Extraction of collinear FFs from SIDIS data

Valerio Bertone
IRFU, CEA, Université Paris-Saclay
In collaboration with Rabah Abdul Khalek and Emanuele R. Nocera

November 18, 2020, IWHSS 2020

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement № 824093

An extended version of the title

Preliminary results on the extraction of collinear unpolarised light-hadron (pion) FFs exploiting SIDIS data

Most recent determinations
Light hadrons

	DEHSS [arXiv:1410.6027]…	**NNFF1.0** [arXiv:1706.07049]	**JAM** [arXiv:1905.03788]
Parameterisation	standard	neural networks	standard
Error propagation	Hessian	Monte Carlo	Monte Carlo
Dataset	SIA, SIDIS, $p\overline{p}$	SIA, ($p\overline{p}$)	SIA, SIDIS, (DIS, DY)
Hadronic species	$\pi^\pm$, $K^\pm$, $p/\overline{p}$, $h^\pm$	$\pi^\pm$, $K^\pm$, $p/\overline{p}$, ($h^\pm$)	$\pi^\pm$, $K^\pm$
Perturbative orders	LO, NLO	LO, NLO, NNLO	NLO

- Heavier hadronic species also available (D^*, Λ, etc.).
- HKKS set [arXiv:1608.04067] not publicly available.
- FF fitters have started releasing their sets in the **LHAPDF format**.

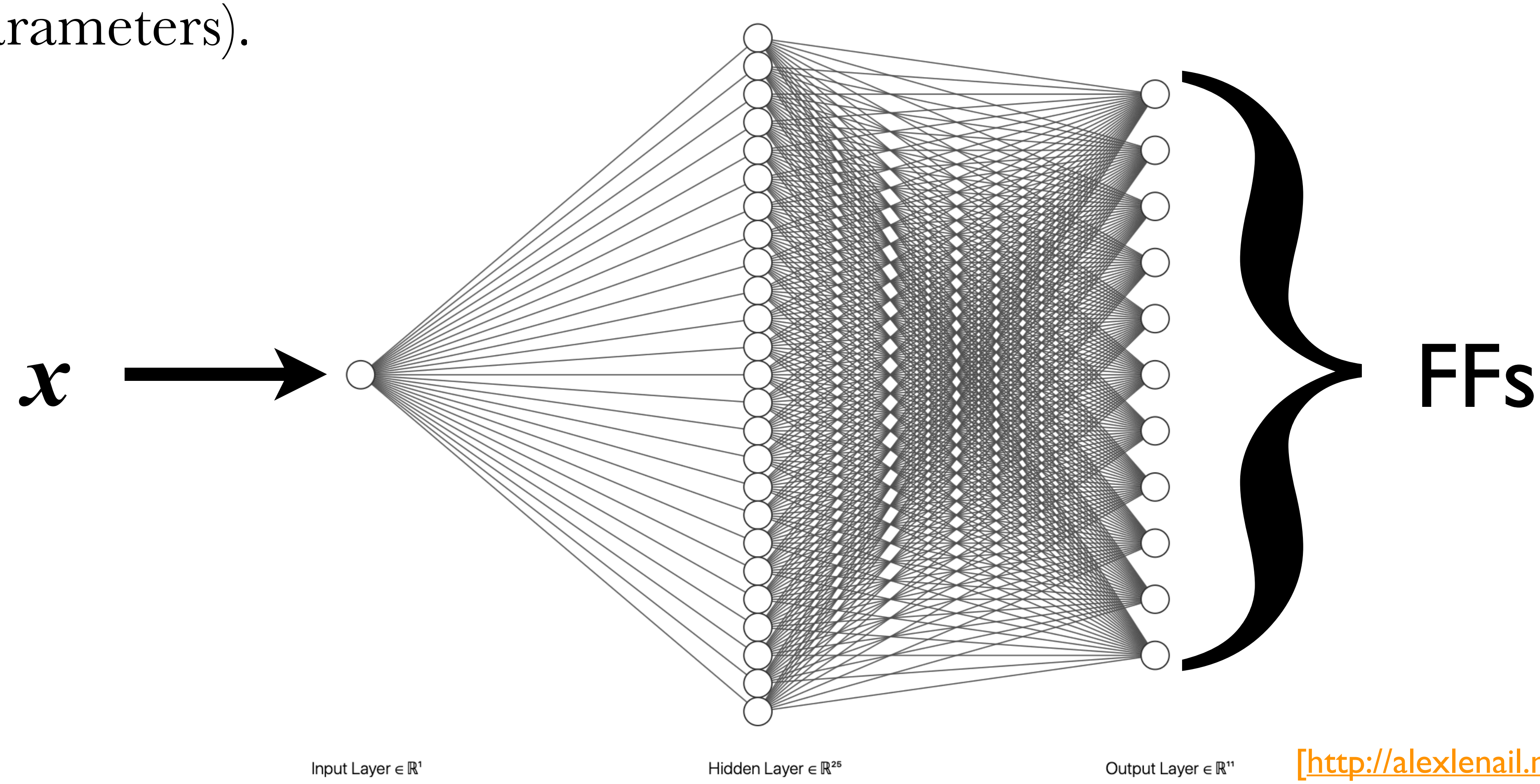
A brand new collaboration
The MAP Collaboration

- **MAP** stands for: Multi-dimensional Analyses of Partonic distributions:
  - it is a newly self-aggregated group of people interested in the hadronic structure (*e.g.* collinear and TMD distributions as well as GPDs).
  - At the moment it comprises people from Amsterdam, Edinburgh, Paris, and Pavia.
- *First* on-going effort aimed at determining unpolarised collinear FFs.
- This particular determination exploits the expertise that some of us have gathered in other environments but within a different context.
- We are all users/developers of public tools such as:
  - **NangaParbat**: a TMD fitting framework [https://github.com/vbertone/NangaParbat],
  - **APFEL++**: A PDF evolution library in C++ [https://github.com/vbertone/apfelxx],
  - **NNAD**: Neural Network Analytic Derivatives [https://github.com/rabah-khalek/NNAD],
  - **PARTONS**: PARtonic Tomography Of Nucleon Software [http://partons.cea.fr/partons/doc/html/index.html],
  - **TMDlib**: a library and plotting tools for TMDs [https://tmdlib.hepforge.org].

The MAP FF set

The parameterisation

- **All** fitted fragmentation functions are parameterised using a **single** NN:
  - architecture still not fixed yet (indicatively, *one single* hidden layer with around 200 free parameters).

