Deep-Inelastic Scattering and Collinear Physics

Second European School on the Physics of the EIC and Related Topics

Emanuele R. Nocera

Università degli Studi di Torino and INFN, Torino

23-24 June 2025

Table of contents

DISCLAIMER

These lectures contain a personal selection of topics and are certainly not exhaustive

Bibliography

1. Textbooks on perturbative QCD and DIS
    - J. Campbell, J. Huston, F. Krauss, *The Black Book of QCD*, Oxford (2018)
    - J.C. Collins, *Foundations of perturbative QCD*, Cambridge (2011)
    - R.K. Ellis, W.J. Stirling, B.R. Webber, *QCD and Collider Physics*, Cambridge (1996)
    - R. Devenish, A. Cooper-Sarkar, *Deep-Inelastic Scattering*, Oxford (2011)
    - E. Leader, *Spin in Particle Physics*, Cambridge (2001)
2. Reviews on Parton Distribution Functions
    - K. Kovarik and P.M. Nadolsky, Rev.Mod.Phys. **92** (2020) 045003
    - J.J. Ethier and E.R. Nocera, Ann.Rev.Nucl.Part.Sci. **70** (2020) 43
    - J. Gao, L. Harland-Lang and J. Rojo, Phys.Rept. **742** (2018) 1
    - S. Forte and G. Watt, Ann.Rev.Nucl.Part.Sci. **63** (2013) 291
    - P. Jimenez-Delgado, W. Melnitchouk and J. F. Owens, J. Phys. G40 (2013) 093102
    - E.C. Aschenauer, R.S. Thorne, R. Yoshida (rev.), *Structure Functions*, PDG, ch. 8
3. Specific topics not addressed above
    - more journal references as we proceed through these lectures

DISCLAIMER

These lectures will focus on collinear leading-twist Parton Distribution Functions

Transverse-momentum-dependent distributions will not be covered here

Deep-Inelastic Scattering and Collinear Physics

Lecture 1: What is Deep-Inelastic Scattering and what are Parton Distributions?

Outline

1.1 DIS as a laboratory of QCD
the unpolarised and polarised DIS cross section
factorisation, evolution
properties of splitting functions, theoretical constraints

1.2 Why Parton Distribution Functions?
the LHC and the quest of precision
understanding the origin of the proton spin

1.3 How can we determine PDFs?
how to formulate the problem and how to solve it

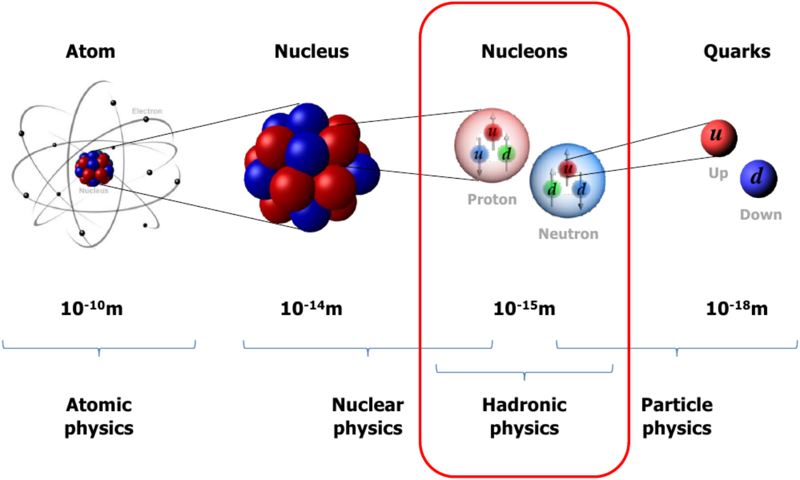
1.4 Which data constrain which PDFs and how?
overview of experimental data (unpolarised): from HERA to the LHC
overview of experimental data (polarised): from EMC to the EIC
which constraints different scattering processes put on PDFs?

I will focus on the phenomenological determination of PDFs
I will not talk about Lattice QCD nor of models of nucleon structure

1.1 DIS as a laboratory of QCD

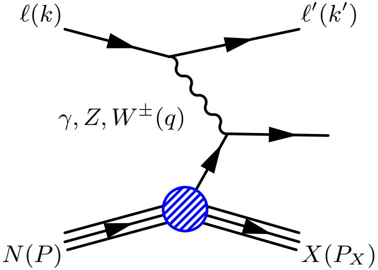
Hadronic physics, or the quest for the nucleon structure

Nucleons make up all nuclei, and hence most of the visible matter in the Universe
They are bound states with internal structure and dynamics



Deep-Inelastic Scattering

$$\ell(k)\,N(P) \to \ell'(k')\,X(P_X)$$



$k(k')$: lepton momentum
M: proton mass
q: gauge boson momentum

P: proton momentum
W: invariant mass of the final state
p parton momentum

DEEP ($Q^2 \gg M^2$) INELASTIC ($W^2 \gg M^2$)

scale: $Q^2 = -q^2$

scaling variable (hadronic): $x = x_B = \frac{Q^2}{2P\cdot q}$

scaling variable (partonic): $z = \frac{Q^2}{2p\cdot q}$

lepton-proton c.m.e. squared: $s = (k+P)^2$

lepton's energy loss: $\nu = \frac{q\cdot P}{M} = E - E'$

inelasticity: $y = \frac{q\cdot P}{k\cdot P} = \frac{\nu}{E}$

Past and present DIS facilities around the world

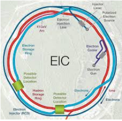
where	who	what		when
SLAC	SLAC	NC DIS (unp.)	$\mu \to p, d$	80s-90s
	NuTeV	CC DIS (unp.)	$\nu(\bar{\nu}) \to$Fe	late 90s
	E142,E143,E154,E155	NC DIS (pol.)	$\mu \to p, d$	80s-90s
CERN	BCDMS	NC DIS (unp.)	$\mu \to p, d$	late 80s
	NMC	NC DIS (unp.)	$\mu \to p, d$	late 90s
	CHORUS	CC DIS (unp.)	$\nu(\bar{\nu}) \to$Pb	late 90s/early 2000s
	EMC, SMC, COMPASS	NC DIS (pol.)	$\mu \to p, d$	late 80s to late 2010s
DESY	H1, ZEUS	NC,CC DIS (unp.)	$e^{\pm}, \mu \to p$	late 90s/early 2010s
JLab	Hall-A	NC DIS (unp., pol.)	$e \to p, d$	late 90s — today
EIC		NC,CC DIS (unp.pol.)	$e^{\pm} \to p, d, A$	2030s(?)


SLAC


CERN


DESY


JLab



DIS cross section: leptonic tensor

The DIS cross sections can be written as

$$\frac{d^2\sigma}{dxdy} = K \sum_{s_{\ell'}} \sum_X \int \Pi_X |\mathcal{M}(\ell N \to \ell' X)|^2 = \frac{2\pi y \alpha^2}{Q^4} \sum_j \eta_j L_j^{\mu\nu} W_{\mu\nu}^j$$

neutral-current (NC) DIS: $j = \gamma, Z$ charged-current (CC) DIS: $j = W^+, W^-$

The corresponding leptonic tensors are (neglecting lepton masses)

$$L^{\gamma}_{\mu\nu} = 2(k_\mu k'_\nu + k'_\mu k_\nu - (k \cdot k')g_{\mu\nu} - i\lambda\epsilon_{\mu\nu\alpha\beta}k^\alpha k'^\beta)$$
$$L^{Z}_{\mu\nu} = (g_V^e + e\lambda g_A^e)^2 L_{\mu\nu\gamma}$$
$$L^{\gamma Z}_{\mu\nu} = (g_V^e + e\lambda g_A^e) L^{\gamma}_{\mu\nu}$$
$$L^{W}_{\mu\nu} = (1 + e\lambda)^2 L^{\gamma}_{\mu\nu}$$

$e = \pm 1$: charge of the incoming lepton $\lambda = \pm 1$ helicity of the incoming lepton

$$g_V^e = -1/2 + 2\sin^2\theta_W \qquad g_A^e = -1/2$$

The ratios of the corresponding propagators and coupling to the photon ones are

$$\eta_\gamma = 1 \quad \eta_Z = \eta_{\gamma Z}^2 \quad \eta_{\gamma Z} = \frac{G_F M_Z^2}{2\sqrt{2}\pi\alpha} \frac{Q^2}{Q^2 + M_Z^2} \quad \eta_W = \frac{1}{2}\left(\frac{G_F M_W^2}{4\pi\alpha} \frac{Q^2}{Q^2 + M_W^2}\right)^2$$

DIS cross section: hadronic tensor

The hadronic tensor describes the interaction of the appropriate currents

$$W_{\mu\nu} = \frac{1}{4\pi}\int d^4 z e^{iq\cdot z}\left\langle P,S|[J^\dagger_\mu(z),J_\nu(0)]|P,S\right\rangle$$

