

# Heavy-flavour production processes relevant for PDF fits

S. Alekhin, G. Bevilacqua, Maria Vittoria Garzelli, A. Kardos,  
S.-O. Moch, O. Zenaiev  
*in collaboration with M. Benzke, B. Kniehl*

Università degli Studi di Firenze, Dipartimento di Fisica e Astronomia  
Eberhard Karls Tuebingen Universität, Institut für Theoretische Physik  
Hamburg Universität, II Institut für Theoretische Physik  
University of Debrecen, MTA-DE Particle Physics Research Group



EBERHARD KARLS  
UNIVERSITÄT  
TÜBINGEN



XXVII International Workshop on DIS and Related Subjects,  
Torino, 8 - 12 April, 2019

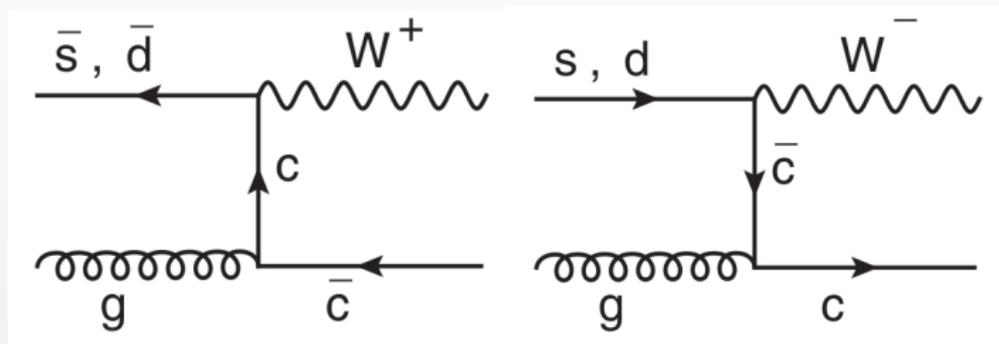
# Outline

Biased selection of processes:

- \*  $W + c$  hadroproduction  $\Rightarrow$  for  $s$  and  $\bar{s}$  quark PDFs.
- \* open heavy flavour (charm and bottom) hadroproduction  
 $\Rightarrow$  for gluon and sea quark PDFs.

I apologize for not discussing all the rest (top, heavy-flavour production in DIS, charmonium, etc...see many other talks at this Workshop).

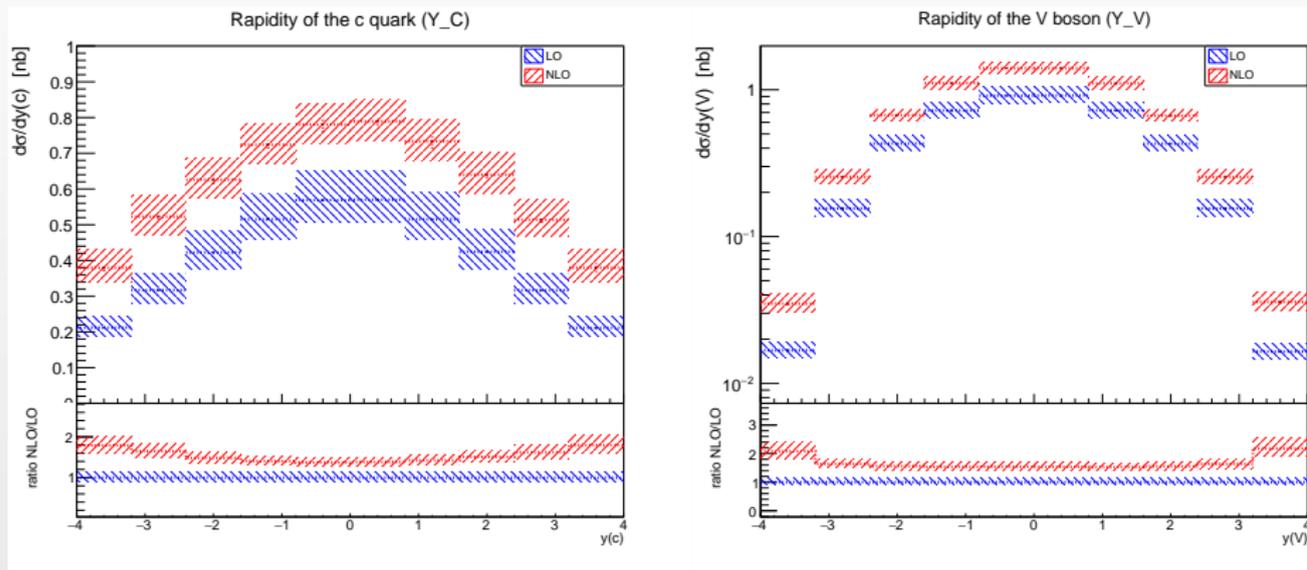
## Wc production and strange quark PDFs



- \* Direct sensitivity to  $s$  and  $\bar{s}$  PDF!
- \* At NLO new channels open up.
- \* Sensitivity to  $s$  PDF in other processes:
  - $W, Z$  hadroproduction (indirect),
  - charm production in  $\nu$ - $N$  DIS (direct).

# $pp \rightarrow W^+ \bar{c}$ production at LO and NLO

Using as ingredients the CT14nlo PDFs,  $m_c = 1.5$  GeV,  $\mu_R = \mu_F = H_T/2$ ,  $\sqrt{s} = 7$  TeV, diagonal  $V_{CKM}$ , and the HELAC-NLO framework:

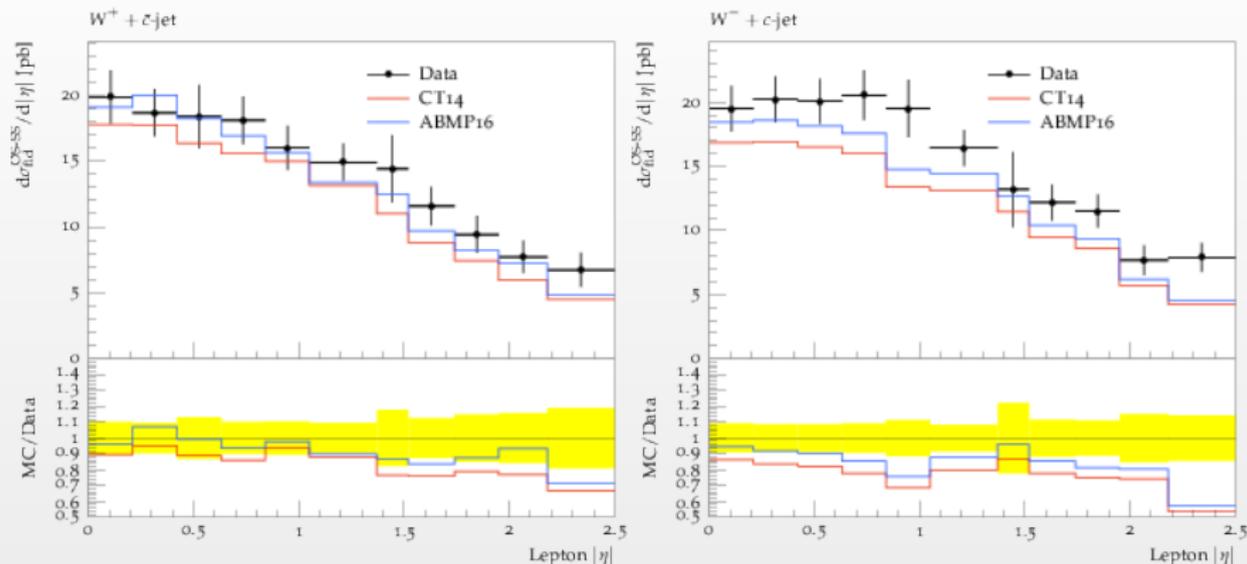


$$\sigma_{TOT}^{LO} = 3.58087 \pm 0.0006 [\text{stat}] -0.40 + 0.52 [\text{scale}] \text{ nb}$$
$$\sigma_{TOT}^{NLO} = 5.552 \pm 0.003 [\text{stat}] -0.48 + 0.58 [\text{scale}] \text{ nb}$$

## $pp \rightarrow Wc$ production at NLO+PS

- \* PowHel + SMC = Helac-NLO + POWHEGBOX + PYTHIA8
- \* Fully exclusive events allow comparisons with the experimental data.
- \* Total  $\sigma_{NLO}$  preserved in absence of cuts (according to the POWHEG method).
- \* Various sources of uncertainties under study.

# Theory predictions vs. ATLAS experimental data from [arXiv:1402.6263] (PRELIMINARY)

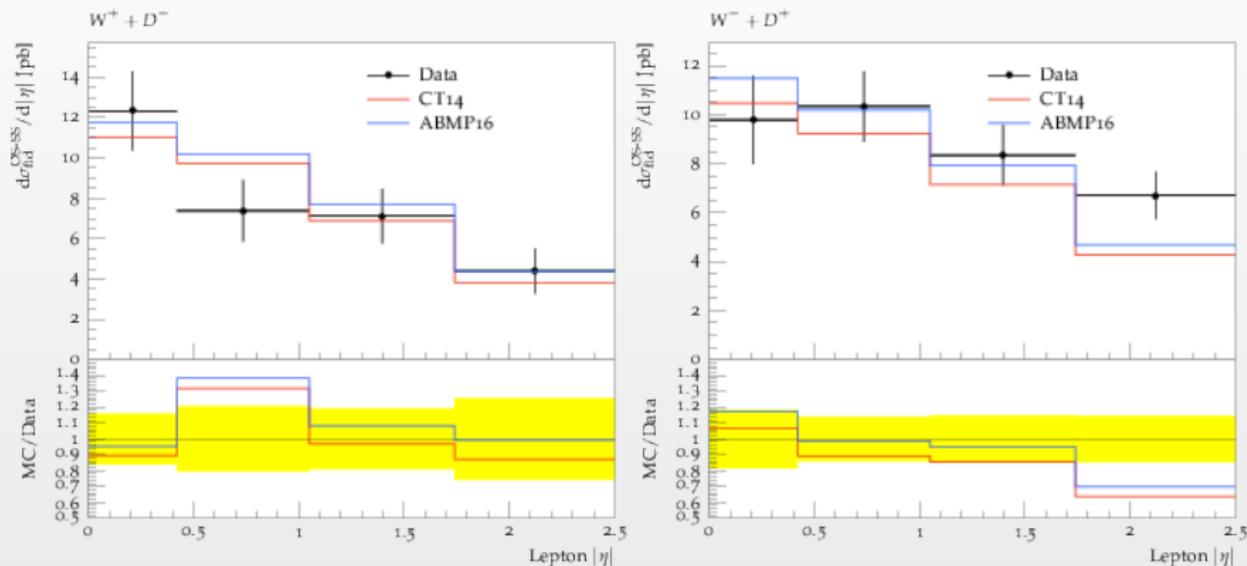


cuts:  $W \rightarrow l\nu$ ,  $p_{T,c\text{-jet}} > 25$  GeV,  $\Delta R = 0.4$

$\sigma^{exp}(W^+\bar{c}) = 33.6 \pm 0.9 \pm 1.8 pb$   $\sigma^{theo}(W^+\bar{c})(ABMP) = 32.1 \pm 0.8 pb$

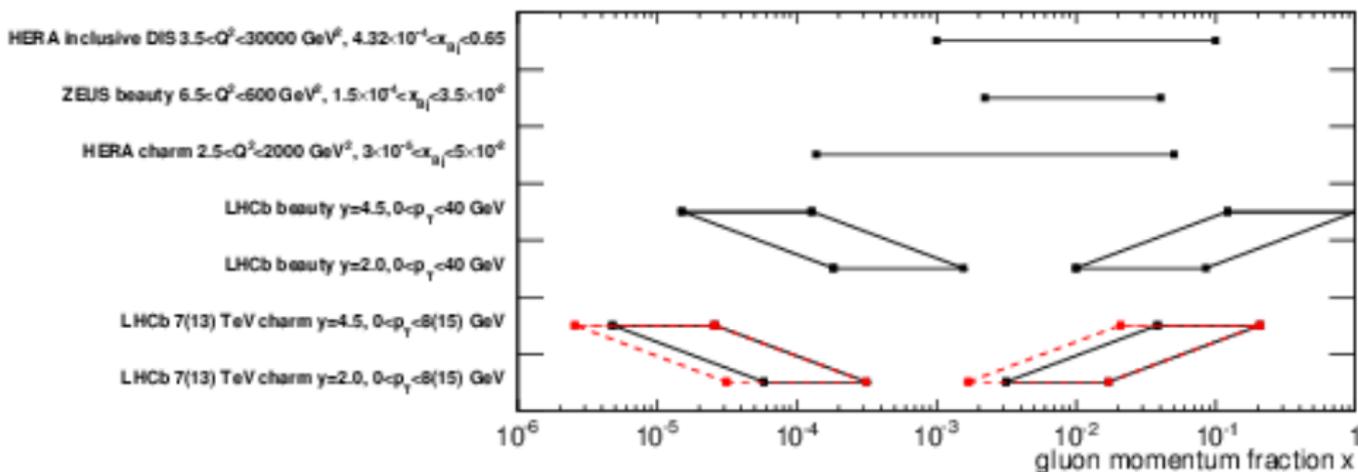
$\sigma^{exp}(W^-c) = 37.3 \pm 0.8 \pm 1.9 pb$   $\sigma^{theo}(W^-c)(ABMP) = 32.1 \pm 0.9 pb$