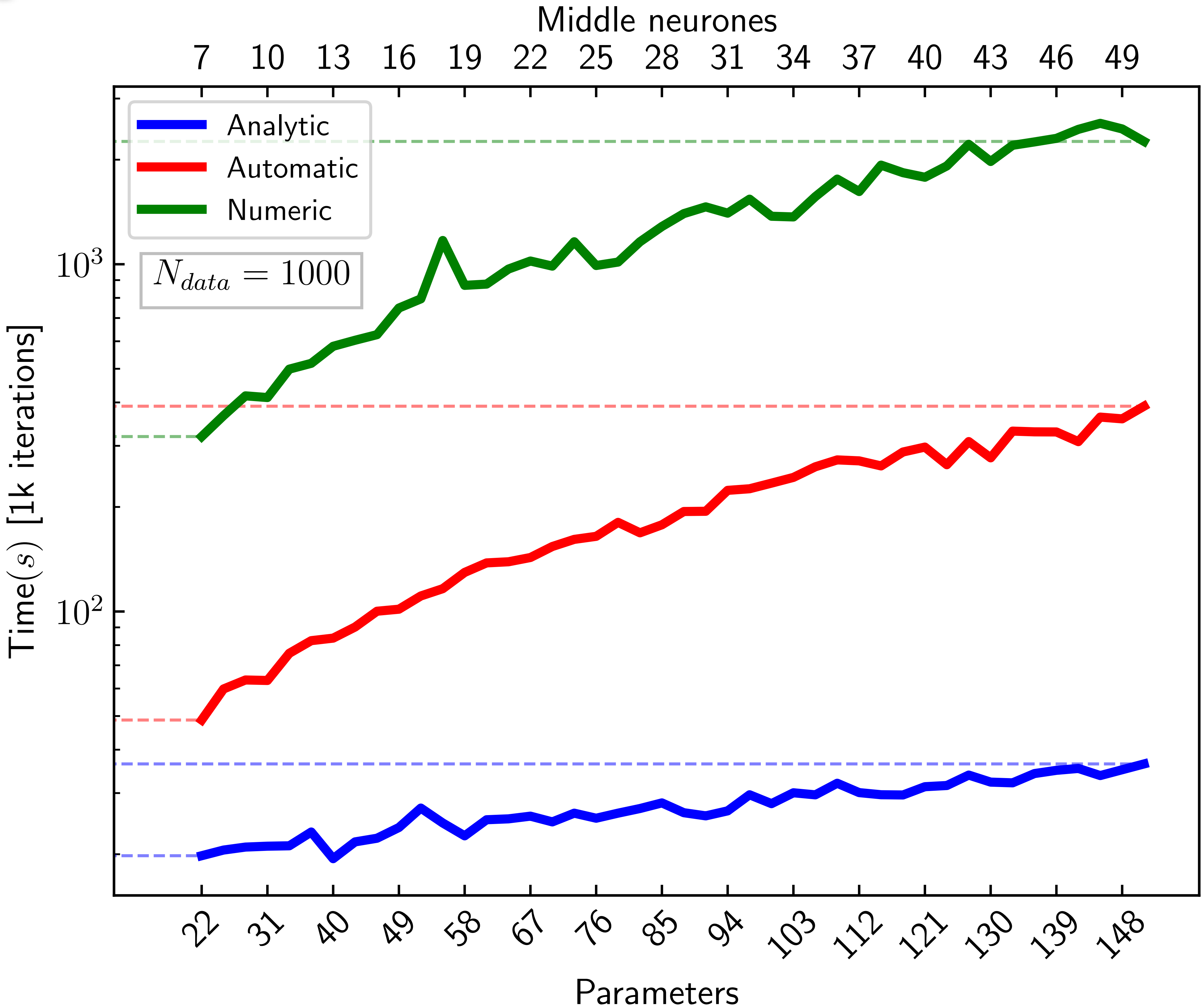


[http://alexlenail.me/NN-SVG/index.html]

  - For the first time, we exploit the fact that we are able to compute the **analytic derivative** 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 because the gradient of the χ^2 can be analytically computed (as opposed to numerical and analytic derivatives).

The MAP FF set

The parameterisation



[Abdul Khalek, Bertone, arXiv:2005.07039]

The MAP FF set

The prediction engine

- Predictions based on **collinear factorisation**:

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

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

- that also allows to derive **evolution equations** (DGLAP):

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

- **Perturbative contributions**, numerical convolutions, solution of the DGLAP equations and numerical integrations provided by **APFEL++**.
- Integration over the final-state phase space fully taken into account.

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

The MAP FF set

The prediction engine

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

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

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

- 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 $\boldsymbol{V}$,
  - computation of the systematic shifts for data-theory comparisons,
  - t_0 prescription (if necessary) to treat normalisation uncertainties.

The MAP FF set

The minimiser

- The use of NNs and the consequent **large number** of free parameters requires and 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]

> 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?.

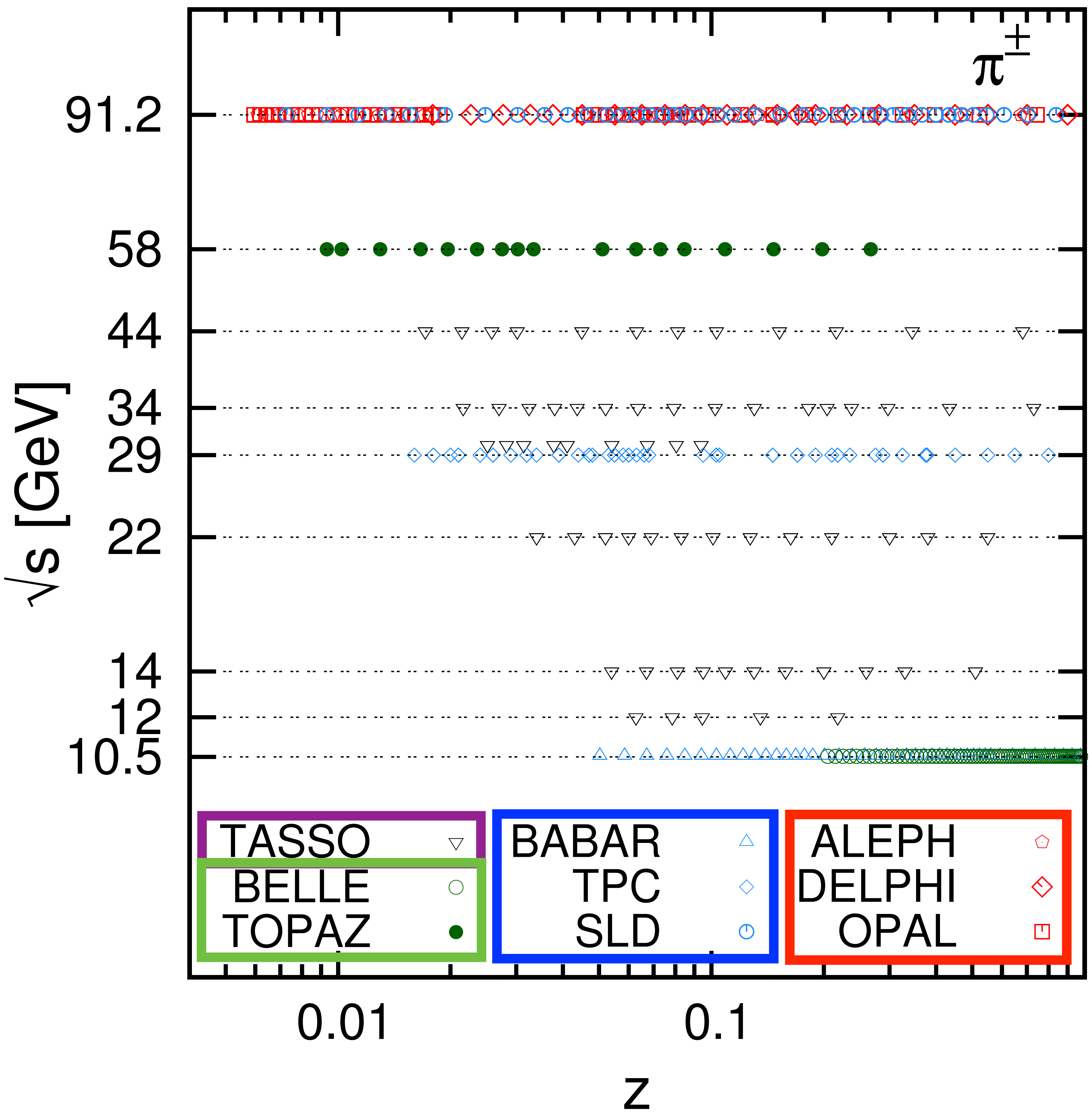
- Ceres solver is very well suited for complicated optimisation problems.

- It allows for the use of **automatic** and **analytic** differentiation.

