Recent results on NNPDF parton distribution and fragmentation functions

Electron Ion Collider User Group Meeting 2017

Emanuele R. Nocera

Rudolf Peierls Centre for Theoretical Physics - University of Oxford

University of Trieste - July 21 2017

Foreword

1. Guiding principle: leading-twist collinear factorisation

$$\mathcal{O}_I = \sum_{f=q,\bar{q},g} C_{If}(y,\alpha_s(\mu^2)) \otimes f(y,\mu^2) + \text{p.s. corrections} \qquad f \otimes g = \int_x^1 \frac{dy}{y} f\left(\frac{x}{y}\right) g(y)$$

2. Parametrisation: general, smooth, flexible at an initial scale Q_0^2

$$x f_i(x, Q_0^2) = A_{f_i}\, x^{a_{f_i}}\, (1-x)^{b_{f_i}}\, \mathscr{F}(x, \{c_{f_i}\}) \qquad \mathscr{F}(x, \{c_{f_i}\}) \text{ is a neural network}$$

3. A prescription to determine/compute expectation values and uncertainties

$$E[\mathcal{O}] = \int \mathcal{D}\Delta f \mathcal{P}(\Delta f|data)\mathcal{O}(\Delta f) \quad V[\mathcal{O}] = \int \mathcal{D}\Delta f \mathcal{P}(\Delta f|data)[\mathcal{O}(\Delta f) - E[\mathcal{O}]]^2$$

Monte Carlo: $\mathcal{P}(\Delta f|data) \longrightarrow \{\Delta f_k\}$

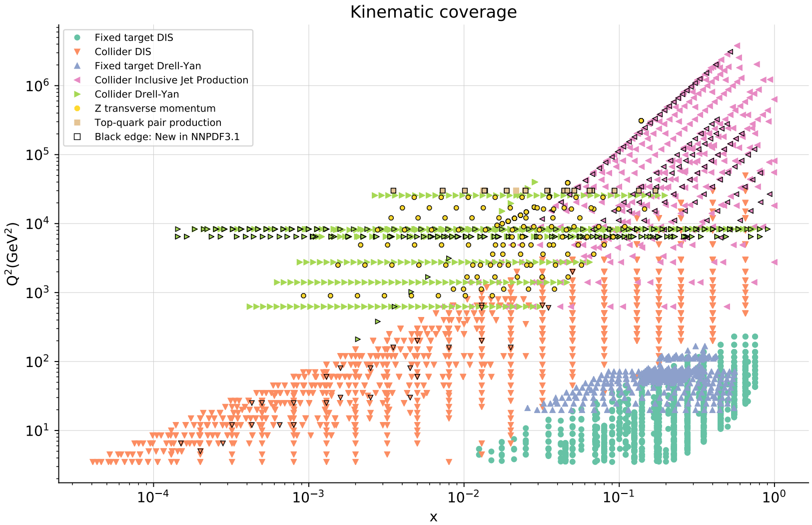
$$E[\mathcal{O}] \approx \tfrac{1}{N}\textstyle\sum_k \mathcal{O}(\Delta f_k) \qquad V[\mathcal{O}] \approx \tfrac{1}{N}\textstyle\sum_k [\mathcal{O}(\Delta f_k) - E[\mathcal{O}]]^2$$

Unpolarised PDFs	Helicity PDFs	Fragmentation functions
NNPDF3.1	NNPDFPOL1.1	NNFF1.0 (π, K, p)
[arXiv:1706:00428]	[NP B887 276]	[arXiv:1706.07049]

ALL PDF/FF sets are available through LHAPDF: http://lhapdf.hepforge.org/

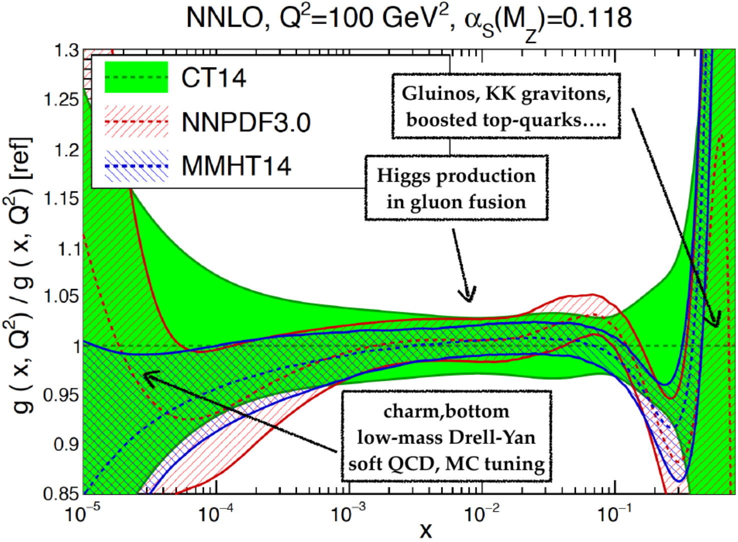
1. NNPDF3.1 [arXiv:1706:00428, submitted to EPJC]

