PROTON STRUCTURE
THE LAST LIGHT PARTON

Gavin Salam, CERN
including work with Aneesh Manohar, Paolo Nason and Giulia Zanderighi

Guido Altarelli Memorial Session
5th International Conference on New Frontiers in Physics, Crete, July 2016
WHAT DO ATLAS & CMS USE MOST FREQUENTLY?

Papers commonly cited by ATLAS and CMS (2014-2016)
as of 2016-06-10, excluding self-citations; all papers > 0.2

fraction of ATLAS & CMS papers that cite them

CTEQ6 PDFs
POWHEG (2007)
POWHEG Box
MSTW2008 PDFs
CL(s) technique (A.Read)
MadGraph 5
Sherpa 1.1
PDF4LHC (2011)
NNLO Itbar
Likelihood tests for new physics
ALPGEN
MC@NLO
FastJet Manual
top+
Pileup subtraction
CL(s) technique (T.Junk)
MC@NLO
WHAT DO ATLAS & CMS USE MOST FREQUENTLY?

Papers commonly cited by ATLAS and CMS (2014-2016)
as of 2016-06-10, excluding self-citations; all papers > 0.2

Nearly all of it is QCD
WHAT DO ATLAS & CMS USE MOST FREQUENTLY?

Papers commonly cited by ATLAS and CMS (2014-2016)

as of 2016-06-10, excluding self-citations; all papers > 0.2

The core of hadron-collider QCD is parton distribution functions (PDFs)
ASYMPTOTIC FREEDOM IN PARTON LANGUAGE

G. ALTARELLI *
Laboratoire de Physique Théorique de l'Ecole Normale Supérieure **, Paris, France

G. PARISI ***
Institut des Hautes Etudes Scientifiques, Bures-sur-Yvette, France

Received 12 April 1977

A novel derivation of the $Q^2$ dependence of quark and gluon densities (of given helicity) as predicted by quantum chromodynamics is presented. The main body of predictions of the theory for deep-inelastic scattering on either unpolarized or polarized targets is re-obtained by a method which only makes use of the simplest tree diagrams and is entirely phrased in parton language with no reference to the conventional operator formalism.
Initial-state splitting

1st order analysis

Summary so far

- Collinear divergence for incoming partons not cancelled by virtuals.
- Real and virtual have different longitudinal momenta.

Situation analogous to renormalization: need to regularize (but in IR instead of UV).

Technically, often done with dimensional regularization.

Physical sense of regularization is to separate (factorize) proton non-perturbative dynamics from perturbative hard cross section.

Choice of factorization scale, $\mu^2$, is arbitrary between $1 \text{ GeV}^2$ and $Q^2$.

In analogy with running coupling, we can vary factorization scale and get a renormalization group equation for parton distribution functions.

Dokshizer Gribov Lipatov Altarelli Parisi equations (DGLAP)
In analogy with running coupling, we can describe initial-state splitting. The physical sense of regularization is to separate non-perturbative dynamics from perturbative hard cross sections. This is done by introducing a factorization scale, $\mu$, which is set to some large value, $Q^2$. Technically, often done with dimensional regularization (but in IR, it is done in terms of $\mu$).

The renormalization group equation is derived by considering the evolution of the quark and antiquark densities, $q^i(x,t)$ and $\bar{q}^i(x,t)$, respectively, and the gluon density, $G(x,t)$.

For the quark densities, we have:

$$\frac{dq^i(x, t)}{dt} = \frac{\alpha(t)}{2\pi} \int_\chi \frac{dy}{y} \left[ \sum_{j=1}^{2f} q^j(y, t) P_{qi}q^j \left( \frac{x}{y} \right) + G(y, t) P_{qi}G \left( \frac{x}{y} \right) \right],$$

where $\alpha(t)$ is the coupling constant.

For the gluon density, we have:

$$\frac{dG(x, t)}{dt} = \frac{\alpha(t)}{2\pi} \int_\chi \frac{dy}{y} \left[ \sum_{j=1}^{2f} q^j(y, t) P_{Gq}j \left( \frac{x}{y} \right) + G(y, t) P_{GG} \left( \frac{x}{y} \right) \right].$$
WHAT DO ATLAS & CMS USE MOST FREQUENTLY?

