to the minimum of χ^2 and therefore the Taylor expansion of the $\Delta\chi^2$ contains a non-vanishing first derivative. The $\Delta\chi^2$ assumes thus the following form:

$$\Delta\chi^2(\vec{a}) = \vec{\bigtriangledown}\chi^2|_{\vec{a}_0} \cdot (\vec{a} - \vec{a}_0) + (\vec{a} - \vec{a}_0)H(\vec{a} - \vec{a}_0). \tag{3.2}$$

A linear term in the expansion of the $\Delta\chi^2$ is therefore tightly correlated to the minimization inefficiency.

3.4 Single parameter model for $\Delta\chi^2$

Once the main sources of $\Delta\chi^2$ fluctuations are provided, we may assess the size of each of these contributions by comparing the $\Delta\chi^2$ results to the theoretical prediction of an appropriate model that describes the expected $\Delta\chi^2$ behavior.

Since the Hessian conversion of a Monte Carlo set is equivalent to a Hessian set with N_{eig} independent parameters that correspond to the basis of the Hessian matrix eigenvectors, we may assume that all these parameters follow the same underlying distribution and therefore it can be possible to describe the $\Delta\chi^2$ in terms of a single parameter model.

We start considering the 'true' Hessian representation that contains the single parameter θ distributed according to a gaussian with mean θ_0 and variance σ^2. Since the $\Delta\chi^2$ is deformed by the aforementioned effects, we may consider the Taylor series expansion of the χ^2 truncated at the fourth order that is given by:

$$\begin{gathered}\chi^2(\theta) = \chi^2(\theta_0) + a\frac{(\theta-\theta_0)}{\sigma} + b\frac{(\theta-\theta_0)^2}{\sigma^2} + c\frac{(\theta-\theta_0)^3}{\sigma^3} + d\frac{(\theta-\theta_0)^4}{\sigma^4} \\ \Delta\chi^2(\theta) = \chi^2(\theta) - \chi^2(\theta_0),\end{gathered} \tag{3.3}$$

where the coefficients a, b, c and d are proportional to the χ^2 derivatives and therefore they directly describe the 'true' shape of the $\Delta\chi^2$. In particular the term a is related to the inefficiency, the term b is the analogous of the tolerance and the terms c and d are related to both non-gaussianity and parabolic deviation. The determination of this coefficients allows us to understand the size of $\Delta\chi^2$ fluctuations and they can be computed by comparing the numerical results to the model predictions.

Moreover we may take into account the finite-size effects by producing a Monte Carlo representation of the parameter θ by drawing a Monte Carlo sample of N replicas $\{\theta_i\}$ from the gaussian distribution of the parameter θ. Due to the statistical fluctuations of the finite-size sample, the mean value μ and the variance s^2 of the Monte Carlo replica sample do not coincide with the 'true' values θ_0 and σ^2 respectively. Once the Monte Carlo representation is provided, we can propagate

the finite-size effects to the parameter θ by producing a Hessian conversion with sigma fraction k of the previous Monte Carlo sample. The procedure of Hessian conversion states that the parameter θ is now distributed according to a gaussian distribution with mean μ and variance s^2/k^2. It can be shown that the sample mean μ and the sample variance s^2 obey the following statements:

- μ and s^2 are independent random variables.
- μ is a random variable that follows a gaussian distribution with central value θ_0 and variance $\frac{1}{N}\frac{\sigma^2}{k^2}$.
- s^2 can be expressed in the following way:

$$s^2 = \frac{\sigma^2}{k^2}\frac{X}{N-1}, \tag{3.4}$$

  where X is a random variable that follows a χ^2_{N-1} probability distribution with $N-1$ degrees of freedom that is given by:

$$p(X; N-1) = \frac{1}{2^{\frac{N-1}{2}}\Gamma(\frac{N-1}{2})}\ X^{\frac{N-1}{2}-1}e^{-\frac{N-1}{2}}. \tag{3.5}$$

Because the Hessian conversion was produced from a finite-size Monte Carlo set, its central values is given by the mean μ of the Monte Carlo sample and the Hessian matrix reduces in the unidimensional case to $\frac{1}{s^2}$, namely the inverse of the sample variance. Moreover, the Hessian conversion assumes a quadratic behavior of the error function near the central value μ given by:

$$\Delta\chi^2_{samp}(\theta) = \frac{(\theta-\mu)^2}{s^2}, \tag{3.6}$$

where the $\Delta\chi^2_{samp}$ refers to the Hessian conversion assumption and must not be confused with the 'true' $\Delta\chi^2$ in Eq. (3.3). The one-sigma contour in the parameter space is thus given by the parameter-fitting criterion $\Delta\chi^2_{samp} = 1$ which corresponds to a shift from the central value μ given by $\mu \pm s$. This shift is the analogous of the variation from the central value along each direction of the rescaled eigenvectors in the multidimensional Hessian representation.