The dataset



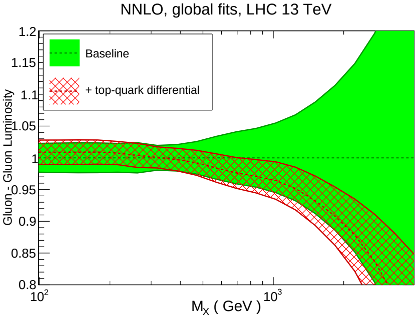
$$N_{\rm dat} = 4295 \qquad \chi^2/N_{\rm dat} = 1.168/1.148 \text{ (NLO/NNLO)}$$

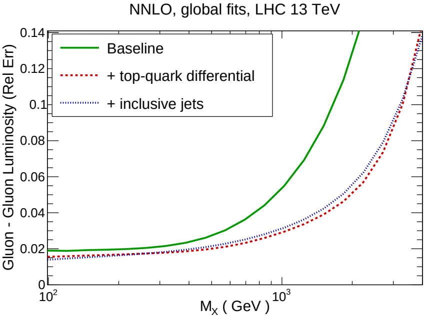
The gluon PDF



A precise knowledge of the gluon PDF is required
over all the range in x to exploit the full potential of the LHC

The gluon PDF at large x: $t\bar{t}$ differential distributions





ATLAS and CMS rapidity distributions at $\sqrt{s} = 8$ TeV

Significant reduction of gg luminosity uncertainties at $M_X \geq \mathcal{O}(1)$ TeV

e.g., at $M_X \sim 2$ TeV, uncertainties decrease from 13% to 5%

Impact of $t\bar{t}$ differential data similar to that of jet data

though jet data analysed neglecting NNLO QCD corrections in the matrix element

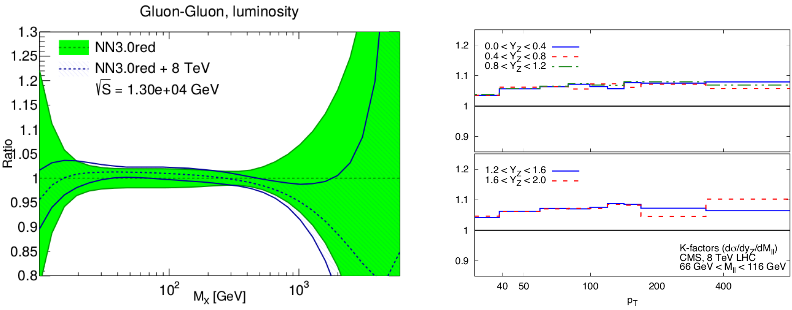
A precision determination of the gluon PDF at large x is now possible at NNLO

the situation should only improve thanks to the recent NNLO jet calculation

$t\bar{t}$ differential distributions are included in the NNPDF3.1 PDF release

[see JHEP 1704 (2017) 044 and 1706.00428 for details]

The gluon PDF at medium x: the Z-boson p_T distribution



ATLAS and CMS p_T distributions at $\sqrt{s} = 8$ TeV

in various rapidity bins in the Z-peak region

NNLO/NLO K-factors 5%-10% depending on the rapidity/invariant mass region

challenge: measurements have sub-percent experimental errors

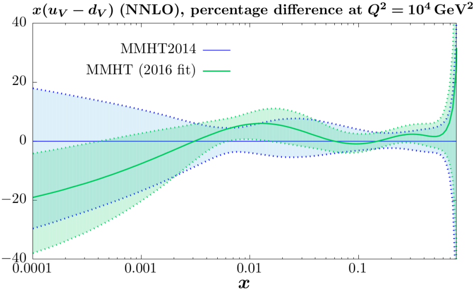
Complementary information on the gluon PDF

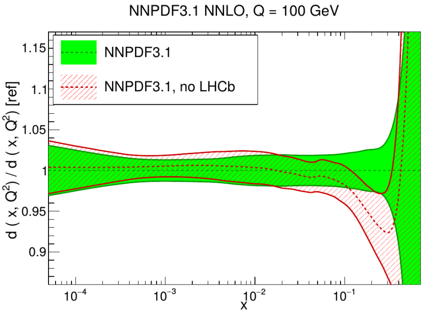
e.g., at $M_X \sim 2$ TeV, uncertainties decrease from 13% to 8%

Z p_T distributions are included in the NNPDF3.1 PDF release

[see 1705.00343 and 1706.00428 for details]

Quark flavour separation from LHC data





High-precision W and Z production data from ATLAS, CMS and LHCb

handle on quark/antiquark flavour separation

Largest impact on light quarks at large x provided by LHCb data

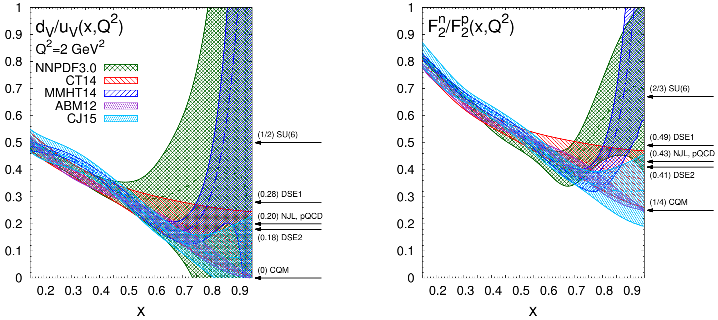
error reduction by a factor 2 in NNPDF3.1 at $x \sim 0.1$