# Theory predictions vs. ATLAS experimental data from [arXiv:1402.6263] (PRELIMINARY)



cuts:  $W \rightarrow l\nu$ ,  $p_{T,D^0} > 8$  GeV

## x coverage of HERA and LHCb experiments

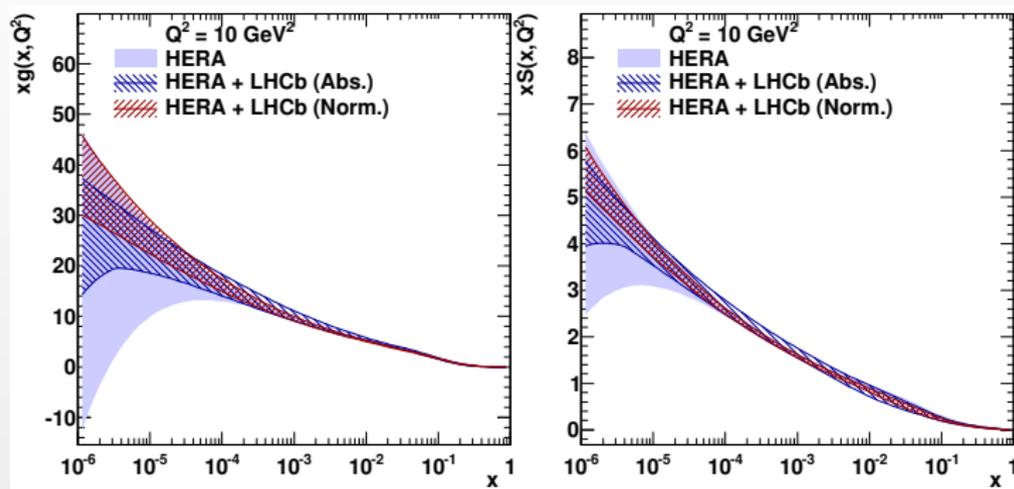


LHCb data allows to cover  $x$  regions uncovered by HERA data, both at low  $x$ 's (especially open charm data) and at large  $x$ 's (especially open bottom data).

Larger rapidities of the emitted quark and/or larger collision energies correspond to more extreme  $x$ 's

# PROSA PDF fit: comparison between three variants

from PROSA collab., EPJC 75 (2015) 471

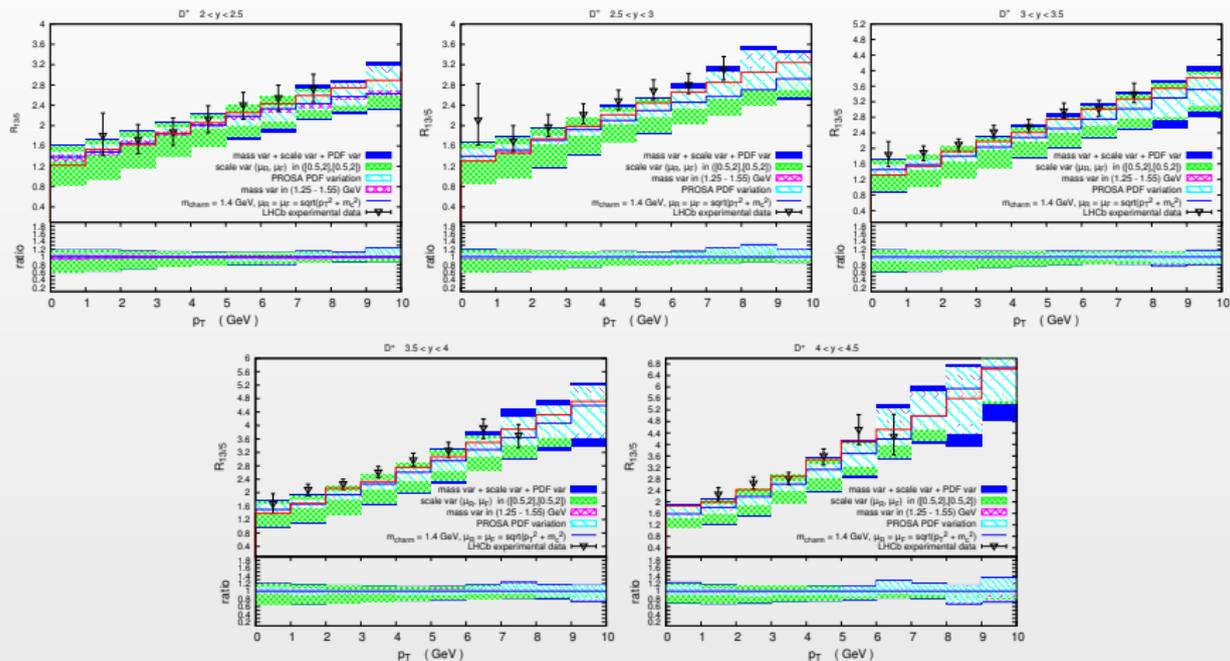


Three variants of the PDF fit:

- 1) one only with HERA data;
- 2) one also including LHCb absolute differential cross-sections;
- 3) another one with reduced uncertainties: for each fixed LHCb  $p_T$  bin, use the ratios of distributions  $(d\sigma/dy)/(d\sigma/dy_0)$  in different rapidity bins (i.e. normalized to the central bin  $3 < y_0 < 3.5$ ):

**in the ratios theoretical uncertainties partly cancel.**

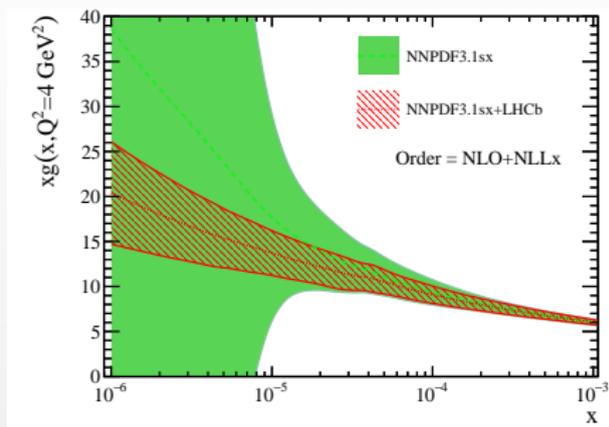
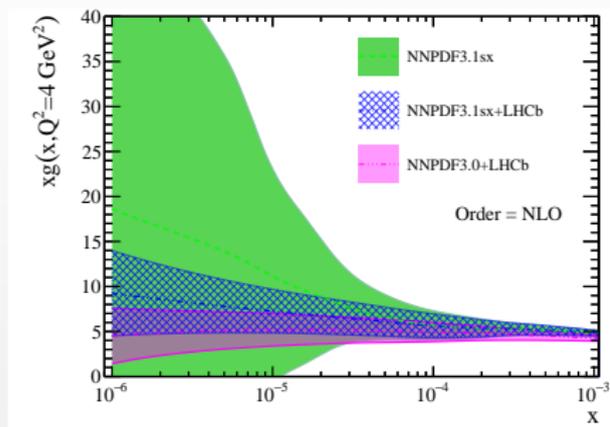
# Performances of PROSA PDF fit wrt to LHCb data on ratios of $p_T$ -distributions of $D^\pm$ at $\sqrt{s} = 13$ and 5 TeV



from PROSA collaboration, [arXiv:1812.02717]

These data are not included in the PROSA PDF fit.