J_α: hadronic contribution to the current $\qquad$ S: nucleon-spin 4-vector

The hadronic tensor is written on the basis of independent four-momenta (combinations) with coefficients F_k (w/o spin) and g_k (w/ spin)

$$\begin{aligned}
W_{\mu\nu} = & \left(-g_{\mu\nu}+\frac{q_\mu q_\nu}{q^2}\right)F_1(x,Q^2)+\frac{\hat{P}_\mu\hat{P}_\nu}{P\cdot q}F_2(x,Q^2)-i\epsilon_{\mu\nu\alpha\beta}\frac{q^\alpha P^\beta}{2P\cdot q}F_3(x,Q^2)\\
&+i\epsilon_{\mu\nu\alpha\beta}\frac{q^\alpha}{P\cdot q}\left[S^\beta g_1(x,Q^2)+\left(S^\beta-\frac{S\cdot q}{P\cdot q}P^\beta\right)g_2(x,Q^2)\right]\\
&+\frac{1}{P\cdot q}\left[\frac{1}{2}\left(\hat{P}_\mu\hat{S}_\nu+\hat{S}_\mu\hat{P}_\nu\right)-\frac{S\cdot q}{P\cdot q}\hat{P}_\mu\hat{P}_\nu\right]g_3(x,Q^2)\\
&+\frac{S\cdot q}{P\cdot q}\left[\frac{\hat{P}_\mu\hat{P}_\nu}{P\cdot q}g_4(x,Q^2)+\left(-g_{\mu\nu}+\frac{q_\mu q_\nu}{q^2}\right)g_5(x,Q^2)\right]
\end{aligned}$$

with $\hat{P}_\mu = P_\mu - \frac{P\cdot q}{q^2}q_\mu \qquad \hat{S}_\mu = S_\mu - \frac{S\cdot q}{q^2}q_\mu$

DIS cross section: unpolarised case

$$\frac{d^2\sigma^i}{dxdy} = \frac{4\pi\alpha^2}{xyQ^2}\eta_i\left[\left(1-y-\frac{x^2y^2M^2}{Q^2}\right)F_2^i + y^2xF_1^i \mp (y-y^2/2)xF_3^i\right] \qquad i = \mathrm{NC}, \mathrm{CC}$$

<u>NC DIS ($e^\pm N \to e^\pm X$)</u>

sign of F_3: $-$ $(+)$ for e^+ (e^-) incoming

$$\eta^{\mathrm{NC}} = 1$$

$$\begin{aligned} F_{1,2}^{\mathrm{NC}} &= F_{1,2}^{\gamma} \\ &\quad - (g_V^e \pm \lambda g_A^e)\eta_{\gamma Z} F_{1,2}^{\gamma Z} \\ &\quad + \left((g_V^e)^2 + (g_A^e)^2 \pm 2\lambda g_V^e g_A^e\right)\eta_Z F_{1,2}^{Z} \\ xF_3^{\mathrm{NC}} &= -(g_A^e \pm \lambda g_V^e)\eta_{\gamma Z} xF_3^{\gamma Z} \\ &\quad + \left[2g_V^e g_A^e \pm \lambda\left((g_V^e)^2 + (g_A^e)^2\right)\right]\eta_Z xF_3^{Z} \end{aligned}$$

<u>CC DIS ($e^\pm(\nu,\bar\nu)N \to \nu,\bar\nu(e^\pm)$)</u>

sign of F_3: $-$ $(+)$ for $\bar\nu$ (ν) incoming

$\eta^{\mathrm{CC}} = (1\pm\lambda)^2\eta_W$ (for $e^\pm$)

$\eta^{\mathrm{CC}} = 4\eta_W$ (for $\nu,\bar\nu$)

$$F_1^{\mathrm{CC}} = F_1^W$$

$$F_2^{\mathrm{CC}} = F_2^W$$

$$xF_3^{\mathrm{CC}} = xF_3^W$$

for incoming $\nu,\bar\nu$, charge and helicity refer to the outgoing $e^\pm$

Neglecting terms proportional to M^2/Q^2

$$\frac{d^2\sigma^i}{dxdy} = \frac{2\pi\alpha^2}{xyQ^2}\eta_i\left[Y_+F_2^i \mp Y_- xF_3^i - y^2F_L^i\right] \qquad Y_\pm = 1\pm(1-y)^2 \qquad F_L^i = F_2^i - 2xF_1^i$$

DIS cross section: longitudinally polarised case

$$\Delta\sigma = \sigma(\lambda_N = -1, \lambda_\ell) - \sigma(\lambda_N = 1, \lambda_\ell)$$

$$\frac{d^2\Delta\sigma^i}{dxdy} = \frac{8\pi\alpha^2}{xyQ^2}\eta^i \left\{ -\lambda_\ell y \left(2 - y - 2x^2y^2\frac{M^2}{Q^2}\right) x g_1^i \right.$$
$$+ \lambda_\ell 4x^3y^2\frac{M^2}{Q^2} g_2^i + 2x^2y\frac{M^2}{Q^2}\left(1 - y - x^2y^2\frac{M^2}{Q^2}\right) g_3^i$$
$$\left. - \left(1 + 2x^2y\frac{M^2}{Q^2}\right)\left[\left(1 - y - x^2y^2\frac{M^2}{Q^2}\right) g_4^i + xy^2 g_5^i\right]\right\} \qquad i = \mathrm{NC}, \mathrm{CC}$$

Neglecting terms proportional to M^2/Q^2

$$\frac{d^2\Delta\sigma^i}{dxdy} = \frac{4\pi\alpha^2}{xyQ^2}\eta_i\left[-Y_+ g_4^i \mp Y_- 2x g_1^i + y^2 g_L^i\right] \qquad Y_\pm = 1 \pm (1-y)^2 \qquad g_L^i = g_4^i - 2x g_5^i$$

which can be obtained from the unpolarised cross section

$$\frac{d^2\sigma^i}{dxdy} = \frac{2\pi\alpha^2}{xyQ^2}\eta_i\left[Y_+ F_2^i \mp Y_- x F_3^i - y^2 F_L^i\right] \qquad Y_\pm = 1 \pm (1-y)^2 \qquad F_L^i = F_2^i - 2x F_1^i$$

from $\quad F_1 \to -g_5 \qquad F_2 \to -g_4 \quad F_3 \to 2g_1$

Structure functions in the naive parton model

Regard the DIS cross section as the incoherent sum of point-like interactions between the lepton and a free, massless, parton

$$\frac{d^2\sigma}{dxdy} = K \sum e_q^2 f_{q/p}(x) \frac{d\hat{\sigma}}{dy} \qquad \frac{d^2\Delta\sigma}{dxdy} = \Delta K \sum e_q^2 \Delta f_{q/p}(x) \frac{d\hat{\sigma}}{dy}$$

with the unpolarised (polarised) parton distribution functions $f_{q/p}$ ($\Delta f_{q/p}$)

$$f_{q/p}(x) = f^{\uparrow}_{q/p}(x) + f^{\downarrow}_{q/p}(x) \qquad \Delta f_{q/p}(x) = f^{\uparrow}_{q/p}(x) - f^{\downarrow}_{/p}(x)$$

In the Bjorken limit $Q^2, \nu \to \infty$ structure functions scale, and moreover $F_L^i = g_L^i = 0$

$$\left[F_2^{\gamma}, F_2^{\gamma Z}, F_2^{Z}\right] = x \sum_q \left[e_q^2, 2e_q g_V^q, (g_V^q)^2 + (g_A^q)^2\right] (q + \bar{q})$$

$$\left[F_3^{\gamma}, F_3^{\gamma Z}, F_3^{Z}\right] = \sum_q \left[0, 2e_q g_A^q, 2 g_V^q g_A^q\right] (q - \bar{q})$$

$$\left[F_2^{W^-}, F_3^{W^-}\right] = \left[2x(u + \bar{d} + \bar{s} + c \ldots), 2(u - \bar{d} - \bar{s} + c \ldots)\right]$$

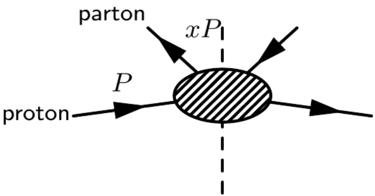
$$\left[g_1^{\gamma}, g_1^{\gamma Z}, g_1^{Z}\right] = \frac{1}{2} \sum_q \left[e_q^2, 2e_q g_V^q, (g_V^q)^2 + (g_A^q)^2\right] (\Delta q + \Delta\bar{q})$$

$$\left[g_5^{\gamma}, g_5^{\gamma Z}, g_5^{Z}\right] = \sum_q \left[0, e_q g_A^q, g_V^q g_A^q\right] (\Delta q - \Delta\bar{q})$$

$$\left[g_1^{W^-}, g_5^{W^-}\right] = \left[\Delta u + \Delta\bar{d} + \Delta\bar{s} + \Delta c \ldots, -\Delta u + \Delta\bar{d} + \Delta\bar{s} - \Delta c \ldots\right]$$

