Evidence for intrinsic charm quarks in the proton

Tommaso Giani

Based on Nature 608 (2022) 7923, 483-487

Richard D. Ball , Alessandro Candido , Juna Cruz Martininez, Stefano Forte,
Tommaso Giani , Felix Hekhorn, Kirill Kudashkin, Giacomo Magni and Juan Rojo

Cracow EPIPHANY Conference

Tuesday, January 17

THE INTRINSIC CHARM OF THE PROTON

S.J. BRODSKY [1]
Stanford Linear Accelerator Center,
Stanford, California 94305, USA

and

P. HOYER, C. PETERSON and N. SAKAI [2]
NORDITA, Copenhagen, Denmark

Received 22 April 1980

Recent data give unexpectedly large cross-sections for charmed particle production at high x_F in hadron collisions. This may imply that the proton has a non-negligible uudc$\bar{c}$ Fock component. The interesting consequences of such a hypothesis are explored.

nature

Explore content About the journal Publish with us

nature > articles > article

Article | Open Access | Published: 17 August 2022

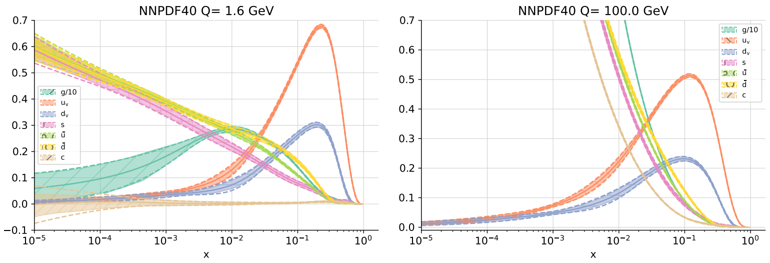
Evidence for intrinsic charm quarks in the proton

The NNPDF Collaboration

Nature **608**, 483–487 (2022) | Cite this article

Introduction: the treatment of the charm PDF

PDFs



$$F\left(x,Q^2\right)=\sum_i f_i\left(\mu^2\right)\otimes \hat{F}_i\left(\frac{Q^2}{\mu^2},\alpha_s\right)+\mathcal{O}\left(\frac{\Lambda_{QCD}^2}{Q^2}\right)$$

$\rightarrow$ $f_i\left(x,\mu^2\right)$ are universal: can be extracted from a set of data for some processes and used to make predictions for others

Heavy quarks PDFs

<u>$N_f = 4$ **scheme (massless scheme)**</u>

- $Q >> m_c$
- computation of $\hat{F}$ does not retain mass effects
- collinear $\log \frac{Q^2}{m_c^2}$ are factored in charm PDF $\rightarrow \quad f_c^{n_f+1}(x, \mu)$
- 4 active flavors in β function and DGLAP

<u>$N_f = 3$ **scheme (massive scheme)**</u>

- $Q \sim m_c$
- computation of $\hat{F}$ retains mass effects, and explicit not resummed $\log \frac{Q^2}{m_c^2}$
- charm decouples: renormalization-group independent charm PDF $\rightarrow f_c^{n_f}(x)$
  [Collins, Wilczek and Zee, Phys.Rev. D18,242, 1978]
- 3 active flavors in β function and DGLAP

$$\text{4FS}: \quad f_q^{n_f+1}\left(x, Q^2\right), \quad q = \{u, d, s, c, \bar{u}, \bar{d}, \bar{s}, \bar{c}\}$$

$$\text{3FS}: \quad f_q^{n_f}\left(x, Q^2\right), \; f_h^{n_f}(x) \quad q = \{u, d, s, \bar{u}, \bar{d}, \bar{s}\} \quad h = \{c, \bar{c}\}$$

Relation between the PDFs and α_s in the two schemes is given by

$$\begin{pmatrix} f_q \\ f_c \end{pmatrix}^{n_f+1} (\mu_h^2) = \begin{pmatrix} A_{qq} & A_{qc} \\ A_{cq} & A_{cc} \end{pmatrix} \begin{pmatrix} f_q \\ f_c \end{pmatrix}^{n_f} (\mu_h^2)$$

