

Extraction of π^\pm and K^\pm fragmentation functions at NNLO

Valerio Bertone

IRFU, CEA, Université Paris-Saclay

In collaboration with R. Abdul Khalek, A. Chiefa and E. R. Nocera

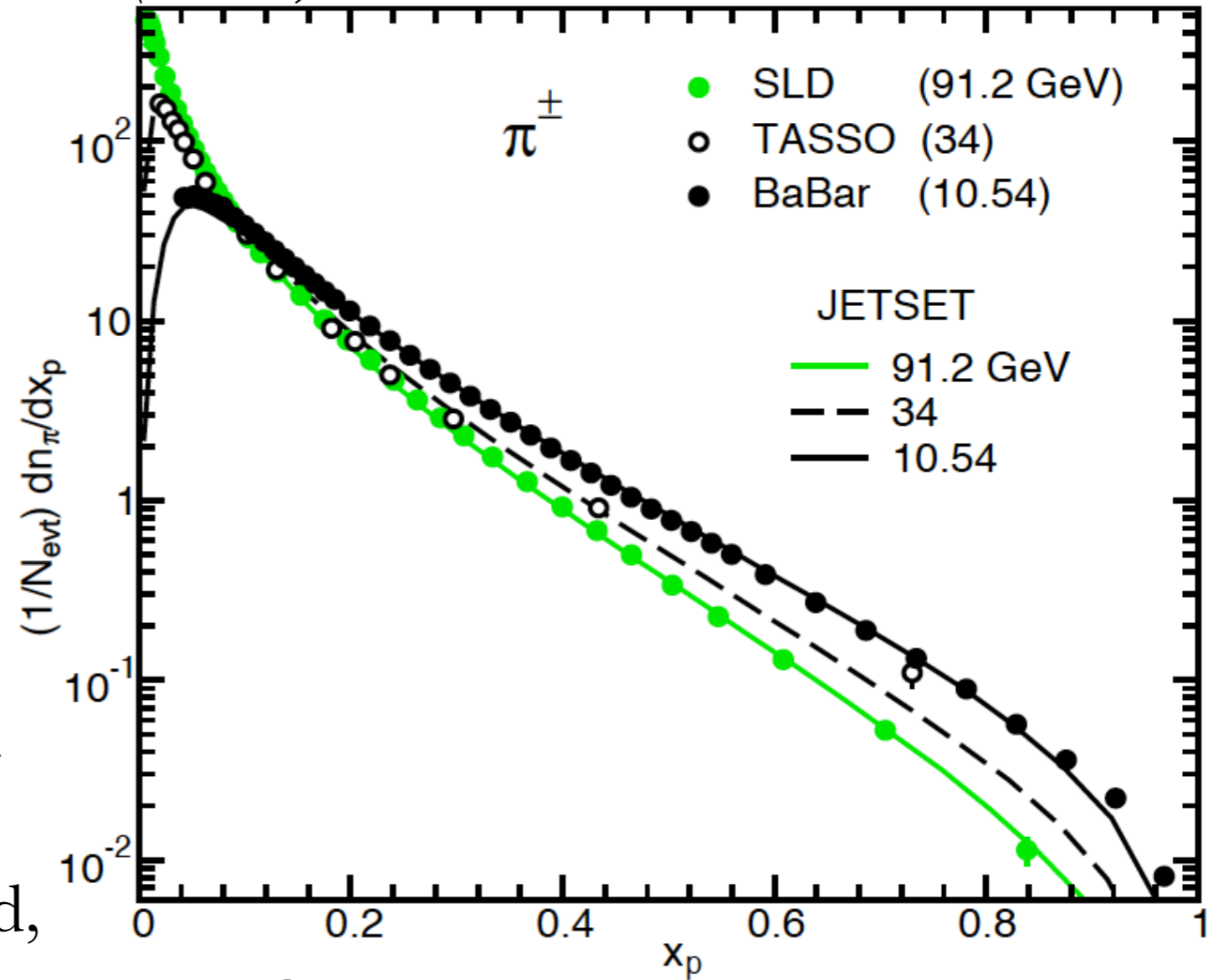
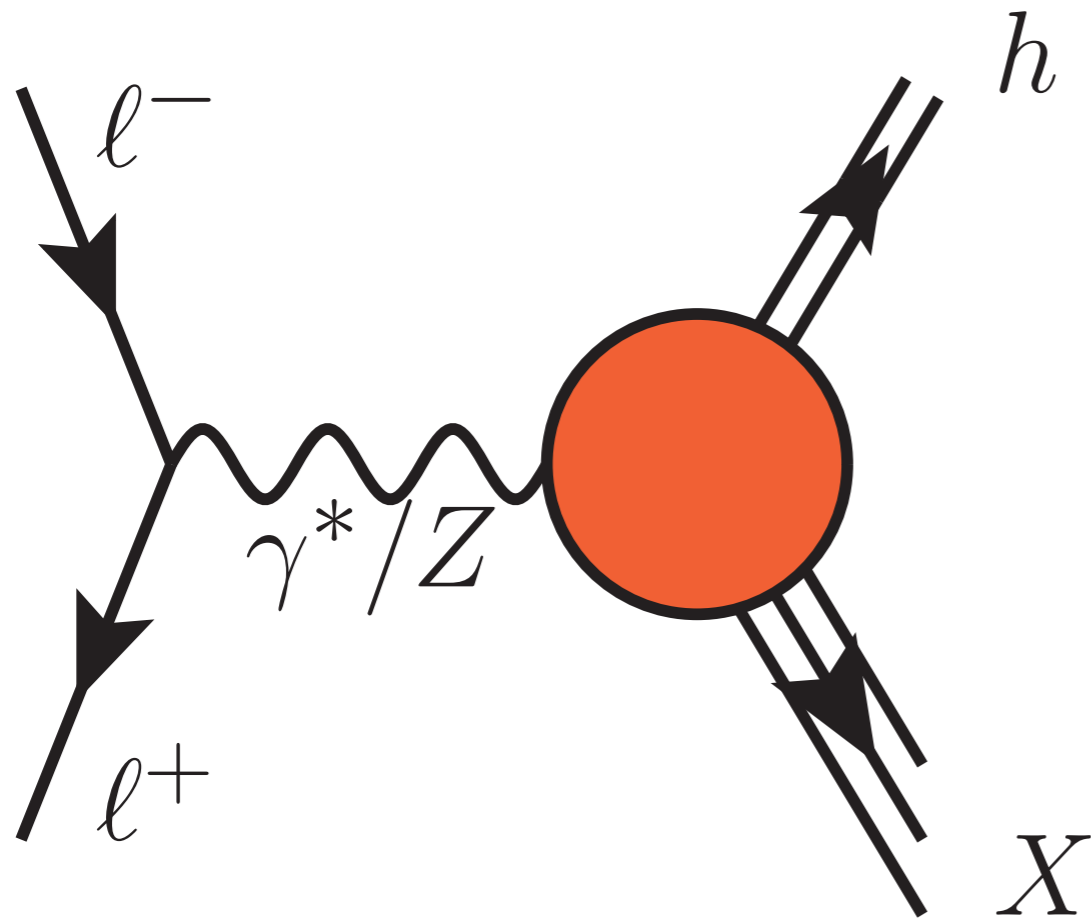
université
PARIS-SACLAY



October 1, 2024, IWHSS 2024, Yerevan, Armenia

Experimental data overview

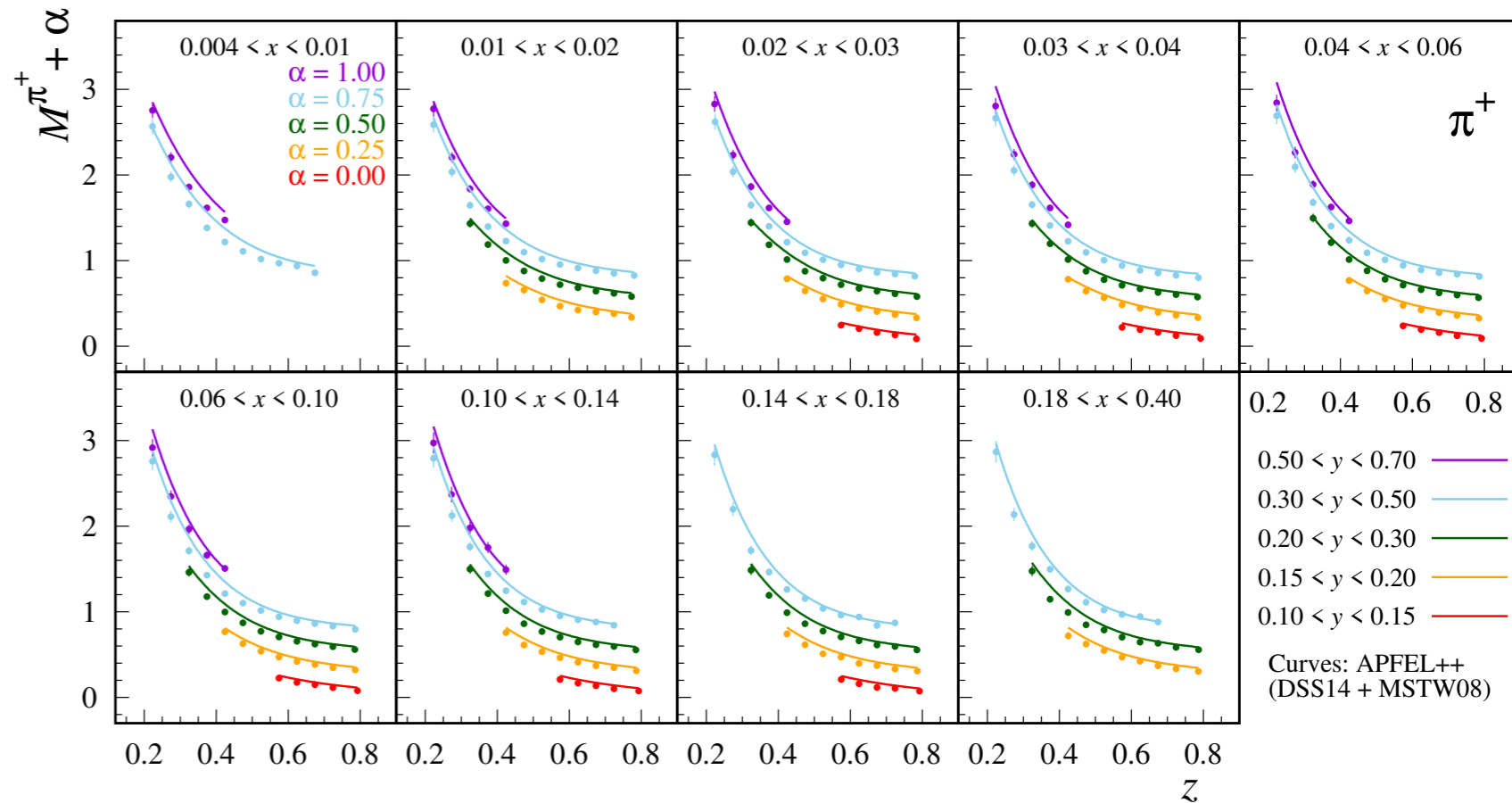
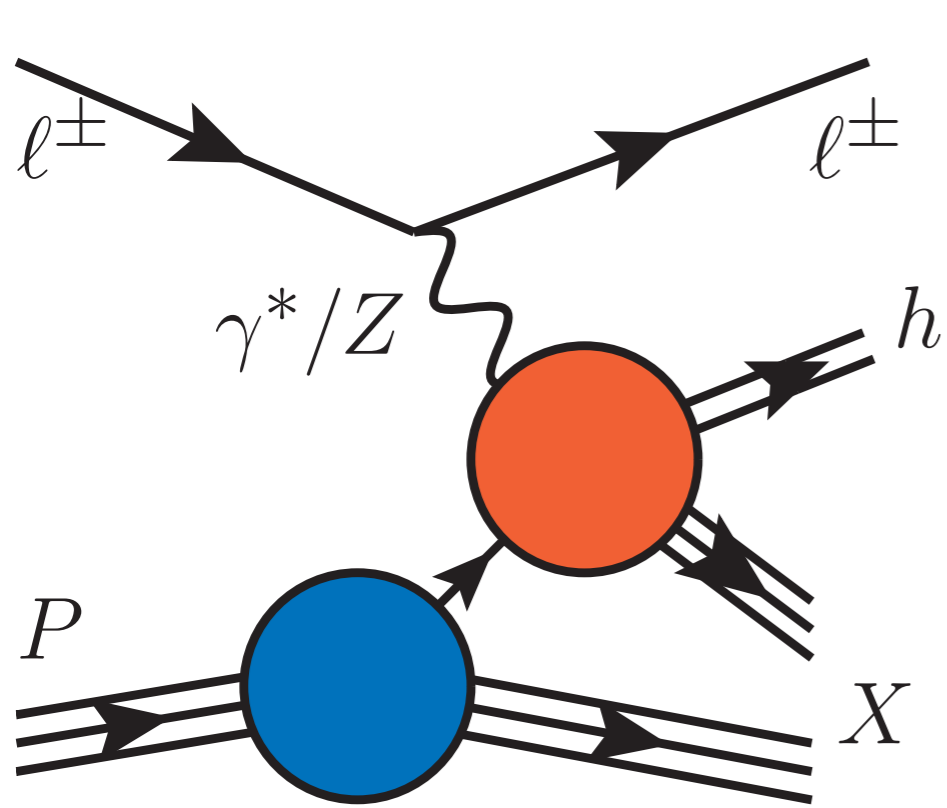
Single-inclusive annihilation (SIA)



- ✓ **Clean** channel: only FFs involved,
- ✓ **higher-order** corrections known to NNLO,
- ✓ **precise data** available (BELLE/BABAR).
- ✗ **No flavour separation,**
 - ✓ tagged data for heavy-quark FFs.
- ✗ Scarcely sensitive to **gluon distribution.**

Experimental data overview

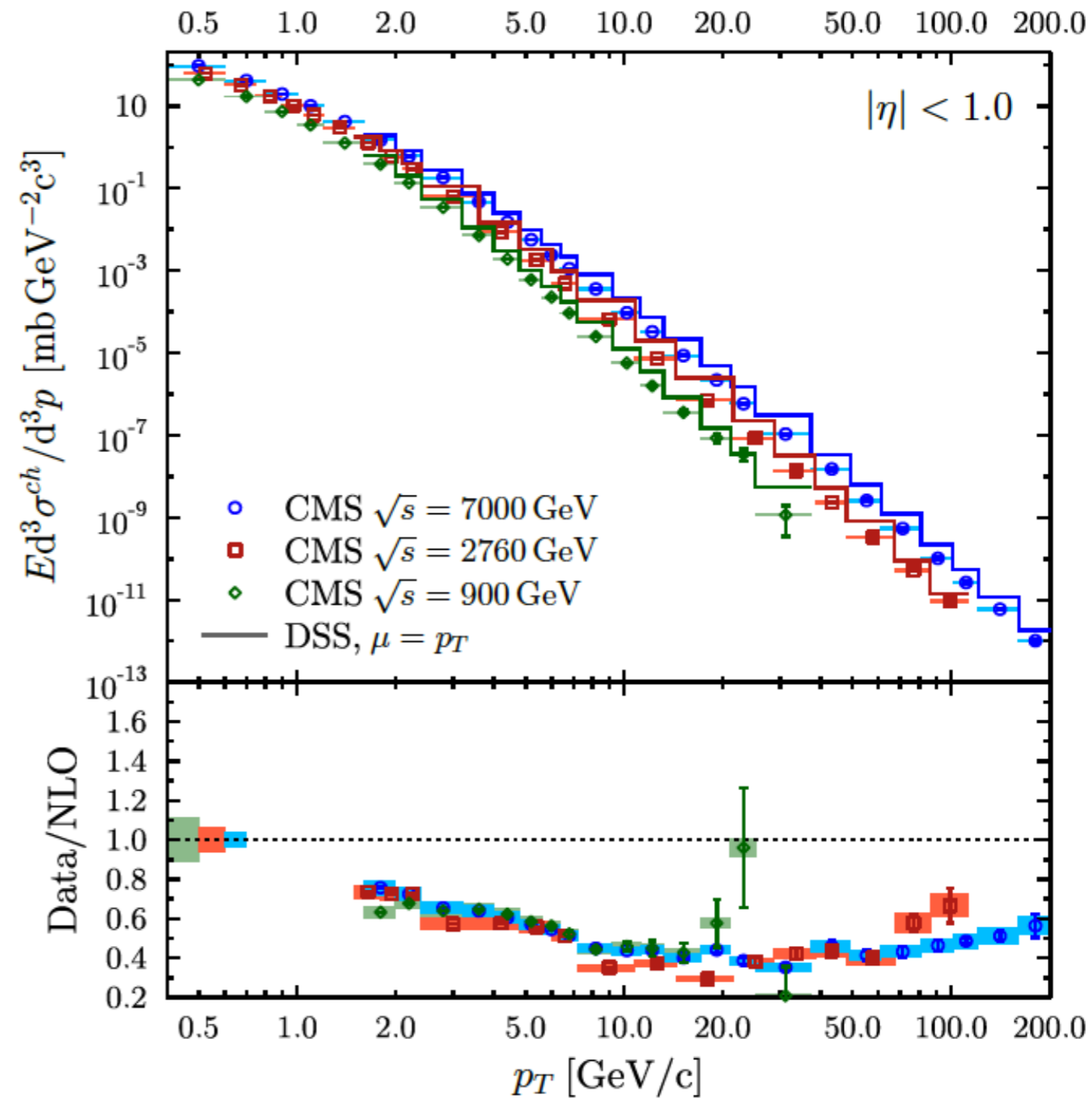
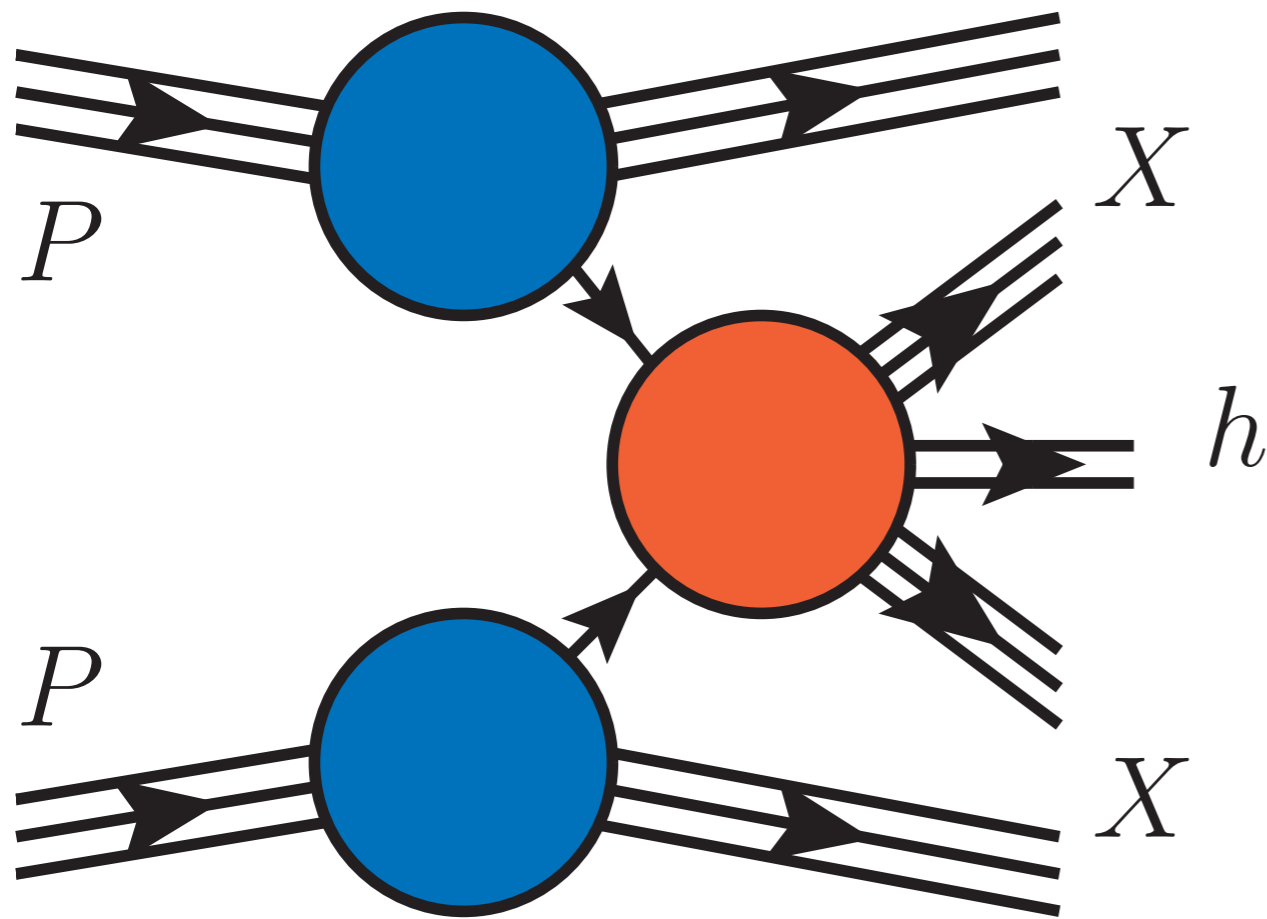
Semi-inclusive deep-inelastic scattering (SIDIS)



- Handle on **flavour separation**,
- precise data** available (HERMES/COMPASS),
- involves both **FFs** and **PDFs**,
- since very recently fully known to $O(\alpha_s^2)$, *i.e.* NNLO:
 - not ready yet to be used in a fit (working on a benchmark and implementation),
 - approximated NNLO have been used in fits so far.