Papers commonly cited by ATLAS and CMS (2014-2016)

as of 2016-06-10, excluding self-citations; all papers > 0.2

This probabilistic “picture”, so clear in the AP paper underpins the rest of QCD at LHC
WHAT DO ATLAS & CMS USE MOST FREQUENTLY?

Papers commonly cited by ATLAS and CMS (2014-2016)

as of 2016-06-10, excluding self-citations; all papers > 0.2

This probabilistic “picture”, so clear in the AP paper underpins the rest of QCD at LHC

Plot by GPI Salam based on data from InspireHEP
WHAT DO ATLAS & CMS USE MOST FREQUENTLY?

Papers commonly cited by ATLAS and CMS (2014-2016)
as of 2016-06-10, excluding self-citations; all papers > 0.2

A summary of DGLAP’s influence at the LHC
impact of DGLAP evolution from $Q_0 = 2 \text{ GeV}$

$g(x, Q)$

CT14nnlo

LHC 13 TeV phys. region
impact of DGLAP evolution from $Q_0 = 2$ GeV

gluon, $Q = 3$ GeV

CT14nnlo

LHC 13 TeV phys. region
impact of DGLAP evolution from $Q_0 = 2$ GeV

gluon, $Q = 4$ GeV

CT14nnlo

LHC 13 TeV phys. region
impact of DGLAP evolution from $Q_0 = 2$ GeV

$g(x, Q)$, $Q = 5$ GeV

CT14nnlo

LHC 13 TeV phys. region
impact of DGLAP evolution from $Q_0 = 2$ GeV

 gluon, $Q = 10$ GeV

CT14nnlo

$LHC$ 13 TeV phys. region
impact of DGLAP evolution from $Q_0 = 2$ GeV

$\text{gluon, } Q = 50$ GeV

CT14nnlo

$LHC 13$ TeV phys. region
impact of DGLAP evolution from $Q_0 = 2$ GeV

$g(x, Q = 100$ GeV

CT14nnlo

$LHC 13$ TeV phys. region
DGLAP evolution changes parton distributions by factors $\sim 10$

Higgs cross section (13 TeV) would be 6x smaller without DGLAP nowadays, used at NNLO, thanks to Moch, Vermaseren & Vogt
EXPERIMENTAL PRECISION TODAY CAN REACH 1%
WHAT ACCURACY DO WE NEED? E.G. FOR LONG-TERM HIGGS PRECISION

**Naive extrapolation suggests LHC has long-term potential to do Higgs (and much other) physics at **1% accuracy**
how well do we know the parton distributions?
PDF uncertainties (Q = 100 GeV)

➤ core partons (up, down, gluon) are quite well known
PDF uncertainties (Q = 100 GeV)

- core partons (up, down, gluon) are quite well known ~2%
- strangenessness ~10%
(core partons (up, down, gluon) are quite well known \(\sim 2\%\))

- strangeness \(\sim 10\%\)

- one other parton, the photon, is debated. The only model-independent determination (NNPDF23qed) has \(O(100\%)\) uncertainty
IT MATTERS FOR DI-LEPTON, DI-BOSON, TTBAR, EW HIGGS, ETC.

**di-lepton spectrum**

\[ \frac{d\sigma}{dM_{ll}} = \text{normal DY contribution} \]

\[ \sqrt{s} = 13 \text{ TeV} \]

Accomando et al, 1606.06646

**photon-induced contribution and uncertainty [NNPDF23]**

\[ \frac{d\sigma}{dM_{ll}} = \text{photon-induced contribution and uncertainty} \]

\[ M_{ll} [\text{TeV}] \]

\[ \frac{d\sigma}{dM_{ll}} [\text{fb} / \text{TeV}] \]
<table>
<thead>
<tr>
<th>Model</th>
<th>Type</th>
<th>Model</th>
<th>in LHAPDF?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gluck Pisano Reya 2002</td>
<td>elastic dipole</td>
<td>inelastic model</td>
<td>✗</td>
</tr>
<tr>
<td>MRST2004qed</td>
<td>✗</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>NNPDF23qed</td>
<td>no separation; fit to data</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>CT14qed</td>
<td>✗</td>
<td>model</td>
<td>✓</td>
</tr>
<tr>
<td>CT14qed_inc</td>
<td></td>
<td>model</td>
<td>✓</td>
</tr>
<tr>
<td>Martin Ryskin 2014</td>
<td>dipole (only electric part)</td>
<td></td>
<td>✗</td>
</tr>
<tr>
<td>Harland-Lang, Khoze Ryskin 2016</td>
<td>dipole</td>
<td>model</td>
<td>✗</td>
</tr>
</tbody>
</table>