However the shift $\mu \pm s$ induces an increase of the actual χ^2 in Eq. (3.3) which differs from the parameter-fitting criterion $\Delta\chi^2_{samp} = 1$ and it is given by:

$$\Delta\chi^2 = \chi^2(\mu \pm s) - \chi^2(\mu). \tag{3.7}$$

The above expression allows the assessment of the behavior of $\Delta\chi^2$ due to statistical fluctuations of both central value and one-sigma contour so that finite-size

effects can be expressed by computing statistical estimators of $\Delta\chi^2$ in Eq. (3.7). Moreover, we can explicitly write Eq. (3.7) in terms of powers of $\frac{s}{\sigma} = \frac{X}{k(N-1)}$ and $(\mu - \theta_0)$. In particular, the sample mean and the sample standard deviation of the $\Delta\chi^2$ require the computation of the expectation value of powers of $(\mu - \theta_0)$ and $\frac{X}{k(N-1)}$ with respect to their probability distributions. The first case is trivial since $(\mu - \theta_0)$ represents a gaussian random variable with zero mean while the expectation value of powers of $\frac{X}{k(N-1)}$ can be carried out using the probability distribution in Eq. (3.5) and thus we obtain:

$$E\left[\left(\frac{s}{\sigma}\right)^l\right] = E\left[\frac{X^l}{k^l(N-1)^l}\right] = \frac{1}{k^l}\frac{\Gamma\left(\frac{m}{2}+l\right)}{\left(\frac{N-1}{2}\right)^l\Gamma\left(\frac{m}{2}\right)} = \frac{1}{k^l}G_N(l), \tag{3.8}$$

where $G_N(l)$ stands for the expression with the gamma functions and $E[\cdot]$ is the expectation value over the sample of $\Delta\chi^2$. We expect that the mean over a $\Delta\chi^2$ sample differs from the sample mean value within the statistical fluctuations described by the sample standard deviation.

We are now able to compute the sample mean and the sample standard deviation of the $\Delta\chi^2$ which may be expressed in the following way:

$$\Delta\chi^2 = a_N\frac{1}{k} + b_N\frac{1}{k^2} + c_N\frac{1}{k^3} + d_N\frac{1}{k^4}, \tag{3.9}$$

where the dependence from the number of replicas N has been included in the coefficients which then assume the following form along with their standard devi-

ation:

$$
\begin{aligned}
a_N &= \pm a_0 G_N\left(\frac{1}{2}\right), \\
\sigma_a &= a_0\sqrt{1 - G_N^2\left(\frac{1}{2}\right)}, \\
b_N &= 1, \\
\sigma_b &= b_0\sqrt{2\frac{3N-2}{N(N-1)}}, \\
c_N &= \pm c_0\left[G_N\left(\frac{3}{2}\right) + \frac{3}{N}G_N\left(\frac{1}{2}\right)\right], \\
\sigma_c &= c_0\sqrt{\frac{N^3+19N^2+3N-15}{N(N-1)^2} - \frac{c_N^2}{c_0^2}}, \\
d_N &= d_0\frac{N^2+7N-6}{N(N-1)}, \\
\sigma_d &= d_0\sqrt{\frac{(N+1)(N+3)(N+5)}{(N-1)^3} + 28\frac{(N+1)(N+3)}{N(N-1)^2} + 140\frac{(N+1)}{N^2(N-1)} + 540\frac{1}{N^3} - \frac{d_N^2}{d_0^2}},
\end{aligned}
\tag{3.10}
$$

where the coefficients a_0, b_0, c_0 and d_0 are real numbers that do not depend on N.

We may notice that in the limit where $N \to +\infty$ the N dependence of the mean values vanishes while the standard deviations tend to zero. Furthermore, the N dependence of b_N appears only in the standard deviation and the coefficients a_N and c_N are sensible to the direction of the shift from the mean value.

This procedure can also be implemented in a numerical simulation that basically repeats the calculation of the $\Delta\chi^2$ in Eq. (3.7) for a huge number of different Monte Carlo samples in order to obtain a conspicuous set of $\Delta\chi^2$ replicas. We may then compute the expectation values directly on this $\Delta\chi^2$ sample; we find that the numerical results are in good agreement with the model predictions as shown in Fig. 3.9 for the coefficients a_N and b_N.

