Progress on helicity-dependent Parton Distributions and related subjects from NNPDF

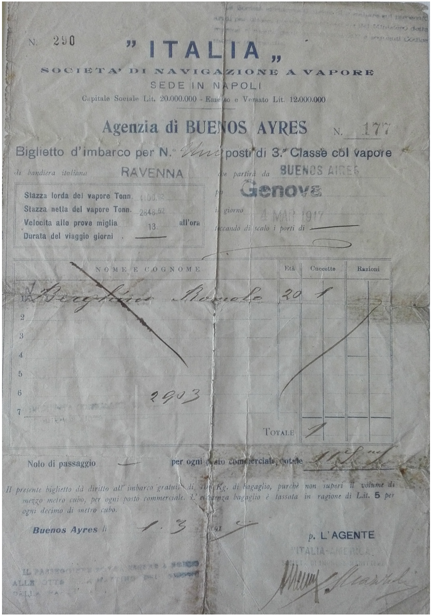
ICAS Institute on the spin of the proton

Emanuele R. Nocera

Rudolf Peierls Centre for Theoretical Physics - University of Oxford

ICAS Buenos Aires – 18th September 2017

Coming back to the origins


N. 290

"ITALIA"

SOCIETA' DI NAVIGAZIONE A VAPORE

SEDE IN NAPOLI

Capitale Sociale Lit. 20.000.000 - Emesso e Versato Lit. 12.000.000

Agenzia di BUENOS AYRES N. 177

Biglietto d'imbarco per N.° Uno posti di 3.ª Classe col vapore

di bandiera italiana RAVENNA che partirà da BUENOS AIRES

per Genova

Stazza lorda del vapore Tonn. 4101.92

Stazza netta del vapore Tonn. 2848.52

Velocità alle prove miglia 13 all'ora

Durata del viaggio giorni

il giorno 4 MAR 1917

toccando di scalo i porti di

	NOME E COGNOME	Età	Cuccette	Razioni
1	Berghino Romolo	20	1	
2				
3				
4				
5				
6	2903			
7				
		TOTALE	1	

Nolo di passaggio per ogni posto commerciale, totale

Il presente biglietto dà diritto all'imbarco gratuito di 100 Kg. di bagaglio, purchè non superi il volume di mezzo metro cubo, per ogni posto commerciale. L'eccedenza bagaglio è tassata in ragione di Lit. 5 per ogni decimo di metro cubo.

Buenos Ayres li 1. 3 191 7

p. L'AGENTE

Outline

1. The NNPDFpol parton sets
   - Recap on the NNPDF methodology
     parametrisation, minimisation, uncertainty representation, reweighting, closure tests
   - NNPDFpol1.0, NNPDFpol1.1 and NNPDFpol1.2
     lessons learnt and impact of new data

---

2. Helicity-dependent PDFs and lattice QCD
   - Suggestions and results from the PDFLattice workshop
     lattice QCD methods, lattice QCD results and their interplay with global PDF fits

---

3. NNFF1.0: NNPDF fragmentation functions
   - Motivation and methods
     new parametrisation and minimisation
   - Results
     perturbative stability, dependence on the data set, comparison with other FF sets

1. The NNPDFpol parton sets

[NPB 874 (2013) 36; PLB 728 (2014) 524; NPB 887 (2014) 276; PLB 742 (2015) 117]

Parametrisation: the NNPDF way

Projection from the infinite-dimensional space of functions
to a finite-dimensional space of parameters

Choose a general, smooth, flexible parametrisation at an initial scale Q_0^2 for each independent parton i (or a combination of them) and hadron h

$$zf_i^h(z,Q_0^2) = A_{f_i^h}\, z^{a_{f_i^h}}\, (1-z)^{b_{f_i^h}}\, \mathscr{F}_i^h(z,\{c_{f_i^h}\})$$

small z

$$zf_i^h(z,Q_0^2) \xrightarrow{z\to 0} z^{a_{f_i^h}}$$

$$\mathscr{F}_i^h(z,\{c_{f_i^h}\}) \xrightarrow[z\to 1]{z\to 0} \text{finite}$$

interpolation in between

large z

$$zf_i^h(x,Q_0^2) \xrightarrow{z\to 1} (1-z)^{b_{f_i^h}}$$

The problem is reduced to the determination of the finite set of parameters $\{c_{f_i^h}\}$

The interpolating function $\mathscr{F}_i^h(z,\{c_{f_i^h}\})$ is chosen to be a neural network, a redundant function with a huge number of parameters ($\mathcal{O}(200)$ per PDF set)

$$\{c_{f_i^h}\} = \{\mathbf{a}\} = \{\omega_{ij}^{(L-1),f_i^h}, \theta_i^{(L),f_i^h}\}$$

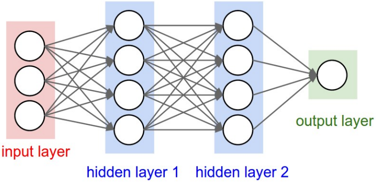
potentially non-smooth; parametrisation bias reduced as much as possible

$z^{a_{f_i^h}}(1-z)^{b_{f_i^h}}$ interpreted as a preprocessing factor

Parametrisation: what a neural network exactly is?

A convenient functional form providing a flexible parametrization
used as a generator of random functions in the FF space

EXAMPLE: MULTY-LAYER FEED-FORWARD PERCEPTRON



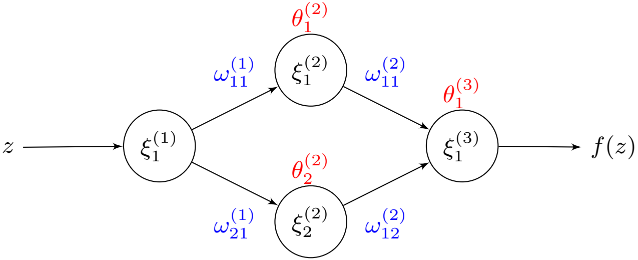
$$\xi_i^{(l)} = g\left(\sum_j^{n_l-1} \omega_{ij}^{(l-1)} \xi_j^{(l-1)} - \theta_i^{(l)}\right)$$

$$g(y) = \frac{1}{1+e^{-y}}$$

- made of neurons grouped into layers (define the architecture)
- each neuron receives input from neurons in the preceding layer (feed-forward NN)
- activation $\xi_i^{(l)}$ determined by a set of parameters (weights and thresholds)
- activation determined according to a non-linear function (except the last layer)

Parametrisation: what a neural network exactly is?

EXAMPLE: THE SIMPLEST 1-2-1 MULTI-LAYER FEED-FORWARD PERCEPTRON



$$f(z) \equiv \xi_1^{(3)} = \left\{ 1 + \exp\left[\theta_1^{(3)} - \frac{\omega_{11}^{(2)}}{1 + e^{\theta_1^{(2)} - x\omega_{11}^{(1)}}} - \frac{\omega_{12}^{(2)}}{1 + e^{\theta_2^{(2)} - x\omega_{21}^{(1)}}} \right] \right\}^{-1}$$

Recall: $$\xi_i^{(l)} = g\left(\sum_j^{n_l - 1} \omega_{ij}^{(l-1)} \xi_j^{(l-1)} - \theta_i^{(l)} \right) \quad ; \quad g(z) = \frac{1}{1 + e^{-z}}$$

Minimisation: goodness of fit

1. Define the fit quality (the χ^2 function)

$$\chi^2 = \sum_{i,j}^{\mathrm{N_{dat}}} (T_i[\{\mathbf{a}\}] - D_i) \left(\mathrm{cov}^{-1}\right)_{ij} (T_j[\{\mathbf{a}\}] - D_j)$$

$$(\mathrm{cov})_{ij} = \delta_{ij} s_i^2 + \left(\sum_{\alpha}^{N_c} \sigma_{i,\alpha}^{(c)} \sigma_{j,\alpha}^{(c)} + \sum_{\alpha}^{N_{\mathcal{L}}} \sigma_{i,\alpha}^{(\mathcal{L})} \sigma_{j,\alpha}^{(\mathcal{L})} \right) D_i D_j$$

s_i are N_{dat} uncorrelated uncertainties (statistic + uncorrealted systematic ucnertainties)
$\sigma_{i,\alpha}^{(c)}$ are $N_{\mathrm{dat}} \times N_c$ additive correlated uncertainties
$\sigma_{i,\alpha}^{(\mathcal{L})}$ are $N_{\mathrm{dat}} \times N_{\mathcal{L}}$ multiplicative uncertainties

2. Find the best-fit configuration of parameters $\{\mathbf{a_0}\}$ which minimise the χ^2

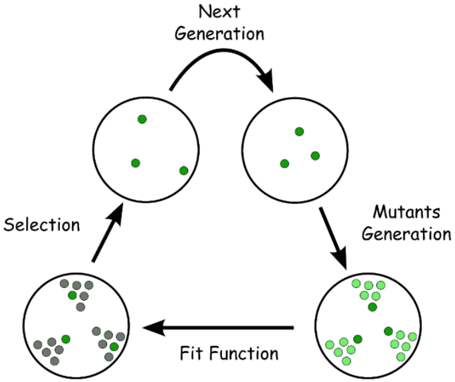
3. Treat conveniently
   - uncorrelated/correlated uncertainties
     need not to overestimate uncertainties and to let the χ^2 be faithful
   - additive/multiplicative uncertainties (the t_0 method [JHEP 1005 (2010) 075])
     need to avoid the D'Agostini bias

$$(\mathrm{cov})_{ij} \to (\mathrm{cov})_{ij}^{t_0} = \delta_{ij} s_i^2 + \left(\sum_{\alpha}^{N_c} \sigma_{i,\alpha}^{(c)} \sigma_{j,\alpha}^{(c)} \right) D_i D_j + \left(\sum_{\alpha}^{N_{\mathcal{L}}} \sigma_{i,\alpha}^{(\mathcal{L})} \sigma_{j,\alpha}^{(\mathcal{L})} \right) T_i^{(0)} T_j^{(0)}$$

Minimisation: genetic algorithm

1. Initial population of NNs
   $\longrightarrow$ pars initialised randomly
2. Mutants generation
   $\longrightarrow$ mutations are introduced
3. Fit function
   $\longrightarrow$ the total χ^2 is computed
4. Selection
   $\longrightarrow$ best pars configs selected
5. Next generation
   $\longrightarrow$ iterate the process

until convergence is achieved



Good exploration of the parameter space

No need to compute gradients

Lower computational efficiency than standard gradient gradient descent

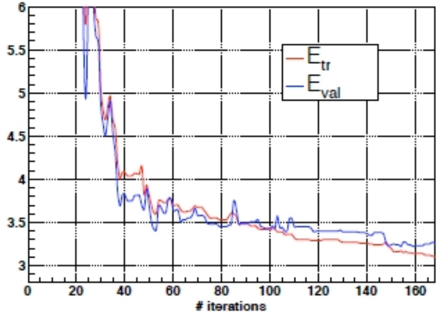
Possible sensitivity to noise in χ^2 driven by noisy data

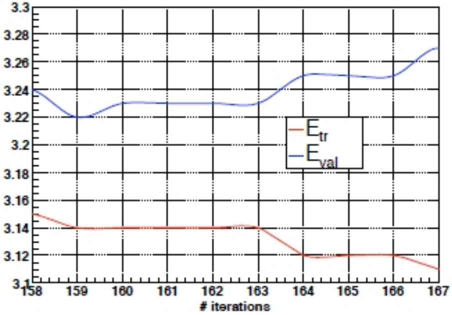
Minimisation: stopping criterion

If the parametrisation is redundant, statistical noise in the data can be learnt

CROSS-VALIDATION METHOD

- divide the data into two subsets (training & validation)
- train the NN on training subset and compute χ^2 for each subset
- stop when the χ^2 of validation subset no longer decreases (NN are learning noise!)