Experimental data overview

Inclusive production of a hadron in pp collisions

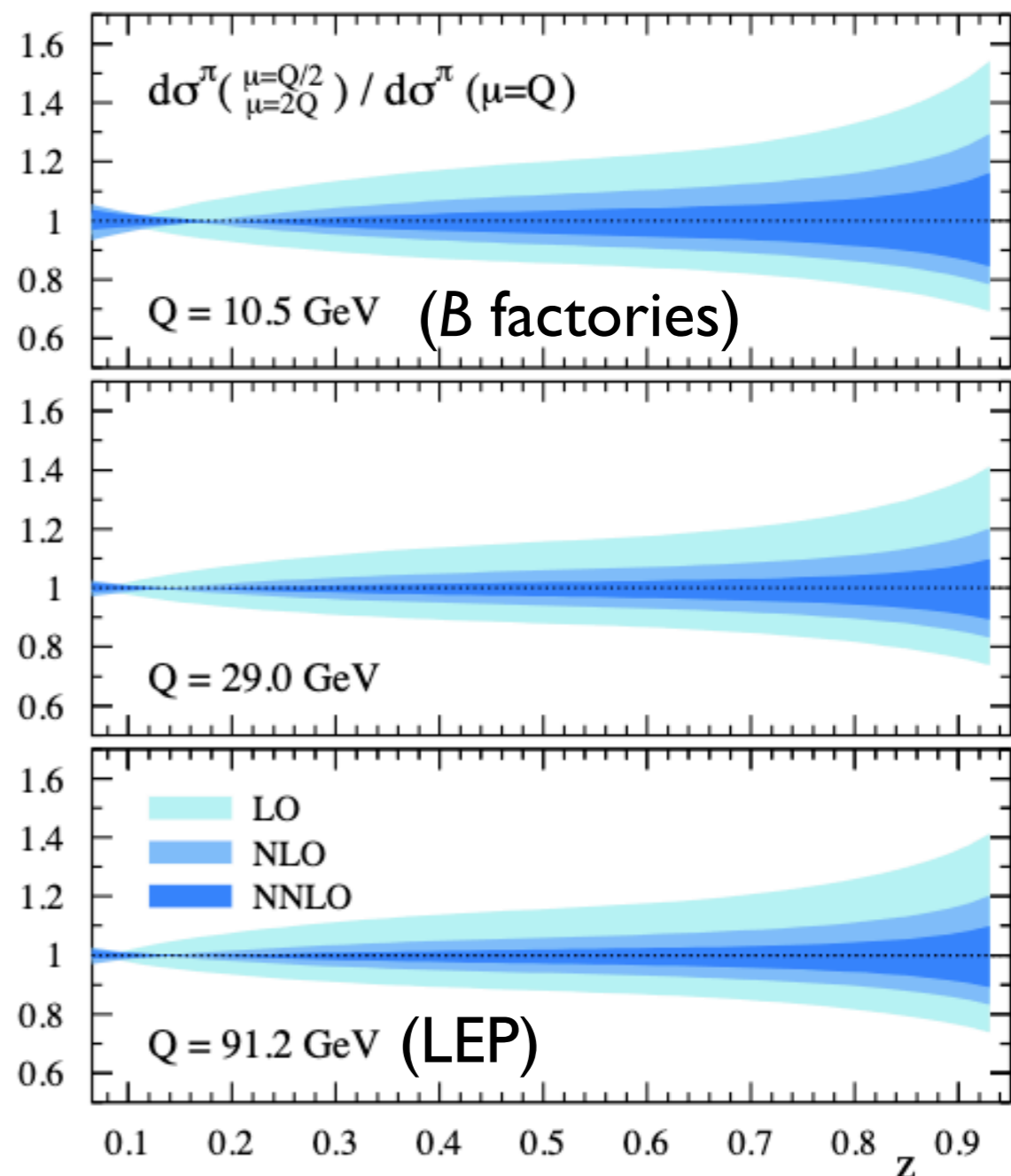
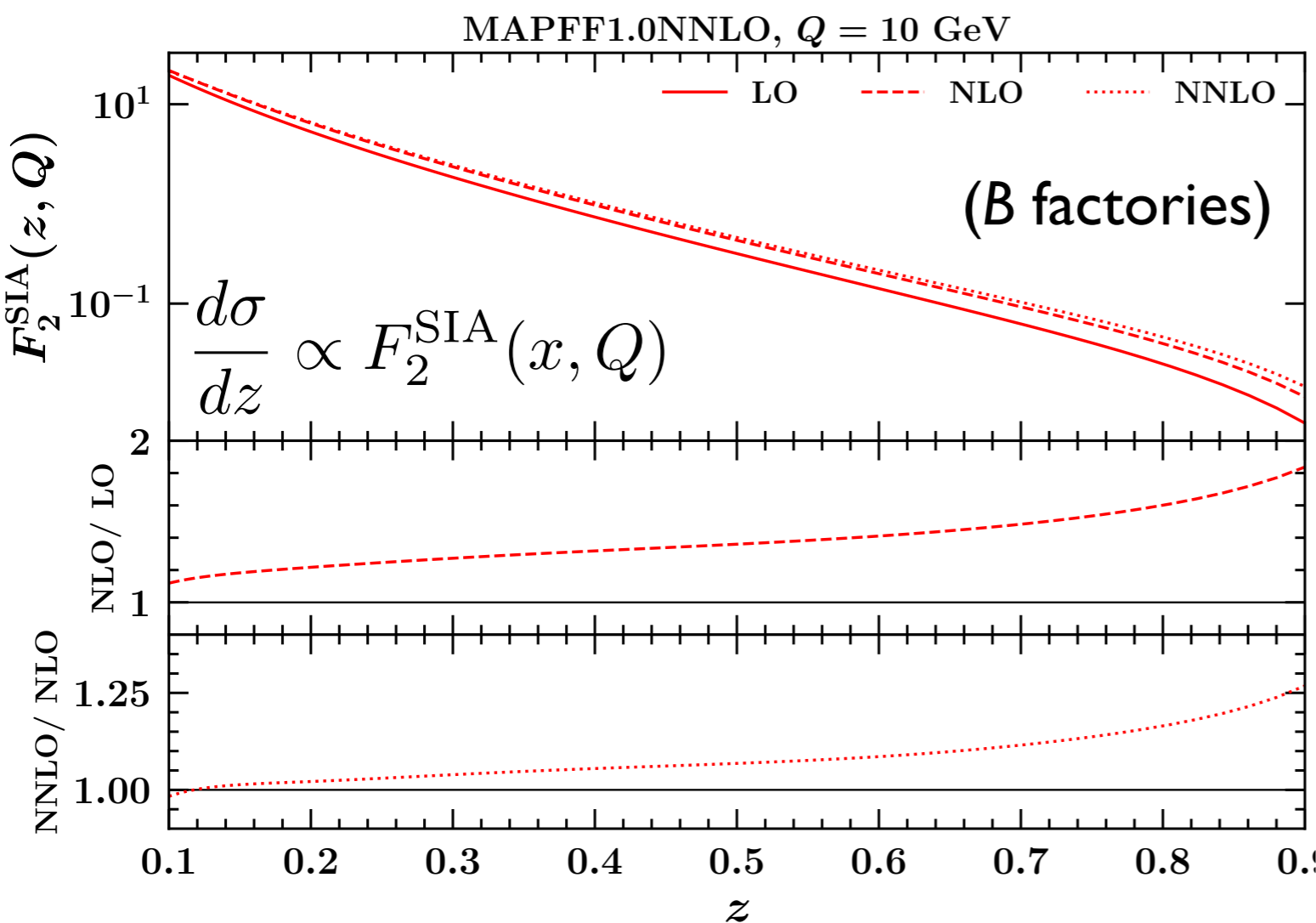


- Direct sensitivity to the **gluon FF**,
- precise data from LHC/Tevatron,
- involves both **FFs** and **PDFs**,
- publicly known so far up to NLO,
- large scale variations at low p_T ,
- cannot be used for NNLO fits yet.**

Impact of NNLO corrections

SIA structure functions

D. Anderle *et al.* [arXiv:1510.05845]



The SIA structure functions show a good **perturbative convergence**:

- NNLO K-factors are typically much smaller than NLO ones,
- however, **NNLO** corrections can be **significant** (up to 15-20%),
- scale-variation bands consistently shrink going from LO to NNLO.

Impact of NNLO corrections

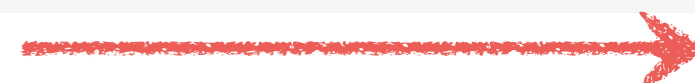
SIA structure functions

V. Bertone *et al.* [arXiv:1706.07049]

Perturbative corrections to SIA improve the agreement with data:

- **perturbative convergence,**
- NLO corrections necessary,
- NNLO beneficial.

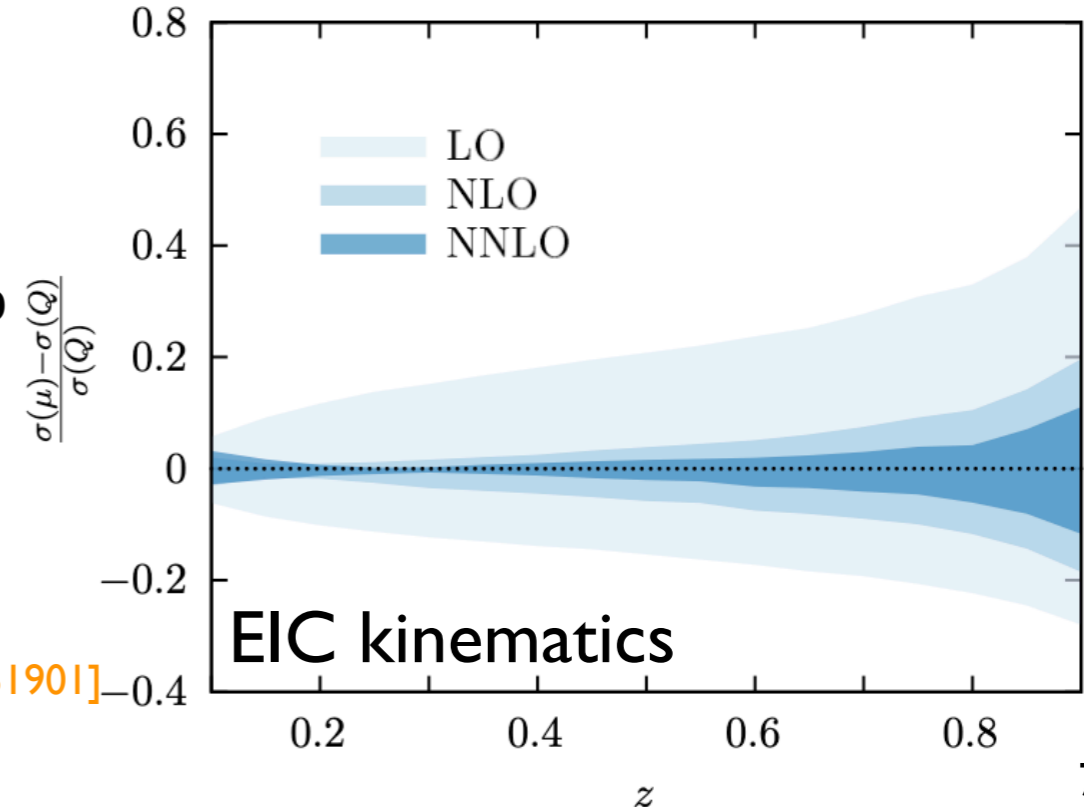
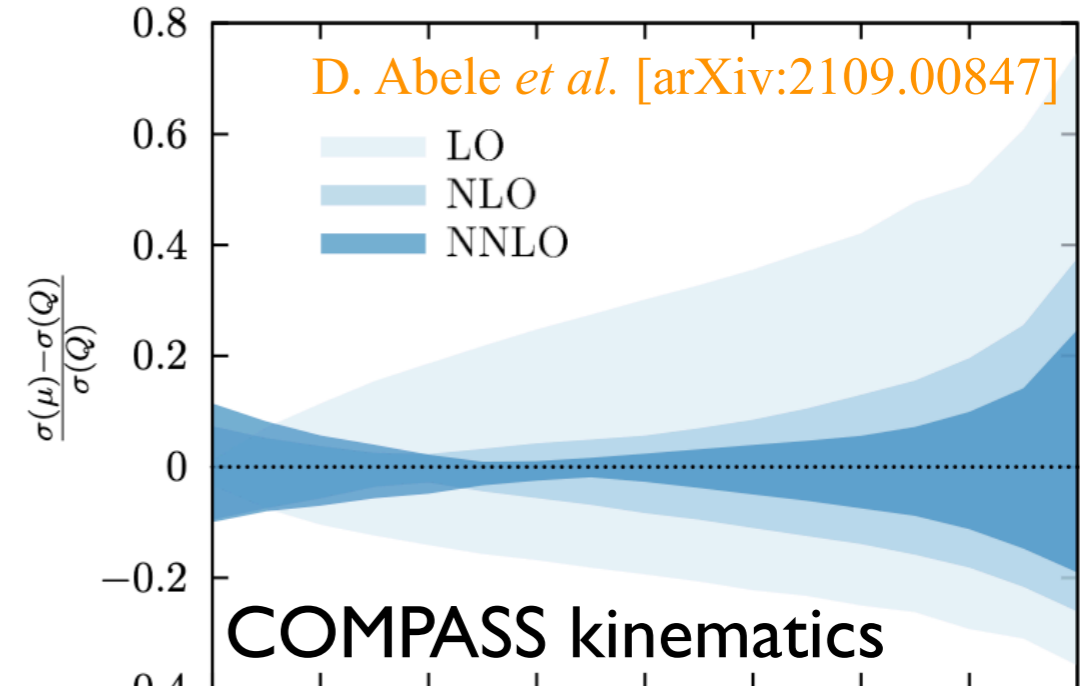
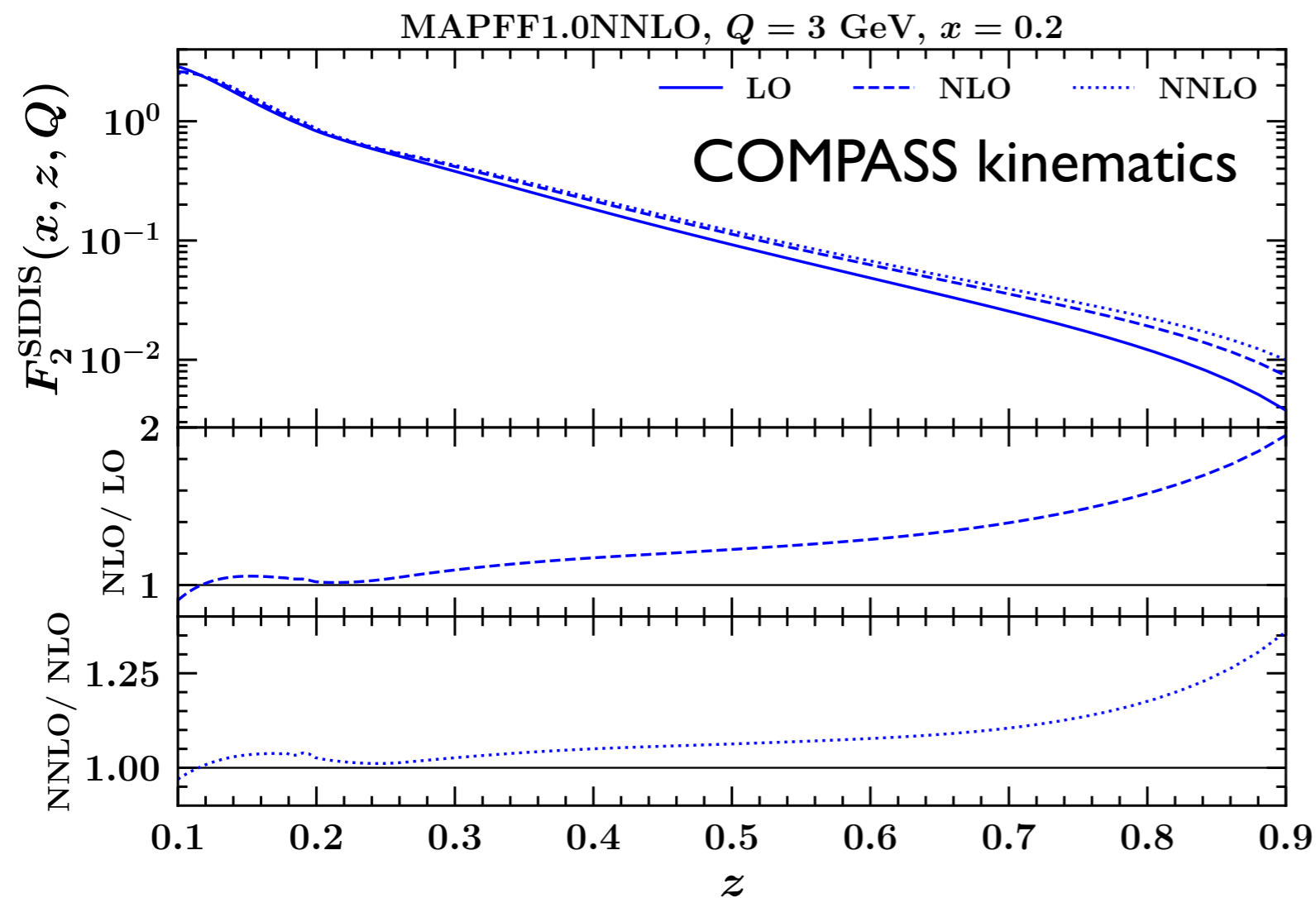
Exp.	N_{dat}	LO χ^2/N_{dat}	NLO χ^2/N_{dat}	NNLO χ^2/N_{dat}
BELLE	70	0.60	0.11	0.09
BABAR	40	1.91	1.77	0.78
TASSO12	4	0.70	0.85	0.87
TASSO14	9	1.55	1.67	1.70
TASSO22	8	1.64	1.91	1.91
TPC	13	0.46	0.65	0.85
TPC-UDS	6	0.78	0.55	0.49
TPC-C	6	0.55	0.53	0.52
TPC-B	6	1.44	1.43	1.43
TASSO34	9	1.16	0.98	1.00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



