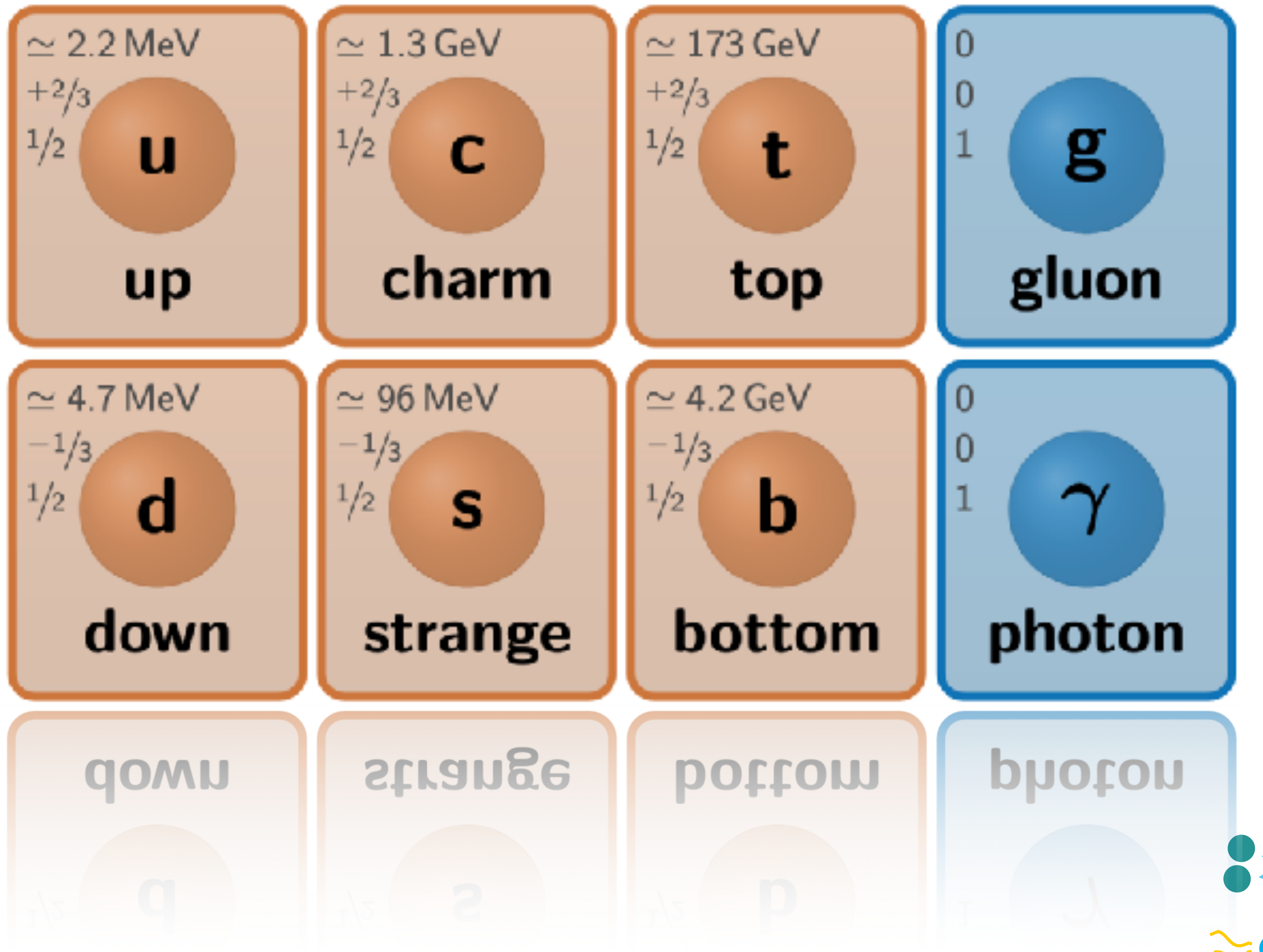
Challenges and developments on PDF determination

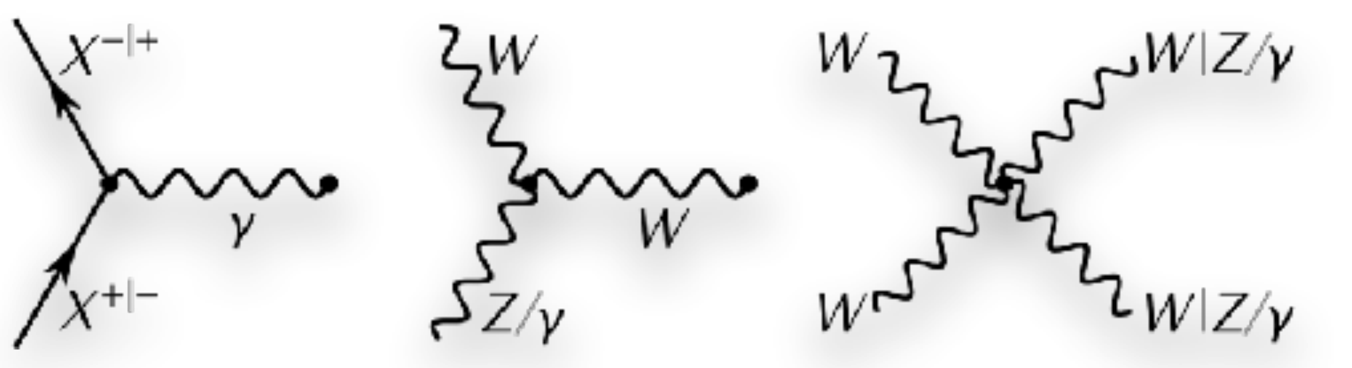
A journey from classical methods to quantum hardware

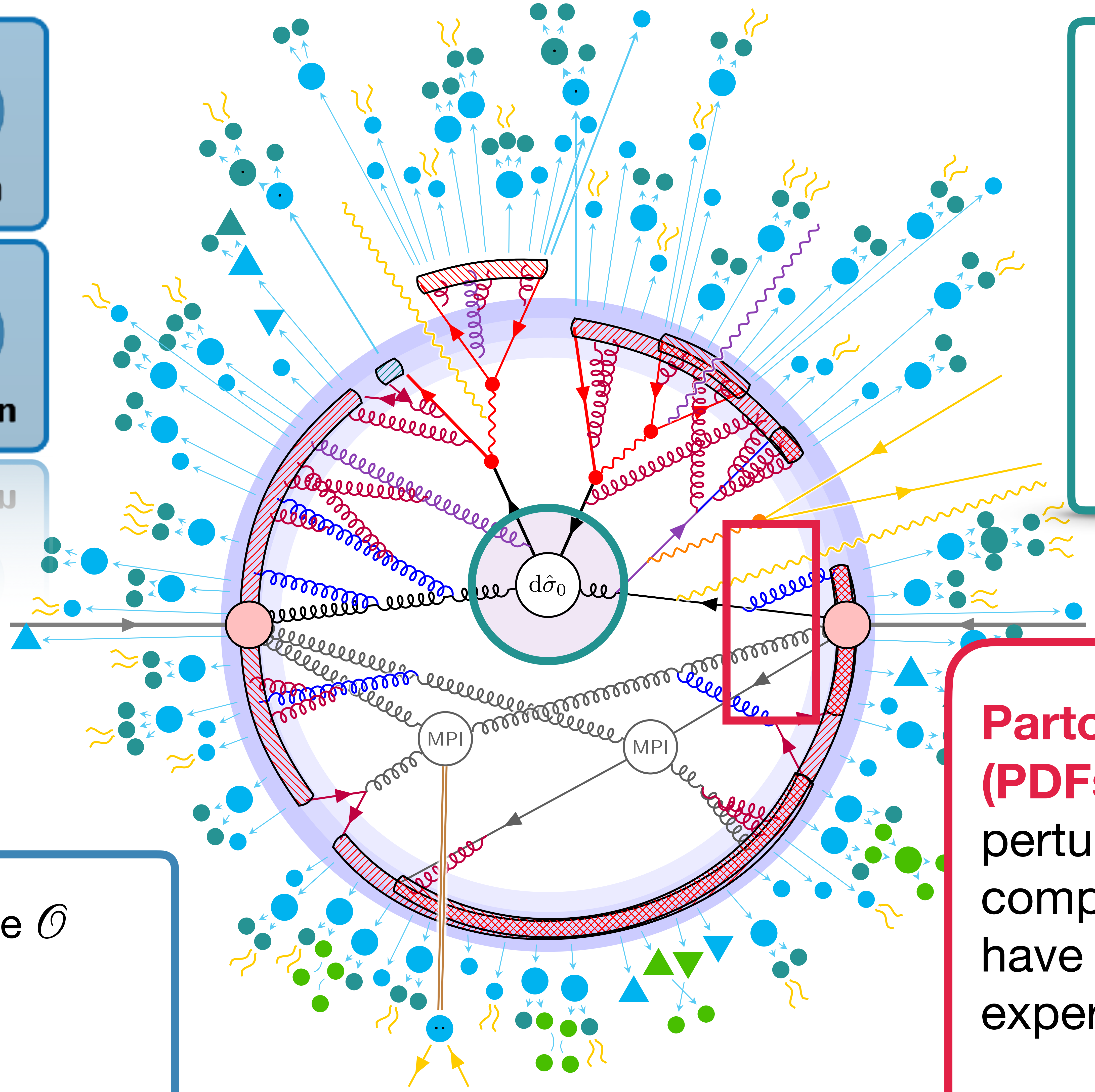
Juan M. Cruz Martinez - CERN TH Department
IFIC Valencia, September 2024

Ingredients for collider predictions in High Energy Physics



Hard scattering $\hat{\sigma}_0$: encodes short-range interactions; computed from first principles.





Parton Distribution Functions (PDFs): encodes long-range non-perturbative interactions; cannot be computed from first principle and have to be determined from experimental Data.

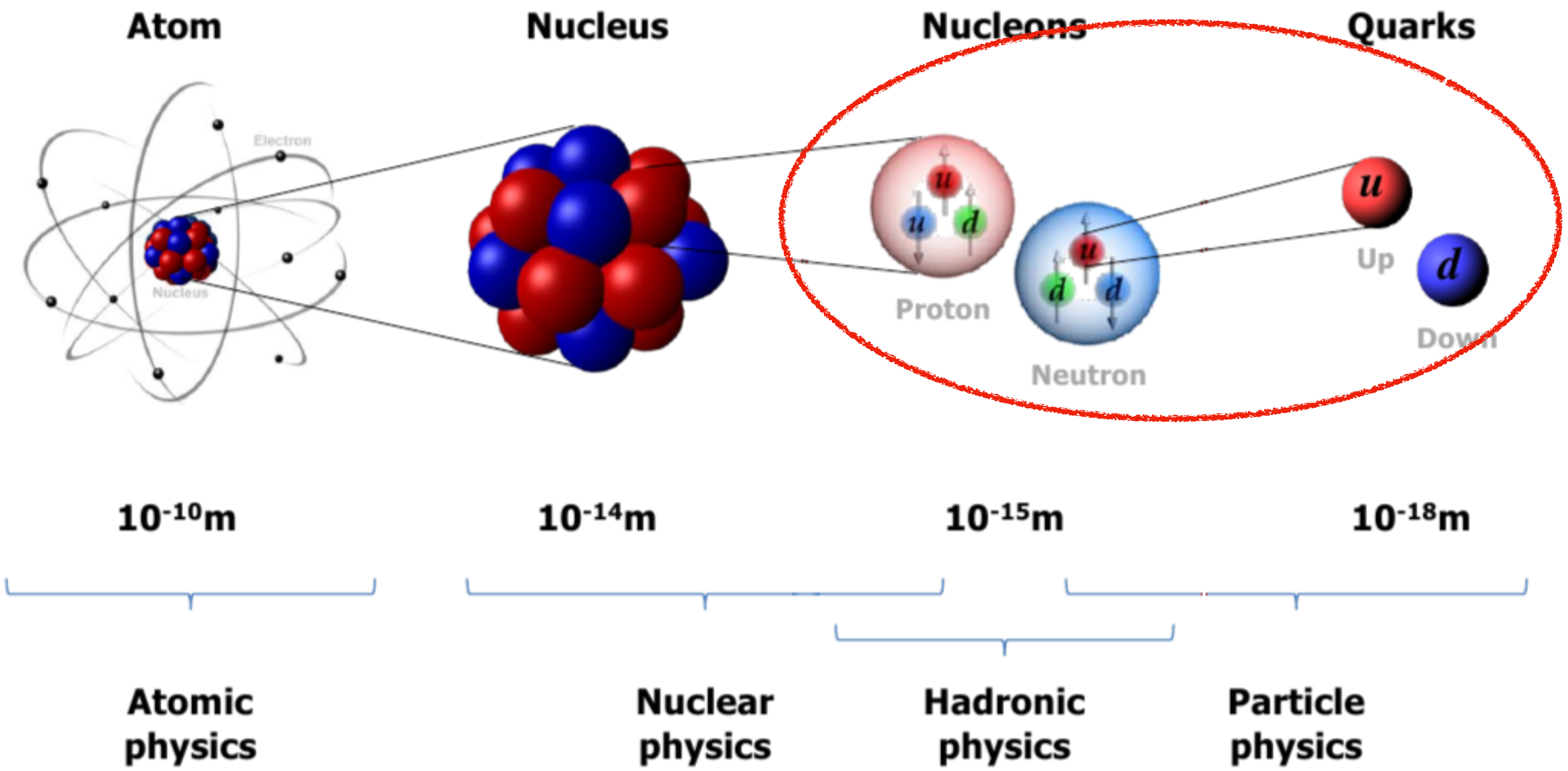
Important: PDFs are Universal

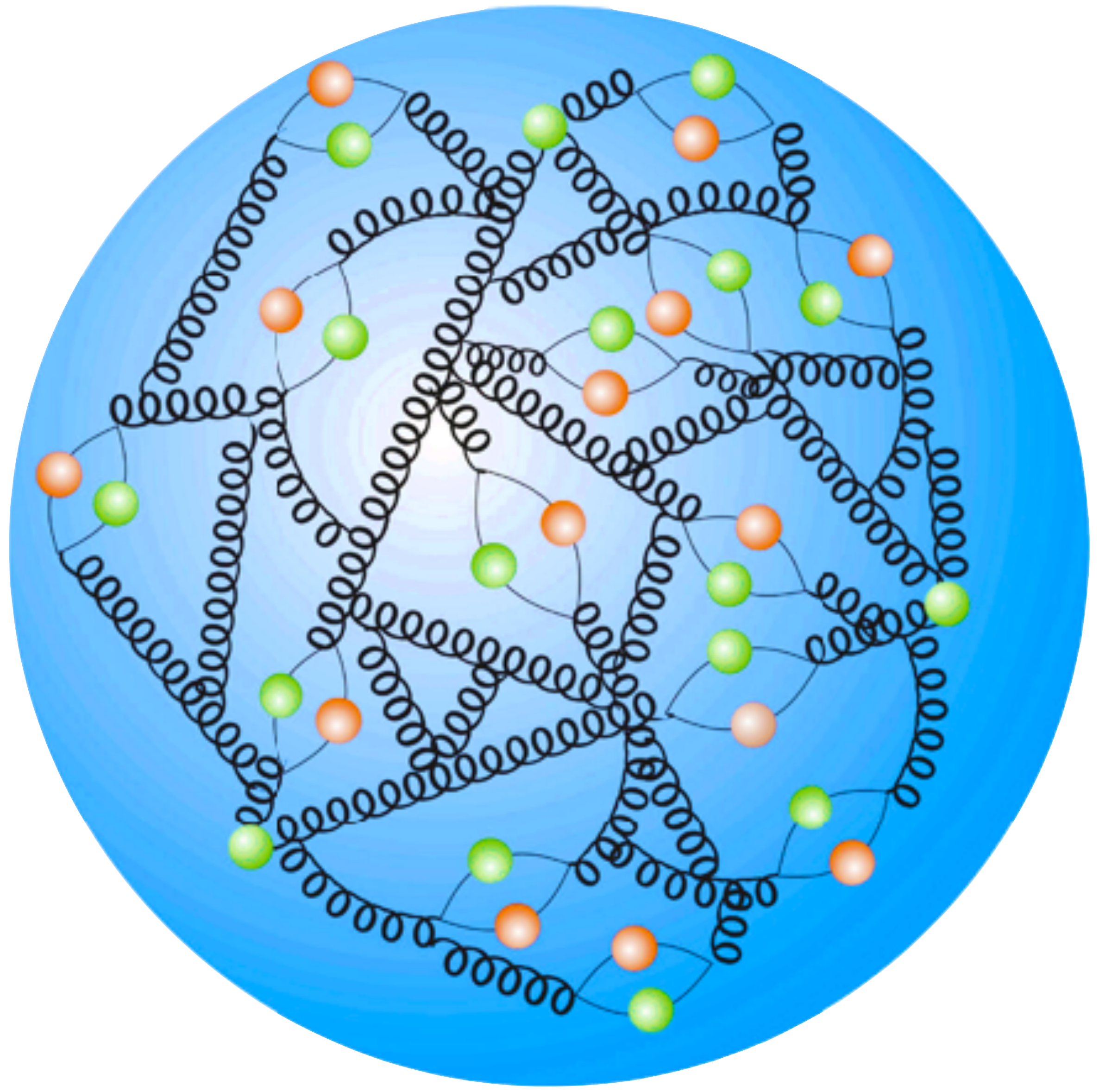
Experiments measure an Observable $\mathcal{O}$ (cross-section, decay rates, etc.):

$$\mathcal{O} = \sum \hat{\sigma}_{ij} \otimes \mathrm{PDF}_{ij}$$

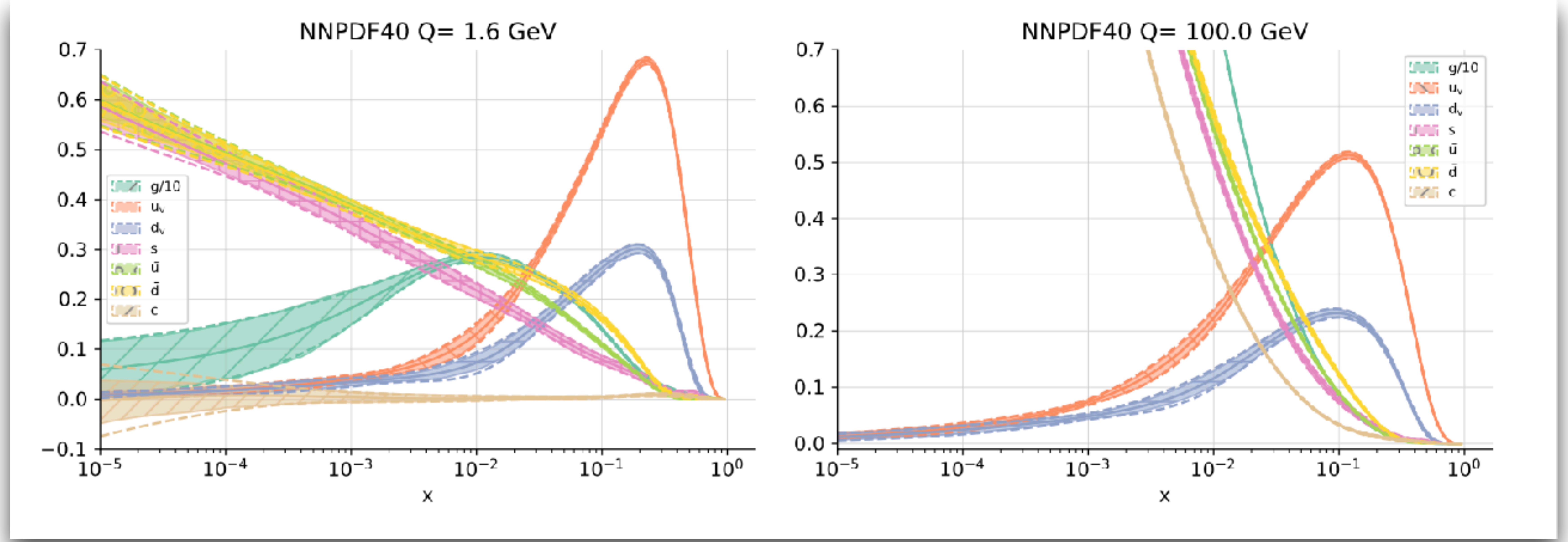
Parton Distribution Function or PDF

A window into the internal structure of the Proton





Parton Distribution Function or PDF



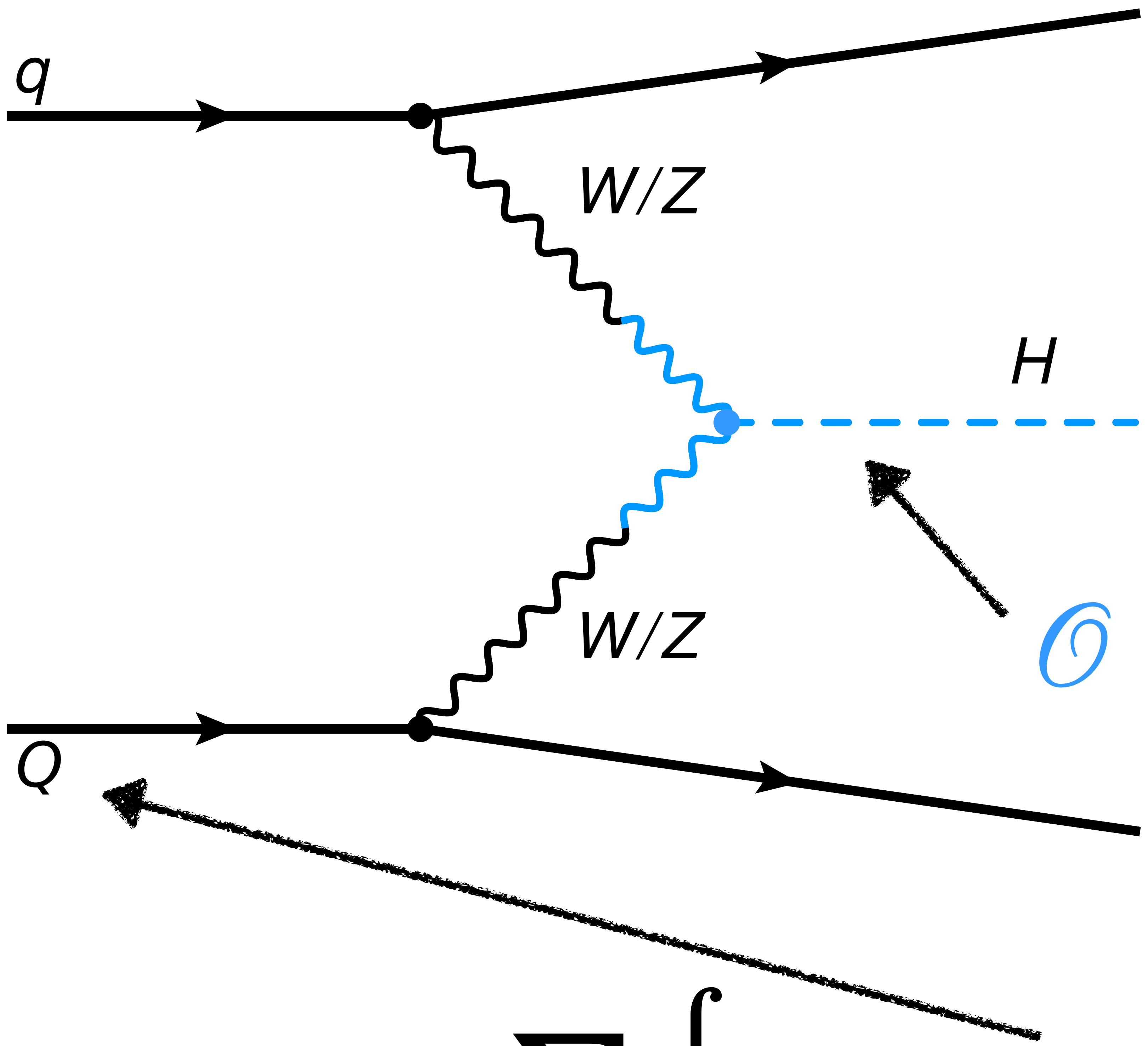
- PDFs are a fuctional probability distribution and depend on the **momentum faction** x **and on an energy scale** Q^2. Crucially, they cannot be directly observed.
- PDFs are non perturbative objects thus can not be computed using standard perturbative methods.
- The PDFs can be extracted from experimental data from colliders involving hadrons and some fixed-target experiments
- Many independent determination by different groups (CTEQ, MSHT, **NNPDF**, HERAPDFs)

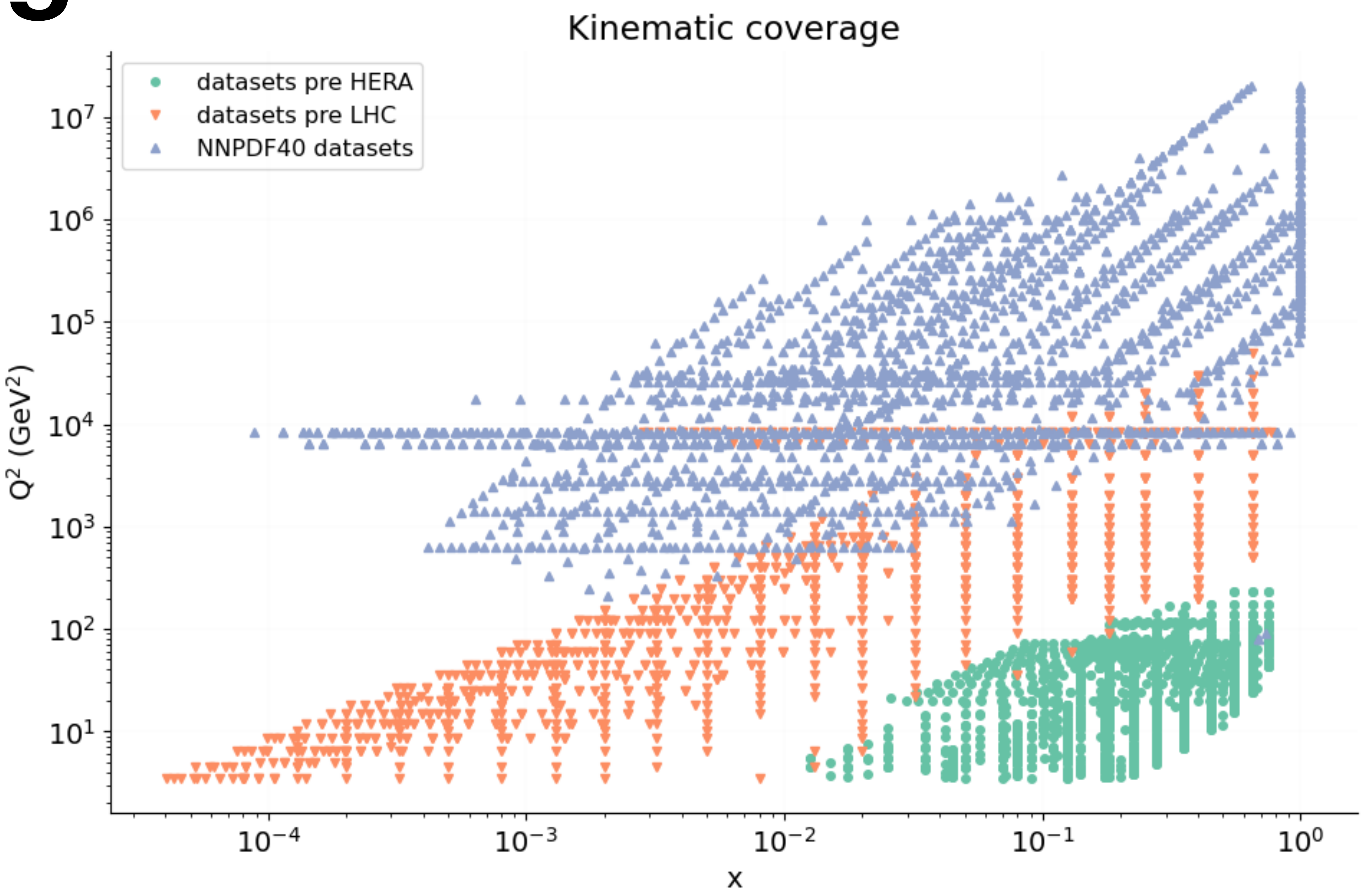
Despite being an empirically-determined object, several theoretical inputs are needed for a PDF fit:

- Hard-scattering cross sections.
- DGLAP evolution that define the dependence in Q^2.
- The accuracy at which these ingredients are computed determines the PDF accuracy.
- State of the art PDFs are determined at **aN3LO** in pQCD.

PDF determination ingredients

An extremely quick summary





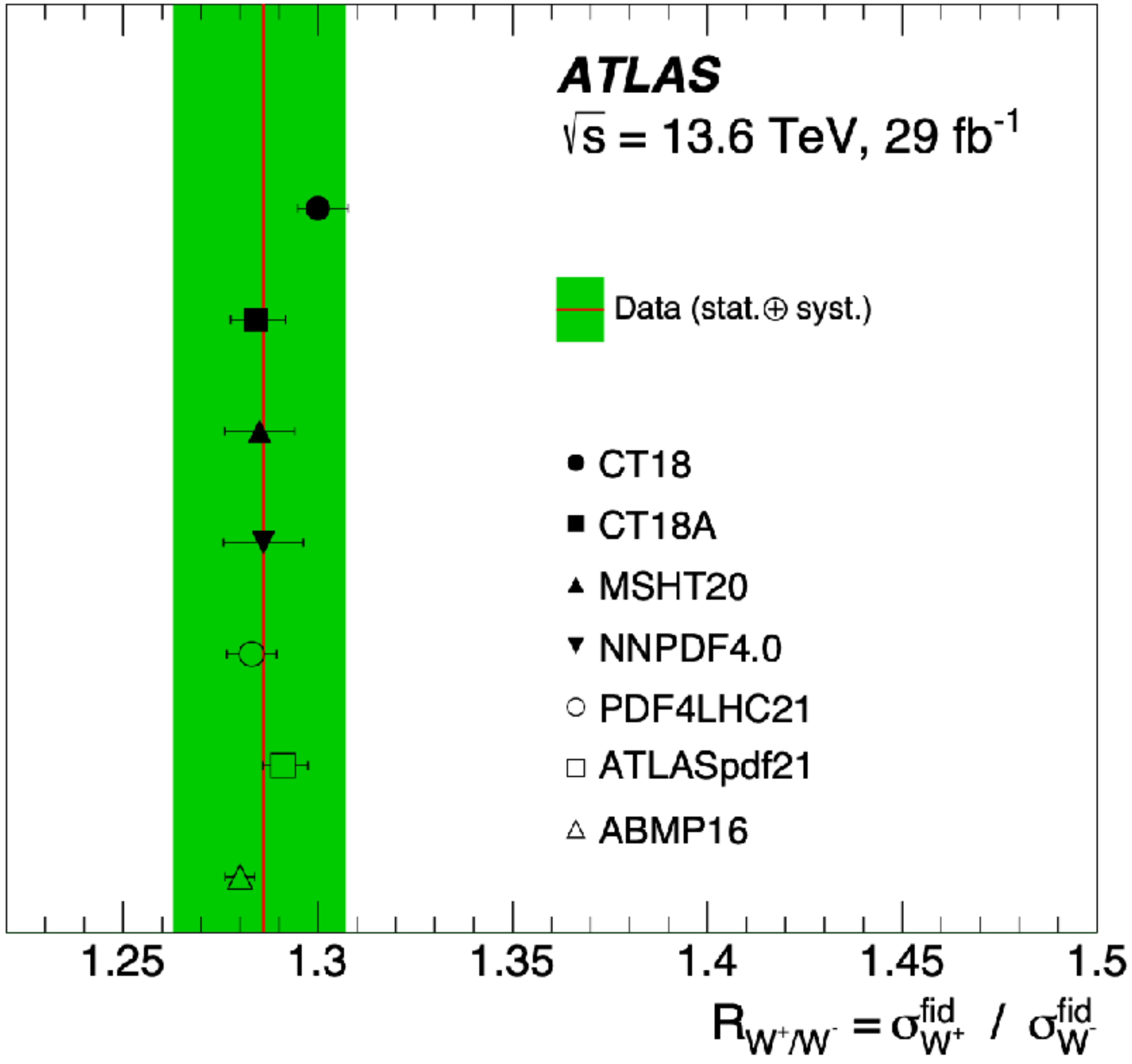
$$\mathcal{O} = \sum_{ij} \int dx_1 dx_2 \; f_i(x_1, \mu_F) \; f_j(x_2, \mu_F) \; \hat{\sigma}_{ij}(x_1, x_2, \mu_R, \mu_F)$$

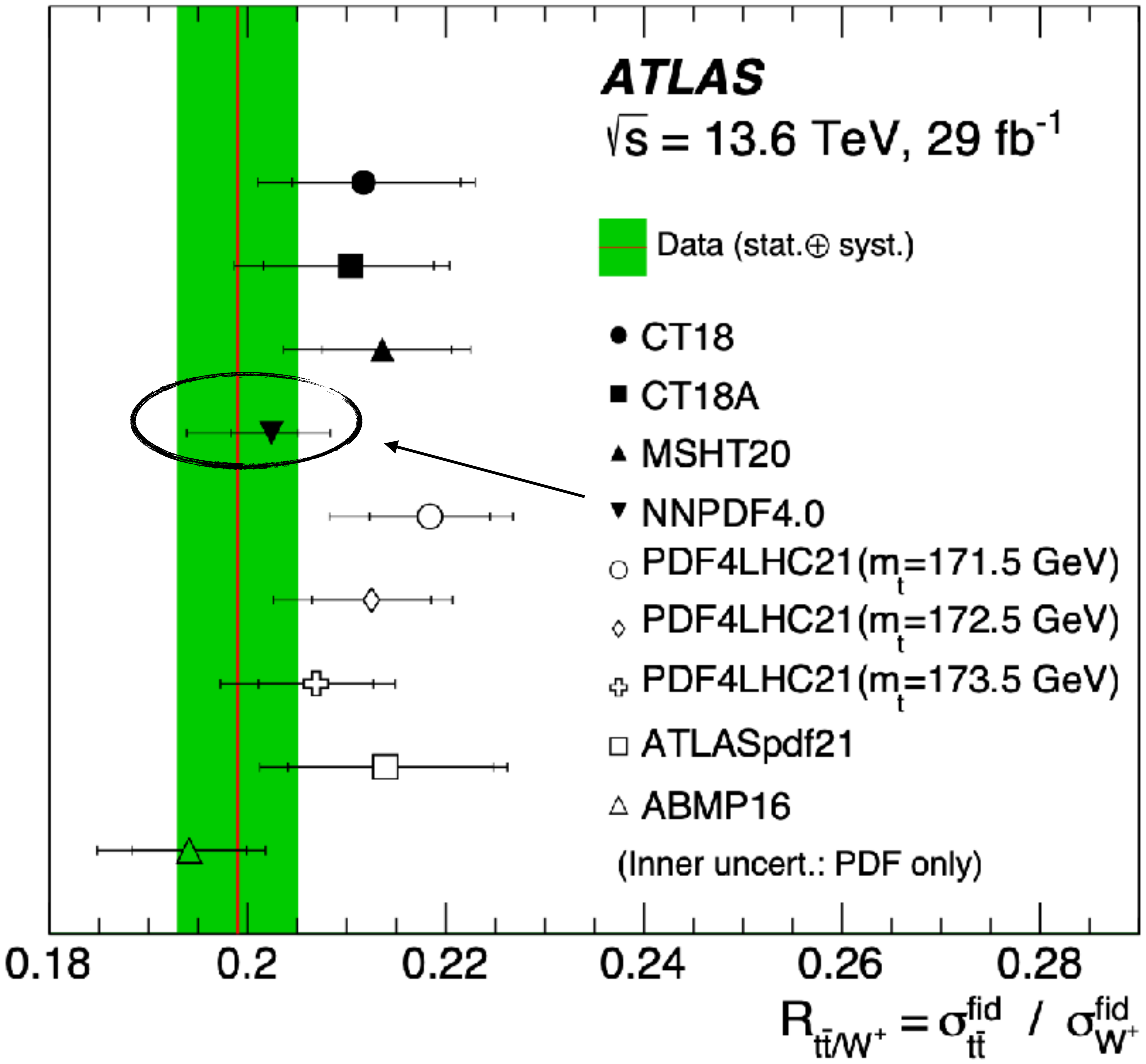
Input: quark type and fraction of the energy participating in the collision

What's the phenomenological impact of the choice of PDF?