Field-theoretic definition of PDFs

PDFs allow for a field-theoretic definition as matrix elements of bilocal operators



collinear transition
of a massles proton h
into a massless parton i
with fractional momentum x
local OPE $\Longrightarrow$ lattice formulation
[See *e.g.* Prog.Part.Nucl.Phys. 121 (2021) 103908]

$$
\begin{aligned}
q(x) &= \frac{1}{4\pi}\int dy^- e^{-iy^- xP^+}\langle P,S|\bar{\psi}(0,y^-,\mathbf{0}_\perp)\gamma^+\psi(0)|P,S\rangle \\
\Delta q(x) &= \frac{1}{4\pi}\int dy^- e^{-iy^- xP^+}\langle P,S|\bar{\psi}(0,y^-,\mathbf{0}_\perp)\gamma^+\gamma^5\psi(0)|P,S\rangle \\
\Delta g(x) &= \frac{1}{4\pi x P^+}\int dy^- e^{-iy^- xP^+}\langle P,S|G^{+\alpha}(0,y^-,\mathbf{0}_\perp)\tilde{G}^+_\alpha(0)|P,S\rangle
\end{aligned}
$$

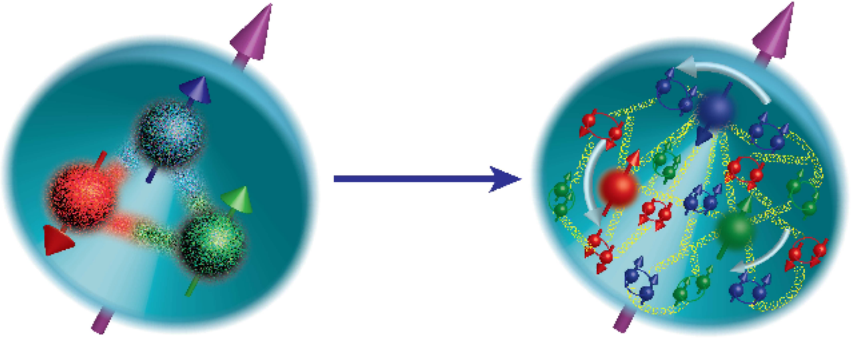
with light-cone coordinates and QCD field-strength tensor G ($A^+ = 0$ gauge)

$$y = (y^+, y^-, \mathbf{y}_\perp)\,, \quad y^+ = (y^0 + y^z)/\sqrt{2}\,, \quad y^- = (y^0 - y^z)/\sqrt{2}\,, \quad \mathbf{y}_\perp = (v^x, v^y)$$

$$G^\alpha_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + f^{abc}A^b_\mu A^c_\nu$$

Structure functions in QCD

naive picture | realistic picture



three non-relativistic quarks $\overset{\text{QCD}}{\underset{\text{factorization,evolution}}{\longleftrightarrow}}$ indefinite number of relativistic quarks and gluons

Increasing Q^2, one should see that each quark is surrounded by a cloud of partons.
The number of resolved partons that share the proton's momentum increases with Q^2.
If quarks were non-interacting, no further structure would be resolved increasing Q^2.

Factorisation of Physical Observables

1. Factorisation theorems apply to sufficiently inclusive scattering processes

short-distance interaction
partonic hard interaction
process-specific kernel

$$\xleftrightarrow[\text{scale }\mu]{\text{factorisation scheme}}$$

long-distance interaction
nucleon structure
universal PDFs

2. Physical observables can be written as convolutions of matrix elements and PDFs

$$F_I(x,\mu^2) = \sum_i \int_x^1 \frac{dz}{z} C_{Ii}(z,\alpha_s(\mu^2)) f_i\left(\frac{x}{z},\mu^2\right)$$ ONE HADRON

$$\sigma(\tau,\mu^2,\mathbf{k}) = \sum_{ij} \int_\tau^1 \frac{dz}{z} \hat{\sigma}_{ij}\left(\frac{\tau}{z},\alpha_s(\mu^2),\mathbf{k}\right) \mathcal{L}_{ij}(z,\mu^2)$$ TWO HADRONS

$$\mathcal{L}_{ij}(z,\mu^2) = (f_i^{h_1} \otimes f_j^{h_2})(z,\mu^2)$$

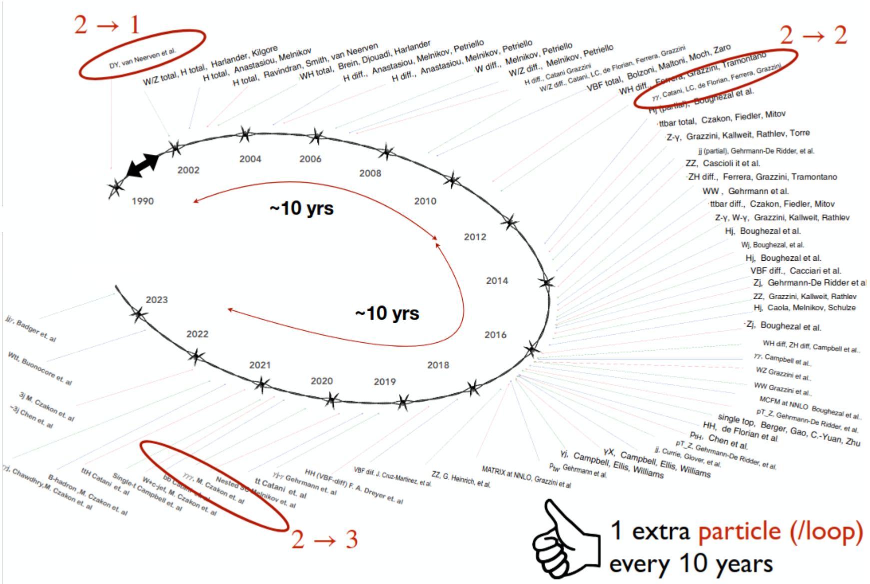
$$f \otimes g = \int_x^1 \frac{dz}{z} f\left(\frac{x}{z}\right) g(z)$$

3. The matrix elements C_{If} and $\hat{\sigma}_{ij}$ can be computed perturbatively

$$C_{Ii}(y,\alpha_s) = \sum_{k=0} a_s^k C_{Ii}^{(k)}(y) \qquad \hat{\sigma}_{ij}(y,\alpha_s) = \sum_{k=0} a_s^k \hat{\sigma}_{ij}^{(k)}(y) \qquad a_s = \alpha_s/(4\pi)$$

4. Because of factorisation, all of these quantities depend on μ^2; usually $Q^2 = \mu^2$

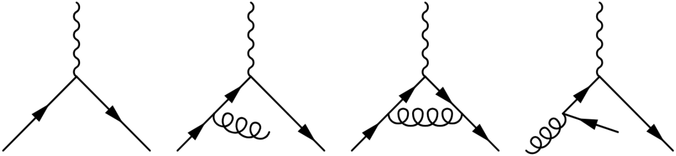
Perturbative corrections



1 extra particle (/loop)
every 10 years

[Figure by courtesy of L. Cieri]

Perturbative corrections



Protons may radiate gluons, that split into a quark-antiquark pair,
that may interact with the vector boson; $F_L \neq 0$, $g_L \neq 0$

Quarks and antiquarks may radiate gluons
that give rise to collinear logarithmic corrections, *e.g.* at leading-log

$$\alpha_s \ln \frac{Q^2}{m^2}$$

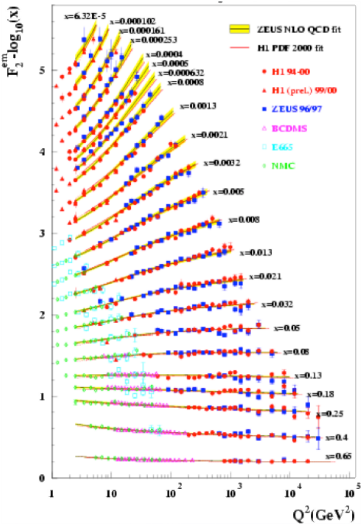
associated to a soft collinear singularity due to the masslessness of quarks.
The factorisation theorem sets a separation between the hard and soft parts of the process; the (arbitrary) scale where this separation occurs is the factorisation scale μ

$$\alpha_s \ln \frac{Q^2}{m^2} = \alpha_s \ln \frac{Q^2}{\mu^2} + \alpha_s \ln \frac{\mu^2}{m^2}$$

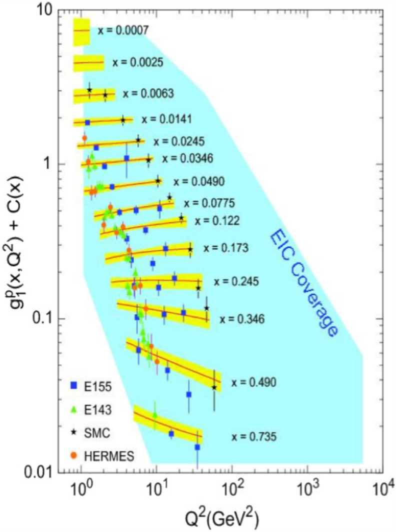
One reasbsorbes the divergent term $\alpha_s \ln \frac{\mu^2}{m^2}$ into the PDFs, which scale logarithmically.