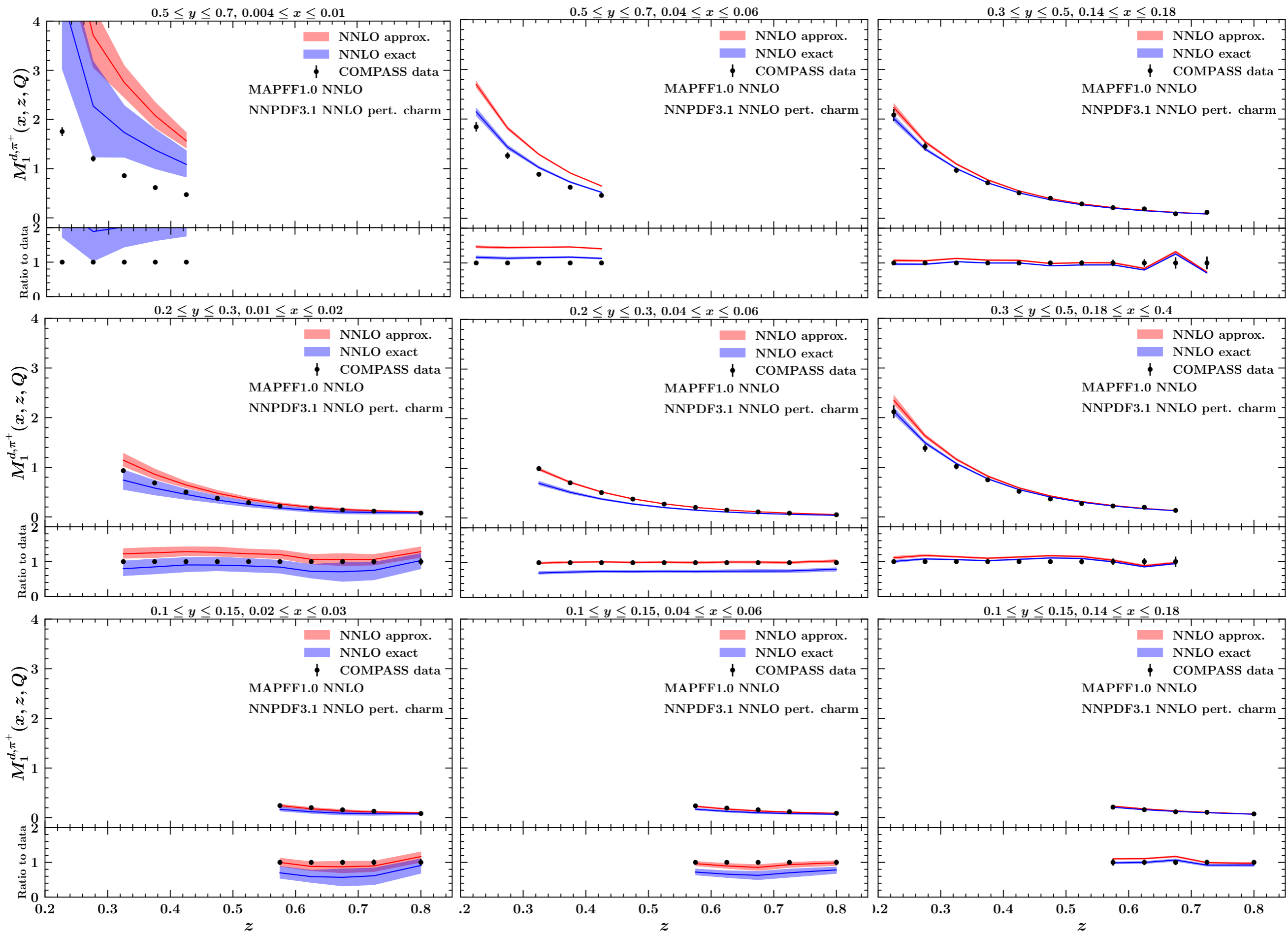
Impact of NNLO corrections

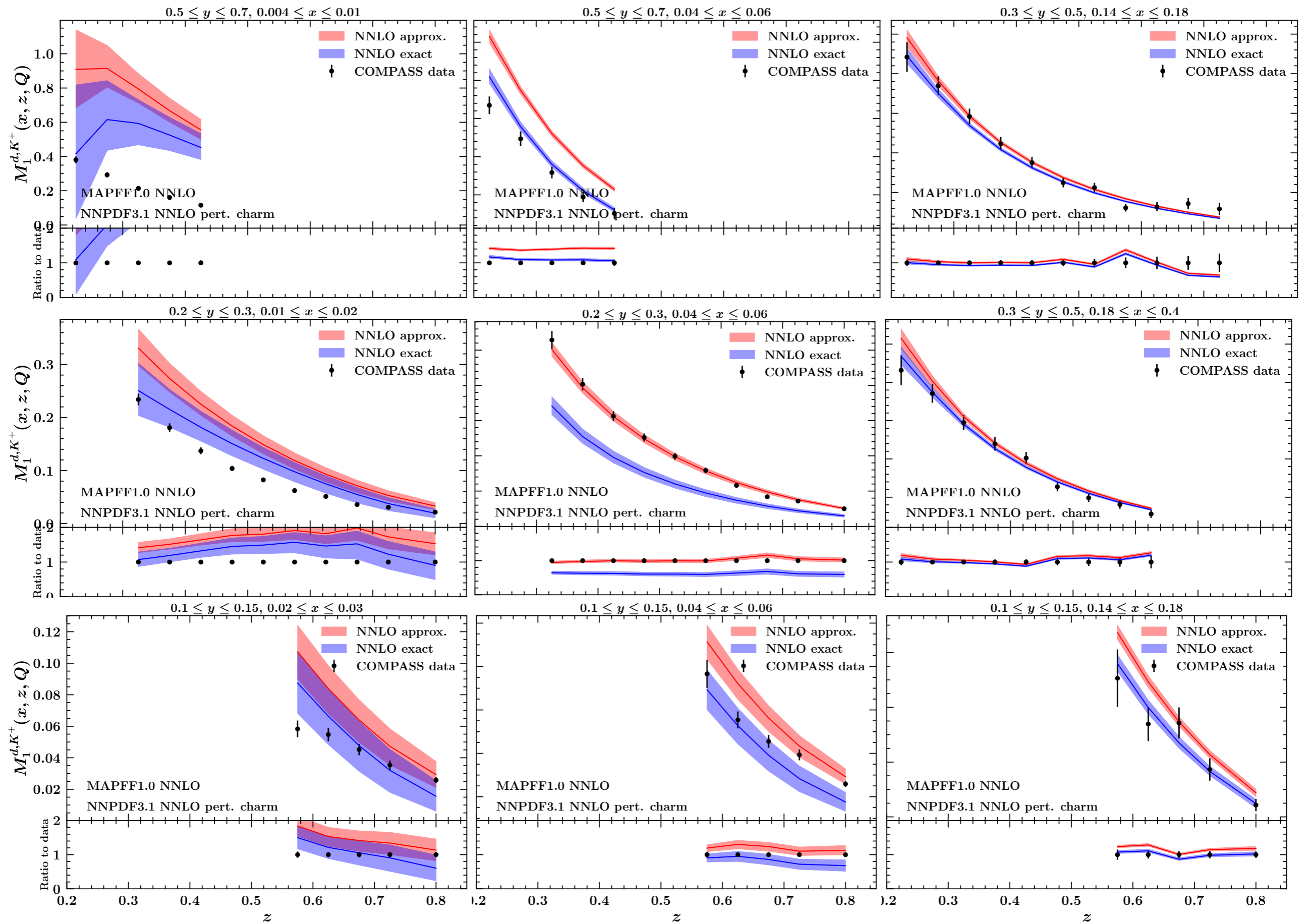
SIDIS structure functions

- Approximated NNLO corrections to SIDIS used so far in fits:
 - expansion to fixed order of **threshold-resummation** results.



- NNLO corrections significant but perturbative series **converges** reasonably well.
- **Exact NNLO** corrections became recently available. [[Phys.Rev.Lett. 132 \(2024\) 25, 251902](#)][[Phys.Rev.Lett. 132 \(2024\) 25, 251901](#)]
- Ongoing benchmark and fit implementation.

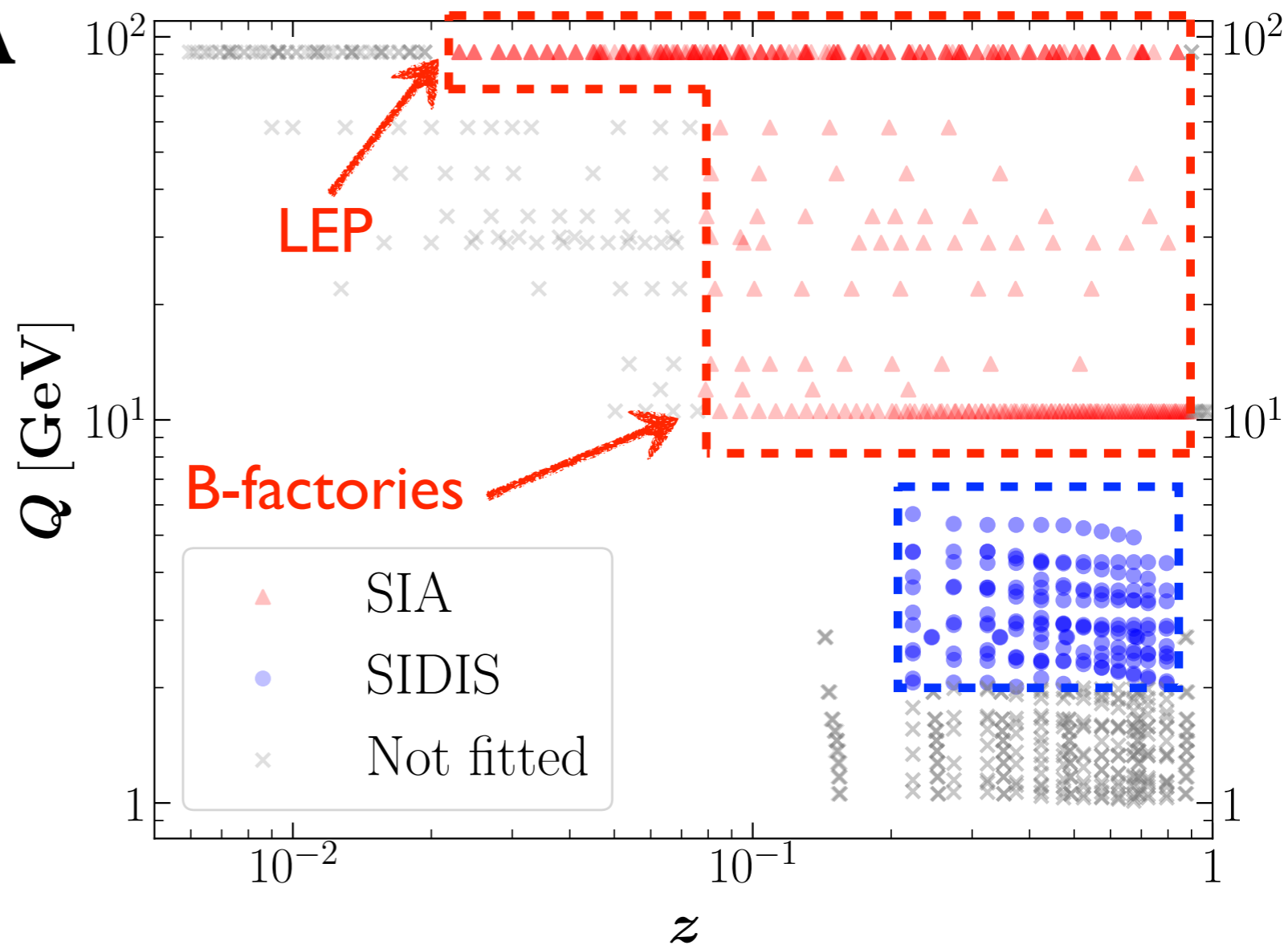
y 

y  x

The MAPFF1.0 NNLO analysis

The data set

- Around 700 data points for **SIA** and **SIDIS** for both pions and kaons.
- **No pp data:** NNLO corrections not known.
- **Wide coverage** in z and Q .
- Cuts on **SIA** $z \in [z_{\min} : 0.9]$:
 $z_{\min} = 0.02$ at $Q = M_Z$ and
 $z_{\min} = 0.075$ otherwise:
 - avoid large small- and large- z (resummation) corrections.
- Cuts on **SIDIS** $z \in [0.2 : 0.8]$ and $Q \geq 2$ GeV:
 - avoid control bins ($z < 0.2$) and contribution from exclusive decays ($z > 0.8$),
 - ensure perturbative convergence ($Q > 2$ GeV),
 - coverage in x limited by the cut in Q .

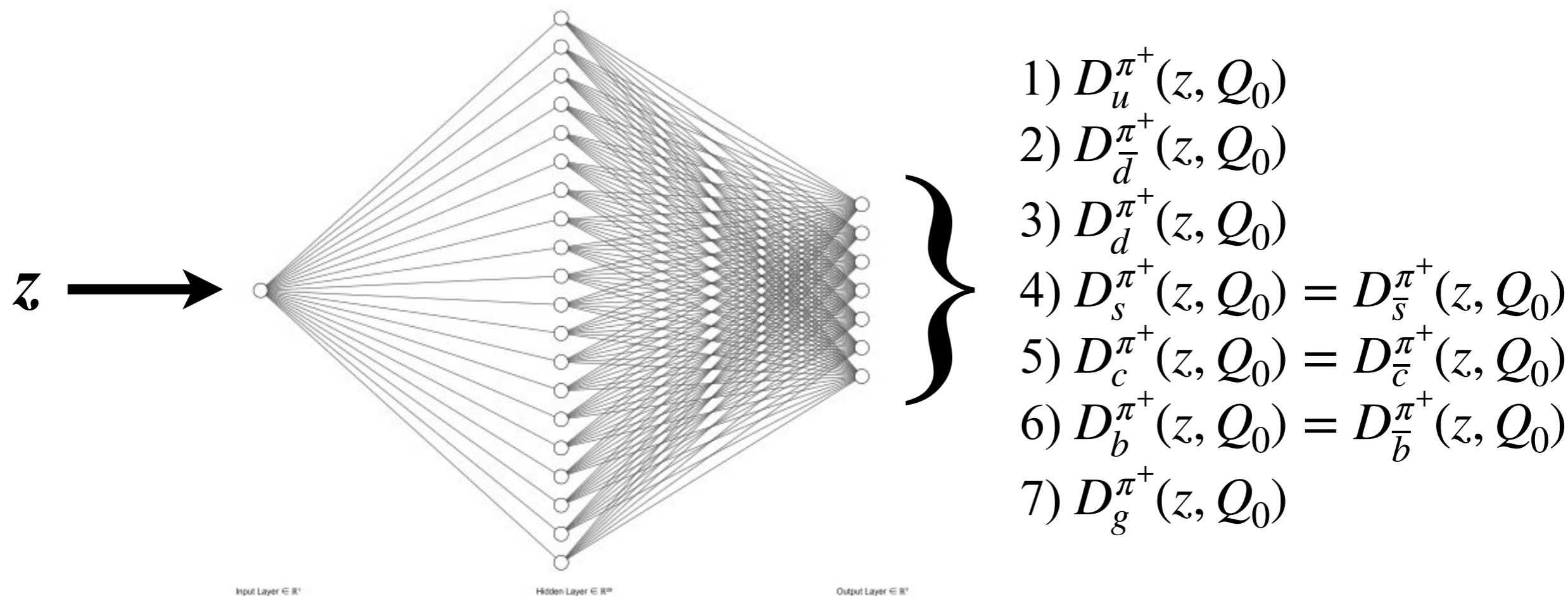


The MAPFF1.0 NNLO analysis

The parameterisation

🍏 **All** fitted FFs are parameterised at $Q_0 = 5$ GeV using a **single** NN:

🍏 architecture 1-20-7 (187 free parameters).

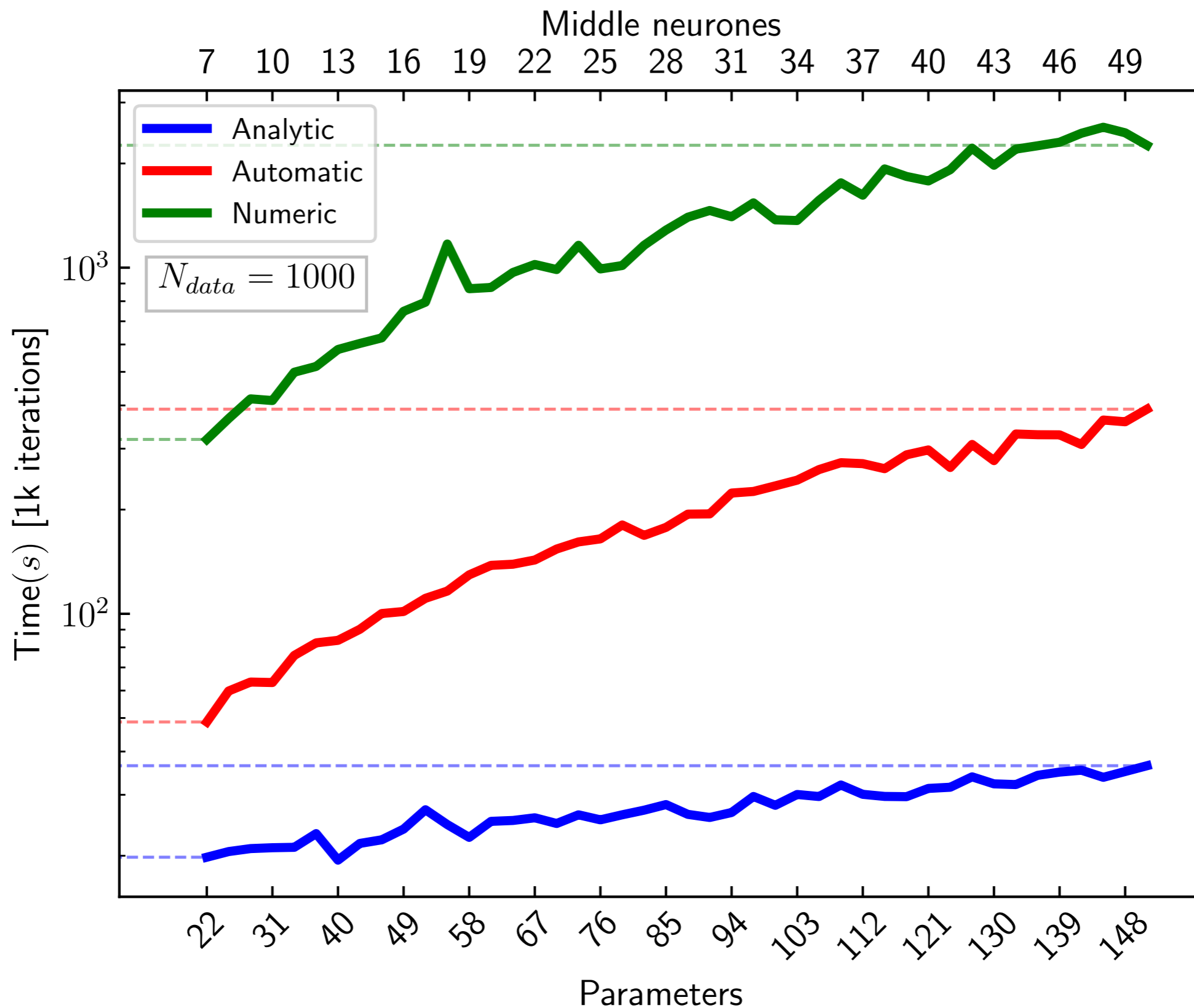


🍏 We exploit the ability to compute the **analytic derivatives** of any NN w.r.t. its free parameters using the **NNAD** library. [R. Abdul Khalek, V. Bertone, arXiv:2005.07039]

🍏 This enormously simplifies the task of the minimiser in that the gradient of the χ^2 can be computed analytically (as opposed to numerical or automatic derivatives). 11

The MAPFF1.0 NNLO analysis

The parameterisation



[Abdul Khalek, Bertone, arXiv:2005.07039]

The MAPFF1.0 NNLO analysis

The minimiser

🍏 The use of NNs and the consequent **large number** of free parameters requires an efficient minimiser.

🍏 MINUIT is not an option:

according to the actual needs and “on demand”. There is no protection against an upper limit on the number of parameters, however the “technological” limitations of MINUIT can be seen around a maximum of 15 free parameters at a time.

🍏 We have chosen to use **Ceres Solver**: [\[http://ceres-solver.org\]](http://ceres-solver.org)

Ceres Solver ^[1] is an open source C++ library for modeling and solving large, complicated optimization problems. It can be used to solve **Non-linear Least Squares** problems with bounds constraints and general unconstrained optimization problems. It is a mature, feature rich, and performant library that has been used in production at Google since 2010. For more, see [Why?](#).

🍏 Ceres Solver is very well suited for complicated optimisation problems.

🍏 It allows for the use of **automatic** and **analytic** differentiations.