Example: Measurement of vector boson production cross section and their ratios at $\sqrt{s} = 13.6$ with the ATLAS detector

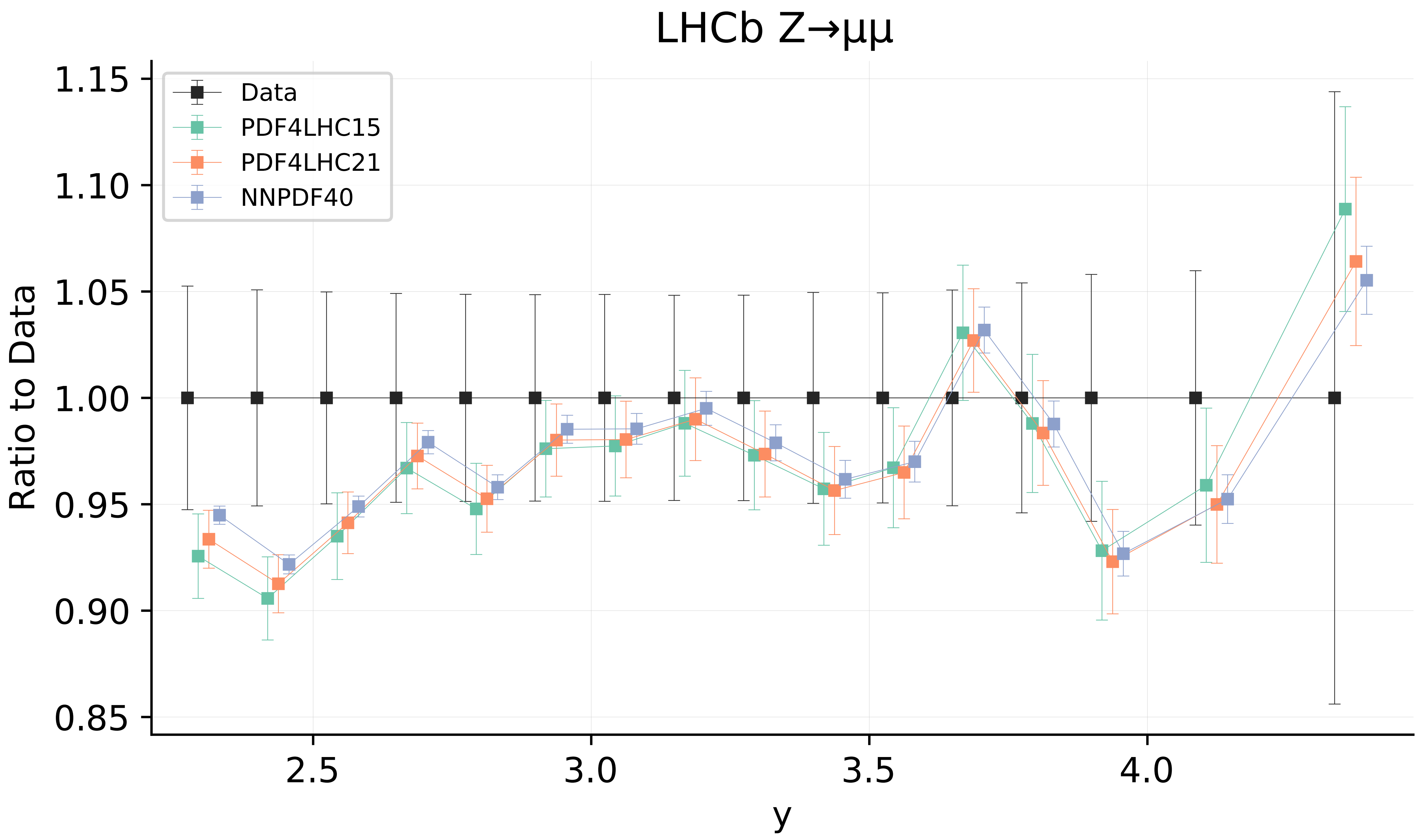
2403.12902

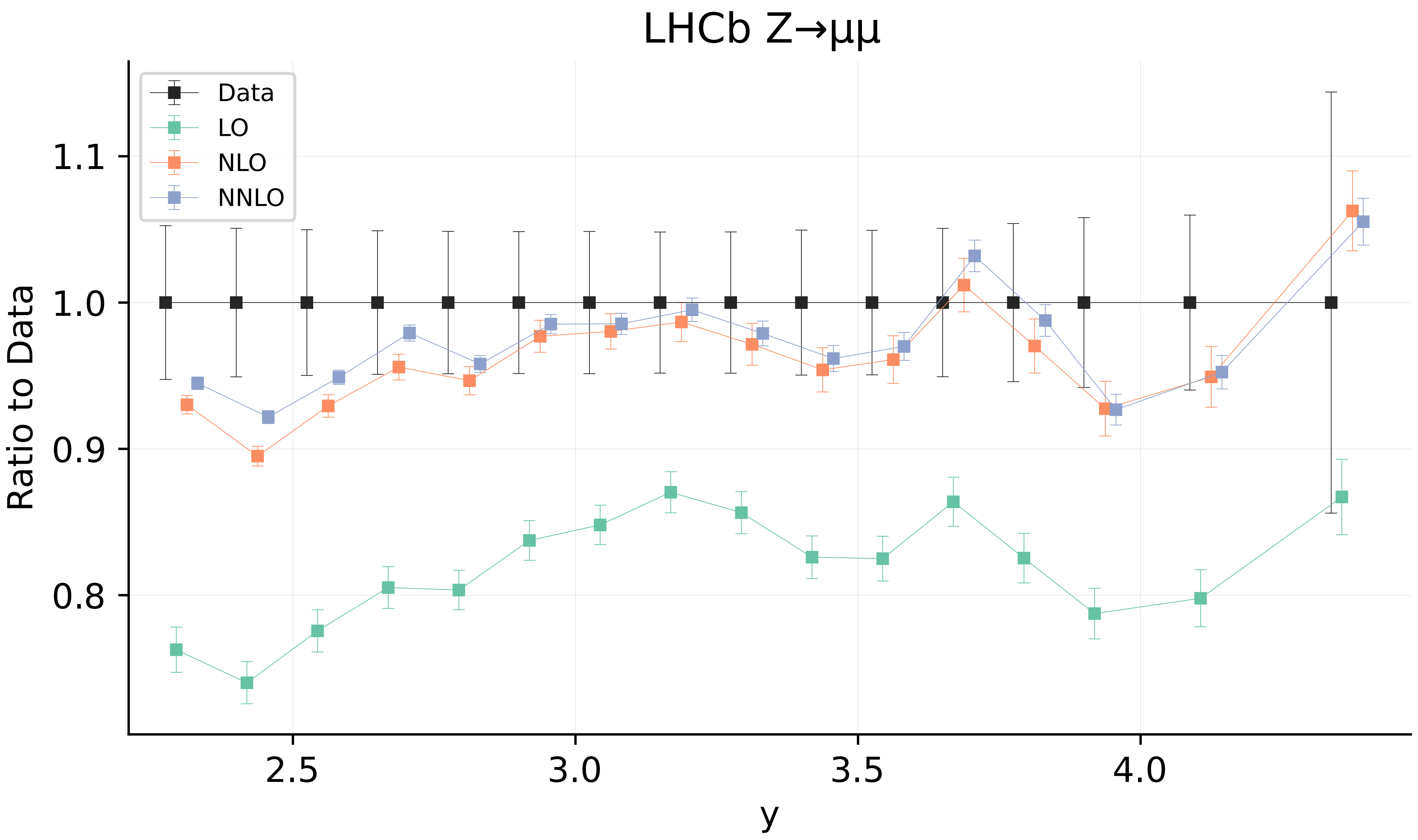




Accuracy and precision, over time and over orders

slowly but surely





PDF4LHC15: combination of NNPDF3.0, MMHT2014, CT14

PDF4LHC21: combination of NNPDF3.1, MSHT20, CT18

NNPDF4.0: updated over NNPDF3.1, with plenty of new data

Predictions for NNPDF4.0 at the corresponding order with Madgraph.

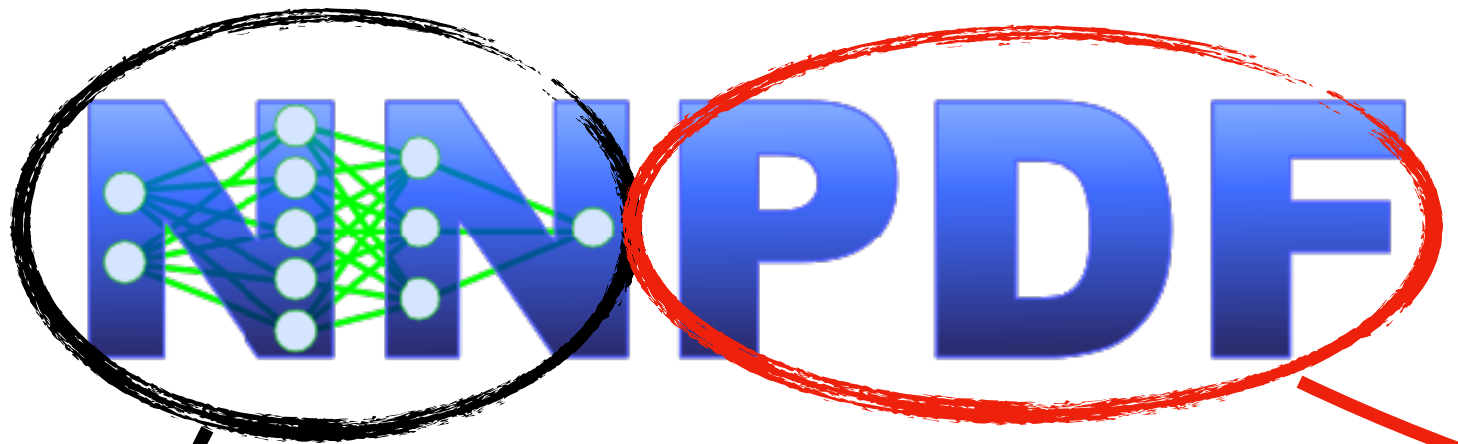
NNLO contribution computed as a k-factor with fewz.

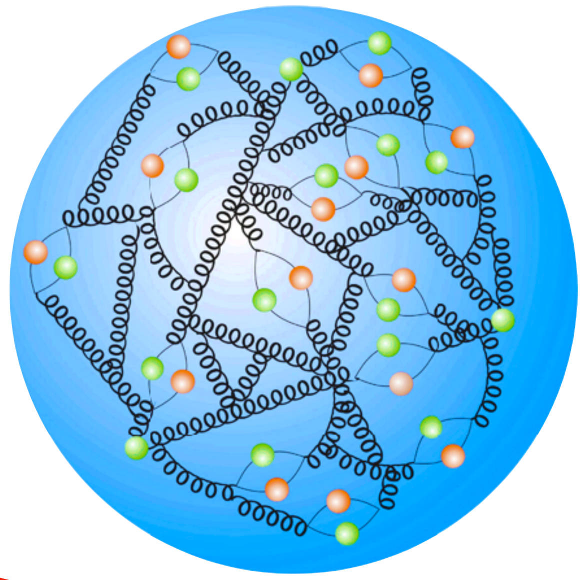
Accurate and trustworthy theory predictions are an essential ingredient of any PDF fit!

The NNPDF methodology

The PDF cannot be computed from first principles and cannot be directly observed, but we can instead ask ourselves:
can we predict the experimental data with this proton?

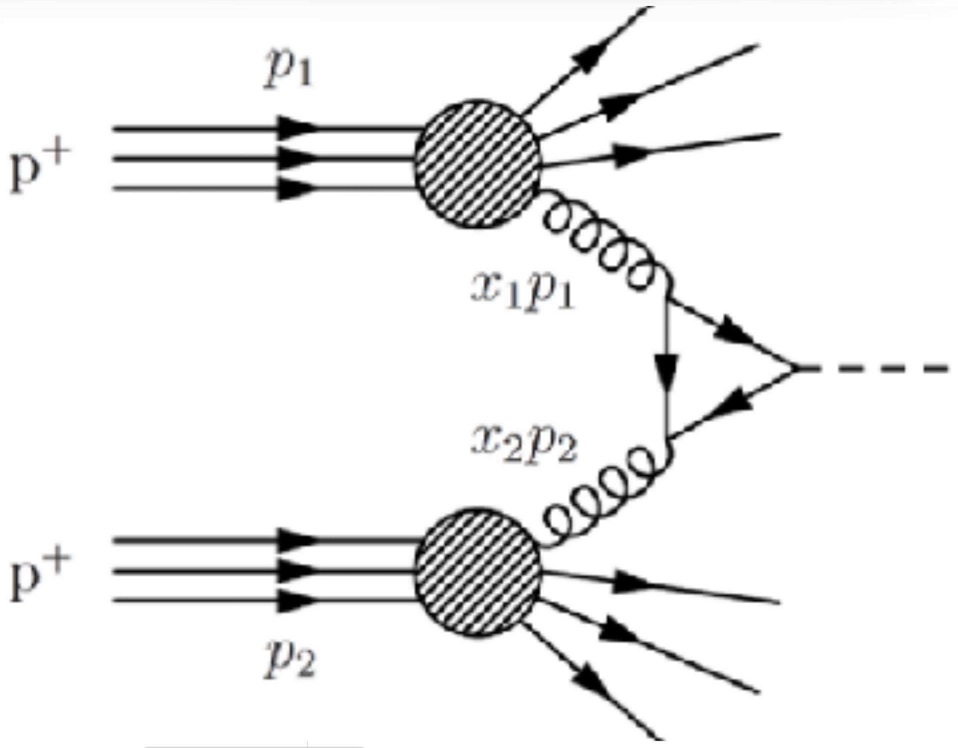
Historically, PDFs have been approximated with polynomial forms, introducing a theoretical (functional) bias. Can we instead use **Neural Networks**?

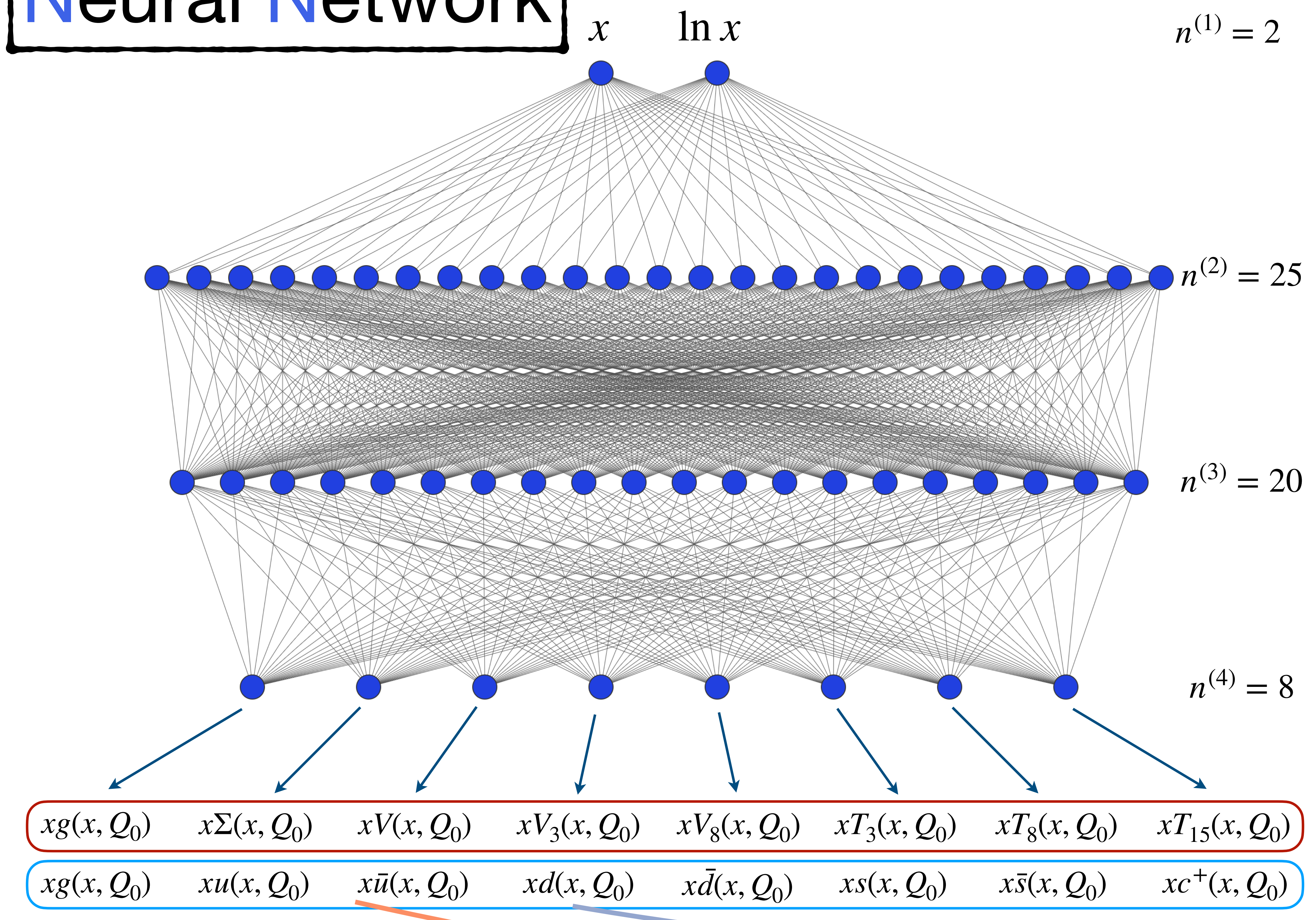


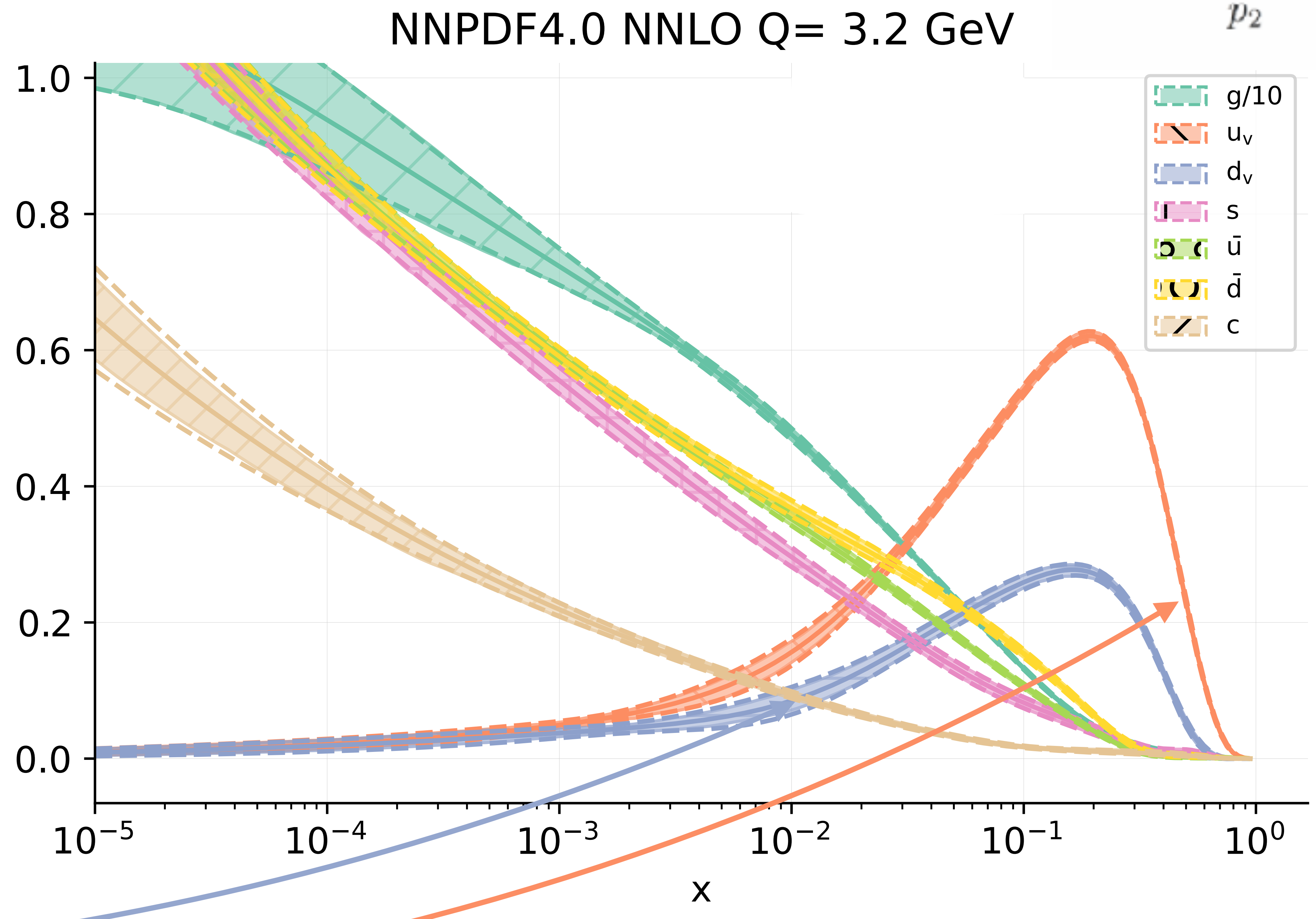


Neural Network

Parton
Distribution
Function

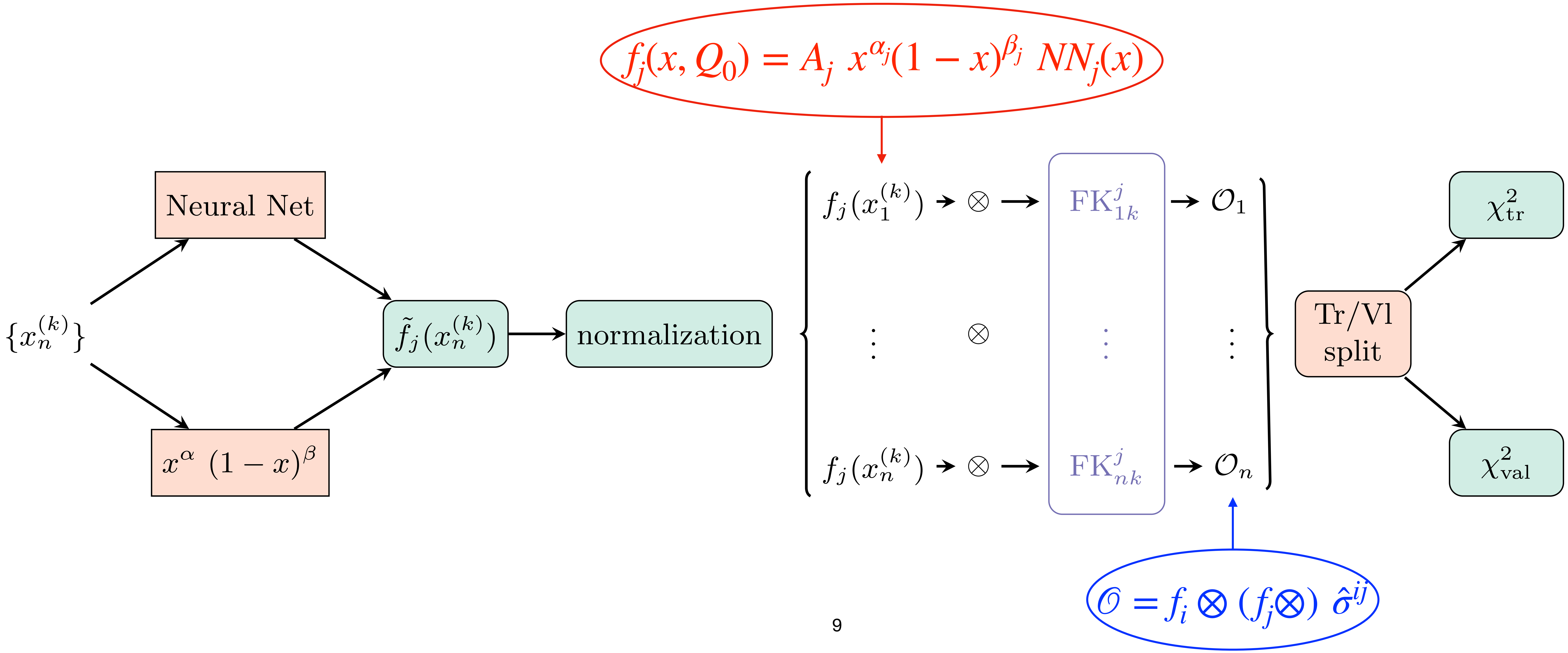






PDF fitting as a Machine learning problem

$$\mathcal{O} = \sum_{ij} \int dx_1 dx_2 \; f_i(x_1, \mu_F)\, f_j(x_2, \mu_F)\; \hat{\sigma}_{ij}(x_1, x_2, \mu_R, \mu_F)$$

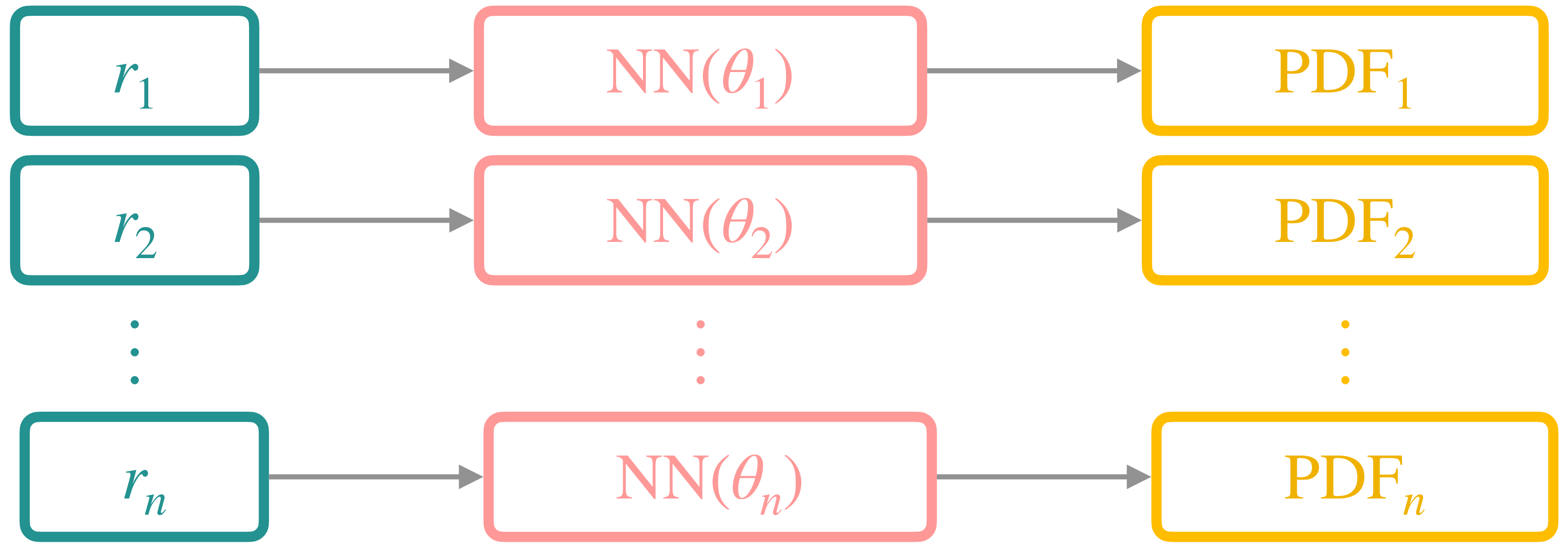


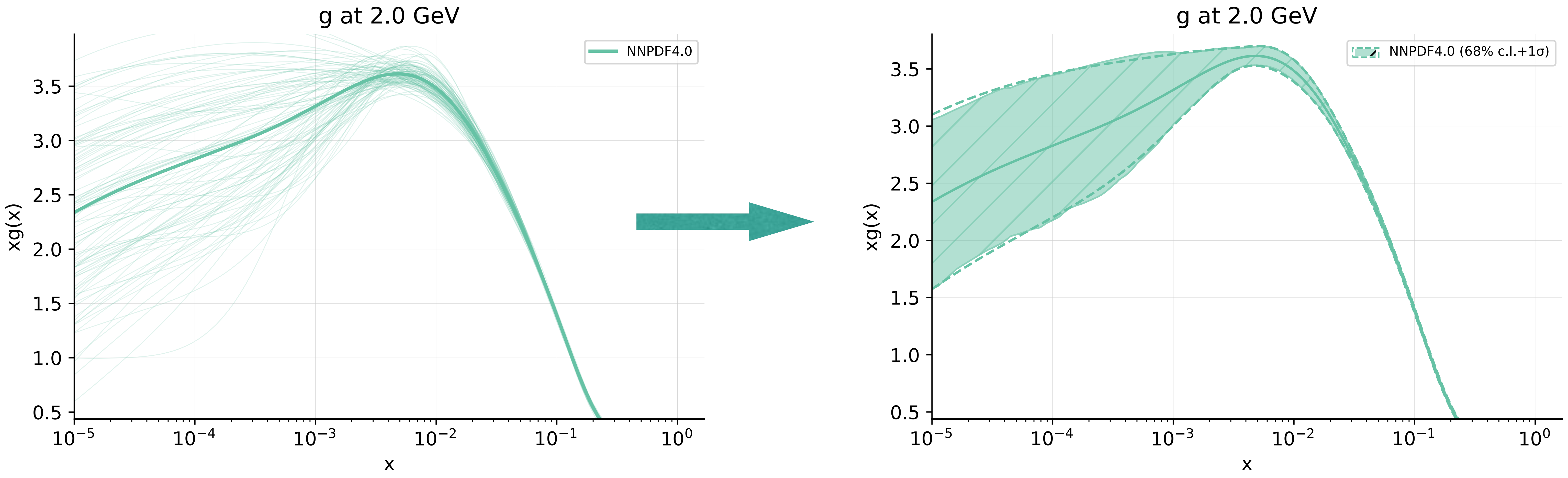
Uncertainties, from data to PDF

Experimental uncertainties are propagated into the fit by **fluctuating the central data:**

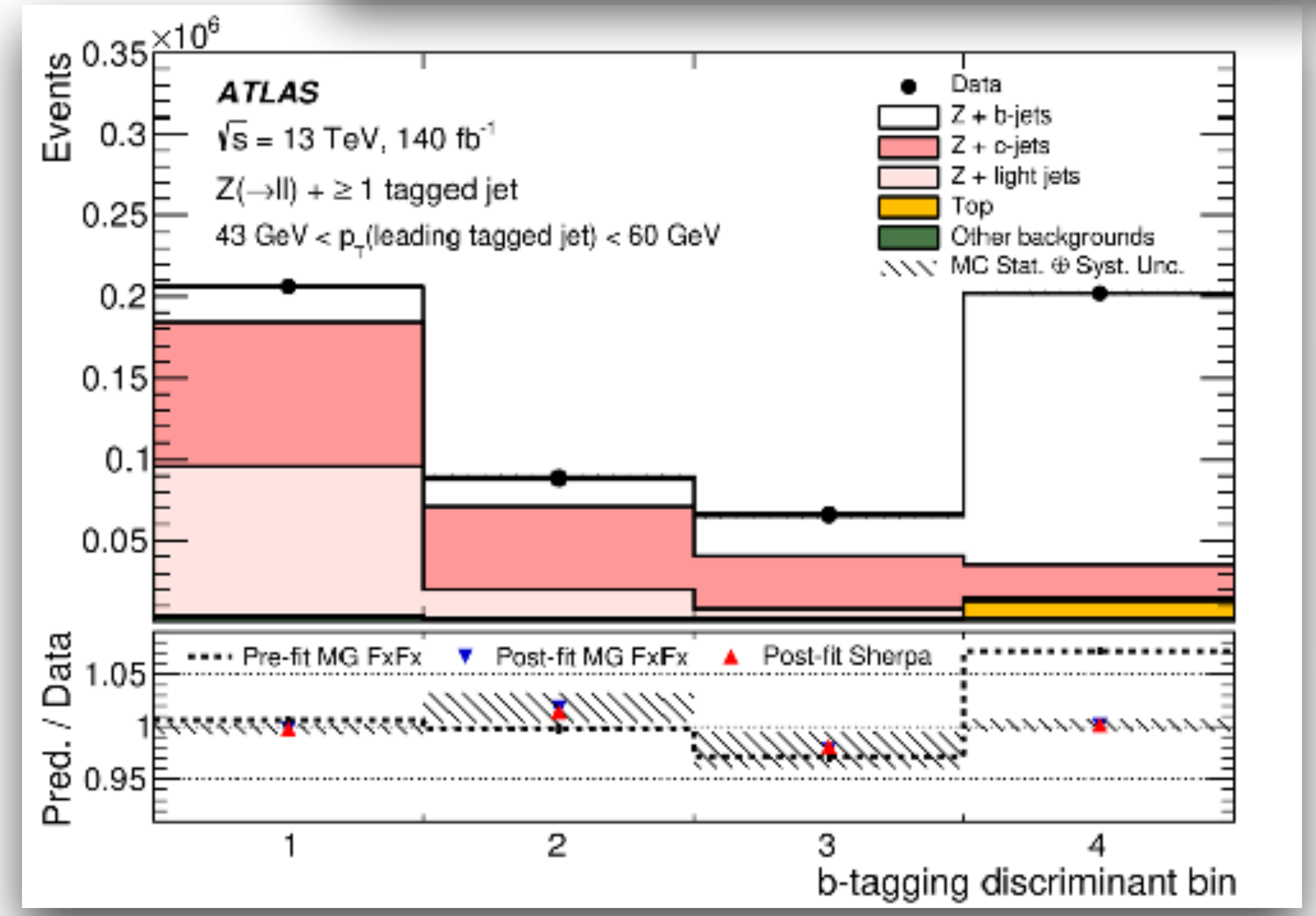
$$D_k = D_k^{(0)} + \sum_{\ell=1}^{n_D} \sqrt{\mathrm{Cov}_{k\ell}} \times \delta_\ell$$

Each of these replicas is then fitted to a separate NN. A PDF replica. The final output, which defines the PDF distribution, is the resulting ensemble of replicas.