- Operator Matrix Elements (OME) A_{ij} known up to N^3LO
- OME depend on $\alpha_s\left(\mu_h\right)$ and $\log \frac{\mu_h}{m_c}$

Perturbative charm

Charm in $N_f = 4$ generated purely by perturbative matching

$$f_c^{n_f}(x) = 0 \quad \rightarrow \quad f_c^{n_f+1}(x, m_c^2) = \sum_{i=q,g} A_{ci}\, f_i^{n_f}(x, m_c^2) = \mathcal{O}\left(\alpha_s^2\right)$$

$$f_c^{n_f+1}(x, Q^2) \propto \alpha_s \log \frac{Q^2}{m_c^2} \left(P_{qg} \otimes f_g^{(n_f+1)}\right) + \mathcal{O}\left(\alpha_s^2\right)$$

- once the light flavors are determined, charm is fixed by matching and evolution

Fitted charm

Non vanishing $N_f = 3$ charm PDF is allowed

$$f_c^{n_f}(x) \neq 0 \quad \rightarrow \quad f_c^{n_f+1}(x, m_c^2) = \sum_{i=q,g} A_{ci}\, f_i^{n_f}(x, m_c^2) + A_{cc}\, f_c^{n_f}(x)$$

- 4FS charm is not fixed by the light flavors PDF
- we can <u>fit charm PDF from data</u> as any other quark

Can we provide a determination of $f_c^{n_f}(x)$?

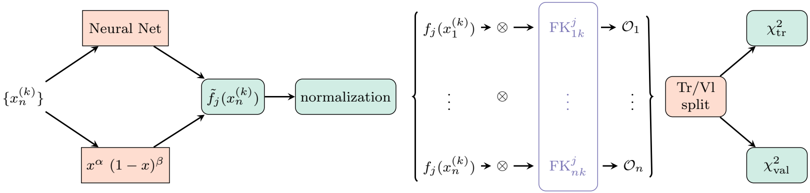
intrinsic charm $\iff \quad f_c^{n_f}(x) \neq 0$

The NNPDF4.0 determination of PDFs

Methodology

→ PDFs are parameterized in $N_f = 4$, at $Q_0 = 1.65$ GeV

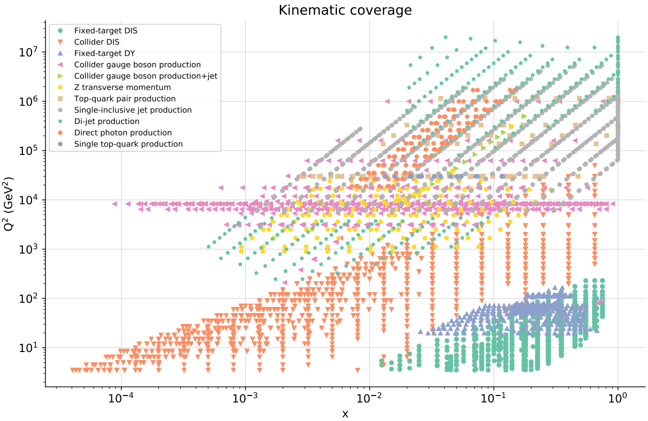
→ PDFs parameterized using a neural network and a preprocessing polynomial factor



→ split of data in training validation sets → χ^2_{tr} , χ^2_{val}

→ minimization performed on χ^2_{tr}, stopping controlled by χ^2_{val}

Data



→ new processes included for the first time: single top, W+jet, isolated photon, di-jets