The best fit does not coincide with the χ^2 absolute minimum

Uncertainty representation: the Monte Carlo method

1. Generate (art) replicas of (exp) data according to the distribution

$$\mathcal{O}_i^{(art)(k)} = \mathcal{O}_i^{(exp)} + r_i^{(k)}\sigma_{\mathcal{O}_i}\,, \qquad i=1,\dots N_{\rm dat}\,, \qquad k=1,\dots,N_{\rm rep}$$

   where $r_i^{(k)}$ are (Gaussianly distributed) random numbers for each k-th replica ($r_i^{(k)}$ can be generated with any distribution, not neccesarily Gaussian)

2. Perform a fit for each replica $k=1,\dots,N_{\rm rep}$
3. Compact computation of observables and their uncertainties (PDF replicas are equally probable members of a statistical ensemble)

$$\langle\mathcal{O}[f(x,Q^2)]\rangle = \frac{1}{N_{\sf rep}}\sum_{k=1}^{N_{\sf rep}}\mathcal{O}[f^{(k)}(x,Q^2)]$$

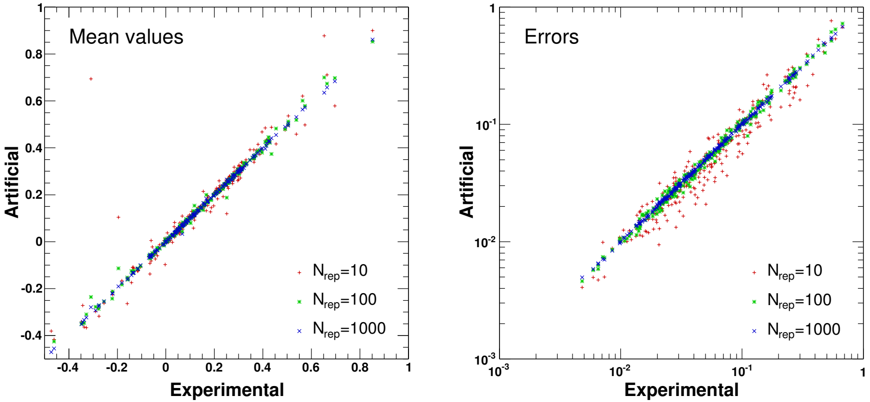
$$\sigma_{\mathcal{O}}[f(x,Q^2)] = \left[\frac{1}{N_{\sf rep}-1}\sum_{k=1}^{N_{\sf rep}}\left(\mathcal{O}[f^{(k)}(x,Q^2)] - \langle\mathcal{O}[f(x,Q^2)]\rangle\right)^2\right]^{1/2}$$

⇒ **no need to rely on linear approximation**

⇒ **computational expensive: need to perform $N_{\rm rep}$ fits instead of one**

The Monte Carlo method: determining the sample size

Require that the average over the replicas reproduces the central value of the original experimental data to a desired accuracy (the standard deviation reproduces the error and so on)



Accuracy of few % requires ~ 100 replicas

Reweighting [PRD 58 (1998) 094023]

Assess the impact of including a new data set $\{y\} = \{y_1, \ldots, y_n\}$ in an old PDF set

Bayesian reweighting [NPB 849 (2011) 112, NPB 855 (2012) 608]

1. Evaluate the agreement between new data and each replica f_k in a prior ensemble

$$\chi^2_k(\{y\},\{f_k\}) = \sum_{i,j}^{n} \{y_i - y_i[f_k]\}\,\sigma_{ij}\,\{y_j - y_j[f_k]\}$$

2. Apply Bayes theorem to determine the conditional probability of PDF upon the inclusion of the new data and update the probability density in the space of PDFs

$$\mathcal{P}_{\rm new} = \mathcal{N}_\chi \mathcal{P}(\chi^2_k|\{f_k\})\mathcal{P}_{\rm old}(\{f_k\}) \quad \mathcal{P}(\chi^2_k|\{f_k\}) = [\chi^2_k(\{y\},\{f_k\}]^{\frac{1}{2}(n-1)} e^{-\frac{1}{2}\chi^2_k(\{y\},\{f_k\})}$$

3. Replicas are no longer equally probable. Expectation values are given by

$$\langle \mathcal{O}[f_i(x,Q^2]\rangle_{\rm new} = \sum_{k=1}^{N_{\rm rep}} w_k \mathcal{O}[f_i^{(k)}(x,Q^2)]$$

$$w_k \propto [\chi^2_k(\{y\},\{f_k\})]^{\frac{1}{2}(n-1)} e^{-\frac{1}{2}\chi^2_k(\{y\},\{f_k\})} \quad \text{with} \quad N_{\rm rep} = \sum_{k=1}^{N_{\rm rep}} w_k$$

Unweighting [NPB 855 (2012) 608]

Unweighting allows for constructing an ensemble of equally probable PDFs
statistically equivalent to a given reweighted set
Hence, the new set can be given without weights

IDEA

Given a weighted set of $N_{\rm rep}$ replicas, select (possibly more than once) replicas carrying relatively hight weight and discard replicas carrying relatively small weight

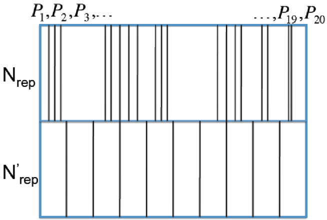
CONSTRUCTION OF THE UNWEIGHTED SET

1. Set the number of replicas $N'_{\rm rep}$ in the unweighted set
(pointless to choose $N'_{\rm rep} > N_{\rm rep}$: no gain of information)
2. Compute, for the k-th replica of the reweighted set, the integer number

$$w'_k = \sum_{j=1}^{N'_{\rm rep}} \theta\left(\frac{j}{N'_{\rm rep}} - P_{k-1}\right)\theta\left(P_k - \frac{j}{N'_{\rm rep}}\right)\,, \quad P_k = \sum_{j=0}^{k}\frac{w_j}{N_{\rm rep}}\,, \quad \sum_{k=1}^{N_{\rm rep}} w'_k = N'_{\rm rep}$$

3. Construct the unweighted set taking w'_k copies of the k-th replica, $k = 1, \ldots, N_{\rm rep}$

Unweighting [NPB 855 (2012) 608]



CONSTRUCTION OF THE UNWEIGHTED SET

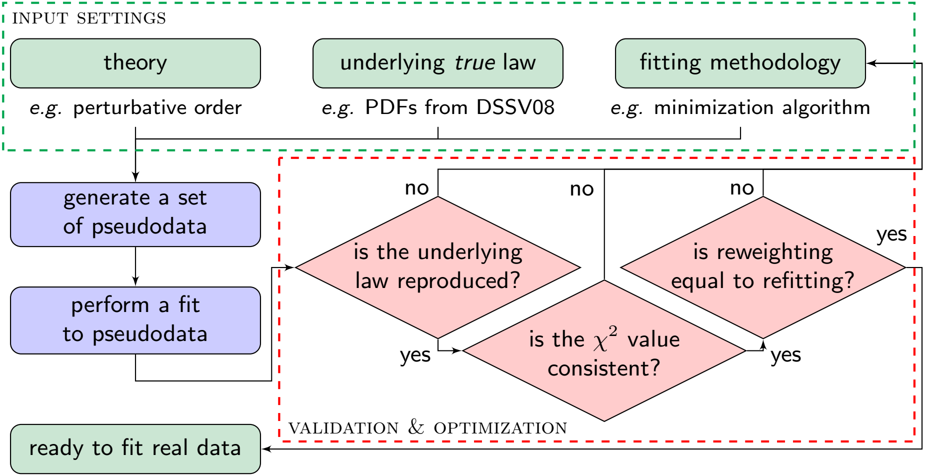
1. Set the number of replicas $N'_{\rm rep}$ in the unweighted set (pointless to choose $N'_{\rm rep} > N_{\rm rep}$: no gain of information)
2. Compute, for the k-th replica of the reweighted set, the integer number

$$w'_k = \sum_{j=1}^{N'_{\rm rep}} \theta\left(\frac{j}{N'_{\rm rep}} - P_{k-1}\right)\theta\left(P_k - \frac{j}{N'_{\rm rep}}\right) \, , \quad P_k = \sum_{j=0}^{k} \frac{w_j}{N_{\rm rep}} \, , \quad \sum_{k=1}^{N_{\rm rep}} w'_k = N'_{\rm rep}$$

3. Construct the unweighted set taking w'_k copies of the k-th replica, $k = 1, \ldots, N_{\rm rep}$

Methodology validation: closure tests [JHEP 1504 (2015) 040]

Validation and optimization of the fitting strategy with known underlying physical law



Full control of procedural uncertainties

Closure tests: levels

1. Level 0: generate pseudodata D_i^0 with zero uncertainty
(but $(\mathrm{cov})_{ij}$ in the χ^2 is the data covariance matrix)
$\rightarrow$ fit quality can be arbitrarily good, if the fitting methodology is efficient: $\chi^2/N_{\mathrm{dat}} \sim 0$
$\rightarrow$ validate fitting methodology (parametrisation, minimisation)
$\rightarrow$ interpolation and extrapolation uncertainty

2. Level 1: generate pseudodata D_i^1 with stochastic fluctuations (no replicas)

$$D_i^1 = (1 + r_i^{\mathrm{nor}} \sigma_i^{\mathrm{nor}}) \left(D_i^0 + \sum_p^{N_{\mathrm{sys}}} r_{i,p}^{\mathrm{sys}} \sigma_{i,p}^{\mathrm{sys}} + r_i^{\mathrm{stat}} \sigma_i^{\mathrm{stat}} \right)$$

$\rightarrow$ experimental uncertainties are not propagated into PDFs: $\chi^2/N_{\mathrm{dat}} \sim 1$
$\rightarrow$ functional uncertainty (a large number of functional forms with equally good χ^2)

3. Level 2: generate N_{rep} Monte Carlo pseudodata replicas $D_i^{2,k}$ on top of Level 2

$$D_i^{2,k} = (1 + r_i^{\mathrm{nor},k} \sigma_i^{\mathrm{nor}}) \left(D_i^1 + \sum_p^{N_{\mathrm{sys}}} r_{i,p}^{\mathrm{sys},k} \sigma_{i,p}^{\mathrm{sys}} + r_i^{\mathrm{stat},k} \sigma_i^{\mathrm{stat}} \right)$$

$\rightarrow$ propagate the fluctuations due to experimental uncertainties into PDFs: $\chi^2/N_{\mathrm{dat}} \sim 1$
$\rightarrow$ input PDFs within the one-sigma band of the fitted PDFs with a probability of $\sim$ 68%
$\rightarrow$ data uncertainty (tolerance supplement data unc. with extrapolation/functional uncs.)

Evolution of NNPDFpol fits

NNPDFpol1.0 [NPB 87 (2013) 36]

- inclusive DIS data from CERN, SLAC and DESY on $g_1^{p,d,n}$

$$g_1(x,Q^2) = \underbrace{\frac{\sum_q^{n_f} e_q^2}{2n_f}\left(\mathcal{C}_{\rm NS}\otimes\Delta q_{\rm NS} + \mathcal{C}_{\rm S}\otimes\Delta\Sigma + 2n_f\mathcal{C}_g\Delta g\right)}_{\text{leading-twist factorization}} + \underbrace{\frac{h^{\rm TMC}}{Q^2} + \frac{h^{\rm HT}}{Q^2} + \mathcal{O}\left(\frac{1}{Q^4}\right)}_{\text{power-suppressed TMCs and HT}}$$

- TMCs included exactly [NP B513 (1998) 301]
- kinematic cut $W^2 \geq 6.25$ GeV^2 to remove sensitivity to dynamical HTs [arXiv:0807.1501]
- inflated uncertainty on a_8 (up to 30% of its exp value) to allow for SU(3) violation
- NLO perturbative accuracy, $\overline{\rm MS}$ renormalization scheme, ZM-VFN scheme

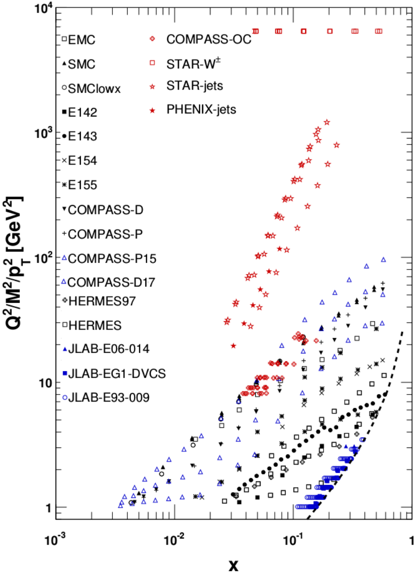
NNPDFpol1.1 [NPB 877 (2014) 276]