The MAP FF set

The SIA data set

- **SIA cross sections** (normalised and absolute)



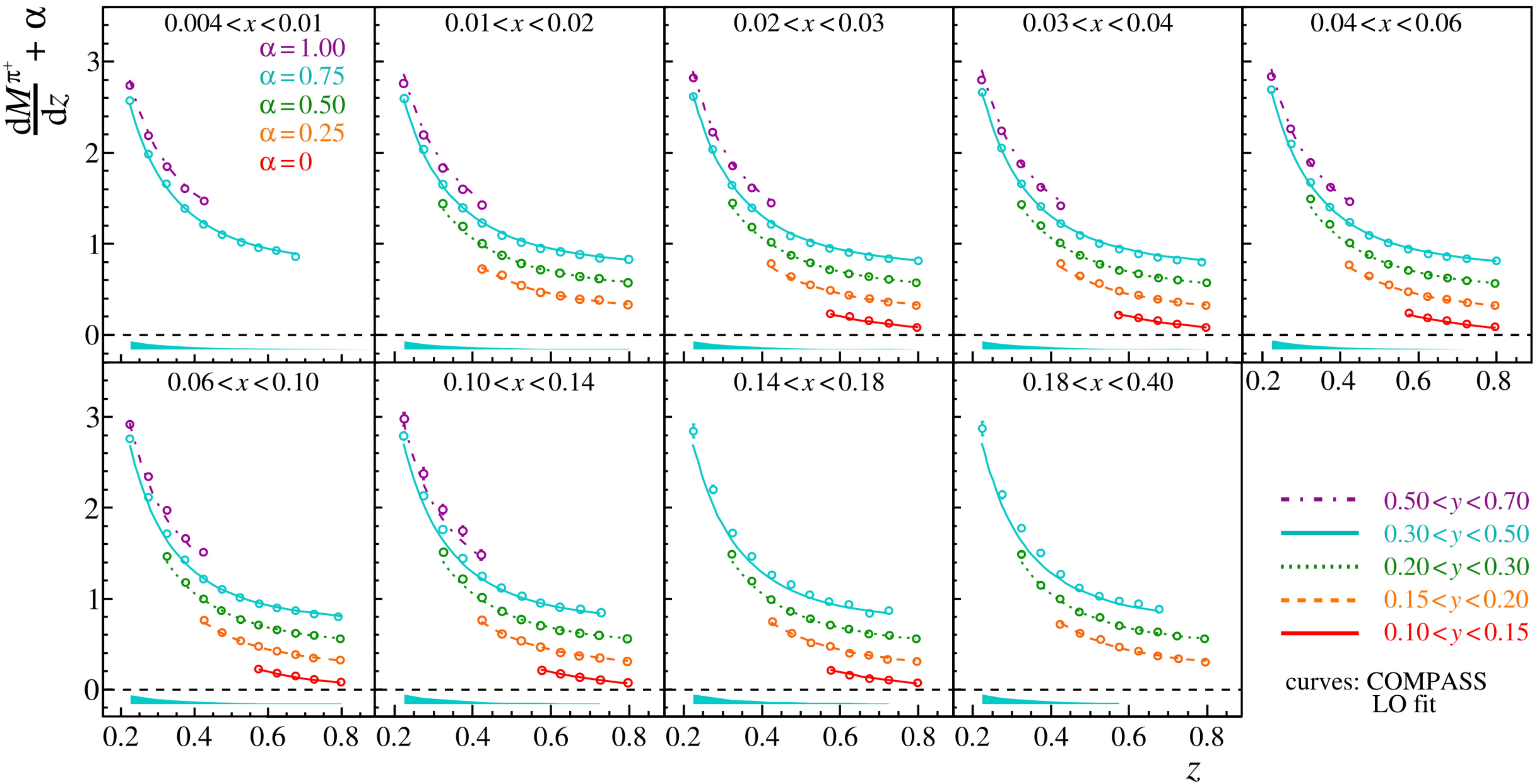
- PETRA (DESY)
  - Phys. Lett. B94 (1980) 444
  - Z.Phys. C17 (1983) 5-15
  - Z.Phys. C42 (1989) 189
- KEK
  - Phys. Rev. Lett. 111 (2013) 062002
  - Phys. Rev. D92 (2015), no. 9 092007
  - Phys. Lett. B345 (1995) 335-342
- SLAC
  - Phys. Rev. D88 (2013) 032011
  - Phys. Rev. Lett. 61 (1988) 1263
  - Phys. Rev. D69 (2004) 072003
  - X.Q. Lu, Ph.D. thesis, Johns Hopkins U., 1986
- LEP (CERN)
  - Z. Phys. C66 (1995) 355-366
  - Eur. Phys. J. C5 (1998) 585-620
  - Z. Phys. C63 (1994) 181-196.
- More than 400 points (after cuts)

The MAP FF set

The preliminary SIDIS data set

- **SIDIS multiplicities** for both π^+ and π^-:

[Physics Letters B 764 (2017) 1–10]

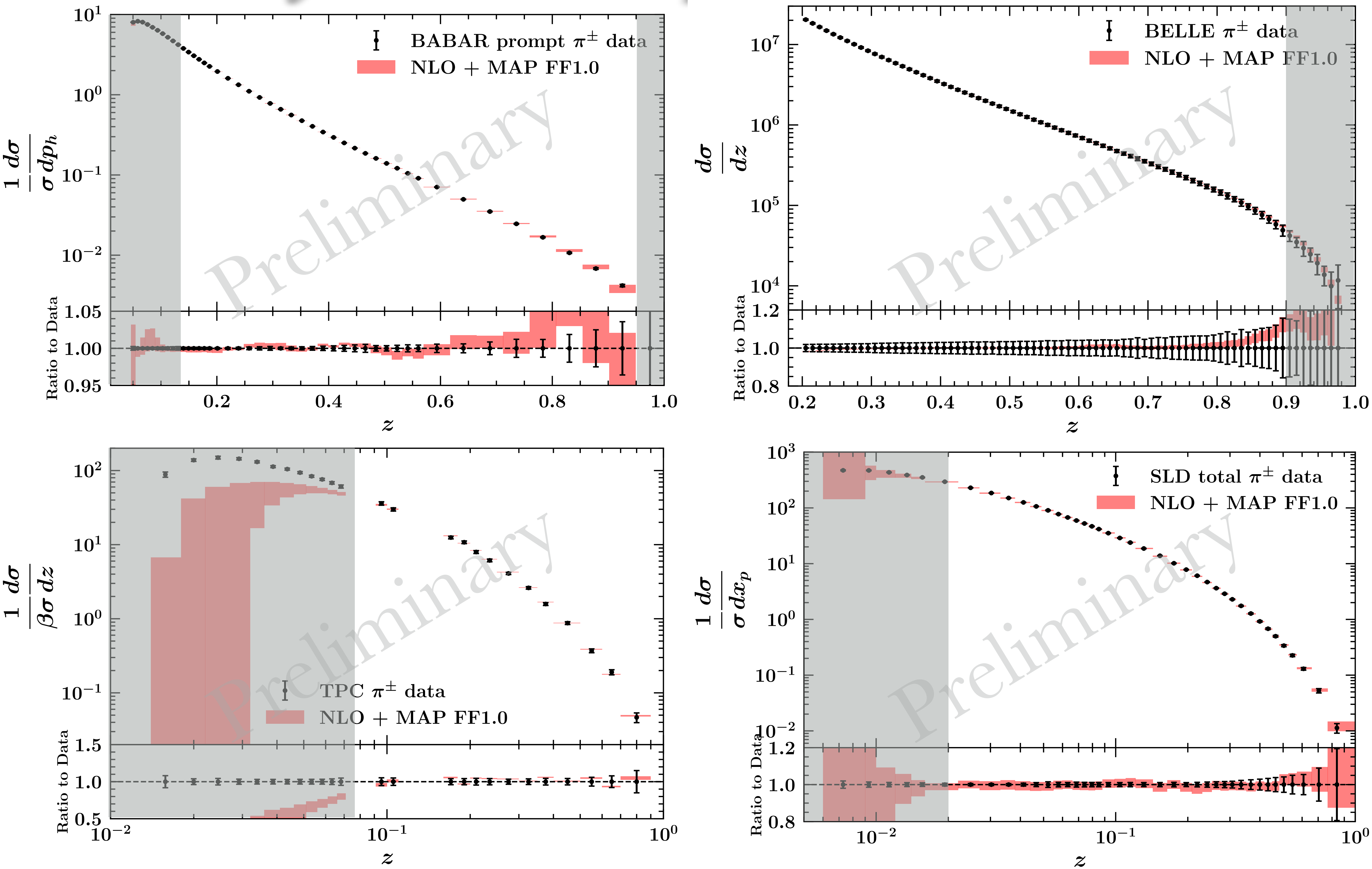


- More than 300 data points after the preliminary cut $Q > 2$ GeV.