→ extensive use of 13 TeV dataset

→ total of $\mathcal{O}(4000)$ datapoints

Data set	NNPDF4.0	NNPDF3.1	ABMP16	CT18	MSHT20
ATLAS W, Z 7 TeV (2010)	✓	✓	✓	✓	✓
ATLAS W, Z 7 TeV (2011)	✓	✓	✗	✓	✓
ATLAS low-mass DY 7 TeV	✓	✓	✗	✗	✗
ATLAS high-mass DY 7 TeV	✓	✓	✗	✗	✓
ATLAS W 8 TeV	✓	✗	✗	✗	✓
ATLAS DY 2D 8 TeV	✓	✗	✗	✗	✓
ATLAS high-mass DY 2D 8 TeV	✓	✗	✗	✗	✓
ATLAS $\sigma_{W,Z}$ 13 TeV	✓	✗	✓	✗	✗
ATLAS W^++jet 8 TeV	✓	✗	✗	✗	✓
ATLAS Z p_T 8 TeV	✓	✓	✗	✓	✓
ATLAS $\sigma_{tt}^{\rm tot}$ 7, 8 TeV	✓	✓	✓	✗	✗
ATLAS $\sigma_{tt}^{\rm tot}$ 13 TeV	✓	✓	✓	✗	✗
ATLAS $t\bar{t}$ lepton+jets 8 TeV	✓	✓	✗	✓	✓
ATLAS $t\bar{t}$ dilepton 8 TeV	✓	✗	✗	✗	✓
ATLAS single-inclusive jets 7 TeV, R=0.6	✗	✓	✗	✓	✓
ATLAS single-inclusive jets 8 TeV, R=0.6	✓	✗	✗	✗	✗
ATLAS dijets 7 TeV, R=0.6	✓	✗	✗	✗	✗
ATLAS direct photon production 13 TeV	✓	✗	✗	✗	✗
ATLAS single top R_t 7, 8, 13 TeV	✓	✗	✓	✗	✗
ATLAS single top diff. 7, 8 TeV	✓	✗	✗	✗	✗
ATLAS single top diff. 8 TeV	✓	✗	✗	✗	✗

Data set	NNPDF4.0	NNPDF3.1	ABMP16	CT18	MSHT20
LHCb Z 940 pb	✓	✓	✗	✗	✓
LHCb $Z \to ee$ 2 fb	✓	✓	✓	✓	✓
LHCb $W, Z \to \mu$ 7 TeV	✓	✓	✓	✓	✓
LHCb $W, Z \to \mu$ 8 TeV	✓	✓	✓	✓	✓
LHCb $Z \to \mu\mu, ee$ 13 TeV	✓	✗	✗	✗	✗

Data set	NNPDF4.0	NNPDF3.1	ABMP16	CT18	MSHT20
CMS W electron asymmetry 7 TeV	✓	✓	✗	✓	✓
CMS W muon asymmetry 7 TeV	✓	✓	✓	✓	✗
CMS Drell-Yan 2D 7 TeV	✓	✓	✗	✗	✓
CMS W rapidity 8 TeV	✓	✓	✓	✓	✓
CMS Z p_T 8 TeV	✓	✓	✗	✓	✗
CMS $W + c$ 7 TeV	✓	✓	✗	✗	✓
CMS $W + c$ 13 TeV	✓	✗	✗	✗	✗
CMS single-inclusive jets 2.76 TeV	✗	✓	✗	✗	✓
CMS single-inclusive jets 7 TeV	✗	✓	✗	✓	✓
CMS dijets 7 TeV	✓	✗	✗	✗	✗
CMS single-inclusive jets 8 TeV	✗	✗	✗	✓	✓
CMS 3D dijets 8 TeV	✓	✗	✗	✗	✗
CMS $\sigma_{tt}^{\rm tot}$ 5 TeV	✓	✗	✓	✗	✗
CMS $\sigma_{tt}^{\rm tot}$ 7, 8 TeV	✓	✓	✓	✗	✓
CMS $\sigma_{tt}^{\rm tot}$ 13 TeV	✓	✓	✓	✗	✗
CMS $t\bar{t}$ lepton+jets 8 TeV	✓	✓	✗	✗	✓
CMS $t\bar{t}$ 2D dilepton 8 TeV	✓	✗	✗	✓	✓
CMS $t\bar{t}$ lepton+jet 13 TeV	✓	✗	✗	✗	✗
CMS $t\bar{t}$ dilepton 13 TeV	✓	✗	✗	✗	✗
CMS single top $\sigma_t + \sigma_{\bar{t}}$ 7 TeV	✓	✗	✓	✗	✗
CMS single top R_t 8, 13 TeV	✓	✗	✓	✗	✗

→ Python object oriented codebase

→ Freedom to use external libraries (default: TensorFlow)