- + new collider data from RHIC, included via reweighting:
  - → jet production: STAR [PRD 86 (2012) 032006, PRL 115 (2015) 092002], PHENIX [PRD 84 (2011) 012006]
  - → W-boson production from STAR [PRL 11 (2014) 072301]
- + open-charm production: COMPASS [PRD 87 (2013) 052018], included via reweighting

NNPDFpol1.2 [in preparation]

- + new inclusive DIS data, included via a complete refit:
  - → COMPASS [PLB 753 (2016) 18] (p) [PLB 769 (2017) 34] (d)
  - → JLAB [PLB 641 (2006) 11, PRC 90 (2014) 025212, PLB 744 (2015) 309, arXiv:1505.07877] (p,d)
- + new STAR data data (reweighting): $A_L^{W^\pm}$ [arXiv:1702.02927] and A_{LL}^{2jets} [PRD 95 (2017) 071103]
- new unpolarised fit NNPDF3.1 [arXiv:1706.00428] as baseline + complete code rewriting

Kinematic coverage and fit quality



EXPERIMENT	$N_{\rm dat}$	$\chi^2/N_{\rm dat}$ 1.0	1.1	1.2
EMC	10	0.44	0.43	0.43
SMC	24	0.93	0.90	0.92
SMClowx	16	0.97	0.97	0.94
E142	8	0.67	0.66	0.55
E143	50	0.64	0.67	0.63
E154	11	0.40	0.45	0.34
E155	40	0.89	0.85	0.98
COMPASS-D	15	0.65	0.70	0.57
COMPASS-P	15	1.31	1.38	0.93
HERMES97	8	0.34	0.34	0.23
HERMES	56	0.79	0.82	0.69
COMPASS-P-15	51	0.98*	0.99*	0.65
COMPASS-D-17	15	1.32*	1.32*	0.80
JLAB-E93-009	148	1.26*	1.23*	0.94
JLAB-EG1-DVCS	18	0.45*	0.59*	0.29
JLAB-E06-014	2	2.81*	3.20*	1.33
COMPASS (OC)	45	1.22*	1.22	1.22
STAR (jets)	41	—	1.05	1.06
PHENIX (jets)	6	—	0.24	0.24
STAR-$A_L^{W^\pm}$	24	—	1.05	1.05
STAR-$A_{LL}^{W^\pm}$	12	—	0.95	0.94
STAR-$A_L^{W^\pm}$ (2013)	8	—	2.76*	1.34
STAR (dijets)	14	—	1.34*	1.00
TOTAL		0.77	1.05	1.01

* data set not included in the corresponding fit

From NNPDFpol1.0: SU(2) and SU(3)

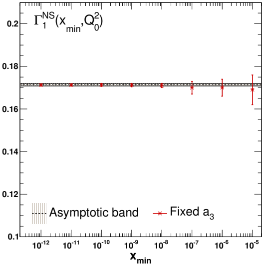
fixed

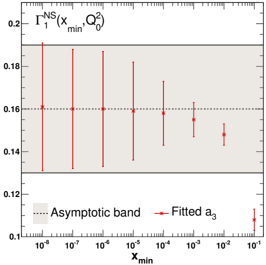
$$a3 = 1.2701 \pm 0.0025$$

fitted

$$a_3 = 1.19 \pm 0.22$$

$$\Gamma_1^{\rm NS}(x_{\rm min}, Q^2) \equiv \int_{x_{\rm min}}^1 dx[g_1^p(x,Q^2) - g_1^n(x,Q^2)] \xrightarrow{x_{\rm min}=0} \frac{1}{6}a_3(Q^2)\Delta C_{\rm NS}[\alpha_s(Q^2)]$$





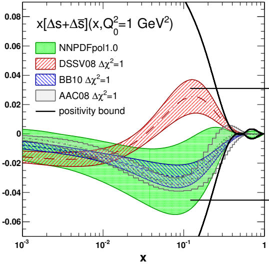
NNPDFpol1.0 [NPB 874 (2013) 36]
$\int_0^1 dx[\Delta s + \Delta\bar{s}] = -0.13 \pm 0.09$

JAM17 [arXiv:1705.05889]
$\int_0^1 dx[\Delta s + \Delta\bar{s}] = -0.03 \pm 0.10$

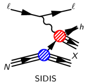
First moment constrained by

$$a_3 = \int_0^1 dx[\Delta u^+ - \Delta d^+] = 1.2701 \pm 0.0025$$

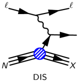
$$a_8 = \int_0^1 dx[\Delta u^+ + \Delta d^+ - 2\Delta s^+] = 0.585 \pm 0.176$$



directly from SIDIS Kaon data



indirectly from DIS + SU(3)

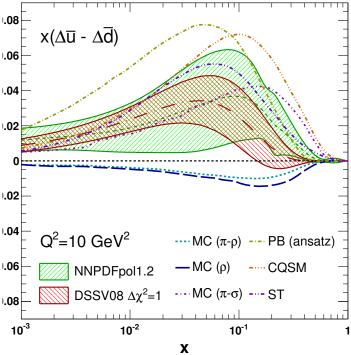


JAM17, first moments fitted: $a_3 = 1.24 \pm 0.04$ $a_8 = 0.46 \pm 0.21$

From NNPDFpol1.1: sea asymmetry and gluon

$W^{\pm}$ boson production

first evidence of broken flavor symmetry for polarized light sea quarks



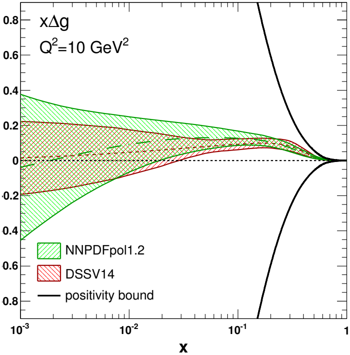
$$\langle x_{1,2}\rangle \simeq \frac{M_W}{\sqrt{s}} e^{-\eta_l/2} \approx [0.04, 0.4]$$

$$\Delta\bar{u} > 0 > \Delta\bar{d},\ |\Delta\bar{d}| > |\Delta\bar{u}|$$

$$\int_{0.04}^{0.4} dx\, \Delta_{sea}(x, Q^2 = 10\,\text{GeV}^2) = +0.06\pm0.03$$

High-p_T jet production

first evidence of a sizable, positive gluon polarization in the proton

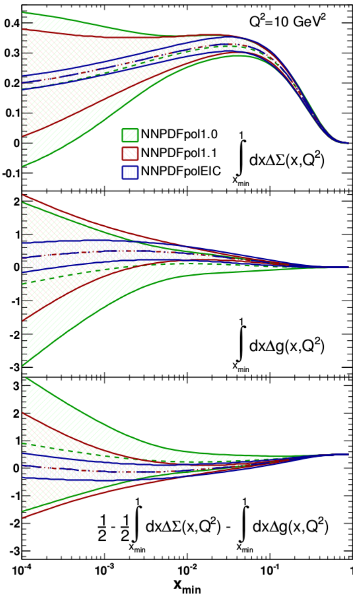


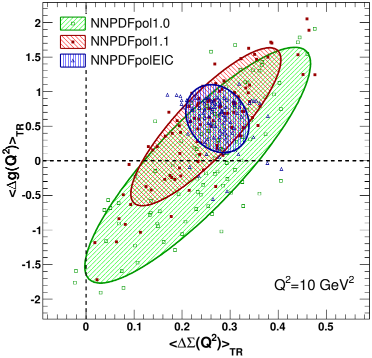
$$\langle x_{1,2}\rangle \simeq \frac{2p_T}{\sqrt{s}} e^{-\eta/2} \approx [0.05, 0.2]$$

NNPDF and DSSV results well compatible

$$\int_{0.05}^{0.2} dx\, \Delta g(x, Q^2 = 10\,\text{GeV}^2) = +0.15\pm0.07$$

The spin content of the proton and NNPDFpolEIC

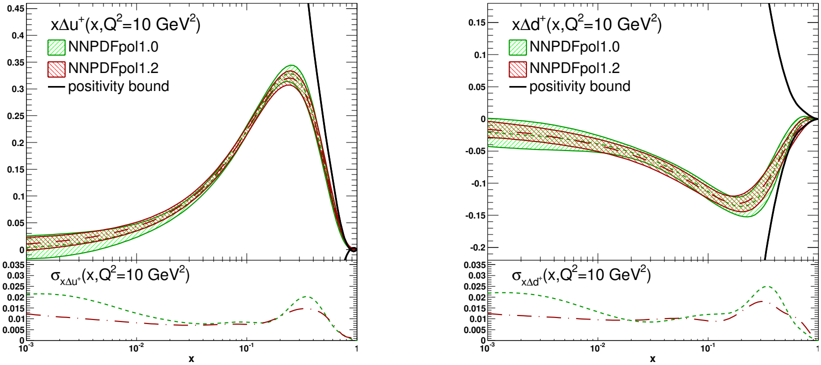




$Q^2 = 10$ GeV^2	$\int_{10^{-3}}^{1} dx\, \Delta\Sigma$	$\int_{10^{-3}}^{1} dx\, \Delta g$
NNPDFpol1.0	$+0.23 \pm 0.15$	-0.06 ± 1.12
NNPDFpol1.1	$+0.25 \pm 0.10$	$+0.49 \pm 0.75$
NNPDFpolEIC	$+0.24 \pm 0.04$	$+0.49 \pm 0.25$

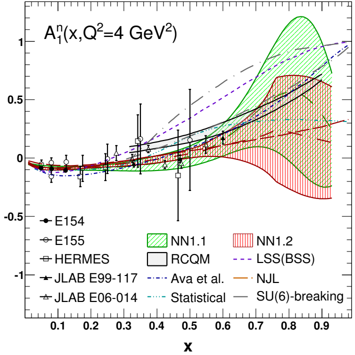
quarks and antiquarks $\sim 20\% - 30\%$
gluons $\sim 50\%$ [PLB 728 (2014) 524] -
$\sim 70\%$ [PRD 92 (2015) 094030]

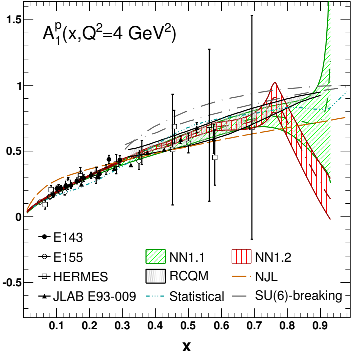
New DIS data in NNPDFpol1.2: total up and down



- Improved accuracy at small x: new COMPASS data
  (+ improved unpolarized F_L and F_2 from NNPDF3.1)
- Improved accuracy at large x: new JLAB data
  (also note that the positivity bound is slightly different)
- A lower cut on W^2 will allow for exploiting the full potential of JLAB data
  (if we replace $W^2 \geq 6.25$ GeV^2 with $W^2 \geq 4.00$ GeV^2 the χ^2 deteriorates significantly)
  (need to include and fit dynamic higher twists, in progress)

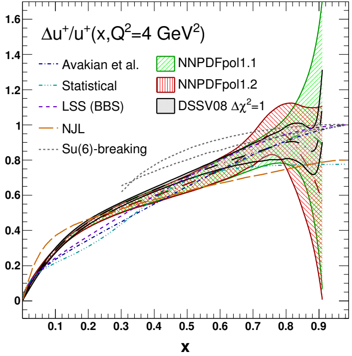
Behavior at large-x values: $A_1^{n,p}$ [PLB 742 (2015) 117]