The MAP FF set

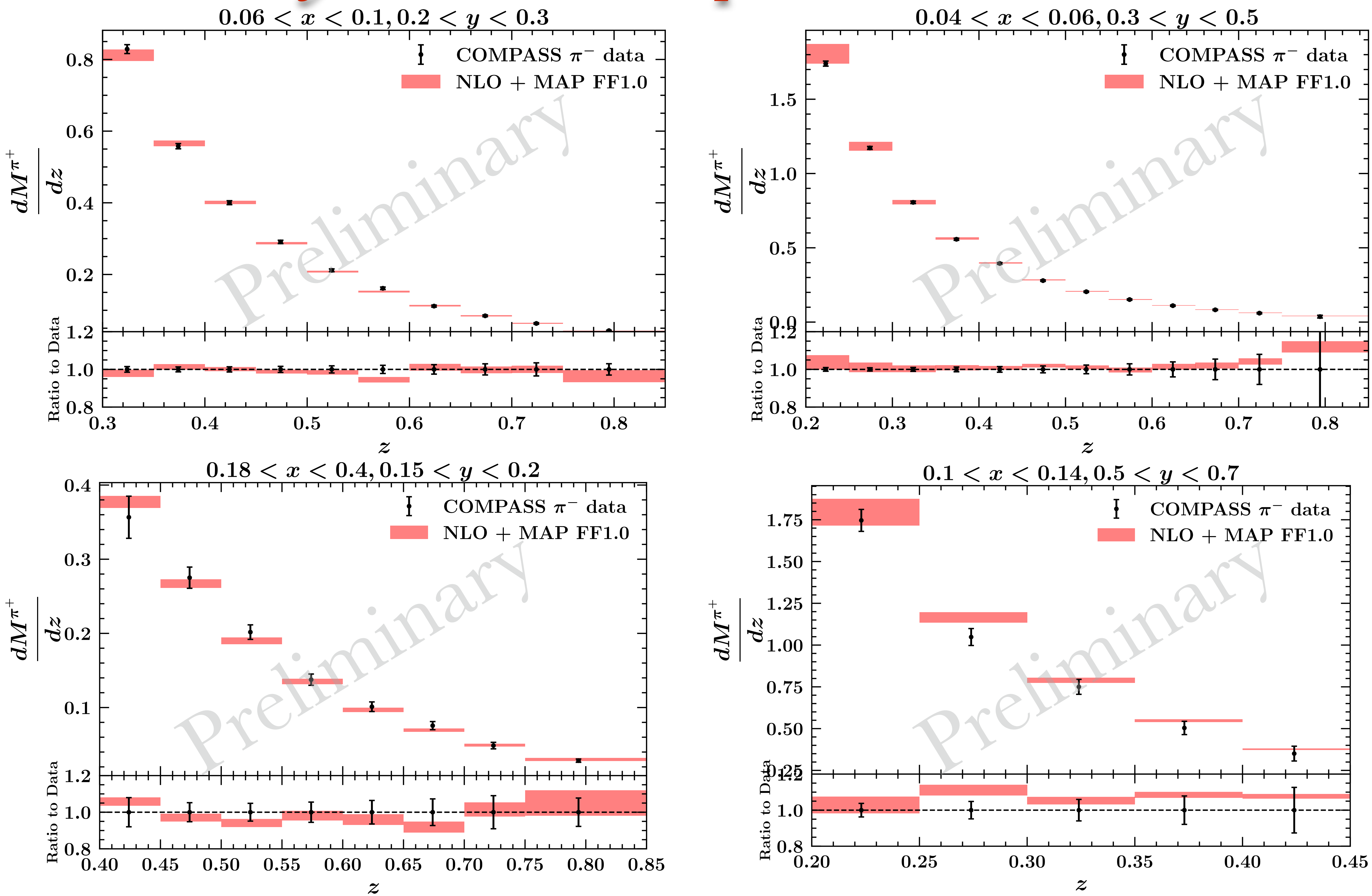
Preliminary results: data vs predictions