*elastic: Budnev, Ginzburg, Meledin, Serbo, 1975*
YOU SHOULDN’T NEED A MODEL
ep scattering (i.e. structure functions) contains all info about proton’s EM field
to extract it, we’ll study a hypothetical (“BSM”) heavy-neutral lepton production process

Manohar, Nason, GPS & Zanderighi, to appear
STEP 1
work out a cross section (exact) in terms of $F_2$ and $F_L$ struct. fns.

\[ \sigma = \frac{1}{4 p \cdot k} \int \frac{d^4q}{(2\pi)^4 q^4} e_{ph}^2(q^2) \left[ 4\pi W_{\mu\nu} L^{\mu\nu}(k, q) \right] \times 2\pi \delta((k - q)^2 - M^2) \]
STEP 2
work out same cross section in terms of a photon distribution

\[ \hat{\sigma}_\gamma \left( \frac{M^2}{xS}, \mu^2 \right) \]

\[ f_{\gamma/p}(x, \mu^2) \]

\[ \sigma = c_0 \sum_a \int \frac{dx}{x} \hat{\sigma}_a \left( \frac{M^2}{xS}, \mu^2 \right) x f_{a/p}(x, \mu^2) \]
STEP 3
equate them to deduce the photon distribution (LUXqed)

\[ x f_{\gamma/p}(x, \mu^2) = \frac{1}{2\pi \alpha(\mu^2)} \int_x^1 \frac{dz}{z} \left\{ \int_{Q_{\text{min}}^2}^{Q_{\text{max}}^2} \frac{dQ^2}{Q^2} \alpha^2(Q^2) \right\} \]

\[ \left[ \left( 2 - 2z + z^2 + \frac{2x^2 m_p^2}{Q^2} \right) F_2(x/z, Q^2) \right] - z^2 F_L \left( \frac{x}{z}, Q^2 \right) - \alpha^2(\mu^2) z^2 F_2 \left( \frac{x}{z}, \mu^2 \right) \right\}, \]

Result is in MSbar scheme & consistent with 2015 de Florian, Rodrigo, Sborlini O(αα_s) P_γx QED split.fns.
# PHOTON PDF ESTIMATES (not exhaustive)

<table>
<thead>
<tr>
<th></th>
<th>elastic</th>
<th>inelastic</th>
<th>in LHAPDF?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gluck Pisano Reya 2002</td>
<td>dipole</td>
<td>model</td>
<td>✗</td>
</tr>
<tr>
<td>MRST2004qed</td>
<td>✗</td>
<td>model</td>
<td>✓</td>
</tr>
<tr>
<td>NNPDF23qed</td>
<td>no separation; fit to data</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>CT14qed</td>
<td>✗</td>
<td>model (data-constrained)</td>
<td>✓</td>
</tr>
<tr>
<td>CT14qed_inc</td>
<td>dipole</td>
<td>model (data-constrained)</td>
<td>✓</td>
</tr>
<tr>
<td>Martin Ryskin 2014</td>
<td>dipole (only electric part)</td>
<td>model</td>
<td>✗</td>
</tr>
<tr>
<td>Harland-Lang, Khoze Ryskin 2016</td>
<td>dipole</td>
<td>model</td>
<td>✗</td>
</tr>
<tr>
<td>LUXqed 2016</td>
<td>data</td>
<td>data</td>
<td>soon</td>
</tr>
</tbody>
</table>
DATA SOURCES – various fits to F2, FL & elastic form factors

high $Q^2$ continuum region (PDFs: PDF4LHC15_nnlo_100)

low $Q^2$ continuum region (Hermes GD11-P)

resonance region (CLAS/CB)

elastic (A1)
DATA SOURCES – various fits to F2, FL & elastic form factors

Figure 1: Elastic form factors (ratio to standard dipole form) as fitted by the A1 collaboration [B14]. Left: electric. Right: magnetic.