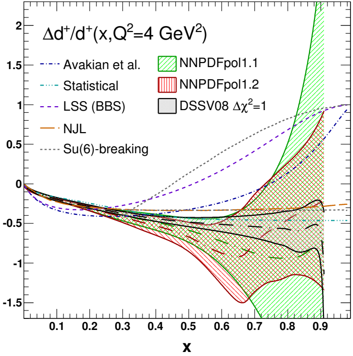




Model	A_1^n	A_1^p	Model	A_1^n	A_1^p
SU(6)	0	$5/9$	NJL	0.35	0.77
RCQM	1	1	DSE (*realistic*)	0.17	0.59
QHD ($\sigma_{1/2}$)	1	1	DSE (*contact*)	0.34	0.88
QHD (ψ_ρ)	1	1	pQCD	1	1
NNPDFpol1.1 ($x = 0.7$)	0.41 ± 0.31	0.75 ± 0.07	NNPDFpol1.1 ($x = 0.9$)	0.36 ± 0.61	0.74 ± 0.34
NNPDFpol1.2 ($x = 0.7$)	0.18 ± 0.26	0.74 ± 0.06	NNPDFpol1.2 ($x = 0.9$)	0.15 ± 0.59	0.24 ± 0.15

Behavior at large-x values: PDF ratios [PLB 742 (2015) 117]



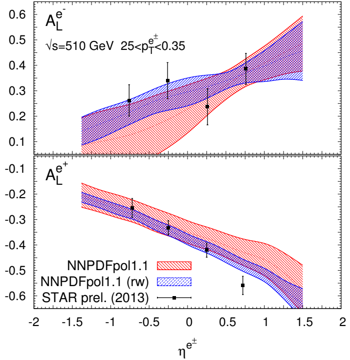


Model	$\Delta u^+/u^+$	$\Delta d^+/d^+$	Model	$\Delta u^+/u^+$	$\Delta d^+/d^+$
SU(6)	$2/3$	$-1/3$	NJL	0.80	-0.25
RCQM	1	$-1/3$	DSE (*realistic*)	0.65	-0.26
QHD ($\sigma_{1/2}$)	1	1	DSE (*contact*)	0.88	-0.33
QHD (ψ_ρ)	1	$-1/3$	pQCD	1	1
NNPDFpol1.1 ($x = 0.7$)	0.82 ± 0.08	-0.88 ± 0.68	NNPDFpol1.1 ($x = 0.9$)	0.91 ± 0.65	-0.74 ± 3.57
NNPDFpol1.2 ($x = 0.7$)	0.86 ± 0.08	-0.75 ± 0.62	NNPDFpol1.2 ($x = 0.9$)	0.62 ± 0.48	-0.23 ± 1.06

New pp data in NNPDFpol1.2: $\Delta\bar{u} - \Delta\bar{d}$ and Δg [arXiv:1702.05077]

$W^{\pm}$ boson production [arXiv:1702.02927]

confirm evidence of broken flavor symmetry for polarized light sea quarks



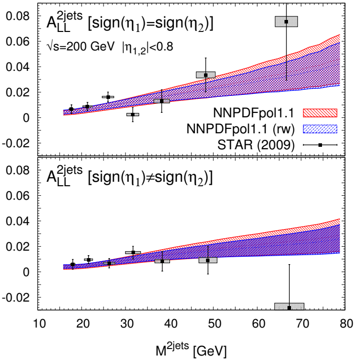
$$\langle x_{1,2} \rangle \simeq \frac{M_W}{\sqrt{s}} e^{-\eta_l/2} \approx [0.04, 0.4]$$

$$\Delta\bar{u} > 0 > \Delta\bar{d},\ |\Delta\bar{d}| > |\Delta\bar{u}|$$

$$\int_{0.04}^{0.4} dx\, \Delta_{sea}(x, Q^2 = 10\,\text{GeV}^2) = +0.07 \pm 0.01$$

High-p_T di-jets [PRD 95 (2017) 071103]

confirm a positive gluon polarization in the proton



$$\langle x_{1,2} \rangle \simeq \frac{p_T}{\sqrt{s}} (e^{\pm\eta_3 \pm \eta_4}) \approx [0.01, 0.2]$$

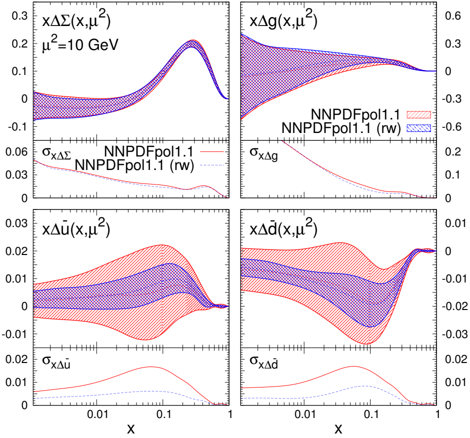
x sensitivity extended down to $x \sim 0.01$

$$\int_{0.01}^{0.2} dx\, \Delta g(x, Q^2 = 10\,\text{GeV}^2) = +0.32 \pm 0.13$$

New pp data in NNPDFpol1.2: $\Delta\bar{u} - \Delta\bar{d}$ and Δg [arXiv:1702.05077]

Data set	$\mathcal{L}$ [pb^{-1}]	$\sqrt{s}$ [GeV]	$\mathcal{A}$
STAR09-2j-ss	21	200	$A_{LL}^{\rm 2jets}$
STAR09-2j-os	21	200	$A_{LL}^{\rm 2jets}$

Data set	$\mathcal{L}$ [pb^{-1}]	$\sqrt{s}$ [GeV]	$\mathcal{A}$
STAR13-W^-	246.2	510	$A_L^{e^-}$
STAR13-W^+	246.2	510	$A_L^{e^+}$



Data set	$N_{\rm dat}$	$\chi^2/N_{\rm dat}$	$\chi^2_{\rm rw}/N_{\rm dat}$
STAR09-2j-ss	7	1.41	1.18
STAR09-2j-os	7	1.26	0.83
STAR13-W^-	4	2.44	0.69
STAR13-W^+	4	3.08	1.30

$$\langle \Delta f(Q^2)\rangle^{[x_{\rm min},x_{\rm max}]} \equiv \int_{x_{\rm min}}^{x_{\rm max}} dx\,\Delta f(x,Q^2)$$

$$\langle \Delta g(Q^2)\rangle^{[0.01,0.2]} = 0.23 \pm 0.15$$

$$\langle \Delta g(Q^2)\rangle^{[0.01,0.2]}_{\rm rw} = 0.32 \pm 0.13$$

$$Q^2 = 10 \text{ GeV}^2$$

2. Helicity-dependent PDFs and lattice QCD

[arXiv:1709.01511; H.-W. Lin, E.R. Nocera, F. Olness, K. Orginos, J. Rojo *et al.*, in preparation]

The PDFLattice2017 workshop

Parton Distributions and Lattice Calculations in the LHC era
22 -24 March 2017, Balliol College, Oxford, UK

http://www.physics.ox.ac.uk/confs/PDFlattice2017/index.asp

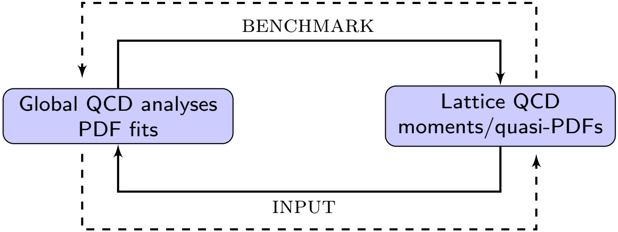
A first joint workshop between global-fitting and lattice PDF communities

Huey-Wen Lin (MSU), Emanuele R. Nocera (Oxford), Fred Olness (SMU),
Kostas Orginos (W&M & JLab), Juan Rojo (VU Amsterdam and NIKHEF)

"The goal of this workshop is to **bring together the global PDF analysis and lattice-QCD PDF communities** *in order to explore ways to improve current PDF determinations. In particular, we plan to* **set precision goals for lattice-QCD calculations** *so that these calculations, together with the experimental input, can achieve improved determinations of PDFs. We discuss what impact such improved determinations of PDFs will have on future new-physics searches."*

Bridging two families together

LHC era broadly denotes high-energy precision physics,
not bound only to the LHC physics program, but including also RHIC, JLAB, ...



LHC (precision physics)	RHIC, JLab, ... (hadron physics)
Higgs boson characterisation	Spin physics, nucleon structure
Precision SM measurements (*e.g.* M_W)	Large-x behaviour
BSM searches, SUSY	Nuclear modifications

GLOBAL QCD FIT AND LATTICE QCD INTERPLAY IN PDF DETERMINATIONS

Conscious focus on collinear unpolarised and longitudinally polarised PDFs only

Lattice QCD methods

Mellin moments of PDFs [see arXiv:1709.01511 for a review]

Determine the matrix elements of local twist-two operators that can be related to the Mellin moments of PDFs

$$4\int_0^1 dx\, x^n g_1(x,Q^2) = \sum_a^{n_f} C_{1,a}^n(\mu^2)\, a_a^n(\mu^2)|_{\mu^2=Q^2} = \sum_a^{n_f} C_{1,a}^n(\mu^2) \int_0^1 dx\, x^n\, 2\Delta f_a(x,Q^2)$$

$$\langle p,s|\mathcal{O}^{5\,i}_{\{\sigma\mu_1,\cdots,\mu_n\}}|p,s\rangle = \frac{1}{n+1} a_i^n\,[s_\sigma p_{\mu_1}\cdots p_{\mu_n} - \text{traces}] \qquad C_{1,i}^n(\mu^2) = \int_0^1 dy\, y^n c_{1,i}(y,\mu^2)$$

Matrix element	Operator	PDF moment
a_q^0	$\bar{q}(x)\gamma_\sigma\gamma_5 q(x)$	$2\,\langle 1\rangle_{\Delta q}$
a_q^1	$(\mathrm{i}/2)\,\bar{q}(x)\gamma_\sigma\gamma_5 \overleftrightarrow{D}_{\mu_1} q(x)$	$2\,\langle x\rangle_{\Delta q}$

Moving beyond the lowest three moments difficult because of power-divergent mixing for lattice QCD twist-two operators

Inversion method (OPE without OPE) [PRD 90 (2014) 014510; PRD 92 (2015) 114517; PRL 118 (2017) 242001]

Relate g_1 to the appropriate Compton amplitude through an integral equation

Compute the Compton amplitude in lattice QCD and solve the equation numerically

Path-integral formulation of the DIS hadronic tensor [PRL 72 (1994) 1790; PRD 62 (2000) 074501]

Carry out the OPE on the Euclidean lattice by evaluation of four-point functions

Convert the hadronic tensor from the Euclidean to the Minkowski space

Lattice QCD methods

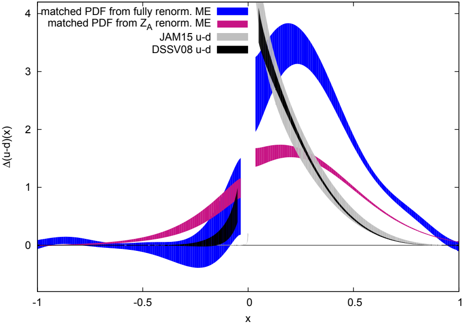
Quasi-PDFs [PRL 110 (2013) 262002]

Defined as momentum-dependent nonlocal static matrix elements for nucleon states at finite momentum, with an ultraviolet cut-off scale $\Lambda \sim 1/a$

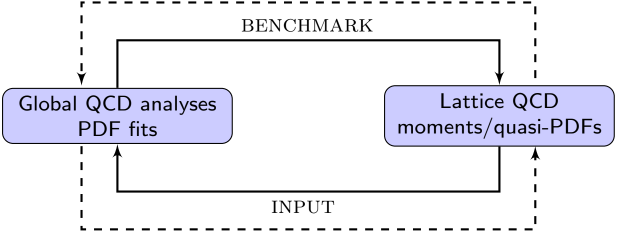
$$\widetilde{q}(x,\Lambda,p_z) = \int \frac{dz}{4\pi} e^{-ixzp_z} \frac{1}{2}\sum_{s=1}^{2} \langle p,s|\, \bar{\psi}(z)\gamma_\alpha e^{ig\int_0^z A_z(z')dz'} \psi(0)\, |p,s\rangle$$

Must be related to the corresponding light-front PDF, usually within LaMET

$$\widetilde{q}(x,\Lambda,p_z) = \int_{-1}^{1} \frac{dy}{|y|} Z\left(\frac{x}{y},\frac{\mu}{p_z},\frac{\Lambda}{p_z}\right)_{\mu^2=Q^2} q(y,Q^2) + \mathcal{O}\left(\frac{\Lambda^2_{\rm QCD}}{p_z^2},\frac{m^2}{p_z^2}\right)$$



List of desiderata



Define a mutually agreed conventional notation
for relevant PDF-related quantities, such as PDF moments.