→ Documentation and tutorials provided

→ Results for a replica available in less than an hour (~ 40mins on 1 CPU)

→ **Code fully public**

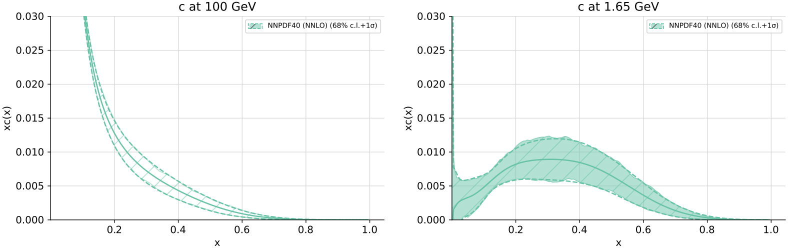
Tests passing DOI 10.5281/zenodo.5362229

NNPDF: An open-source machine learning framework for global analyses of parton distributions

The NNPDF collaboration determines the structure of the proton using Machine Learning methods. This is the main repository of the fitting and analysis frameworks. In particular it contains all the necessary tools to reproduce the NNPDF4.0 PDF determinations.

Extraction of intrinsic c

NNPDF4.0 charm PDF: parameterized and determined from experimental data in $N_f = 4$.



$$\tilde{\boldsymbol{f}}^{n_f+1}\left(Q_1^2\right) = \tilde{\boldsymbol{E}}^{(n_f+1)}\left(Q_1^2 \leftarrow \mu_h\right) \tilde{\boldsymbol{A}}^{n_f}\left(\mu_h^2\right) \tilde{\boldsymbol{E}}^{(n_f)}\left(\mu_h \leftarrow Q_0^2\right) \tilde{\boldsymbol{f}}^{n_f}\left(Q_0^2\right)$$

EKO
Evolution Kernel Operators

tests passing | docs passing | codecov 100% | codefactor A

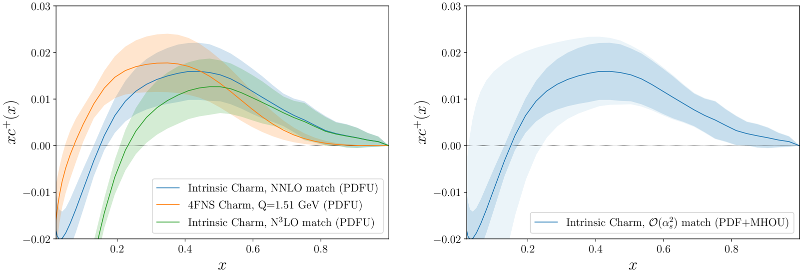
EKO is a Python module to solve the DGLAP equations in N-space in terms of Evolution Kernel Operators in x-space.

[Candido, Hekhorn and Magni, Eur.Phys.J.C 82 (2022)]

- → implementation of DGLAP solutions at LO, NLO, NNLO
- → various solution methods implemented
- → implementation of matching in Mellin space
- → inverse Matching implemented both exactly and expanding and α_s

Intrinsic charm

$$f_i^{n_f+1}(x, \mu = 1.65\,\text{GeV}) \;\rightarrow\; f_i^{n_f+1}(x, \mu = 1.51\,\text{GeV}) \;\rightarrow\; f_c^{n_f}(x)$$

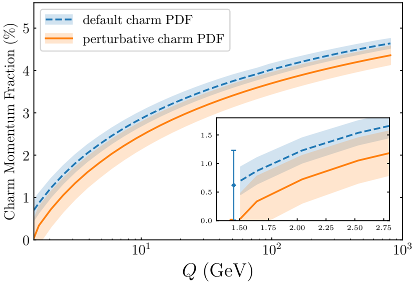


- $\rightarrow$ valence like peak in the region $0.3 < x < 0.6$
- $\rightarrow$ in the region $0.3 < x < 0.6$ PDF uncertainty is the dominant one
- $\rightarrow$ large perturbative uncertainties for $x < 0.2$