Figure 2: Elastic contribution to \( f/p \) \((x, Q^2)\) with various fits for the form factors, normalised to the result obtained with the A1 world fit, including polarised data. The ratio freezes above \( x = 0.9 \) because the A1 fits extend only up to \( Q^2 = 10 \text{GeV}^2 \) and beyond that scale we simply extrapolate the results for \( G_E/M_1 \) using the standard-dipole shape, normalised to \( G_E/M_1(10 \text{GeV}^2) \).

[Should we try to do this better? Maybe not so critical for now.]

Figure 9: HERMES data for the photon-proton cross section \( \sigma_p \) \(L+T\) as a function of \( W^2 \), together with world data and the results from the GD11-P fit (central curve) and its uncertainties (outer curves), in bins of \( Q^2 \). The data points denoted 'real photon' are for photoproduction. Inner error bars are statistical uncertainties, while outer error bars are total uncertainties calculated as the sum in quadrature of all statistical and systematic uncertainties including normalization.
the results

ratio of some widely used PDFs to LUXqed (red)
PHOTON UNCERTAINTY (1–2%) COMPARED TO OTHER FLAVOURS

PDF uncertainties (Q = 100 GeV)

- **photon** (LUXqed)
- **strange** (PDF4LHC15)
- **up** (PDF4LHC15)
$\gamma \gamma$ luminosity for $E_{CM} = 13$ TeV

d$\gamma \gamma$ / dln $M^2$

$LUX\text{qed}$

$NNPDF30$
### APPLICATION TO HIGGS PHYSICS

**Process:**

\[ pp \rightarrow H W^+ (\rightarrow l^+\nu) + X \text{ at 13 TeV} \]

<table>
<thead>
<tr>
<th>Contribution</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-photon induced contributions</td>
<td>91.2 ± 1.8 fb</td>
</tr>
<tr>
<td>Photon-induced contribs (NNPDF23)</td>
<td>6.0 ^+4.4_{-2.9} fb</td>
</tr>
<tr>
<td>Photon-induced contribs (LUXqed)</td>
<td>4.4 ± 0.1 fb</td>
</tr>
</tbody>
</table>

*non-photon numbers from LHCHXSWG*
CLOSING REMARKS

➤ LHC physics would be unrecognisable without Guido’s contributions, first and foremost the simple physical picture contained in the DGLAP equations.

➤ Parton distribution functions are among the crucial inputs to LHC physics, with significant open problems still to solve today.

➤ More generally, Guido’s dedication, his combination of breadth and attention to detail, all serve as a model for what a physicist may aspire to.
extra slides
the DY process initiated by a quark-antiquark interaction looks reasonably under control over a large portion of the invariant masses, compared to RunI. More in detail, the relative PDF error grows above 10% for a dilepton invariant mass of 4 TeV, as shown in Fig. 5c. The theoretical error on the central value of the DY dilepton spectrum is estimated from the standard deviation of a limited number of replicas. Following this approach, we have evaluated the differential cross section versus forward-backward asymmetry studies.

The theoretical error on the central value. The methods are basically two. CTEQ and MRST apply the Hessian method that exploits the PDF uncertainties with (magenta line) and without (blue line) the PI contribution. Standard acceptance cuts are applied.

The impact of these uncertainties in the context of resonant and non-resonant searches for a neutral massive vector boson ($Z_0$) is further consolidated. The theoretical error on the central value. The methods are basically two. CTEQ and MRST apply the Hessian method that exploits the PDF uncertainties with (magenta line) and without (blue line) the PI contribution. Standard acceptance cuts are applied.

At the LHC RunII with 13 TeV, the PDF uncertainties coming from the large-resonant and non-resonant mode is further consolidated. The theoretical error on the central value. The methods are basically two. CTEQ and MRST apply the Hessian method that exploits the PDF uncertainties with (magenta line) and without (blue line) the PI contribution. Standard acceptance cuts are applied.

The method is based on the standard deviation of a large set of replicas (order 100) that represent other possible fits of the data. For any observable, the central value is defined as the average of the different replicas. Following this approach, we have evaluated the difference between the central values coming from the different replicas and the PDF eigenvalues. In this approach, the error is estimated from the standard deviation of a limited number of replicas. The other procedure consists of paired PDF fits (order 20 pair of fits). The other procedure consists of paired PDF fits (order 20 pair of fits). The other procedure consists of paired PDF fits (order 20 pair of fits). The other procedure consists of paired PDF fits (order 20 pair of fits).