This model allows us to reduce the problem of the $\Delta\chi^2$ fluctuations due to finite-size, inefficiency, non-gaussianity and parabolic deviation to the determination of the four coefficients a_0, b_0, c_0 and d_0. In particular, there are two contributions due to finite-size effects: the explicit dependence on N of the coefficients a_N, b_N, c_N and d_N and the statistical fluctuations of these coefficients which are described by their standard deviations in Eq. (3.10). Once the finite-size effects are removed from the N dependent coefficients, we find that a_0 describes the contribution due to the inefficiency, b_0 is the analogous of the tolerance and c_0 and d_0

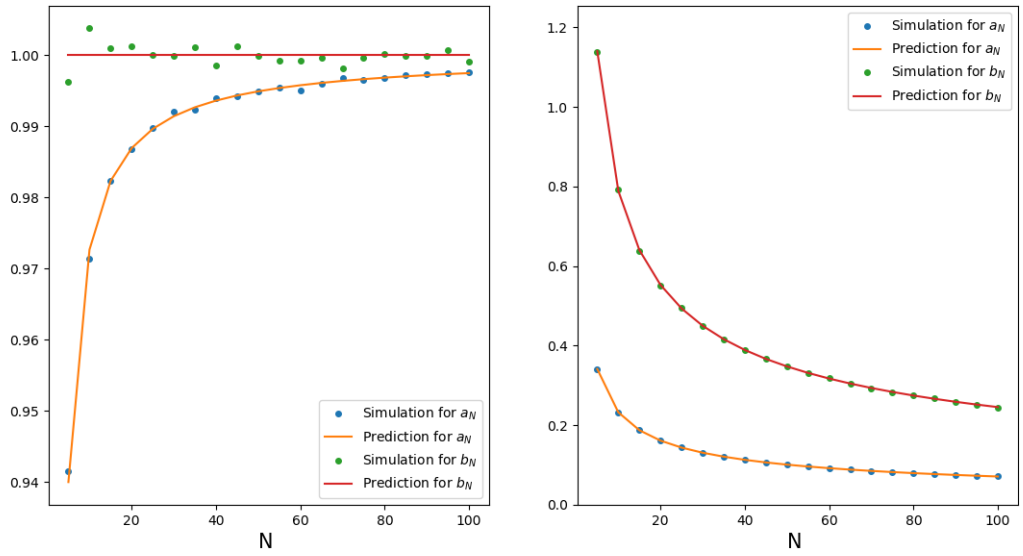

Figure 3.9: *Comparison of the N dependence between the model prediction and the numerical simulation results for the mean value and standard deviation of the coefficients a_N and b_N.*

are related to both non-gaussianity and parabolic deviation.

Since the χ^2 in Eq. (3.3) is a polynomial with respect to $\frac{1}{k}$, we may determine the coefficients by fitting the dependence of $\frac{1}{k}$ for the χ^2 obtained as follows: we may produce from the Monte Carlo global set with 1000 replicas the maximum number of Monte Carlo subsets with a given fixed number of replicas N_{rep}. We then define N_{batch} as the number of Monte Carlo subsets and it represents the quotient between 1000 and N_{rep}.

We therefore produce for each batch of Monte Carlo subsets a Hessian conversion with $N_{eig} = 50$ for several values of the sigma fraction k rejecting the same x points of the aforementioned conversion of the set with 1000 replicas with the optimal value $\varepsilon = 0.3$. We do not impose the ε criterion directly on the Monte Carlo batches because potentially large statistical fluctuations of PDF uncertainties may be mistaken for non-gaussian effects. The final value of χ^2 is thus defined as the average over all the Hessian eigenvectors of all the batches and an estimation of its statistical fluctuations is provided by the sample standard deviation. We adopt this specific strategy because we made the assumption that all the eigenvectors are described by the same distribution.

We calculate the χ^2 with this strategy for different values of k both smaller and greater than 1, namely $k = \{7, 6, 5, 4, 3, 2, 1, 0.66, 0.5, 0.4, 0.33, 0.2, 0.166,$

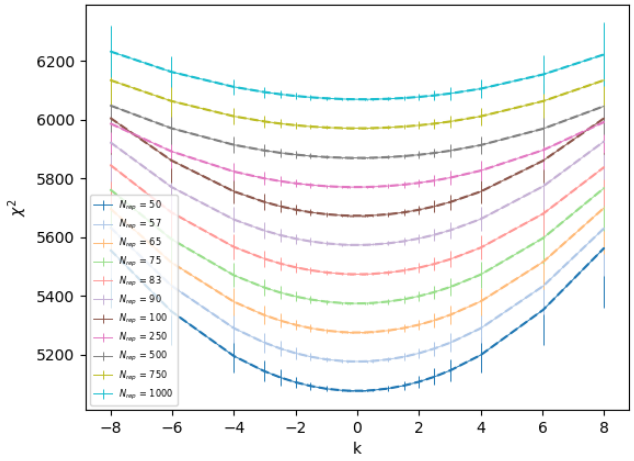

Figure 3.10: *Plots of the χ^2 (solid line) and the fit results (dashed line) as a function of $1/k$ for each value of N_{rep}. Each plot is lifted by 100 in order to improve the readability.*