# NNPDF fits reweighted by LHCb $D$ -meson data



from V. Bertone, R. Gauld and J. Rojo, [arXiv:1808.02034v2]

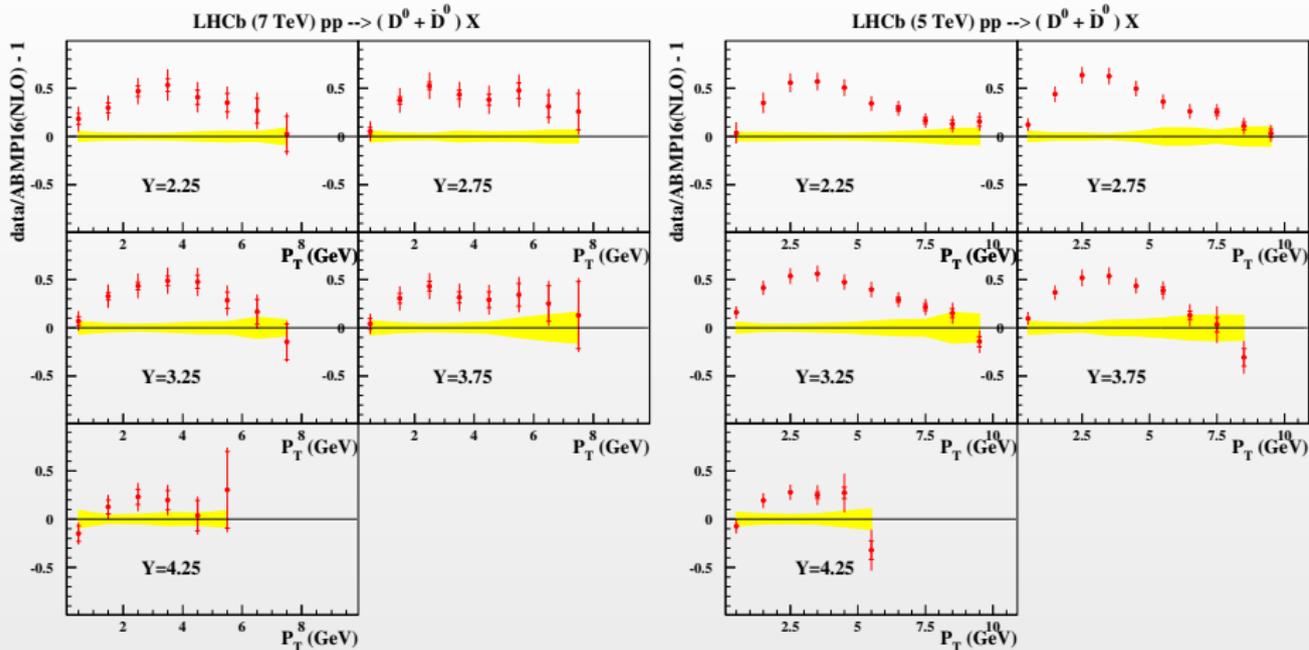
- \* gluon in the NNPDF3.1 + LHCb NLO PDF fit more positive but still compatible with the gluon in the NNPDF3.0 + LHCb NLO PDF fit.
- \* gluon in the fit partially including small- $x$  NLL resummation effects even more positive.

## Further investigations: considering other PDF fits

- \* Several **updated comprehensive PDF fits** are present on the market: ABMP16, CT14, NNPDF3.1, JR14, MMHT, CJ15....
- \* So far, besides the PROSA fit, only very few PDF fits have included the LHCb open-charm data. While the PROSA fit includes LHCb data from the very beginning, these fits have been reweighted a posteriori with LHCb data.
- \* **Further studies** are certainly **welcome!**
- \* In the following we consider the case of the ABMP16 PDF fit, which includes:
  - (HERA I + II) combined H1 + ZEUS inclusive NC + CC DIS data.
  - heavy-flavour DIS data.
  - Drell-Yan and top production data at the LHC.
  - higher-twist effects.

See *S. Alekhin, J. Bluemlein, S.O. Moch, R. Placakyte, PRD 96 (2017) 014011*  
, *EPJC 78 (2018) 477*

# Comparison data/theory for the $pp \rightarrow D^0 + \bar{D}^0 + X$ LHCb data

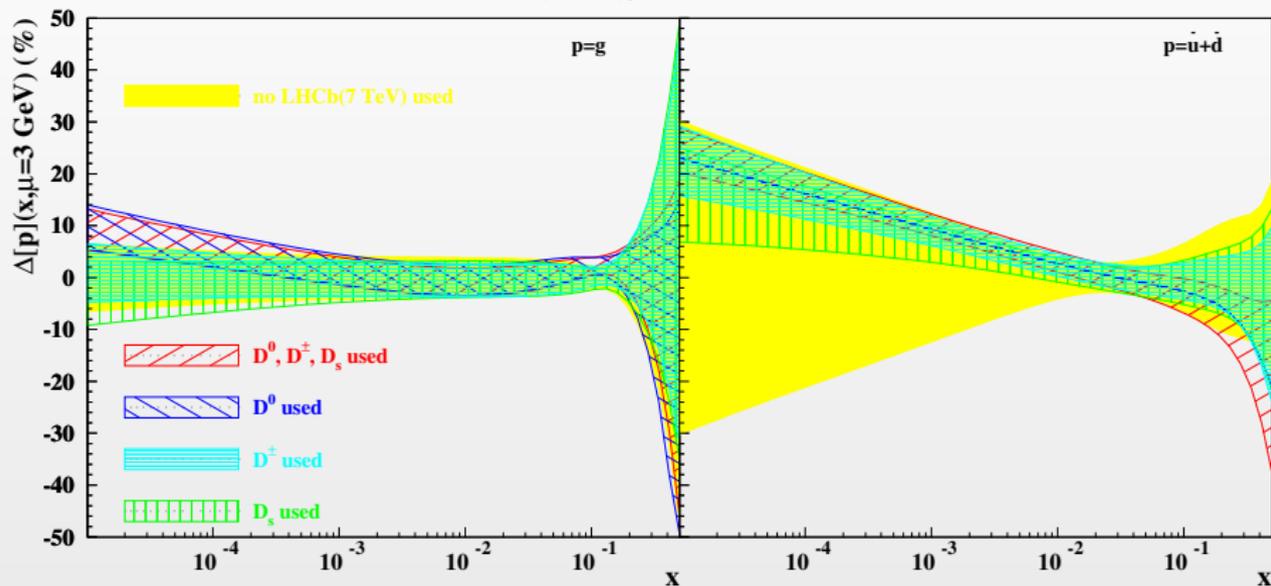


\* The pulls show a similar trend at different  $\sqrt{s}$ .