The MAPFF1.0 NNLO analysis

The theory setup

- 🍏 Predictions based on **collinear factorisation**:

$$d\sigma^{e^+e^- \rightarrow h+X} \propto \sum_{i=q,g} d\hat{\sigma}_i(z, \alpha_s(Q)) \otimes D_i^h(z, Q)$$

$$d\sigma^{ep \rightarrow h+X} \propto \sum_{j,i=q,g} f_j(x, Q) \otimes d\hat{\sigma}_{ji}(x, z, \alpha_s(Q)) \otimes D_i^h(z, Q)$$

- 🍏 FFs obey the **DGLAP evolution equations**:

$$\frac{d}{d \ln Q^2} D_i^h(z, Q) = \sum_{i=q,g} P_{ij}(z, \alpha_s(Q)) \otimes D_j^h(z, Q)$$

- 🍏 Integration over the final-state phase space fully taken into account.

$$\sigma^{ep \rightarrow h+X} = \int_{y_{\min}}^{y_{\max}} dy \int_{x_{\min}}^{x_{\max}} dx \int_{Q_{\min}}^{Q_{\max}} dQ \frac{d^3 \sigma^{ep \rightarrow h+X}}{dy dx dQ}$$

- 🍏 **Perturbative contributions**, numerical convolutions, solution of the DGLAP equations and numerical integrations provided by **APFEL++**.

V. Bertone [arXiv:1708.00911]

- 🍏 All perturbative ingredients are **NNLO** accurate (approx. for SIDIS).

The MAPFF1.0 NNLO analysis

The theory setup

- 🍏 The χ^2 is computed exploiting **all** possible sources of uncertainties:

$$\chi^2 = \sum_{i,j} (m_i - t_i) V_{ij}^{-1} (m_j - t_j)$$

$$V_{ij} = \delta_{ij} \sigma_{unc}^2 + \left(\sum_{k=1}^{n_{sys}} \delta_i^{(k)} \delta_j^{(k)} \right) m_i m_j$$

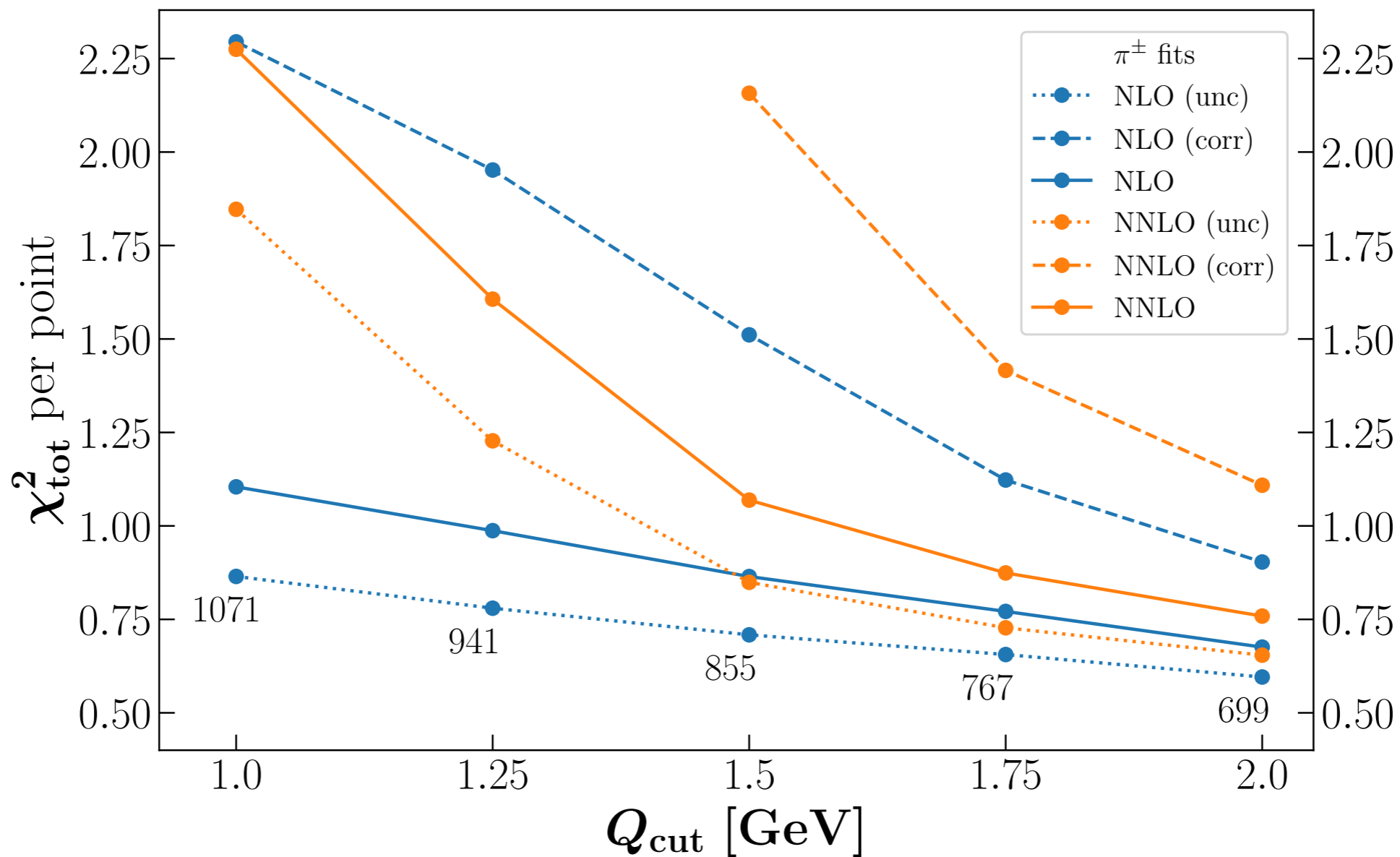
A. Bacchetta *et al.* [arXiv:1912.07550]

- 🍏 Data handling and computation of the χ^2 is delegated to **NangaParbat**:
 - 🍏 Monte-Carlo replica generation consistent with the covariance matrix,
 - 🍏 efficient computation of the χ^2 based on the Cholesky decomposition of V ,
 - 🍏 computation of the systematic shifts for data-theory comparisons,
 - 🍏 t_0 prescription (if necessary) to treat normalisation uncertainties.

Experimental data

Importance of experimental correlations

Bertone *et al.* [arXiv:2204.10331]

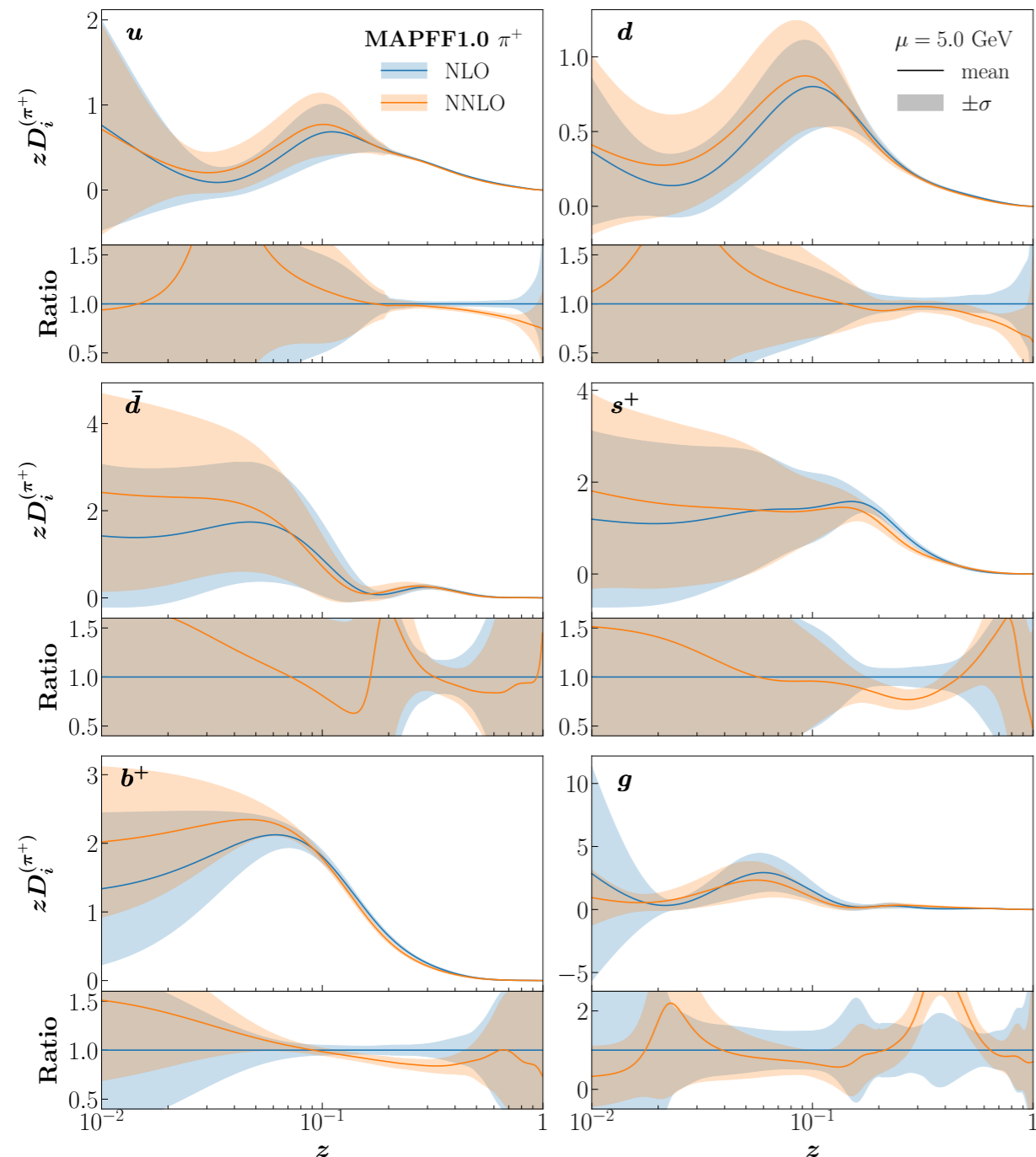


The MAPFF1.0 analysis observed that the fit quality depends significantly on the treatment of **systematic uncertainties** of COMPASS:

- uncorrelated (unc), *i.e.* 0% correlated,
- 100% correlated (corr),
- 80% correlated (**default**: experimental prescription).

FFs at NNLO

Fit quality for pions



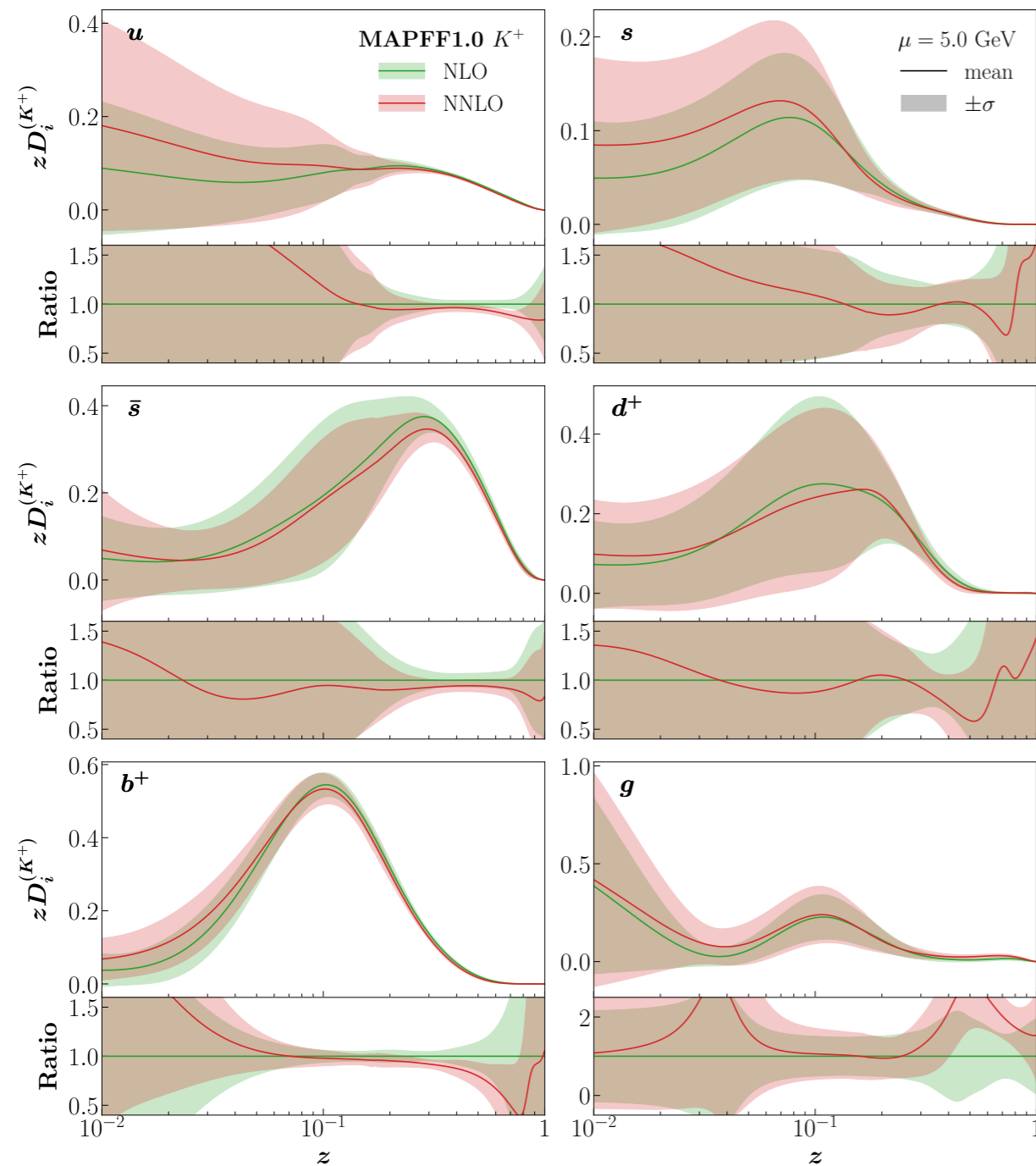
Bertone et al. [arXiv:2204.10331]

Data set	N_{dat}	NLO	NNLO
BELLE h^\pm	70	0.14	0.13
BABAR h^\pm	39	0.91	0.76
TASSO 12 GeV h^\pm	4	0.90	0.92
TASSO 14 GeV h^\pm	9	1.33	1.35
TASSO 22 GeV h^\pm	8	1.65	1.81
TPC h^\pm	13	0.23	0.25
TASSO 30 GeV h^\pm	2	0.30	0.34
TASSO 34 GeV h^\pm	9	1.08	1.48
TASSO 44 GeV h^\pm	6	1.13	1.37
TOPAZ h^\pm	5	0.24	0.37
ALEPH h^\pm	23	1.24	1.46
DELPHI (inclusive) h^\pm	21	1.31	1.25
DELPHI (uds tagged) h^\pm	21	2.68	2.89
DELPHI (b tagged) h^\pm	21	1.58	1.73
OPAL h^\pm	24	1.63	1.79
SLD (inclusive) h^\pm	34	1.05	1.13
SLD (uds tagged) h^\pm	34	1.59	2.16
SLD (b tagged) h^\pm	34	0.55	0.68
HERMES $h^- d$	2	0.41	0.32
HERMES $h^+ p$	2	0.01	0.02
HERMES $h^- d$	2	0.17	0.11
HERMES $h^+ p$	2	0.35	0.32
COMPASS h^-	157	0.48	0.55
COMPASS h^+	157	0.62	0.72
Global data set	699	0.68	0.76

- Pion case:
 - NLO better than NNLO,
 - appreciable effects on FFs.