Not in the fit



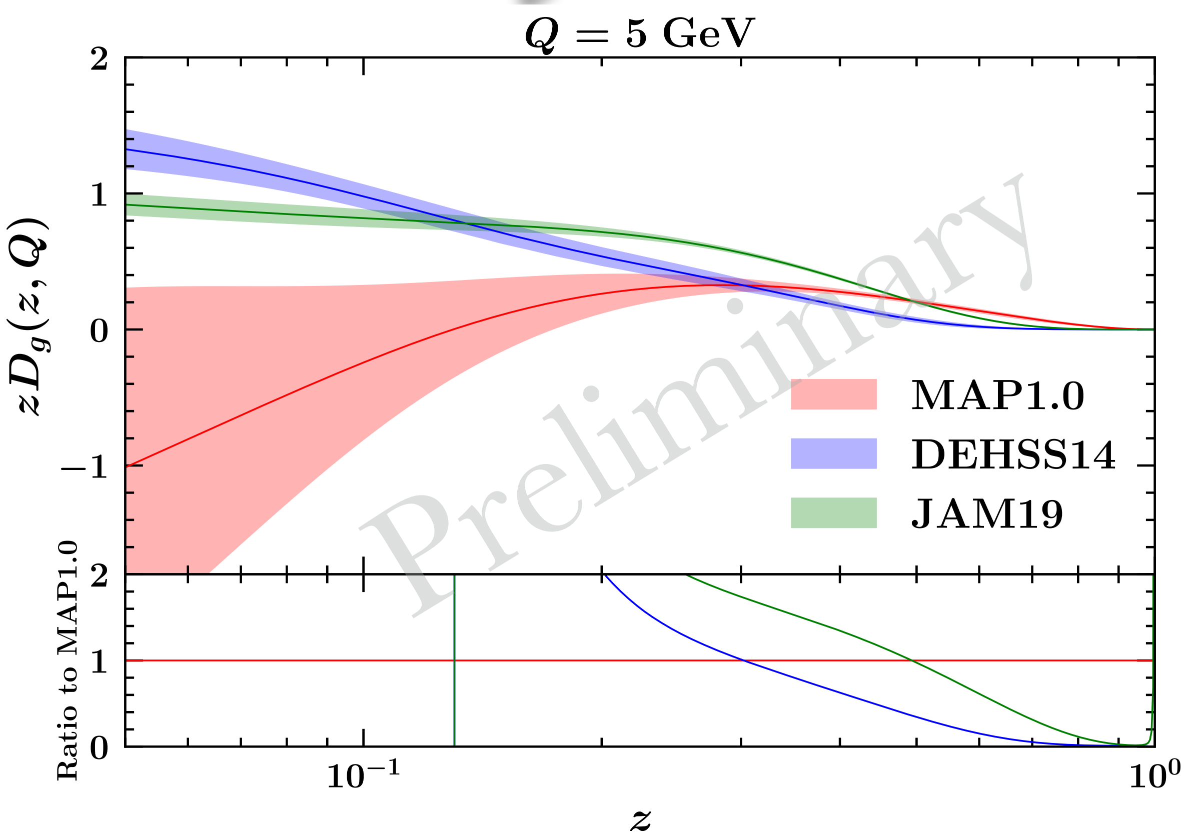
The MAP FF set

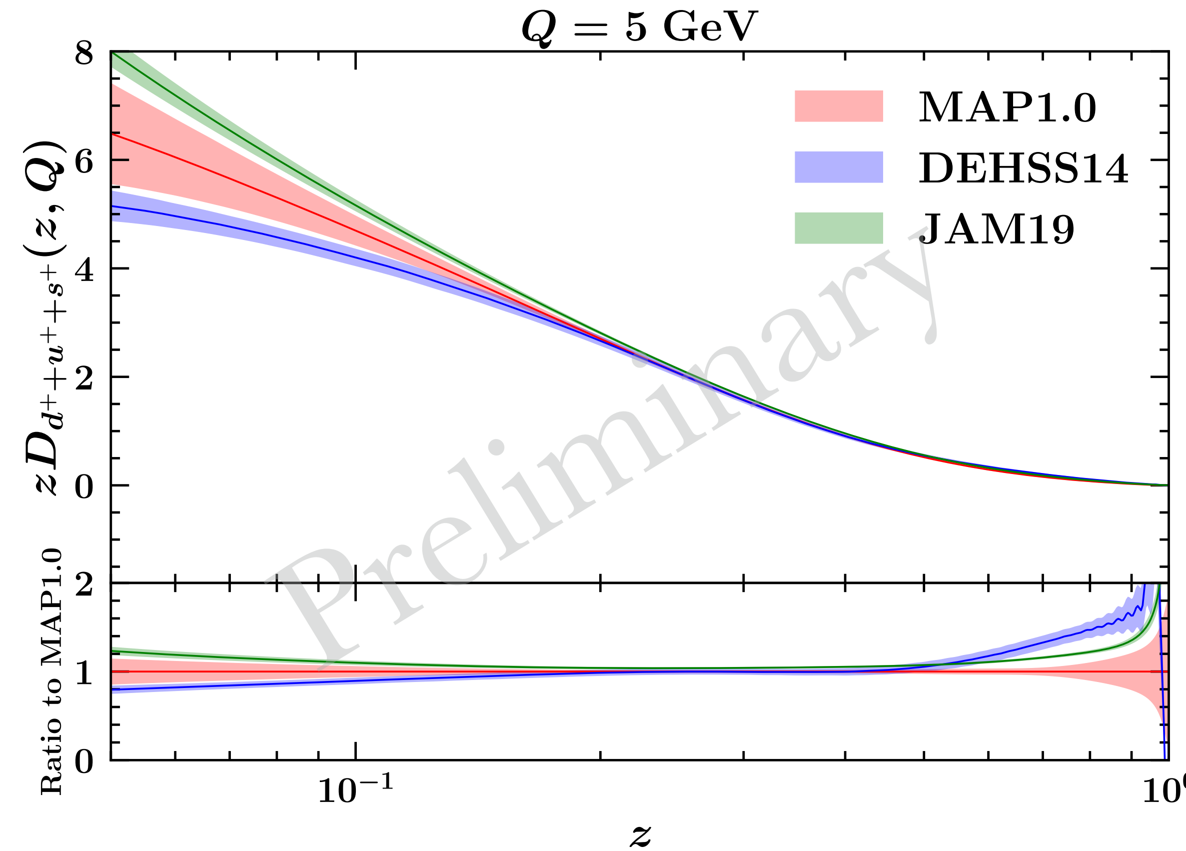
Preliminary results: data vs. predictions