\* Data overestimate the theory predictions for  $p_T \sim 3 - 4 \text{ GeV}$ .

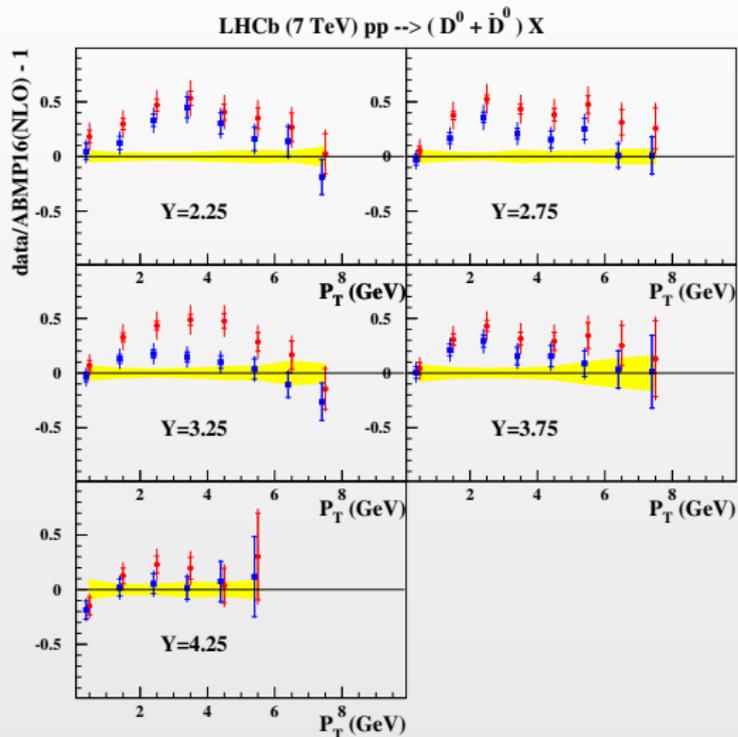
# Modifications of the ABMP16 gluon and sea PDFs after inclusion of the LHCb open-charm data at $\sqrt{s} = 7$ TeV

LHCb(7 TeV), HERA included



- \* Modifications of the fit driven by the inclusion of the  $D^0$  data.
- \* Slight increase of the gluon and sea quark PDFs at small  $x$ 's.

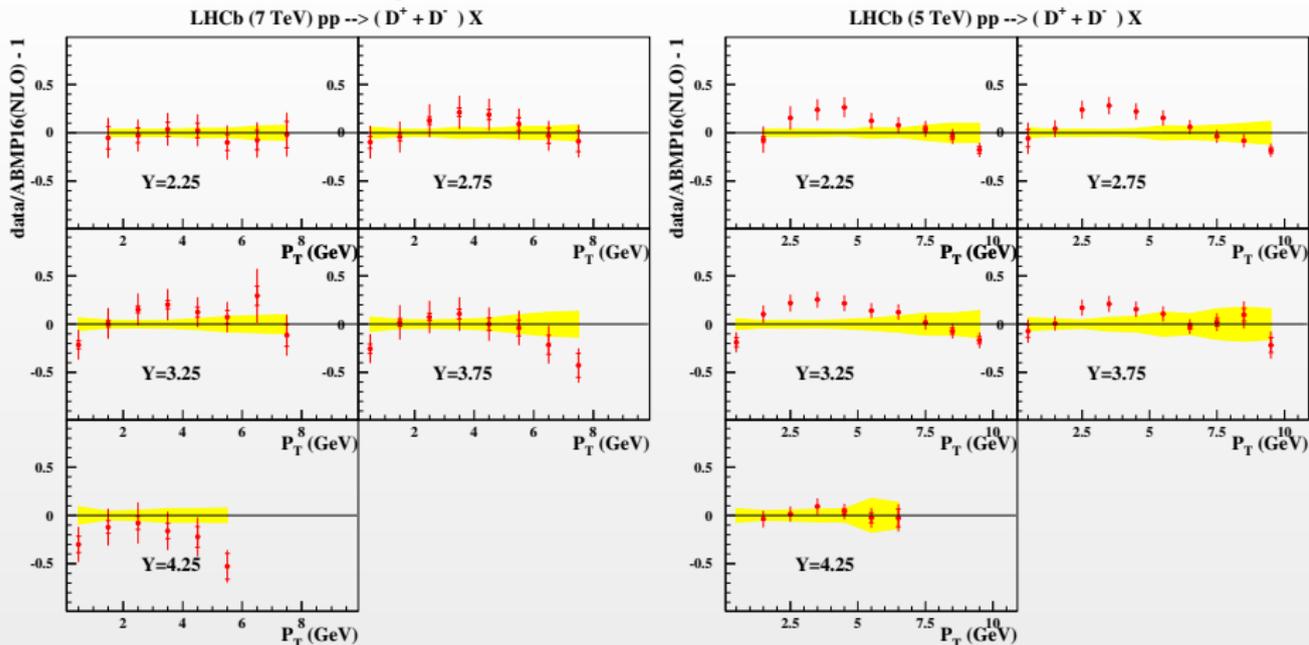
# Pulls for the $pp \rightarrow D^0 + X$ LHCb data at $\sqrt{s} = 7$ TeV



\* The ABMP16 PDF fit including the  $D^0$  LHCb data performs slightly better than the one before inclusion of these data.

\* But discrepancies are still present at small  $y$ .

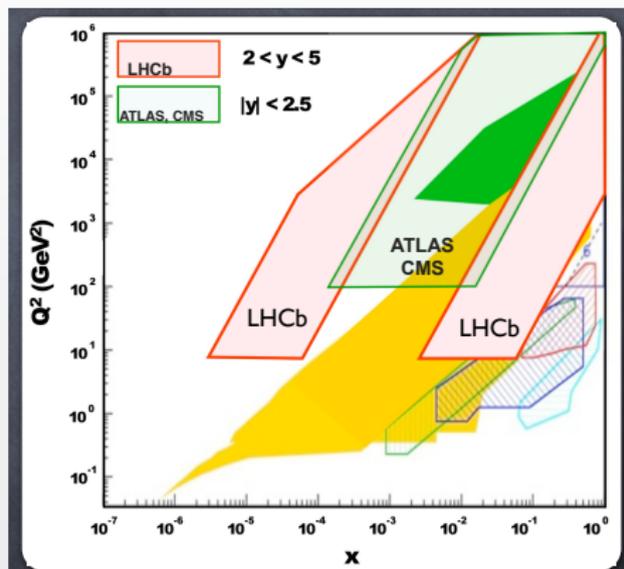
# Comparison data/theory for the $pp \rightarrow D^\pm + X$ LHCb data



\* Puzzle: at small rapidities the  $D^\pm$  data at  $\sqrt{s} = 7$  TeV turn out to be described better than those at 5 TeV, whereas we do not expect significant modifications of the physics: are the experimental data at different energies compatible among each other ?