FFs at NNLO

Fit quality for kaons



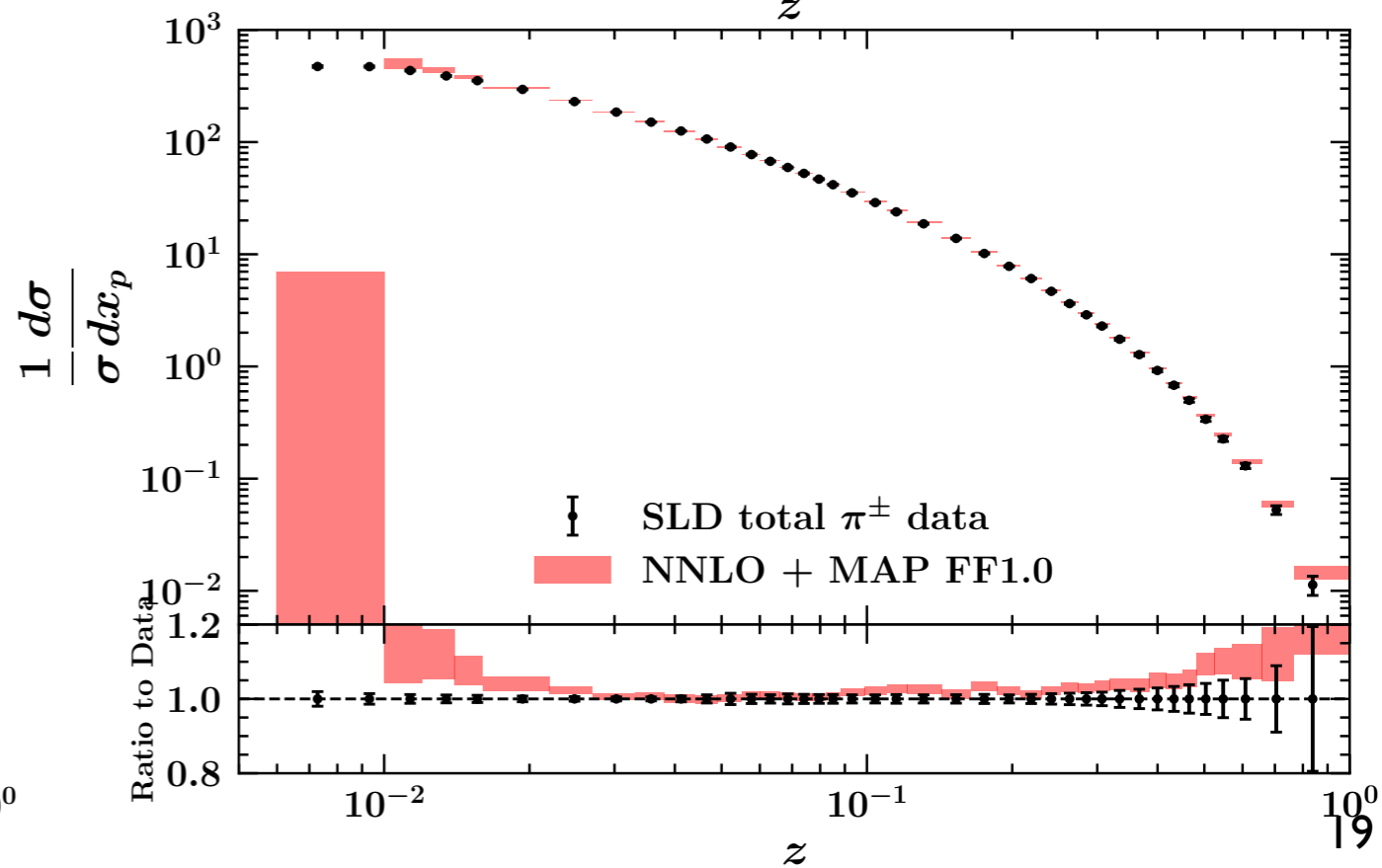
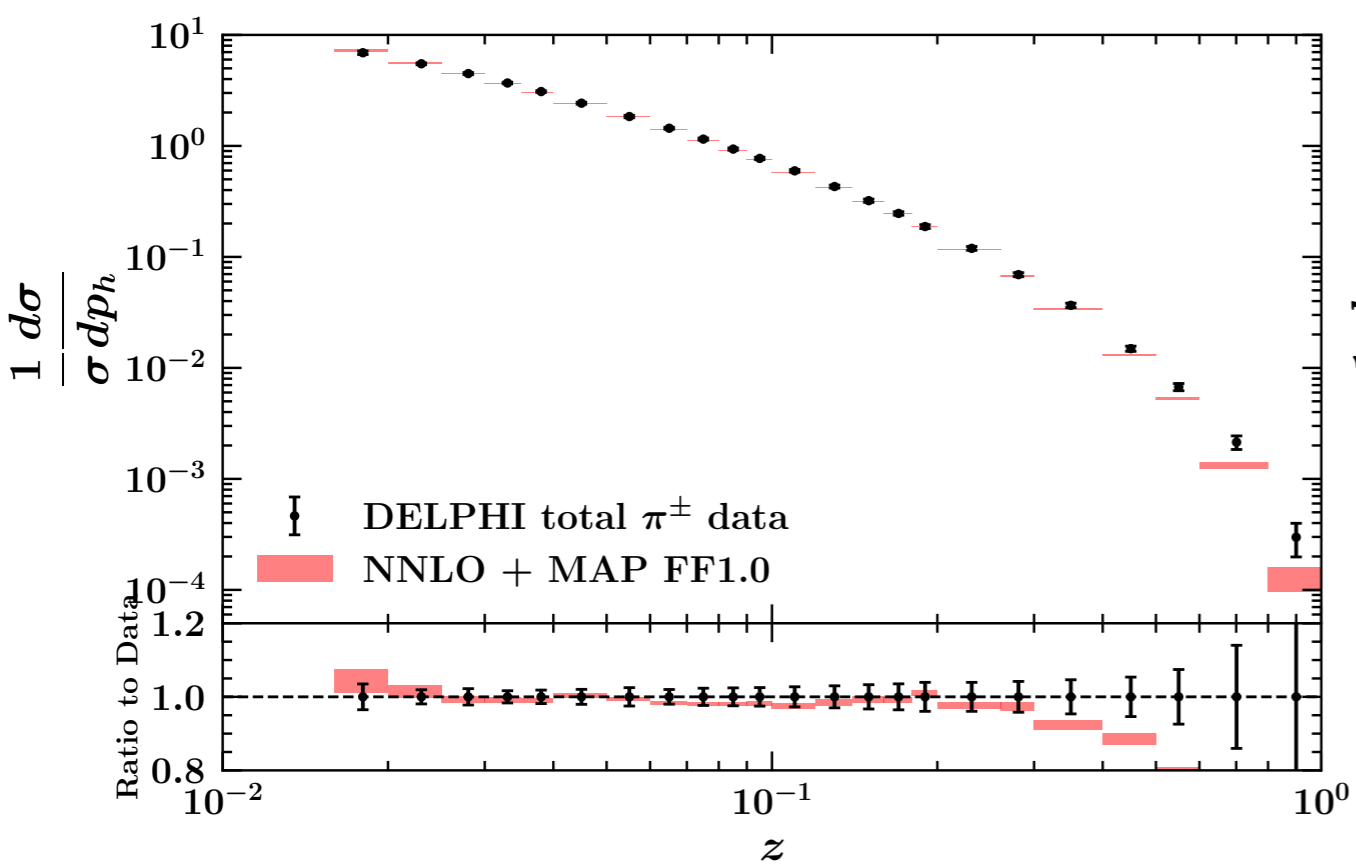
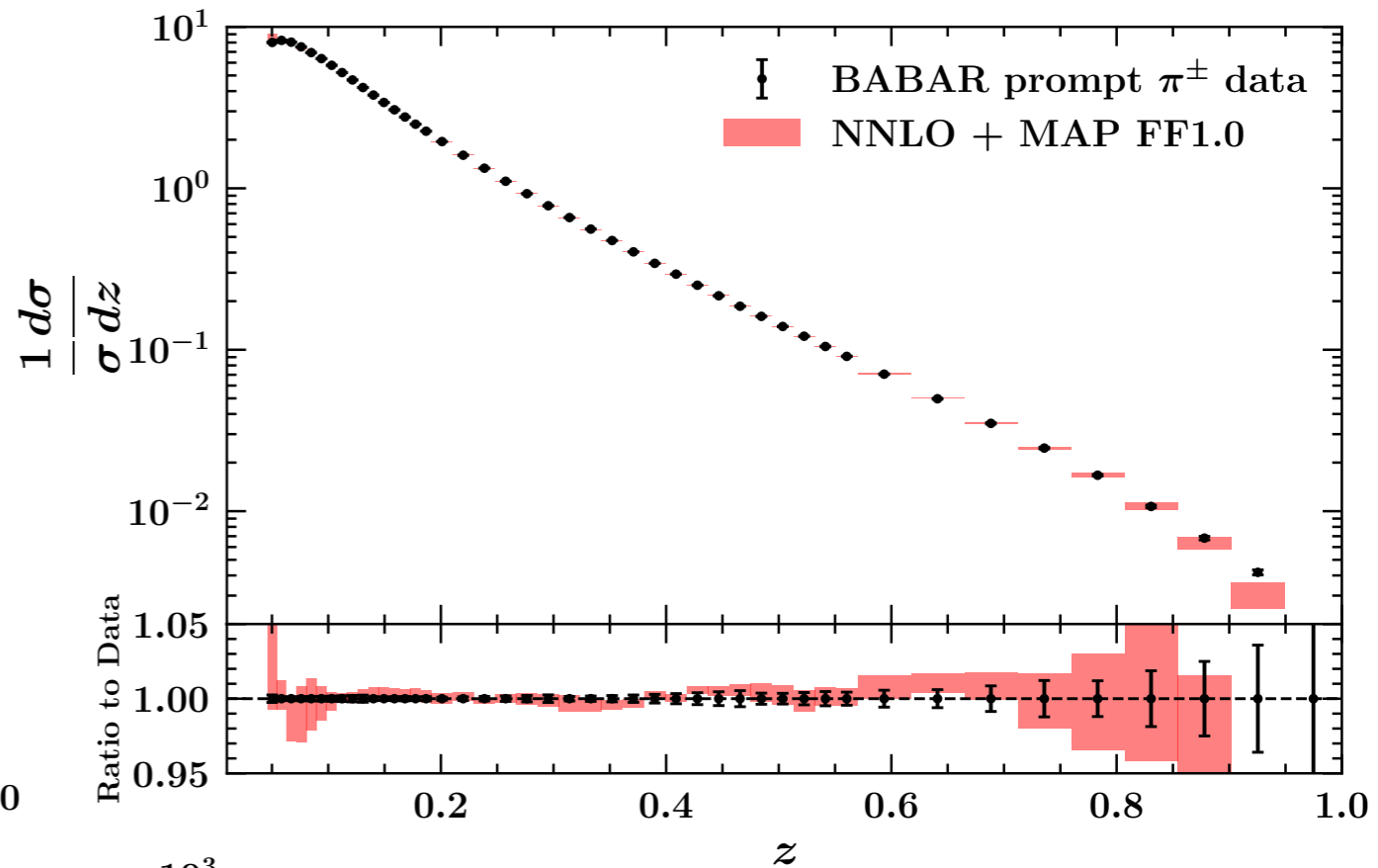
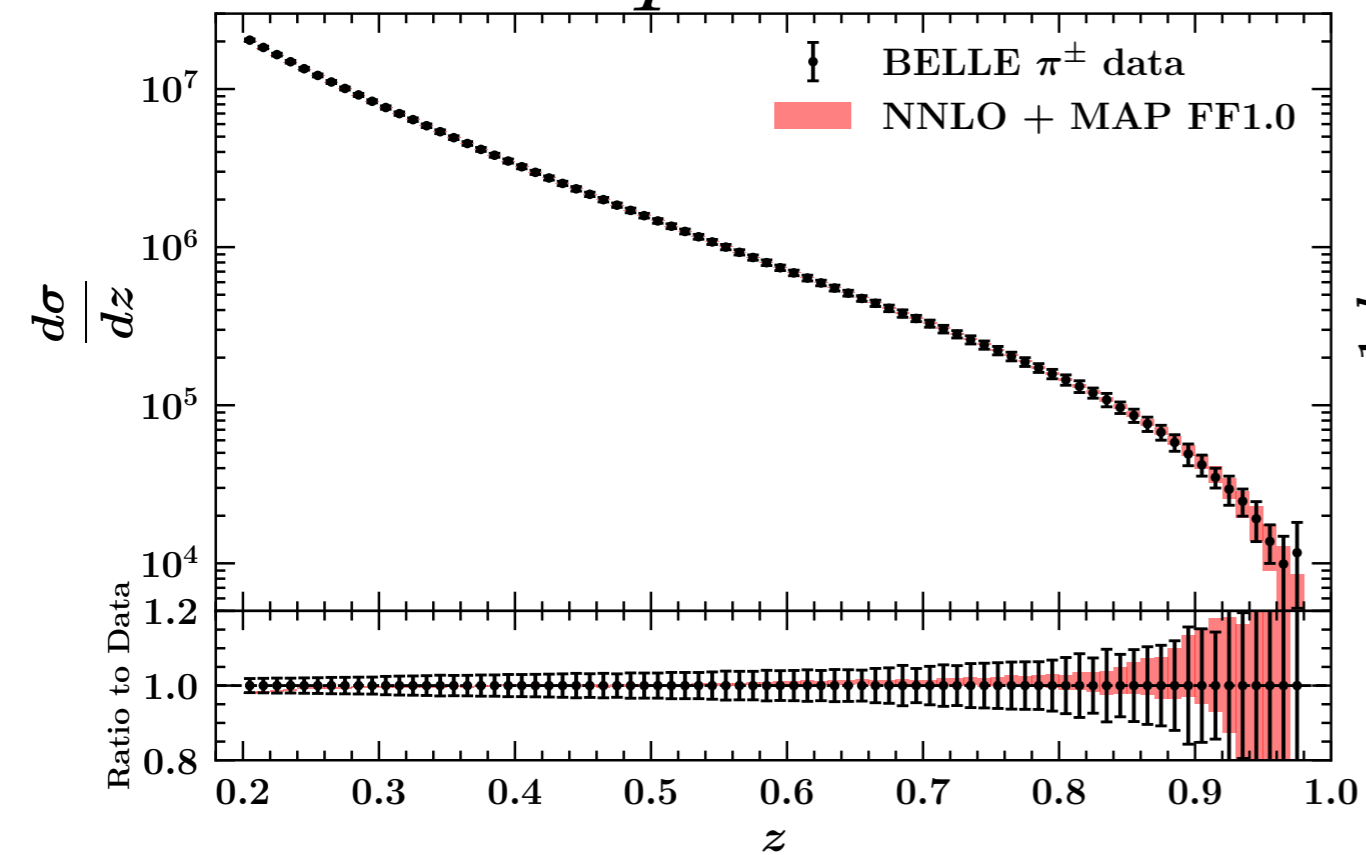
Bertone et al. [arXiv:2204.10331]

Data set	N_{dat}	NLO	NNLO
BELLE h^\pm	70	0.39	0.41
BABAR h^\pm	28	0.36	0.25
TASSO 12 GeV h^\pm	3	0.85	0.87
TASSO 14 GeV h^\pm	9	1.24	1.22
TASSO 22 GeV h^\pm	6	0.89	0.90
TPC h^\pm	13	0.38	0.40
TASSO 30 GeV h^\pm	—	—	—
TASSO 34 GeV h^\pm	5	0.07	0.06
TASSO 44 GeV h^\pm	—	—	—
TOPAZ h^\pm	3	0.10	0.11
ALEPH h^\pm	18	0.49	0.48
DELPHI (inclusive) h^\pm	23	0.97	0.99
DELPHI (uds tagged) h^\pm	23	0.44	0.38
DELPHI (b tagged) h^\pm	23	0.42	0.45
OPAL h^\pm	10	0.39	0.36
SLD (inclusive) h^\pm	35	0.83	0.67
SLD (uds tagged) h^\pm	35	1.37	1.52
SLD (b tagged) h^\pm	35	0.75	0.77
HERMES $h^- d$	2	0.18	0.13
HERMES $h^+ p$	2	0.05	0.04
HERMES $h^- d$	2	0.58	0.48
HERMES $h^+ p$	2	0.56	0.43
COMPASS h^-	156	0.74	0.59
COMPASS h^+	156	0.76	0.67
Global data set	659	0.62	0.55

- Kaon case:
- NNLO better than NLO,
- moderate effects on FFs.

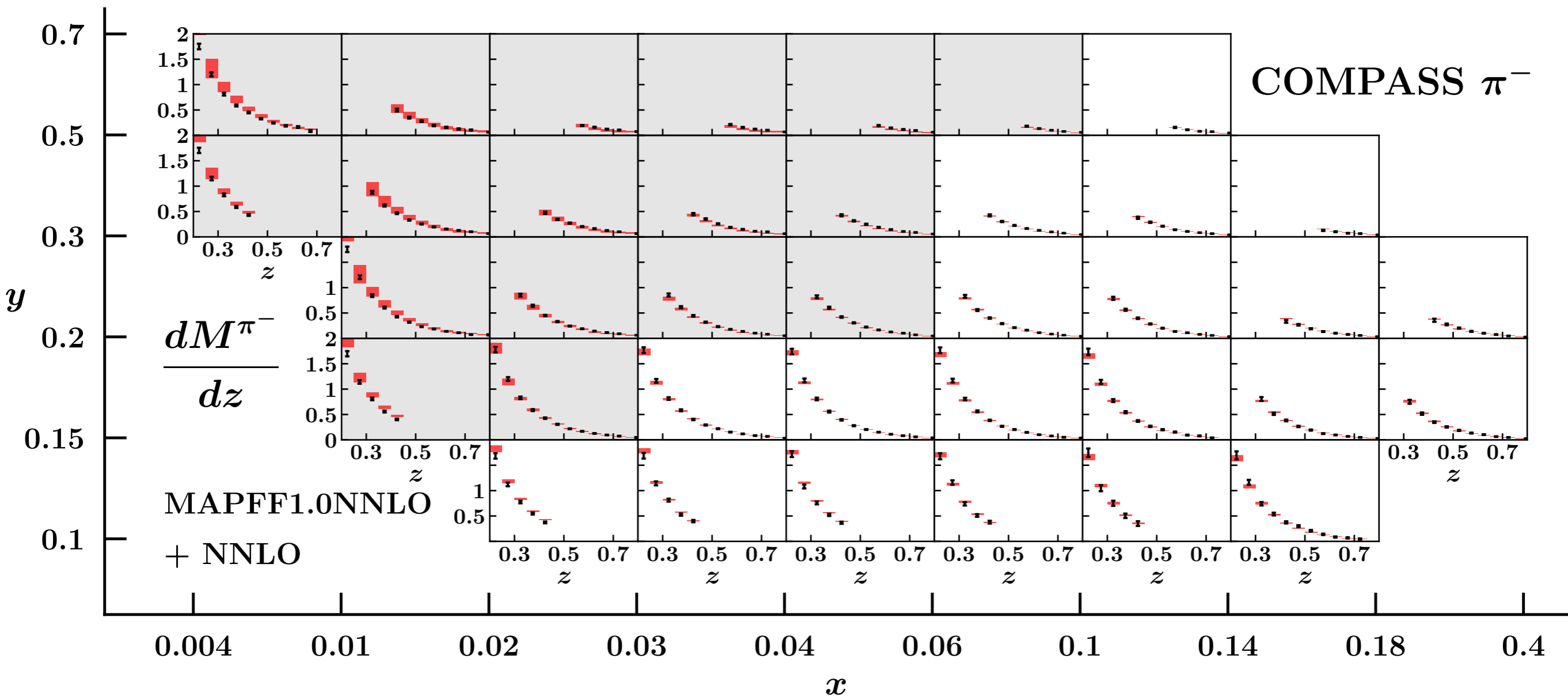
FFs at NNLO

Data description: SIA



FFs at NNLO

Data description: SIDIS



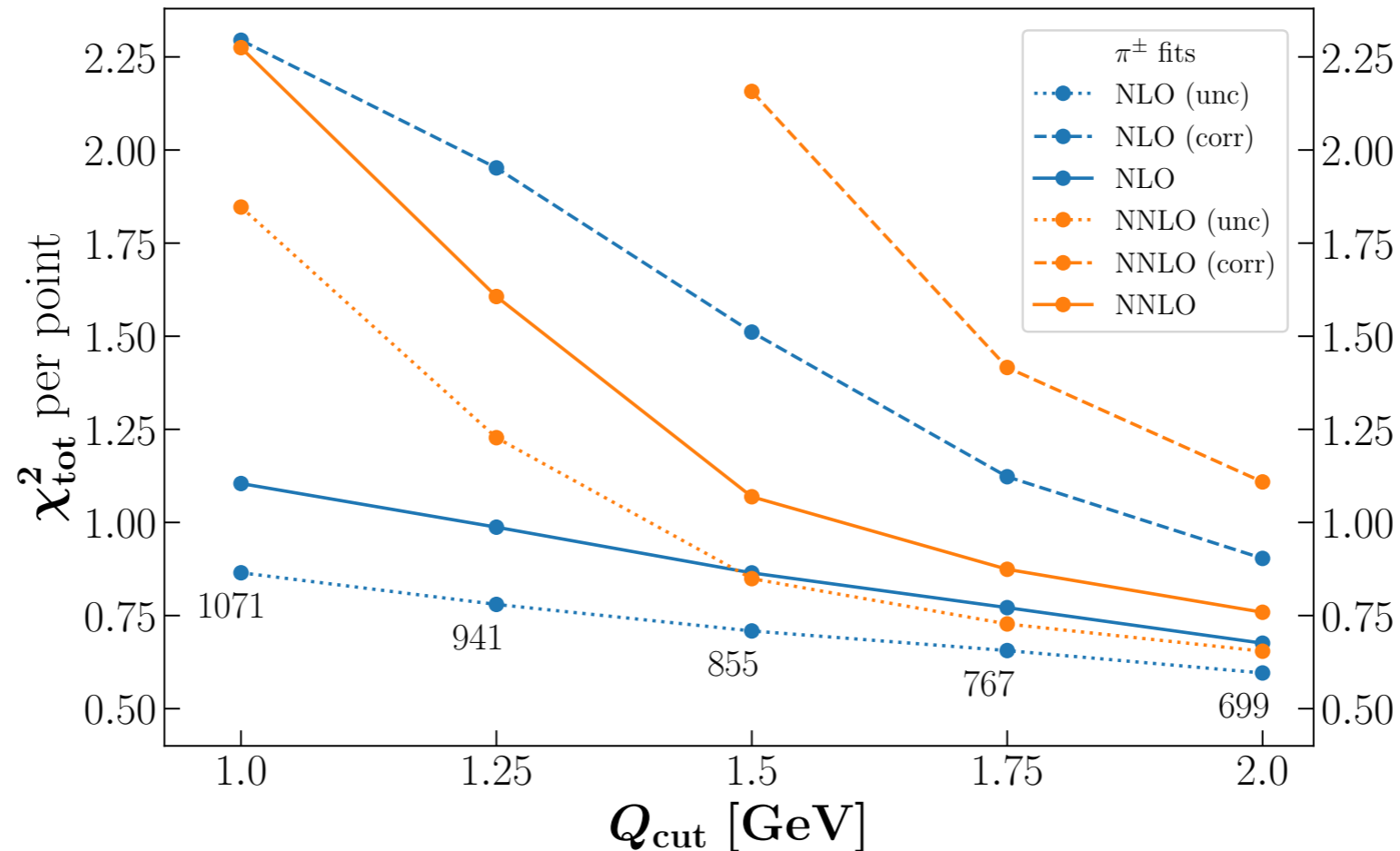
- Good description of the data included in the fit:
 - even for bins that are not included because of kinematic cuts.

FFs at NNLO

NLO vs. approximated NNLO

- While both MAPFF1.0 and BDSS confirm that COMPASS high- Q data is better described by NNLO, it is not clear as yet where NNLO starts doing better than NLO.