0.125}. As regards the number of Monte Carlo replicas, we choose to vary N_{rep} in the range {50, 57, 65, 75, 83, 90, 100, 250, 500, 750, 1000}. Since the number of batches N_{batch} is inversely proportional to N_{rep} and the χ^2 sample contains $N_{eig}N_{batch}$ elements, when N_{rep} increases the size of the χ^2 sample becomes small and thus it may be subjected to statistical fluctuations. So far we have considered the χ^2 variation along only one direction of the Hessian eigenvectors and now we can repeat the $\Delta\chi^2$ calculation for the shift in the opposite direction.

Finally the coefficients a_N, b_N, c_N and d_N are found for each value of N_{rep} by fitting the values of χ^2 with a polynomial of degree four as shown in Fig. 3.10. We may then remove the dependence from N_{rep} calculated in Eq. (3.10) and thus we obtain the N-independent coefficients a_0, b_0, c_0 and d_0 for each value of N_{rep}. The final results are then given by the average of the N-independent coefficients over the values obtained varying N_{rep} and thus we find $a_0 = 0.12565 \pm 0.00009$, $b_0 = 5.90682 \pm 0.00007$, $c_0 = 0.000448 \pm 0.000008$ and $d_0 = 0.000412 \pm 0.0000002$. The uncertainties of these coefficients are calculated from the standard deviations provided by the fits.

Once a_0, b_0, c_0 and d_0 are provided, we may check the predicting power of the model by comparing the N-dependent coefficients a_N, b_N, c_N and d_N with the fit results. From the comparisons in Fig. 3.11, we may deduce that a_N and c_N are subjected to large fluctuations that are underestimated by the model predictions. The coefficient d_N presents statistical fluctuations that are compatible with the model prediction. However the fit results deeply deviate from the model prediction

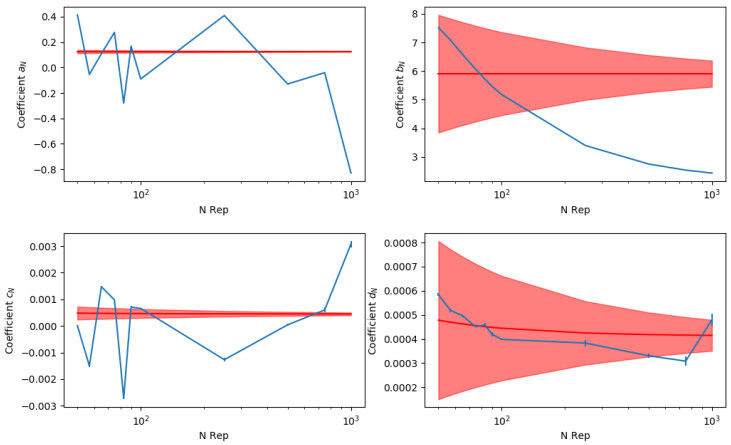

Figure 3.11: *Comparison between the fit results and the model prediction for the N-dependent coefficients a_N, b_N, c_N and d_N.*

for the coefficient b_N which represents the main contribution to $\Delta\chi^2$ fluctuations since it is at least an order of magnitude greater than the other coefficients. The strong N dependence of the coefficient b_N implies that an important contribution of finite-size effects is not considered by the model.

Since the Hessian conversion neglects the eigenvectors of the PDF covariance matrix with small eigenvalues, we may suggest that the loss of information introduced by this approximation can not be neglected and thus represents a further contribution to $\Delta\chi^2$ fluctuations. We therefore test the stability of the fit results by varying the number of elements involved in the χ^2 average.

We first test the stability of the fit results considering the χ^2 average over disjointed subsets of eigenvectors for each Monte Carlo batch. Since the eigenvectors are ordered in terms of the corresponding eigenvalue, we evenly divide the eigenvectors in four subsets which give the results shown in Fig. 3.12. While a_N and c_N present the same fluctuations of the previous case, the coefficients b_N and d_N are clearly not stable when we modified the number of eigenvectors. This behavior may support the hypothesis that the missing eigenvector contribution can not be neglected. This hypothesis can be tested in the following way: in the previous case we divide the total set of eigenvectors in four parts and thus we obtain four disjointed subsets. For $N_{eig} = 50$ the first subset thus contains the first 12 eigenvectors, the second subset contains the second 12 eigenvectors etc. We can now produce four cumulative subsets by taking the first subset and adding each time the next subset. In the case $N_{eig} = 50$, the first cumulative subset contains the first 12 eigenvectors, the second contains the first 24 eigenvectors, the third contains the first 36 eigenvectors etc. Assuming the hypothesis of the missing eigenvectors