Combined effect of (LHC) CMS, LHCb and (Tevatron) D0 W, Z data

improved determination of $x(u_V - d_V)$

[see R. Thorne's talk at DIS2017 and 1706.00428 for details]

The PDF ratio d_V/u_V at large x [EPJC 76 (2016) 383]



$$\frac{d_V}{u_V} \xrightarrow{x\to 1} (1-x)^{b_{d_V}-b_{u_V}} \xrightarrow[\text{c. r.}]{b_{d_V}=b_{u_V}} k \qquad \frac{F_2^n}{F_2^p} \xrightarrow{x\to 1} \frac{4(1-x)^{b_{u_V}}+(1-x)^{b_{d_V}}}{(1-x)^{b_{u_V}}+4(1-x)^{b_{d_V}}} \xrightarrow[\text{c. r.}]{b_{d_V}=b_{u_V}} 1$$

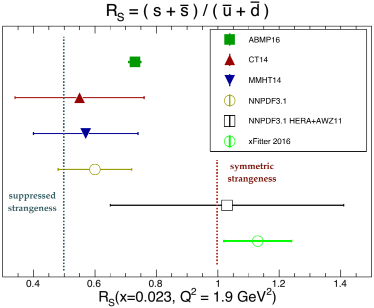
case $b_{u_V} \gg b_{d_V}$: $\frac{d_V}{u_V} \xrightarrow{x\to 1} \infty$; $\frac{F_2^n}{F_2^p} \xrightarrow{x\to 1} 4$ case $b_{u_V} \ll b_{d_V}$: $\frac{d_V}{u_V} \xrightarrow{x\to 1} 0$; $\frac{F_2^n}{F_2^p} \xrightarrow{x\to 1} \frac{1}{4}$

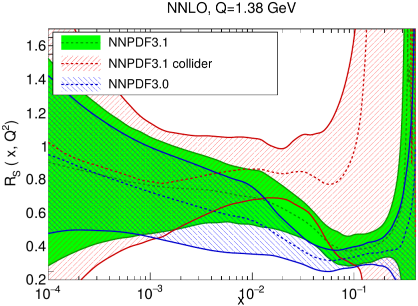
No predictive power from current PDF determinations, no discrimination among models unless $\frac{d_V}{u_V} \xrightarrow{x\to 1} k$ is built in the parametrisation (CT14, CJ16, ABM12)

The EIC may measure the ratio F_2^n/F_2^p with high accuracy, provided neutron beams expected to be less prone to nuclear and/or higher twist corrections than fixed-target DIS

Complementary measurements from the LHC (DY) and (particularly) the LHeC (DIS)

The strange PDF from collider data





In most PDF fits the strange PDF is suppressed w.r.t up and down sea quark PDFs

effect mostly driven by neutrino dimuon data

A symmetric strange sea PDF is preferred by recent collider data

in particular by ATLAS W, Z rapidity distributions (2011)

$$R_s(x,Q^2) = \frac{s(x,Q^2)+\bar{s}(x,Q^2)}{\bar{u}(x,Q^2)+\bar{d}(x,Q^2)} \begin{cases} \sim 0.5 \text{from neutrino and CMS } W+c \text{ data} \\ \sim 1.0 \text{from ATLAS } W, Z \end{cases}$$

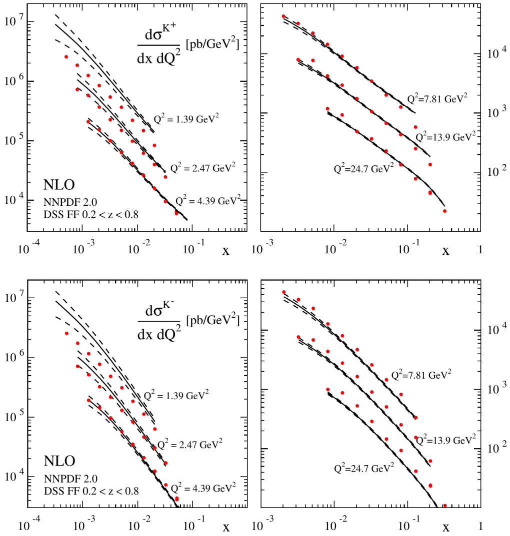
The new ATLAS data can be accommodated easily in the global fit

increased strangeness, though not as much as in a collider-only fit

some tension remains between collider and neutrino data

[see R. Thorne's talk at DIS2017 and 1706.00428 for details]

The strange PDF: $K^{\pm}$ production in SIDIS at an EIC



[figure taken from arXiv:1108:1713]

red points: pseudodata at an EIC (based on PYTHIA + JETSET)

black curves: theory predictions (NNPDF2.0 + DSS07, NLO)

$0.01 \leq y \leq 0.95$, $\sqrt{s} = 70.7$ GeV
z integrated in the range $[0.2, 0.8]$