Bertone *et al.* [arXiv:2204.10331]



Borsa *et al.* [arXiv:2202.05060]

Experiment	$Q^2 \geq 1.5 \text{ GeV}^2$			$Q^2 \geq 2.0 \text{ GeV}^2$			$Q^2 \geq 2.3 \text{ GeV}^2$			$Q^2 \geq 3.0 \text{ GeV}^2$		
	#data	NLO	NNLO	#data	NLO	NNLO	#data	NLO	NNLO	#data	NLO	NNLO
SIA	288	1.05	0.96	288	0.91	0.87	288	0.90	0.91	288	0.93	0.86
COMPASS	510	0.98	1.14	456	0.91	1.04	446	0.91	0.92	376	0.94	0.93
HERMES	224	2.24	2.27	160	2.40	2.08	128	2.71	2.35	96	2.75	2.26
TOTAL	1022	1.27	1.33	904	1.17	1.17	862	1.17	1.13	760	1.16	1.07

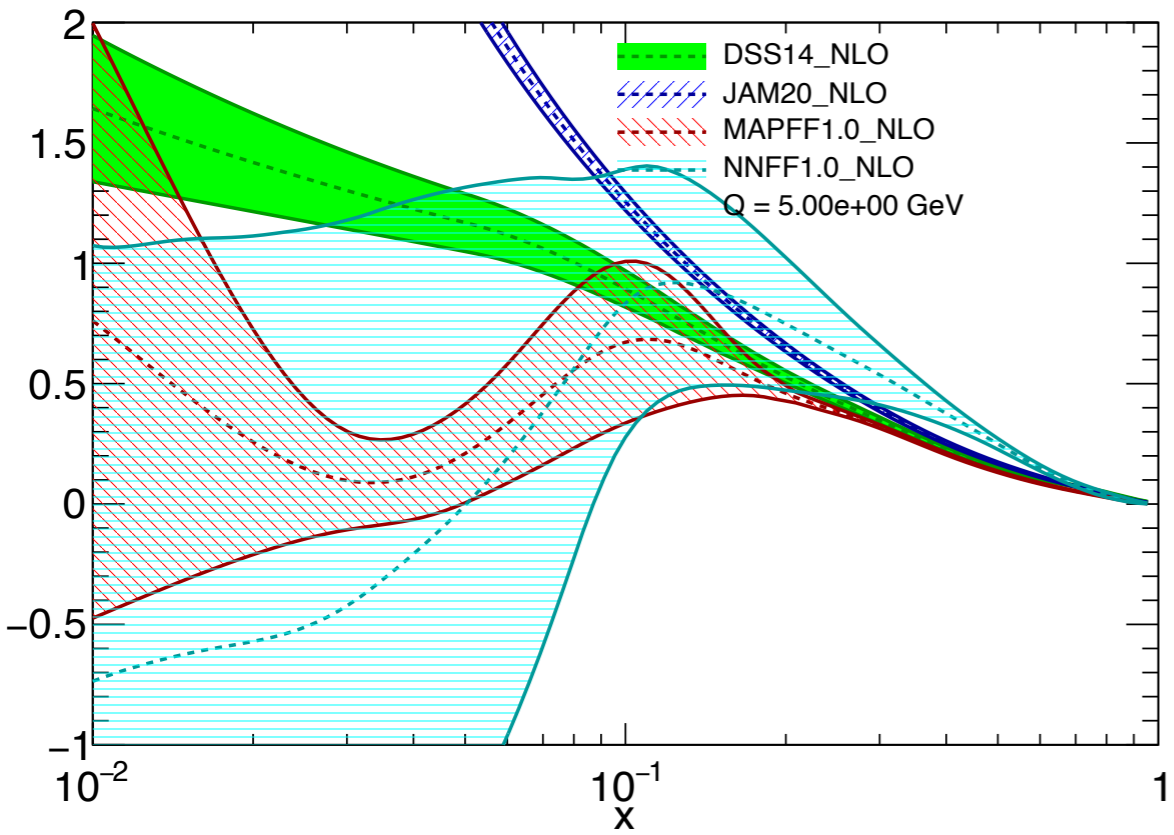
Summary

- Present extractions of FFs have reached a high sophistication:
 - inclusive data sets,
 - high theoretical accuracy,
 - COMPASS is playing a central role in this quest.
- Still a long way to go:
 - still large discrepancies amongst active FF collaborations,
 - waiting for the EIC to help us constrain FFs even further,
 - a synergic effort along the same lines of the PDF4LHC working group (FF4EIC?) would be extremely welcome.

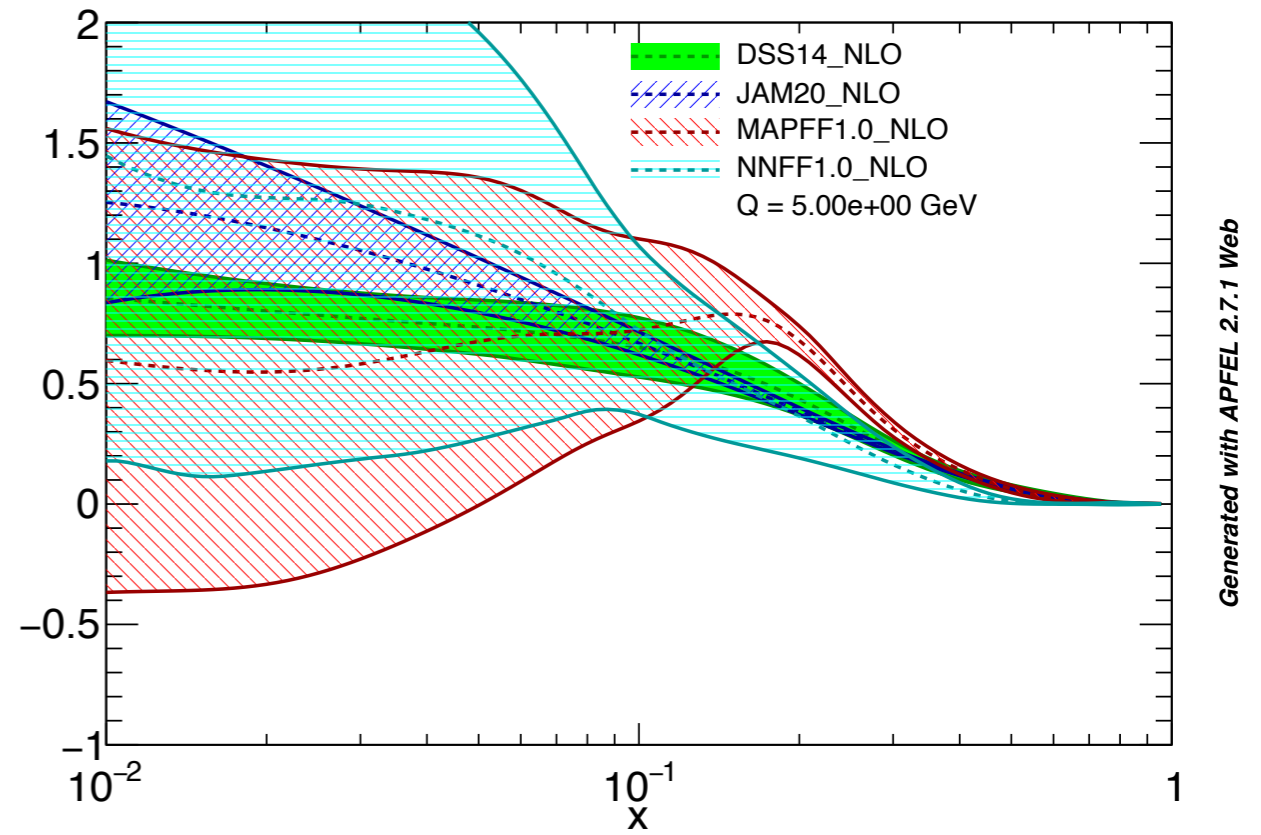
Backup

Comparison to other extractions

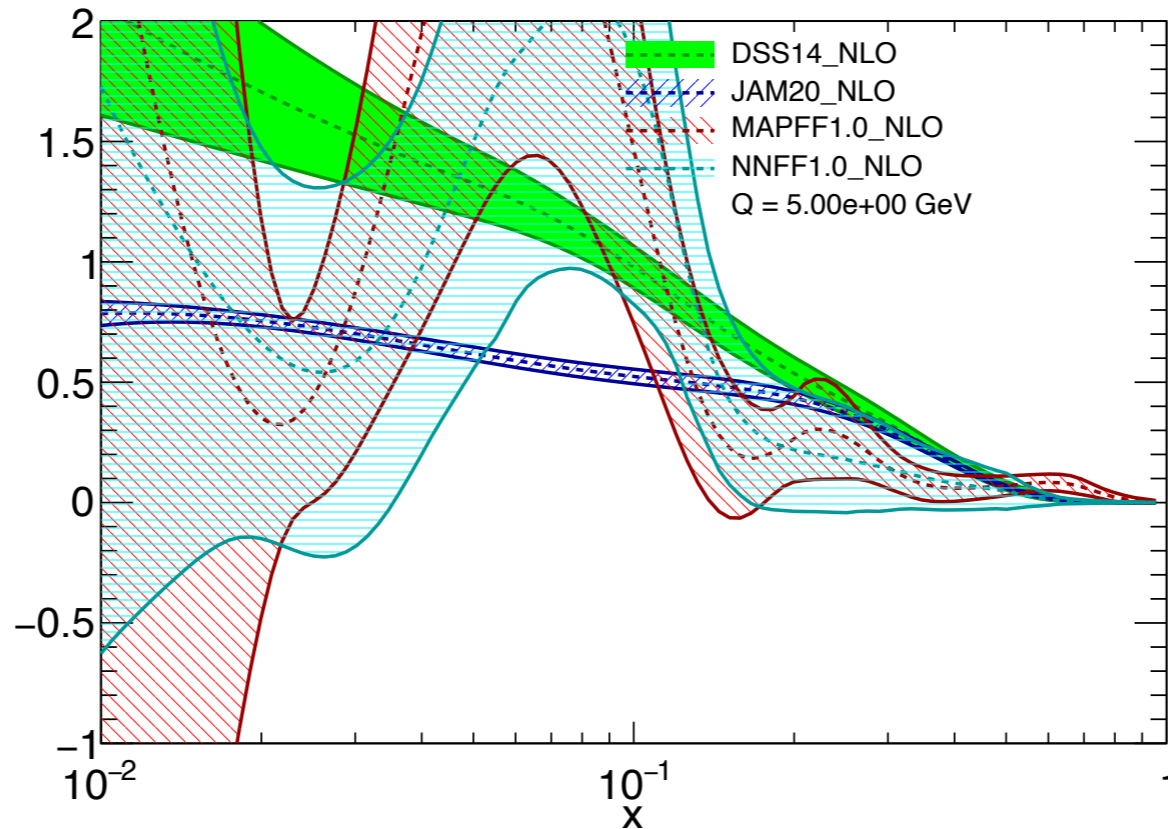
$xu(x,Q)$, comparison



$xs(x,Q)$, comparison



$xg(x,Q)$, comparison

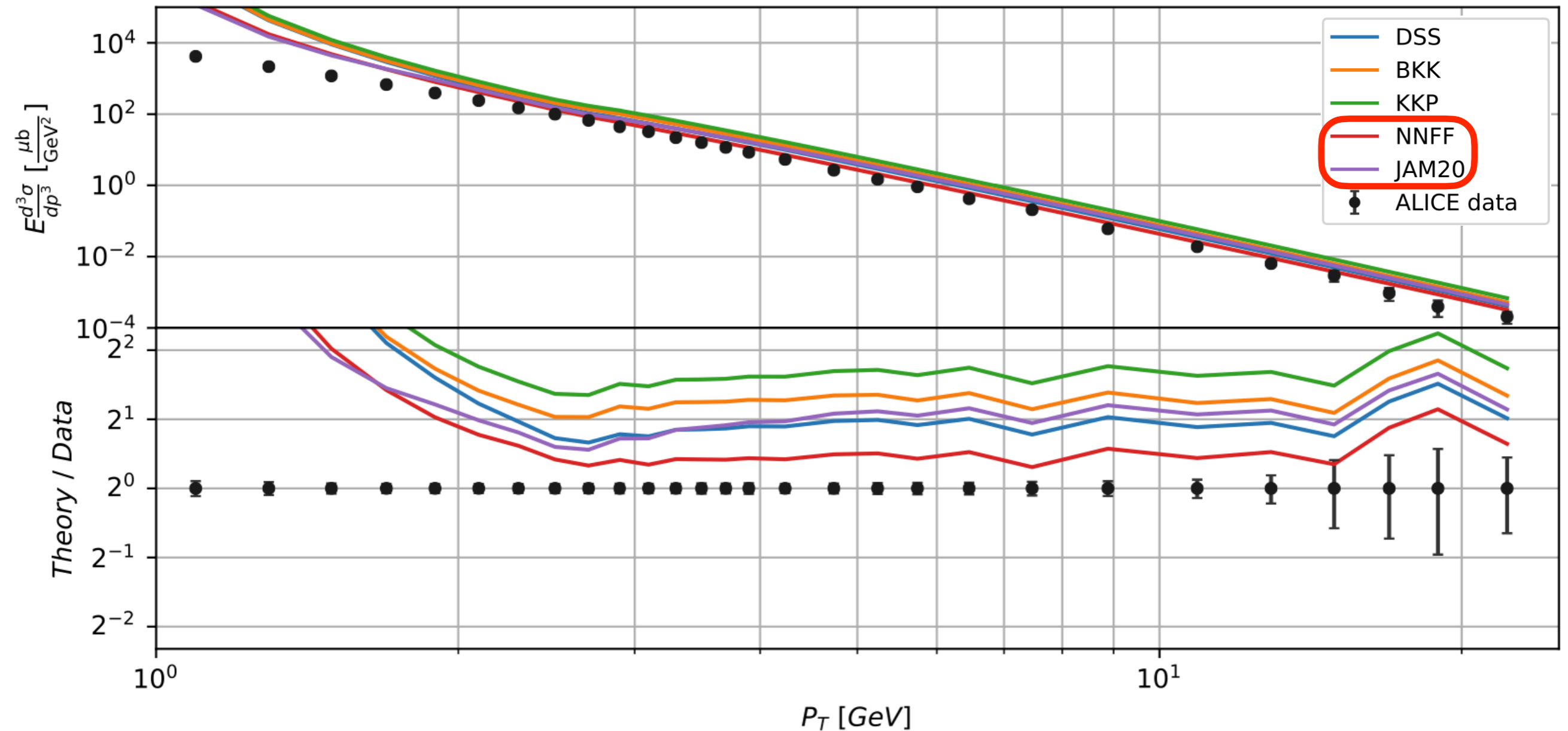


Impact of NNLO corrections

pp observables

Comparison data/theory for inclusive **pion p_T spectrum at the LHC:**

Duwentäster *et al.* [ArXiv:2105.09873] $p + p \rightarrow \pi^0 + X$ at $\sqrt{s_{NN}} = 7$ TeV



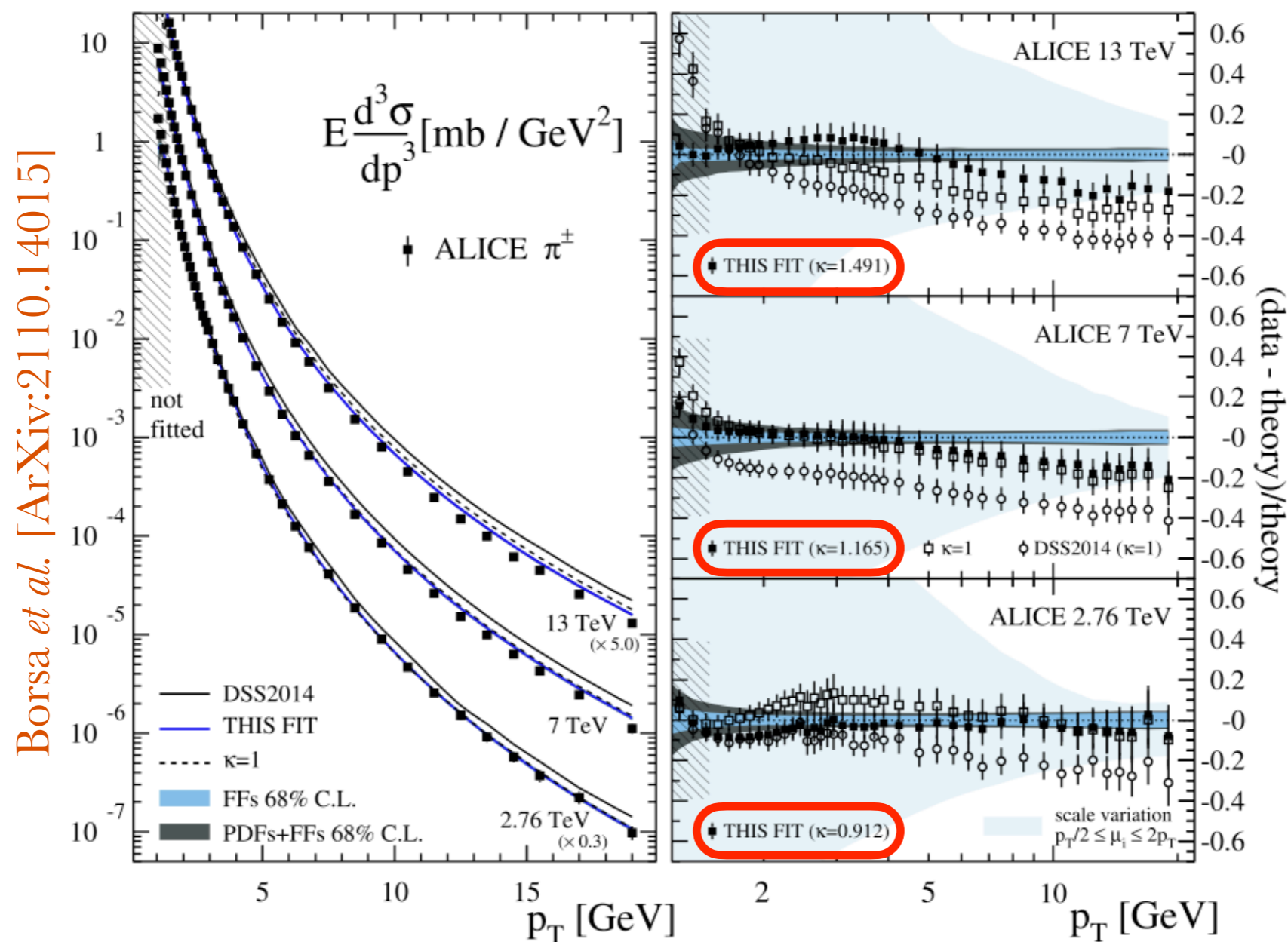
Data is largely **overshot by NLO predictions** obtained with all FF sets.

Even most **modern FF sets** (NNFF and JAM) do not do a great job:

- however neither NNFF nor JAM include *pp* data.