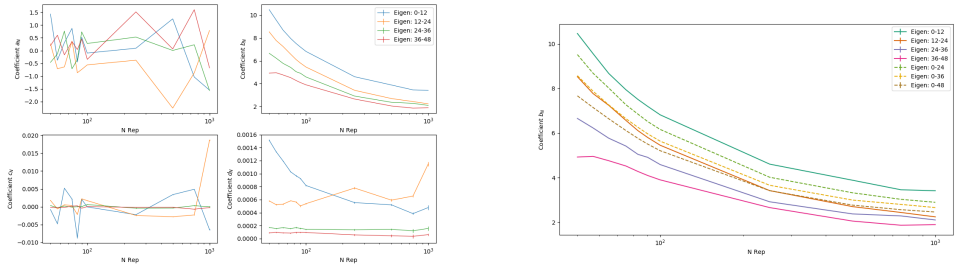

Figure 3.12: *The plot on the left shows the coefficients related to the χ^2 average over four disjointed subset of eigenvectors. The plot on the right shows in addition to previous results for the coefficient b_N the fit results with the cumulative subsets of eigenvectors (dashed lines).*

contribution, the cumulative subsets provide a better estimation of the χ^2 than the single subset which they are made of and therefore we expect that the fit results for b_N produced with the cumulative subsets lay under the corresponding results of the disjointed subsets. However, the right plot of Fig. 3.12 shows the opposite behaviour and therefore the dependence on N_{eig} is not related to the neglected eigenvectors in the Hessian conversion.

We then produce the χ^2 for fixed sigma fraction k and number of Monte Carlo replicas N_{rep} by averaging over each Hessian eigenvector for a smaller number of Monte Carlo batches. The results for the fits with four different choices of batches are shown in Fig. 3.13: the fluctuations of the coefficients a_N and c_N are larger by a factor two and thus are compatible with the statistical fluctuation effects which typically scales with the squared root of the sample size. Indeed the statistical fluctuations are expected to increase by a factor 2 since we reduce the χ^2 sample by a factor 4. As regards the coefficients b_N and d_N, the stability test confirms the decreasing N_{rep} dependence although d_N is subject to statistical fluctuations too.

Since there is a clear dependence of the fit coefficients from the number of eigenvectors, we deduce that the assumption according to which all the parameters of the Hessian conversion follow the same distribution must take into account further finite-size contributions that are related to the multidimensional problem of the diagonalization of the PDF covariance matrix. For example, this model does not consider the variation of the eigenvector directions due to statistical fluctuations. Since the one-sigma contour in the space of parameters is the hypesphere defined by Eq. (2.20) which is subjected to statistical fluctuations, the finite-size effects involve a N_{eig} dimensional region in the space of parameters and therefore we expect that the sample size must scale with the volume in this N_{eig} dimensional space in order to provide a good coverage of the one-sigma contour.

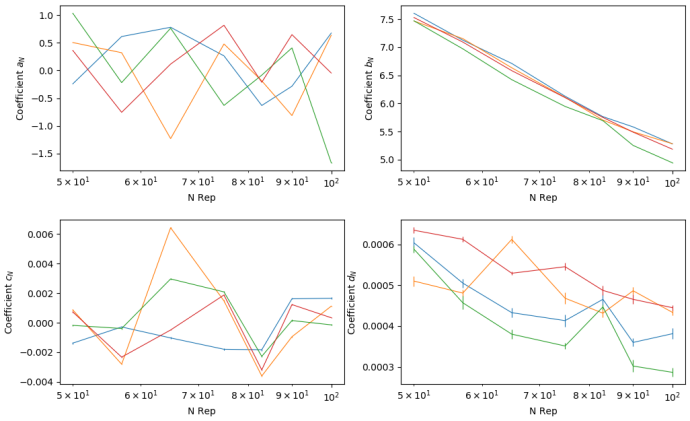

Figure 3.13: *Fit results for the N-dependent coefficients a_N, b_N, c_N and d_N obtained with four different choices of batches in the χ^2 calculation.*

3.5 Model independent approach

The quartic approximation of the $\Delta\chi^2$ provide a really good description of the sigma fraction k dependence since the quartic parameter d_N is small and on turn the fits of the χ^2 shape provide an excellent agreement with the numerical results as shown in Fig. 3.10. We may thus conclude that the functional form of the $\Delta\chi^2$ is well described by Eq. (3.9). Furthermore the a_N, b_N, c_N and d_N coefficients can be expressed by factorizing the N dependence in the following way:

$$z_N = z_0 f_z(N), \tag{3.11}$$

where $z = \{a, b, c, d\}$, $f_z(N)$ is a function that tends to 1 when $N \to +\infty$ since the finite-size effects are supposed to vanish in the limit of infinite size and z_0 represents the asymptotic coefficient.