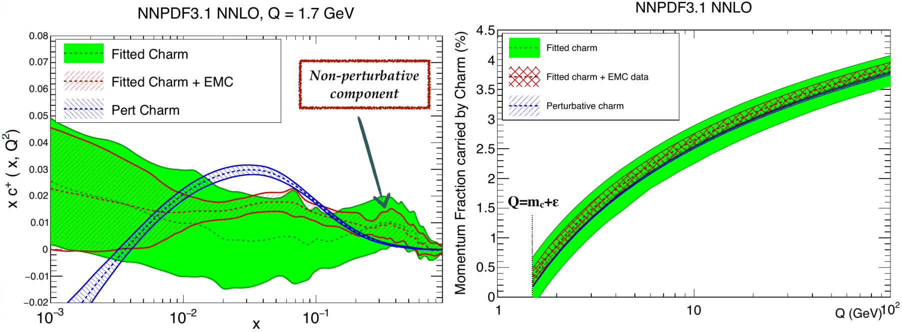
small x: $d\sigma^{K^+} \approx d\sigma^{K^-}$
large x: $d\sigma^{K^+} \gg d\sigma^{K^-}$
may constrain s^+ and s^-

drawback: $K^{\pm}$ fragmentation
a) study FFs separately
b) analyse PDFs and FFs simultaneously

LHeC: direct sensitivity to s
charm tagging in CC DIS ($W + s \to c$)

$\pi^{\pm}$ production in SIDIS at an EIC
allow for a determination of $\bar{u} - \bar{d}$

The charm PDF: perturbative vs fitted



Parametrise the $c^+(x,Q_0^2)$, quark and gluon PDFs on the same footing

take into account massive charm-initiated contribution to the DIS structure functions

stabilise the dependence of LHC processes upon variations of m_c

quantify the nonperturbative charm component in the proton (BHPS? sea-like?)

Fitted charm found to differ from perturbative charm at scales $Q \sim m_c$ in NNPDF3.1

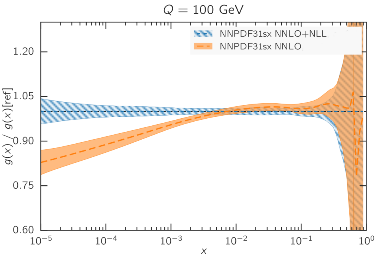
preference for a BHPS-like shape

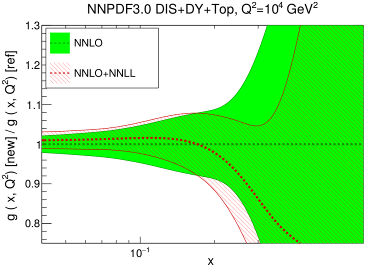
shape driven by LHCb W, Z data + EMC data

At $Q = 1.65$ GeV charm carry 0.26 ± 0.42 % of the proton momentum

but it is affected by large uncertainties, especially if no EMC data are included

Beyond fixed-order accuracy





small x: $\frac{1}{x}\ln^k x$
high-energy gluon emission: single logs

large x: $\left(\frac{\ln^k(1-x)}{(1-x)}\right)_+$
soft gluon emission: double logs

Large logs $\alpha_s \ln \sim 1$ spoil the convergence of the perturbative series

PDFs with threshold resummation [JHEP 1509 (2015) 191] (only DIS, DY Z/γ, total $t\bar{t}$)
suppression in PDFs partially or totally compensates enhancements in partonic cross-sections
accuracy of the resummed fit competitive with the fixed-order fit, except for the large-x gluon
large uncertainties for MSSM particle resummed cross-sections [EPJ C76 (2016) 53]

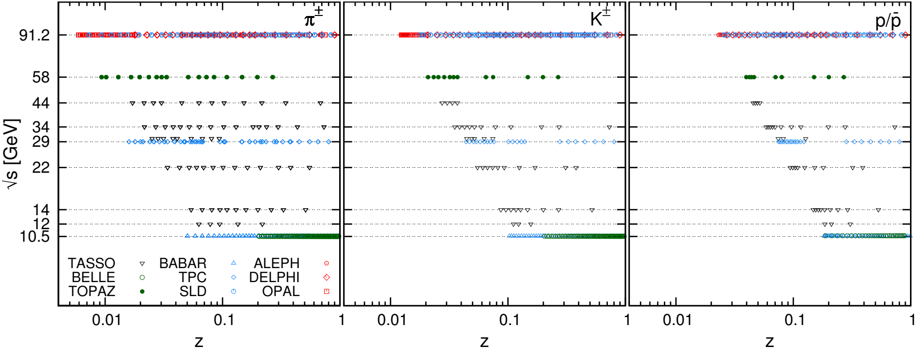
Towards PDFs with high-energy resummation [NNPDF, in progress] (DIS only)

$N_{\rm dat}$	NLO+NLL	NLO	NNLO+NLL	NNLO
3102	1.111	1.113	1.108	1.126

Resummed PDFs enhanced at small x, uncertainties reduced

2. NNFF1.0 [arXiv:1706.07049, accepted for publication in EPJC]

The dataset



CERN-LEP: ALEPH [ZP C66 (1995) 353] DELPHI [EPJ C18 (2000) 203] OPAL [ZP C63 (1994) 181]

KEK: BELLE ($n_f = 4$) [PRL 111 (2013) 062002] TOPAZ [PL B345 (1995) 335]

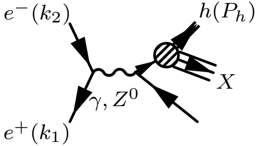
DESY-PETRA: TASSO [PL B94 (1980) 444, ZP C17 (1983) 5, ZP C42 (1989) 189]