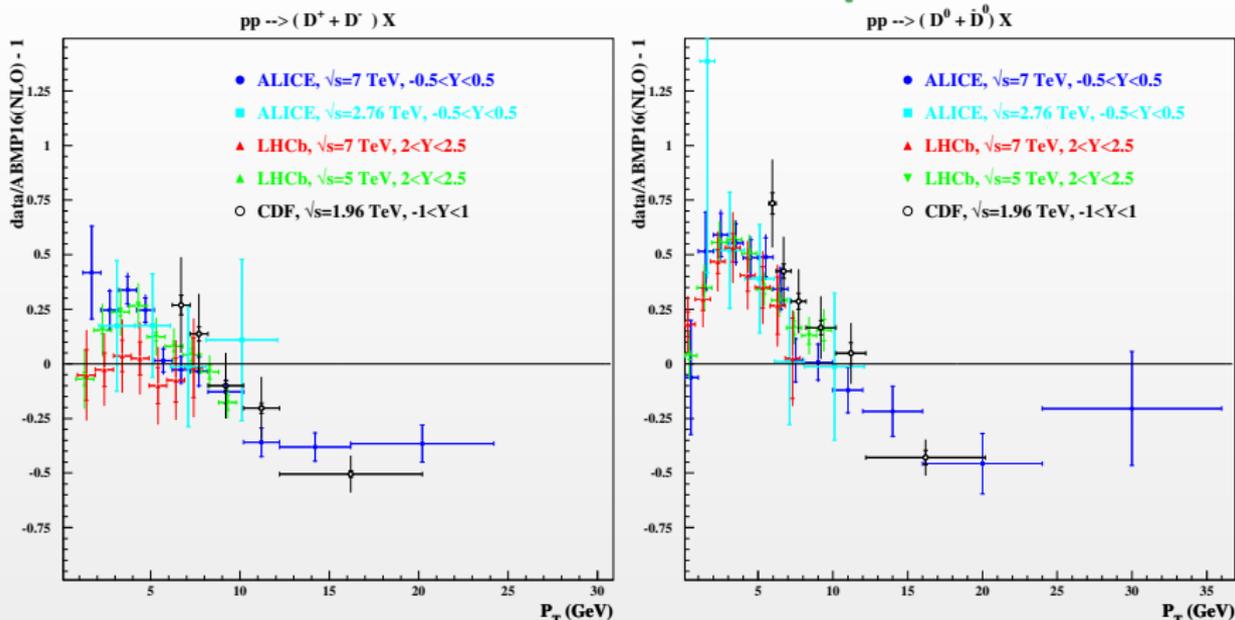
## Further investigations: adding data from other experiments

- \* LHCb open-charm data ( $2 < y < 4.5$ )
- \* ATLAS (and CMS) open-charm data ( $|y| < 2.5$ )
- \* CDF open-charm data ( $|y| < 1$ )
- \* ALICE open-charm data ( $|y| < 0.5$ )
- + further open-bottom data



Different experiments span ( $Q^2$ ,  $x$ ) regions partially overlapping: good for verifying their compatibility and for cross-checking their theoretical description.

# Pulls for the LHCb, ALICE, CDF open-charm data



\* Fluctuations for  $D^\pm$ , while a trend is visible for  $D^0$ .

\* In case of  $D^0$ , data at a fixed  $p_T$  seem to be reproduced similarly well/bad, independently of the  $\sqrt{s}$  and of the  $y$  probed.

\* This implies that the difference in shape between theory predictions and exp. data can not be washed out by modifying PDFs at low  $x$ 's.

## Temptative explanations of the discrepancies data/theory at $p_T \sim$ a few GeV

- \* PDF effect ?

No, because for the  $D^0$  we see similar discrepancies at different  $\sqrt{s}$ .

- \* FFNS vs. VFNS effect ?

Unlikely to be so large for  $p_T \sim$  few GeV, see the next slides.

- \* Fragmentation Function / Hadronization effect ?

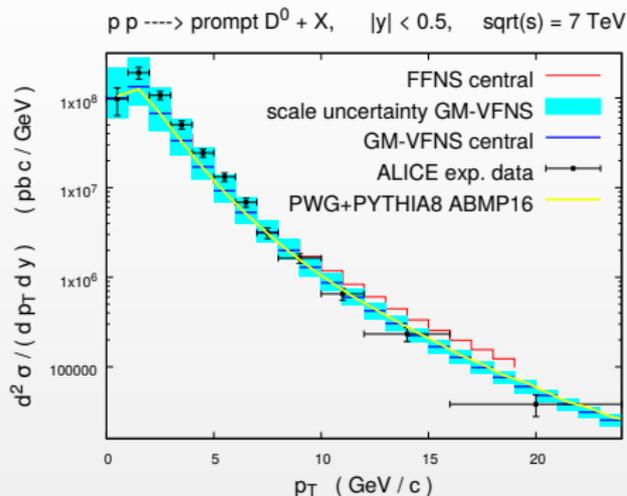
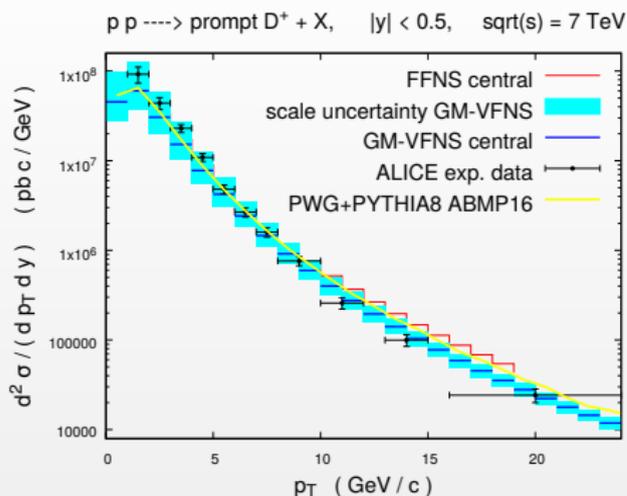
Unlikely to be so large, see the next slides.

- \* Higher-order effects ?

Possible modification of shapes, because scale variation uncertainty is very large.

- \* Alternatively, one could think to a combination of small effects.