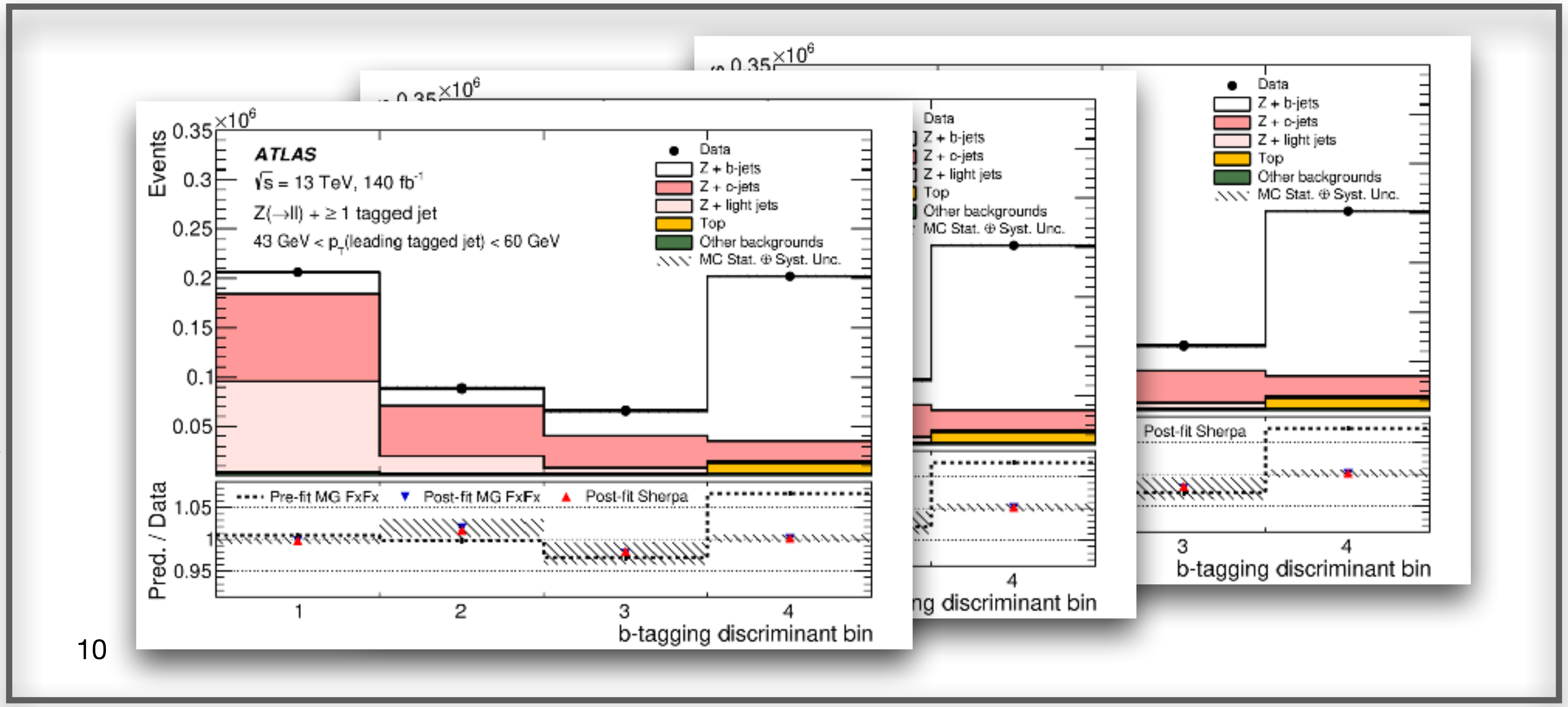




Monte Carlo Representation

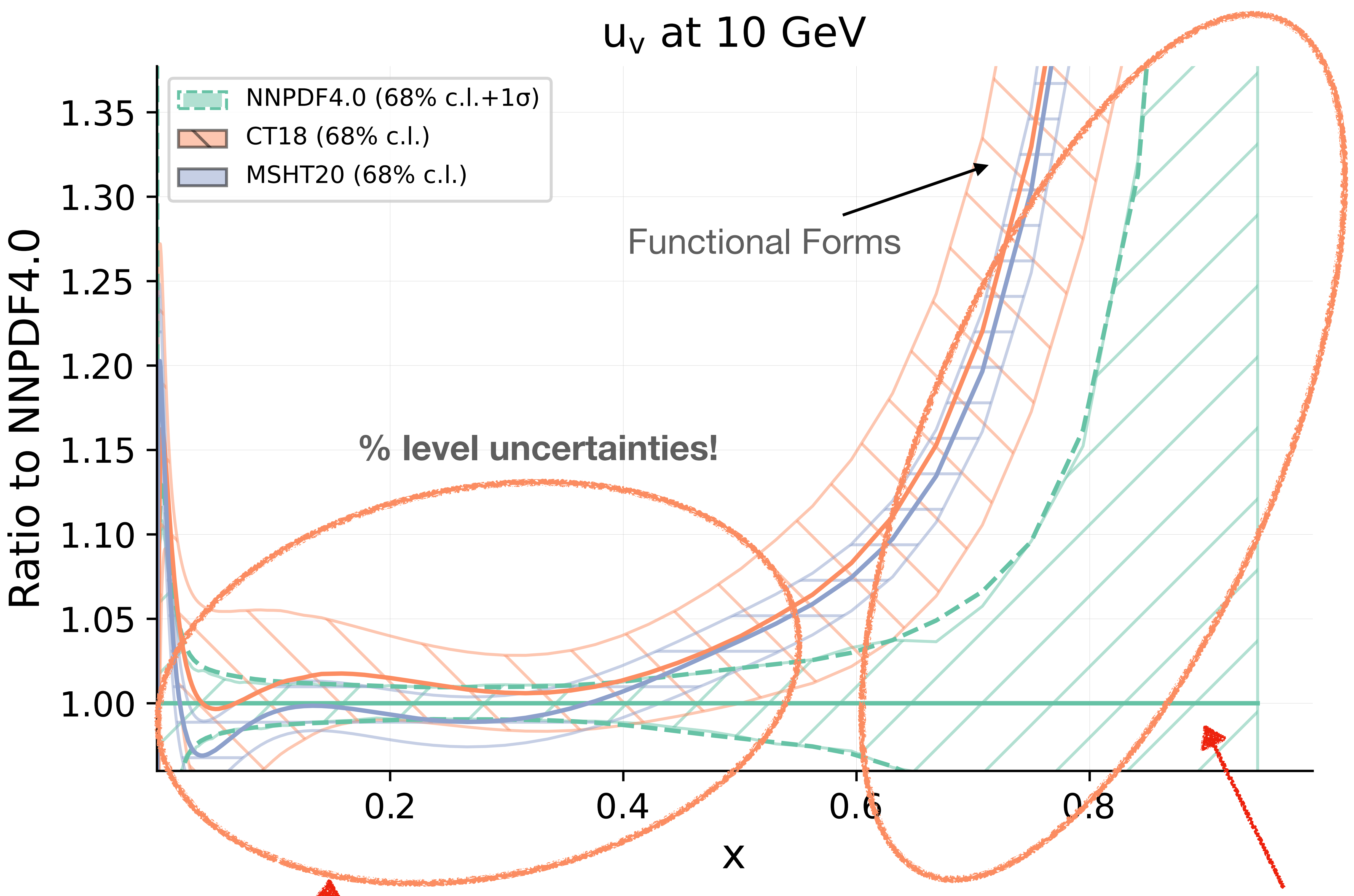


Generate Replicas of the Datasets



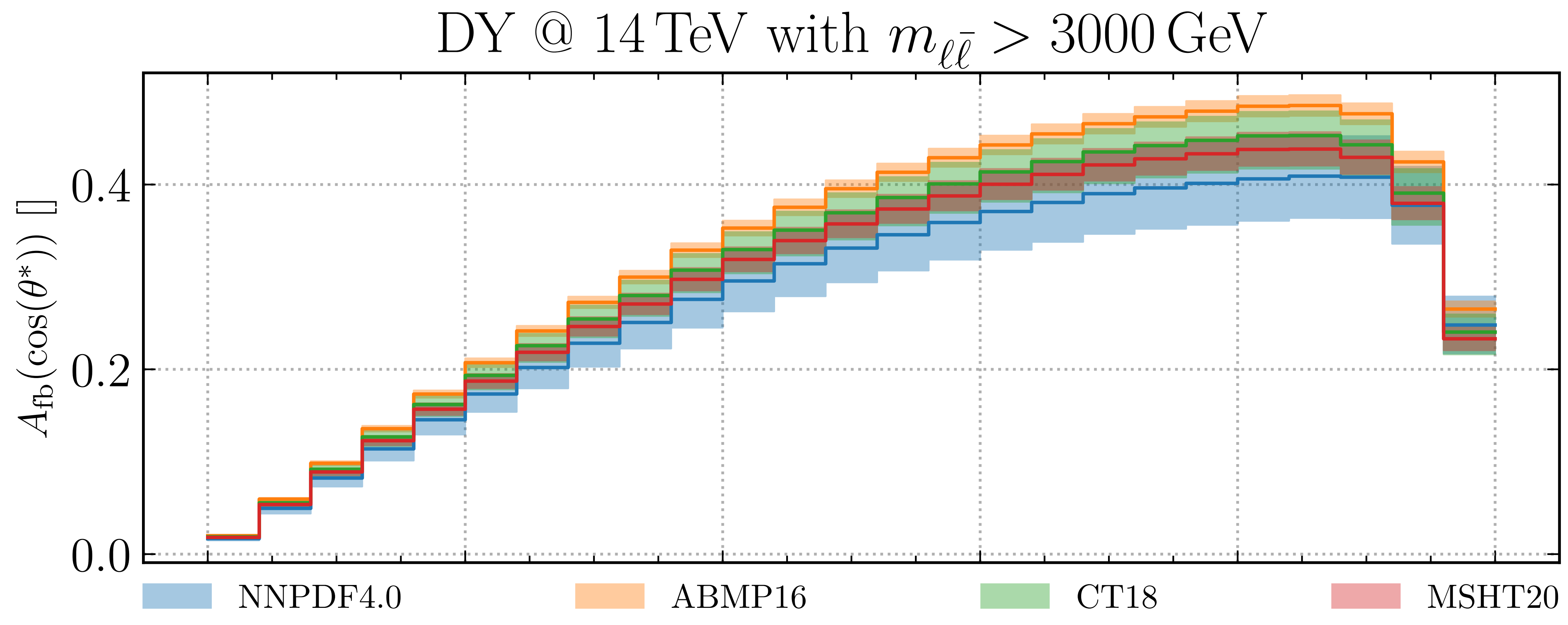
The precision follows the data

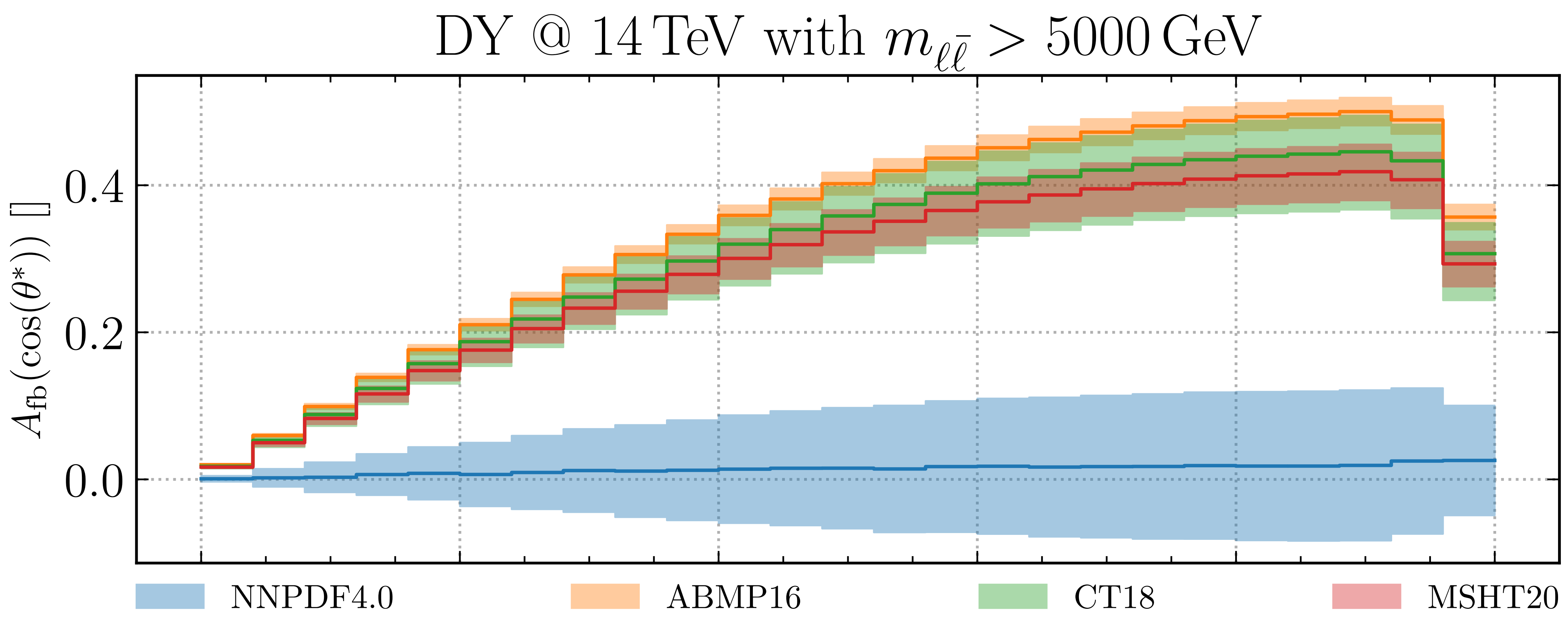
Not all regions are equally well determined, for PDFs the “data region” ends at around x~0.5



Data region: reasonable agreement aiming for both accuracy & precision

Extrapolation region hic sunt dracones!





In hep-ph/2209.08115 it was demonstrated how a too restrictive parametrization can lead to the extrapolation behaviours not justified by the available data!

The loss function

The PDF parameters are optimized to minimize the χ^2 that compares the experimental data D_i and the theoretical predictions $\mathcal{O}_i$.

$$\boxed{\chi^2 = \sum_N (\mathcal{O}_i - D_i)\; cov^{-1}_{ij}\; (\mathcal{O}_j - D_j)}$$

The number of datapoints in NNPDF4.0 is ~4500 separated in ~100 datasets. Each datasets is compared to a NXLO calculation. Each replica (~100) requires ~15000 iterations. If we want to estimate scale uncertainties we requite 7 or 9 variations for each of the theory calculations.

A single PDF fit might need just about 1500000000 integrals to complete.

We need a practical solution to this problem: Fast Kernel tables.

NB: the χ^2 in the loss function optimized in NNPDF fits is not the χ^2 to the experimental data, but a modified form to account for multiplicative uncertainties

$$cov_{ij} \longrightarrow cov_{ij} + t_{0i} t_{0j} s_i s_j$$

arXiv:0912.2276

Fast Kernel Tables (FKTables)

Since the PDF depends only on the values of x an Q: bin the cross section on the relevant variables.

Note that we also single out μ_R in order to perform scale variations or α_s determinations.

$$\frac{d^4\hat{\sigma}_{ij}}{d\mu_F d\mu_R dx_1 dx_2}$$

The evolution on the μ scales is exact (O(α^2)) so the grid needed during the fit can be further simplified:

$$\frac{d^2\hat{\sigma}_{ij}}{dx_1 dx_2}$$

x-grid

$$\mathcal{O} = \sum_{ij} \int dx_1 dx_2 \; f_i(x_1, \mu_F)\, f_j(x_2, \mu_F)\; \hat{\sigma}_{ij}(x_1, x_2, \mu_R, \mu_F) = f_i^{\alpha} f_j^{\beta} \hat{\sigma}^{ij}_{\alpha\beta}$$

flavours

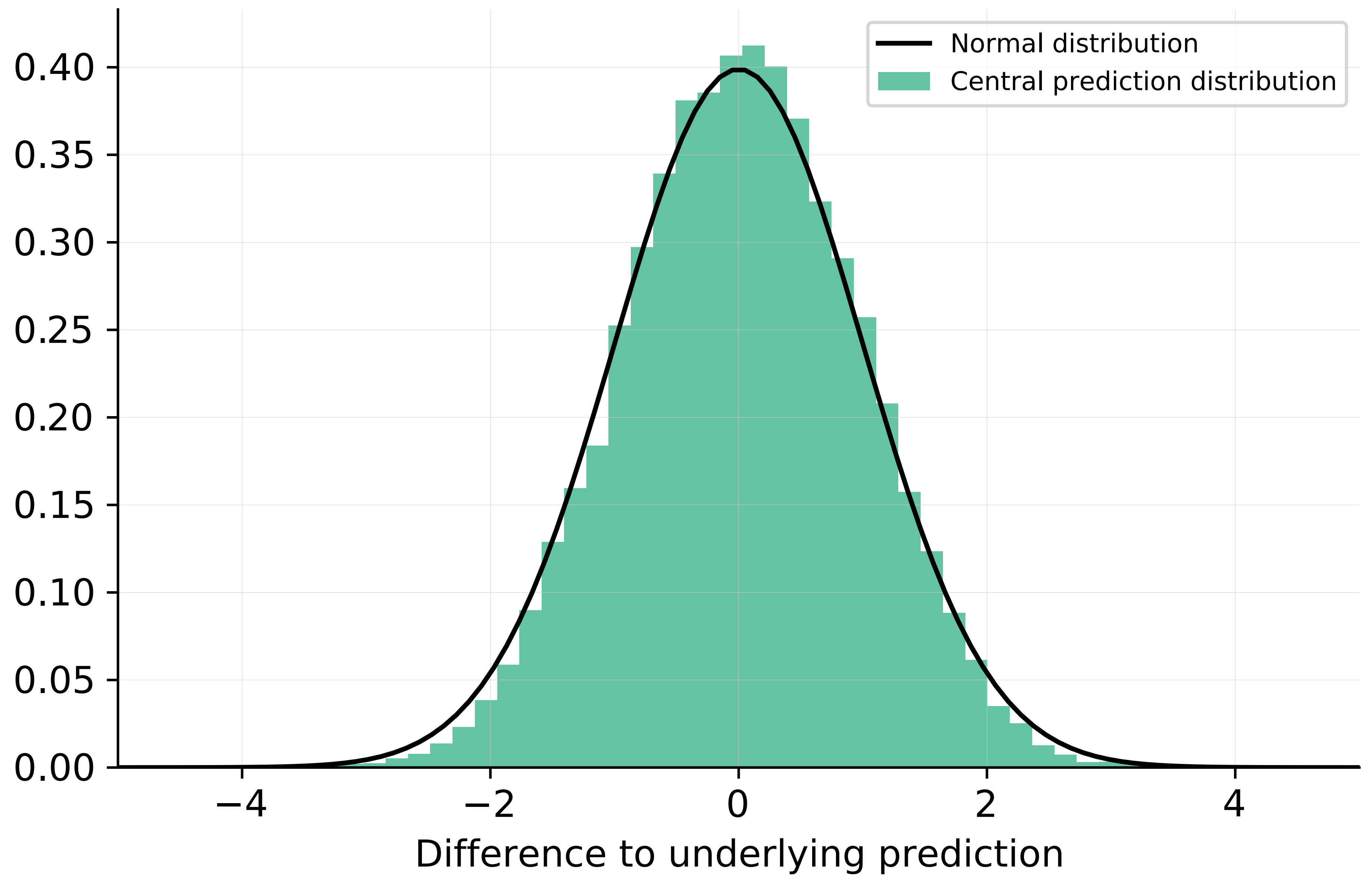
With O(50) points we can get a good representation of most observables, i.e., for each step of the fitting process we *just* need to contract the PDF with a tensor of *only* $4500 \times 50 \times 50 \times 14 \times 14 \simeq 10^9$ elements. Easy!

Pineline: Industrialization of high-energy theory predictions
A. Barontini, A. Candido, **JCM,** F. Hekhorn, C. Schwan - [hep-ph] 2302.12124

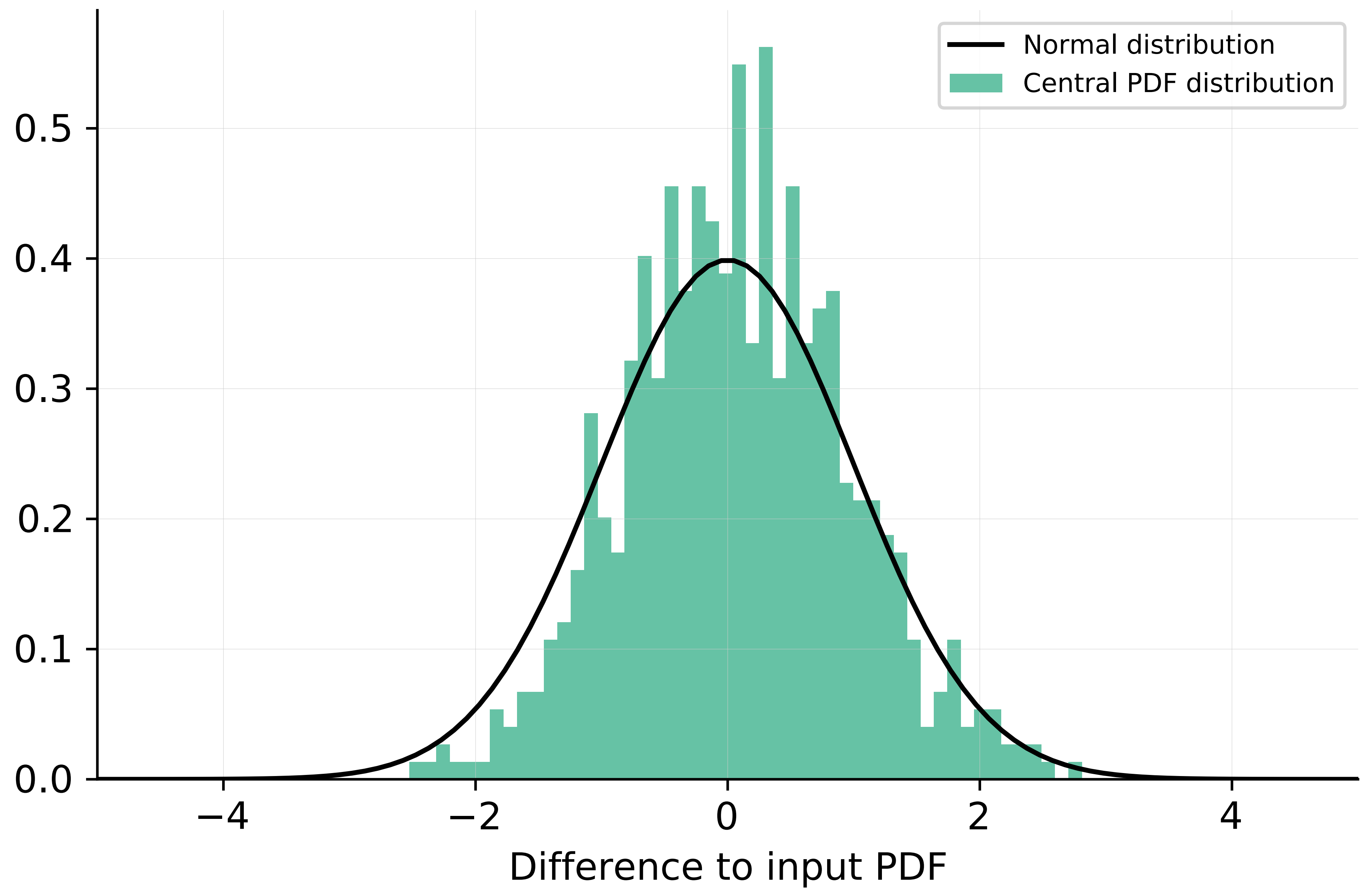
Validation and testing

Closure tests

1. Select some other PDF as the truth (an NNPDF replica or a fit from another group)
2. Generate fake data according to the theoretical predictions used in the fit
3. Generate variations of the data using the experimental uncertainties



- ➡ Check whether the parametrization is flexible enough
- ➡ Check whether we can reproduce the “true” PDF if it were known
- ➡ Do all of that in an environment in which everything is consistent and no theoretical knowledge is missing (no MHOU needed)

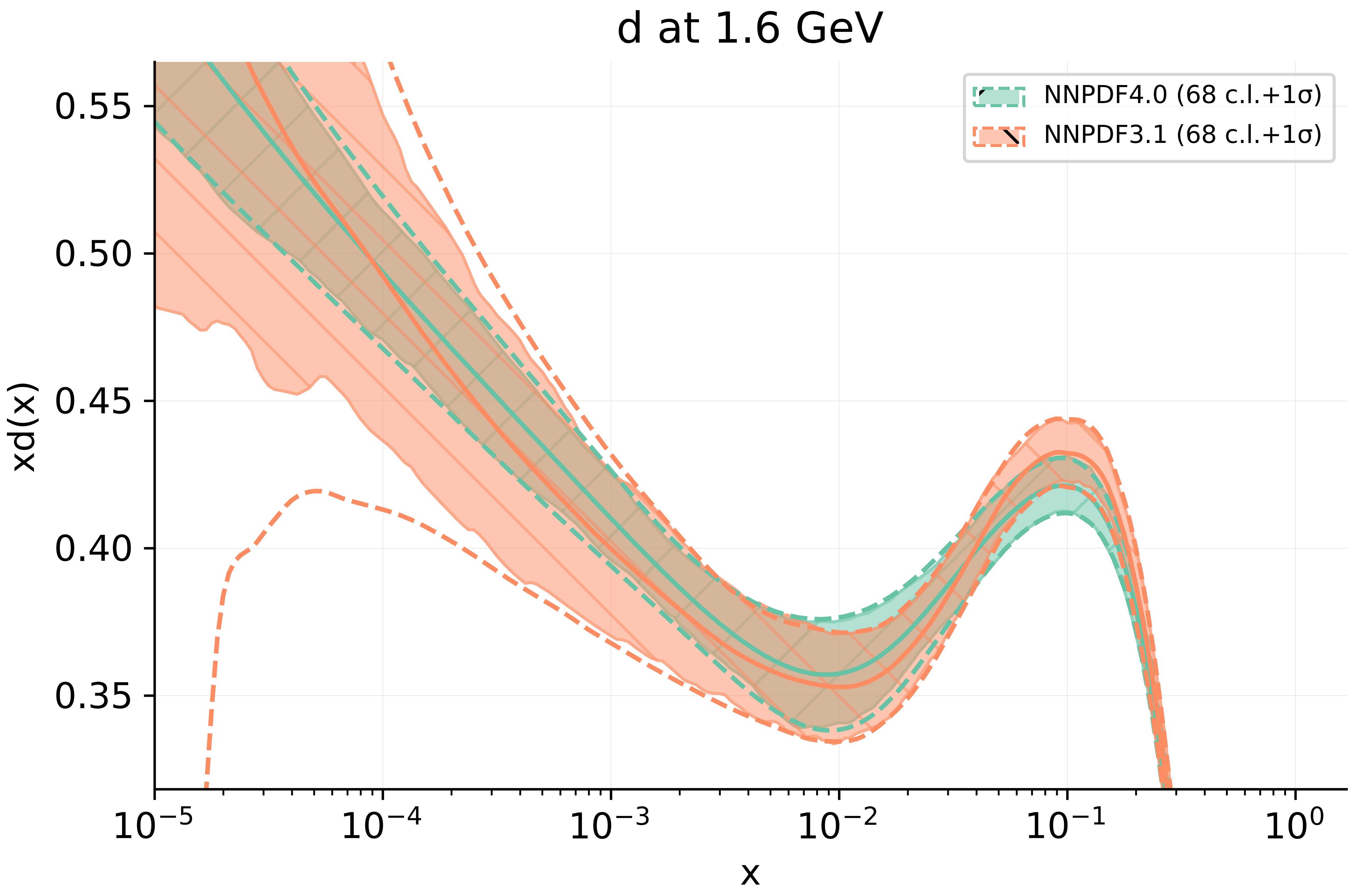


Validation and testing

Future tests

From the NNPDF3.1 family of fits to NNPDF4.0, the uncertainty bands of the PDF have shrinked considerably. How can we know whether this is only due to the new data?

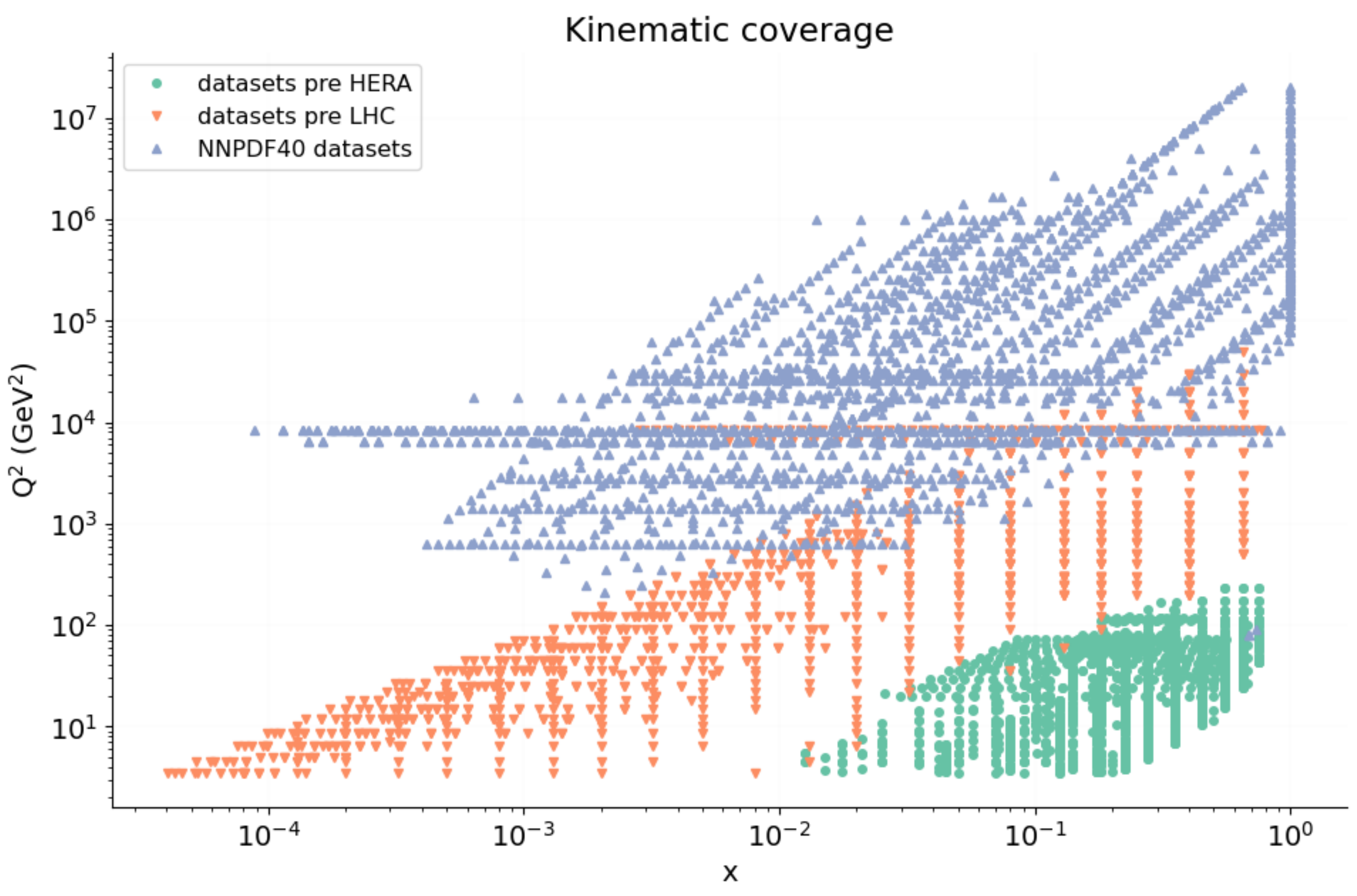
Will the NNPDF4.0 fit be able to “predict” (or rather, accommodate) new data from future experiments?