Momentum fraction

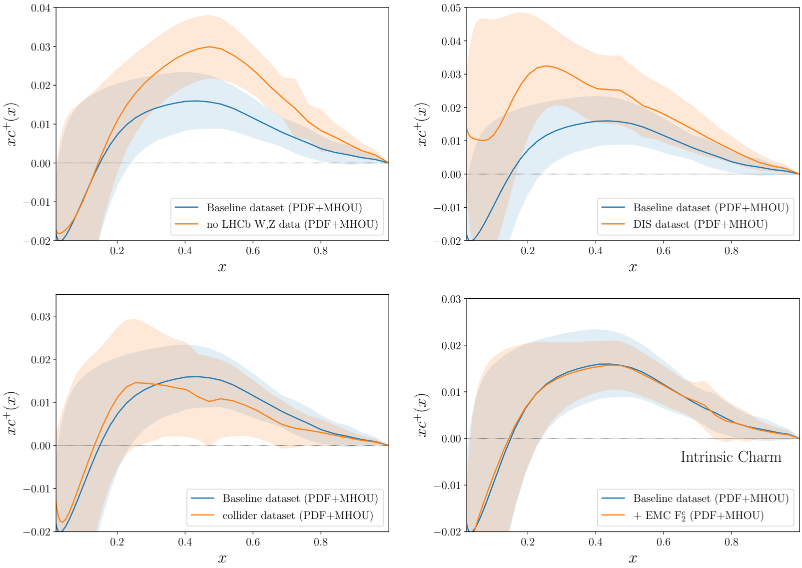
$$[c] = \int_0^1 dx\, xc^+(x,Q)$$

Scheme	Q	Charm PDF	m_c	$[c]$ (%)
3FNS	–	default	1.51 GeV	$0.62 \pm 0.28_{\rm pdf} \pm 0.54_{\rm mhou}$
4FNS	1.65 GeV	default	1.51 GeV	$0.87 \pm 0.23_{\rm pdf}$
4FNS	1.65 GeV	perturbative	1.51 GeV	$0.346 \pm 0.005_{\rm pdf} \pm 0.44_{\rm mhou}$



- → momentum fraction of 4FNS PDF different from 0 at 3σ level (PDF uncertainty only)
- → big MHO uncertainty on IC momentum fraction (due to region $x < 0.2$)
- → not possible to tell whether 4FS momentum fraction is of perturbative or intrinsic origin

Dataset dependence



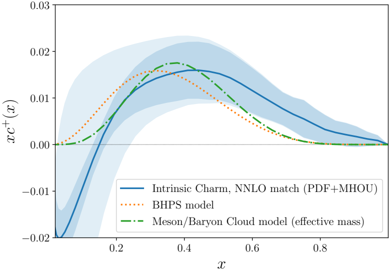
Comparison with models

- BHPS model: [Phy. Letter B (1980) 451-455]

$$xc^{+} = \frac{1}{2} N x^{3} \left[\frac{1}{3} (1 - x) \left(1 + 10x + x^{2}\right) + 2x \left(1 + x^{2}\right) \ln x \right]$$

- Meson Baryon model: [arxiv:1311.1578]

$$xc^{+} = \frac{N}{B(\alpha + 2, \beta + 1)} x^{1+\alpha} (1 - x)^{\beta}$$

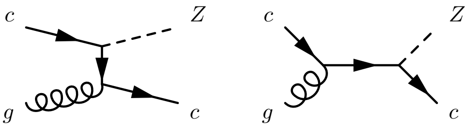


Phenomenology: Z+c @ LHCb

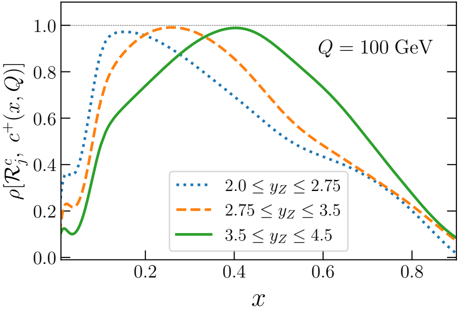
Z+c measurements from LHCb

Measurement of Z bosons produced in association with charm in the forward region
[Phys.Rev.Lett. 128 (2022)]

$$\mathcal{R}_j^c = \frac{\sigma\,(pp \to Z + \text{charm jet})}{\sigma\,(pp \to \text{Z+jet})}$$