Breaking of the Bjorken scaling

UNPOLARISED



POLARISED



PDF evolution: DGLAP equations

1. A set of $(2n_f + 1)$ integro-differential equations, n_f is the number of active partons

$$\frac{\partial}{\partial \ln \mu^2} f_i(x, \mu^2) = \sum_j^{n_f} \int_x^1 \frac{dz}{z} P_{ji}\left(z, \alpha_s(\mu^2)\right) f_j\left(\frac{x}{z}, \mu^2\right)$$

2. They are often written in a convenient PDF basis

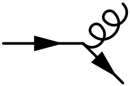
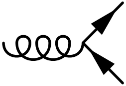
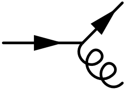
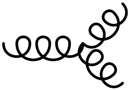
$$f_\pm = (f_q \pm f_{\bar{q}}) - (f_{q'} \pm f_{\bar{q}'}) \qquad f_v = \sum_i^{n_f} (f_q - f_{\bar{q}}) \qquad f_\Sigma = \sum_i^{n_f} (f_q + f_{\bar{q}})$$

$$\frac{\partial}{\partial \ln \mu^2} f_{\pm,v}(x, \mu^2) = P^{\pm,v}(x, \mu_F^2) \otimes f_{\pm,v}(x, \mu^2)$$

$$\frac{\partial}{\partial \ln \mu^2} \begin{pmatrix} f_\Sigma(x, \mu^2) \\ f_g(x, \mu^2) \end{pmatrix} = \begin{pmatrix} P_{qq} & 2n_f P_{qg} \\ P_{gq} & P_{gg} \end{pmatrix} \otimes \begin{pmatrix} f_\Sigma(x, \mu^2) \\ f_g(x, \mu^2) \end{pmatrix}$$

3. The splitting functions P can be computed perturbatively

$$P_{ji}(z, \alpha_s) = \sum_{k=0} a_s^{k+1} P_{ji}^{(k)}(z), \qquad a_s = \alpha_s/(4\pi)$$

$P_{qq}^{(0)}$  $P_{gq}^{(0)}$  $P_{qg}^{(0)}$  $P_{gg}^{(0)}$ 

Splitting Functions and Anomalous Dimensions

Perform the Mellin transform $\gamma_{ji}(N,\alpha_s(\mu^2)) \equiv \int_0^1 dx x^{N-1} P_{ji}(x,\alpha_s(\mu^2))$

$$\frac{\partial}{\partial \ln \mu^2} f_{\pm,v}(N,\mu^2) = \gamma_{\pm,v}(N,\mu_F^2) \cdot f_{\pm,v}(N,\mu^2)$$

$$\frac{\partial}{\partial \ln \mu^2} \begin{pmatrix} f_\Sigma(N,\mu^2) \\ f_g(N,\mu^2) \end{pmatrix} = \begin{pmatrix} \gamma_{qq} & 2n_f\gamma_{qg} \\ \gamma_{gq} & \gamma_{gg} \end{pmatrix} \cdot \begin{pmatrix} \Sigma(N,\mu^2) \\ g(N,\mu^2) \end{pmatrix}$$

How many different anomalous dimensions are there?

LO: $\gamma_{qq} = \gamma_{\pm,v} \rightarrow$ 4 independent splitting functions
NLO: $\gamma_{qq} \neq \gamma_+ \neq \gamma_- \rightarrow$ 6 independent splitting functions
NNLO: $\gamma_- \neq \gamma_v \rightarrow$ 7 independent splitting functions

Which PDF combinations evolve independently?

LO: f_g, f_Σ, and any $2n_f - 1$ linear combinations of f_q and $f_{\bar{q}}$
NLO: f_g, f_Σ, any $n_f - 1$ linear combinations of $f_q - f_{\bar{q}}$, and of $f_q + f_{\bar{q}}$
NNLO: as NLO, and $f_V = \sum_q^{n_f}(f_q - f_{\bar{q}})$

A common choice

$$f_g,\ f_\Sigma = \sum_q^{n_f}(f_q + f_{\bar{q}}),\ f_V = \sum_q^{n_f}(f_q - f_{\bar{q}})$$

iterative NS combinations of $f_{q^+} = f_q + f_{\bar{q}}$ and of $f_{q^-} = f_q - f_{\bar{q}}$

$$T_3 = f_{u^+} - f_{d^+} \qquad T_8 = f_{u^+} + f_{d^+} - 2f_{s^+} \qquad T_{15} = f_{u^+} + f_{d^+} + f_{s^+} - 3f_{c^+} \ \dots$$

$$V_3 = f_{u^-} - f_{d^-} \qquad V_8 = f_{u^-} + f_{d^-} - 2f_{s^-} \qquad V_{15} = f_{u^-} + f_{d^-} + f_{s^-} - 3f_{c^-} \ \dots$$

Anomalous dimensions: perturbative accuracy

LO (1977)

[NPB 126 (1977) 298]

$$
\begin{aligned}
\gamma_{\rm ps}^{(0)}(N) &= 0 \\
\gamma_{\rm qg}^{(0)}(N) &= 2n_f\,(\mathbf{N}_- + 4\mathbf{N}_+ - 2\mathbf{N}_{+2} - 3)\,S_1 \\
\gamma_{\rm gq}^{(0)}(N) &= 2C_F\,(2\mathbf{N}_{-2} - 4\mathbf{N}_- - \mathbf{N}_+ + 3)\,S_1 \\
\gamma_{\rm gg}^{(0)}(N) &= C_A\left(4(\mathbf{N}_{-2} - 2\mathbf{N}_- - 2\mathbf{N}_+ + \mathbf{N}_{+2} + 3)\,S_1 - \frac{11}{3}\right) + \frac{2}{3}n_f
\end{aligned}
$$

NLO (1980)

[NPB 175 (1980) 27; PLB 97 (1980) 437]