Future tests of parton distributions
JCM, S. Forte, E. Nocera - [hep-ph] 2103.08606

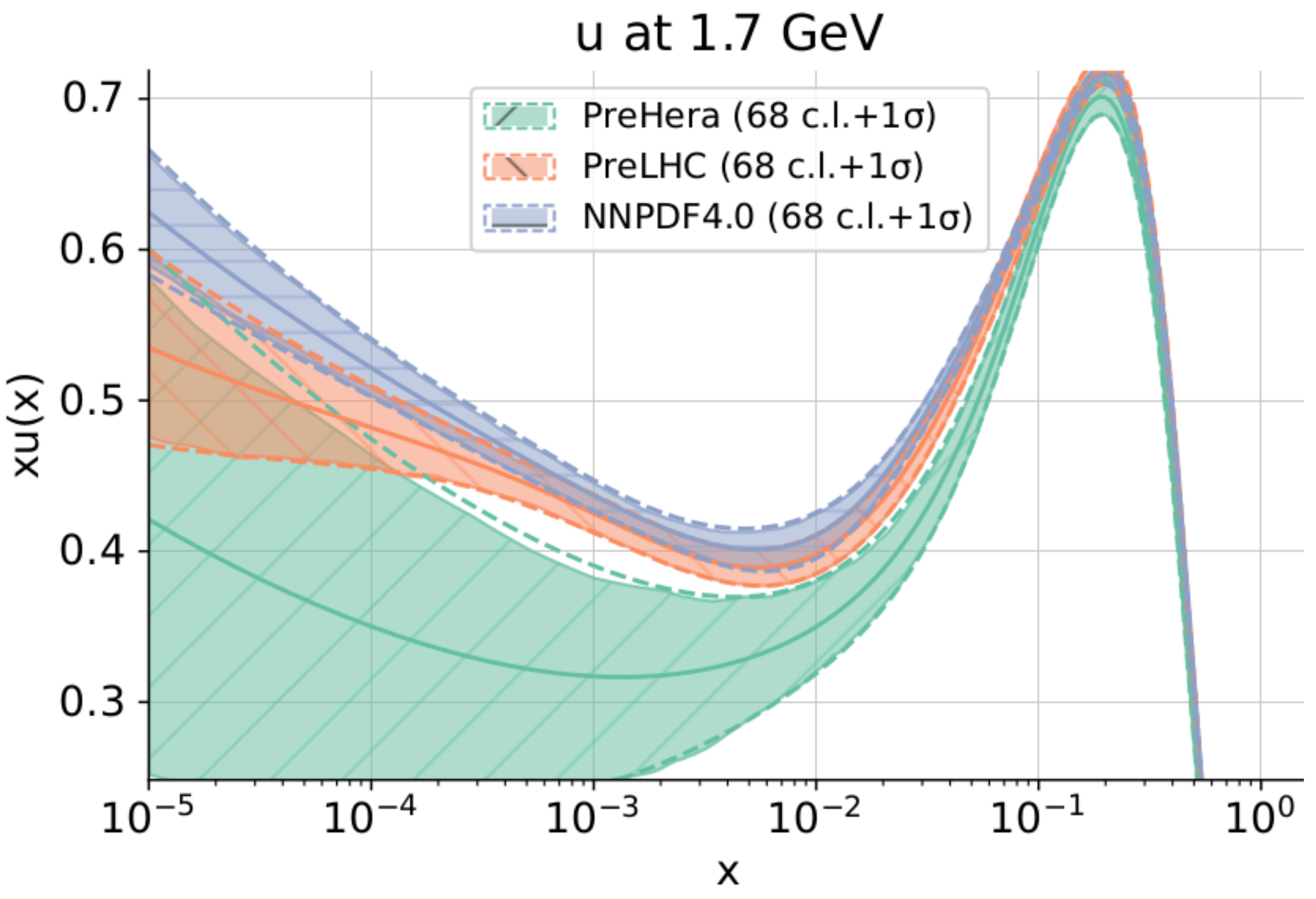
Validation and testing

Future tests



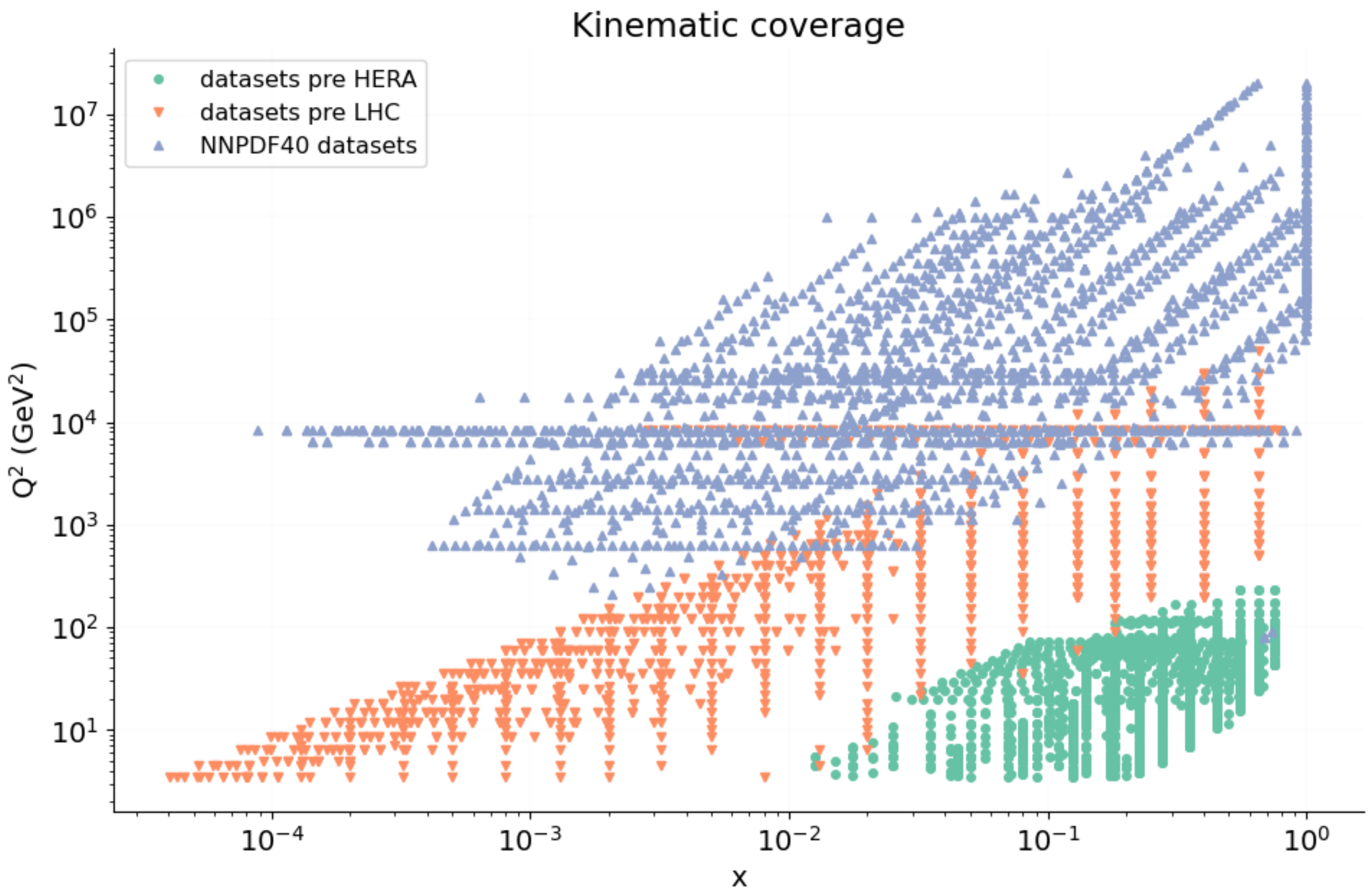
fit / dataset	NNPDF4.0	pre-LHC	pre-HERA
pre-HERA	1.06	1.01	0.91
pre-LHC	1.20	1.21	**26.1**
NNPDF4.0	1.29	**2.15**	**22.57**

χ^2/N



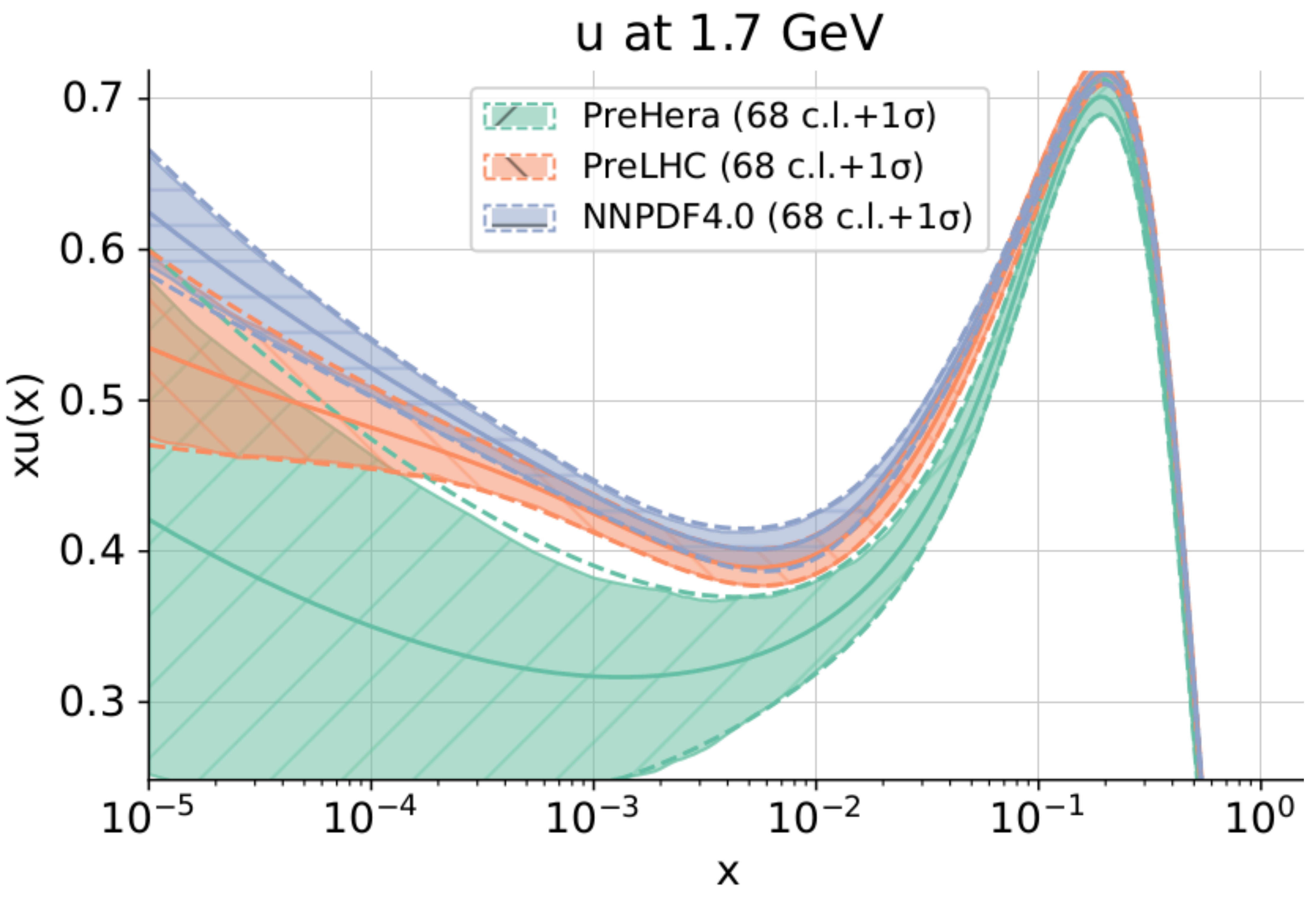
Validation and testing

Future tests

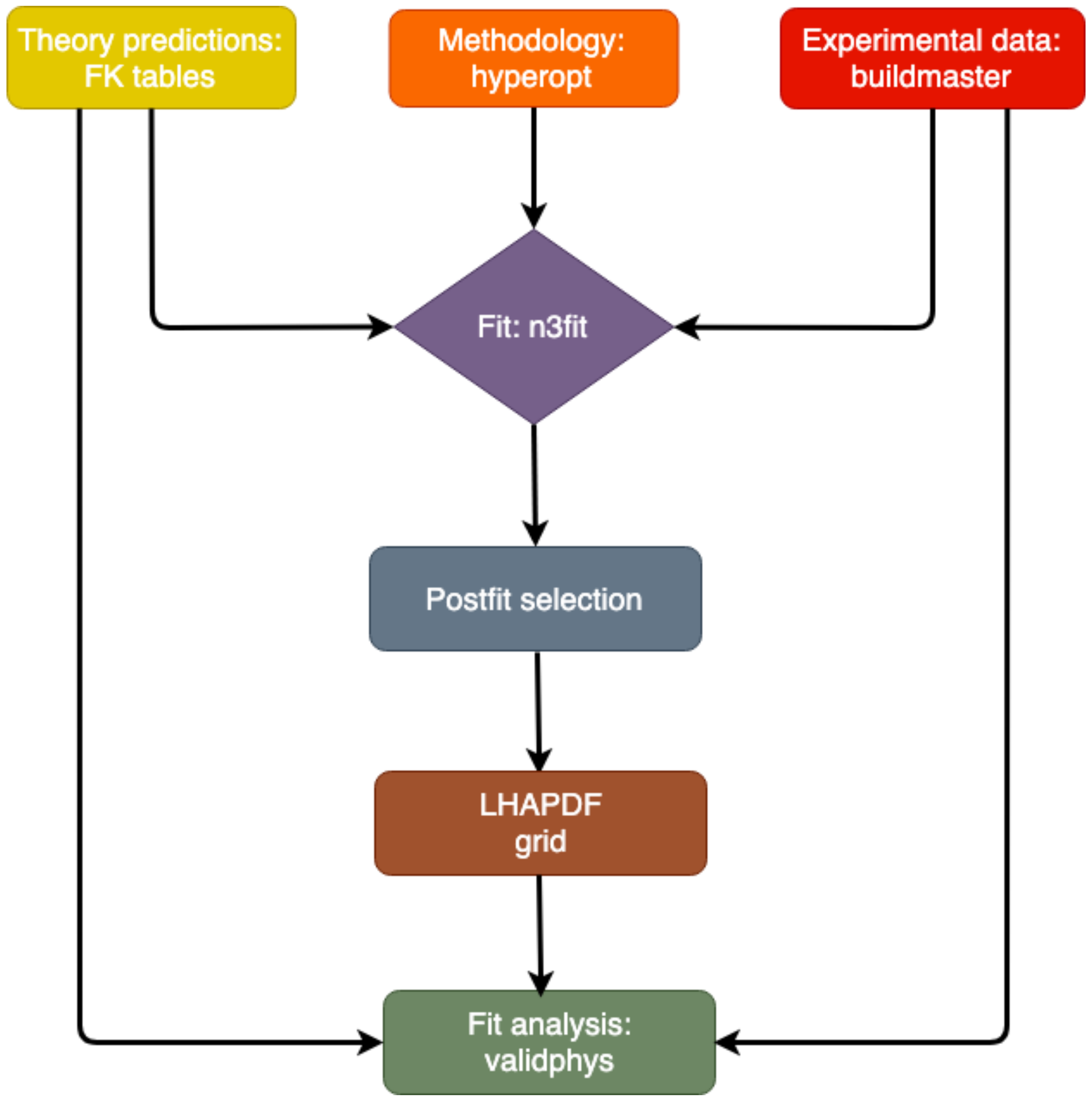


fit / dataset	NNPDF4.0	pre-LHC	pre-HERA
pre-HERA			0.87
pre-LHC		1.18	**1.22**
NNPDF4.0	1.12	**1.30**	**1.38**

χ^2/N



NNPDF fitting framework summary



The ingredients necessary to complete a global PDF fits are:

- Experimental data and uncertainties (hepdata)
- Theory predictions in the form of interpolation tables (plougshare, madgraph): Fast Kernel Tables
- Fitting framework (n3fit) -> PDF at scale Q_0
- DGLAP evolution for any value of Q (Apfel, EKO, Apfel++)
- Postfit selection (eliminate outliers, underlearnt or wiggly replicas and double-check physical constraints)
- Final output: LHAPDF grid
- (optional) an analysis framework to facilitate creating nice plots and presentations

An open-source machine learning framework for global analyses of parton distributions
NNPDF collaboration - [hep-ph] 2109.02671

Open source

The whole NNPDF fitting framework is open source, documented and available to be used for all your PDF fitting needs!

- Code
- Data
- Theory Predictions
- Documentation
- Tutorials

https://github.com/NNPDF/nnpdf

https://docs.nnpdf.science/

Towards a new generation of parton densities with deep learning models
S. Carrazza, **JCM** - [hep-ph] 1907.05075

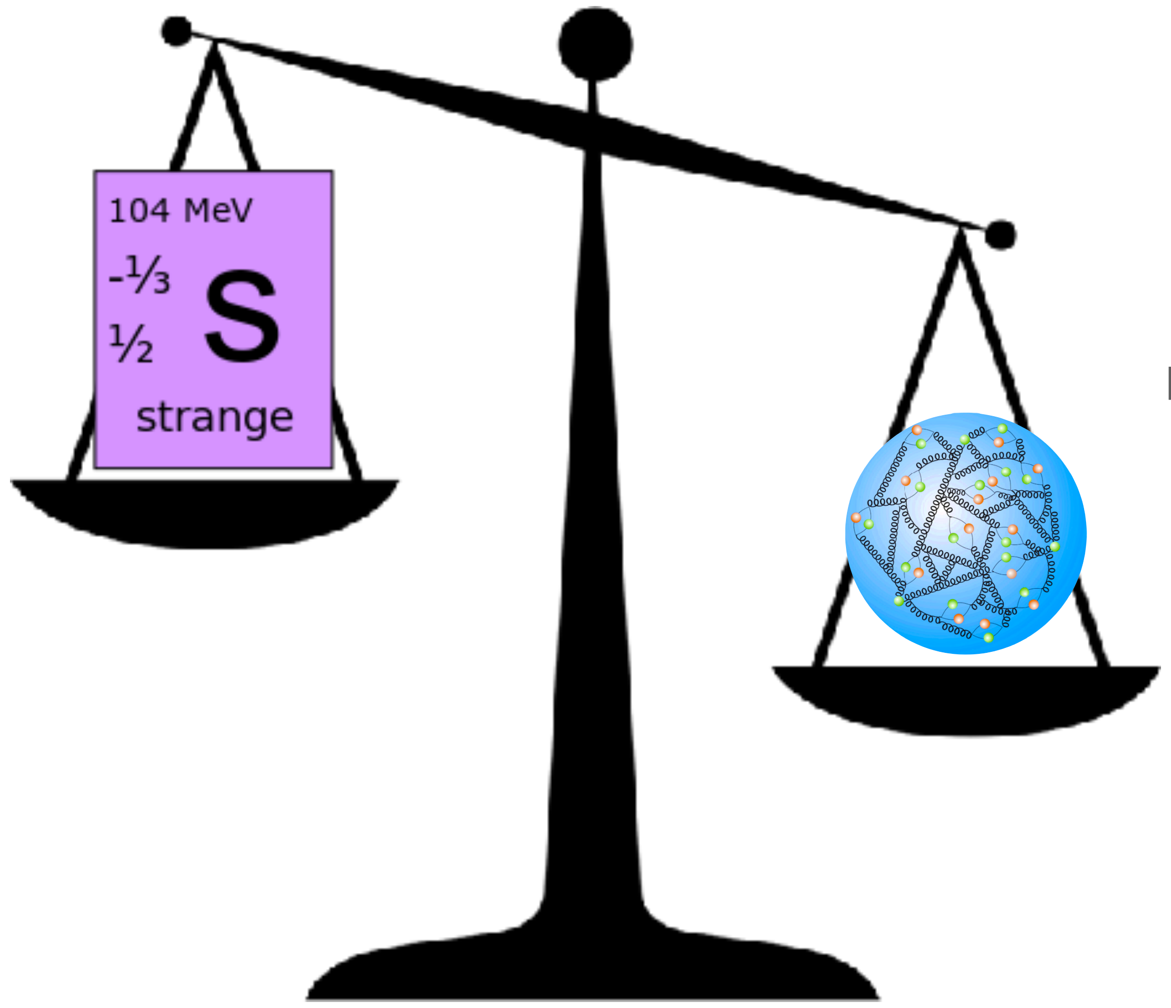
An open-source machine learning framework for global analyses of parton distributions
NNPDF collaboration - [hep-ph] 2109.02671

Results, beyond NNPDF4.0

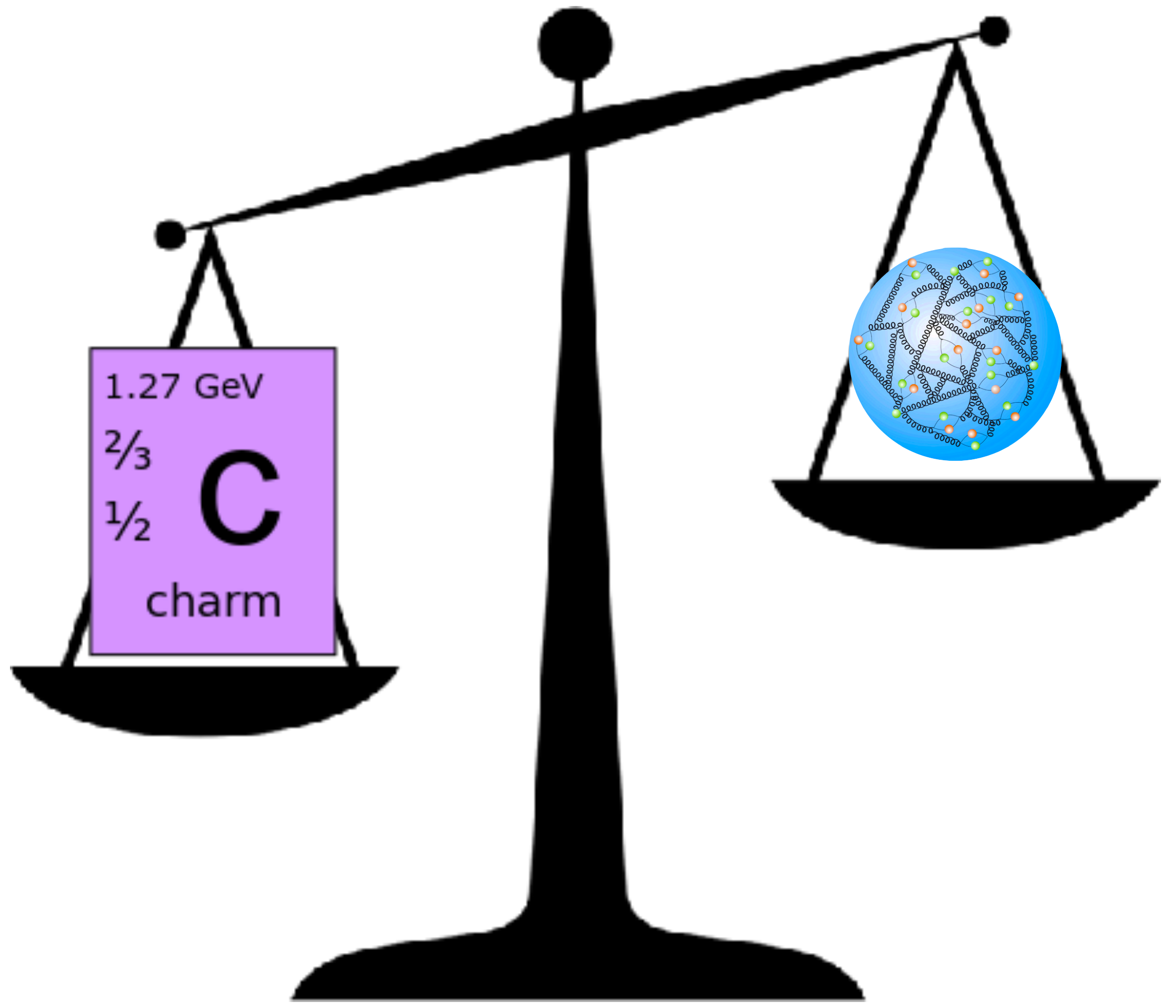
1. Charm in the proton
2. Missing Higher Order Corrections
3. aN3LO PDFs
4. and what now?

A most charming proton

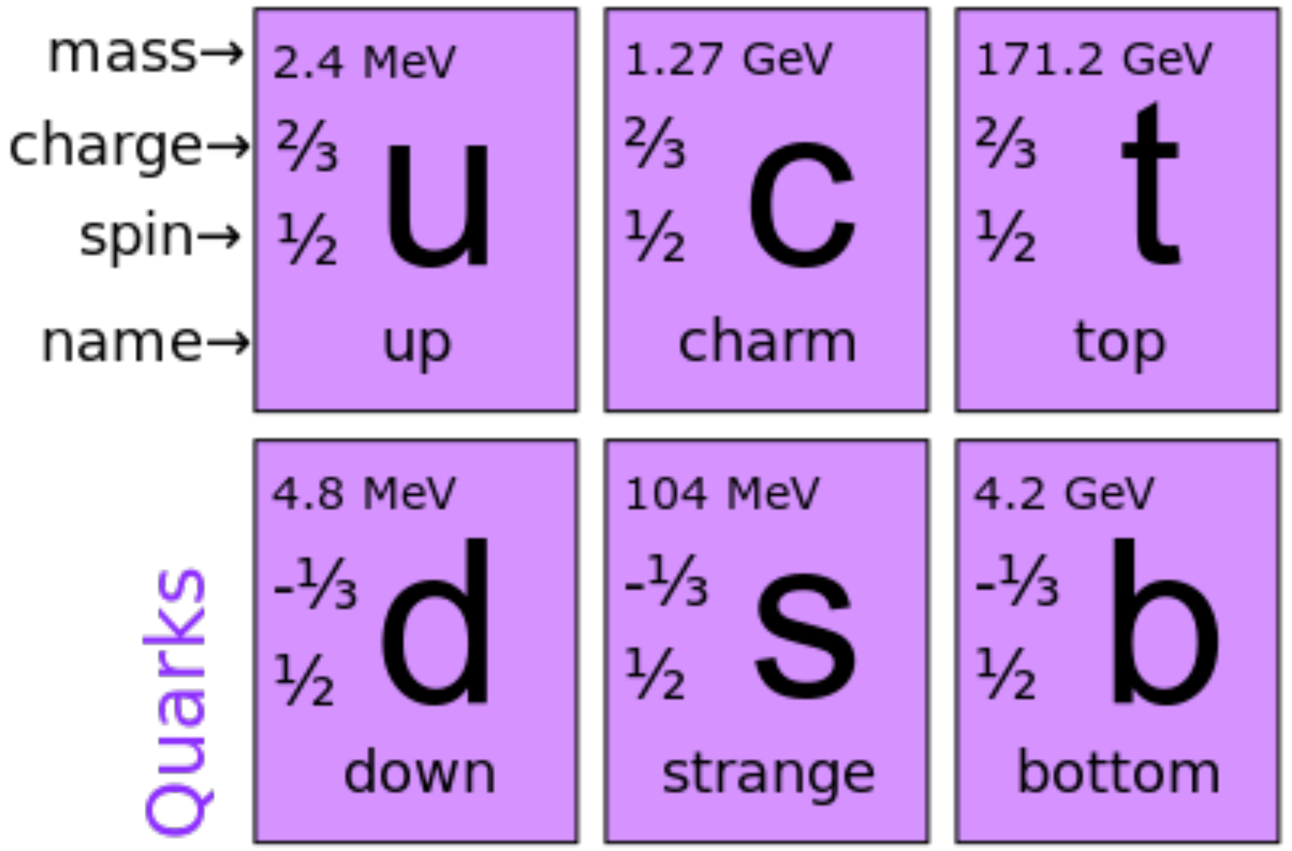
but what about heavier ones?

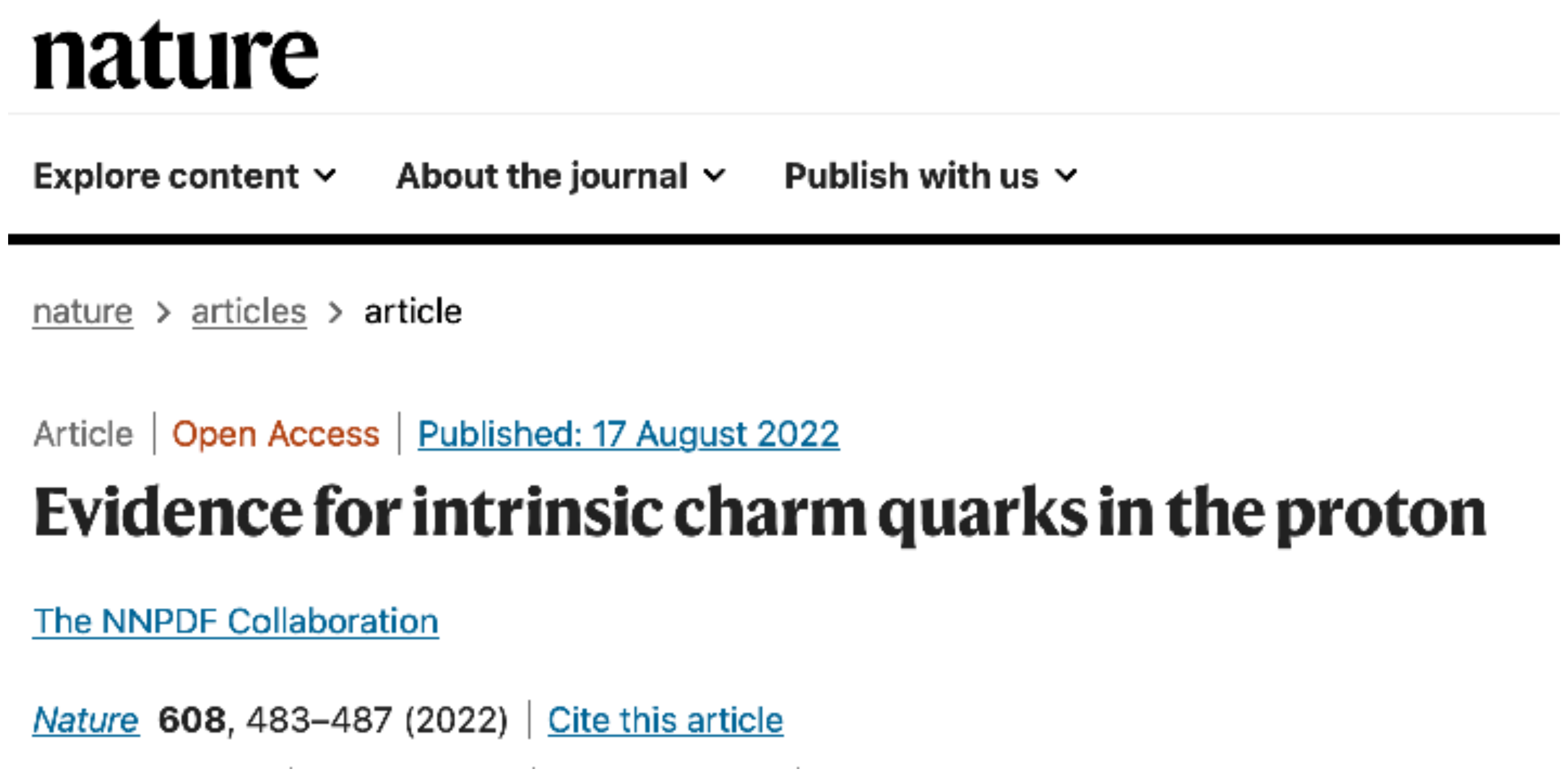


mass of the proton ~ 1 GeV



It's no surprise that we can find quarks lighter than the proton inside of it…




nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature > articles > article

Article | Open Access | Published: 17 August 2022

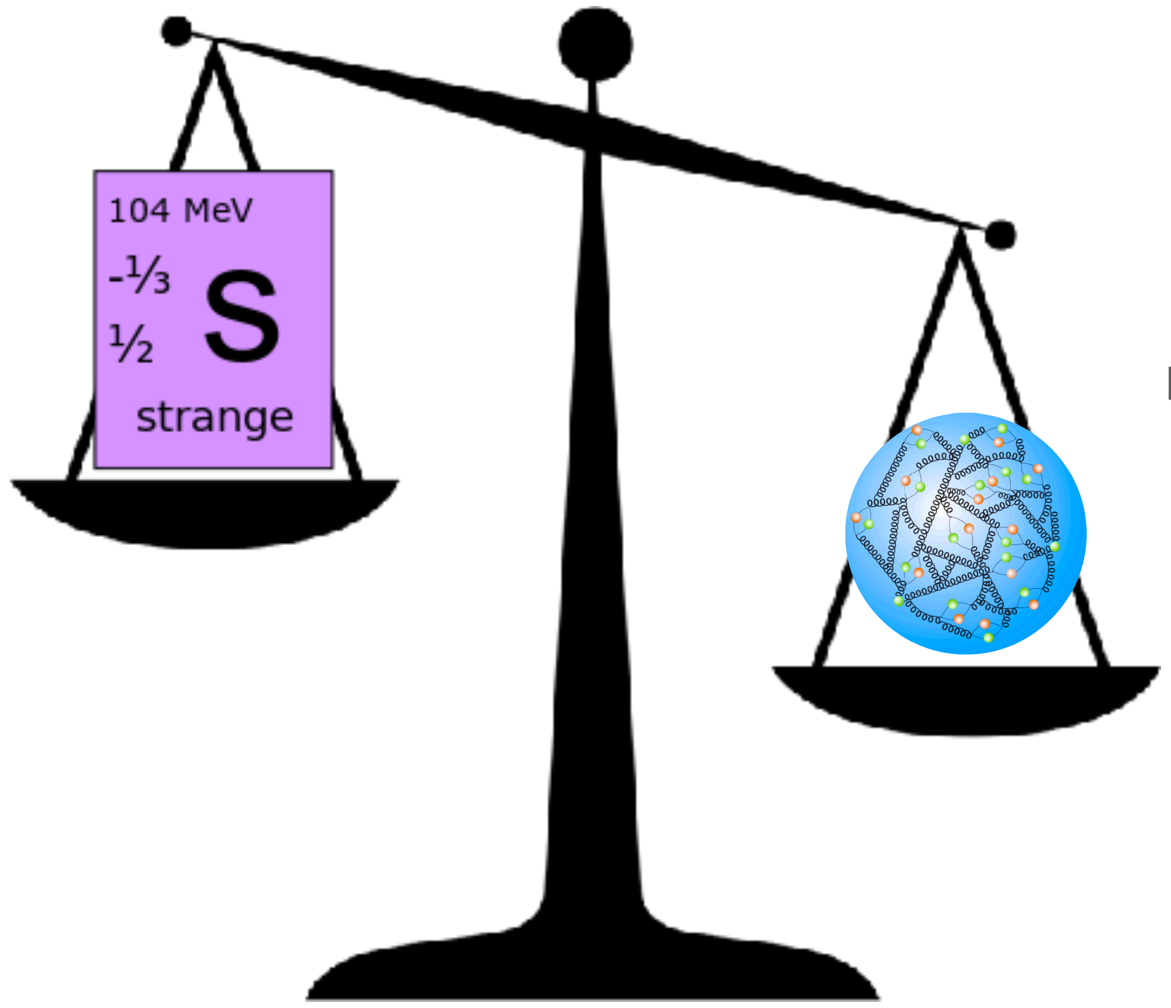
Evidence for intrinsic charm quarks in the proton

The NNPDF Collaboration

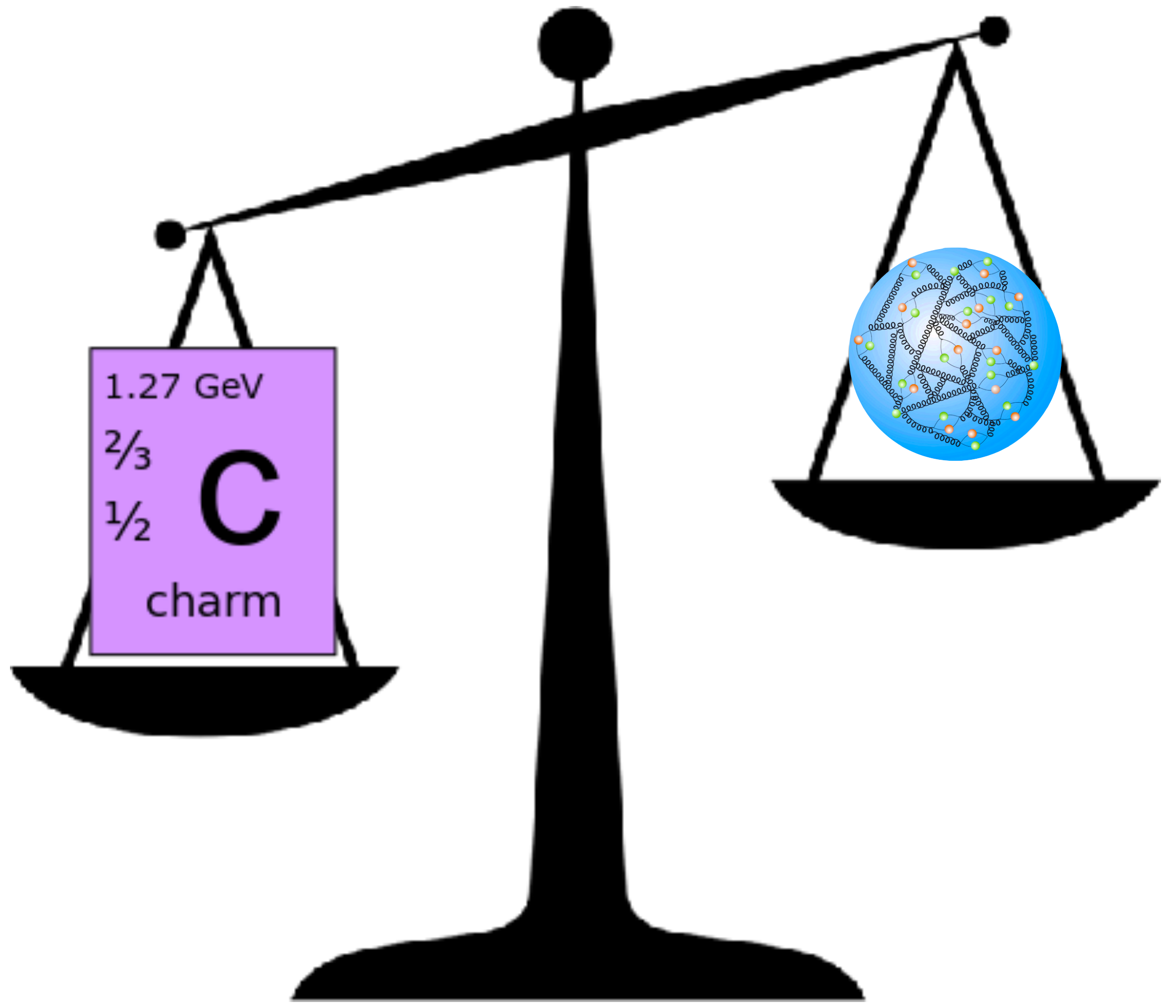
Nature **608**, 483–487 (2022) | Cite this article

A most charming proton

but what about heavier ones?



mass of the proton ~ 1 GeV



It's no surprise that we can find quarks lighter than the proton inside of it…

mass→ charge→ spin→ name→	2.4 MeV ⅔ ½ **u** up	1.27 GeV ⅔ ½ **c** charm	171.2 GeV ⅔ ½ **t** top
Quarks	4.8 MeV -⅓ ½ **d** down	104 MeV -⅓ ½ **s** strange	4.2 GeV -⅓ ½ **b** bottom

New Scientist

Physics

Physicists surprised to discover the proton contains a charm quark

The textbook description of a proton says it contains three smaller particles - two up quarks and a down quark - but a new analysis has found strong evidence that it also

It's not a new idea, but it has not been easy to prove it.