Assess the sources of systematic uncertainties in lattice-QCD calculations.

Identify a best-set of quantities
to benchmark lattice-QCD calculations against global-fit determinations.

Set precision targets for lattice-QCD calculations
with respect to global-fit determinations.

Assess the impact of lattice-QCD calculations
on global-fit determinations within their current/projected precision.

DESIDERATA

Appraising lattice QCD calculations

Mom.	Collab.	Ref.	N_f	discretisation	quark mass	finite volume	renormalisation	excited states		Value
g_A	CalLat 17	[arXiv:1704.01114]	2+1+1	■	★	■	★	★	◇	1.278(21)(26)
	PNDME 16	[PRD 94 (2016) 054508]	2+1+1	○	★	○	★	★		1.195(33)(20)
	LHPC 14	[PLB 734 (2014) 290]	2+1	■	★	★	★	★		0.97(8)
	Mainz 17	[arXiv:1705.06186]	2	★	○	★	★	★		$1.278(68)(^{+0}_{-0.087})$
	ETMC 17	[arXiv:1705.03399]	2	■	★	■	★	★	*	1.212(33)(22)
	RQCD 15	[PRD 91 (2015) 054501]	2	○	○	○	★	○	‡	1.280(44)(46)
	QCDSF 14	[PLB 732 (2014) 41]	2	○	○	○	★	■	‡	1.29(5)(3)
$\langle 1\rangle_{\Delta u^+}$	ETMC 17	[arXiv:1706.02973]	2	■	★	■	★	★	*	$0.830(26)(4)$
$\langle 1\rangle_{\Delta d^+}$	ETMC 17	[arXiv:1706.02973]	2	■	★	■	★	★	*	$-0.386(16)(6)$
$\langle 1\rangle_{\Delta s^+}$	χQCD 17	[PRD 95 (2017) 114509]	2+1	■	○	○	★	★	†,◁	-0.0403(44)(78)
	Engelhardt 12	[PRD 86 (2012) 114510]	2+1	■	■	○	★	★	◁	-0.031(17)
	ETMC 17	[arXiv:1706.02973]	2	■	★	■	★	★	*	-0.042(10)(2)
$\langle x\rangle_{\Delta u^- - \Delta d^-}$	RBC/ UKQCD 10	[PRD 82 (2010) 014501]	2+1	■	■	★	★	■		0.256(23)/ 0.205(59)
	LHPC 10	[PRD 82 (2010) 094502]	2+1	■	■	○	○	■		0.1972(55)
	ETMC 15	[PRD 92 (2015) 114513]	2	■	★	■	★	★	*	0.229(33)

* Study employing a single physical pion mass ensemble.

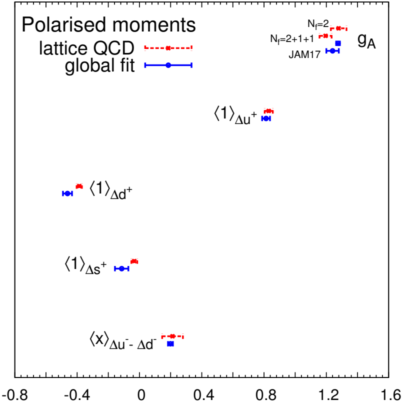
‡ g_A is determined via the ratio g_A/f_π employing the physical value for f_π.

◇ Approach inspired by the Feynman-Hellmann method is employed.

† Partially quenched simulation with $m_\pi = 330$ MeV.

◁ Some parts of the renormalisation are estimated.

Comparing lattice QCD and global fit PDF moments



Moment	Lattice QCD	Global Fit
g_A	1.195(39)*	1.275(12)
	1.279(50)**	1.24(4)‡
$\langle 1\rangle_{\Delta u^+}$	0.830(26)†	0.813(25)
$\langle 1\rangle_{\Delta d^+}$	-0.386(17)†	-0.462(29)
$\langle 1\rangle_{\Delta s^+}$	-0.052−-0.014	-0.114(43)
$\langle x\rangle_{\Delta u^- - \Delta d^-}$	0.146–0.279	0.199(16)

* $N_f = 2$.
** $N_f = 2+1+1$.
‡ JAM17
† Single lattice result available.
$\Delta q^{\pm} + \Delta q \pm \Delta\bar{q}$, $q = u, d, s$; $Q = 2$ GeV.

$$g_A = \langle 1\rangle_{\Delta u^+ - \Delta d^+} = \int_0^1 dx\ \left[\Delta u^+(x,Q^2) - \Delta d^+(x,Q^2)\right]$$

$$\langle 1\rangle_{\Delta q^+} = \int_0^1 dx\, \Delta q^+(x,Q^2)$$

$$\langle x\rangle_{\Delta u^- - \Delta d^-} = \int_0^1 x dx\ \left[\Delta u^-(x,Q^2) - \Delta d^-(x,Q^2)\right]$$

Which precision shall we require to lattice QCD?

Genrate lattice QCD pseudodata assuming NNPDFpol1.1 central values for

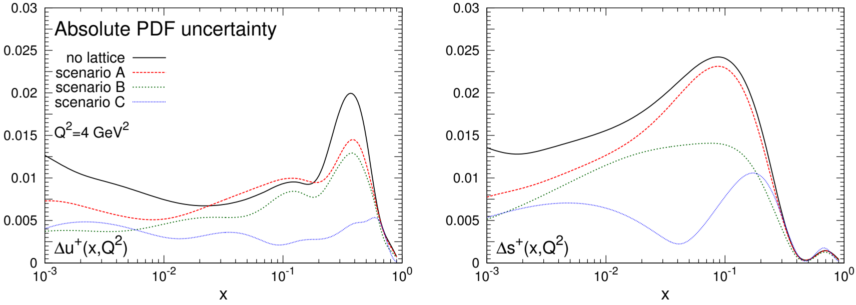
$$g_A \equiv \langle 1 \rangle_{\Delta u^+ - \Delta d^+},\ \langle 1 \rangle_{\Delta u^+},\ \langle 1 \rangle_{\Delta d^+},\ \langle 1 \rangle_{\Delta s^+},\ \langle x \rangle_{\Delta u^- - \Delta d^-}$$

Assume percentage uncertainties according to three scenarios

scenario	g_A	$\langle 1 \rangle_{\Delta u^+}$	$\langle 1 \rangle_{\Delta d^+}$	$\langle 1 \rangle_{\Delta s^+}$	$\langle x \rangle_{\Delta u^- - \Delta d^-}$
A	4%	3%	4%	60%	30%
B	2%	2%	2%	30%	15%
C	1%	1%	1%	10%	10%

Reweight NNPDFpol1.1 with these lattice pseudodata

Assess the impact on the uncertainty of the reweighted PDFs

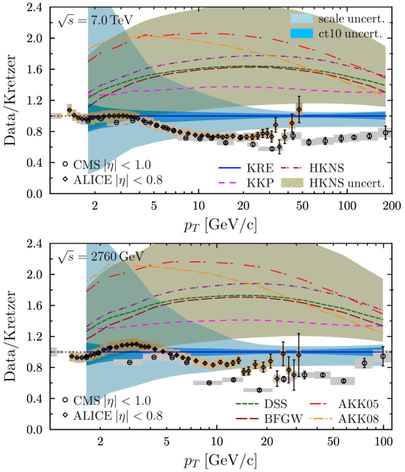


3. NNFF1.0: NNPDF Fragmentation Functions

[EPJC 77 (2017) 516]

Fragmentation functions: why should we bother?

Example 1: Ratio of the inclusive charged-hadron spectra measured by CMS and ALICE



Figures taken from [NPB 883 (2014) 615]

Example 2: The strange polarised parton distribution at $Q^2 = 2.5$ GeV^2 ($\Delta s = \Delta \bar{s}$)

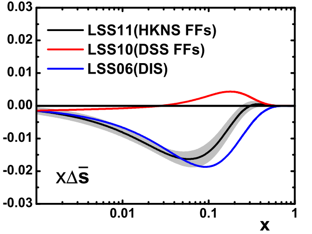
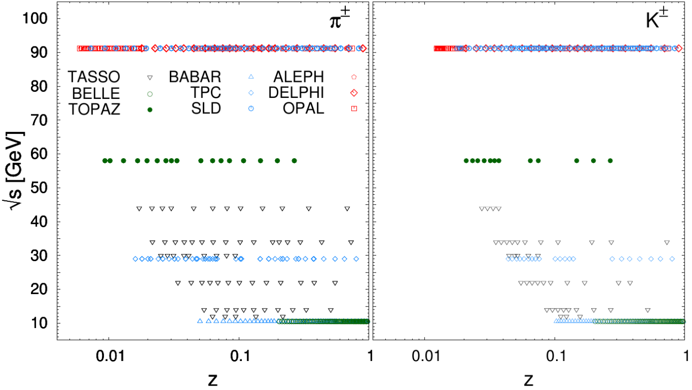


Figure taken from [PRD D84 (2011) 014002]

1 Predictions from all available FF sets are not compatible with CMS and ALICE data, not even within scale and PDF/FF uncertainties
$\longrightarrow$ input for nuclear medium modifications

2 If SIDIS data are used to determine Δs, $K^{\pm}$ FFs for different sets lead to different results. Such results may differ significantly among them and w.r.t. the results obtained from DIS
$\longrightarrow$ input for polarised PDFs and TMDs

The NNFF1.0 data set



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]

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

From observables to fragmentation functions

$$\boxed{\mathcal{F}_2^h = \langle e^2 \rangle \left\{ C_{2,q}^{\rm S} \otimes D_\Sigma^h + n_f \mathcal{C}_{2,g}^{\rm S} \otimes D_g^h + \mathcal{C}_{2,q}^{\rm NS} \otimes D_{\rm NS}^h \right\}}$$

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

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]

$$\begin{aligned} F_2^{h,n_f=5} =& \frac{1}{5} \left[\left(2\hat{e}_u^2 + 3\hat{e}_d^2\right) C_{2,q}^{\rm S} + 3\left(\hat{e}_u^2 - \hat{e}_d^2\right) C_{2,q}^{\rm NS} \right] \otimes \left(D_{u^+}^h + D_{c^+}^h \right) \\ &+ \frac{1}{5} \left[\left(2\hat{e}_u^2 + 3\hat{e}_d^2\right) C_{2,q}^{\rm S} - 2\left(\hat{e}_u^2 - \hat{e}_d^2\right) C_{2,q}^{\rm NS} \right] \otimes \left(D_{d^+}^h + D_{s^+}^h + D_{b^+}^h \right) \\ &+ \left(2\hat{e}_u^2 + 3\hat{e}_d^2\right) C_{2,g}^{\rm S} \otimes D_g^h \end{aligned}$$

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 $C_{2,g}^{\rm S}$ is $\mathcal{O}(\alpha_s^2)$, and tenous

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 [arXiv:1706.00428]

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

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)

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

PIONS	KAONS
$h = \pi^+ + \pi^-,\ i = u^+, d^+ + s^+, c^+, b^+, g$	$h = K^+ + K^-,\ i = u^+, d^+ + s^+, c^+, b^+, g$

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

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

Hadron mass corrections: included exactly à la Albino-Kniehl-Kramer [NPB 803 (2008) 42]

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

PIONS	KAONS
$z_{\min} = 0.1,\ z_{\min} = 0.05\ (\sqrt{s} = M_Z);\ z_{\max} = 0.90$	$z_{\min} = 0.2,\ z_{\min} = 0.1\ (\sqrt{s} = M_Z);\ z_{\max} = 0.90$
$z_{\min} = 0.075,\ z_{\min} = 0.01\ (\sqrt{s} = M_Z);\ z_{\max} = 0.90$	$z_{\min} = 0.1,\ z_{\min} = 0.05\ (\sqrt{s} = M_Z);\ z_{\max} = 0.90$

Minimisation: adaptive algorithms [N. Hansen, Springer (2016)]

The Covariance Matrix Adaption - Evolution Strategy (CMA-ES)