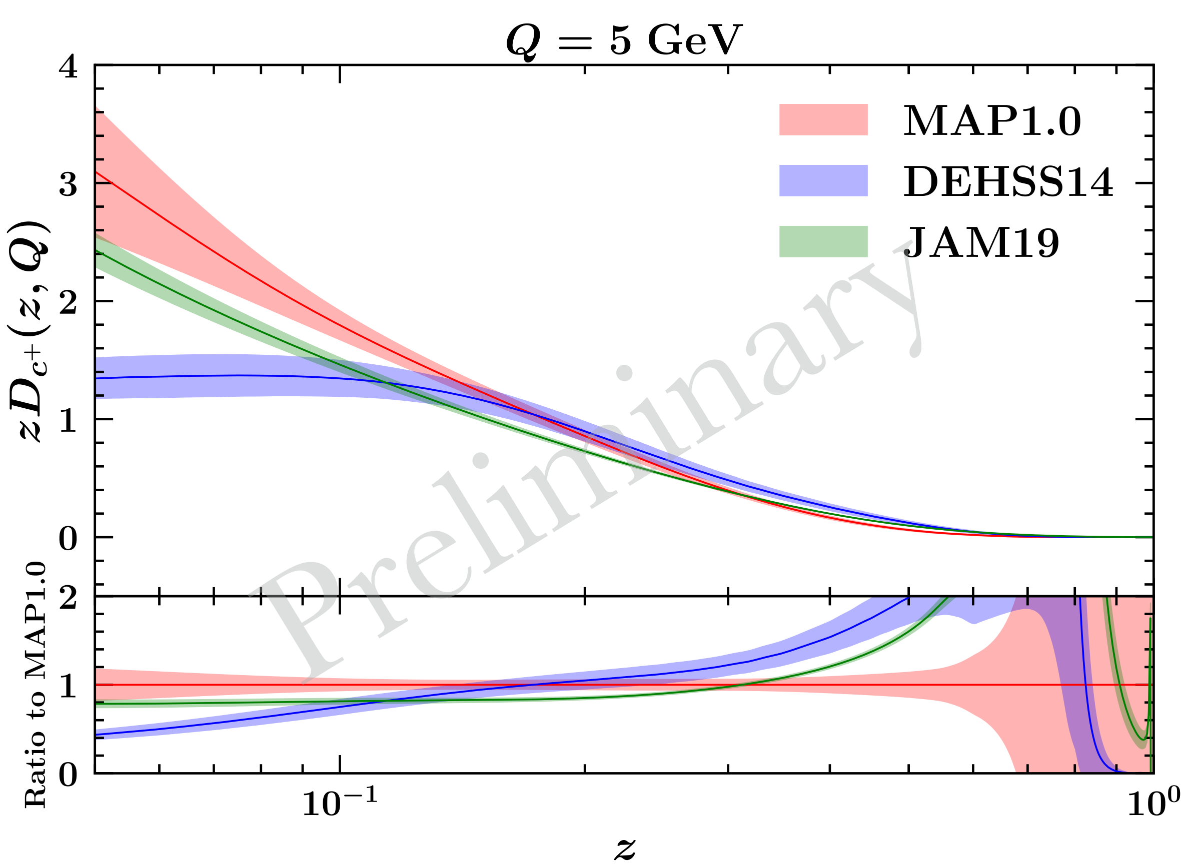


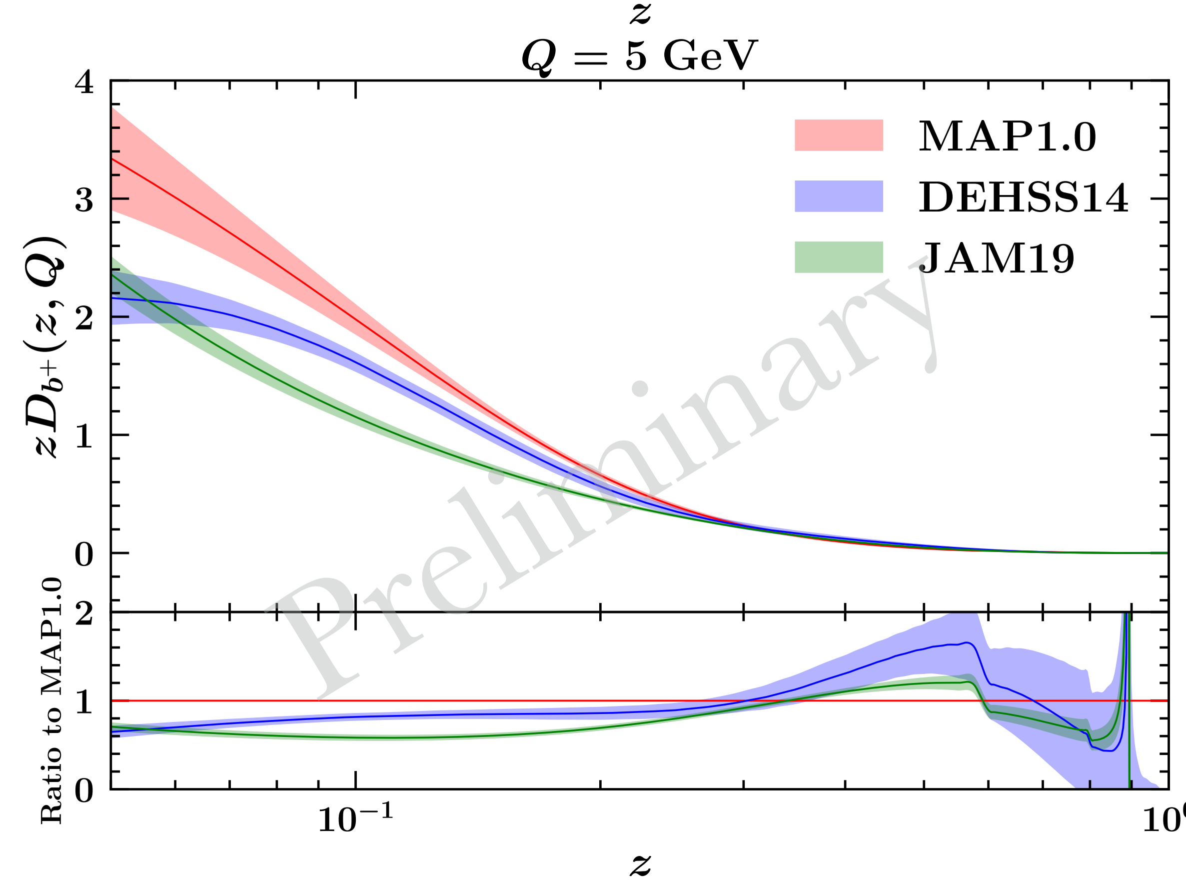
The MAP FF set

Preliminary results: the FFs









The MAP FF set

Steps left to the completion

- Inclusion on the **HERMES** data sets,
- Extend the analysis to **kaon** and **proton** FFs,
- Optimisation of the **kinematic cuts**.
- Optimisation of the **hyper-parameters**
  (NN architecture, training-validation fraction, etc.).
- Considering kaon and proton **multiplicity ratios** from COMPASS
  (thanks to Fabienne Kunne and Yann Bedfer for pointing these data sets out).
- Run the final fits, write the paper, and release the LHAPDF grids.

Conclusions

- We are in the process of completing a determination of **collinear unpolarised light-hadrons FFs**.

- The main features are:

  - totally based on **publicly available tools** for the management of the different aspects of the fit (NangaParbat, APFEL++, NNAD, Ceres-solver).

  - based on **SIA** and **SIDIS** data,

  - full treatment of the **experimental information**, particularly of the experimental correlated and uncorrelated uncertainties,

  - **full integration** over the final state kinematics,

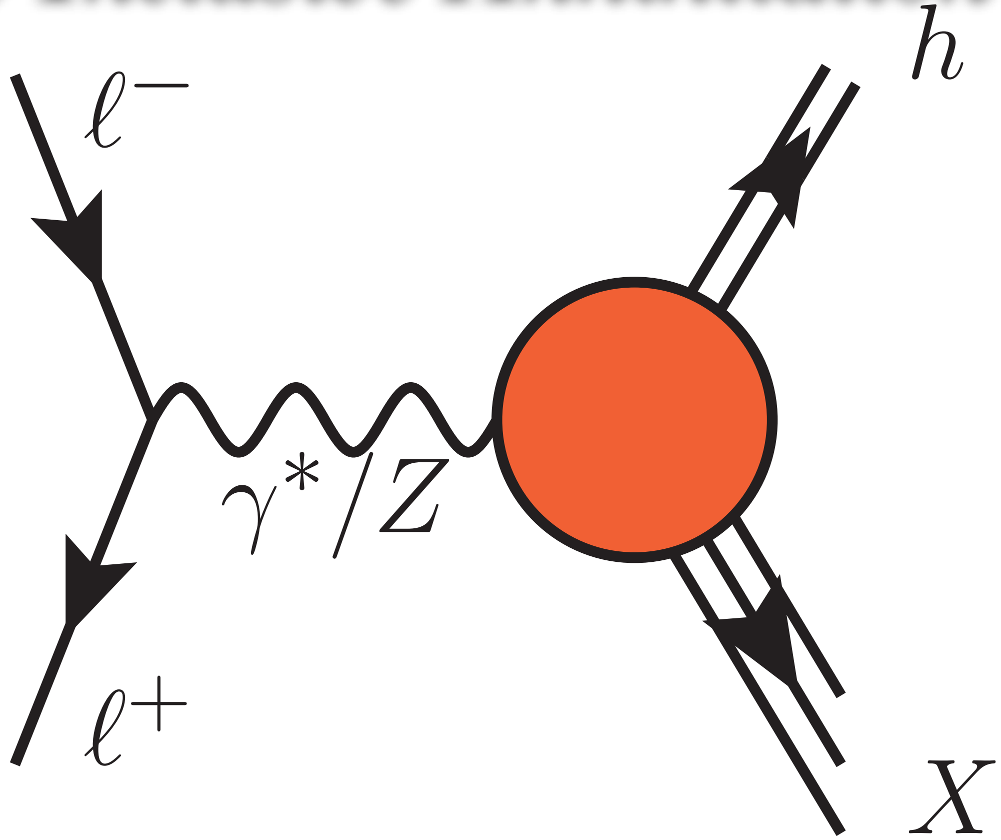
  - FFs parameterised in terms of NNs:

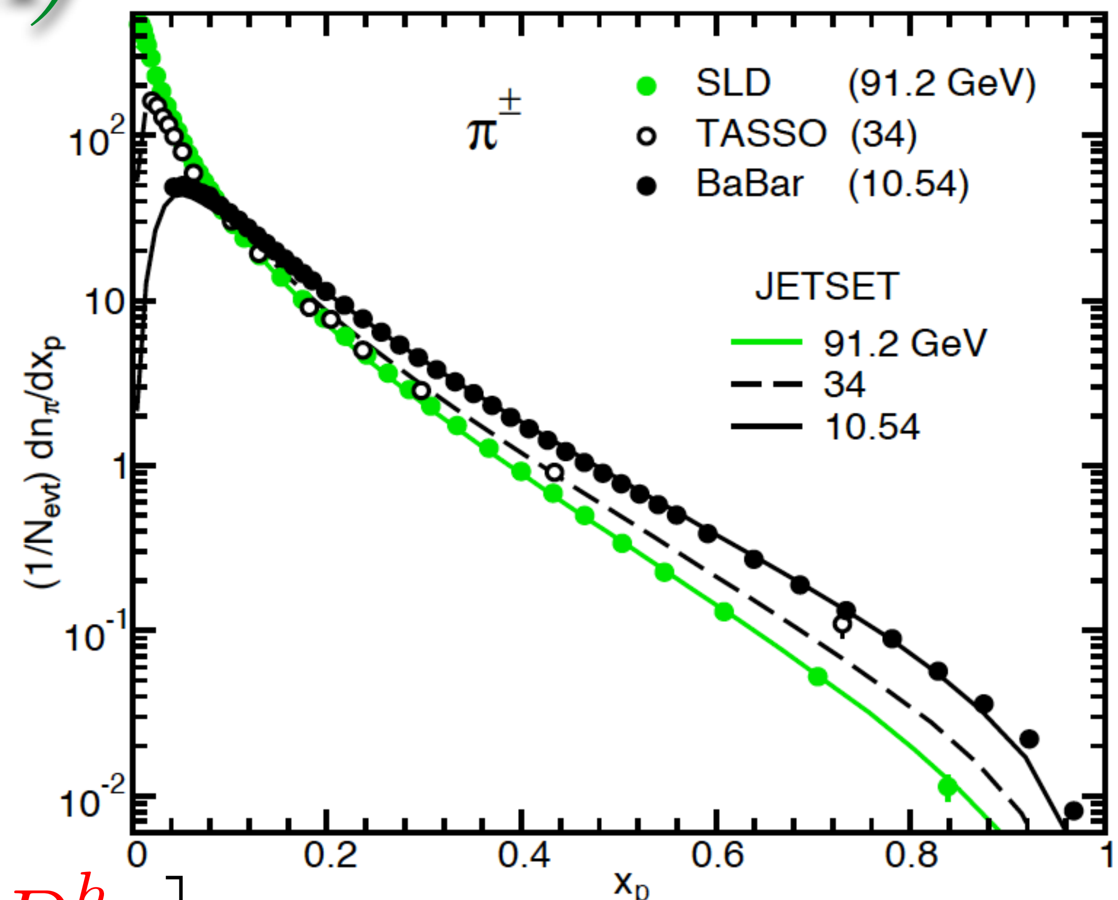
    - exploitation of the analytic gradient.