We can now adopt a model independent approach based on the analysis of $\Delta\chi^2$ data in order to evaluate the coefficients and their asymptotic values. The N dependence of the coefficients can be found in the same way of the previous analysis: we calculate the χ^2 for fixed k and N_{rep} by averaging over all the eigenvectors for each batch. We than fit the k dependence for each value of N_{rep} and we obtain the values of a_N, b_N, c_N and d_N for different values of N_{rep}. Since we know that the χ^2 depends on the number of Hessian eigenvectors involved in the average, we may calculate the uncertainties due to these effects by computing the standard deviation of the coefficient sample obtained by the average of the χ^2 over the set of batches only. Fig. 3.14 shows the dependence of the coefficients from N_{rep} with the one-sigma band for two values of N_{eig}, namely $N_{rep} = 50$ and $N_{rep} = 30$. We may draw the following conclusions:

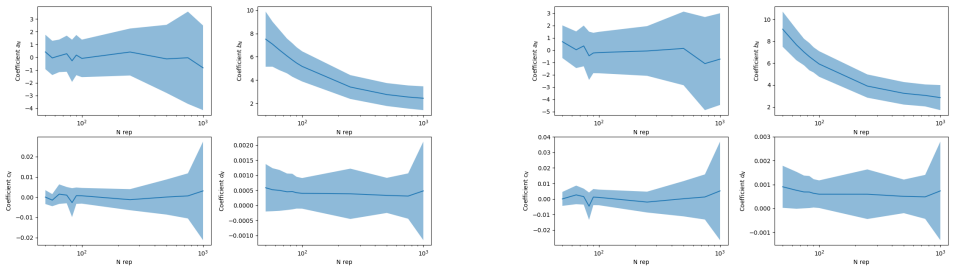

Figure 3.14: *Coefficient results with error band obtained from the standard deviation over the eigenvector sample. The plots on the left is produced with* $N_{eig} = 50$ *while the plots on the right with* $N_{eig} = 30$.

- The coefficients a_N and c_N do not show a particular dependence from N_{rep} and therefore we can assume $f_a(N) = f_c(N) = 1$. Furthermore the asymptotic values are both compatible with zero and are subjected to large uncertainties that we may assess as follows:

$$\begin{aligned} a_0 &= 0 \pm 2, \\ c_0 &= 0.0 \pm 0.1. \end{aligned} \tag{3.12}$$

- The coefficient b_N represents the dominating contribution and shows a clear N_{rep} dependence while the uncertainties are about 30% of the central value. We may thus introduce a functional form in order to estimate $f_b(N)$ and b_0.

- The coefficient d_N is really small but it is not compatible to zero within the one-sigma band. Moreover it has an extremely weak dependence on N_{rep}. Due to these facts, we can reasonably assume that $f_d(N) = 1$ and thus we find $d_0 = 0.007 \pm 0.008$

We may therefore estimate the N_{rep} dependence of the coefficient b_N by introducing a functional form that must provide a good description of both the results obtained with $N_{eig} = 50$ and $N_{eig} = 30$. We thus find that the b_N values are well described by the following function:

$$b_N = \frac{n_0 + n_1\sqrt{N} + n_2 N}{d_0 + d_1\sqrt{N} + d_2 N} \tag{3.13}$$

where the coefficients n_0, n_1, n_2, d_0, d_1 and d_2 are determined by an appropriate fit for both $N_{eig} = 50$ and $N_{eig} = 30$.

Fig. 3.15 shows a good agreement between the fit results and the values of b_N. We may then express b_N as in Eq. (3.11) by rewriting its functional form in the

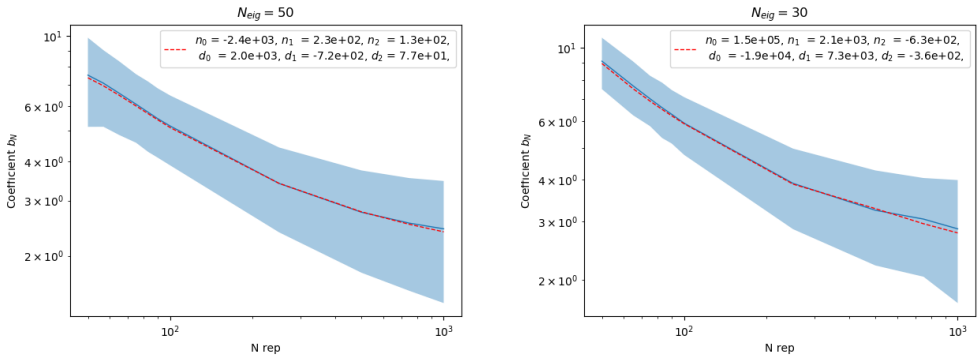

Figure 3.15: *Coefficient results with error band obtained from the standard deviation over the eigenvectors sample. The plots on the left is produced with* $N_{eig} = 50$ *while the plots on the right with* $N_{eig} = 30$ *.*

following way:

$$b_N = b_0 \frac{1 + \frac{n'_1}{\sqrt{N}} + \frac{n'_2}{N}}{1 + \frac{d'_1}{\sqrt{N}} + \frac{d'_2}{N}} = b_0 f_b(N) \tag{3.14}$$