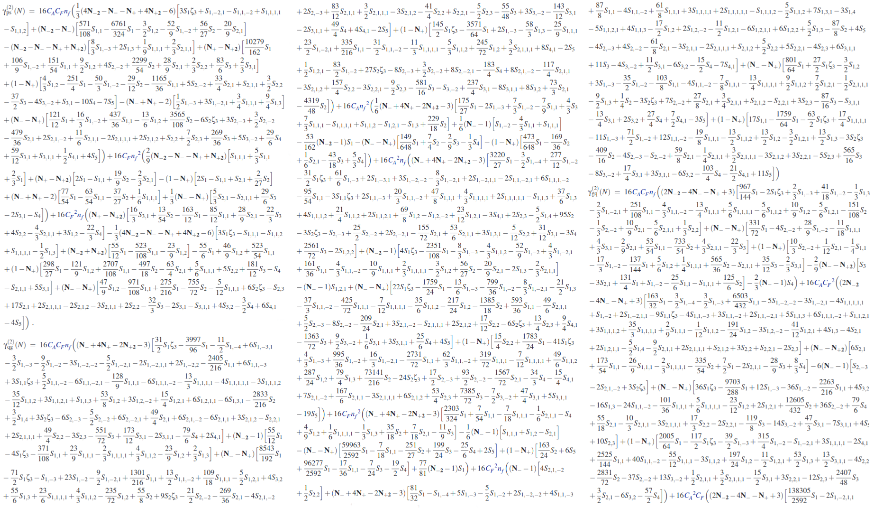
$$
\begin{aligned}
\gamma_{\rm ps}^{(1)}(N) &= 4C_F n_f\Big(\frac{20}{9}(\mathbf{N}_{-2}-\mathbf{N}_-)S_1 - (\mathbf{N}_+-\mathbf{N}_{+2})\Big[\frac{56}{9}S_1+\frac{8}{3}S_2\Big] + (1-\mathbf{N}_+)\Big[8S_1-4S_2\Big] \\
&\quad - (\mathbf{N}_--\mathbf{N}_+)\Big[2S_1+S_2+2S_3\Big]\Big) \\
\gamma_{\rm qg}^{(1)}(N) &= 4C_A n_f\Big(\frac{20}{9}(\mathbf{N}_{-2}-\mathbf{N}_-)S_1 - (\mathbf{N}_--\mathbf{N}_+)\Big[2S_1+S_2+2S_3\Big] - (\mathbf{N}_+-\mathbf{N}_{+2})\Big[\frac{218}{9}S_1 \\
&\quad + 4S_{1,1} + \frac{44}{3}S_2\Big] + (1-\mathbf{N}_+)\Big[27S_1+4S_{1,1}-7S_2-2S_3\Big] - 2(\mathbf{N}_-+4\mathbf{N}_+-2\mathbf{N}_{+2}-3)\Big[S_{1,-2} \\
&\quad + S_{1,1,1}\Big]\Big) + 4C_F n_f\Big(2(\mathbf{N}_+-\mathbf{N}_{+2})\Big[5S_1+2S_{1,1}-2S_2+S_3\Big] - (1-\mathbf{N}_+)\Big[\frac{43}{2}S_1+4S_{1,1}-\frac{7}{2}S_2\Big] \\
&\quad + (\mathbf{N}_--\mathbf{N}_+)\Big[7S_1-\frac{3}{2}S_2\Big] + 2(\mathbf{N}_-+4\mathbf{N}_+-2\mathbf{N}_{+2}-3)\Big[S_{1,1,1}-S_{1,2}-S_{2,1}+\frac{1}{2}S_3\Big]\Big) \qquad (3.7) \\
\gamma_{\rm gq}^{(1)}(N) &= 4C_A C_F\Big(2(2\mathbf{N}_{-2}-4\mathbf{N}_--\mathbf{N}_++3)\Big[S_{1,1,1}-S_{1,-2}-S_{1,2}-S_{2,1}\Big] + (1-\mathbf{N}_+)\Big[2S_1 \\
&\quad - 13S_{1,1}-7S_2-2S_3\Big] + (\mathbf{N}_{-2}-2\mathbf{N}_-+\mathbf{N}_+)\Big[S_1-\frac{22}{3}S_{1,1}\Big] + 4(\mathbf{N}_--\mathbf{N}_+)\Big[\frac{7}{9}S_1+3S_2+S_3\Big] \\
&\quad + (\mathbf{N}_+-\mathbf{N}_{+2})\Big[\frac{44}{9}S_1+\frac{8}{3}S_2\Big]\Big) + 4C_F n_f\Big((\mathbf{N}_{-2}-2\mathbf{N}_-+\mathbf{N}_+)\Big[\frac{4}{3}S_{1,1}-\frac{20}{9}S_1\Big] - (1-\mathbf{N}_+)\Big[4S_1 \\
&\quad - 2S_{1,1}\Big]\Big) + 4C_F^2\Big((2\mathbf{N}_{-2}-4\mathbf{N}_--\mathbf{N}_++3)\Big[3S_{1,1}-2S_{1,1,1}\Big] - (1-\mathbf{N}_+)\Big[S_1-2S_{1,1}+\frac{3}{2}S_2 \\
&\quad - 3S_3\Big] - (\mathbf{N}_--\mathbf{N}_+)\Big[\frac{5}{2}S_1+2S_2+2S_3\Big]\Big) \\
\gamma_{\rm gg}^{(1)}(N) &= 4C_A n_f\Big(\frac{2}{3}-\frac{16}{3}S_1-\frac{23}{9}(\mathbf{N}_{-2}+\mathbf{N}_{+2})S_1+\frac{14}{3}(\mathbf{N}_-+\mathbf{N}_+)S_1+\frac{2}{3}(\mathbf{N}_--\mathbf{N}_+)S_2\Big) \\
&\quad + 4C_A^2\Big(2S_{-3}-\frac{8}{3}-\frac{14}{3}S_1+2S_3-(\mathbf{N}_{-2}-2\mathbf{N}_--2\mathbf{N}_++\mathbf{N}_{+2}+3)\Big[4S_{1,-2}+4S_{1,2}+4S_{2,1}\Big] \\
&\quad + \frac{8}{3}(\mathbf{N}_+-\mathbf{N}_{+2})S_2 - 4(\mathbf{N}_--3\mathbf{N}_++\mathbf{N}_{+2}+1)\Big[3S_2-S_3\Big] + \frac{109}{18}(\mathbf{N}_-+\mathbf{N}_+)S_1 + \frac{61}{3}(\mathbf{N}_- \\
&\quad - \mathbf{N}_+)S_2\Big) + 4C_F n_f\Big(\frac{1}{2}+\frac{2}{3}(\mathbf{N}_{-2}-13\mathbf{N}_--\mathbf{N}_+-5\mathbf{N}_{+2}+18)S_1+(3\mathbf{N}_--5\mathbf{N}_++2)S_2 \\
&\quad - 2(\mathbf{N}_--\mathbf{N}_+)S_3\Big)\,.
\end{aligned}
$$

Numerical solution (LO, NLO , NNLO and aN^{3}LO) of DGLAP implemented in open-source software: EKO [EPJ C82 (2022) 976] and APFEL++ [CPC 185 (2014) 1647]

Anomalous dimensions: perturbative accuracy

NNLO (2004)

[NPB 691 (2004) 129]

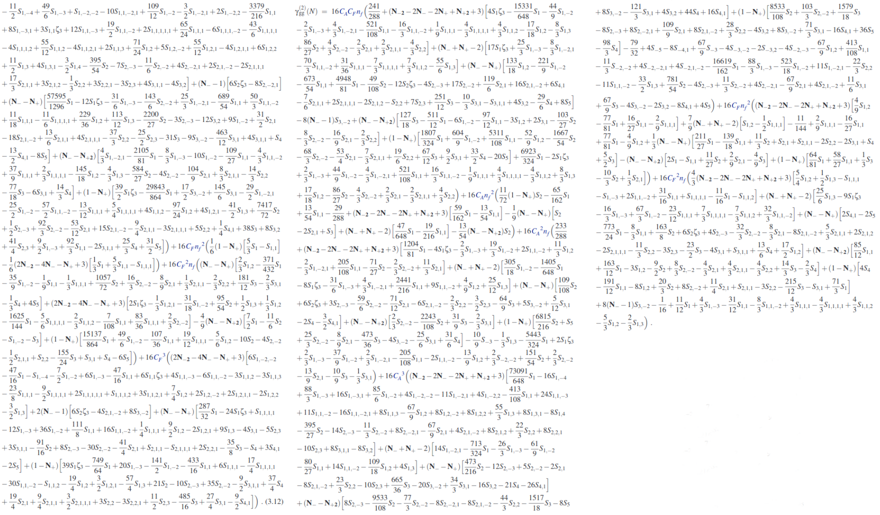


aN^{3}LO (2020 - ongoing)

Anomalous dimensions: perturbative accuracy

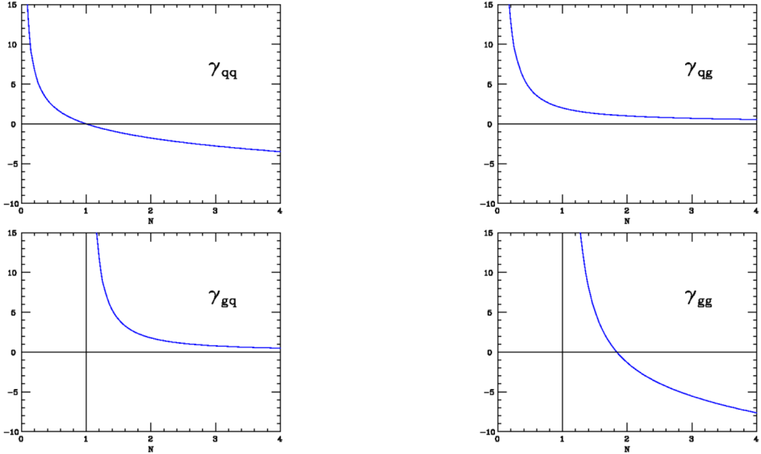
NNLO cont'd (2004)

[NPB 691 (2004) 129]



aN^{3}LO (2020 - ongoing)

Anomalous Dimensions: Scale Dependence at LO



As Q^2 increases, PDFs decrease at large x and increase at small x due to radiation

Gluon sector singular at $N = 1$, therefore the gluon grows fater at small x

$\gamma_{qq}(1) = 0$ follows from baryon number conservation (beyond LO, $\gamma_{qq}(1) = \gamma_{q\bar{q}}(1)$)

$\gamma_{qq}(2) + \gamma_{qg}(2) = \gamma_{gq}(2) + \gamma_{gg}(2) = 0$ follows from momentum conservation

Theoretical constraints

1. Momentum sum rule (momentum conservation)

$$\int_0^1 dx x \left[\sum_{q=1}^{n_f} \left(f_q(x,Q^2) + f_{\bar{q}}(x,Q^2) \right) + f_g(x,Q^2) \right] = 1$$