THE INTRINSIC CHARM OF THE PROTON

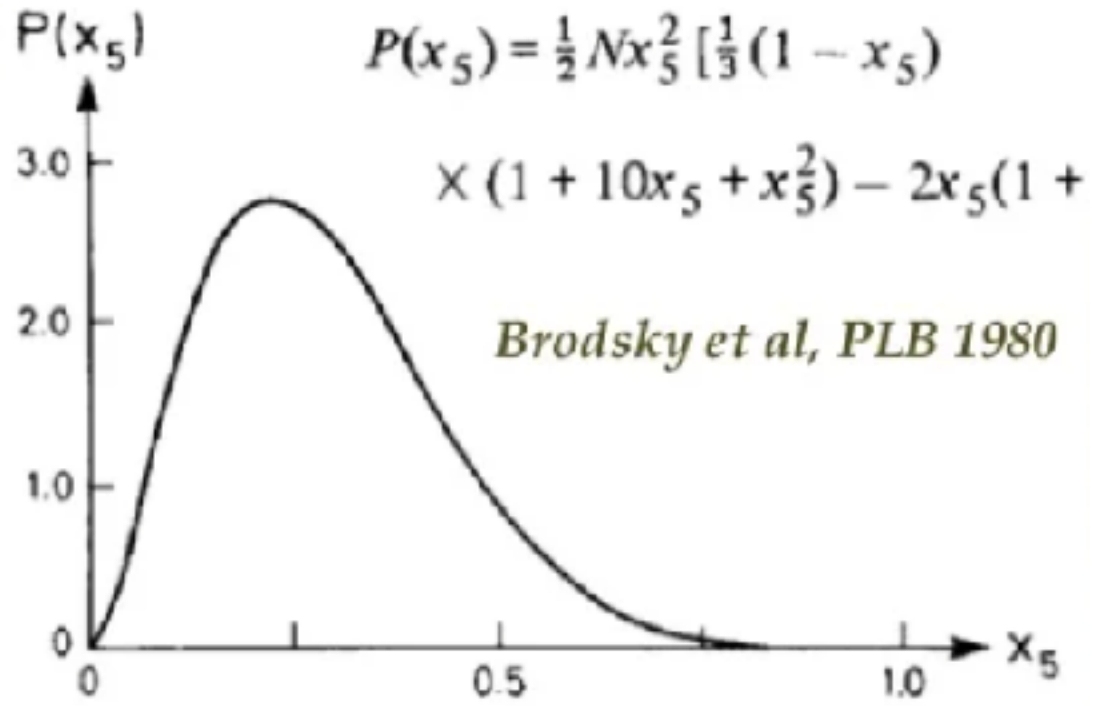
S.J. BRODSKY [1]
Stanford Linear Accelerator Center,
Stanford, California 94305, USA

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.

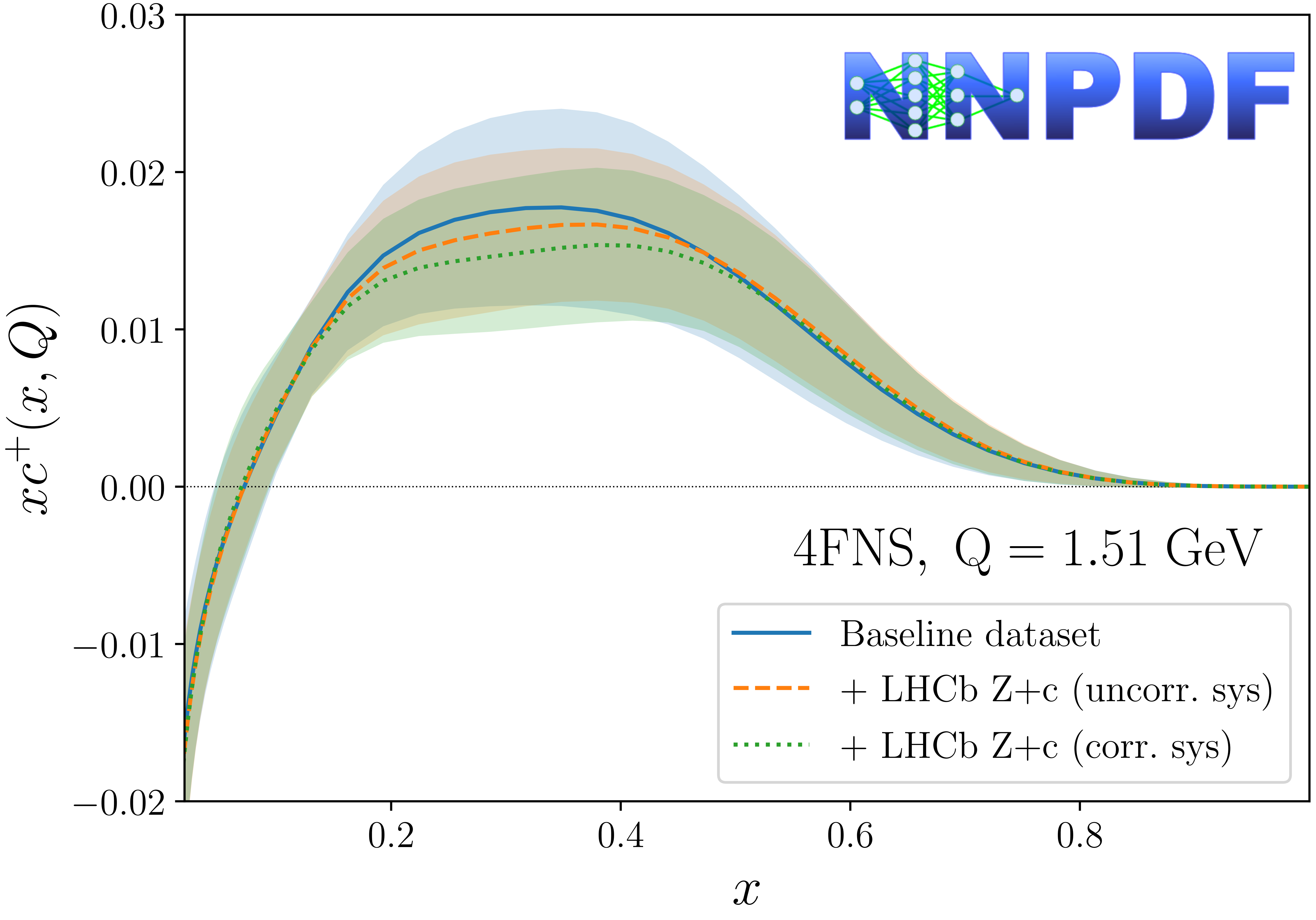


$P(x_5) = \frac{1}{2} N x_5^2 [\frac{1}{3}(1 - x_5)$

$\times (1 + 10x_5 + x_5^2) - 2x_5(1 +$

Brodsky et al, PLB 1980

Intrinsically charming



Evidence for intrinsic charm quarks in the proton
NNPDF Collaboration — [hep-ph] 2208.08372

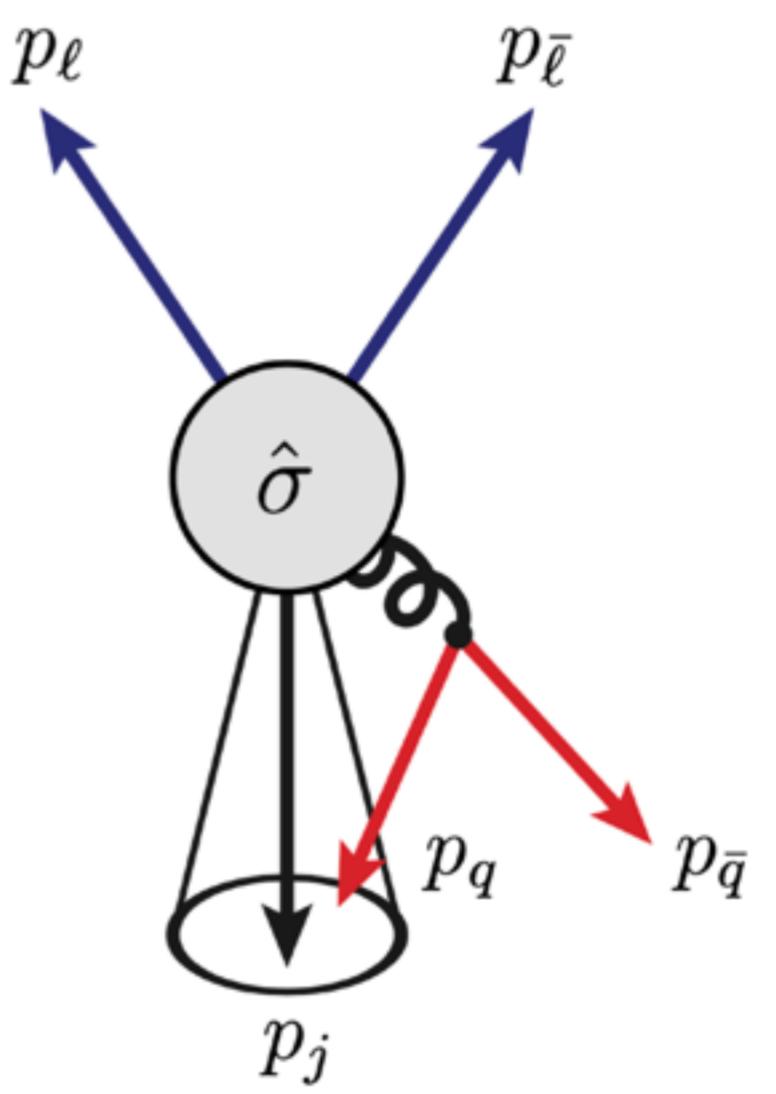
NNPDF is the only collaboration which fits charm by default, i.e., $c^+ = c + \bar{c} \neq 0$ at the fitting scale which means the contribution is not limited to DGLAP evolution

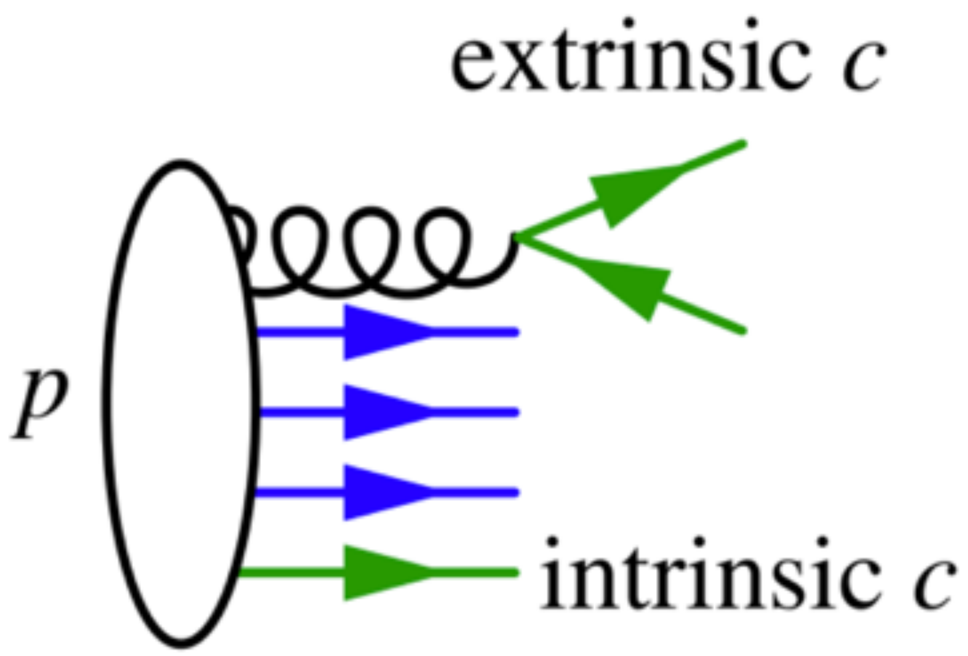
Open challenges:

- Better grasp of MHOU
- Improved jet algorithms

  in order to match data and predictions…

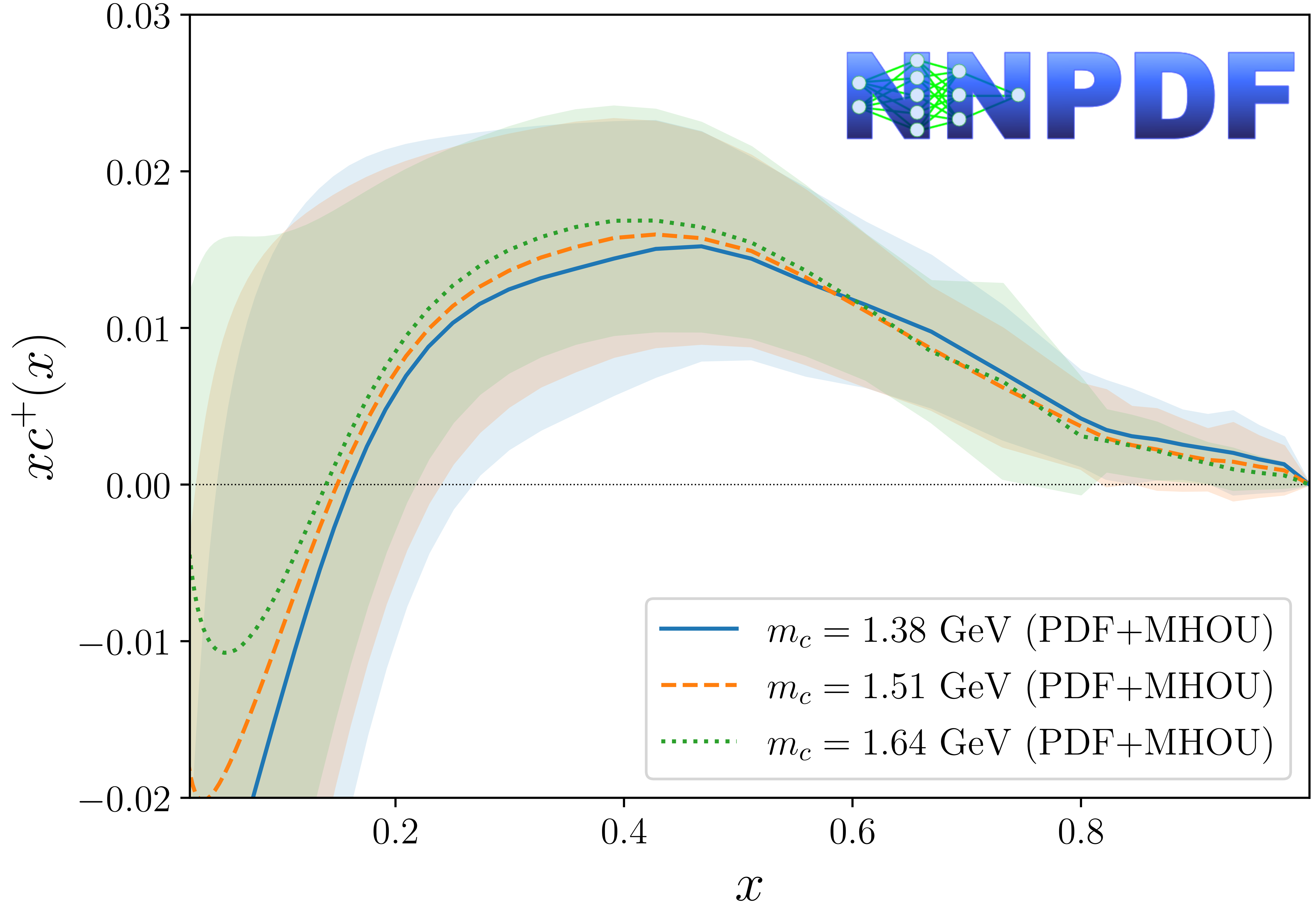
  collinear-safe jet algorithms need to be used





Diagrams taken from talks by G. Stagnitto and D. Zuliani

Intrinsically charming



Evidence for intrinsic charm quarks in the proton
NNPDF Collaboration — [hep-ph] 2208.08372

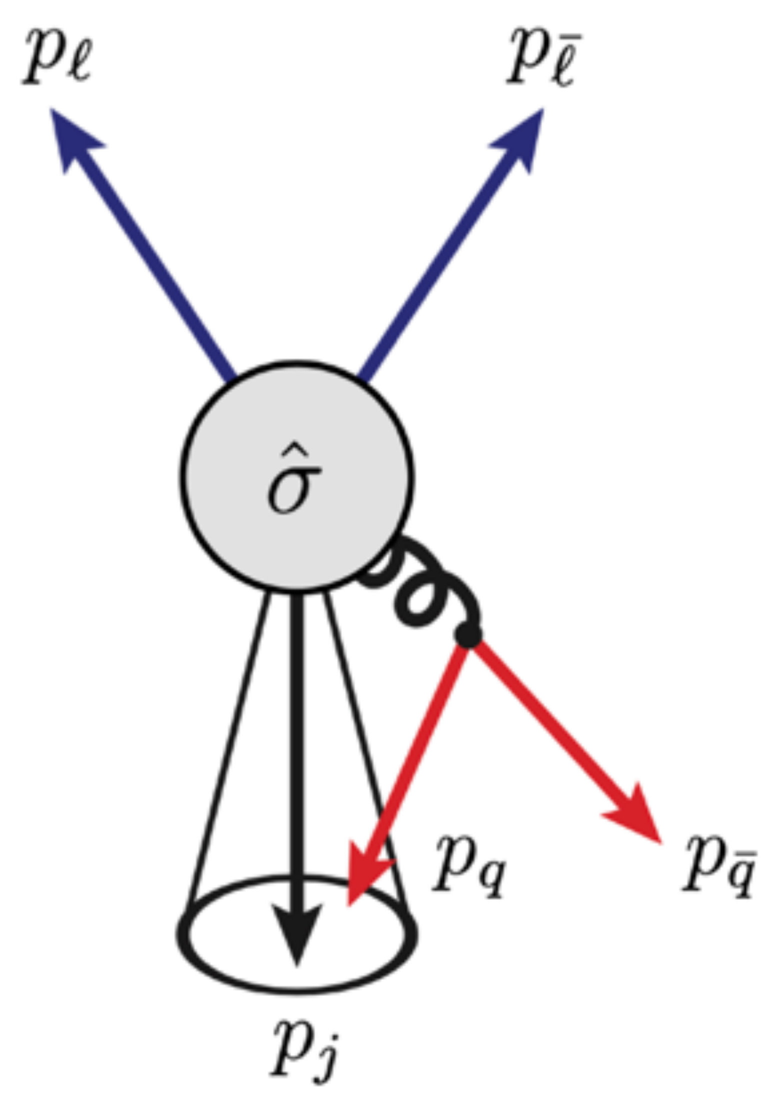
NNPDF is the only collaboration which fits charm by default, i.e., $c^+ = c + \bar{c} \neq 0$ at the fitting scale which means the contribution is not limited to DGLAP evolution

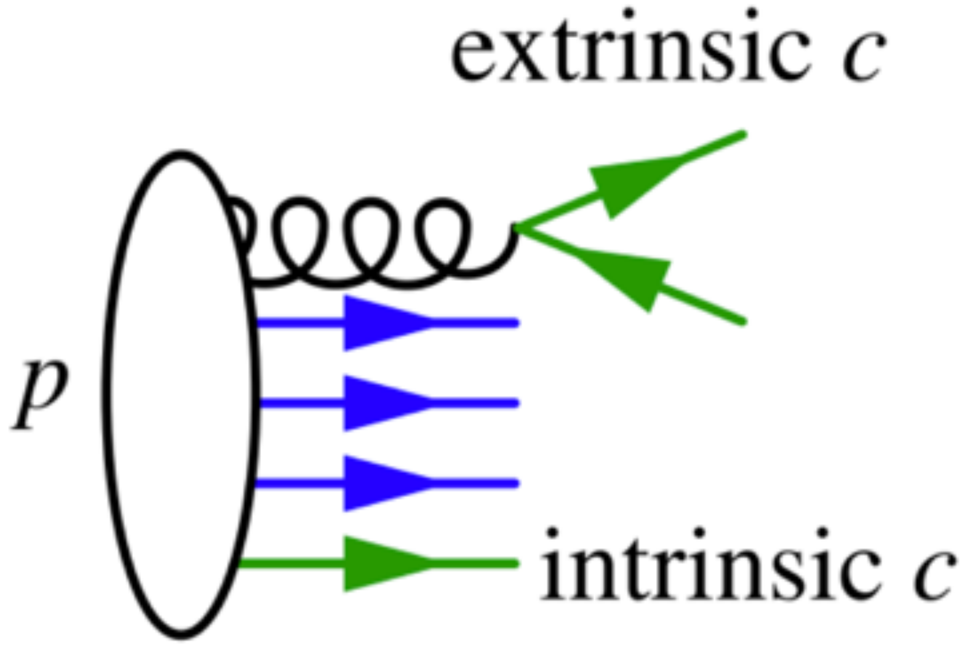
Open challenges:

- Better grasp of MHOU
- Improved jet algorithms

in order to match data and predictions…

collinear-safe jet algorithms need to be used

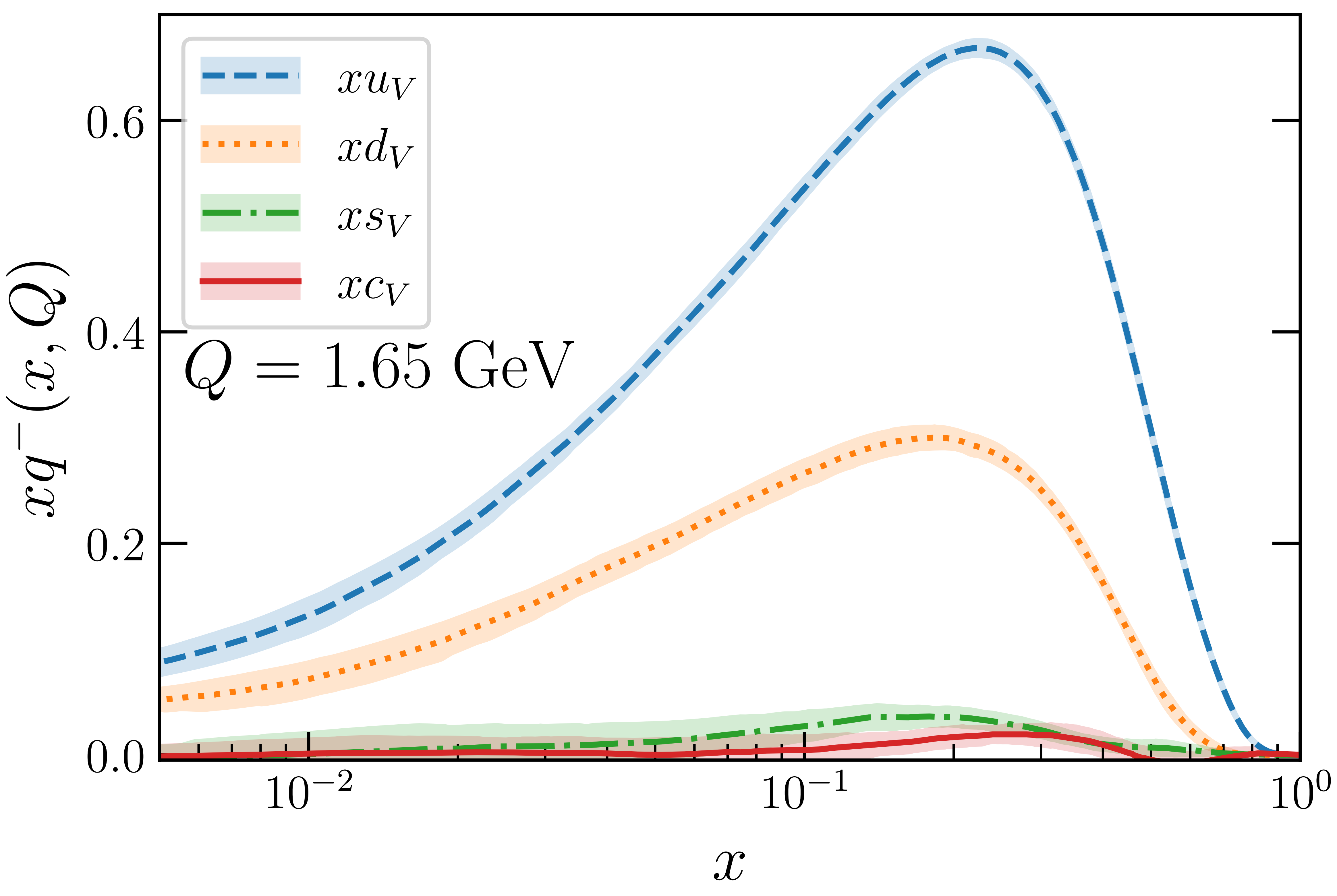


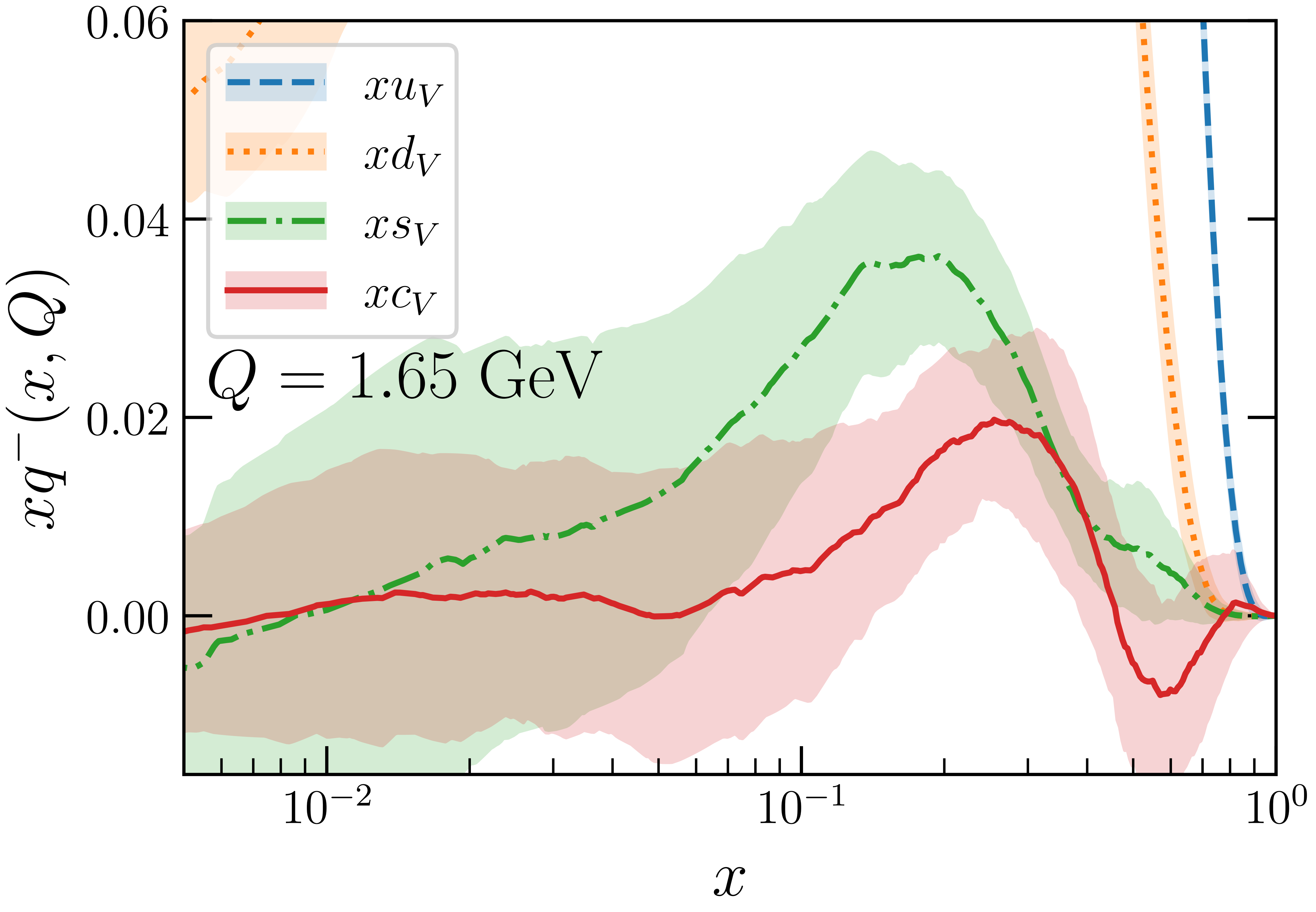


Diagrams taken from talks by G. Stagnitto and D. Zuliani

~~Intrinsically~~ Asymmetrically charming

The determination of the charm content of the proton assumed 0 charm asymmetry ($c_v = c - \bar{c} = 0$) for purely practical reasons… however there's no reason why the charm should behave differently than other quarks.





Proving the charm-anticharm asymmetry would prove the non-perturbative nature of the charm component of the proton!

Intrinsic charm quark valence distribution of the proton
NNPDF collaboration - hep-ph/2311.00743

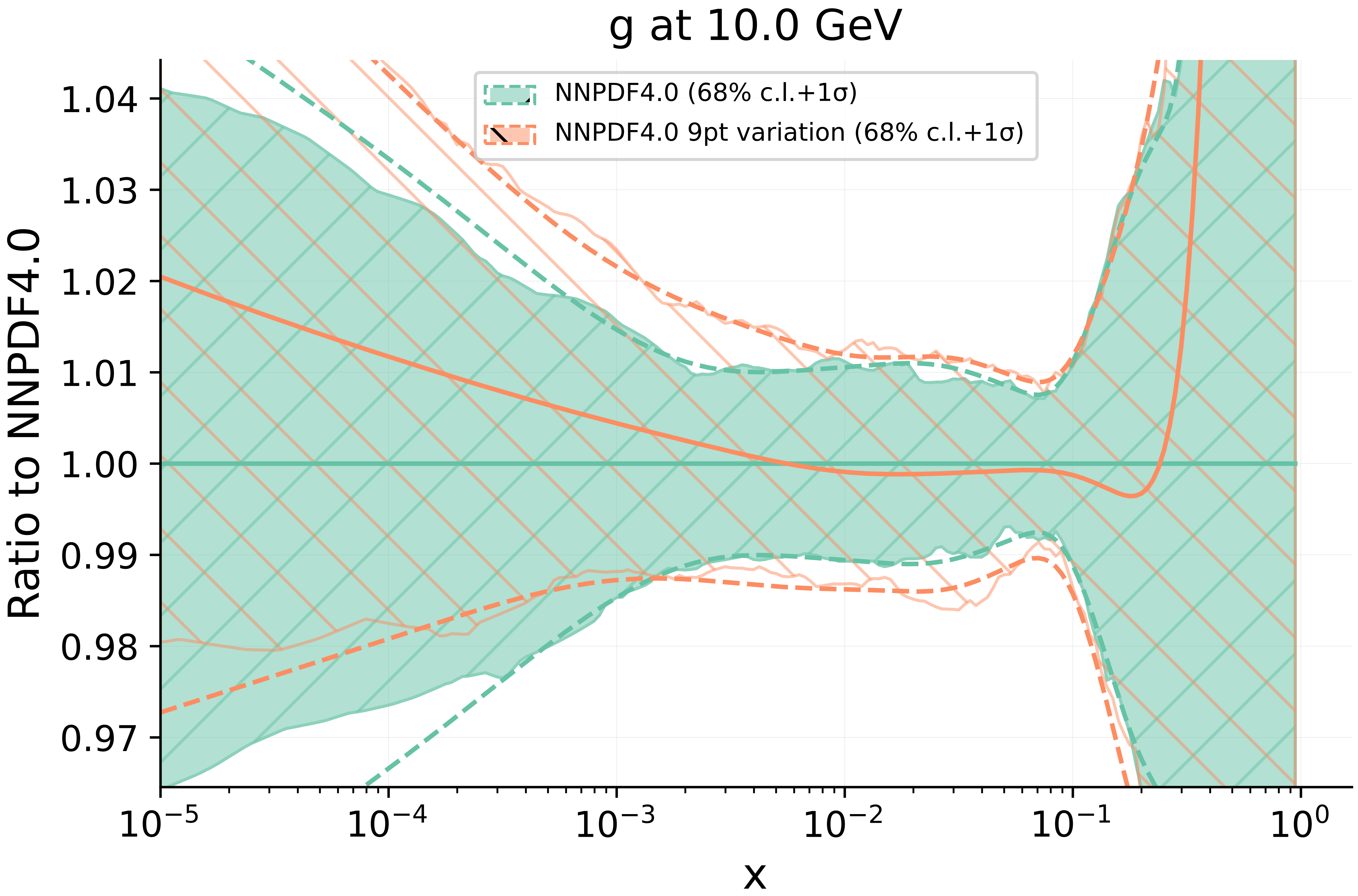
Missing Higher Orders

Uncertainties beyond the data

PDF uncertainties are propagated only from the data but this is just half of the story, fixed-order predictions also contain uncertainties:

$$\sigma_{NNLO} = \sigma_0 + \alpha_s \sigma_1 + \alpha_s^2 \sigma_2 + \cancel{\mathcal{O}}(\alpha_s^3)$$

A spurious dependence on unphysical scales (renormalization, factorization) is kept. This is exploited to generate a “theory uncertainty” to estimate missing higher orders.



Theory uncertainties are included in the fit by constructing a “theory covariance matrix”

$$cov_{ij} = cov_{ij}^{\text{exp}} + cov_{ij}^{\text{th}}$$

Determination of the theory uncertainties from missing higher orders on NNLO parton distributions with percent accuracy
NNPDF collaboration - hep-ph/2401.10319

N3LO: the next frontier

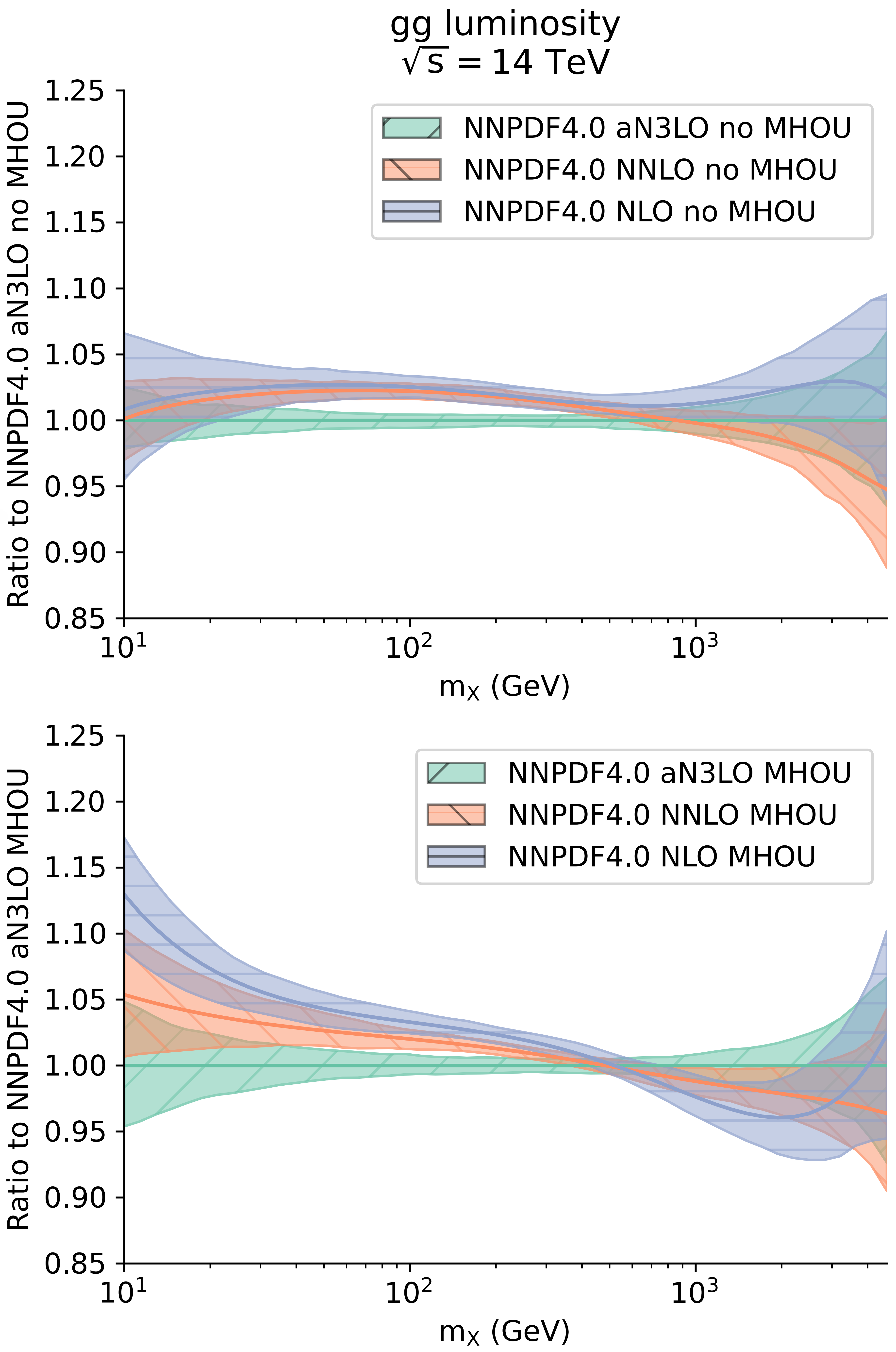
First N3LO results by the NNPDF and MSHT collaborations:

- Order-by-order converge improved in both PDFs and phenomenological predictions.
- Good agreement between both collaborations' results, despite using very different strategies for the approximation.
- Necessary to match the order of theoretical calculations.