Back-up

Experimental data

Single Inclusive Annihilation (SIA)





$$\frac{d\sigma^h}{dz} = \hat{\sigma}_0^h \left[C_q \otimes D_\Sigma^h + C_g \otimes D_g^h + C_{\rm NS} \otimes D_{\rm NS}^h \right]$$

- ✅ **Clean** channel: only FFs involved,
- ✅ **higher-order** corrections to NNLO,
- ✅ **precise data** available (BELLE/BABAR).
- ❌ **No flavour separation**,
  - ✅ tagged data for heavy-quark FFs.
- ❌ gluon distribution **suppressed** by α_s.

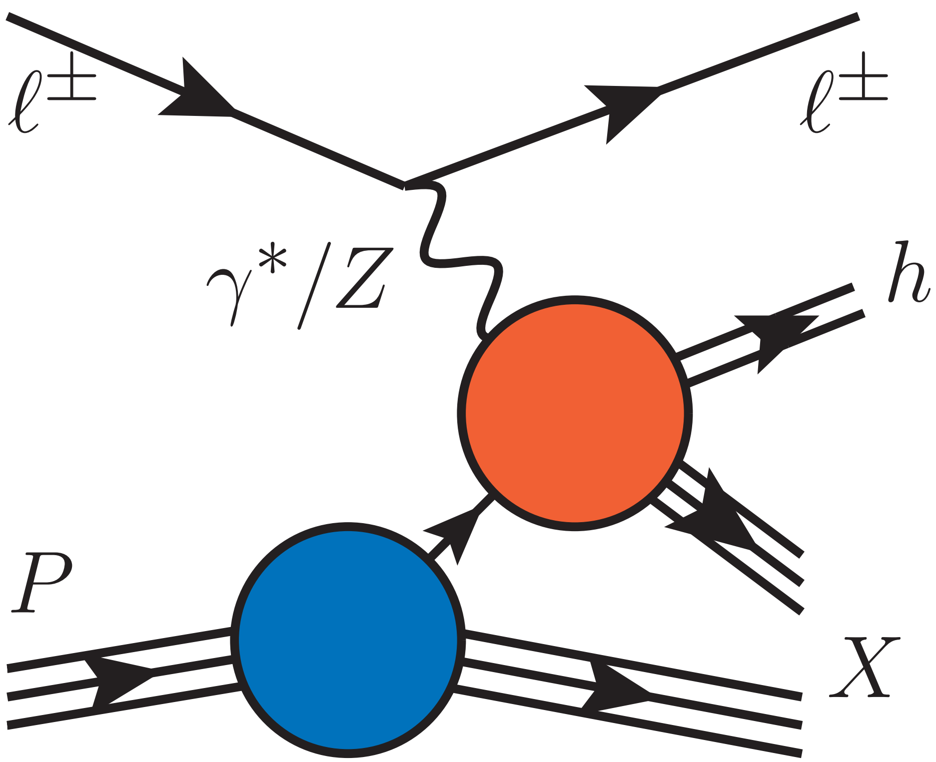
$$D_\Sigma^h = \sum_q \left(D_q^h + D_{\bar{q}}^h \right) = \sum_q D_{q^+}^h$$

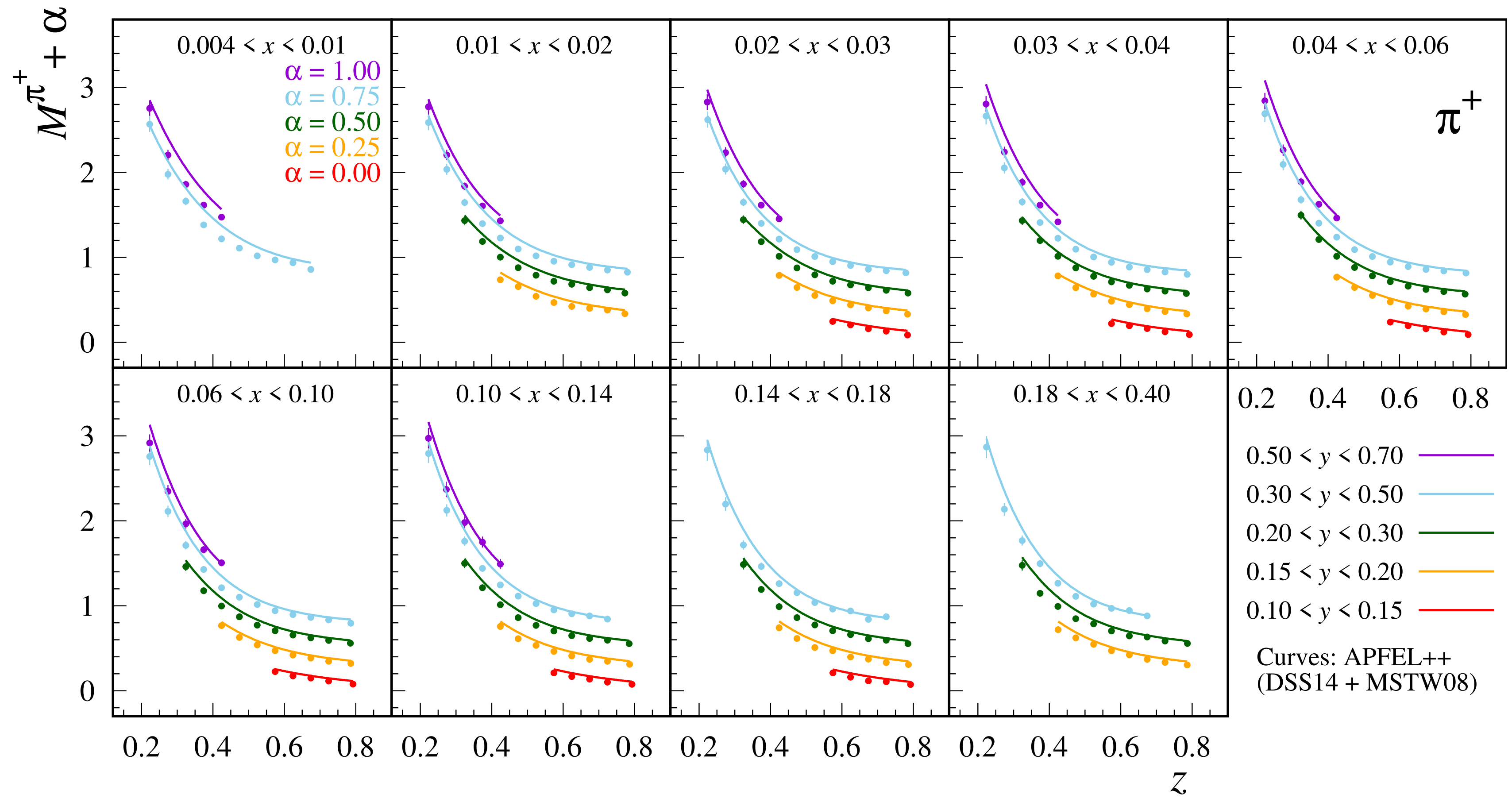
$$D_{\rm NS}^h = \sum_q \left(\frac{\hat{e}_q^2}{\langle \hat{e}_q^2 \rangle} - 1 \right) D_{q^+}^h$$

$$C_q, C_{\rm NS} \propto \mathcal{O}(1) \quad \text{while} \quad C_g \propto \mathcal{O}(\alpha_s)$$

Experimental data

Semi Inclusive Deep Inelastic Scattering (SIDIS)