2. Valence sum rules (baryon number conservation)

$$\int_0^1 dx \left[f_u(x,Q^2) - f_{\bar{u}}(x,Q^2) \right] = 2$$

$$\int_0^1 dx \left[f_d(x,Q^2) - f_{\bar{d}}(x,Q^2) \right] = 1$$

$$\int_0^1 dx \left[f_q(x,Q^2) - f_{\bar{q}}(x,Q^2) \right] = 0 \qquad q = s, c, b, t$$

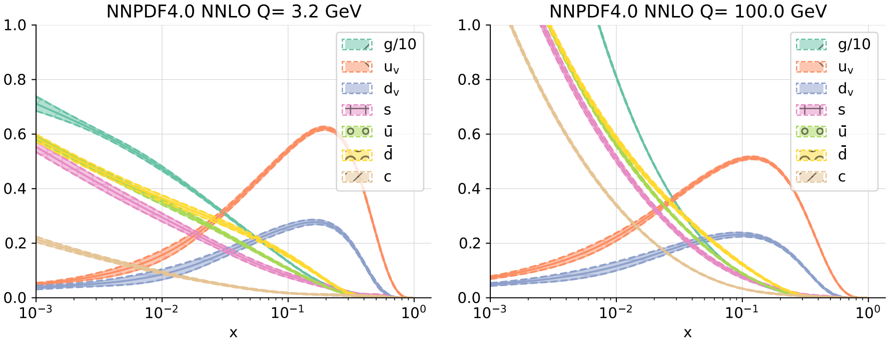
3. Isospin symmetry of the strong interaction

$$f_u^p = f_d^n \qquad f_{\bar{u}}^p = f_{\bar{d}}^n$$

4. Positivity of cross sections [PRD 105 (2022) 076010; EPJ C84 (2024) 335]
   → PDFs should be positive-definite at LO
   → beyond LO, PDFs ought not be positive, however they are positive for Q^2 large

5. Integrability of non-singlet PDFs
   → follows from operator product expansion

Unpolarised PDFs: Qualitative features



The valence bump follows from sum rules

The small-x growth of the gluon PDF follows from singularity of γ_{gg} at $N=1$

The similar small-x rise of all PDFs follows from singlet-gluon mixing

PDF depletion at large x and Q^2 follows from sign change of anomalous dimensions

Valence does not evolve multiplicatively because $\gamma_- \neq \gamma_v$

Valence does not vanish at all scales

What about the polarised case?

1. Coefficient functions

$$\Delta\mathcal{C}_{If}(y,\alpha_s)=\sum_{k=0} a_s^k \Delta C_{If}^{(k)}(y)$$

- DIS (up to NNLO) [NPB 417 (1994) 61]
- SIDIS (up to NNLO) [arXiv:2404.08597; arXiv:2404.09959]
- *pp* (up to (N)NLO)
  - [PRD 70 (2004) 034010]
  - [PLB 817 (2021) 136333]
  - [PRD 67 (2003) 054004, ibidem 054005]

2. Splitting functions

$$\Delta P_{ji}(z,\alpha_s)=\sum_{k=0} a_s^{k+1} \Delta P_{ji}^{(k)}(z)$$

- LO [NP B126 (1977) 298]
- NLO [ZP C70 (1996) 637, PR D54 (1996) 2023]
- NNLO [NP B889 (2014) 351]

$$P^{(0)}_{\mathrm{NS},qq}=P^{(0)}_{\mathrm{S},qq}=2C_F\left[\frac{1+x^2}{(1-x)_+}+\frac{3}{2}\delta(1-x)\right]$$

$$P^{(0)}_{qg}=2T_R\left[x^2+(1-x)^2\right]$$

$$P^{(0)}_{gq}=2C_F\left[\frac{1+(1-x)^2}{x}\right]$$

$$P^{(0)}_{gg}=4C_A\left[\frac{x}{(1-x)_+}+\frac{1-x}{x}+x(1-x)\right]+\delta(1-x)\frac{11C_A-4n_fT_R}{3}$$

$$\Delta P^{(0)}_{\mathrm{NS},qq}=\Delta P^{(0)}_{\mathrm{S},qq}=P^{(0)}_{\mathrm{NS},qq}$$

$$\Delta P^{(0)}_{qg}=2T_R(2-x)$$

$$\Delta P^{(0)}_{gq}=C_F(2-x)$$

$$\Delta P^{(0)}_{gg}=2C_A\left[\frac{1}{(1-x)_+}+1-2x\right]+\delta(1-x)\frac{11C_A-4n_fT_R}{6}$$

Theoretical constraints

1. Spin sum rule (Jaffe and Manohar [NPB 337 (1990) 509])

$$\int_0^1 dx \left[\sum_{q=1}^{n_f} \left(\Delta f_q(x,Q^2) + \Delta f_{\bar{q}}(x,Q^2)\right) + \Delta f_g(x,Q^2)\right] + \mathcal{L}_q(Q) + \mathcal{L}_g(Q) = \frac{1}{2}$$

2. β decays of the octet baryon, assuming SU(2) and SU(3) symmetry

$$\int_0^1 dx \Delta T_3(x,Q^2) = a_3 = 1.2756 \pm 0.0013 \qquad \int_0^1 dx \Delta T_8(x,Q^2) = a_8 = 0.585 \pm 0.025$$

3. Isospin symmetry of the strong interaction

$$\Delta f_u^p = \Delta f_d^n \qquad \Delta f_{\bar{u}}^p = \Delta f_{\bar{d}}^n$$

4. Positivity of cross sections
   $\rightarrow$ at LO, polarised PDFs are bound by unpolarised PDFs

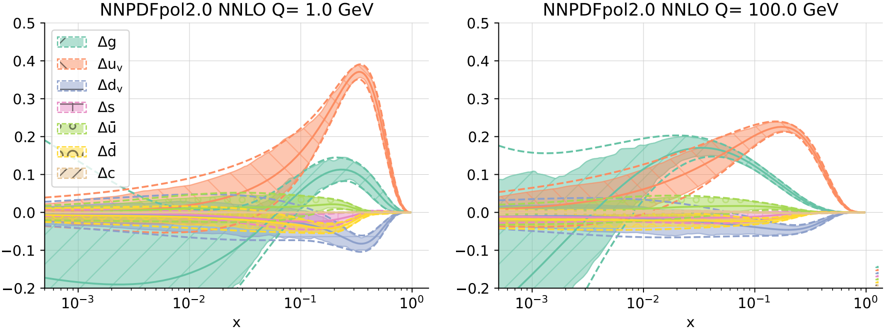
$$|\Delta f(x,Q^2)| < f(x,Q^2) \qquad \text{which follows from} \qquad |g_1(x,Q^2)| < F_1(x,Q^2)$$

   $\rightarrow$ beyond LO, other relations hold, but are of limited effect [NP B534 (1998) 277]

5. Polarised PDFs ought to be integrable (nucleon axial matrix elements are finite

$$\langle P,S|\bar{\Psi}_i\gamma^\mu\Psi_i|P,S\rangle \longrightarrow \text{finite for each parton } i$$

Polarised PDFs: Qualitative features



Polarised PDFs can be negative, as they are defined as spin differences

The small-x behaviour of polarised PDFs is suppressed by parton evolution

The valence peak moves towards large x as Q increases, as for unpolarised PDFs

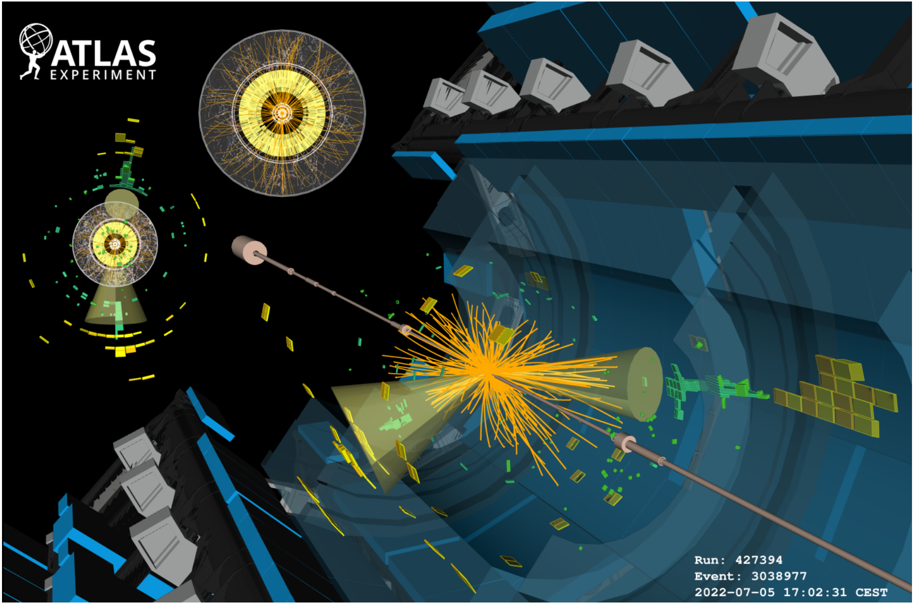
PDF depletion at large x and Q^2 follows from sign change of anomalous dimensions

Valence does not evolve multiplicatively because $\gamma_- \neq \gamma_v$

Valence does not vanish at all scales

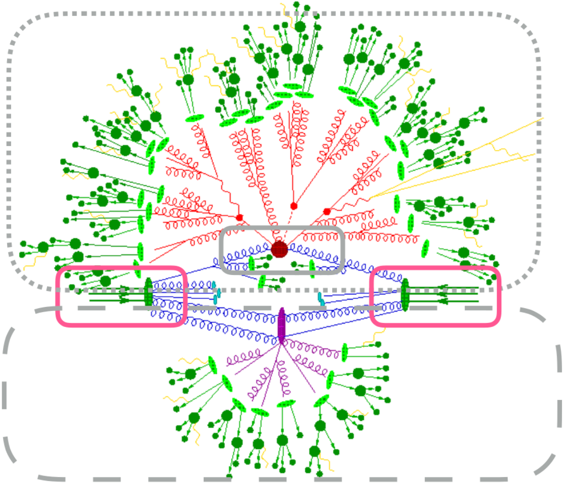
1.2 Why Parton Distribution Functions?