SLAC: BABAR ($n_f = 4$) [PR D88 (2013) 032011] SLD [PR D58 (1999) 052001] TPC [PRL 61 (1988) 1263]

$$\frac{d\sigma^h}{dz} = \frac{4\pi\alpha^2(Q^2)}{Q^2} F_2^h(z,Q^2) \quad h = \pi^+ + \pi^-,\ K^+ + K^-, p + \bar{p}$$ possibly normalised to $\sigma_{\rm tot}$

$N_{\rm dat} = 428$ (pions) $\quad N_{\rm dat} = 385$ (kaons) $\quad N_{\rm dat} = 360$ (protons)

From observables to fragmentation functions



$$e^+(k_1) + e^-(k_2) \xrightarrow{\gamma, Z^0} h(P_h) + X$$

$$q = k_1 + k_2 \qquad q^2 = Q^2 > 0 \qquad z = \frac{2P_h \cdot q}{Q^2}$$

$$\boxed{F_2^h = \langle e^2 \rangle \left\{ C_{2,q}^{\mathrm{S}} \otimes D_\Sigma^h + n_f \mathcal{C}_{2,g}^{\mathrm{S}} \otimes D_g^h + \mathcal{C}_{2,q}^{\mathrm{NS}} \otimes D_{\mathrm{NS}}^h \right\}}$$

$$\langle e^2 \rangle = \frac{1}{n_f} \sum_{q=1}^{n_f} \hat{e}_q^2 \qquad D_\Sigma^h = \sum_{q=1}^{n_f} D_{q^+}^h \qquad D_{\mathrm{NS}}^h = \sum_{q=1}^{n_f} \left(\frac{\hat{e}_q^2}{\langle e^2 \rangle} - 1 \right) D_{q^+}^h \qquad D_{q^+}^h = D_q^h + D_{\bar{q}}^h$$

Coefficient functions and splitting functions known up to NNLO

[NPB 751 (2006) 18; NPB 749 (2006) 1; PLB 638 (2006) 61; NPB 845 (2012) 133]

No sensitivity to individual quark and antiquark FFs

Limited sensitivity to flavour separation via the variation of $\hat{e}_q$ with Q^2

$\hat{e}_u^2/\hat{e}_d^2(Q^2 = 10\,\mathrm{GeV}) \sim 4 \Rightarrow D_{u^+}^h,\ D_{d^+}^h + D_{s^+}^h;\quad \hat{e}_u^2/\hat{e}_d^2(Q^2 = M_Z) \sim 0.8 \Rightarrow D_\Sigma^h$

Flavor separation between uds and c, b quarks achieved thanks to tagged data

Direct sensitivity to D_g^h only beyond LO, as $C_{2,g}^{\mathrm{S}}$ is $\mathcal{O}(\alpha_s^2)$

Indirect sensitivity to D_g^h via scale violations in the time-like DGLAP evolution

Fit settings

Physical parameters: consistent with the NNPDF3.1 PDF set

$$\alpha_s(M_Z) = 0.118,\ \alpha(M_Z) = 1/128,\ m_c = 1.51 \text{ GeV},\ m_b = 4.92 \text{ GeV}$$

Solution of DGLAP equations: numerical solution in z-space as implemented in APFEL

extensive benchmark performed up to NNLO [JHEP 1503 (2015) 046]

Parametrisation: each FF is parametrised with a feed-forward neural network (2-5-3-1)

$$D_i^h(Q_0, z) = \text{NN}(x) - \text{NN}(1),\quad Q_0 = 5 \text{ GeV}$$

$h = \pi^+ + \pi^-$
$i = u^+, s^+, c^+, b^+, g$
$D_{u^+}^{\pi^\pm} = D_{d^+}^{\pi^\pm}$ (isospin symmetry)

$h = K^+ + K^-, h = p + \bar{p}$
$i = u^+, d^+ + s^+, c^+, b^+, g$
no further theoretical assumptions

we assume charge conjugation, from which $D_{q^+}^{\pi^+} = D_{q^+}^{\pi^-}$

we enforce positivity by construction, assuming quadratic NNs

initial scale above m_b, but below the lowest c.m. energy of the data, avoid threshold crossing

Heavy flavours: heavy-quark FFs are parametrised independently at the initial scale Q_0

a matched GM-VFNS (like FONLL) may be required if $Q_0 < m_c$ [PRD 94 (2016) 034037]

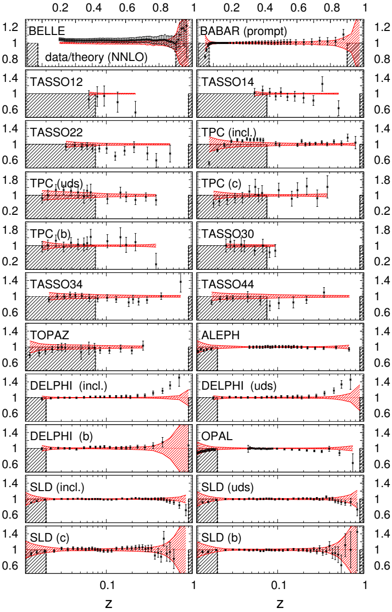
Kinematic cuts: $z \to 0$: contributions $\propto \ln z$; $z \to 1$: contributions $\propto \ln(1-z)$

$$z_{\min} = 0.075\ (\sqrt{s} = M_Z);\ z_{\min} = 0.02\ (\sqrt{s} = M_Z);\ z_{\max} = 0.9$$

kinematic corrections $\propto M_h/(sz^2)$ included exactly in the cross sections [PRD 73 (2006) 054020]