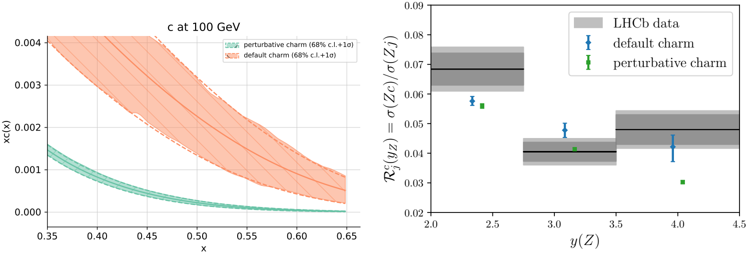
$y(Z)$	$\mathcal{R}_j^c$ (%)
2.00–2.75	$6.84 \pm 0.54 \pm 0.51$
2.75–3.50	$4.05 \pm 0.32 \pm 0.31$
3.50–4.50	$4.80 \pm 0.50 \pm 0.39$
2.00–4.50	$4.98 \pm 0.25 \pm 0.35$



→ The most forward rapidity bin is sensitive to charm PDF in the region of the valence peak

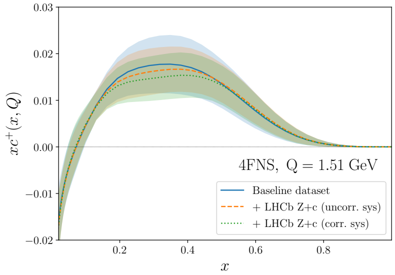
Theory prediction for R_j^c computed using POWHEG @ NLO + PS, using both perturbative and default (fitted) charm

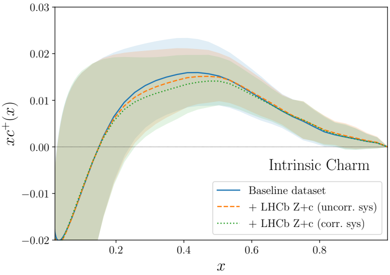
Z bosons	$p_{\mathrm{T}}(\mu) > 20\,\mathrm{GeV}$, $2.0 < \eta(\mu) < 4.5$, $60 < m(\mu^+\mu^-) < 120\,\mathrm{GeV}$
Jets	$20 < p_{\mathrm{T}}(j) < 100\,\mathrm{GeV}$, $2.2 < \eta(j) < 4.2$
Charm jets	$p_{\mathrm{T}}(c \text{ hadron}) > 5\,\mathrm{GeV}$, $\Delta R(j, c \text{ hadron}) < 0.5$
Events	$\Delta R(\mu, j) > 0.5$

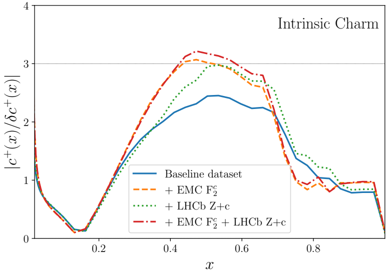


→ theory prediction based on perturbative charm in disagreement with LHCb data

→ better agreement found using NNPDF4.0 baseline (fitted charm)







→ inclusion of LHCb (by bayesian RW) and EMC data gives compatible results with moderate reduction of PDF error

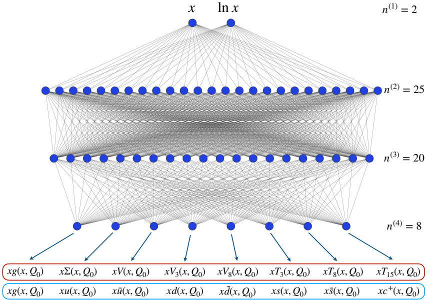
→ local significance for IC at 2.5σ in $0.3 < x < 0.6$, getting to 3σ when including LHCb and EMC data

Summary

- starting point: NNPDF4.0 charm PDF, fitted in the 4FS
- determination of charm PDF in the 3FNS
  - → local evidence for non-vanishing valence peak
  - → large uncertainties at small x
- future data
  - → HL-LHC data
  - → EIC data for F_c^2
- to do:
  - → $c - \bar{c}$ asymmetry