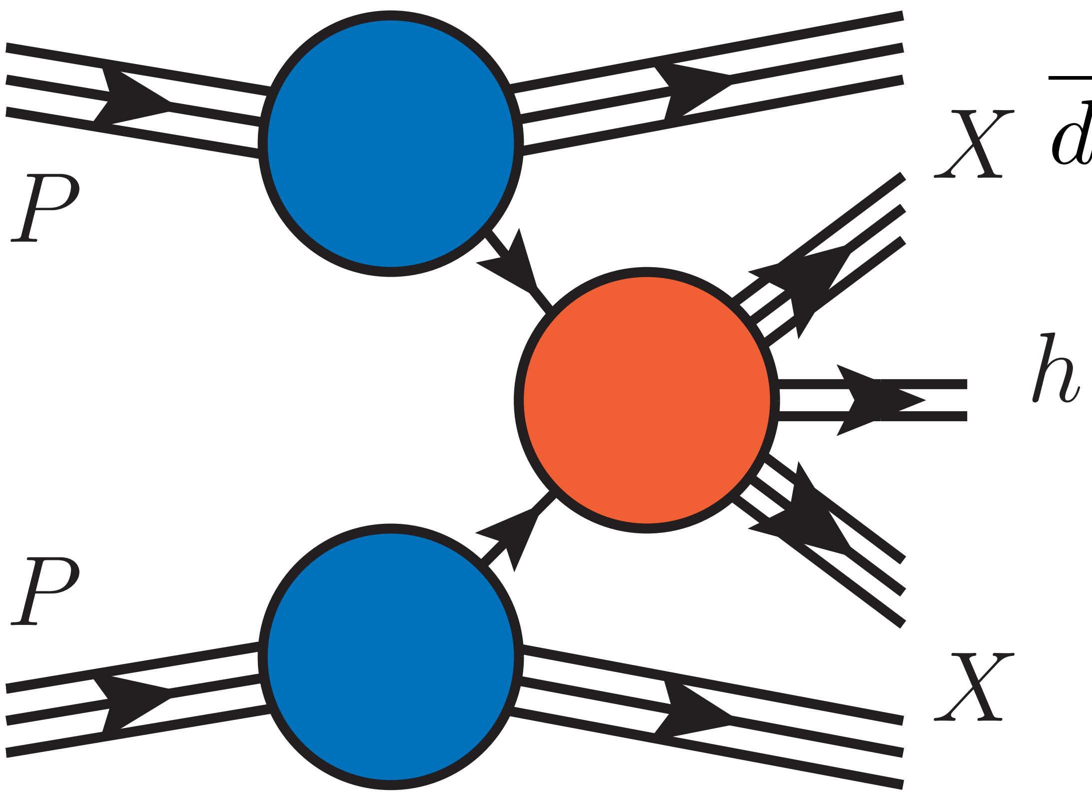


$$\frac{d\sigma^h}{dxdydz} = \hat{\sigma}_0^h \sum_{q,\overline{q}} e_q^2 \left[f_q \otimes C_{qq} \otimes D_q^h + f_g \otimes C_{gq} \otimes D_q^h + f_q \otimes C_{qg} \otimes D_g^h \right]$$

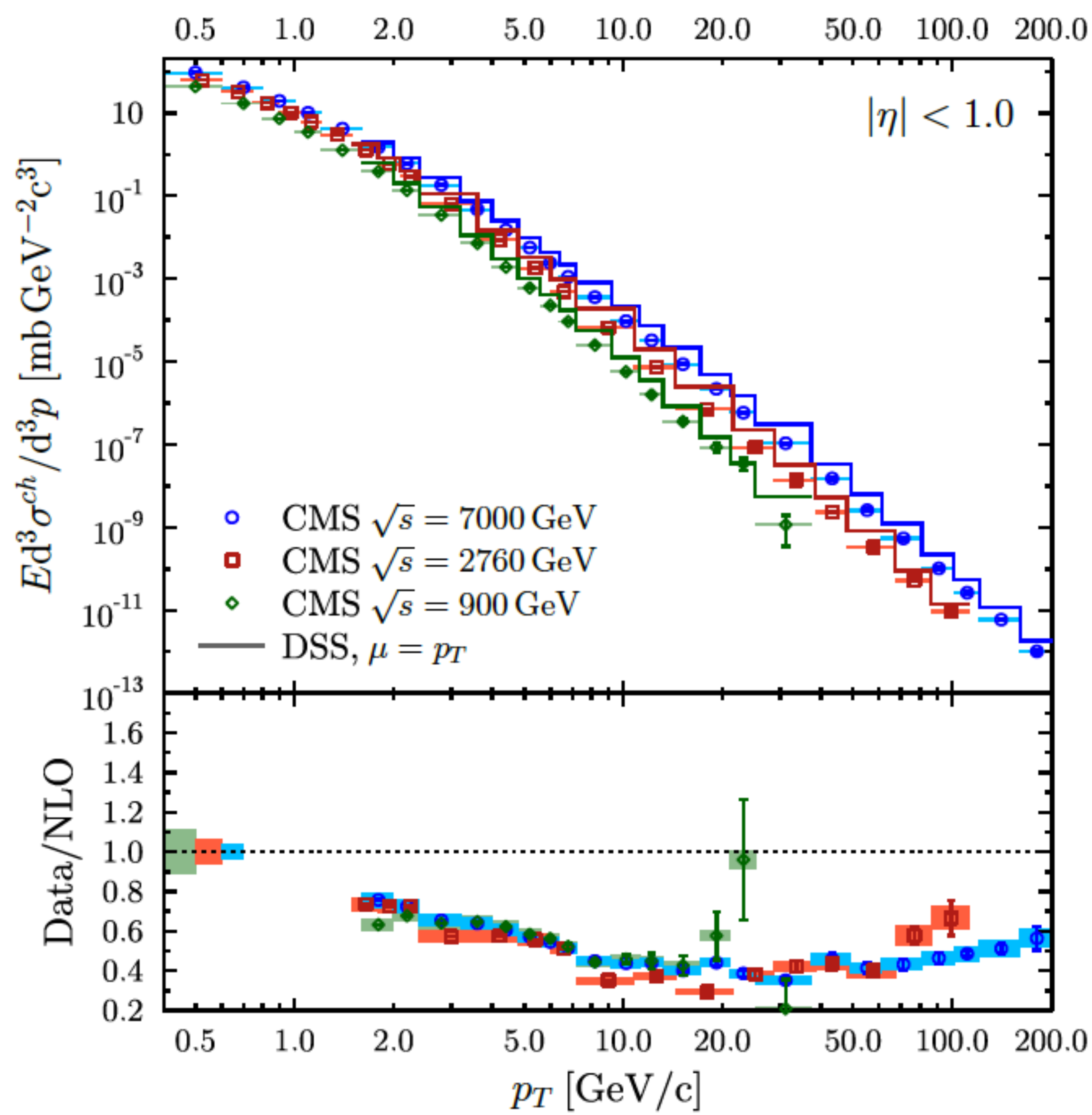
- ✅ handle on **flavour separation**,
- ✅ **precise data** available (HERMES/COMPASS).
- ❌ Involves both **FFs** and **PDFs**,
- ❌ Fully known so far up to $O(\alpha_s)$, *i.e.* NLO.

Experimental data

Hadroproduction in proton-proton collisions (pp)



$$\frac{d\sigma^h}{dp_T d\eta} = \sum_{ijk} f_i^{(1)} \otimes f_j^{(2)} \otimes \frac{d\hat{\sigma}_{ijk}}{dp_T d\eta} \otimes D_k^h$$



- ✅ Direct sensitivity to the **gluon FF**,
- ✅ **Scale scan** ($\mu_F \propto p_T$),
- ✅ Precise data from LHC/Tevatron
- ❌ Involves both **FFs** and **PDFs**,
- ❌ Known so far up to NLO,
  - ❌ large scale variations at low p_T,
- ❌ cumbersome to compute.

The MAP FF set

The minimiser