1. Initialisation at the (0)-th generation

$$\mathbf{a}^{(0)} \sim \mathcal{N}(0, \mathbf{C}^{(0)})\,, \qquad \mathbf{C}^{(0)} = \mathbf{I}$$

2. Mutation at the (i)-th generation, λ mutants, step-size $\sigma^{(i-1)}$

$$\mathbf{x}_k^{(i)} \sim \mathbf{a}^{(i-1)} + \sigma^{(i-1)}\mathcal{N}(0, \mathbf{C}^{(i-1)})\,, \qquad \text{for } k = 1, \dots, \lambda$$

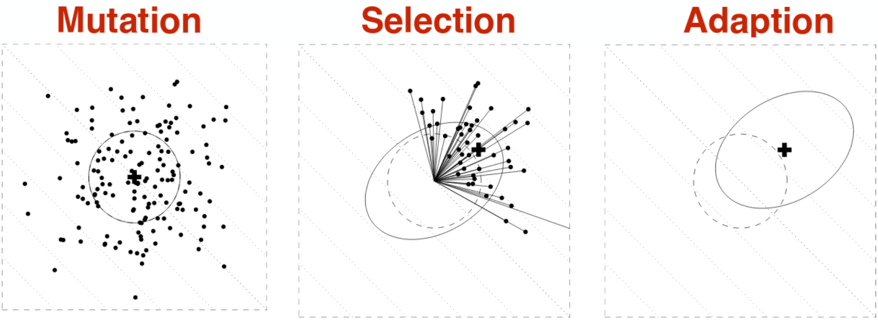
compute the fitness of each mutant and rank them such that $\chi^2(\mathbf{x}_k) < \chi^2(\mathbf{x}_{k+1})$

3. (Non-elitist) recombination
compute the new search centre as a weighted average over the $\mu = \lambda/2$ best mutants

$$\mathbf{a}^{(i)} = \mathbf{a}^{(i-1)} + \sum_{i=1}^{\mu} w_i \left(\mathbf{x}_k^{(i)} - \mathbf{a}^{(i-1)} \right)$$

update $\mathbf{C}$ using information on the parameter space learnt from the mutants
iterate until convergence is reached

Minimisation: the CMA-ES algorithms [N. Hansen, Springer (2016)]

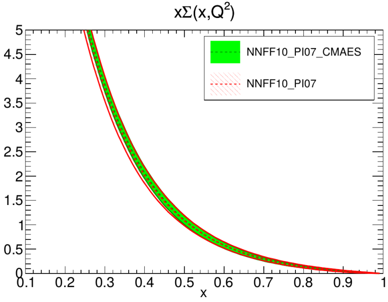


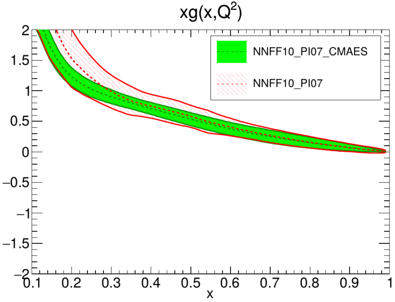
The key features of the CMA-ES family of algorithms are
the determination of the search distribution covariance matrix $\mathbf{C}^{(i)}$ (and possibly of the step-size σ^i)

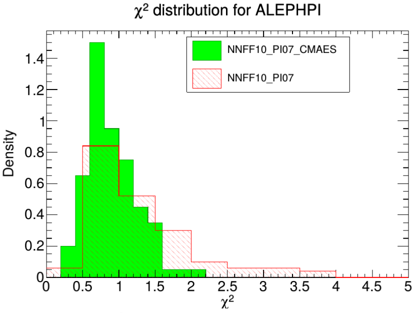
These features are optimised by the fit procedure,
making use of the information present in the ensemble of mutants
to learn preferred directions in parameter space

Internal parameters ($\sigma^{(0)}$, λ, w_i) tuned by trial and error

Minimisation: dependence on the algorithm



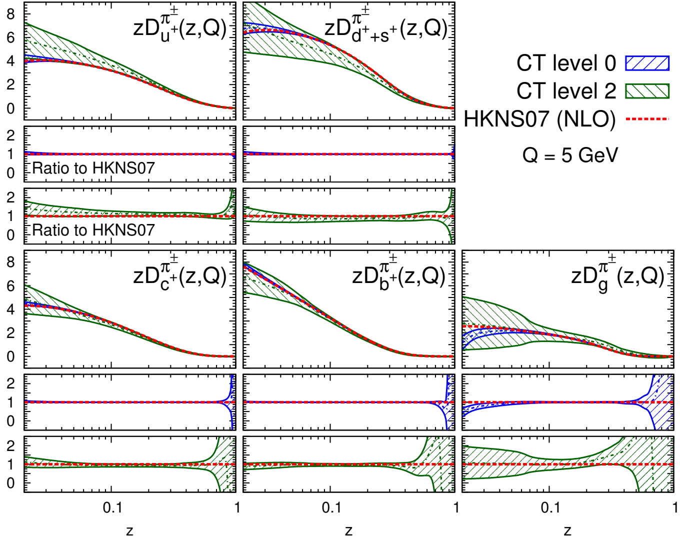




Singlet and Gluon FFs at $Q = 5$ GeV

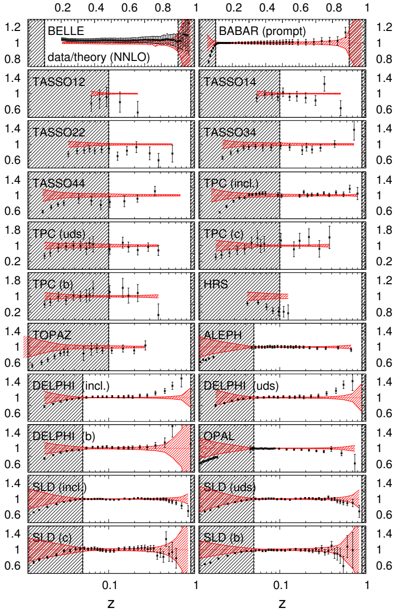
	NNFF10_PI07_CMAES	NNFF10_PI07
$\chi^2_{\rm tot}/N_{\rm dat}$	0.93	0.92
$\langle E_{\rm tr}\rangle \pm \sigma_{\rm tr}$	1.97±0.28	1.97±0.71
$\langle E_{\rm val}\rangle \pm \sigma_{\rm val}$	2.16±0.37	2.91±1.72
$\langle {\rm TL}\rangle \pm \sigma_{\rm TL}$	3065±1673	5560±9394

Closure testing NNFF1.0



$\chi^2/N_{\rm dat} = 0.0001$ (L0) $\qquad \chi^2/N_{\rm dat} = 1.0262$ (L1)

Fit quality: π^+



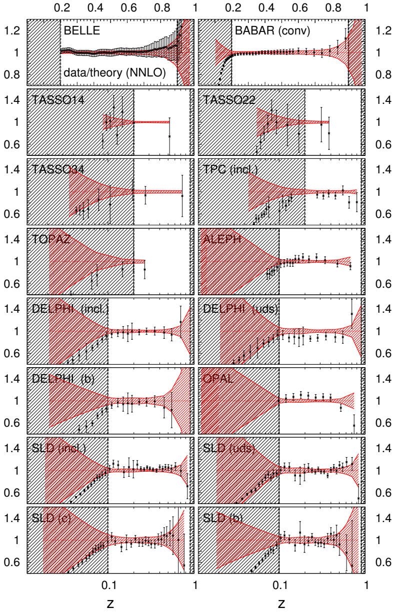
Exp.	$N_{\rm dat}$	NNLO theory $\chi^2/N_{\rm dat}$	remarks
BELLE	70	0.08	lack of correlations
BABAR	37	1.17	☑
TASSO12	2	1.61	small sample
TASSO14	7	1.83	} data fluctuations
TASSO22	7	2.16	} data fluctuations
TASSO34	8	1.09	☑
TASSO44	5	1.95	data fluctuations
TPC	12	0.98	☑
TPC-UDS	6	0.45	☑
TPC-C	6	0.50	☑
TPC-B	6	1.41	☑
TOPAZ	4	0.66	☑
ALEPH	22	0.88	☑
DELPHI	16	2.32	tension with OPAL
DELPHI-UDS	16	1.90	tension with OPAL
DELPHI-B	16	1.09	☑
OPAL	22	2.05	tension with DELPHI
SLD	29	1.09	☑
SLD-UDS	29	0.80	☑
SLD-C	29	0.97	☑
SLD-B	29	0.44	☑
TOTAL	378	0.99	☑

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

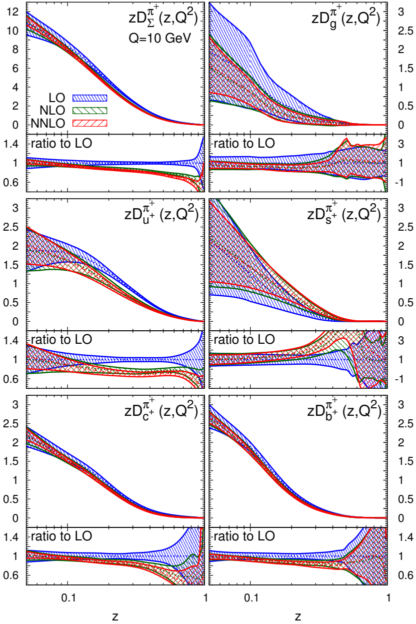
Fit quality: K^+

Exp.	$N_{\rm dat}$	NNLO theory $\chi^2/N_{\rm dat}$	remarks
BELLE	70	0.19	lack of correlations
BABAR	28	0.77	☑
TASSO14	3	1.30	small sample
TASSO22	2	0.29	small sample
TASSO34	2	0.09	small sample
TPC	7	1.19	☑
ALEPH	13	0.72	☑
DELPHI	11	0.17	☑
DELPHI-UDS	11	1.97	☑
DELPHI-B	11	0.41	☑
OPAL	9	2.10	tension with other M_Z data
SLD	21	0.77	☑
SLD-UDS	21	1.11	☑
SLD-C	20	0.42	☑
SLD-B	21	0.71	☑
TOTAL	250	0.67	☑

Overall good description of the dataset
Excellent BELLE/BABAR consistency
Signs of tension OPAL vs DELPHI (inclusive)
Anomalously small $\chi^2/N_{\rm dat}$ for BELLE
Data description rapidly deteriorates at low z
Prediction uncertainties blow up at low z



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	0.67	0.12	0.08
BABAR	37	1.64	1.26	1.17
TASSO12	2	1.19	1.57	1.61
TASSO14	7	1.60	1.81	1.83
TASSO22	7	1.81	2.19	2.16
TASSO34	8	1.26	1.11	1.09
TASSO44	5	2.05	2.00	1.95
TPC	12	0.51	0.69	0.98
TPC-UDS	6	0.79	0.52	0.45
TPC-C	6	0.54	0.51	0.50
TPC-B	6	1.45	1.41	1.41
TOPAZ	4	1.25	0.75	0.66
ALEPH	22	2.25	1.10	0.88
DELPHI	16	1.63	2.17	2.32
DELPHI-UDS	16	1.40	1.75	1.90
DELPHI-B	16	1.45	1.18	1.09
OPAL	22	2.68	2.19	2.05
SLD	29	2.66	1.35	1.09
SLD-UDS	29	1.63	0.98	0.80
SLD-C	29	2.39	1.15	0.97
SLD-B	29	0.45	0.43	0.44
TOTAL	378	1.50	1.05	0.99

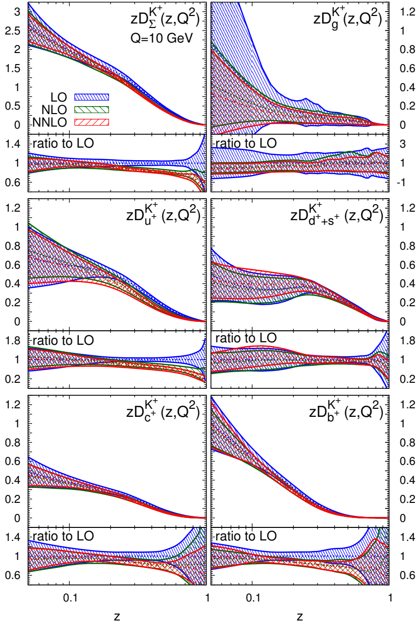
Excellent perturbative convergence
FFs almost stable from NLO to NNLO
LO FF uncertainties larger than HO