Hyperoptimization

Parameter	NNPDF4.0
Architecture	2-25-20-8
Activation function	hyperbolic tangent
Initializer	`glorot_normal`
Optimizer	`Nadam`
Clipnorm	6.0×10^{-6}
Learning rate	2.6×10^{-3}
Maximum # epochs	17×10^{3}
Stopping patience	10% of max epochs
Initial positivity $\Lambda^{(\mathrm{pos})}$	185
Initial integrability $\Lambda^{(\mathrm{int})}$	10



✗ Selecting manually the best set of parameters is a slow process and systematic success is not guaranteed

✓ Hyperparameter scan: let the computer decide automatically

- Define a methodology (a specific hyperparameter combination)
- Define a reward function to grade the methodology
- Scan over thousands of hyperparameter combinations and select the best one

DGLAP

Evolution equations

$$\mu_F^2 \frac{d\boldsymbol{f}}{d\mu_F^2}\left(x, \mu_F^2\right) = \boldsymbol{P}\left(\alpha_s\left(\mu_R^2\right), \mu_F^2\right) \otimes \boldsymbol{f}\left(\mu_F^2\right) ,$$

can be written in Mellin space as

$$\frac{d\tilde{\boldsymbol{f}}}{d\alpha_s}\left(N, \alpha_s\right) = -\frac{\gamma\left(N, \alpha_s\right)}{\beta\left(\alpha_s\right)}\tilde{\boldsymbol{f}}\left(N, \alpha_s\right) ,$$

and solved as

$$\tilde{\boldsymbol{f}}\left(N, \alpha_s\right) = \tilde{\boldsymbol{E}}\left(\alpha_s \leftarrow \alpha_s^0\right)\tilde{\boldsymbol{f}}\left(N, \alpha_s^0\right) , \quad \tilde{\boldsymbol{E}}\left(\alpha_s \leftarrow \alpha_s^0\right) = \mathcal{P}\exp\left[-\int_{\alpha_s^0}^{\alpha_s}\frac{\gamma\left(\alpha_s'\right)}{\beta\left(\alpha_s'\right)}d\alpha_s'\right] .$$

In x-space

$$\boldsymbol{f}\left(x, \mu\right) = \boldsymbol{E}\left(\mu \leftarrow \mu_0\right) \otimes \boldsymbol{f}\left(x, \mu_0\right) .$$

Matching

When we cross the threshold for heavy quark production we have to apply the matching

$$\tilde{\boldsymbol{f}}^{n_f+1}\left(Q_1^2\right) = \tilde{\boldsymbol{E}}^{(n_f+1)}\left(Q_1^2 \leftarrow \mu_h\right) \tilde{\boldsymbol{A}}^{n_f}\left(\mu_h^2\right) \tilde{\boldsymbol{E}}^{(n_f)}\left(\mu_h \leftarrow Q_0^2\right) \tilde{\boldsymbol{f}}^{n_f}\left(Q_0^2\right)$$

$$\begin{pmatrix} \tilde{V} \\ \tilde{h}_- \end{pmatrix}^{n_f+1} \left(\mu_h^2\right) = \tilde{\boldsymbol{A}}^{n_f}_{NS,h_-}\left(\mu_h^2\right) \begin{pmatrix} \tilde{V} \\ \tilde{h}_- \end{pmatrix}^{n_f} \left(\mu_h^2\right)$$

$$\begin{pmatrix} \tilde{g} \\ \tilde{\Sigma} \\ \tilde{h}_+ \end{pmatrix}^{n_f+1} \left(\mu_h^2\right) = \tilde{\boldsymbol{A}}^{n_f}_{S,h_+}\left(\mu_h^2\right) \begin{pmatrix} \tilde{g} \\ \tilde{\Sigma} \\ \tilde{h}_+ \end{pmatrix}^{n_f} \left(\mu_h^2\right)$$

OME depend on $\alpha_s^{n_f+1}\left(\mu_h^2\right)$ and $\log\left(\mu_h^2/m_h^2\right)$