Elena Accomando,1, 2, * Juri Fiaschi,1, 2, † Francesco Hautmann,2, 3, ‡ Stefano Moretti,1, 2, § and C.H. Shepherd-Themistocleous1, 2, ¶

E-mail: claire.shepherd@stfc.ac.uk
E-mail: S.Moretti@soton.ac.uk
E-mail: hautmann@thphys.ox.ac.uk
E-mail: Juri.Fiaschi@soton.ac.uk
E-mail: E.Accomando@soton.ac.uk
The same result in terms of parton distributions can be written as
\[
\hat{F}(x) = \frac{1}{z \hat{p}} \hat{f}(x, \mu^2, E, M) + \hat{g}(x, \mu^2, E, M),
\]
where \( z \) is the photon self energy and \( \hat{f}(x, \mu^2) \) is the lepton vertex are renormalised.
SEPARATE CONTRIBUTIONS TO PHOTON PDF

FIG. 1. Our breakup of the contribution for large $x$, or equivalently $Q^2/x < 100$ GeV. There is a sizeable contribution arising from the elastic contribution, with an important magnetic competitive scale choice of $e$ fits [39–41] together with the known massless NNLO contributions.

The leading twist contribution to $x F_2$ for which we use the HERMES parametrization [36]. Both the GD11-P and Christy and Bosted (CB) [35] fits are constrained by photoproduction data over a large range of $W$ regions, as illustrated in Fig. 1. In the resonance region, we also consider an alternative fit, to the world data, by CLAS [34], and a fit by Hermes [36] based on W 2/3 that we use in each region.

We determine $(x,F_2)$ = $(x,R)$ = $(x,L)$ = $(x,G_1, G_2)$ at scales $Q^2 < 0.8 GeV^2$ because of the rapid drop-off. We also require $Q^2 < 5 GeV^2$.

PDF4LHC15 (PDF) = $1 + R$, which are related by $R$ = 10.5 GeV. They are stacked in [33] (E); an estimate of the uncertainty in the resonance region taken as the di 9. These fits are 0.5 GeV away from the CLAS fit and CB fits (RES); a systematic uncertainty due to the merger of global PDF contributions that we have considered and are shown in Fig. 2.

For the uncertainty on $x F_2$, or equivalently $R$, standard 68%CL uncertainties on the PDFs, applied linearly and consist of: a conservative estimate of the uncertainty on our calculation of the photon parton density function thanks to a long history of scattering studies. We also require $x F_2$ = 0.5 GeV.

We also use a fit to data by CLAS [34], and we use a fit to data by HERMES [36]. Both the GD11-P and Christy and Bosted (CB) [35] fits are constrained by photoproduction data over a large range of $W$ regions, as illustrated in Fig. 1. In the resonance region, we also consider an alternative fit, to the world data, by CLAS [34], and a fit by HERMES [36] based on $W 2/3$ that we use in each region.

In the resonance region, we also consider an alternative fit, to the world data, by CLAS [34], and a fit by HERMES [36] based on $W 2/3$ that we use in each region.

At high $Q^2$, $x F_2$ = $1 + R$, which are related by $R$ = 10.5 GeV. They are stacked in [33] (E); an estimate of the uncertainty in the resonance region taken as the di 9. These fits are 0.5 GeV away from the CLAS fit and CB fits (RES); a systematic uncertainty due to the merger of global PDF contributions that we have considered and are shown in Fig. 2.

In Fig. 3, we show the sources contributing to the uncertainty on the parton density function $x F_2$, or equivalently $R$, at our reference scale $Q^2 = 100$ GeV. They are stacked in [33] (E); an estimate of the uncertainty in the resonance region taken as the di 9. These fits are 0.5 GeV away from the CLAS fit and CB fits (RES); a systematic uncertainty due to the merger of global PDF contributions that we have considered and are shown in Fig. 2.
breakdown of uncertainties

uncertainties on R (R)  
higher orders (HO)  
pdf errors (PDF)  
error on elastic component (E)  
error on resonance region (RES)  
matching PDF and fits (M)  
twist 4 correction to R in PDF (T)  
sum in quadrature

LUXqed, \( \mu = 100 \) GeV
$\gamma\gamma$ luminosity for $E_{CM} = 8$ TeV

$dL_{\gamma\gamma} / d\ln M^2$

M [TeV]

- LUXqed
- NNPDF30