Dependence upon perturbative order: K^+

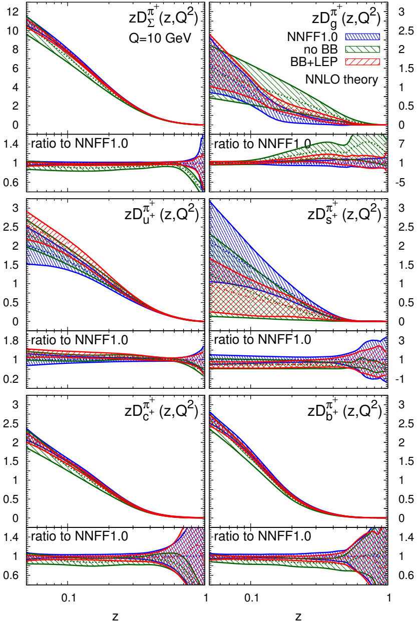
Exp.	$N_{\rm dat}$	LO $\chi^2/N_{\rm dat}$	NLO $\chi^2/N_{\rm dat}$	NNLO $\chi^2/N_{\rm dat}$
BELLE	70	0.37	0.34	0.19
BABAR	28	1.02	0.96	0.77
TASSO14	3	1.23	1.24	1.30
TASSO22	2	0.29	0.32	0.29
TASSO34	2	0.02	0.03	0.09
TPC	7	0.41	0.49	1.19
ALEPH	13	0.66	0.71	0.72
DELPHI	11	0.17	0.16	0.17
DELPHI-UDS	11	2.01	1.94	1.97
DELPHI-B	11	0.51	0.44	0.41
OPAL	9	2.02	2.08	2.10
SLD	21	0.81	0.80	0.77
SLD-UDS	21	1.16	1.19	1.11
SLD-C	20	0.49	0.46	0.42
SLD-B	21	0.71	0.68	0.71
TOTAL	250	0.73	0.72	0.67

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

i	$\mathrm{N}^{i+1}\mathrm{LO}/\mathrm{N}^{i}\mathrm{LO}$	D_g	D_Σ	D_{c^+}	D_{b^+}
0	NLO/LO [%]	95-300	70-80	65-80	70-85
1	NNLO/NLO [%]	70-130	90-100	90-110	95-115



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	0.08	(5.95)	0.08
BABAR	37	1.17	(82.2)	1.22
TASSO12	2	1.61	0.84	(1.61)
TASSO14	7	1.83	1.77	(1.85)
TASSO22	7	2.16	1.55	(2.48)
TASSO34	8	1.09	1.35	(1.55)
TASSO44	5	1.95	2.22	(2.60)
TPC	12	0.98	1.94	(0.87)
TPC-UDS	6	0.45	0.56	(0.79)
TPC-C	6	0.50	0.73	(0.57)
TPC-B	6	1.41	1.59	(1.47)
TOPAZ	4	0.66	0.75	(1.50)
ALEPH	22	0.88	0.69	0.71
DELPHI	16	2.32	2.50	2.38
DELPHI-UDS	16	1.90	1.98	1.91
DELPHI-B	16	1.09	1.10	1.13
OPAL	22	2.05	1.87	1.98
SLD	29	1.09	0.72	1.07
SLD-UDS	29	0.80	0.60	0.73
SLD-C	29	0.97	0.80	1.10
SLD-B	29	0.44	0.43	0.43
TOTAL	378	0.99	1.14	0.93

no BB: larger uncertainties; different gluon shape and different light flavour separation

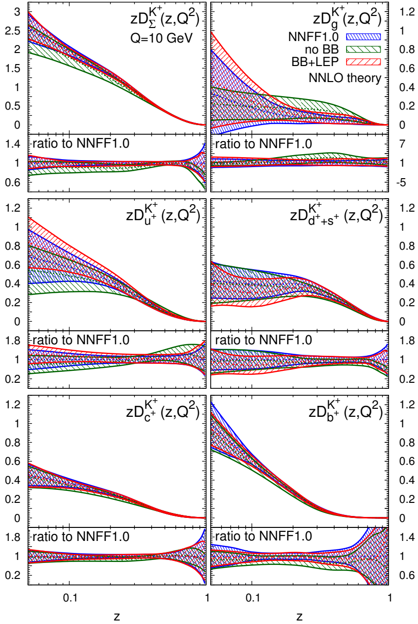
BB+LEP: comparable uncertainties; slightly different size of gluon and light flavoured quarks

Dependence upon the dataset: K^+

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	0.19	(16.3)	0.37
BABAR	28	0.77	(190)	0.99
TASSO14	3	1.30	1.80	(1.23)
TASSO22	2	0.29	0.23	(0.33)
TASSO34	2	0.09	0.02	(0.04)
TPC	7	1.19	0.61	(0.45)
ALEPH	13	0.72	0.75	0.63
DELPHI	11	0.17	0.23	0.16
DELPHI-UDS	11	1.97	2.05	2.00
DELPHI-B	11	0.41	0.45	0.48
OPAL	9	2.10	2.01	2.01
SLD	21	0.77	0.76	0.77
SLD-UDS	21	1.11	1.12	1.19
SLD-C	20	0.42	0.36	0.47
SLD-B	21	0.71	0.76	0.70
TOTAL	250	0.67	0.86	0.74

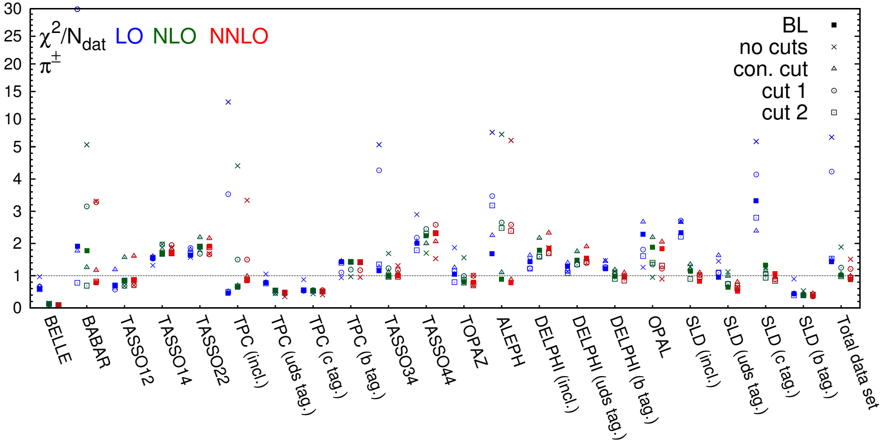
no BB: larger uncertainties; different gluon shape and different light flavour separation; significant degradation in the description of BELLE and BABAR data

BB+LEP: comparable uncertainties; FFs stable; no significant degradation in fit quality; fair description of the data not included in the fit



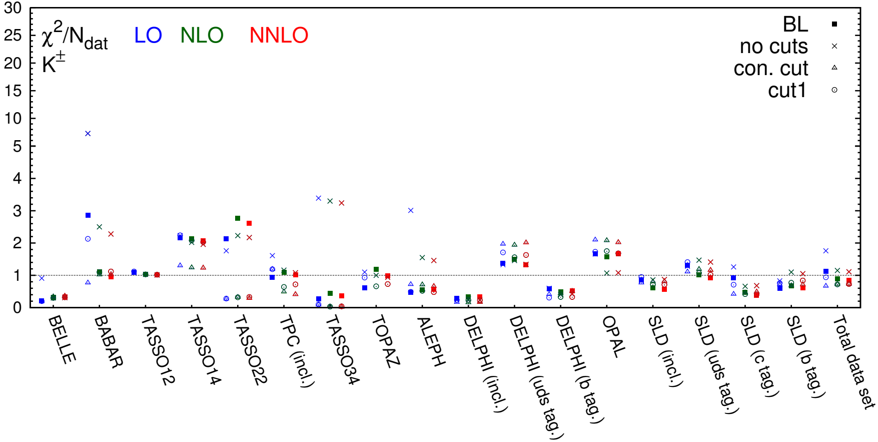
Dependence upon kinematic cuts

Hadron	BL		no cuts		con. cut		cut1		cut2	
	$z_{\rm min}^{(m_Z)}$	$z_{\rm min}$	$z_{\rm min}^{(m_Z)}$	$z_{\rm min}$	$z_{\rm min}^{(m_Z)}$	$z_{\rm min}$	$z_{\rm min}^{(m_Z)}$	$z_{\rm min}$	$z_{\rm min}^{(m_Z)}$	$z_{\rm min}$
$\pi^\pm$	0.02	0.075	0.00	0.00	0.05	0.10	0.01	0.05	0.01	0.075
$K^\pm$	0.02	0.075	0.00	0.00	0.10	0.20	0.05	0.10	—	—



Dependence upon kinematic cuts

Hadron	BL		no cuts		con. cut		cut1		cut2	
	$z_{\min}^{(m_Z)}$	$z_{\min}$	$z_{\min}^{(m_Z)}$	$z_{\min}$	$z_{\min}^{(m_Z)}$	$z_{\min}$	$z_{\min}^{(m_Z)}$	$z_{\min}$	$z_{\min}^{(m_Z)}$	$z_{\min}$
$\pi^\pm$	0.02	0.075	0.00	0.00	0.05	0.10	0.01	0.05	0.01	0.075
$K^\pm$	0.02	0.075	0.00	0.00	0.10	0.20	0.05	0.10	—	—

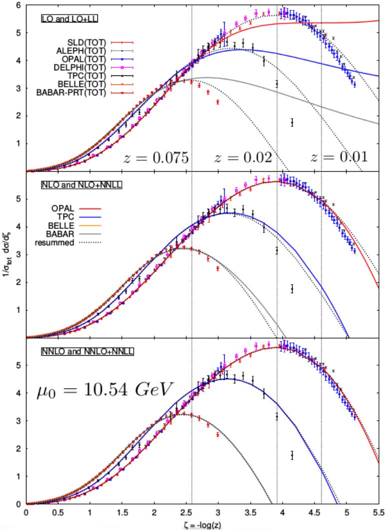


Small-z resummed Fragmentation Functions [PRD 95 (2017) 054003]

436 Total data Points:

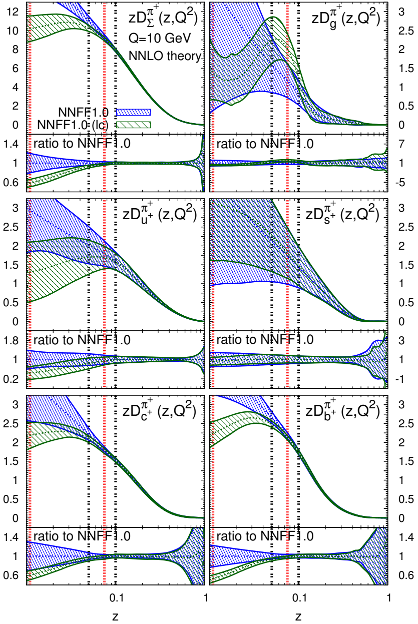
- LEP cut (z = 0.01) due to inconsistency between OPAL and ALEPH
- TPC lower cut (z = 0.02) based on difference of energy fraction $z = 2\,E_h/Q$ and three momentum fraction $x_p = z - 2m_h^2/(zQ^2) + \mathcal{O}(1/Q^4)$ in c.m.s being less than at least 15%

accuracy	χ^2	χ^2/dof
LO	1260.78	2.89
NLO	354.10	0.81
NNLO	330.08	0.76
LO+LL	405.54	0.93
NLO+NNLL	352.28	0.81
NNLO+NNLL	329.96	0.76