Matching: more in details

OME are available up to N3LO with only NLO for heavy quark entries

$$\tilde{\boldsymbol{A}}^{n_f}\left(\mu_h^2\right) = \boldsymbol{I} + \alpha_s\left(\mu_h^2\right)\tilde{\boldsymbol{A}}^{n_f,(1)} + \alpha_s^2\left(\mu_h^2\right)\tilde{\boldsymbol{A}}^{n_f,(2)} + \alpha_s^3\left(\mu_h^2\right)\tilde{\boldsymbol{A}}^{n_f,(3)} + \dots$$

$$\tilde{\boldsymbol{A}}^{n_f,(1)}_{S,h_+} = \begin{pmatrix} A^{S,(1)}_{gg} & 0 & A^{S,(1)}_{gH} \\ 0 & 0 & 0 \\ A^{S,(1)}_{Hg} & 0 & A^{S,(1)}_{HH} \end{pmatrix}, \quad \tilde{\boldsymbol{A}}^{n_f,(2)}_{S,h_+} = \begin{pmatrix} A^{S,(2)}_{gg} & A^{S,(2)}_{gq} & 0 \\ 0 & A^{(2)}_{qq} & 0 \\ A^{S,(2)}_{Hg} & A^{S,(2)}_{Hq} & 0 \end{pmatrix}, \quad \tilde{\boldsymbol{A}}^{n_f,(3)}_{S,h_+} = \begin{pmatrix} A^{S,(2)}_{gg} & A^{S,(2)}_{gq} & 0 \\ 0 & A^{(2)}_{qq} & 0 \\ A^{S,(2)}_{Hg} & A^{S,(2)}_{Hq} & 0 \end{pmatrix}$$

$$\tilde{\boldsymbol{A}}^{n_f,(1)}_{NS,h_-} = \begin{pmatrix} 0 & 0 \\ 0 & A^{NS,(1)}_{HH} \end{pmatrix}, \quad \tilde{\boldsymbol{A}}^{n_f,(2)}_{NS,h_-} = \begin{pmatrix} A^{NS,(2)}_{qq} & 0 \\ 0 & 0 \end{pmatrix}, \quad \tilde{\boldsymbol{A}}^{n_f,(3)}_{NS,h_-} = \begin{pmatrix} A^{NS,(3)}_{qq} & 0 \\ 0 & 0 \end{pmatrix}$$

NNLO [Eur.Phys.J.C 1 (1998) 301-320],
NLO massive [Phys.Lett.B 754 (2016) 49-58],
N^{3}LO [Ablinger, Blumlein et al, 2009-2017]

The inversion can be done either perturbatively or exactly

$$\left(\tilde{\boldsymbol{A}}^{n_f}\right)^{-1}\left(\mu_h^2\right) = \boldsymbol{I} - \alpha_s\left(\mu_h^2\right)\tilde{\boldsymbol{A}}^{n_f,(1)} + \alpha_s^2\left(\mu_h^2\right)\left(\tilde{\boldsymbol{A}}^{n_f,(2)} - \left(\tilde{\boldsymbol{A}}^{n_f,(1)}\right)^2\right) + \mathcal{O}\left(\alpha_s^3\right)$$

Fitted vs perturbative charm

Fitted charm

$\rightarrow$ agnostic about $f_c^{n_f}$

$\rightarrow$ independent of matching conditions: no dependence on perturbative order, μ_h, m_c

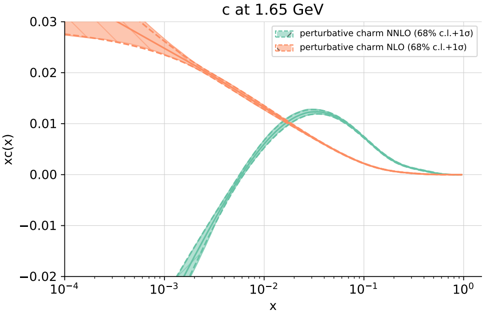
$\rightarrow$ more realistic PDF error given by the data

Perturbative charm

✗ fix $f_c^{n_f} = 0$

✗ c PDF built using matching + evolution: perturbative unstable, strong dependence on μ_h, m_c

✗ big MHO uncertainty, unrealistic PDF error



Mass dependence