First Collisions of LHC Run III



[Image credit: ATLAS collaboration]

A Laboratory for Quantum Chromodynamics



Hard scattering of partons (Perturbative QCD+EW)

Parton Distribution Functions

Parton Showering and Hadronisation

Multi-Parton Interactions Underlying Events

[Plot by courtesy of SHERPA]

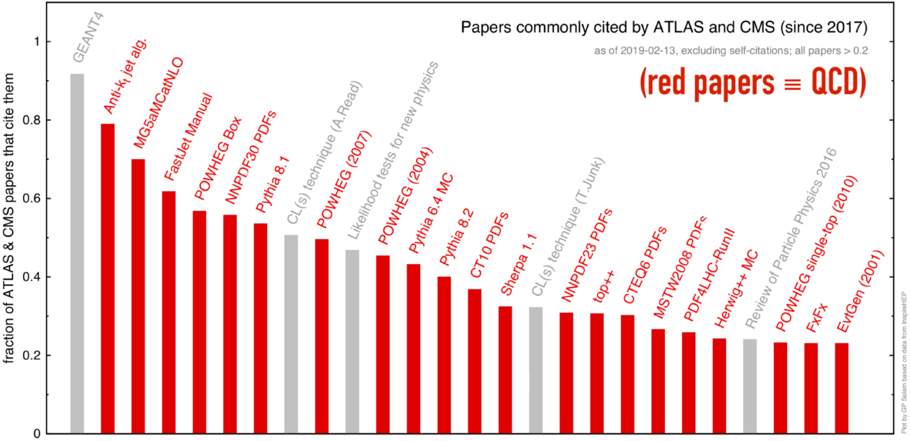
$$\sigma(\tau, Q^2, \mathbf{k}) = \sum_{ij} \int_\tau^1 \frac{dz}{z} \hat{\sigma}_{ij} \left(\frac{\tau}{z}, \alpha_s(Q^2), \mathbf{k} \right) \mathcal{L}_{ij}(z, Q^2) \quad \mathcal{L}_{ij}(z, Q^2) = (f_i^{h_1} \otimes f_j^{h_2})(z, Q^2)$$

LHC, QCD and unpolarised PDFs

The LHC is a Proton Collider – Any interaction contains a strong interaction

Quantum Chromodynamics (QCD) is the main actor

Within QCD, Parton Distribution Functions (PDFs) play a leading role



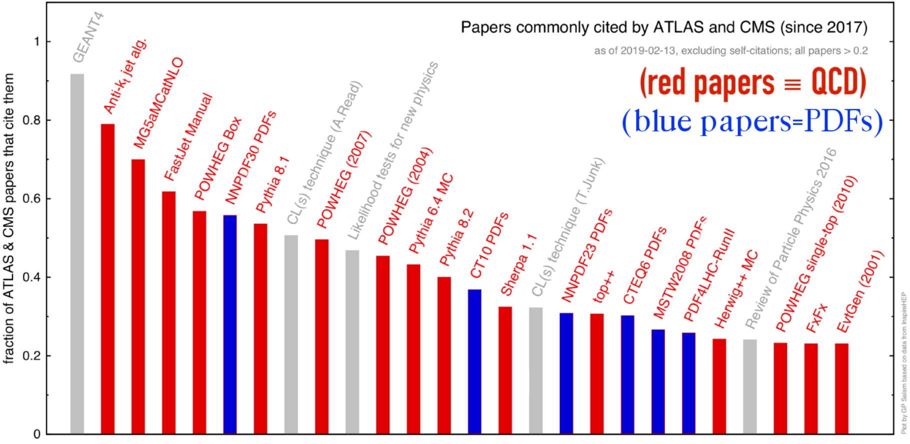
[Plot by courtesy of G. Salam]

LHC, QCD and unpolarised PDFs

The LHC is a Proton Collider – Any interaction contains a strong interaction

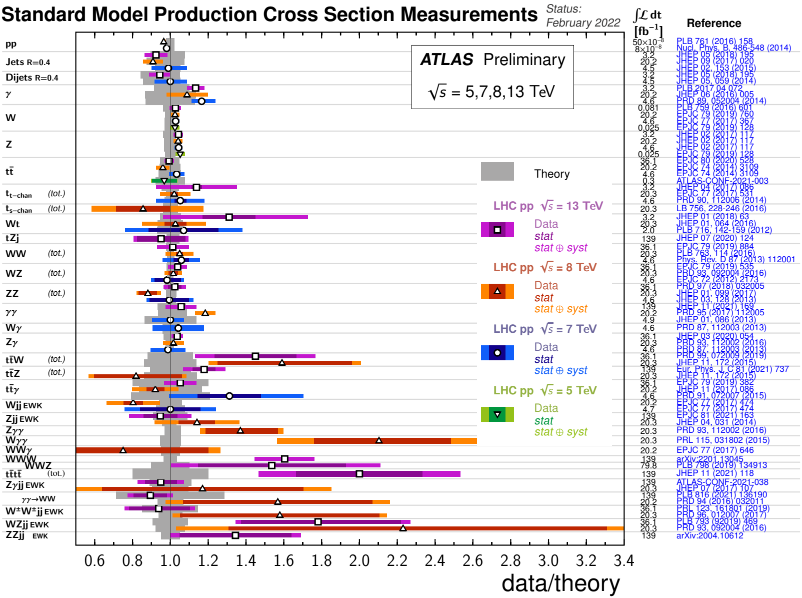
Quantum Chromodynamics (QCD) is the main actor

Within QCD, Parton Distribution Functions (PDFs) play a leading role



[Plot by courtesy of G. Salam]

Physics at the LHC as Precision Physics

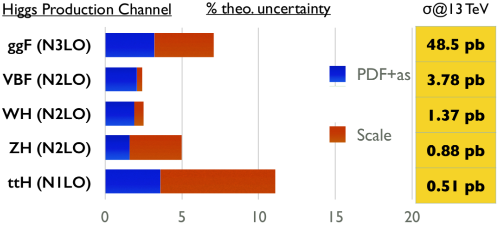


[Plot from ATLAS Collaboration web page]

PDFs as a Tool: Making Predictions with PDFs

PDF uncertainty is often the dominant source of uncertainty in LHC cross sections

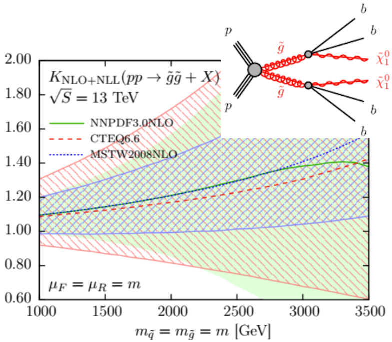
Precision



Channel	$m_{W^+} - m_{W^-}$ [MeV]	Stat. Unc.	Muon Unc.	Elec. Unc.	Recoil Unc.	Bckg. Unc.	QCD Unc.	EW Unc.	PDF Unc.	Total Unc.
$W \to e\nu$	−29.7	17.5	0.0	4.9	0.9	5.4	0.5	0.0	24.1	30.7
$W \to \mu\nu$	−28.6	16.3	11.7	0.0	1.1	5.0	0.4	0.0	26.0	33.2
Combined	−29.2	12.8	3.3	4.1	1.0	4.5	0.4	0.0	23.9	28.0

[Plot from the CERN Yellow Report 2016]

Discovery



[EPJC 76 (2016) 53]

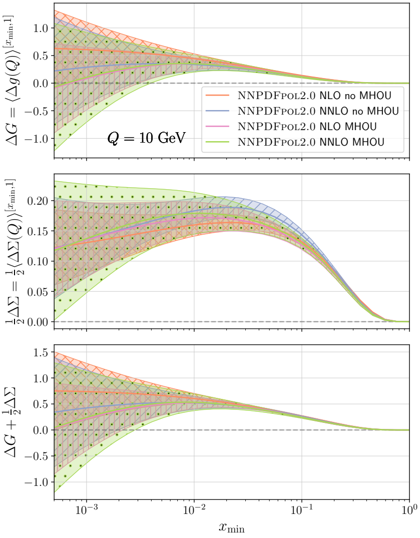
Higgs boson characterisation

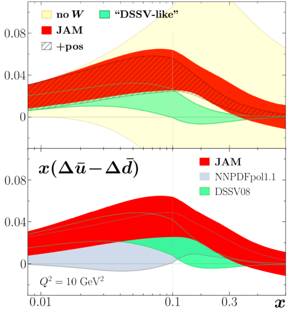
Determination of SM parameters, such as the mass of the W boson

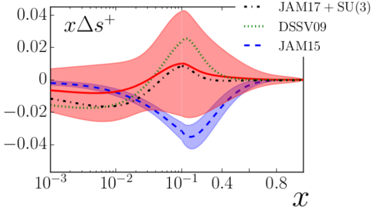
Searches for beyond SM physics at large invariant mass of the final state

And polarised PDFs?