Slide: courtesy of D. P. Anderle

Dependence upon kinematic cuts: π^+



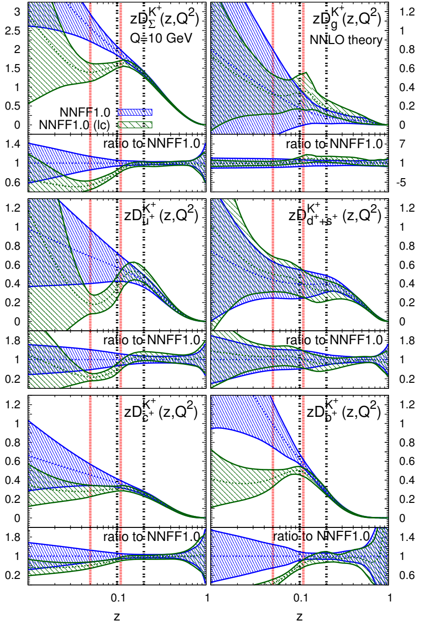
NNLO theory	NNFF1.0		NNFF1.0 (lc)	
Exp.	$N_{\rm dat}$	$\chi^2/N_{\rm dat}$	$N_{\rm dat}$	$\chi^2/N_{\rm dat}$
BELLE	70	0.08	70	0.09
BABAR	37	1.17	40	0.82
TASSO12	2	1.61	4	0.87
TASSO14	7	1.83	9	1.69
TASSO22	7	2.16	8	1.88
TASSO34	8	1.09	9	0.97
TASSO44	5	1.95	6	2.32
TPC	12	0.98	13	0.88
TPC-UDS	6	0.45	6	0.47
TPC-C	6	0.50	6	0.52
TPC-B	6	1.41	6	1.42
TOPAZ	4	0.66	5	0.75
ALEPH	22	0.88	30	2.39
DELPHI	16	2.32	22	1.70
DELPHI-UDS	16	1.90	22	1.43
DELPHI-B	16	1.09	22	0.85
OPAL	22	2.05	38	1.31
SLD	29	1.09	38	0.97
SLD-UDS	29	0.80	38	0.61
SLD-C	29	0.97	38	0.84
SLD-B	29	0.44	38	0.41
TOTAL	378	0.99	468	0.94

Slight improvement of the overall fit quality
Excellent consistency in the overlapping region
Significantly varied FF shapes at low z
Possible tensions with ALEPH at small z

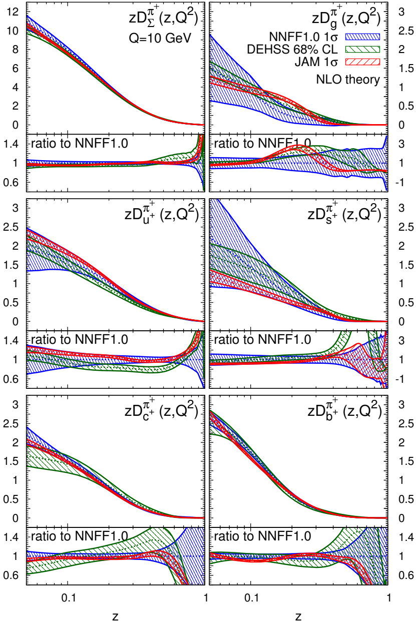
Dependence upon kinematic cuts: K^+

NNLO theory	NNFF1.0		NNFF1.0 (lc)	
Exp.	$N_{\rm dat}$	$\chi^2/N_{\rm dat}$	$N_{\rm dat}$	$\chi^2/N_{\rm dat}$
BELLE	70	0.19	70	0.32
BABAR	28	0.77	43	1.12
TASSO12	—	—	3	1.02
TASSO14	3	1.30	7	2.03
TASSO22	2	0.29	4	0.33
TASSO34	2	0.09	4	0.04
TPC	7	1.19	12	0.72
TOPAZ	—	—	3	0.73
ALEPH	13	0.72	18	0.48
DELPHI	11	0.17	16	0.23
DELPHI-UDS	11	1.97	16	1.63
DELPHI-B	11	0.41	16	0.33
OPAL	9	2.10	10	1.68
SLD	21	0.77	29	0.71
SLD-UDS	21	1.11	29	1.02
SLD-C	20	0.42	29	0.41
SLD-B	21	0.71	29	0.84
TOTAL	250	0.67	338	0.73

Slight deterioration of the overall fit quality
Excellent consistency in the overlapping region
Significantly varied FF shapes at low z



Comparison among various FF determinations (pions)



DEHSS [PRD 91 (2015) 014035]
(+SIDIS +PP)

JAM [PRD 94 (2016) 114004]
(almost same dataset as NNFF1.0)

$D_\Sigma^{\pi^+}$: excellent mutual agreement
both c.v. and unc. (bulk of the dataset)

$D_g^{\pi^+}$: slight disagreement
different shapes, larger uncertainties
DEHSS: data; JAM: parametrisation

$D_{u^+}^{\pi^+}$, $D_{s^+}^{\pi^+}$: good overall agreement
excellent with JAM, though larger uncertainties
slight different shape w.r.t. DHESS (dataset)

$D_{c^+}^{\pi^+}$, $D_{b^+}^{\pi^+}$: good overall agreement
excellent with JAM, same uncertainties
slight different shape w.r.t. DHESS (dataset)

Comparison among various FF determinations (kaons)

DEHSS [PRD 95 (2017) 094019]
(+SIDIS +PP)

JAM [PRD 94 (2016) 114004]
(almost same dataset as NNFF1.0)

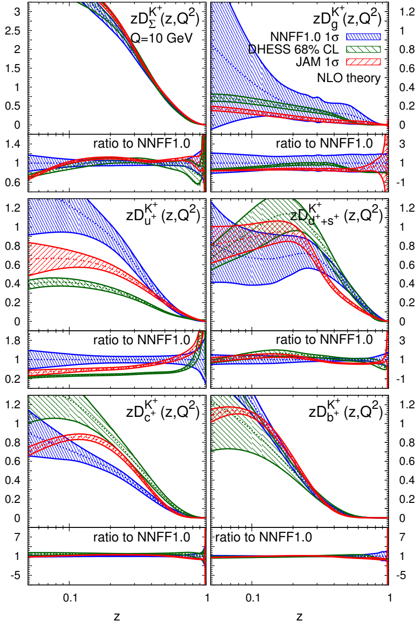
$D^{\pi^+}_{\Sigma}$: excellent agreement (both c.v. and unc.)
bulk of the dataset

$D^{\pi^+}_{g}$: good mutual agreement
similar shapes, larger uncertainties
DEHSS: data; JAM: parametrisation

$D^{\pi^+}_{u^+}$: mutual sizable disagreement
differences in dataset and parametrisation
comparable uncertainties in the data region

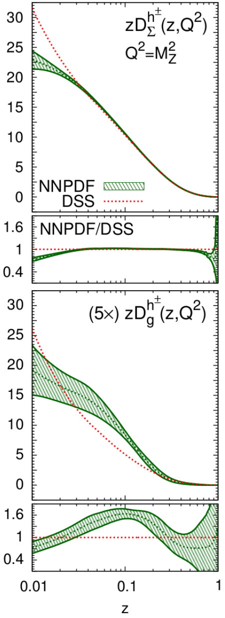
$D^{\pi^+}_{d^+} + D^{\pi^+}_{s^+}$: fair mutual agreement
differences in dataset and parametrisation
comparable uncertainties in the data region

$D^{\pi^+}_{c^+}$, $D^{\pi^+}_{b^+}$: excellent mutual agreement
uncertainties similar to JAM
DHESS shows inflated uncertainties

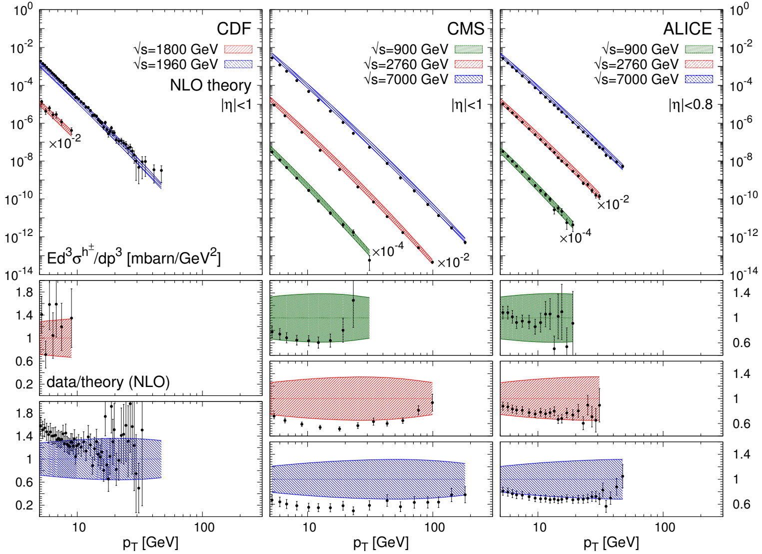


Fragmentation functions of unidentified charged hadrons

Experiment	Observable	$\sqrt{s}$ [GeV]	$N_{\rm dat}$	$\chi^2/N_{\rm dat}$
TASSO14	F_2 (incl.)	14.00	15 (20)	1.23
TASSO22	F_2 (incl.)	22.00	15 (20)	0.51
TPC	F_2 (incl.)	29.00	21 (34)	1.65
TASSO35	F_2 (incl.)	35.00	15 (20)	1.14
TASSO44	F_2 (incl.)	44.00	15 (20)	0.68
ALEPH	F_2 (incl.)	91.20	32 (35)	1.04
	F_L (incl.)	91.20	19 (21)	0.36
DELPHI	F_2	91.20	21 (27)	0.65
	F_2 (uds tagged)	91.20	21 (27)	0.17
	F_2 (b tagged)	91.20	21 (27)	0.82
	F_L (incl.)	91.20	20 (22)	0.72
	F_L (b tagged)	91.20	20 (22)	0.44
OPAL	F_2 (incl.)	91.20	20 (22)	2.41
	F_2 (uds tagged)	91.20	20 (22)	0.90
	F_2 (c tagged)	91.20	20 (22)	0.61
	F_2 (b tagged)	91.20	20 (22)	0.21
	F_L (incl.)	91.20	20 (22)	0.31
SLD	F_2	91.28	34 (40)	0.75
	F_2 (uds tagged)	91.28	34 (40)	1.03
	F_2 (c tagged)	91.28	34 (40)	0.62
	F_2 (b tagged)	91.28	34 (40)	0.97
Total dataset			471 (527)	0.83



Fragmentation functions of unidentified charged hadrons



4. Conclusions

Summary

1. The NNPDFpol parton sets
   $\longrightarrow$ continuous effort in the inclusion of new data
   $\longrightarrow$ sometime in 2018: new release with all DIS and pp data (except πs)
2. Helicity-dependent PDFs and lattice QCD
   $\longrightarrow$ joint effort between the global-fit and lattice-QCD communities
   $\longrightarrow$ lattice QCD could provide a valuable input in the future
3. NNFF1.0: NNPDF fragmentation functions
   $\longrightarrow$ recent effort in the determination of the fragmentation functions
   $\longrightarrow$ unpolarised/polarised PDF and FF sets available from a common methodology now

Outlook

1. Extend the `APFEL` framework to include SIDIS (time-like DY)
   $\longrightarrow$ mid-term goal, work in progress
2. Extend the `NNPDF` framework to a simultaneous fit of PDFs and FFs
   $\longrightarrow$ long-term goal, requires a careful code thinking

Summary

1. The NNPDFpol parton sets
   $\longrightarrow$ continuous effort in the inclusion of new data
   $\longrightarrow$ sometime in 2018: new release with all DIS and pp data (except πs)
2. Helicity-dependent PDFs and lattice QCD
   $\longrightarrow$ joint effort between the global-fit and lattice-QCD communities
   $\longrightarrow$ lattice QCD could provide a valuable input in the future
3. NNFF1.0: NNPDF fragmentation functions
   $\longrightarrow$ recent effort in the determination of the fragmentation functions
   $\longrightarrow$ unpolarised/polarised PDF and FF sets available from a common methodology now

Outlook

1. Extend the `APFEL` framework to include SIDIS (time-like DY)
   $\longrightarrow$ mid-term goal, work in progress
2. Extend the `NNPDF` framework to a simultaneous fit of PDFs and FFs
   $\longrightarrow$ long-term goal, requires a careful code thinking

Thank you