Impact of NNLO corrections

pp observables

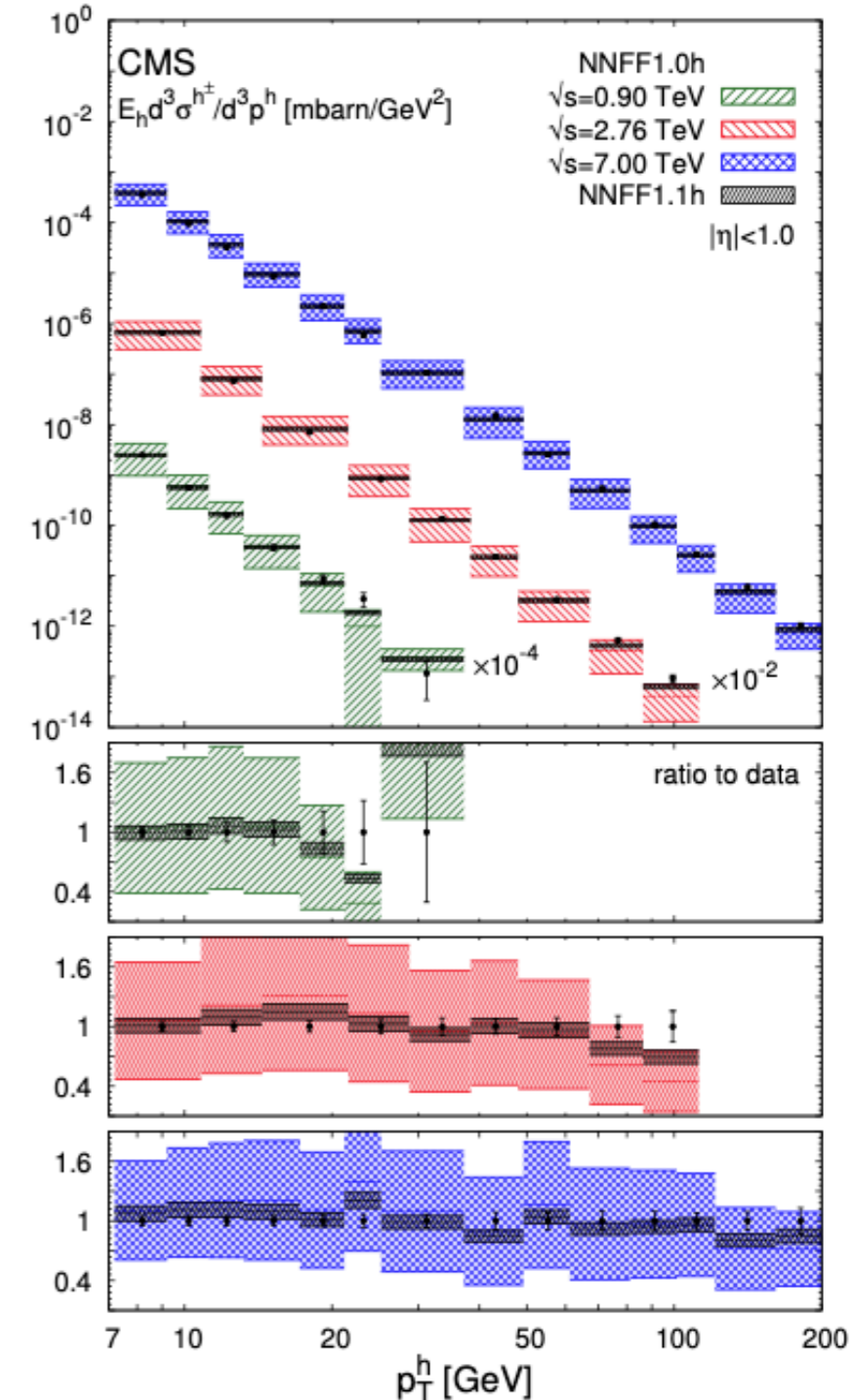
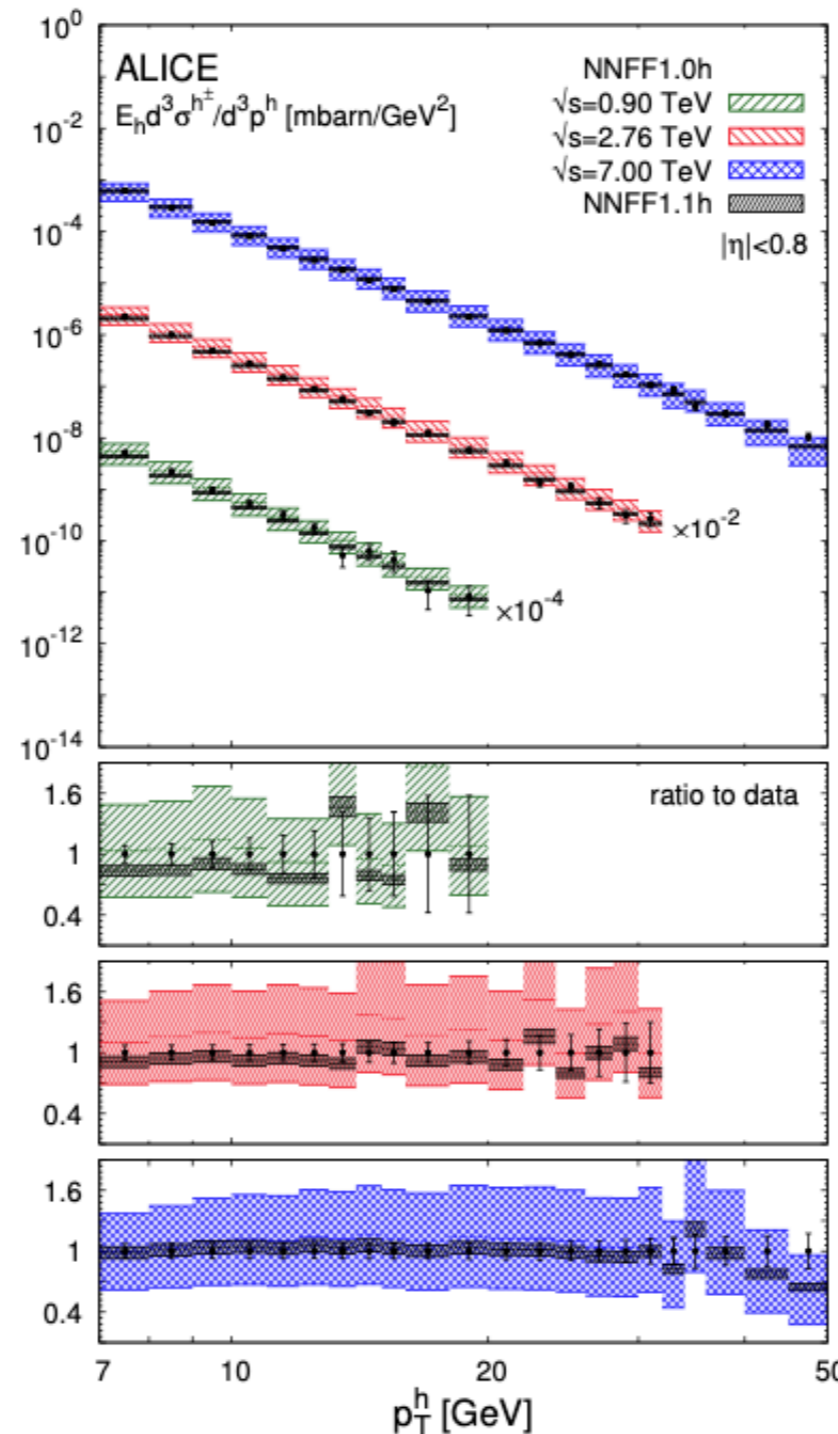
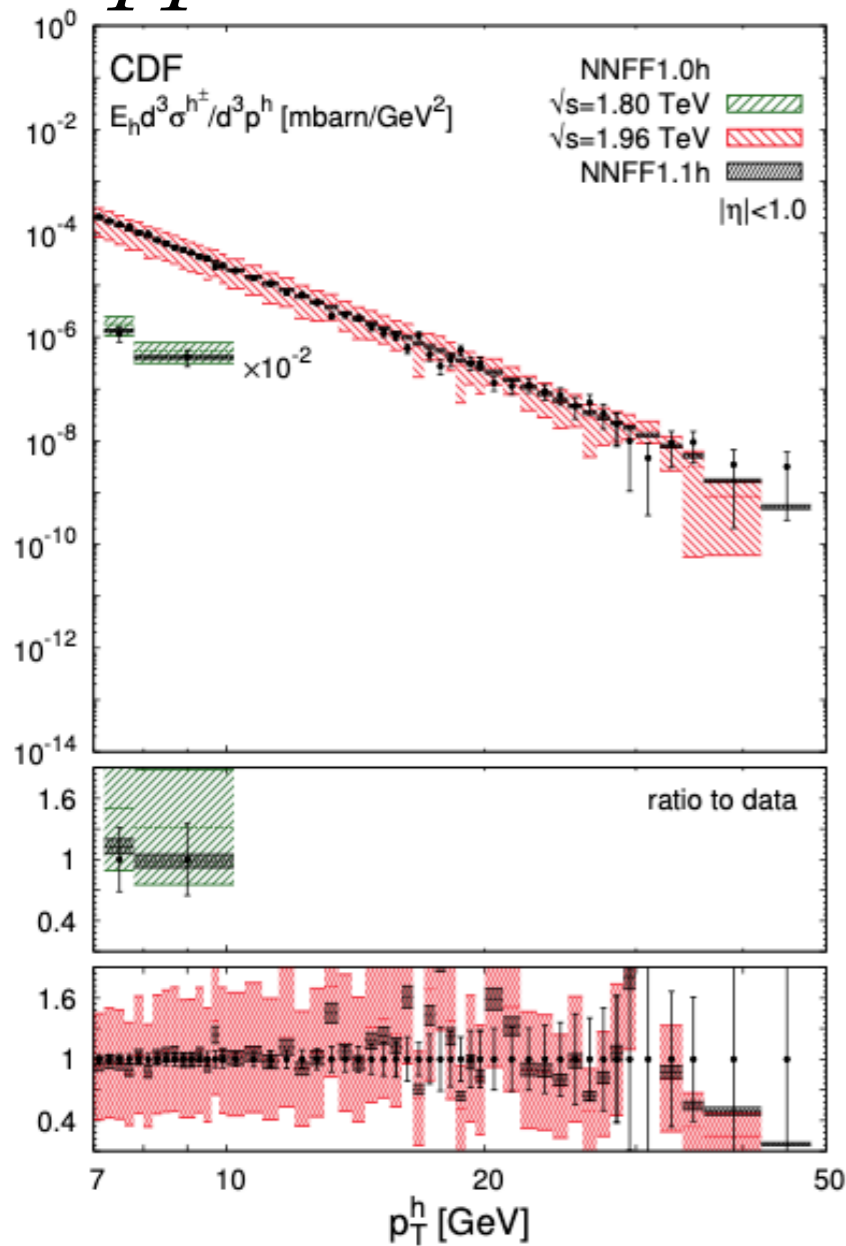


Of course, including pp data in a fit of FFs helps improve the agreement:

- however, **adjustments** of μ_R and μ_F significantly improves the description,
- reflection of **large NNLO corrections**.
- Same conclusion found in a more recent analysis of h^\pm FFs. [Phys.Rev.D 109 (2024) 5, 052004]

Impact of NNLO corrections

pp observables



Bertone *et al.* [ArXiv:1807.03310]

In the case of charged-hadron FFs, *pp* data has a **huge impact**:

- consequence of a better constraint on the gluon FF.
- Analysis limited to $p_T > 7$ GeV to avoid **large NNLO corrections**.

Perturbative content

DGLAP evolution

- FFs obey the standard collinear DGLAP evolution equations:

$$\mu^2 \frac{\partial}{\partial \mu^2} D_{\text{NS}}^h = P_{\text{NS}} \otimes D_{\text{NS}}^h$$

$$\mu^2 \frac{\partial}{\partial \mu^2} \begin{pmatrix} D_{\Sigma}^h \\ D_g^h \end{pmatrix} = \begin{pmatrix} P_{qq} & P_{qg} \\ P_{gq} & P_{gg} \end{pmatrix} \otimes \begin{pmatrix} D_{\Sigma}^h \\ D_g^h \end{pmatrix}$$

- Time-like splitting functions known up to NNLO.
[A. Mitov and S. O. Moch \[hep-ph/0604160\]](#), [M. Gluck, E. Reya, and A. Vogt \[Phys.Rev. D48 \(1993\)\]](#)
- Numerical implementation in the APFEL/APFEL++ codes:
[V. Bertone, C. Carrazza, J. Rojo \[arXiv:1310.1394\]](#), [V. Bertone \[arXiv:170800911\]](#)
 - careful benchmark against in the the \mathcal{N} -space MELA code,
[V. Bertone, S. Carrazza, E. R. Nocera \[arXiv:1501.00494\]](#)
 - perfect agreement with QCDNUN(after a correction of a bug in the latter).
[M. Botje \[arXiv:1602.08383\]](#)
- Matching conditions known up to NLO.

Perturbative content

Hard cross sections

- **Single-inclusive annihilation:**

- currently known up to $O(\alpha_s^2)$, *i.e.* NNLO, in the **zero-mass** scheme.
A. Mitov and S. O. Moch [hep-ph/0604160]

- **mass corrections** known up to to $O(\alpha_s)$,
T. Kneesch [desy-thesis-10-049]

- **small- z resummation** corrections up to NNLL,
D. Anderle [arXiv:1611.03371]

- **hadron-mass** corrections (particular relevant for kaons and protons),
e.g. A. Accardi [arXiv:1411.3649]

- **threshold (large- z) resummation** corrections up to N^3LL .
S. O. Moch and A. Vogt [arXiv:0908.2746]

- **Semi-Inclusive Deep-Inelastic-Scattering:**

- currently fully know up to $O(\alpha_s)$, *i.e.* NLO, in the **zero-mass** scheme,

- partial knowledge of the $O(\alpha_s^2)$ corrections to F_L ,
D. Anderle [arXiv:1612.01293]

- approximated **NNLO** and **N^3LO** recently computed,
M. Abele *et al.* [arXiv:2109.00847], M. Abele *et al.* [arXiv:2203.07928]

- **threshold resummation** corrections to N^3LL .

- **Hadroproduction in proton-proton collisions:**

- currently fully know up to $O(\alpha_s^3)$, *i.e.* NLO, in the **zero-mass** scheme.
P. Aurenche *et al.* [hep-ph/9910252]

- (NNLO corrections available for single jet production, but not public).
J. Currie *et al.* [arXiv:1611.01460]

Experimental data

experiment		data type	N_i	#data in fit	experiment		data type	N_i	#data in fit
TPC [35]	29 GeV	incl.	1.038	17	BRAHMS [41]	0.20 TeV	π^+, π^-	1.313	26
		<i>uds, c, b</i> tag	1.038	27	STAR [16–19]	0.20 TeV	π^0	1.190	12
TASSO [36]	34 GeV	incl.	1.038	11		0.20 TeV	π^0	0.921	7
	44 GeV	incl.	1.038	7		0.20 TeV	π^+, π^-	1.029	26
SLD [37]	91.2 GeV	incl.	0.977	28		0.20 TeV	π^+, π^-	1.158	34
		<i>uds, c, b</i> tag	0.977	51	PHENIX [20, 21]	0.20 TeV	π^0	1.177	22
ALEPH [38]	91.2 GeV	incl.	1.012	22		0.51 TeV	π^0	1.178	27
DELPHI [39]	91.2 GeV	incl.	1.000	17	ALICE [12–15]	0.90 TeV	π^0	1.012	7
		<i>uds, b</i> tag	1.000	34		2.76 TeV	π^0	1.002	24
OPAL [40]	91.2 GeV	incl.	1.000	21		2.76 TeV	π^\pm	0.959	38
		<i>u, d, s, c, b</i> tag	0.793	25		7 TeV	π^0	1.016	25
BABAR [10]	10.54 GeV	incl.	1.060	45		7 TeV	π^\pm	0.976	32
BELLE [9]	10.52 GeV	incl.	1.067	78		8 TeV	π^0, π^\pm	1.048	36
SIA data (sum)				378		13 TeV	π^\pm	0.981	32
HERMES [33]	π^+, π^-	(p- Q^2)	0.984	56	PP data (sum)				348
	π^+, π^-	(d- Q^2)	0.988	56	TOTAL:				1460
	π^+, π^-	(p- x)	1.007	56					
	π^+, π^-	(d- x)	1.009	56					
COMPASS [11]	π^+, π^-	(d- z)	1.004	510					
SIDIS data (sum)				734					

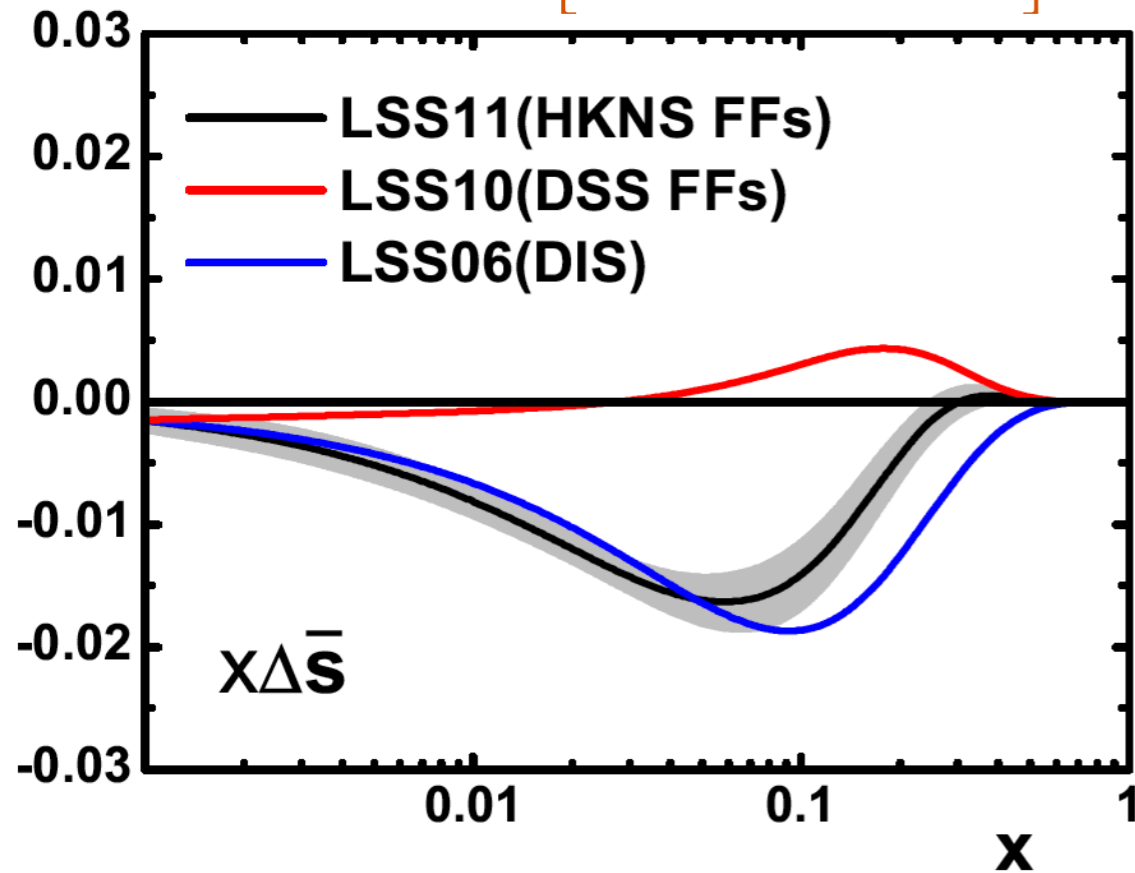
Borsa *et al.* [ArXiv:2110.14015]

BDSS analysis: almost 1500 data points for **SIA**, **SIDIS**, and **pp** for pions.