Results are not *exact* N3LO but rather **aN3LO** (for approximated)

- Splitting functions not fully known, but known in enough limits to build a reasonable approximation. Details and benchmarks in 2406.16188 and references therein.
- Fiducial double-hadronic predictions are only accessible at NNLO (and many only through k-factors).
- The effect of missing contributions estimated through scale variations, greatly improving convergence.

The path to N3LO parton distributions
NNPDF collaboration - hep-ph/2402.18635



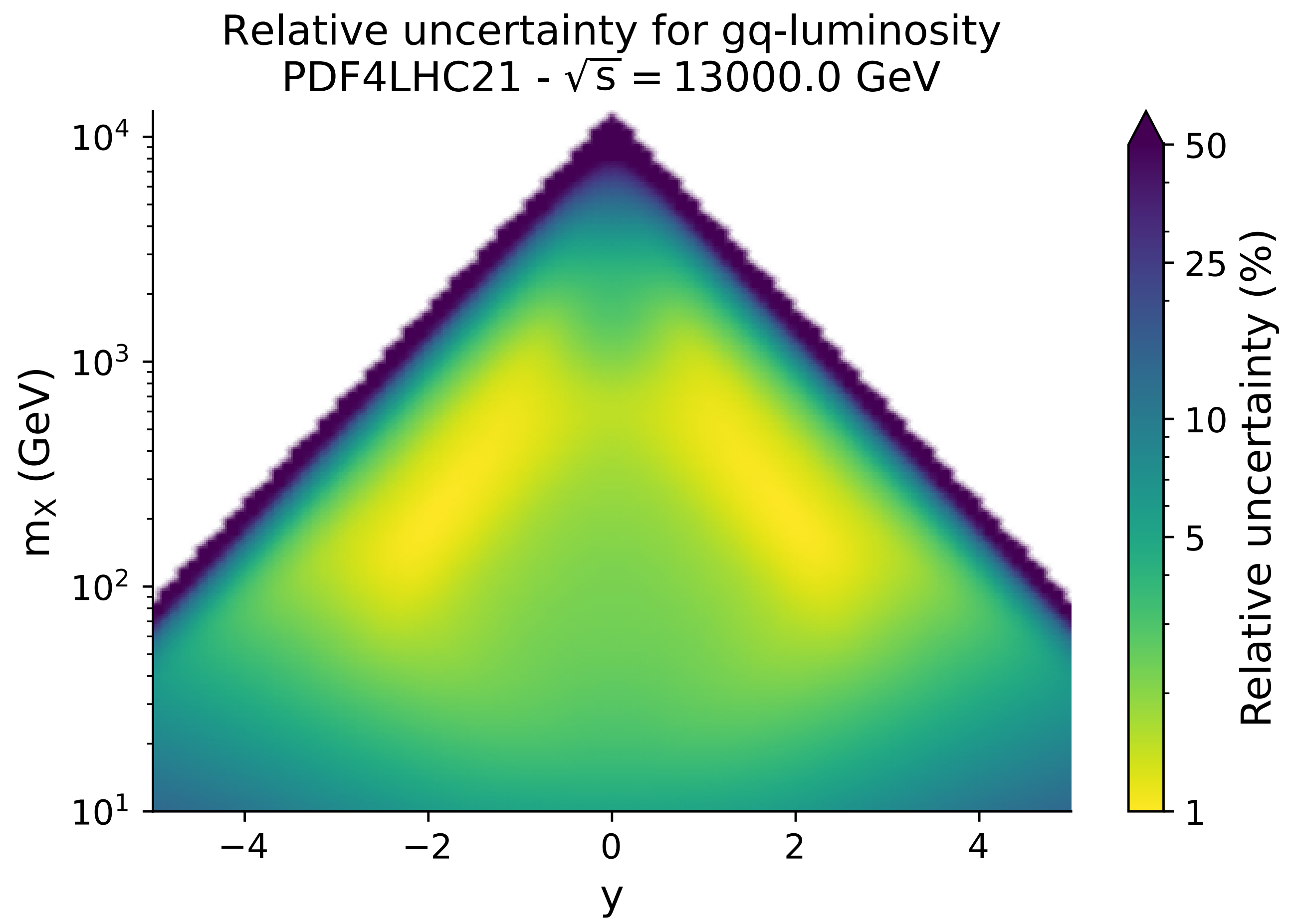
And what now?

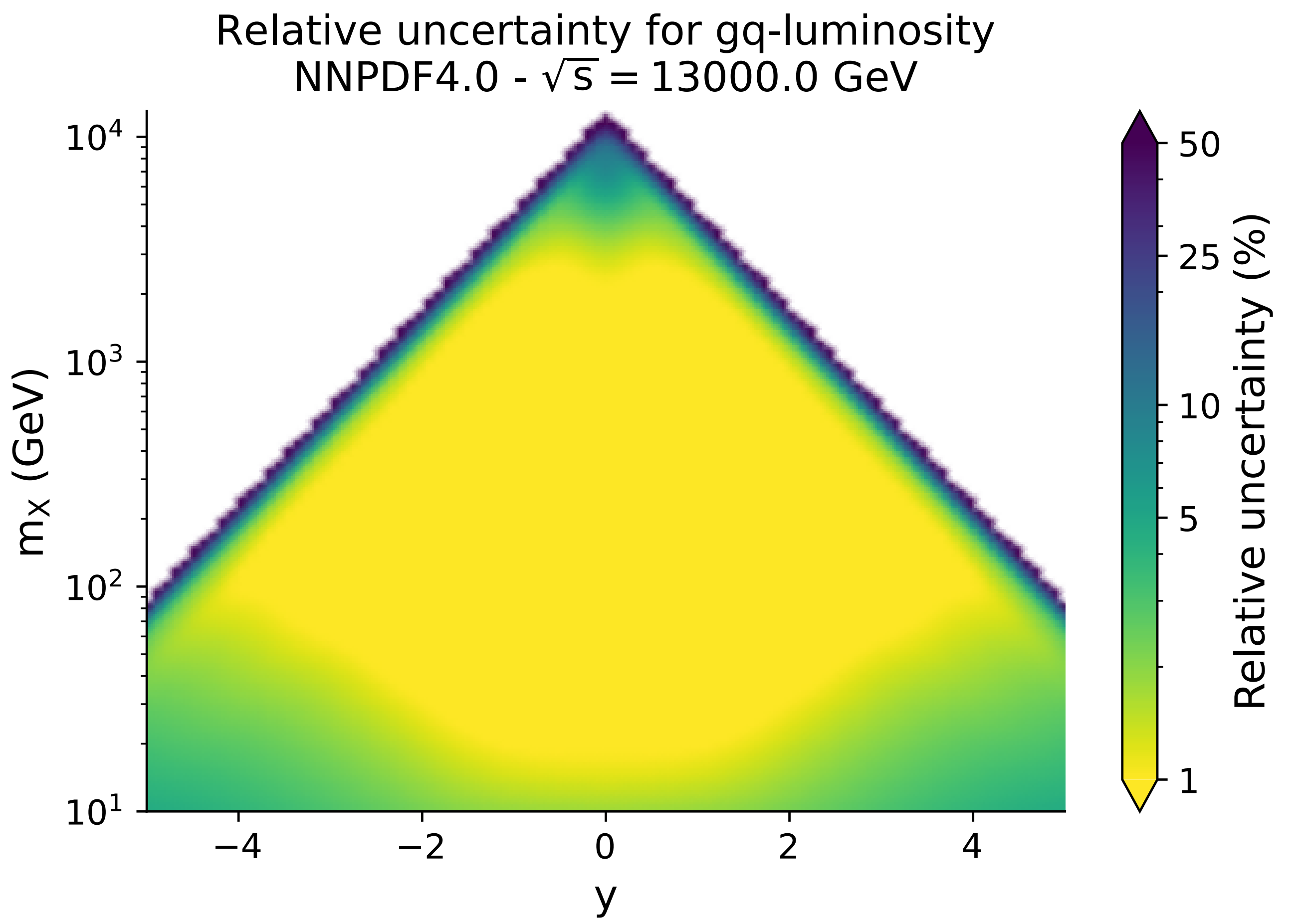
Open challenges in PDF fitting:

- Exact N3LO evolution
- Exact NNLO grids for hadronic coefficients and N3LO k-factors
- Diminishing returns: every step of the way becomes computationally more costly and complex for a smaller and smaller improvements. **But it is very much needed! Missing effects are competitive with the current uncertainty limits!** (and many other effects still not under consideration: TMDs, resummation, higher twists)

Maybe try to change the computing paradigm?

R&D in PDF determination and beyond





Percent-level uncertainties in PDF determination is only a recent achievement, made possible by the large amount of data from runs I and II of the LHC!
Plenty of work is still to be done.

Can the PDF fitting methodologies be improved?

Is it possible to fit PDFs with exact NNLO/N3LO coefficients?

Simultaneous multiparametric fits

NNLO grids are a necessity

Is feasible to compute fiducial cross sections at N3LO?
can we even afford it!?

Quantum Computing

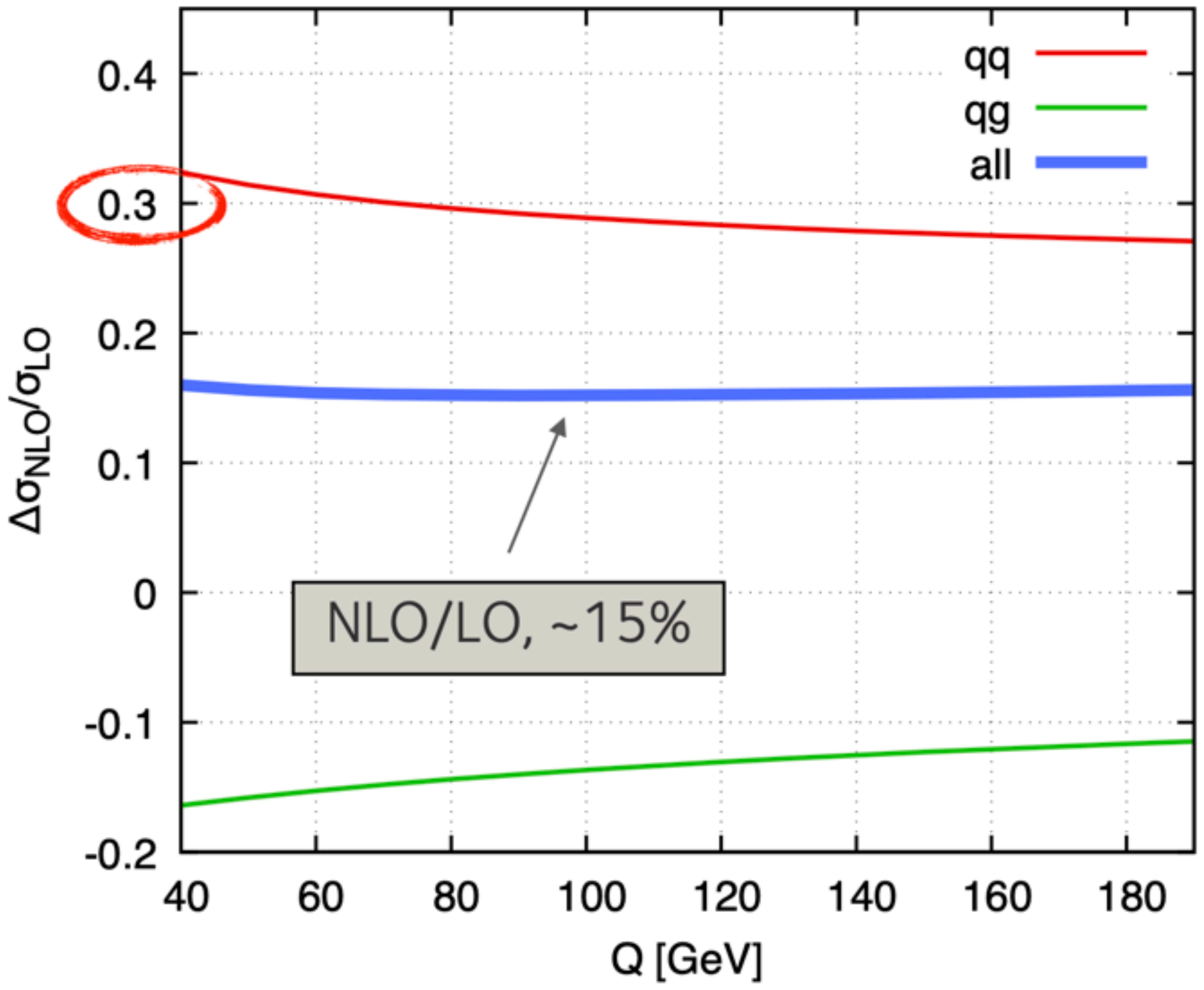
Hardware accelerators

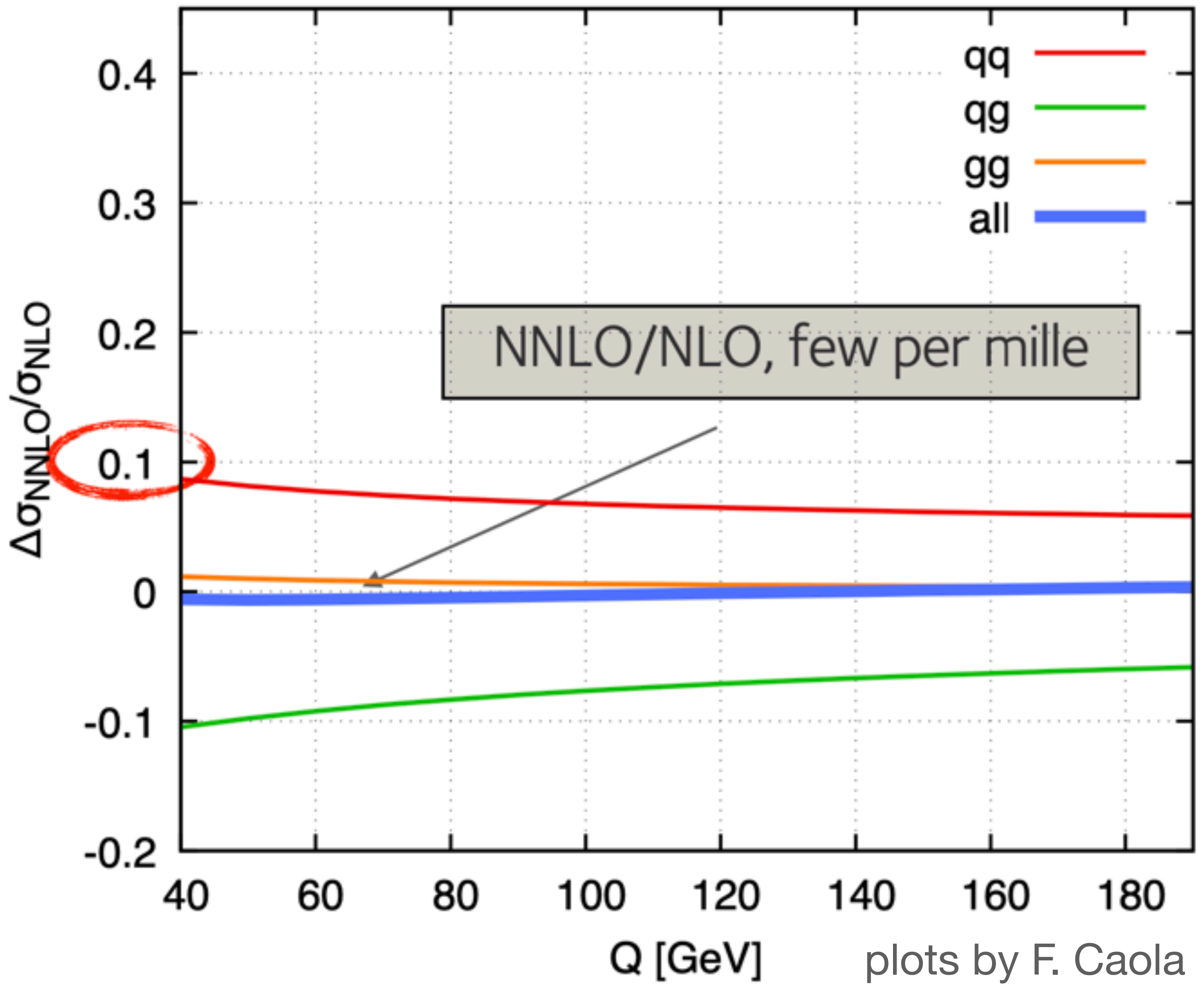
The importance of Theory Predictions

is the problem really solved?

Current frontier: N3LO

PDF Fits, however, are still struggling with implementing (already existing) NNLO corrections (most are implemented through k-factors!) or NLO EW





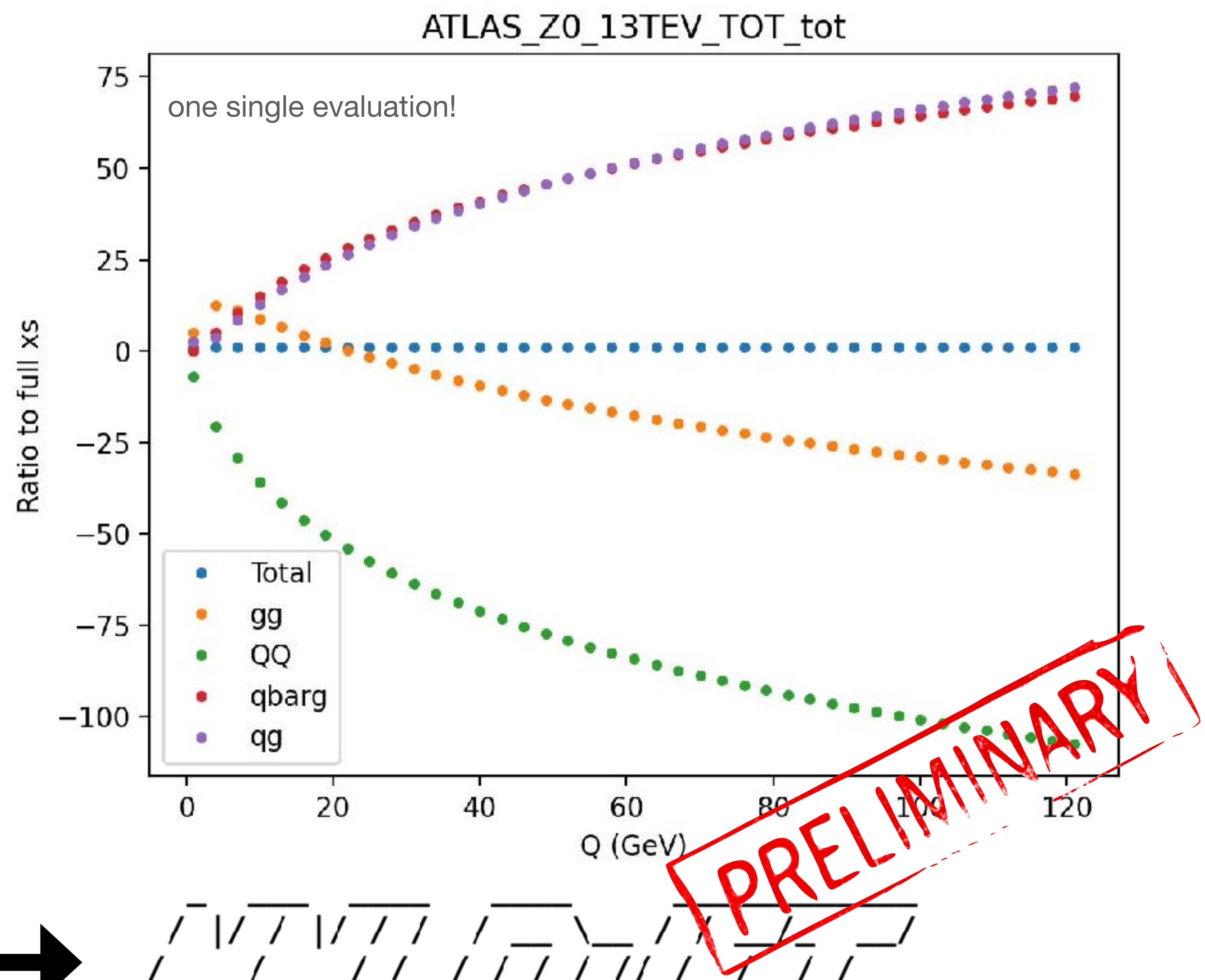
plots by F. Caola

Despite the recent N3LO PDFs, a lot of exciting work still to be done at NNLO!

Exact NNLO predictions now in grid form!

While N3LO corrections exists for many processes, only NLO exact calculations for hadron collisions have until now been used in PDF extractions: **due to the computational cost of higher order corrections**

While observable-differential k-factors are a good approximation, interpolation grids are essential to get the exact channel-by-channel break-down necessary for an accurate determination of Parton Distribution Functions.



PineAPPL

PineAPPL: combining EW and QCD corrections for fast evaluation of LHC processes
C. Schwan et al. - [hep-ph] 2008.12789

PineAPPL: combining EW and QCD corrections for fast evaluation of LHC processes
NNLOJET collaboration - in preparation

Fiducial grids for vector boson production in hadron collisions
JCM, A. Huss, C. Schwan - in preparation

PDFs vs New Data

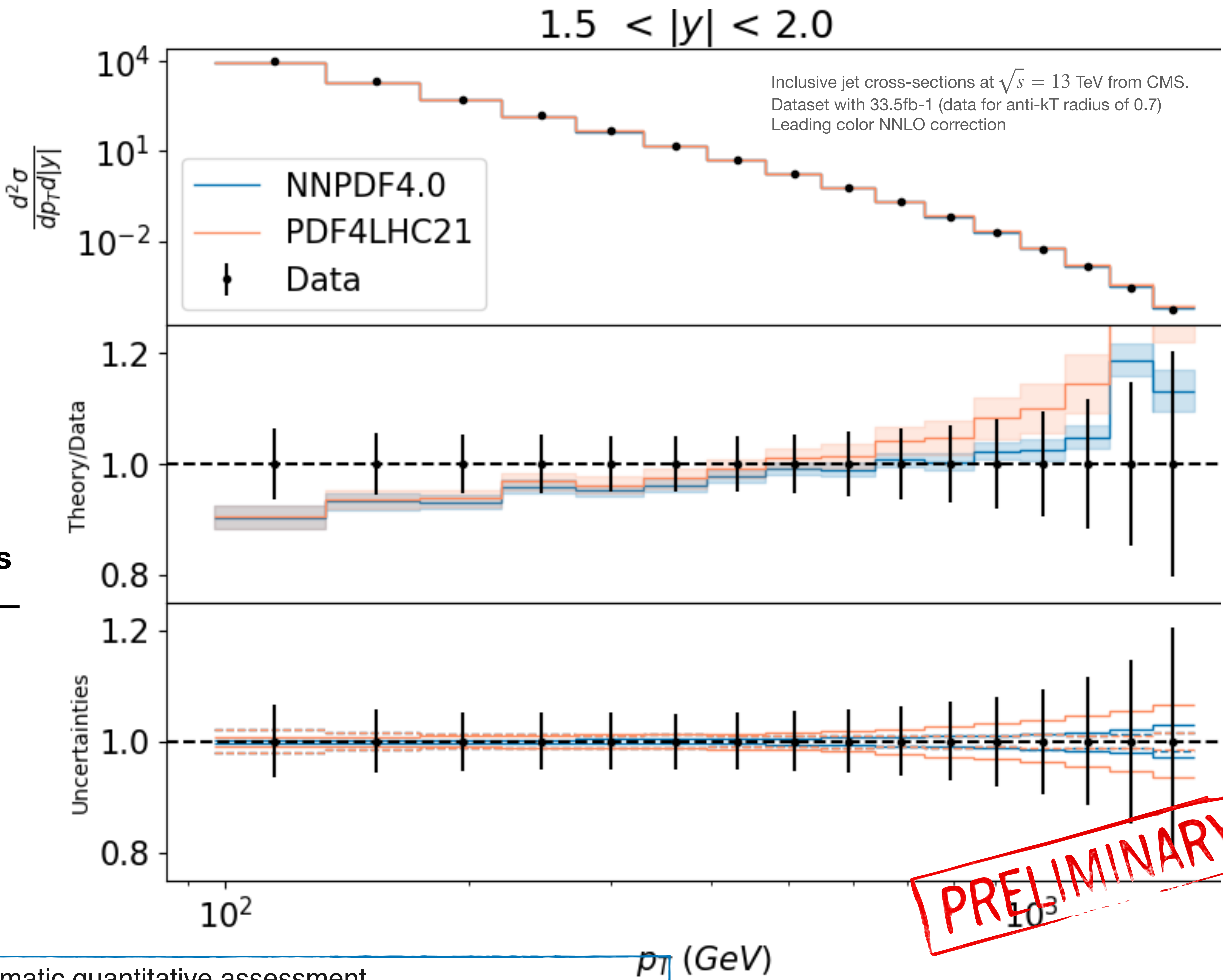
The final goal of PDFs determination is to construct objects that enable us to predict observables.

Hence, a systematic data-theory comparison is the only true test of PDF.

Predictions: NNLOJET

Hepdata: 10.17182/hepdata.115022.v2

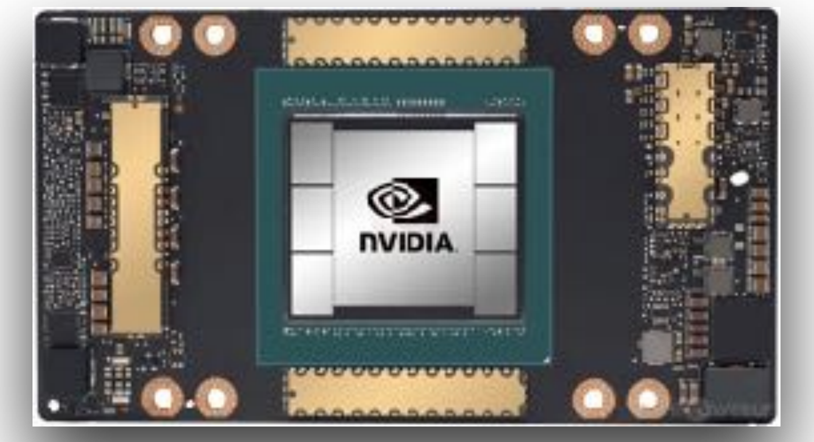
χ^2/N	Only exp. and th. unc.	All uncertainties
PDF4LHC21	4.76	2.85
NNPDF40	3.81	3.23

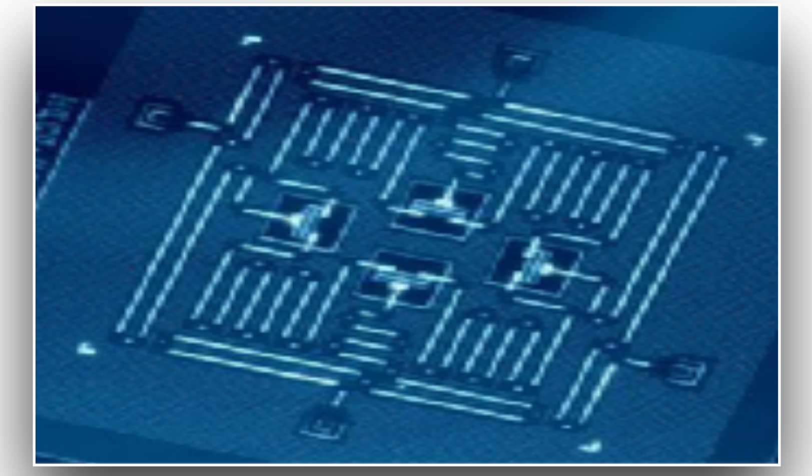


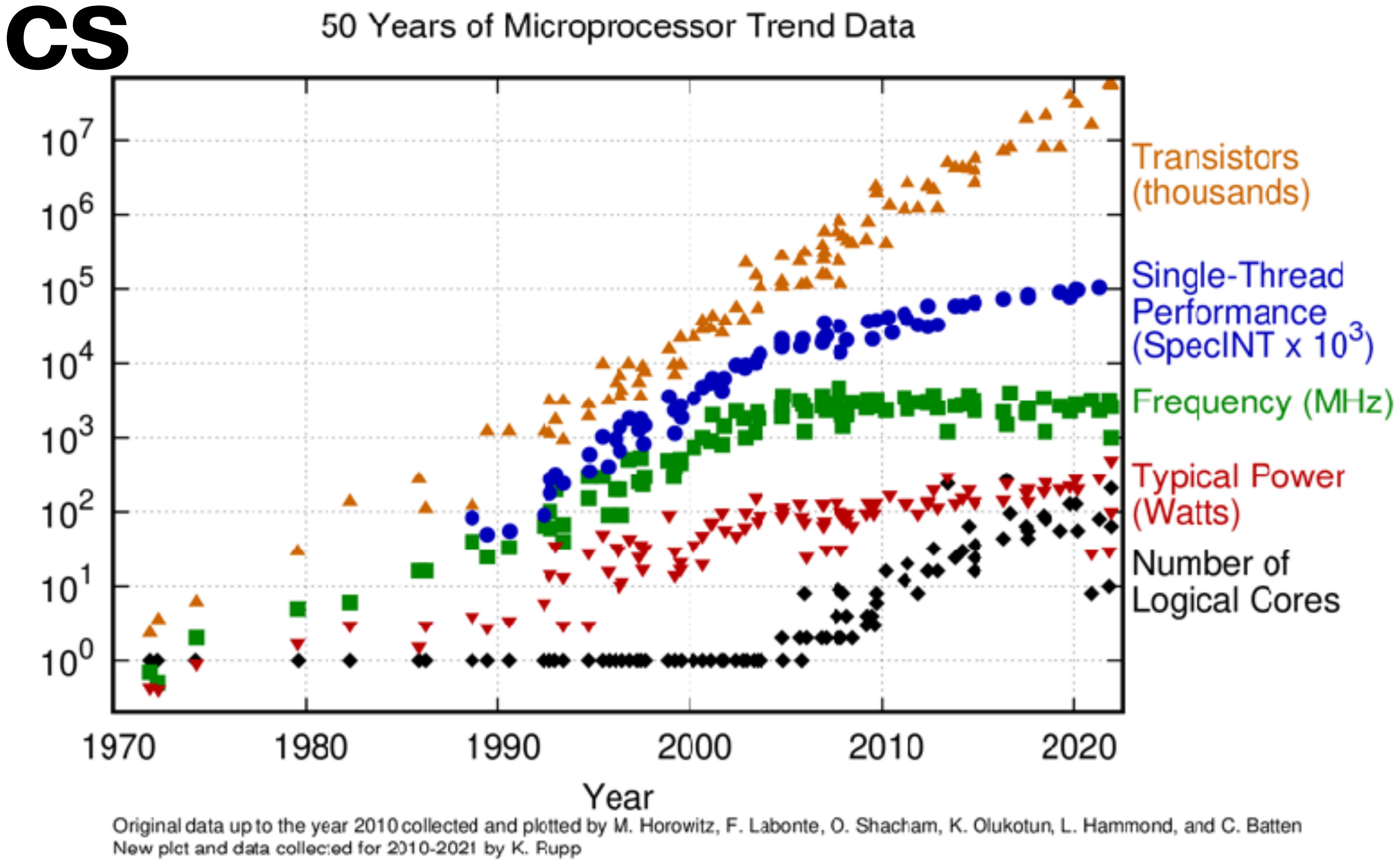
Parton distributions confront LHC Run II data: a systematic quantitative assessment
A. Chiesa, M. Constantini, JCM, E. Nocera, T. Rabemananjara, J. Rojo, T. Sharma, M. Ubiali - in preparation

New hardware in High Energy Physics

The computational footprint of High Energy Physics is growing at an uncontrolled pace. Going forward, new algorithms, architectures and computing paradigms will be needed to bridge the ever-increasing gap between theory and experiment.







GPU computing

- Different computing paradigm: vectorization is key
- Existing theoretical frameworks can be leveraged
- **Gains are guaranteed**, but ultimately, **similar scaling as current paradigm**

Quantum computing

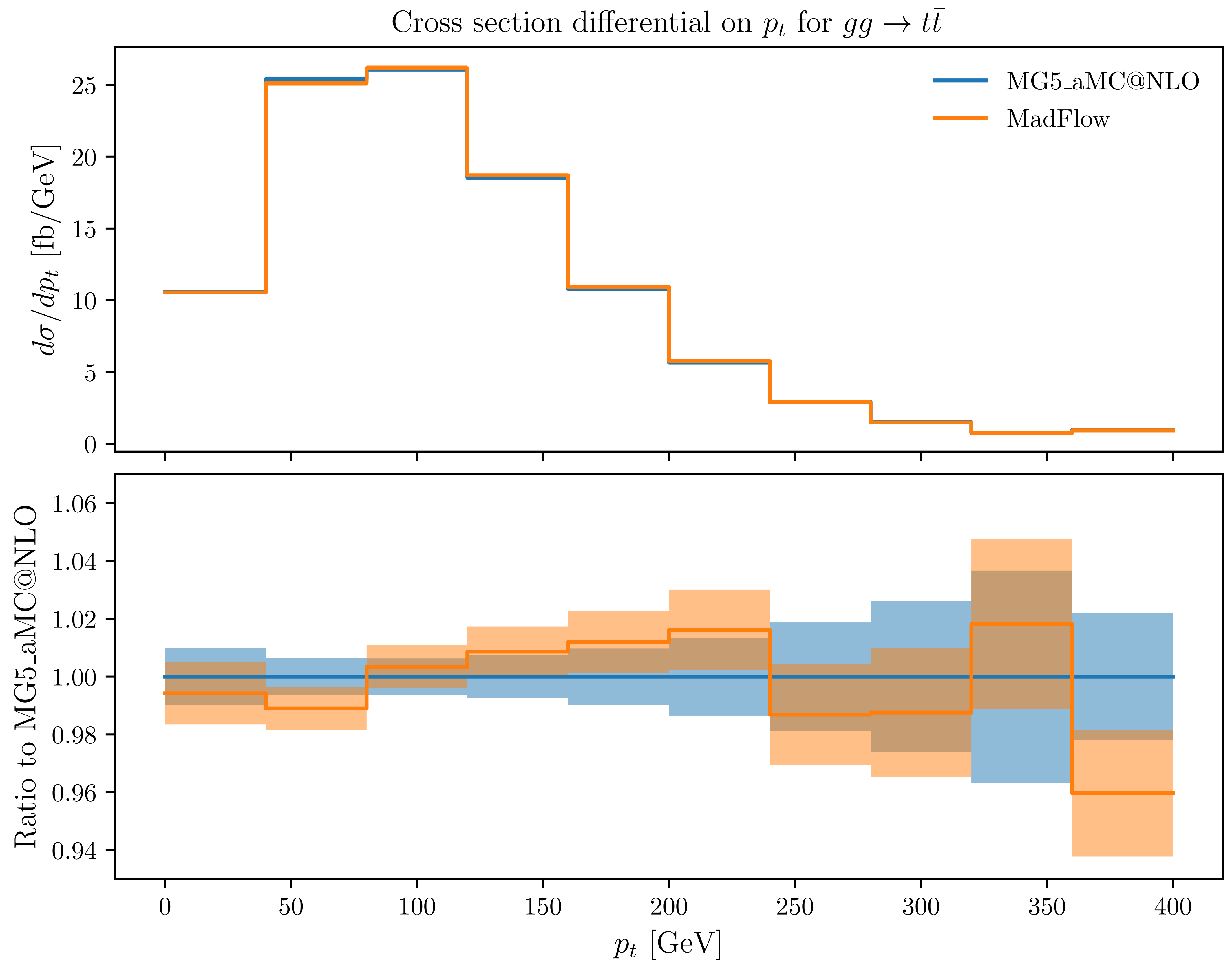
- Different computing and **theoretical paradigm**
- Need to develop completely new strategies
- **Gains potentially much greater,** but they require associated **algorithmic developments**

Access through collaborators J.I. Latorre and S. Carrazza to real hardware at TII Abu Dhabi.
Many opportunities for experimenting in a real-life scenario.