Minimisation: CMA-ES, verfied with closure tests

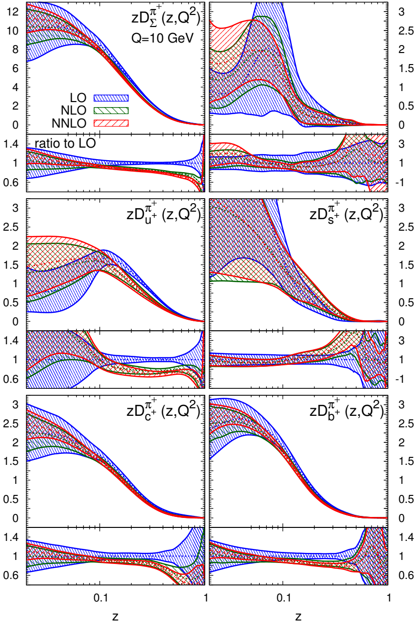
Fit quality: π^+



Exp.	$N_{\rm dat}$	NNLO theory $\chi^2/N_{\rm dat}$	remarks
BELLE	70	0.09	lack of correlations
BABAR	40	0.78	☑
TASSO12	4	0.87	small sample
TASSO14	9	1.70	} data fluctuations
TASSO22	8	1.91	} data fluctuations
TPC	13	0.85	☑
TPC-UDS	6	0.49	☑
TPC-C	6	0.52	☑
TPC-B	6	1.43	☑
TASSO30	–	–	not fitted
TASSO34	9	1.00	☑
TASSO44	6	2.34	data fluctuations
TOPAZ	5	0.80	☑
ALEPH	23	0.78	☑
DELPHI	21	1.86	tension with OPAL
DELPHI-UDS	21	1.54	tension with OPAL
DELPHI-B	21	0.95	☑
OPAL	24	1.84	tension with DELPHI
SLD	34	0.83	☑
SLD-UDS	34	0.52	☑
SLD-C	34	1.06	☑
SLD-B	34	0.36	☑
TOTAL	428	0.87	☑

Overall good description of the dataset
Signs of tension OPAL vs DELPHI (inclusive)
Anomalously small $\chi^2/N_{\rm dat}$ for BELLE

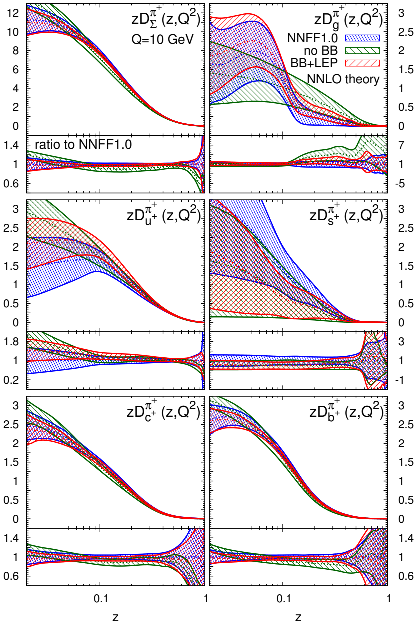
Dependence upon perturbative order: π^+



Exp.	$N_{\rm dat}$	LO $\chi^2/N_{\rm dat}$	NLO $\chi^2/N_{\rm dat}$	NNLO $\chi^2/N_{\rm dat}$
BELLE	70	0.60	0.11	0.09
BABAR	40	1.91	1.77	0.78
TASSO12	4	0.70	0.85	0.87
TASSO14	9	1.55	1.67	1.70
TASSO22	8	1.64	1.91	1.91
TPC	13	0.46	0.65	0.85
TPC-UDS	6	0.78	0.55	0.49
TPC-C	6	0.55	0.53	0.52
TPC-B	6	1.44	1.43	1.43
TASSO30	–	–	–	–
TASSO34	9	1.16	0.98	1.00
TASSO44	6	2.01	2.24	2.34
TOPAZ	5	1.04	0.82	0.80
ALEPH	23	1.68	0.90	0.78
DELPHI	21	1.44	1.79	1.86
DELPHI-UDS	21	1.30	1.48	1.54
DELPHI-B	21	1.21	0.99	0.95
OPAL	24	2.29	1.88	1.84
SLD	34	2.33	1.14	0.83
SLD-UDS	34	0.95	0.65	0.52
SLD-C	34	3.33	1.33	1.06
SLD-B	34	0.45	0.38	0.36
TOTAL	428	1.44	1.02	0.87

Excellent perturbative convergence
FFs almost stable from NLO to NNLO
LO FF uncertainties larger than HO