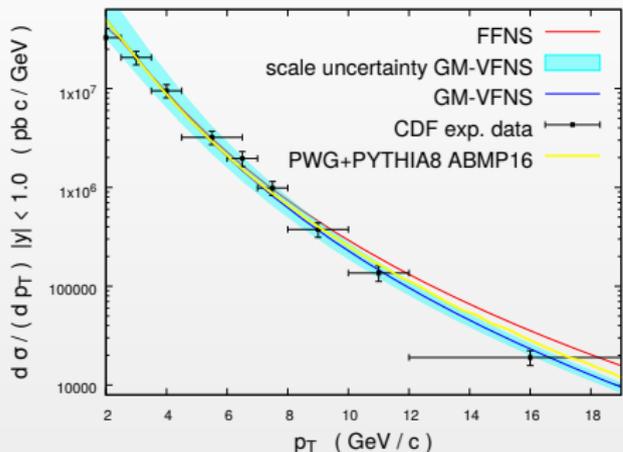
# Theory predictions vs. ALICE exp. data at $\sqrt{s} = 7$ TeV



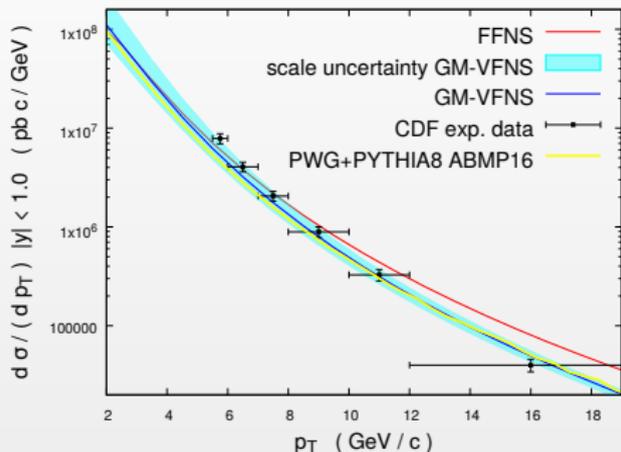
- \* GM-VFNS + FF predictions and POWHEGBOX + PYTHIA predictions compatible among each other up to  $p_T \sim 20$  TeV.
- \* Very different treatment of fragmentation.
- \* FFNS + FF predictions overestimate data at large  $p_T$ .
- \* POWHEGBOX + PYTHIA agrees better with  $D^\pm$  than with  $D^0$  data.

# Theory predictions vs. CDF exp. data at $\sqrt{s} = 1.96$ TeV

p anti-p  $\rightarrow$  prompt  $D^+ + X$ ,  $|y| < 1.0$ ,  $\sqrt{s} = 1.96$  TeV



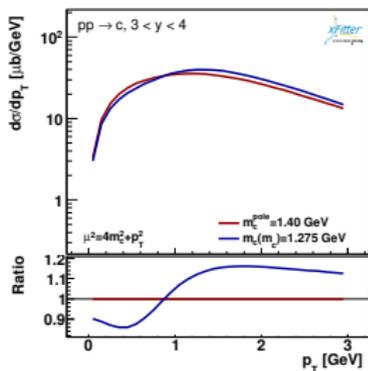
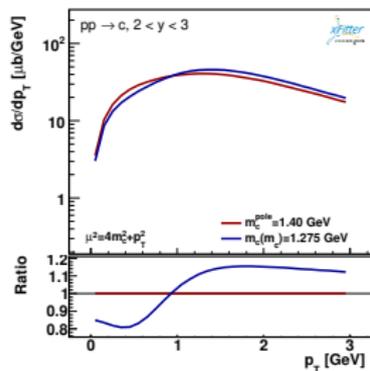
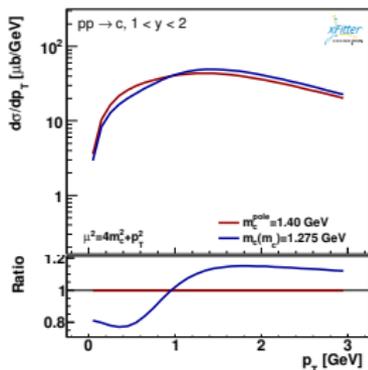
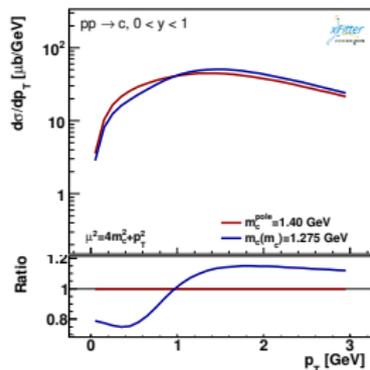
p anti-p  $\rightarrow$  prompt  $D^0 + X$ ,  $|y| < 1.0$ ,  $\sqrt{s} = 1.96$  TeV



\* Similar observations as before.

\* Analogous picture for other experiments...

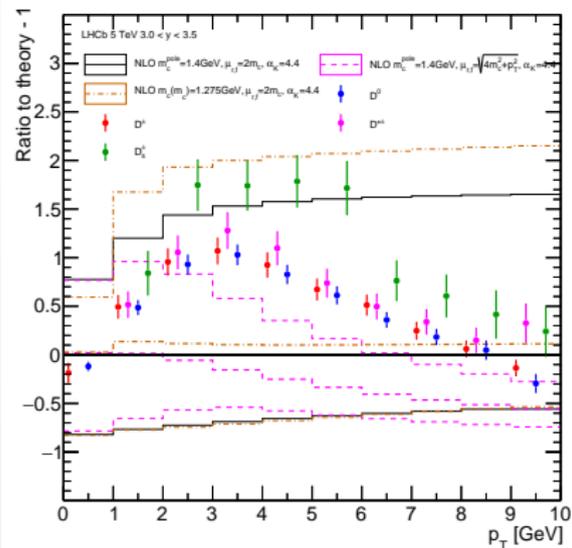
# Theory predictions with different charm mass renormalization schemes



Shape modification when using  $m_c^{\text{MSbar}}$  instead of  $m_c^{\text{pole}}$  is within 20%.

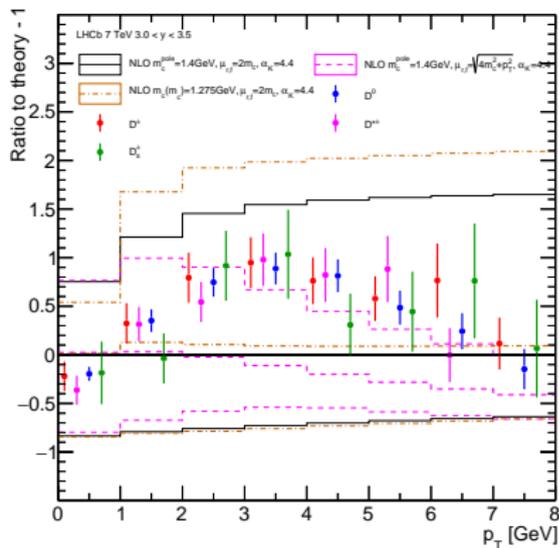
# Comparison data/theory for the $pp \rightarrow D^\pm, D^0, D_s^\pm, D^{*,\pm} + X$

## LHCb data



$$\sqrt{s} = 5\text{TeV}$$

\* Small shape modification when using  $m_c^{\text{MSbar}}$  instead of  $m_c^{\text{pole}}$ .



$$\sqrt{s} = 7\text{TeV}$$

## $m_c$ values from PROSA PDF fit

\* Heavy-quark masses are fitted simultaneously with PDFs in the PROSA PDF fit.