How is the spin of the proton distributed across the spin of partons?







[arXiv:2503.11814; PRD 106 (2022) L031502; PRL 119 (2017) 132001]

1.3 How can we determine PDFs?

PDF determination in statistical language

Inverse problem

Given a set of data D, determine $p(f|D)$ in the space of functions $f : [0,1] \to \mathbb{R}$.

The expectation value and uncertainty of each physical observable $\mathcal{O}$ that depends on a PDF set $[f]$ are functional integrals of the PDFs

$$\langle \mathcal{O}[f] \rangle = \int \mathcal{D}f \, p(f|D) \, \mathcal{O}[f] \qquad \text{expectation value}$$

$$\sigma_{\mathcal{O}}[f] = \left[\int \mathcal{D}f \, p(f|D) \, \left(\mathcal{O}[f] - \langle \mathcal{O}[f] \rangle \right)^2 \right]^{1/2} \qquad \text{uncertainty}$$

THE PROBLEM IS ILL-DEFINED

We want to determine infinite-dimensional objects, the PDFs, from a finite set of data

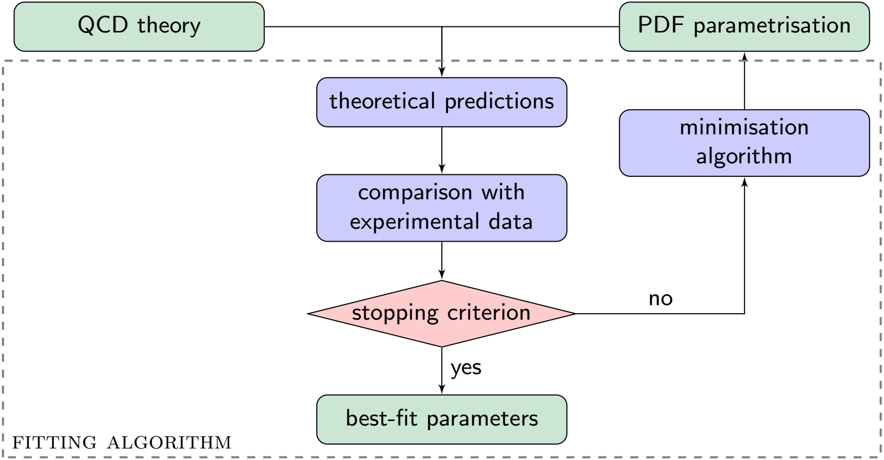
Solution: parametric regression

Approximate $p(f|D)$ with its projection in the space of parameters $p(\boldsymbol{\theta}|D,\mathcal{H})$

Determine $p(\boldsymbol{\theta}|D,\mathcal{H}) \propto p(D|\boldsymbol{\theta},\mathcal{H}) p(\boldsymbol{\theta}|\mathcal{H})$ as MAP $\boldsymbol{\theta}^* = \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|D,\mathcal{H})$

Determining PDFs from (LHC) experimental data

$$\sigma(\tau, Q^2\mathbf{k}) = \sum_{ij}\int_\tau^1 \frac{dz}{z}\hat{\sigma}_{ij}\left(\frac{\tau}{z}, \alpha_s(Q^2), \mathbf{k}\right)\mathcal{L}_{ij}(z, Q^2) \quad \mathcal{L}_{ij}(z, Q^2) = (f_i^{h_1}\otimes f_j^{h_2})(z, Q^2)$$



$$\chi^2 = \sum_{i,j}^{N_{\rm dat}}[T_i[\{\vec{a}\}] - D_i](\text{cov}^{-1})_{ij}[T_j[\{\vec{a}\}] - D_j]$$ with $\{\vec{a}\}$ the set of parameters

The ingredients of PDF determination

DATA
or
the set of observables included in the analysis

experimentalists' input of the QCD analysis

THEORY
or
the theoretical details of the QCD analysis

theorists' tools for precision physics

METHODOLOGY
or
a prescription to determine PDFs

uncertainty representation

parametrisation

optimisation

validation

Each of these ingredients is a source of uncertainty in the PDF determination
Each of these ingredients require to make choices which lead to different PDF sets

Overview of current unpolarised PDF determinations

	NNPDF4.0	MSHT20	CT18	HERAPDF2.0	CJ22	ABMP16
Fixed-target DIS	☑	☑	☑	☒	☑	☑
JLAB	☒	☒	☒	☒	☑	☒
HERA I+II	☑	☑	☑	☑	☑	☑
HERA jets	☑	☒	☒	☑	☒	☒
Fixed target DY	☑	☑	☑	☒	☑	☑
Tevatron W, Z	☑	☑	☑	☒	☑	☑
LHC vector boson	☑	☑	☑	☒	☑	☑
LHC $W+c$ $Z+c$	☑	☒	☒	☒	☒	☒
Tevatron jets	☑	☑	☑	☒	☑	☒
LHC jets	☑	☑	☑	☒	☒	☒
LHC top	☑	☑	☒	☒	☒	☑
LHC single t	☑	☒	☒	☒	☒	☒
LHC prompt γ	☑	☒	☒	☒	☒	☒
statistical treatment	Monte Carlo	Hessian $\Delta\chi^2$ dynamical	Hessian $\Delta\chi^2$ dynamical	Hessian $\Delta\chi^2 = 1$	Hessian $\Delta\chi^2 = 1.645$	Hessian $\Delta\chi^2 = 1$
parametrisation	Neural Network	Chebyschev pol.	Bernstein pol.	polynomial	polynomial	polynomial
HQ scheme	FONLL	TR$'$	ACOT-χ	TR$'$	ACOT-χ	FFN
accuracy	aN^{3}LO	aN^{3}LO	NNLO	NNLO	NLO	NNLO
latest update	EPJ C82 (2022) 428	EPJ C81 (2021) 341	PRD 103 (2021) 014013	EPJ C82 (2022) 243	PRD 107 (2023) 113005	PRD 96 (2017) 014011

All PDF sets are available as (x, Q^2) interpolation grids through the LHAPDF library

1.4 Summary of Lecture 1

Summary of Lecture 1

1. Deep Inelastic Scattering has been, is, and will be a crucial laboratory of QCD
   $\longrightarrow$ hadronic structure encoded in unpolarised and polarised PDFs
   $\longrightarrow$ PDFs are related to physical observales via factorisation and evolution
   $\longrightarrow$ (critically different) qualitative PDF features are driven by this theoretical framework
2. PDFs are a limiting factor for precision and discovery
   $\longrightarrow$ unpolarised PDFs: SM and BSM physics at the LHC
   $\longrightarrow$ polarised PDFs: contribution of partons' spin to the proton spin
3. PDFs are determined from experimental data by means of parametric regression
   $\longrightarrow$ need to define data, theory, and methodology
4. Different physical observables constrain different PDF combinations
   $\longrightarrow$ fixed-target NC DIS: u and d
   $\longrightarrow$ fixed-target CC DIS: s and $\bar{s}$
   $\longrightarrow$ HERA NC and CC DIS: u, $\bar{u}$, d, $\bar{d}$, g (scaling violations and tagged DIS)
   $\longrightarrow$ fixed-target DY: u and d at large x
   $\longrightarrow$ collider DY: u, $\bar{u}$, d, $\bar{d}$, s
   $\longrightarrow$ collider DY+c: s (W) and c (Z)
   $\longrightarrow$ Zp_T, $t\bar{t}$, jets: g
   $\longrightarrow$ only a small fraction of the above is available for polarised PDFs

Lecture 2: Data, theoretical, and methodological accuracy in PDF determination