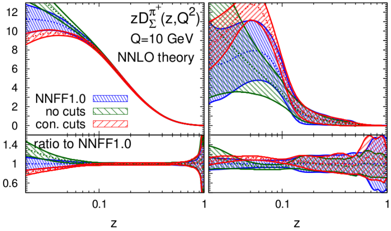
Dependence upon the dataset: π^+



NNLO theory Exp.	$N_{\rm dat}$	NNFF1.0 $\chi^2/N_{\rm dat}$	no BB $\chi^2/N_{\rm dat}$	BB+LEP $\chi^2/N_{\rm dat}$
BELLE	70	0.09	(4.92)	0.09
BABAR	40	0.78	(144)	0.88
TASSO12	4	0.87	0.52	(0.87)
TASSO14	9	1.70	1.38	(1.71)
TASSO22	8	1.91	1.29	(2.15)
TPC	13	0.85	2.12	(0.81)
TPC-UDS	6	0.49	0.54	(0.77)
TPC-C	6	0.52	0.74	(0.58)
TPC-B	6	1.43	1.60	(1.48)
TASSO30	–	–	–	–
TASSO34	9	1.00	1.17	(1.38)
TASSO44	6	2.34	2.52	(2.97)
TOPAZ	5	0.80	0.92	(1.72)
ALEPH	23	0.78	0.57	0.74
DELPHI	21	1.86	1.97	1.82
DELPHI-UDS	21	1.54	1.56	1.42
DELPHI-B	21	0.95	1.01	0.95
OPAL	24	1.84	1.75	1.92
SLD	34	0.83	0.87	0.95
SLD-UDS	34	0.52	0.53	0.63
SLD-C	34	1.06	0.69	0.96
SLD-B	34	0.36	0.49	0.37
TOTAL	428	0.87	1.06	0.82

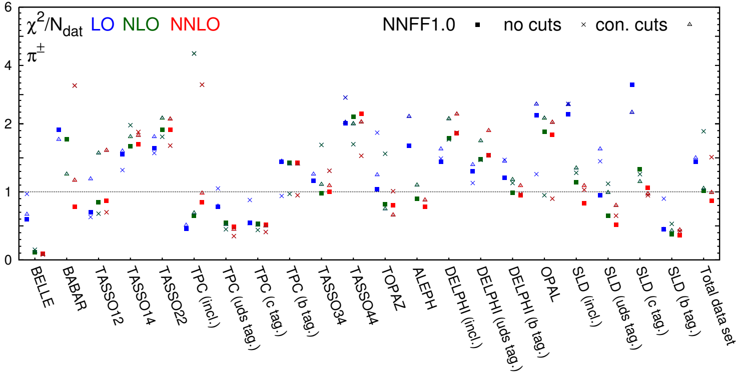
no BB: larger uncertainties; different gluon shape and different light flavour separation
BB+LEP: comparable uncertainties; slightly different size of gluon and light quarks

Dependence upon kinematic cuts: π^+

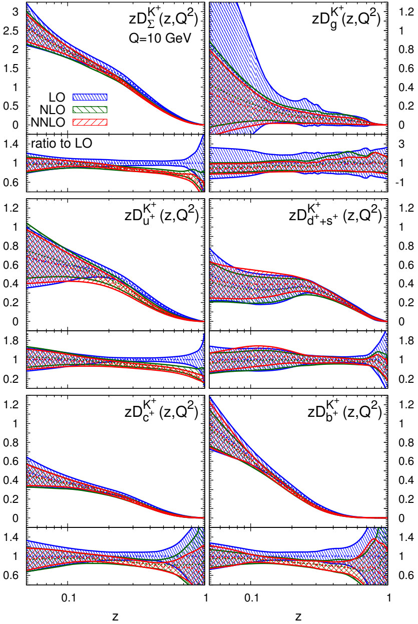


Excellent consistency in the overlapping region
Significantly varied FF shapes at low z
Stability upon inclusion of NNLO corrections

NNFF1.0		no cuts		con. cuts	
$z_{\min}^{(M_Z)}$	$z_{\min}$	$z_{\min}^{(M_Z)}$	$z_{\min}$	$z_{\min}^{(M_Z)}$	$z_{\min}$
0.02	0.075	0.00	0.00	0.05	0.10



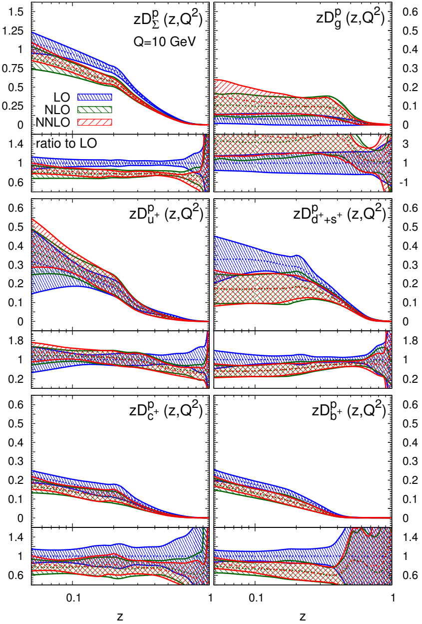
Fit quality vs perturbative order: K^+



Exp.	$N_{\rm dat}$	LO $\chi^2/N_{\rm dat}$	NLO $\chi^2/N_{\rm dat}$	NNLO $\chi^2/N_{\rm dat}$
BELLE	70	0.21	0.32	0.33
BABAR	43	2.86	1.11	0.95
TASSO12	3	1.10	1.03	1.02
TASSO14	9	2.17	2.13	2.07
TASSO22	6	2.14	2.77	2.62
TPC	13	0.94	1.09	1.01
TPC-UDS	–	–	–	–
TPC-C	–	–	–	–
TPC-B	–	–	–	–
TASSO30	–	–	–	–
TASSO34	5	0.27	0.44	0.36
TASSO44	–	–	–	–
TOPAZ	3	0.61	1.19	0.99
ALEPH	18	0.47	0.55	0.56
DELPHI	22	0.28	0.33	0.34
DELPHI-UDS	22	1.38	1.49	1.32
DELPHI-B	22	0.58	0.49	0.52
OPAL	10	1.67	1.57	1.66
SLD	35	0.86	0.62	0.57
SLD-UDS	35	1.31	1.02	0.93
SLD-C	34	0.92	0.47	0.38
SLD-B	35	0.59	0.67	0.62
TOTAL	395	1.02	0.78	0.73