\* The PROSA fit with NLO theory predictions on the basis of  $m_c^{pole}$  leads to:

$$m_c^{pole} = 1.265 \pm 0.067 \text{ GeV}, \quad m_b^{pole} = 4.178 \pm 0.140 \text{ GeV}$$

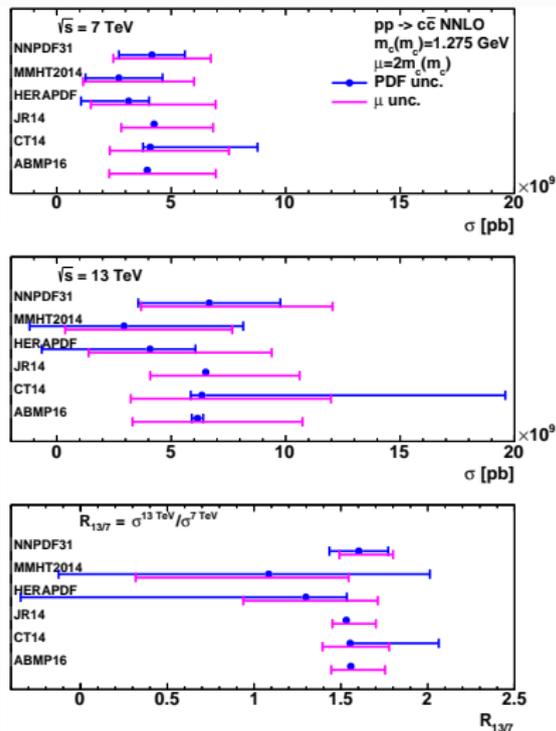
\* The PROSA fit with NLO theory predictions on the basis of  $m_c^{MSbar}$  leads to:

$$m_c(m_c) = 1.170 \pm 0.048 \text{ GeV}, \quad m_b(m_b) = 3.983 \pm 0.249 \text{ GeV}$$

\* The latter values are in more reasonable agreement with running mass extractions quoted in PDG:

$$m_c(m_c) = 1.27 \pm 0.03 \text{ GeV}, \quad m_b(m_b) = 4.18 + 0.04 - 0.03 \text{ GeV}$$

# Total NNLO $\sigma(pp \rightarrow c\bar{c})$ for different PDF sets



\* Scale uncertainties larger than 50% even at NNLO.

\* Consistency between PDF sets in the values of the ratio  $R_{13/7}$  within uncertainties.

## Summary and Conclusions

- \* LHCb open charm and bottom data have the potentiality to constrain gluon and sea quark PDF at low and large  $x$ 's.
- \* Incorporation in PDF fits so far limited to very few cases (PROSA, recent NNPDF variants, ABMP preliminary, nCTEQ15).
- \* Compatibility with other open charm and bottom data under investigation.
- \* Theory predictions (and PDF fits) plagued by large scale uncertainty.
- \* Similar uncertainties when using  $\overline{\text{MS}}$  scheme for charm mass renormalization, but more reasonable charm mass values from PROSA fit.
- \* Still under investigation: how to reconcile shapes of absolute distributions with experimental data ?
- \* Ratios of theory predictions wash out differences which, however, still need to be clarified!

# Back-up slides

## Experimental data used in $p$ PDF fits: examples

PDF sets	$\Delta\chi^2$ criterion	data sets used in analysis
ABM12	1	incl. DIS, DIS charm, DY
CJ15	1	incl. DIS, DY (incl. $p\bar{p} \rightarrow W^\pm X$ ), $p\bar{p}$ jets, $\gamma$ +jet
CT14	100	incl. DIS, DIS charm, DY, $p\bar{p}$ jets, $pp$ jets
HERAPDF2.0	1	incl. DIS, DIS charm, DIS jets [only HERA data]
JR14	1	incl. DIS, DIS charm, DY, $p\bar{p}$ jets, DIS jets
MMHT14	2.3 ... 42.3 (dynamical)	incl. DIS, DIS charm, DY, $p\bar{p}$ jets, $pp$ jets, $t\bar{t}$
NNPDF3.0	n.a.	incl. DIS, DIS charm, DY, $p\bar{p}$ jets, $pp$ jets, $t\bar{t}$ , $W + c$

**Table :** Summary of major hard processes used in the various PDF analyses and the confidence level criteria employed.

**Different PDF fits differ for the use of different data sets.**

# Flavour Number Schemes

$m_u, m_d, m_s = 0$  always,  $m_t \neq 0$  always

$m_c, m_b$  treatment depends on scheme

And, depending on the kinematics and  $E_{cm}$ , it may happen that  $p_{T,c} \gg m_c, m_b$

## \* **Fixed flavour number scheme (FFNS):**

- $c$  ( $b$ ) masses are retained at all scales in all components of the calculation.

These quarks can only be produced as final states.

They are excluded from initial states.

**problem** at high  $p_T$ :  $\log(p_{T,c}^2/m_c^2)$  ( $\log(p_{T,b}^2/m_b^2)$ ) may become so big that they may spoil the convergence of the perturbative series!

## \* **Zero-mass variable flavour number scheme (ZM-VFNS):**

- $c$  ( $b$ ) massless quarks at all scales in all components of the calculation.

These quarks are present in the initial state above fixed thresholds.

**problem** at low  $p_T$ : powers of  $(m_c^2/p_{T,c}^2)$  missing!

## \* **General-mass variable flavour number schemes (GM-VFNS):**

- $c$  ( $b$ ) mass retained in part of the calculation;

meant to combine optimal features of FFNS and ZM-VFNS at different  $p_T$ .

**advantage:** logs resummed and powers (at least the leading ones) present.

**problem:** some arbitrariness in the combination (different variants possible)

# GM-VFNS

- Combining the ZM-VFN (high  $p_T$ ) and FFN (small  $p_T$ ) schemes yields the GM-VFN scheme [Kniehl, Kramer, Schienbein, Spiesberger]
- Combine massive and massless results and subtract terms to avoid double counting
- Radiative corrections give rise to IR divergences which cancel in the sum of virtual and real diagrams
- In the massive calculation there remain finite terms including some containing  $\log(m^2/s)$
- These logs correspond to the  $\log(\mu_{I/F}^2/s)$  of the massless calculation
- Also, taking the limit  $m \rightarrow 0$  of the massive result does not reduce to the massless one (dimreg and finite mass regulators yield different finite terms)

$$\lim_{m \rightarrow 0} d\sigma_{\text{FFN}} = d\sigma_{\text{ZM}}(\mu_I = \mu_F = m) + d\sigma_{\text{sub}}$$

→ Subtract these terms in the combination

$$\frac{d\sigma}{dp_T dy} = \frac{d\sigma_{\text{FFN}}}{dp_T dy} - \lim_{m \rightarrow 0} \frac{d\sigma_{\text{FFN}}}{dp_T dy} + \frac{d\sigma_{\text{ZM}}}{dp_T dy}$$