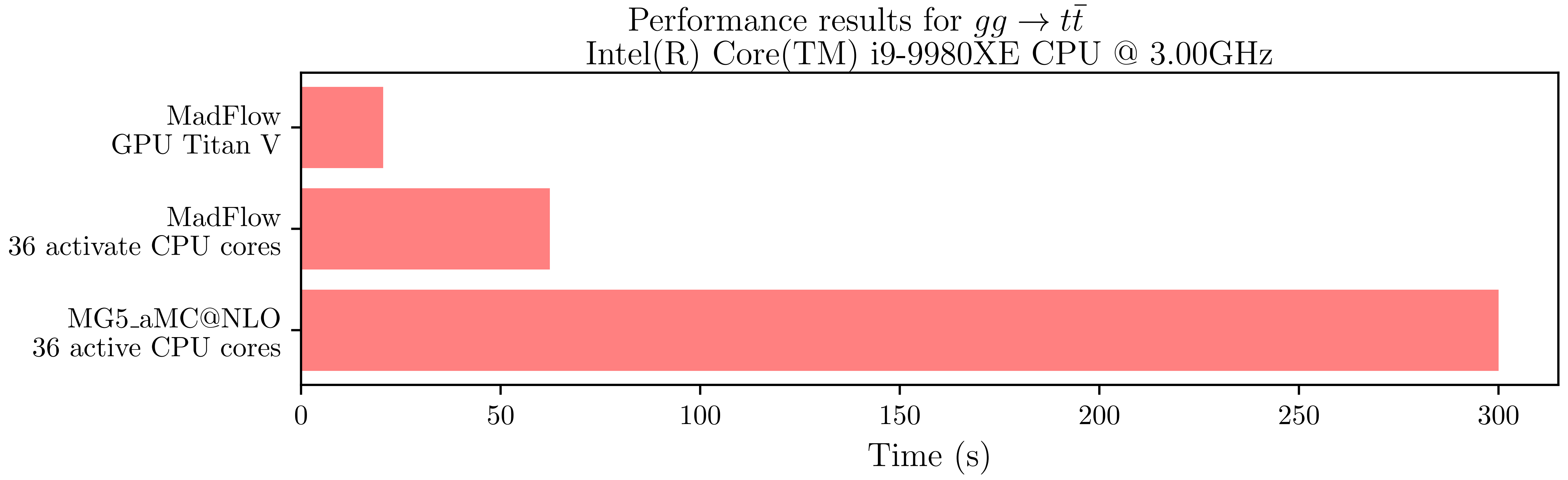
QIBO TII Technology Innovation Institute

Note: Qibo is also the chosen library by Qilimanjaro in the framework of the **Quantum Spain initiative.**

GPU-aware Monte Carlo integration



- Exploit MadGraph interface to automate diagram generation, extended to write them in a vectorized way and using GPU-friendly kernels.
- PDF interpolation using a *tensorized* version of LHAPDF
- Write a phase space generator that's completely general (vectorized version of Rambo).
- *Automagically* generate the matrix elements. Only at Leading Order.



VegasFlow: accelerating Monte Carlo simulation across multiple hardware platforms
S. Carrazza, **JCM** - [comp-ph] 2002.12921

PDFFlow: Parton distribution functions on GPU
S. Carrazza, **JCM**, M. Rossi - [hep-ph] 2009.06635

MadFlow: automating Monte Carlo simulation on GPU for particle physics processes
S. Carrazza, **JCM,** M. Rossi, M. Zaro - [comp-ph] 2002.12921

Accelerating Berends-Giele recursion for gluons in arbitrary dimension for finite fields
JCM, G. De Laurentis, M. Pellen - in preparation

Accelerated Fits using GPUs

PDF Fits are not good candidates for vectorization outside of the typical ML GPU-usage. However, PDF fits applications are!

- Simultaneous fit of multiple replicas:

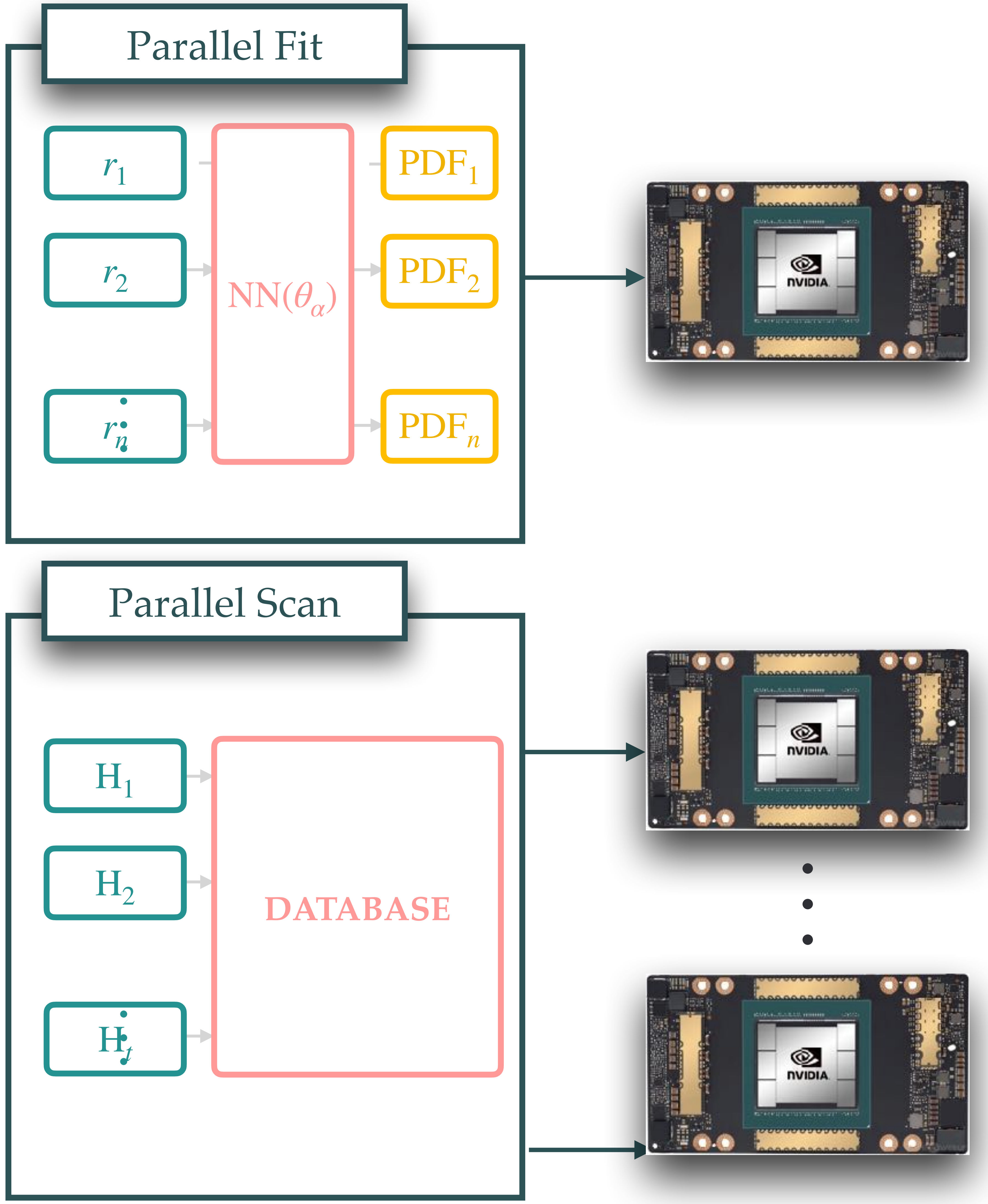
  Tensorflow allows the exact same codebase to be used for both CPU and GPU

  Redesign of the framework in order to share memory-heavy objects across all the replicas

$\Longrightarrow$ **Running hundreds of replicas at once on a GPU in the time it would take to run a single replica**

- **Distributed asynchronous scans**
  - Different fits can share a single database for a scan of parameters (e.g., simultaneously fit of PDF and W mass).
  - First example implementation for a hyperparameter scan.

Opens the door to systematic PDF (+ parameters) fits.



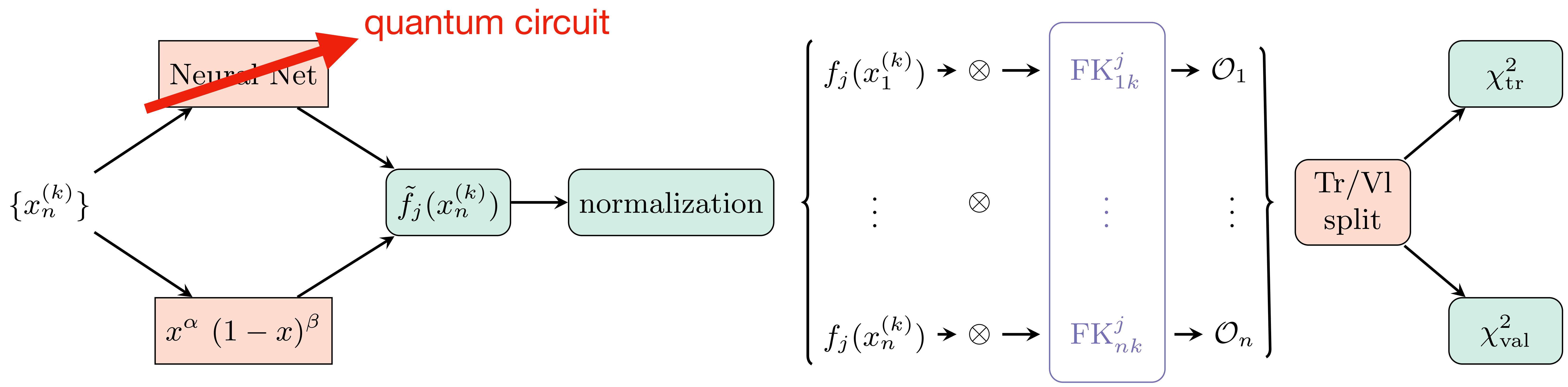
Hyperparameter Optimisation in Deep Learning Models with Ensemble Methods and Application to Proton Structure Fits
JCM, A. Jansen, G. van Oord, T. Rabemanjara, C. Rocha, J. Rojo, R. Stegeman - in preparation

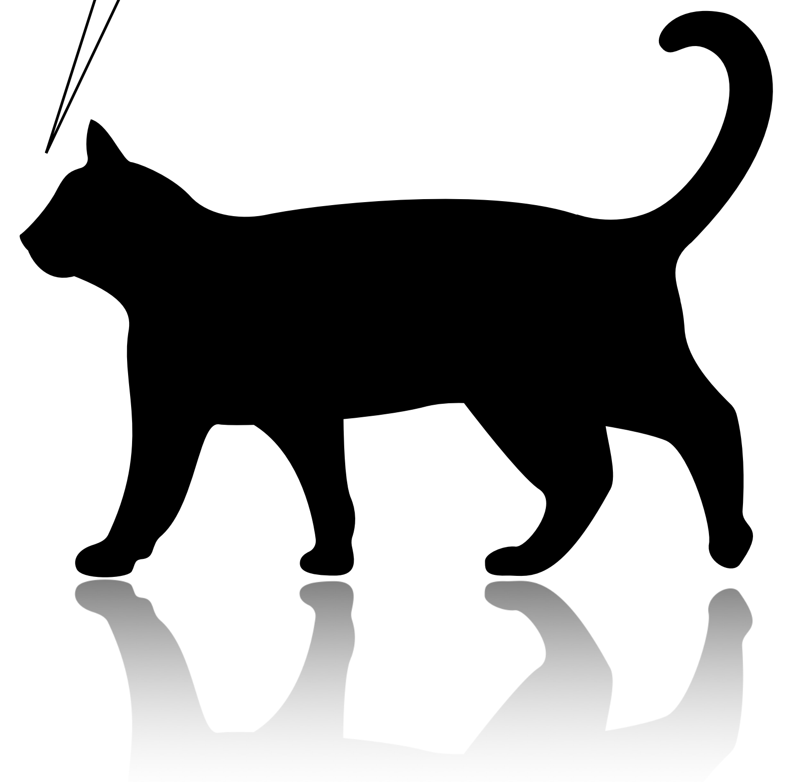
Beyond classical hardware

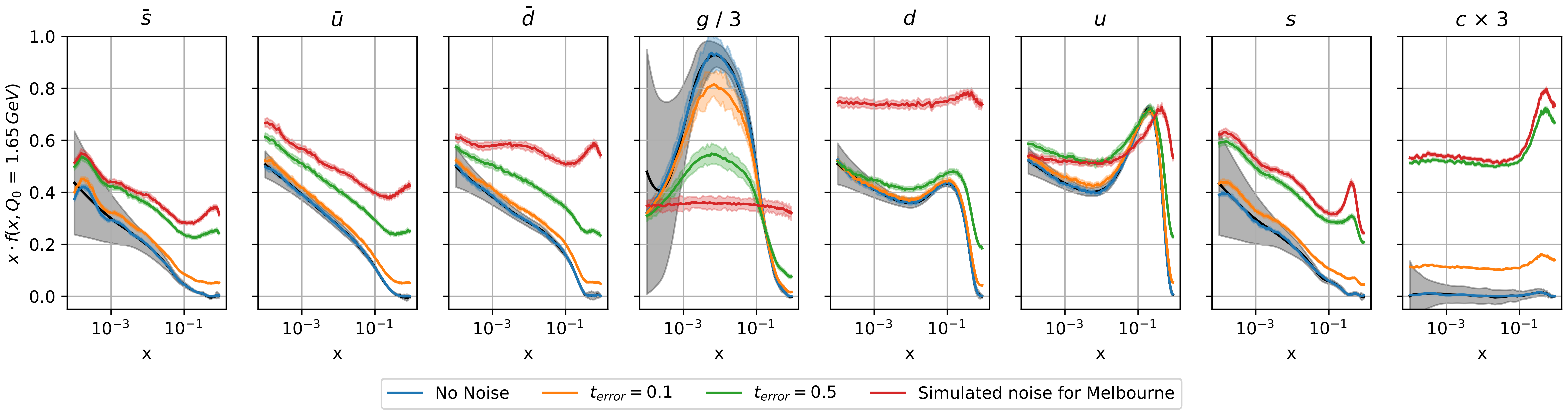
Quantum PDFs

Determining the proton content with a quantum computer
A. Perez-Salinas, **JCM**, A. Alhajri, S. Carrazza [hep-ph] 2011.13934

Quantum Machine Learning seems interesting. Can a Quantum Circuit be expressive enough for a PDF fit?





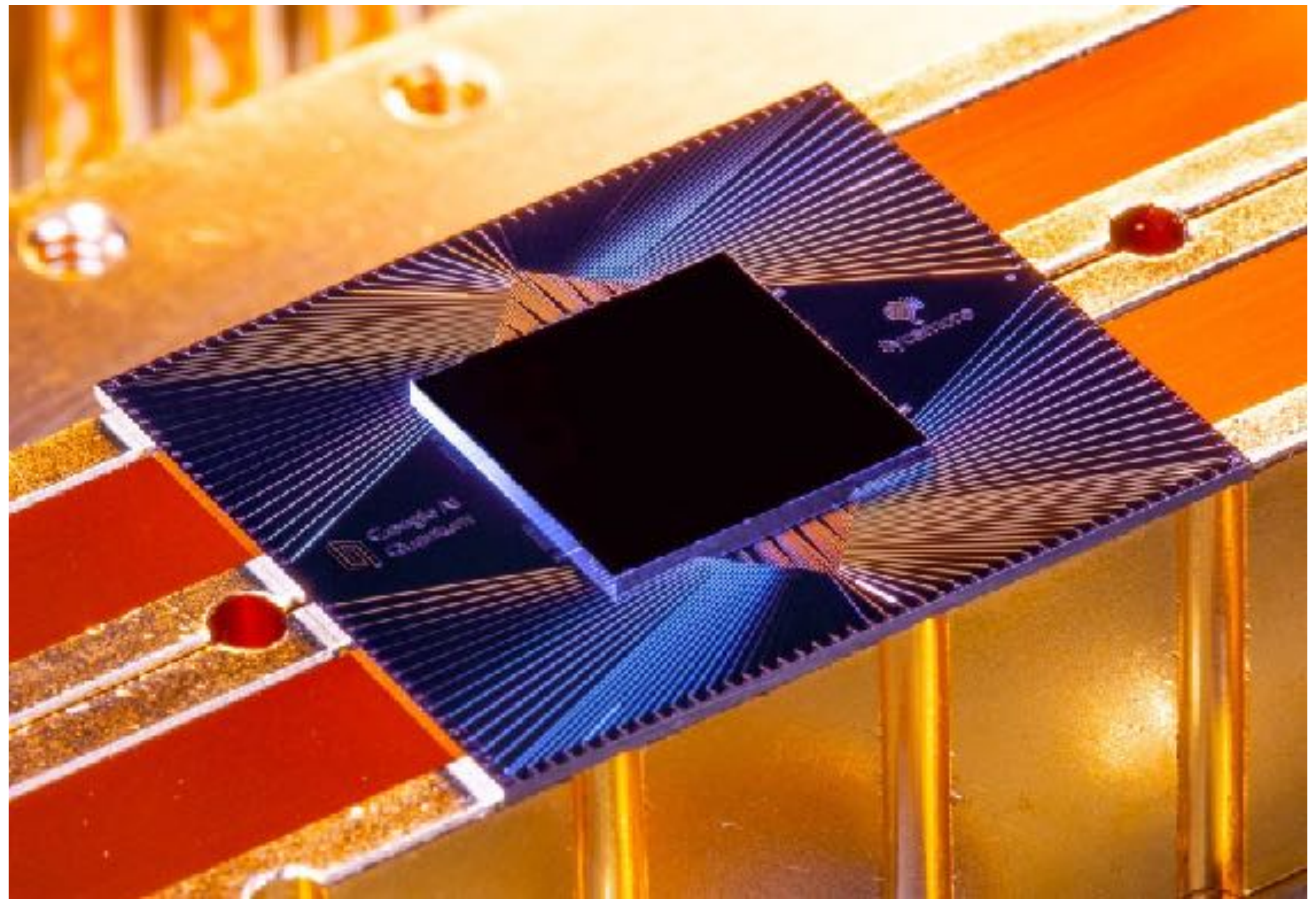


Beyond classical hardware

Quantum integration

Determining probability density functions with adiabatic quantum computing
M. Robbiati, **JCM**, S. Carrazza, [quant-ph] 2303.11346

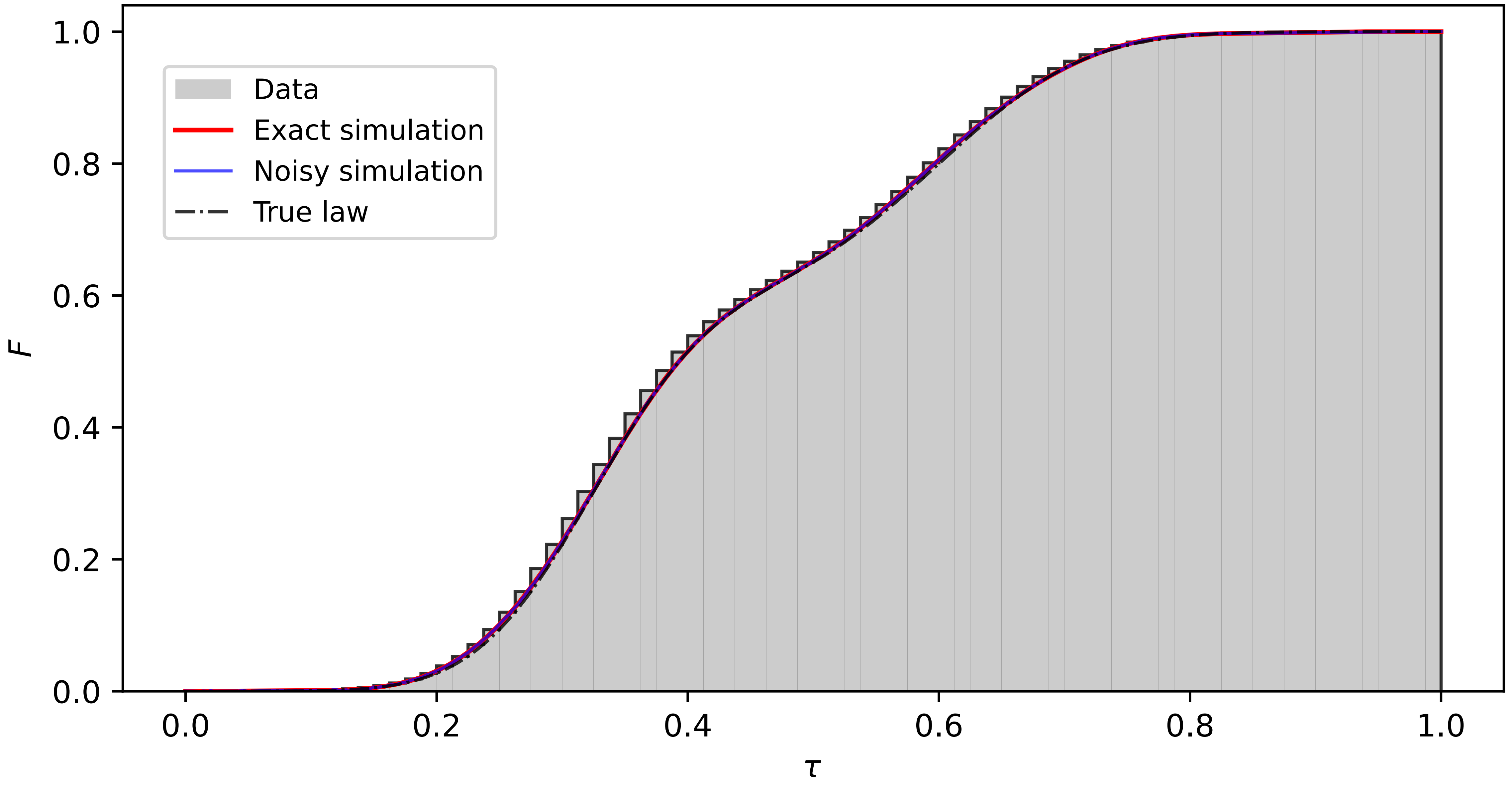
QiNNtegrate: Multi-variable integration with a variational quantum circuit
JCM, M. Robbiati, S. Carrazza [quant-ph] 2308.05657



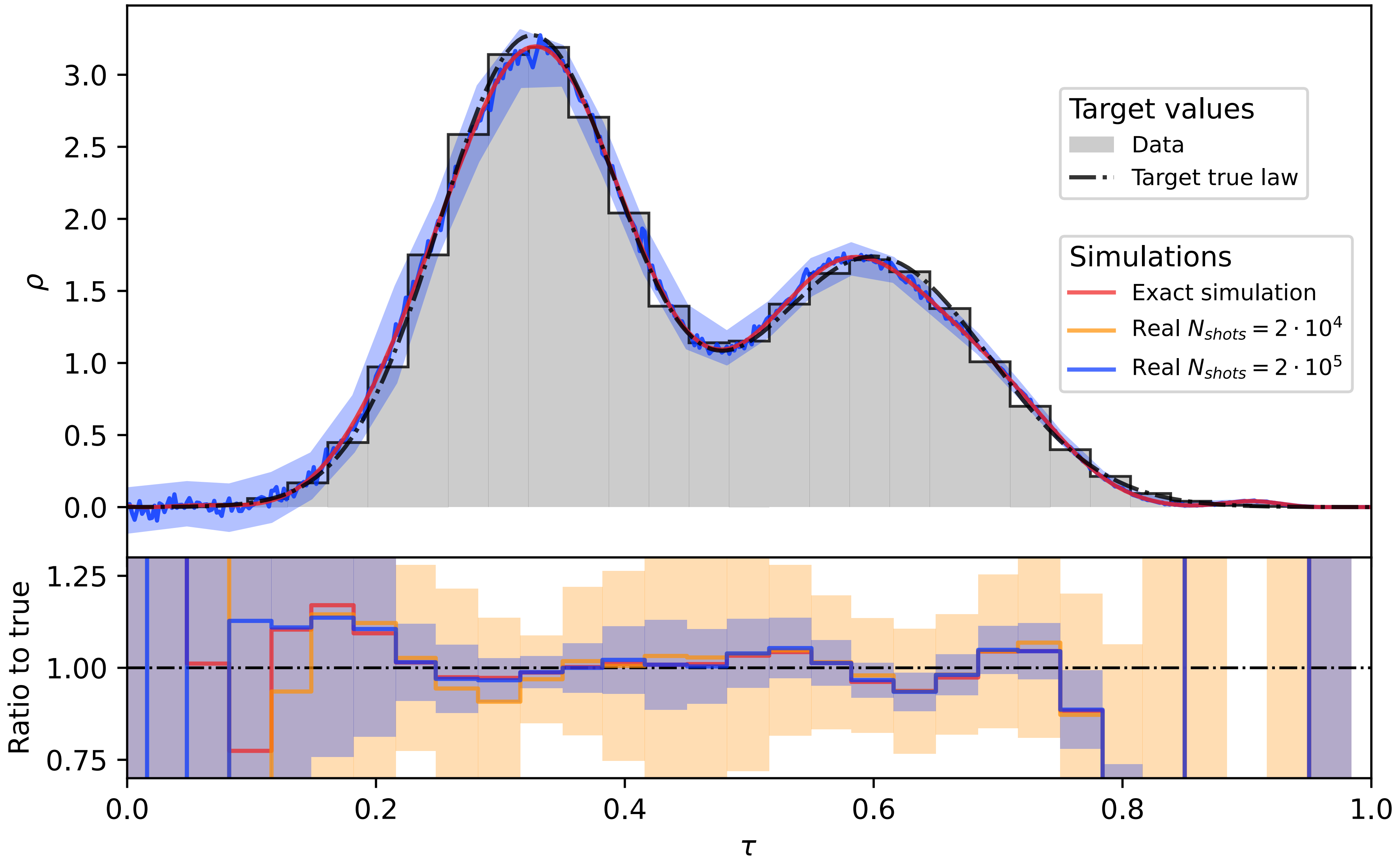
Exploiting the Parameter Shift Rule we can trivially move between an observable and its derivative:

$$\partial_\mu F = r\left[F(\mu^+) - F(\mu^-)\right]$$

CDF estimation - $\rho(x) = 0.6\mathcal{N}(x; -10, 4) + 0.4\mathcal{N}(x; 5, 5)$



PDF estimation - $\rho(x) = 0.6\mathcal{N}(x; -10, 4) + 0.4\mathcal{N}(x; 5, 5)$

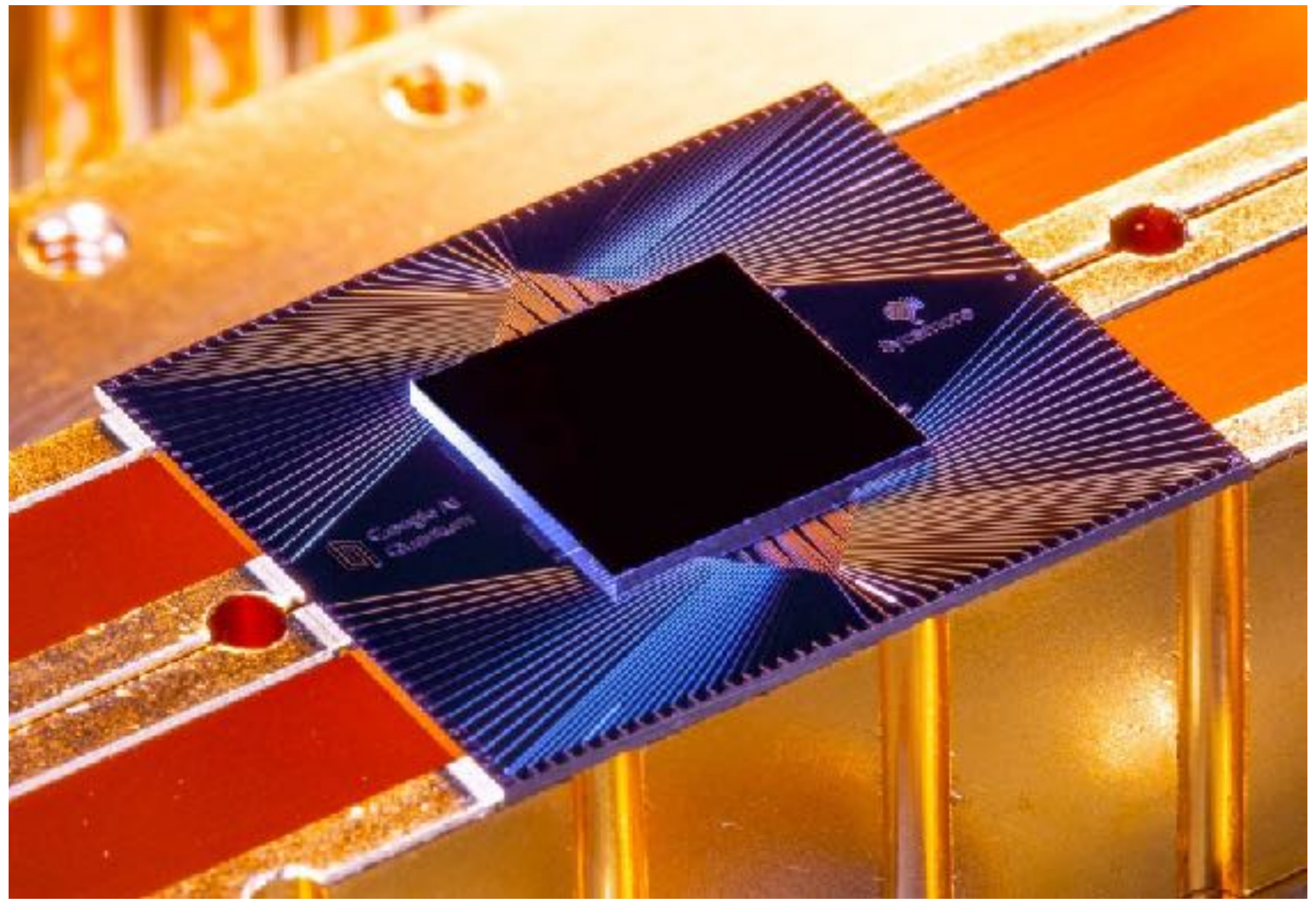


Beyond classical hardware

Quantum integration

Determining probability density functions with adiabatic quantum computing
M. Robbiati, **JCM**, S. Carrazza, [quant-ph] 2303.11346

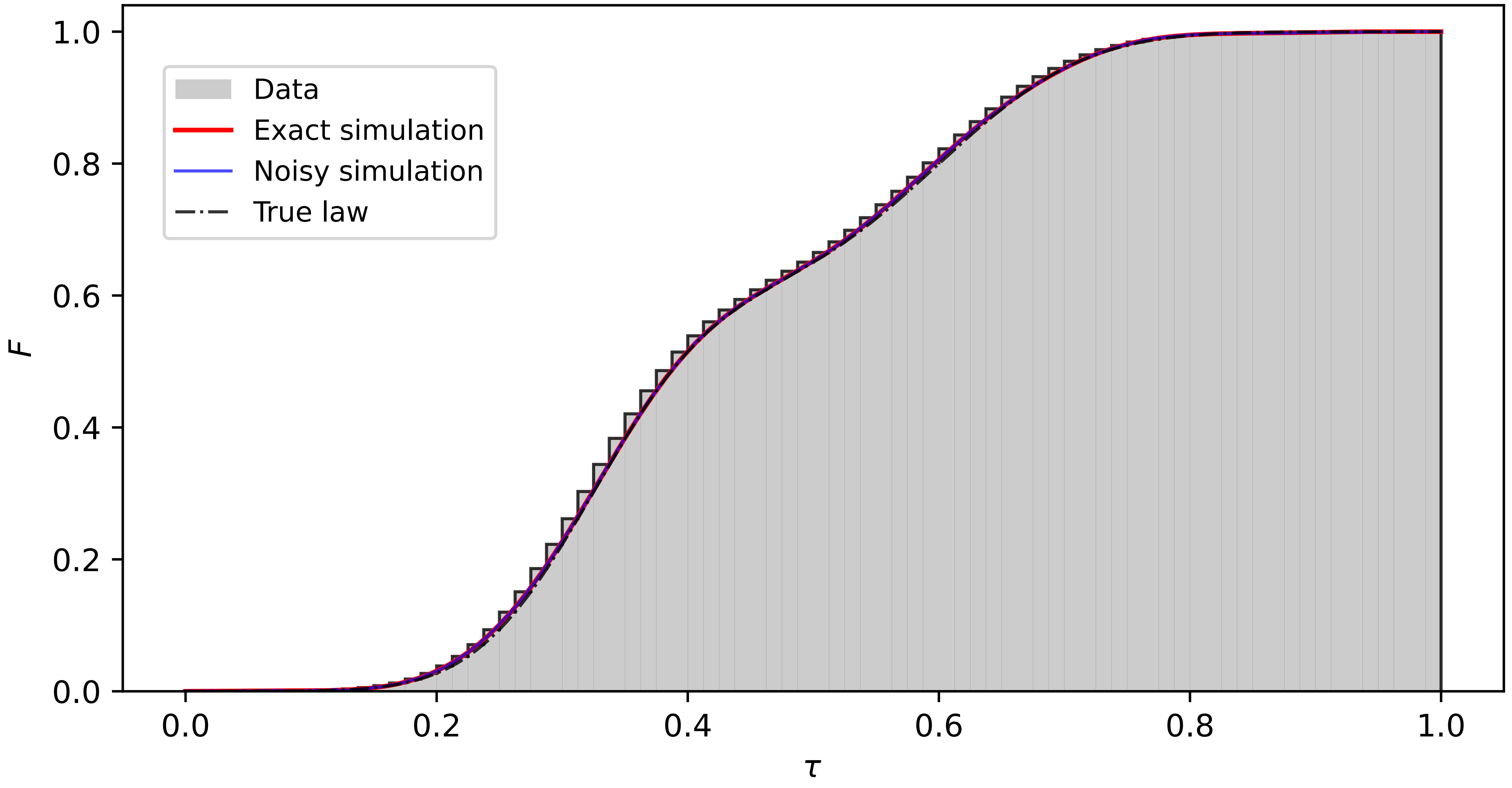
QiNNtegrate: Multi-variable integration with a variational quantum circuit
JCM, M. Robbiati, S. Carrazza [quant-ph] 2308.05657