We can therefore factorize the dependence of N_{rep} contained in $f_b(N)$ from b_N and its uncertainty and we finally obtain the following values for b_0:

$$\begin{aligned} b_0 &= 1.68 \pm 0.72 \quad \text{for } N_{eig} = 50 \\ b_0 &= 1.74 \pm 0.63 \quad \text{for } N_{eig} = 30 \end{aligned} \tag{3.15}$$

Since these two values are compatible within their uncertainties, we may thus conclude that the asymptotic values is $b_0 = 1.7 \pm 0.7$ and it is independent from the number of eigenvectors.

Since the other coefficients provide a poor contribution to $\Delta\chi^2$, the coefficient b_0 thus coincides with the variation $\Delta\chi^2$ of the one-sigma contour in the limit of infinite size of the sample. Since the tolerance is defined as the squared roots of this $\Delta\chi^2$, we conclude that the value of b_0 is completely analogous to the introduction of a tolerance $t = \sqrt{b_0} = 1.3 \pm 0.3$. Such a value of the tolerance is compatible with 1 and the deviation from 1 of the central value of t can reasonably be related to both experimental uncertainty underestimation and theoretical uncertainty contribution. Therefore the introduction of the tolerance for the Monte Carlo PDF sets provided by the NNPDF collaboration introduces only a small deviation from the parameter-fitting criterion $\Delta\chi^2 = 1$ which is not compatible with the large values of the tolerance ($t \simeq 5$) required by the existing Hessian PDF sets.

Chapter 4

Conclusions

In this thesis we discuss the problem of the tolerance for the Monte Carlo PDF sets provided by the NNPDF collaboration in terms of an appropriate Hessian conversion. For this purpose we provide a covariance matrix in the space of Monte Carlo PDF replicas and then we express the Hessian eigenvectors as a linear combination of Monte Carlo replicas. We then compute the $\Delta\chi^2$ and we find that $\Delta\chi^2$ values range in $-7 \lesssim \Delta\chi^2 \lesssim 15$ and the t_0 prescription must be adopted for the definition of χ^2. We interpret this behavior in terms of large fluctuations of the one-sigma contour which can depend on the following effects: finite-size effects, non gaussianity, inefficiency of the minimization algorithm and parabolic deviation of the χ^2 shape.

We approach the finite-size effects by increasing the number of Monte Carlo replicas fitted from deep inelastic scattering dataset. However this Monte Carlo set suffers from large non-gaussian deviations that invalidate the Hessian conversion. We thus adopt a global Monte Carlo set with 1000 replicas that presents more gaussian uncertainties and we improve the Hessian conversion rejecting non gaussian contributions. We also provide an Hessian conversion with fractional covariance matrix that describes different size of fluctuations in order to probe the underlying shape of the $\Delta\chi^2$.

All the aforementioned effects are then included in the quartic expansion of the χ^2 with coefficients that depend on the replica sample size and can be calculated by fitting the Hessian conversion results produced with different fractions of the covariance matrix from several Monte Carlo sets with different sizes.

We compute the coefficients of the χ^2 expansion with the assumption that the $\Delta\chi^2$ for each eigenvector follows the same underlying distribution given by a single parameter model. However we find that the single parameter model predictions for the χ^2 coefficients do not consider further finite-size effects related to the multidimensional problem of one-sigma contour fluctuations in the N_{eig} dimensional space of parameters. We thus conclude that the replica sample size must scale

with the volume of the hypersphere in the space of parameters in order to provide a good coverage of $\Delta\chi^2$ fluctuations.

We therefore compute the finite-size dependence of the coefficients from the numerical data for the χ^2. We find that the linear, cubic and quartic coefficients are related to inefficiency, non-gaussianity and parabolic deviation. In particular these coefficients are subjected to large statistical fluctuations but provide a small contribution to $\Delta\chi^2$. Moreover, the dominant contribution to $\Delta\chi^2$ is driven by the quadratic coefficient that shows a strong dependence from the replica size. Once the finite-size effects are factorized, we find that the quadratic coefficient is analogous to the introduction of a tolerance $t = 1.3 \pm 0.3$. We then conclude that such a value of the tolerance is compatible with 1 and the deviation from 1 of the central value of t can reasonably be related to both experimental uncertainty underestimation and theoretical uncertainty contribution.