Excellent perturbative convergence
FFs almost stable from NLO to NNLO
LO FF uncertainties larger than HO

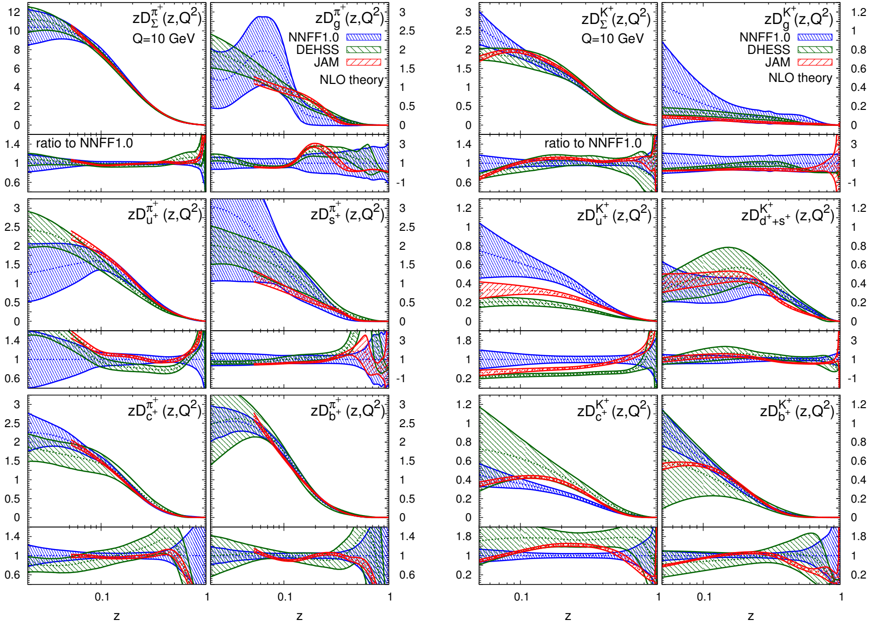
Fit quality vs perturbative order: p



Exp.	$N_{\rm dat}$	LO $\chi^2/N_{\rm dat}$	NLO $\chi^2/N_{\rm dat}$	NNLO $\chi^2/N_{\rm dat}$
BELLE	29	0.10	0.31	0.50
BABAR	43	4.74	3.75	3.25
TASSO12	3	0.69	0.70	0.72
TASSO14	9	1.32	1.25	1.22
TASSO22	9	0.98	0.92	0.93
TPC	20	1.04	1.10	1.08
TPC-UDS	–	–	–	–
TPC-C	–	–	–	–
TPC-B	–	–	–	–
TASSO30	2	0.25	0.19	0.18
TASSO34	6	0.82	0.81	0.78
TASSO44	–	–	–	–
TOPAZ	4	0.79	1.21	1.19
ALEPH	26	1.36	1.43	1.28
DELPHI	22	0.48	0.49	0.49
DELPHI-UDS	22	0.47	0.46	0.45
DELPHI-B	22	0.89	0.89	0.91
OPAL	–	–	–	–
SLD	36	0.66	0.65	0.64
SLD-UDS	36	0.77	0.76	0.78
SLD-C	36	1.22	1.22	1.21
SLD-B	35	1.12	1.29	1.33
TOTAL	360	1.31	1.23	1.17

Excellent perturbative convergence
FFs almost stable from NLO to NNLO
LO FF uncertainties larger than HO

Comparison with other recent FF sets: π^+ and K^+



4. Conclusions

Summary

1. Unpolarised PDFs: full exploitation of the LHC harvest
   - increased precision over an extended kinematic region
   - theoretical improvements (NNLO, resummations, EW corrections . . .)
   - complementarity with the EIC program

2. Polarised PDFs: only a polarised EIC could provide a significant advancement
   - unprecedented kinematic reach
   - check modified evolution at small x
   - meanwhile, make the most from the RHIC spin physics program

3. Fragmentation Functions: significant phenomenological effort
   - improve the methodological sophistication of current analyses
   - (ongoing) make the most from LHC data
   - (ongoing) cross-talk between FFs and PDFs requires simultaneous analyses of both

Summary

1. Unpolarised PDFs: full exploitation of the LHC harvest
   - increased precision over an extended kinematic region
   - theoretical improvements (NNLO, resummations, EW corrections ...)
   - complementarity with the EIC program

2. Polarised PDFs: only a polarised EIC could provide a significant advancement
   - unprecedented kinematic reach
   - check modified evolution at small x
   - meanwhile, make the most from the RHIC spin physics program

3. Fragmentation Functions: significant phenomenological effort
   - improve the methodological sophistication of current analyses
   - (ongoing) make the most from LHC data
   - (ongoing) cross-talk between FFs and PDFs requires simultaneous analyses of both

Thank you