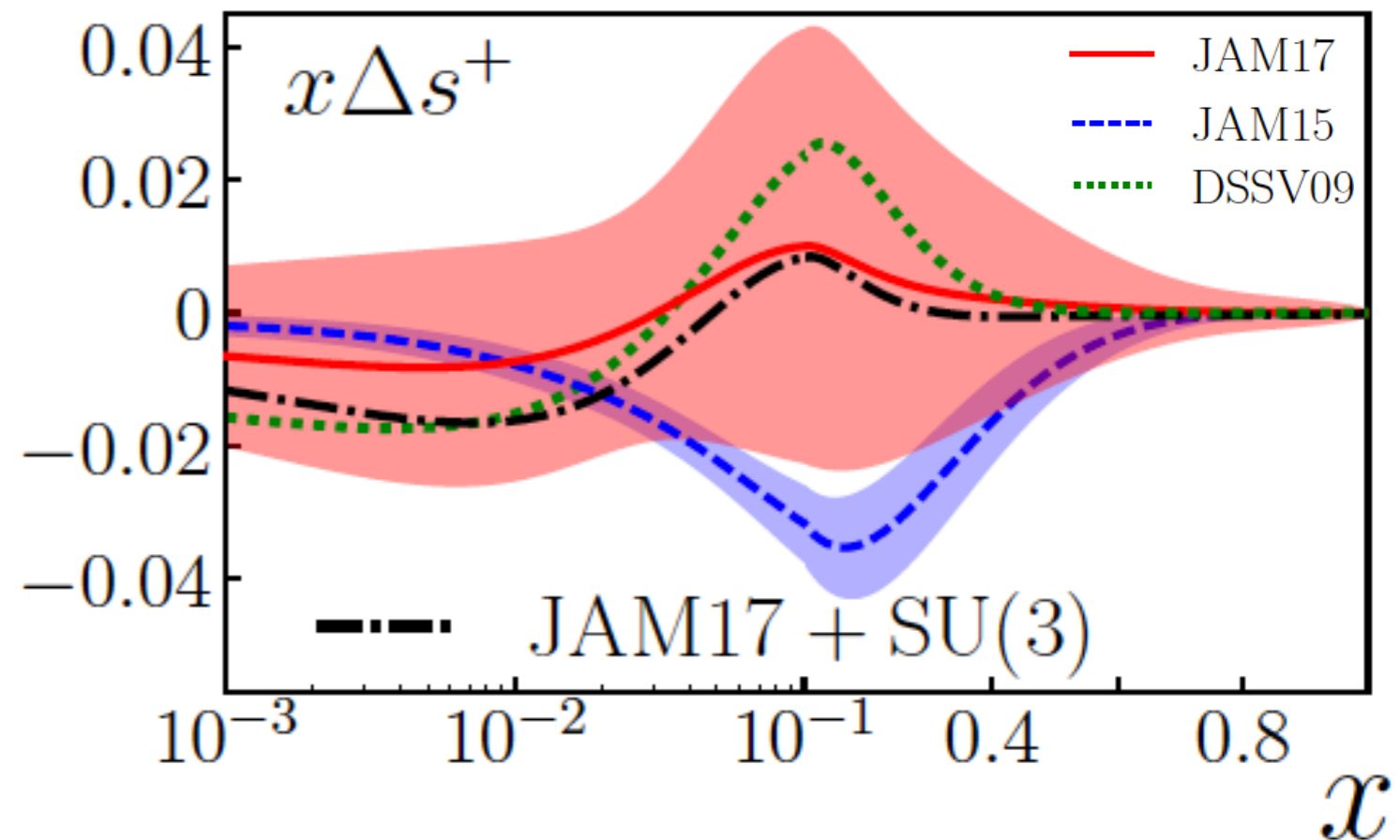
FFs and polarised PDFs

Extraction of the **longitudinally polarised parton distribution functions**:

Laeder *et al.* [ArXiv:1103.5979]



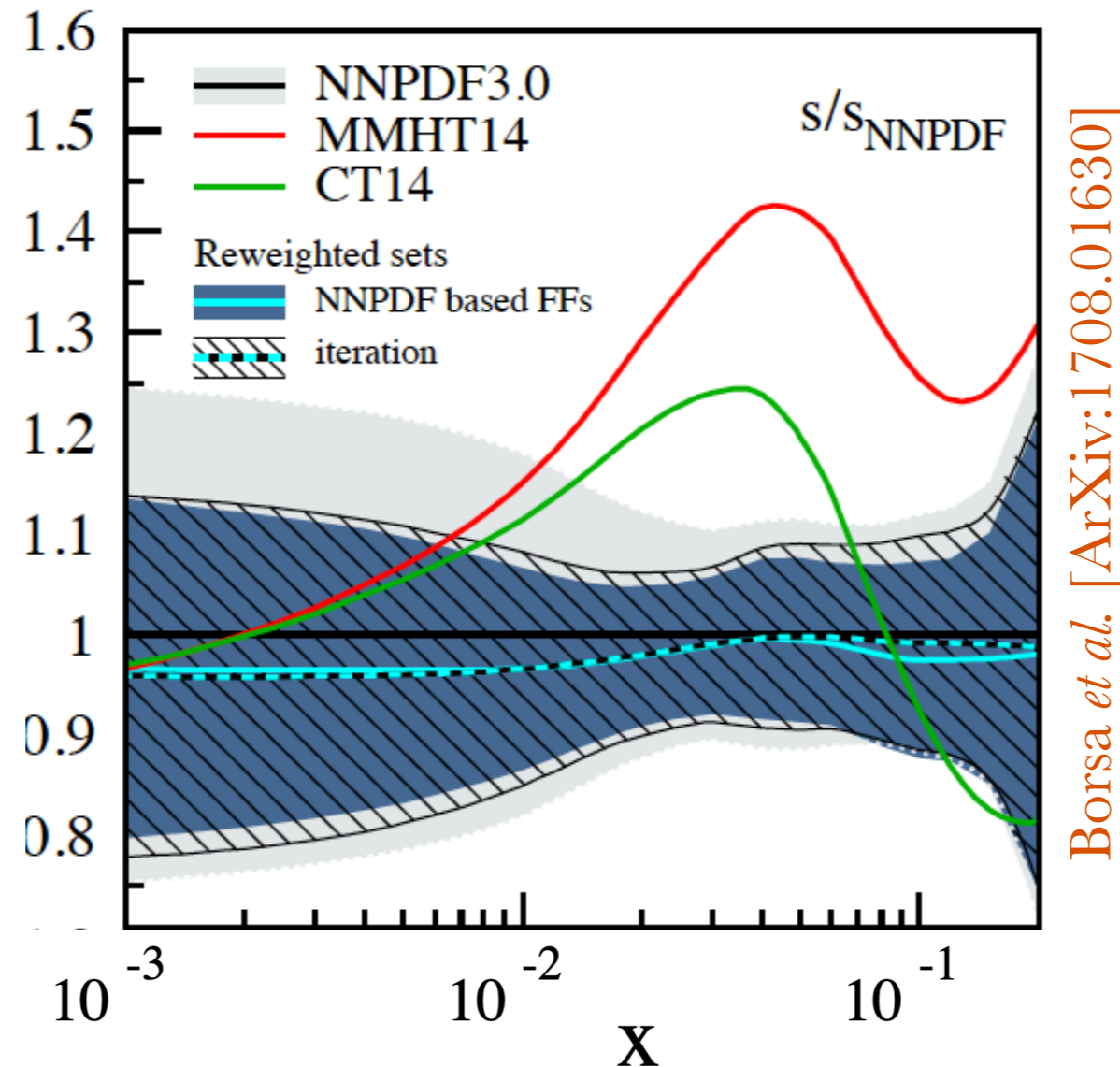
J. J. Ethier *et al.* [ArXiv:1705.05889]



- In the presence of semi-inclusive DIS (SIDIS) data the **strange quark distribution** is very sensitive to the choice of the FF set used in the analysis.
- Fitting PDFs and FFs simultaneously does not lead to a definitive answer.

FFs and unpolarised PDFs

SIDIS multiplicities depend on **PDFs** and thus the precise data from HERMES and COMPASS are expected to help constrain PDFs (noticeably the **strange** PDFs if a kaon is produced in the final state).



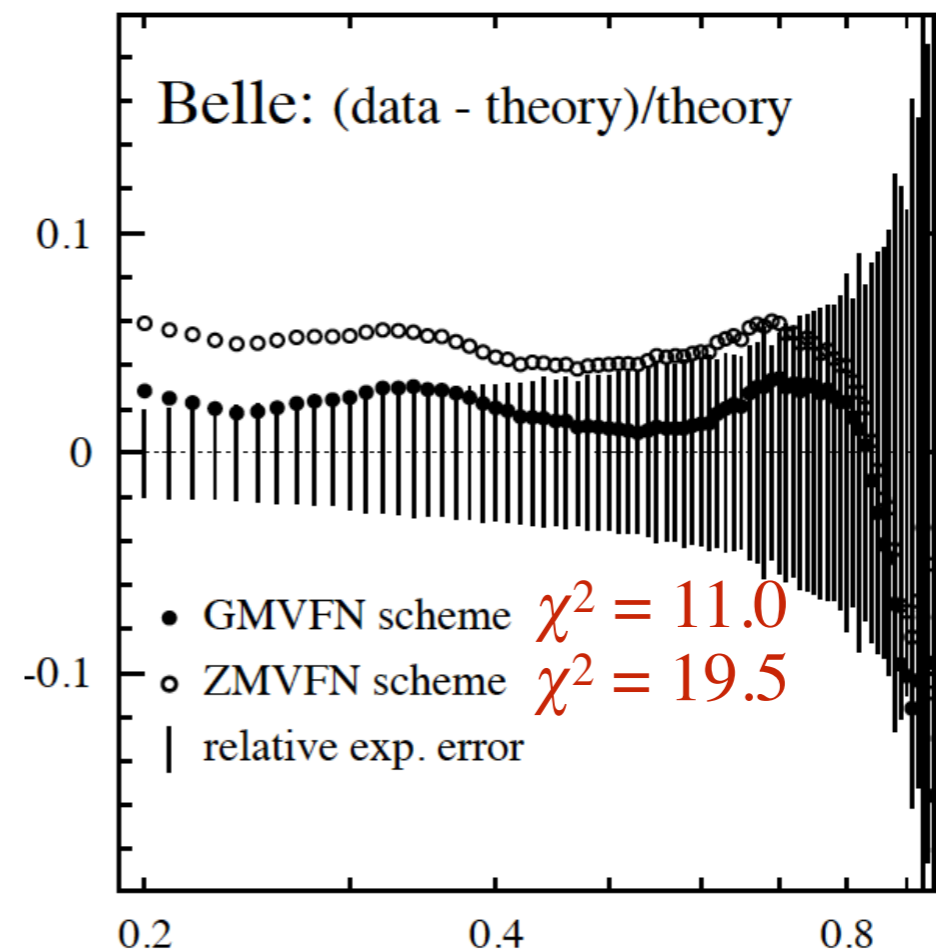
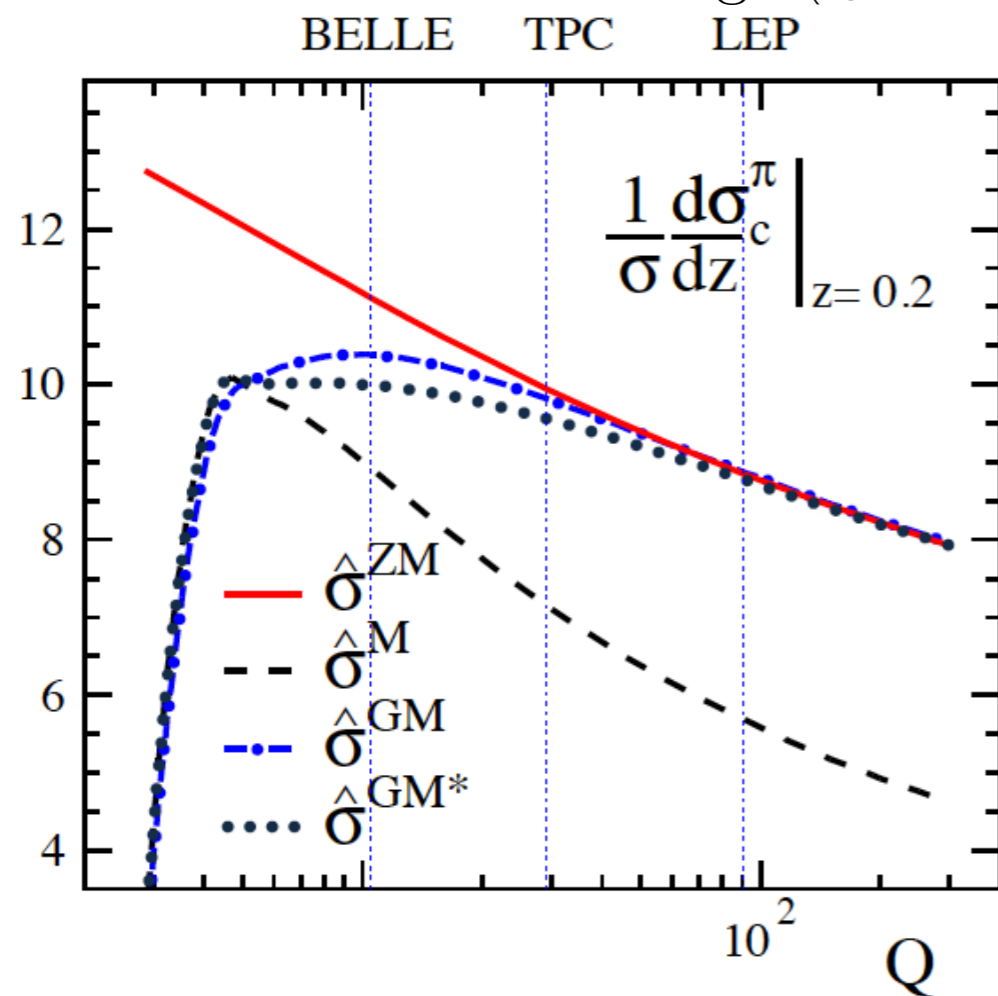
- On the other hand, SIDIS multiplicities also depend on **FFs** whose determination in turn often depends on PDFs.

A **combined** extraction of PDFs and FFs is a possible way to overcome this limitation.

Recent results

Impact of mass corrections on FFs

Heavy-quark mass corrections relevant for $Q \gtrsim m_h$. The precision of the B-factories data and the kinematic coverage ($Q \approx 2m_b$) is such that these corrections are significant.



Epele *et al.* [ArXiv:1604.08427]

- Mass corrections known to $O(\alpha_s)$ for **SIA**.
- Use of **FONLL** to smoothly interpolate between massive and massless scheme:
 - appropriate both for B-factories ($Q \approx 10.5$ GeV) and LEP ($Q = M_Z$) data.
- Marked improvement of the χ^2 of **BELLE** (and BABAR).

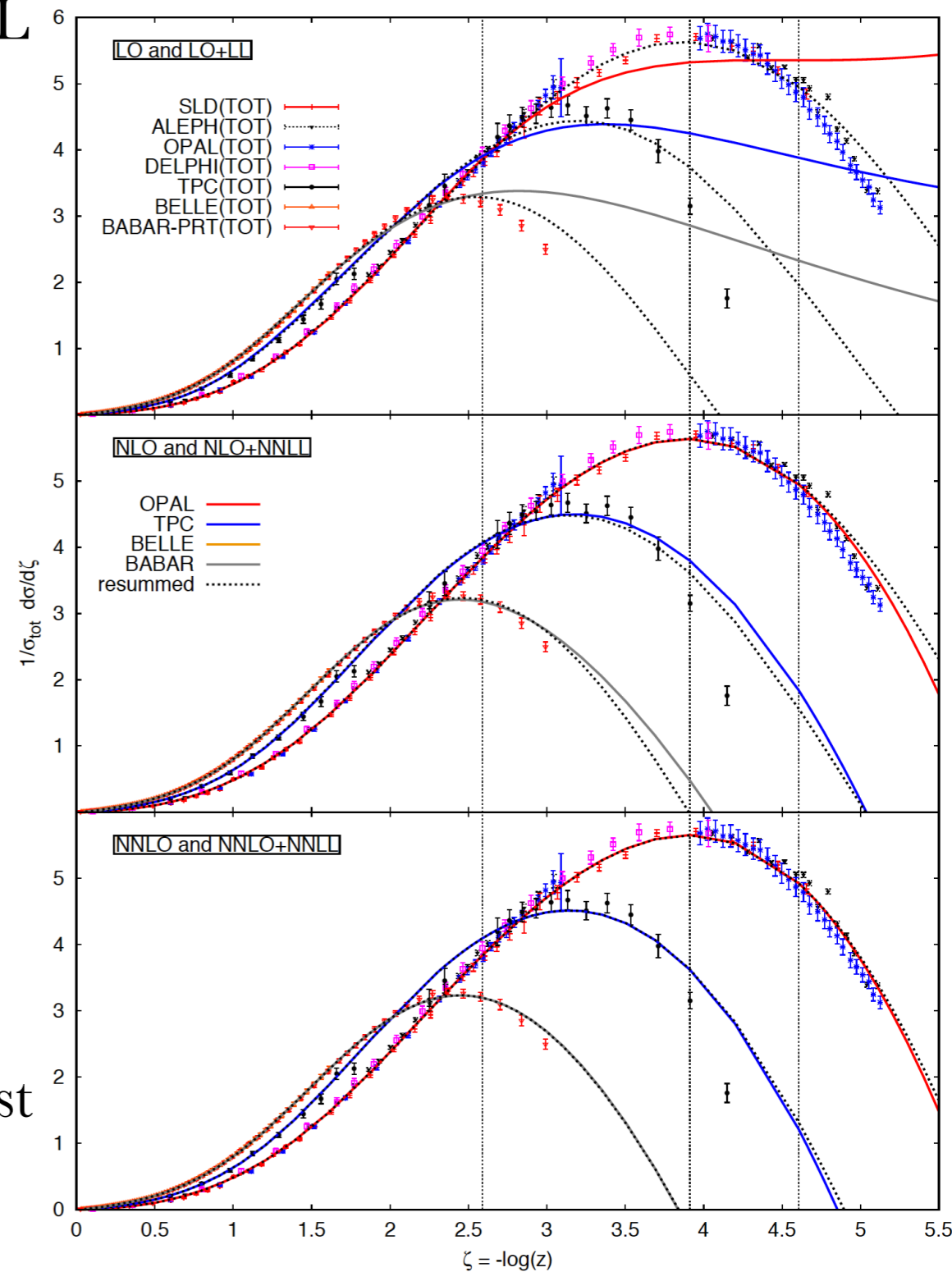
Recent results

The impact of small- z resummation corrections

- **Small- z resummation** corrections up to NNLL are implemented in the **SIA** hard cross sections and the **DGLAP** splitting functions.
- **Fits** at the different orders are performed to a variety of **pion** SIA data.

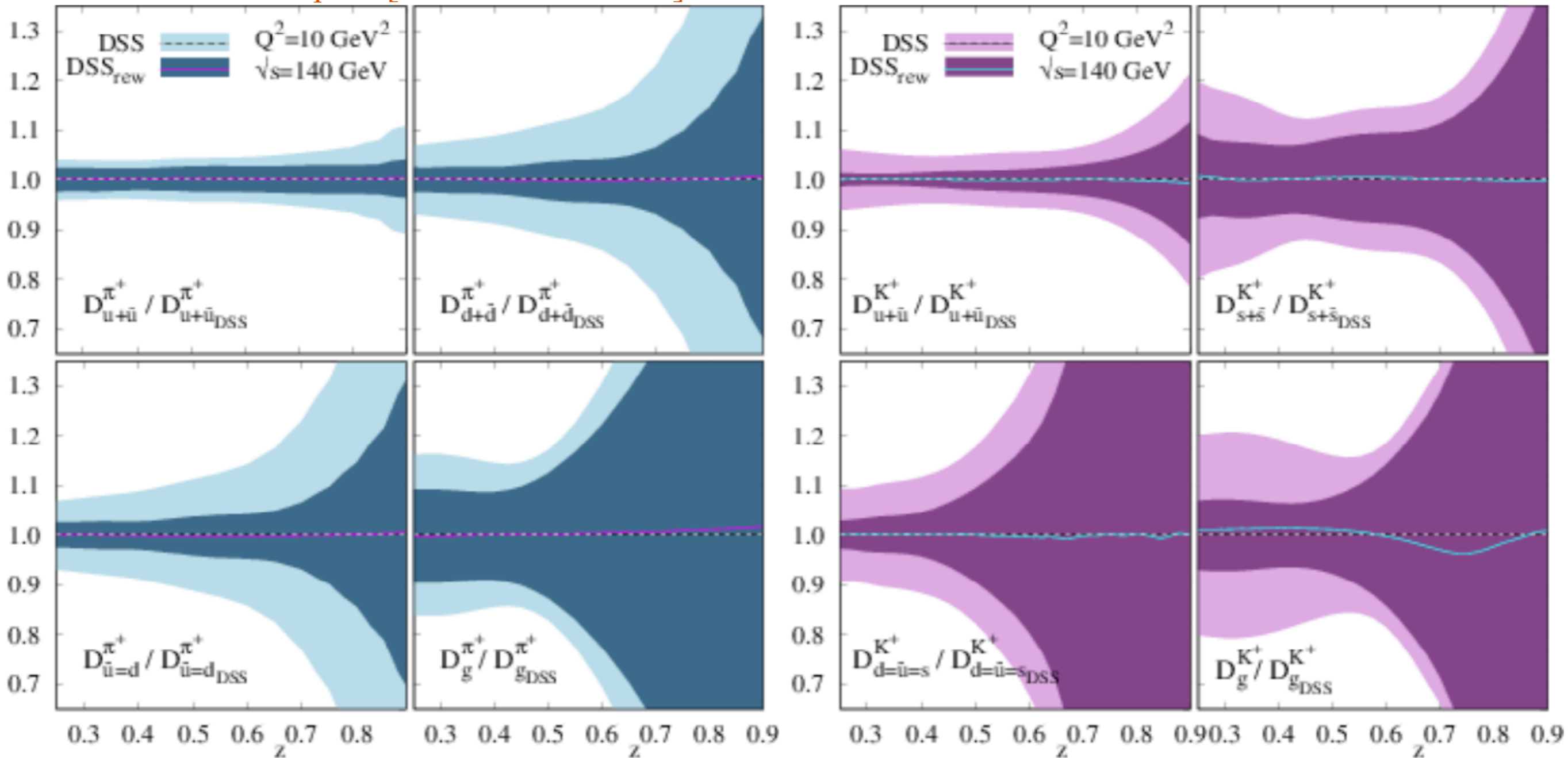
accuracy	χ^2
LO	1260.78
NLO	354.10
NNLO	330.08
LO+LL	405.54
NLO+NNLL	352.28
NNLO+NNLL	329.96

- Beyond LO, resummation provides only a **small improvement** in the kinematic region of interest w.r.t. the to fixed-order (particularly at NNLO).
- Important to study the interplay with **hadron mass corrections**, effective in the same region.



Impact of the EIC on FFs

EIC Yellow Report [ArXiv:2103.05419]



EIC data is expected to have a significant impact on pion (left) and kaon (right) FFs:

- particularly pronounced for pion unfavoured FFs and kaon favoured FFs.