Since the NNPDF collaboration provides Monte Carlo sets of PDFs with 100 replicas optimized from the prior set of 1000 replicas, the large finite-size effects discussed in this work may suggest a more careful assessment of the size of the Monte Carlo replica sample.

The methodologies discussed in this work may be improved by providing a more accurate assessment of the non gaussian deviation in the Hessian conversion. Furthermore the explicit calculation of the finite-size effects for the N_{eig}-dimensional case could explain the strong dependence from the Monte Carlo replica sample size of the parameter b_N and provide a more accurate assessment of the effects due to inefficiency, non-gaussianity and parabolic deviation. Moreover, the calculation of the $\Delta\chi^2$ results of the Hessian conversion produced from a larger Monte Carlo replica sample represents the most direct way to remove the large finite-size effects as we supposed when the DIS only Monte Carlo sets were provided.

Bibliography

[1] D. J. Gross and F. Wilczek, "Ultraviolet Behavior of Nonabelian Gauge Theories," Phys. Rev. Lett. **30** (1973) 1343. doi:10.1103/PhysRevLett.30.1343

[2] H. D. Politzer, "Reliable Perturbative Results for Strong Interactions?," Phys. Rev. Lett. **30** (1973) 1346. doi:10.1103/PhysRevLett.30.1346

[3] R. K. Ellis, W. J. Stirling and B. R. Webber, "QCD and collider physics," Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **8** (1996) 1.

[4] W. Schroers, "Parton distributions from the lattice," Nucl. Phys. A **755** (2005) 333 doi:10.1016/j.nuclphysa.2005.03.034 [hep-ph/0501156].

[5] R. D. Ball *et al.* [NNPDF Collaboration], "Parton distributions from high-precision collider data," Eur. Phys. J. C **77** (2017) no.10, 663 doi:10.1140/epjc/s10052-017-5199-5 [arXiv:1706.00428 [hep-ph]].

[6] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, "Better jet clustering algorithms," JHEP **9708** (1997) 001 doi:10.1088/1126-6708/1997/08/001 [hep-ph/9707323].

[7] M. Wobisch and T. Wengler, "Hadronization corrections to jet cross-sections in deep inelastic scattering," In *Hamburg 1998/1999, Monte Carlo generators for HERA physics* 270-279 [hep-ph/9907280].

[8] G. Soyez, "The SISCone and anti-k(t) jet algorithms," doi:10.3360/dis.2008.178 arXiv:0807.0021 [hep-ph].

[9] G. D'Agostini, "On the use of the covariance matrix to fit correlated data," Nucl. Instrum. Meth. A **346** (1994) 306. doi:10.1016/0168-9002(94)90719-6

[10] R. D. Ball *et al.* [NNPDF Collaboration], "Fitting Parton Distribution Data with Multiplicative Normalization Uncertainties," JHEP **1005** (2010) 075 doi:10.1007/JHEP05(2010)075 [arXiv:0912.2276 [hep-ph]].

[11] R. A. Khalek *et al.* [NNPDF Collaboration], "Nuclear Parton Distributions from Neural Networks," arXiv:1811.05858 [hep-ph].

[12] J. Rojo, "Machine Learning tools for global PDF fits," arXiv:1809.04392 [hep-ph].

[13] J. C. Collins and J. Pumplin, "Tests of goodness of fit to multiple data sets," hep-ph/0105207.

[14] G. Watt and R. S. Thorne, "Study of Monte Carlo approach to experimental uncertainty propagation with MSTW 2008 PDFs," JHEP **1208** (2012) 052 doi:10.1007/JHEP08(2012)052 [arXiv:1205.4024 [hep-ph]].

[15] S. Carrazza, S. Forte, Z. Kassabov, J. I. Latorre and J. Rojo, "An Unbiased Hessian Representation for Monte Carlo PDFs," Eur. Phys. J. C **75** (2015) no.8, 369 doi:10.1140/epjc/s10052-015-3590-7 [arXiv:1505.06736 [hep-ph]].

Acknowledgements

Ringrazio il Prof. Stefano Forte e il Prof. Stefano Carrazza per la disponibilità, la pazienza, le spiegazioni e le discussioni. Lavorare sotto la loro guida è stato estremamente stimolante.

Ringrazio la mia famiglia, i miei genitori, Martina, Alessandro per il sostegno durante questi anni.

Un ringraziamento speciale va ad Adriano per essere stato il mio principale complice in questi anni universitari. Ringrazio inoltre Guglielmo, Ilaria, Seba, Matteo, Laura, Silvia, Stefano per i fantastici momenti passati insieme.

Grazie infine ad Arianna per esserci sempre.