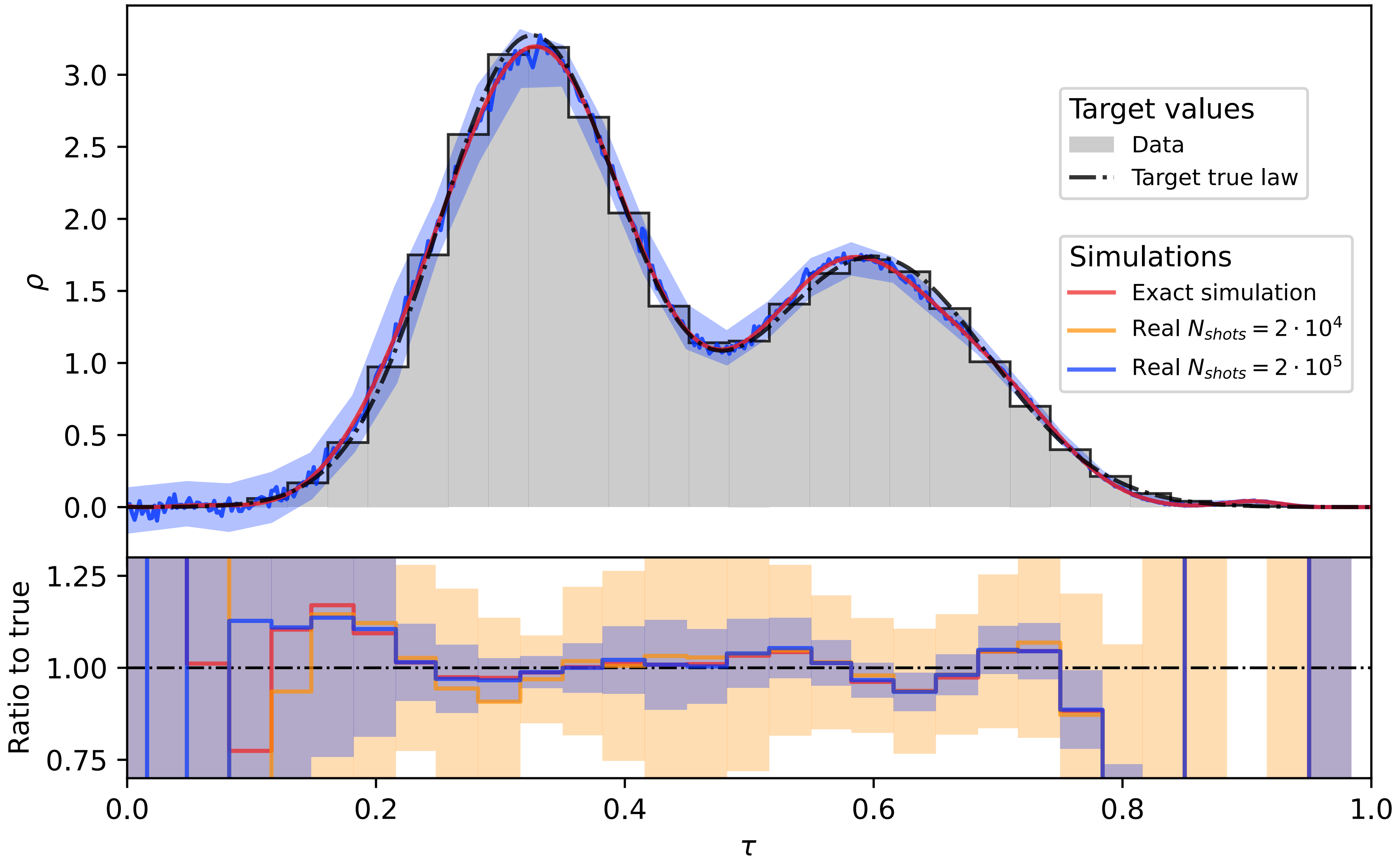
Exploiting the Parameter Shift Rule we can trivially move between an observable and its derivative:

$$\partial_\mu F = r\left[F(\mu^+) - F(\mu^-)\right]$$

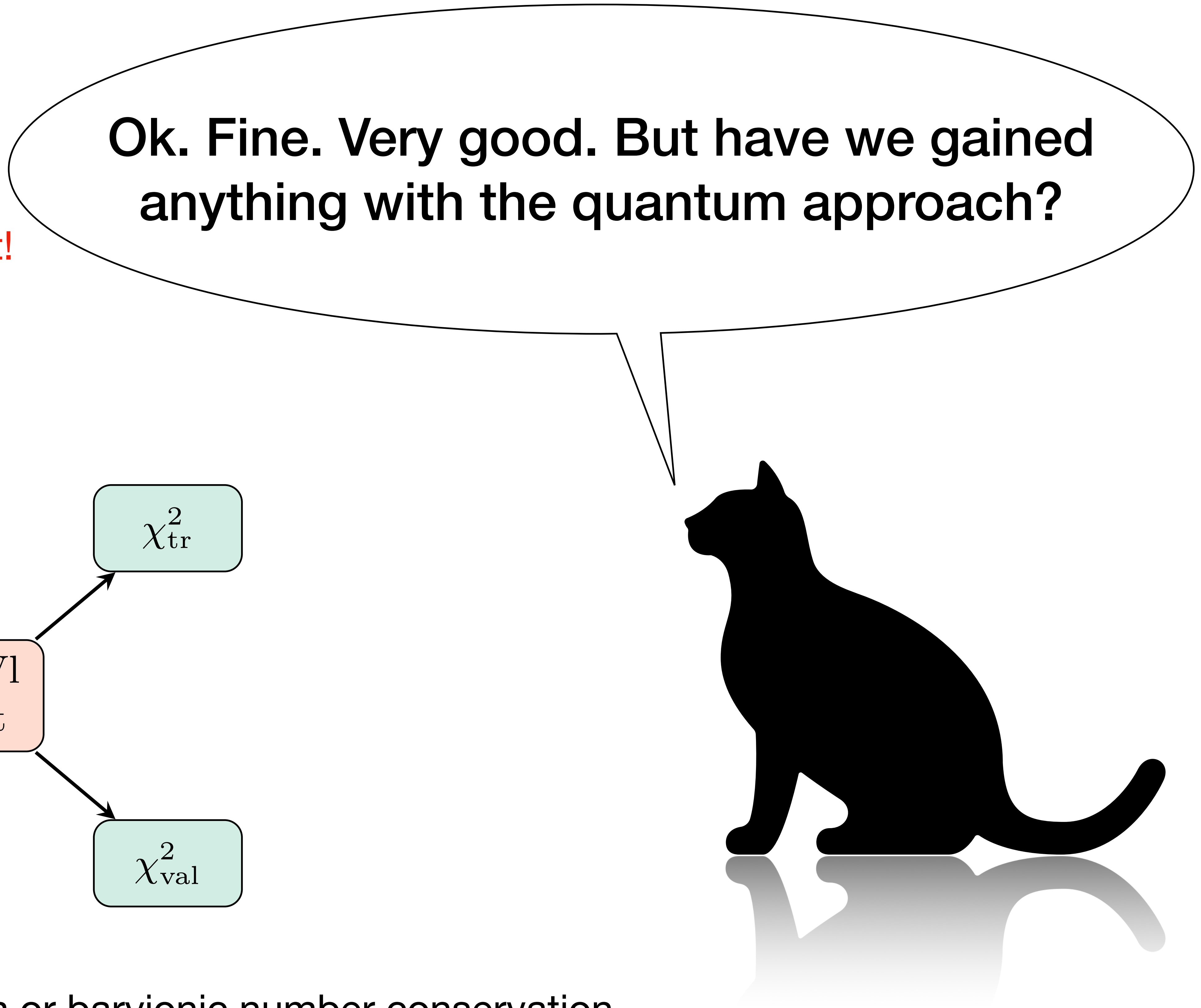
CDF estimation - $\rho(x) = 0.6\mathcal{N}(x; -10, 4) + 0.4\mathcal{N}(x; 5, 5)$



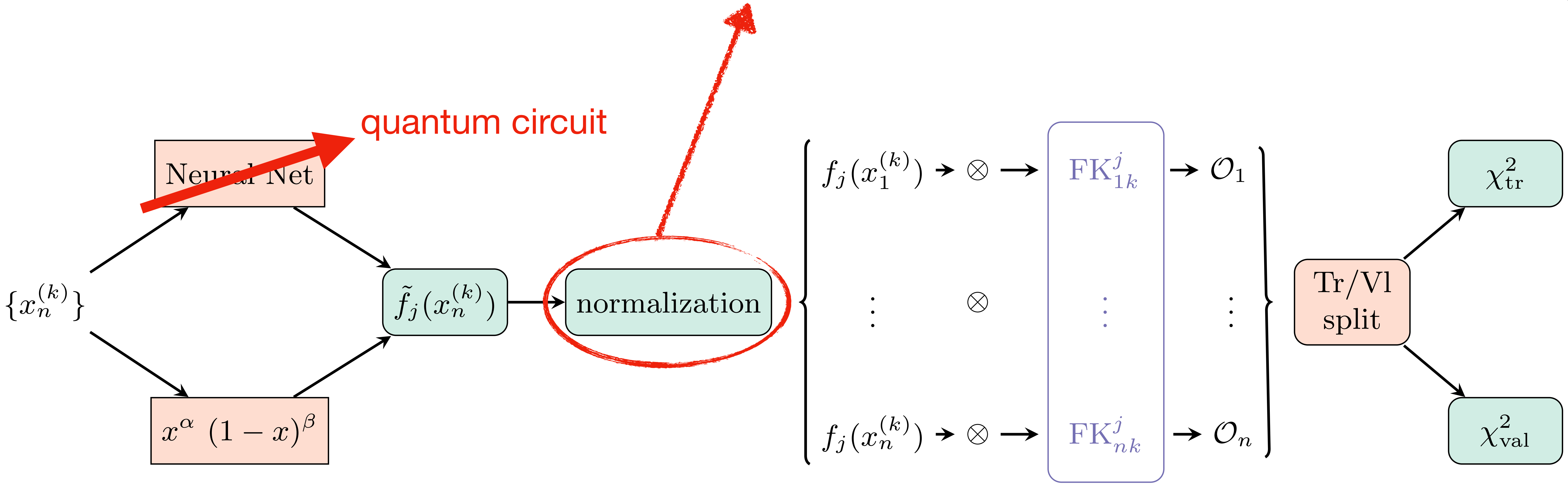
PDF estimation - $\rho(x) = 0.6\mathcal{N}(x; -10, 4) + 0.4\mathcal{N}(x; 5, 5)$



Back to the PDF fit



The normalization in this diagram is actually an integral of the NN over the input!



The normalization of the PDF ensures that the sum rules are satisfied: e.g., momentum or baryionic number conservation.

This requires an integral of the unnormalized PDF. In other words: $f_j(x) \approx \dfrac{\hat{f}_j(x)}{\int_0^1 dx \hat{f}_j(x)}$ which during a fit is done numerically for every step of the training (i..e, thousands of extra evaluations).

Instead, with the parameter shift rule, the circuit becomes the integral of the PDF while we fit the derivative (i.e, the shifted circuit):

$$f_j \approx \sum_{layers} \frac{C(x^+) - C(x^-)}{C(1) - C(0)}$$

Conclusions

Amazing developments of the last few years

- Percent-level uncertainties in the data region
- Estimation of theory uncertainties up to NNLO
- Approximated N3LO PDFs
- First evidences of an intrinsic charm component
- Hardware acceleration can be used for systematic and simultaneous parameter scans
- Quantum Computing as a tool to enhance HEP

Amazing developments of the years to come

- Exact NNLO (and beyond) QCD PDFs
  - Improved N3LO approximation -> exact?
  - Exact predictions up to NNLO. N3LO k-factors
  - EW corrections to hard coefficients
  - Polarized PDFs, Nuclear PDFs
  - TMDs?
- Searches of new physics with PDFs
  - Is the charm non-perturbative?
  - Simultaneous fits of PDFs with SM parameters
- New hardware opens new horizons
  - Can quantum computers gives us short-cuts in pQCD?
  - Hardware accelerators as a way of reducing the frictions to incorporate new theoretical insights

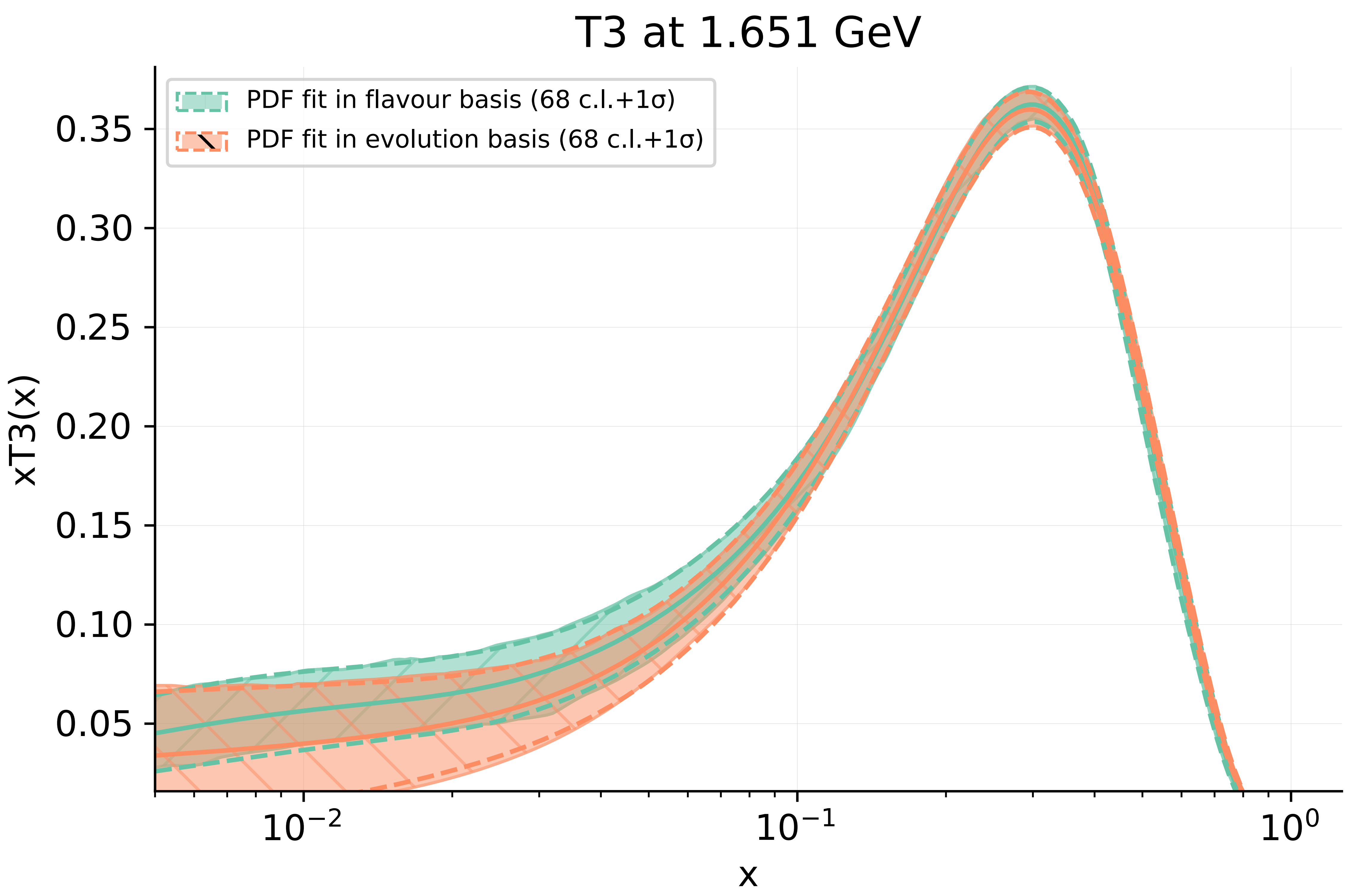
Thank you!

Backup

PDF Parametrization

To ease the fit, the output is parametrized by default in the "evolution" basis at the input scale Q_0

$$
\begin{aligned}
g &= g, \\
\Sigma &= u + \bar{u} + d + \bar{d} + s + \bar{s} + 2c\,, \\
T_3 &= (u + \bar{u}) - \left(d + \bar{d}\right)\,, \\
T_8 &= \left(u + \bar{u} + d + \bar{d}\right) - 2\,(s + \bar{s})\,, \\
T_{15} &= \left(u + \bar{u} + d + \bar{d} + s + \bar{s}\right) - 3\,(c + \bar{c})\,, \\
V &= (u - \bar{u}) + \left(d - \bar{d}\right) + (s - \bar{s})\,, \\
V_3 &= (u - \bar{u}) - \left(d - \bar{d}\right)\,, \\
V_8 &= \left(u - \bar{u} + d - \bar{d}\right) - 2\,(s - \bar{s})\,.
\end{aligned}
$$



Different parametrization achieve similar results: basis independence

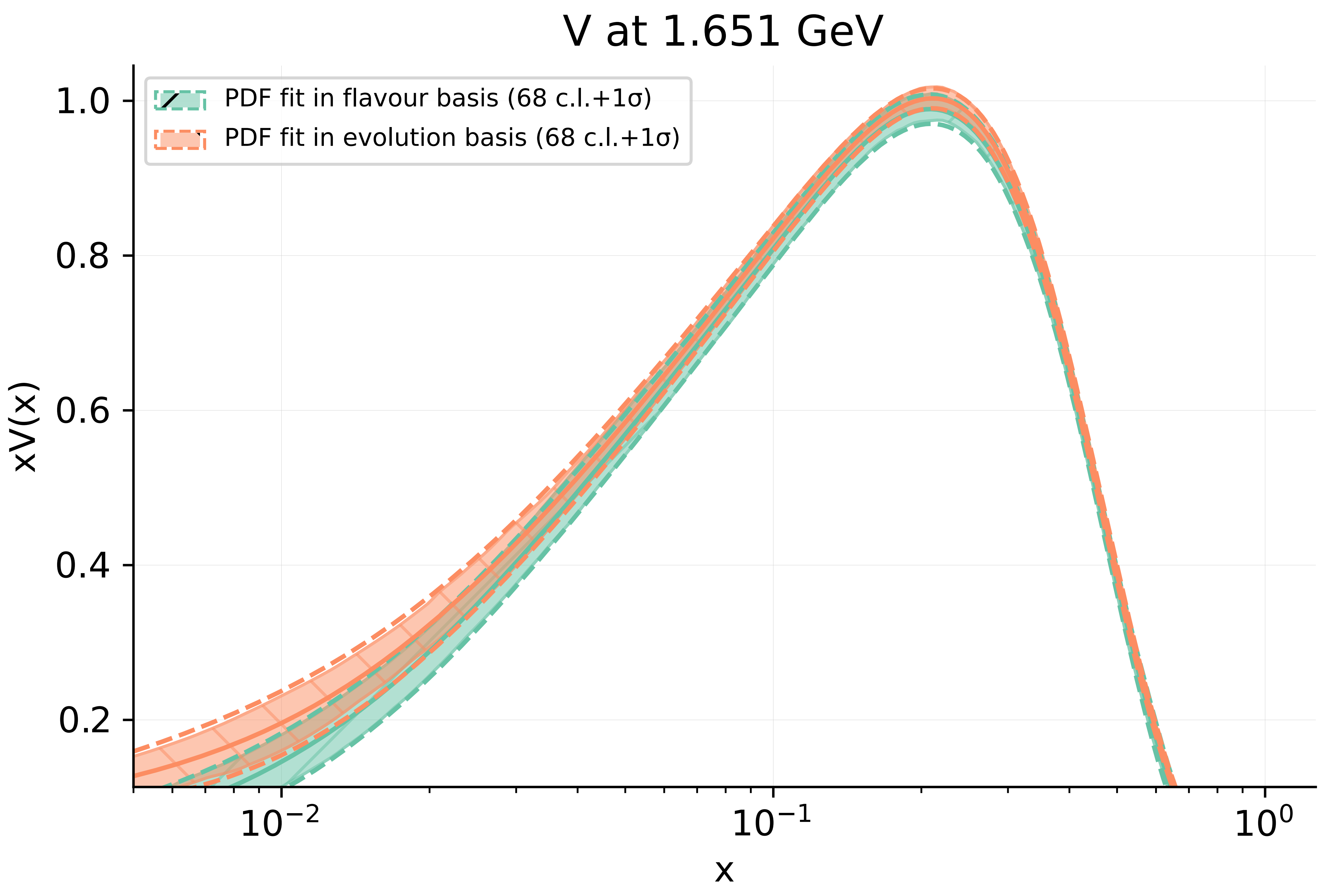
$$xV\left(x, Q_0\right) \propto \mathrm{NN}_V(x)$$

$$xV\left(x, Q_0\right) \propto \left(\mathrm{NN}_u(x) - \mathrm{NN}_{\bar{u}}(x) + \mathrm{NN}_d(x) - \mathrm{NN}_{\bar{d}}(x) + \mathrm{NN}_s(x) - \mathrm{NN}_{\bar{s}}(x)\right)$$

PDF Parametrization

To ease the fit, the output is parametrized by default in the “evolution” basis at the input scale Q_0

$$
\begin{aligned}
g &= g, \\
\Sigma &= u + \bar{u} + d + \bar{d} + s + \bar{s} + 2c\,, \\
T_3 &= (u + \bar{u}) - (d + \bar{d})\ , \\
T_8 &= (u + \bar{u} + d + \bar{d}) - 2\,(s + \bar{s})\ , \\
T_{15} &= (u + \bar{u} + d + \bar{d} + s + \bar{s}) - 3\,(c + \bar{c})\ , \\
V &= (u - \bar{u}) + (d - \bar{d}) + (s - \bar{s})\ , \\
V_3 &= (u - \bar{u}) - (d - \bar{d})\ , \\
V_8 &= (u - \bar{u} + d - \bar{d}) - 2\,(s - \bar{s})\ .
\end{aligned}
$$

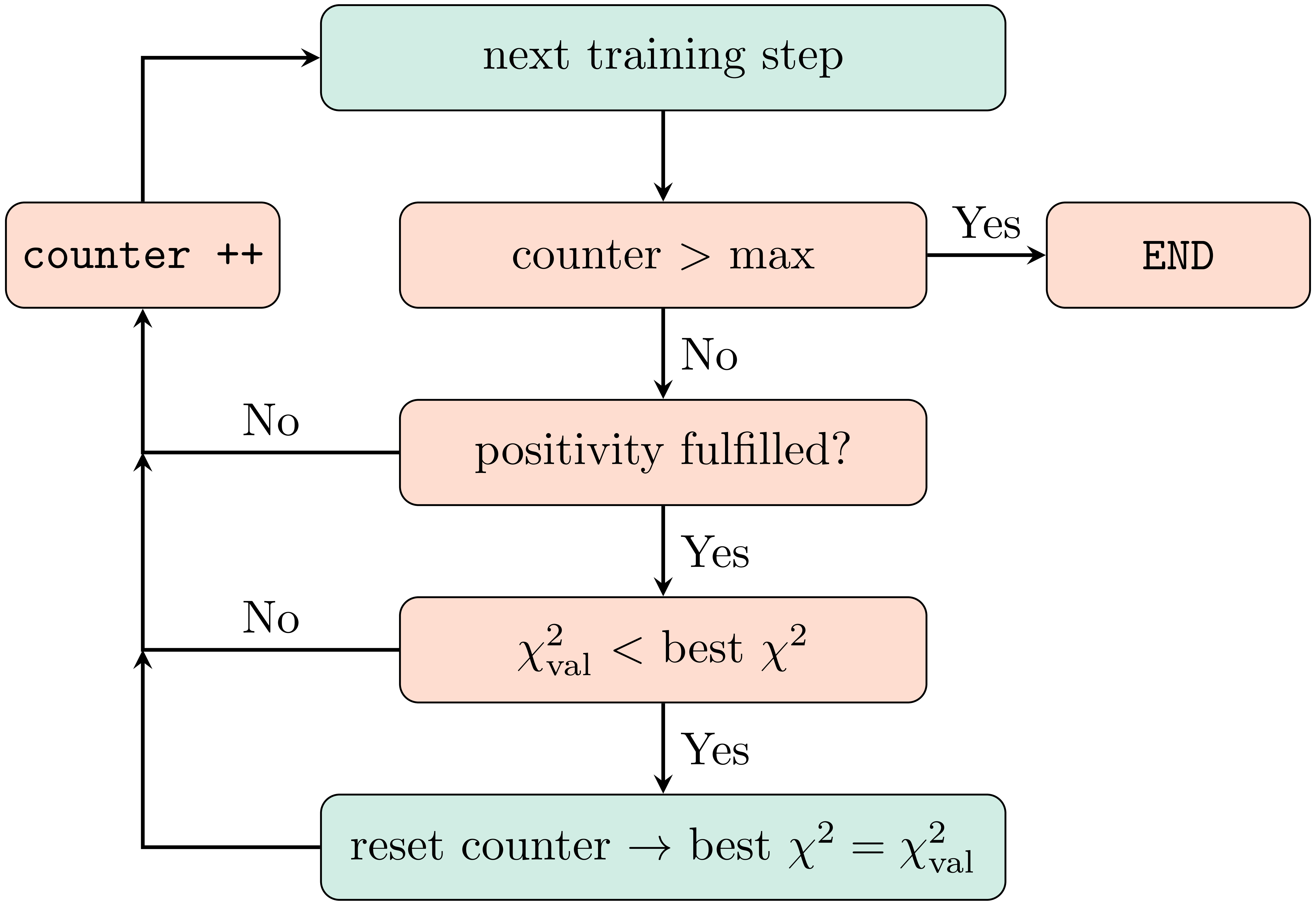


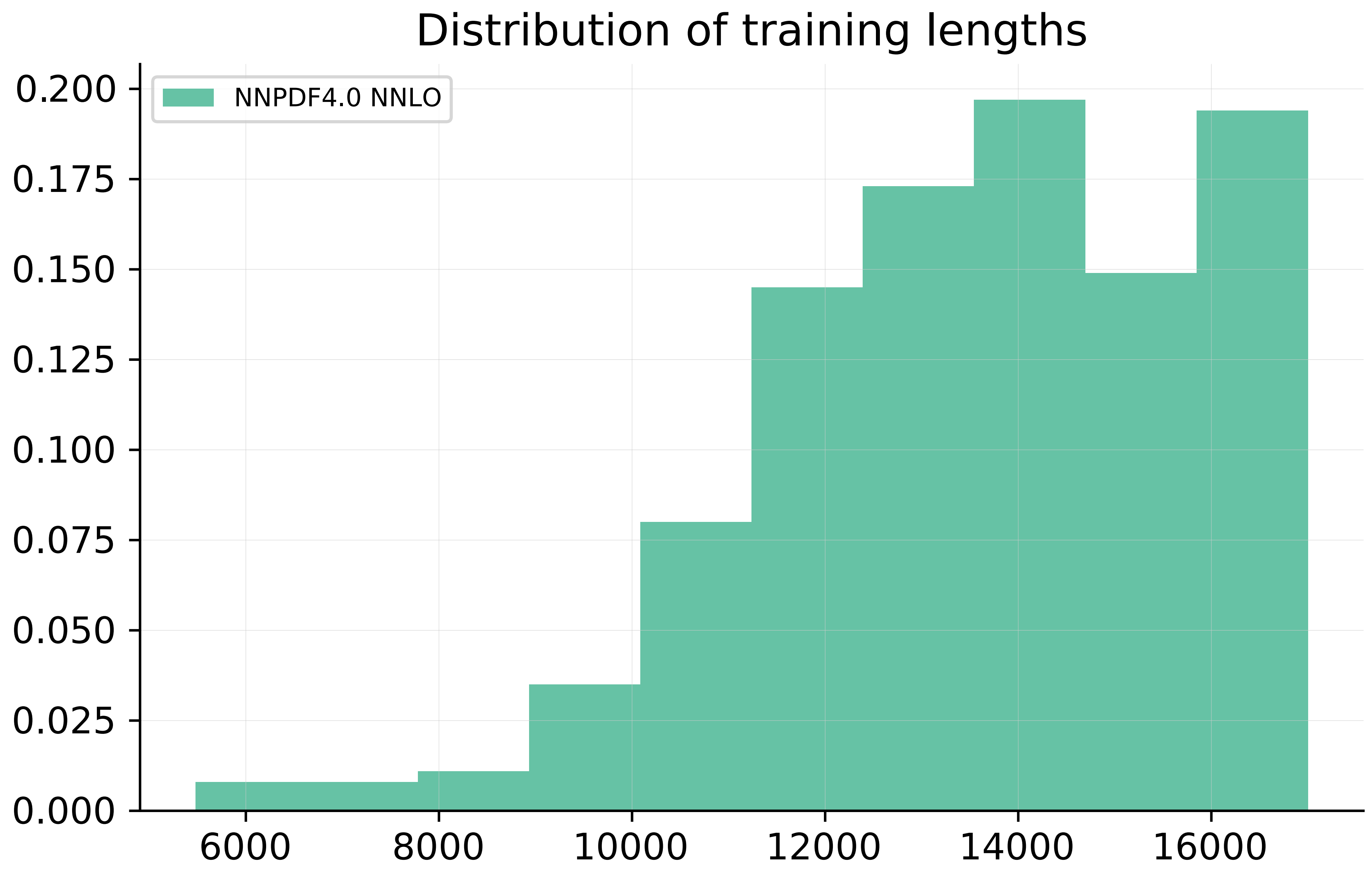
Different parametrization achieve similar results: basis independence

$$xV\left(x, Q_0\right) \propto \mathrm{NN}_V(x)$$

$$xV\left(x, Q_0\right) \propto \left(\mathrm{NN}_u(x) - \mathrm{NN}_{\bar{u}}(x) + \mathrm{NN}_d(x) - \mathrm{NN}_{\bar{d}}(x) + \mathrm{NN}_s(x) - \mathrm{NN}_{\bar{s}}(x)\right)$$

Stopping algorithm

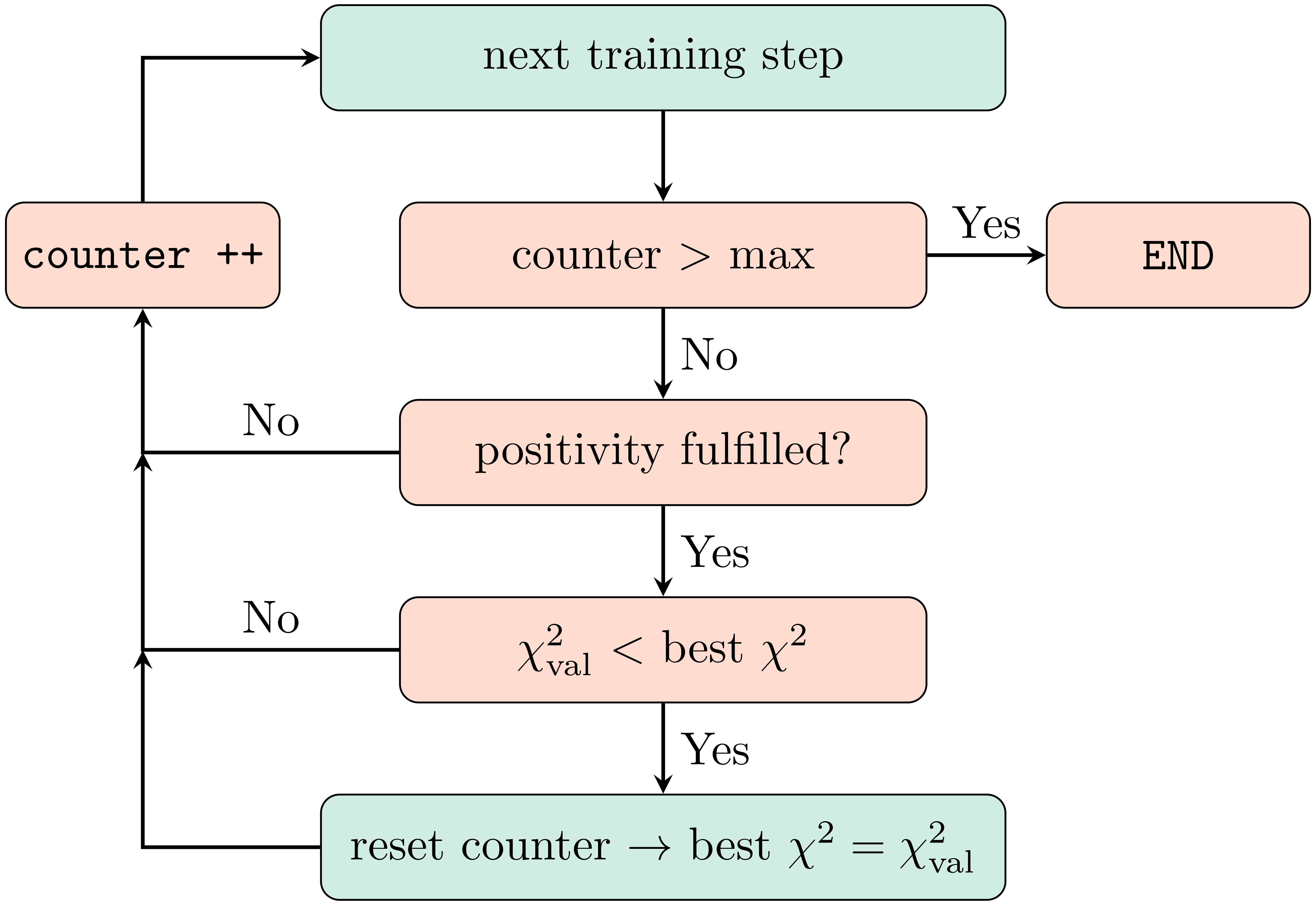


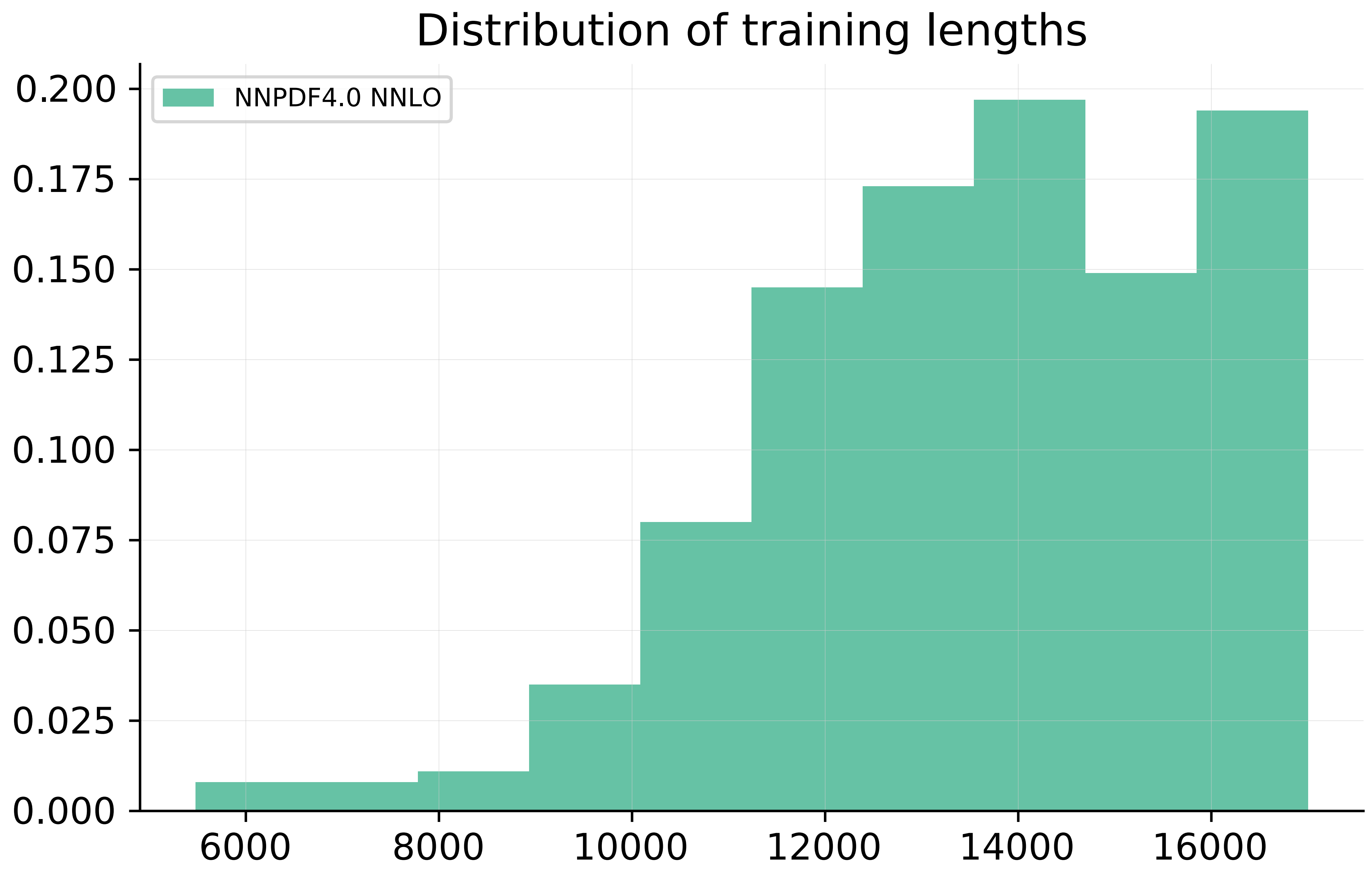


Regardless of the training algorithm or frameworks used the fitting method consist on

1. Reducing the loss function
2. Check the constraints are fulfilled
3. Continue until the validation metric stops improving.

Stopping algorithm



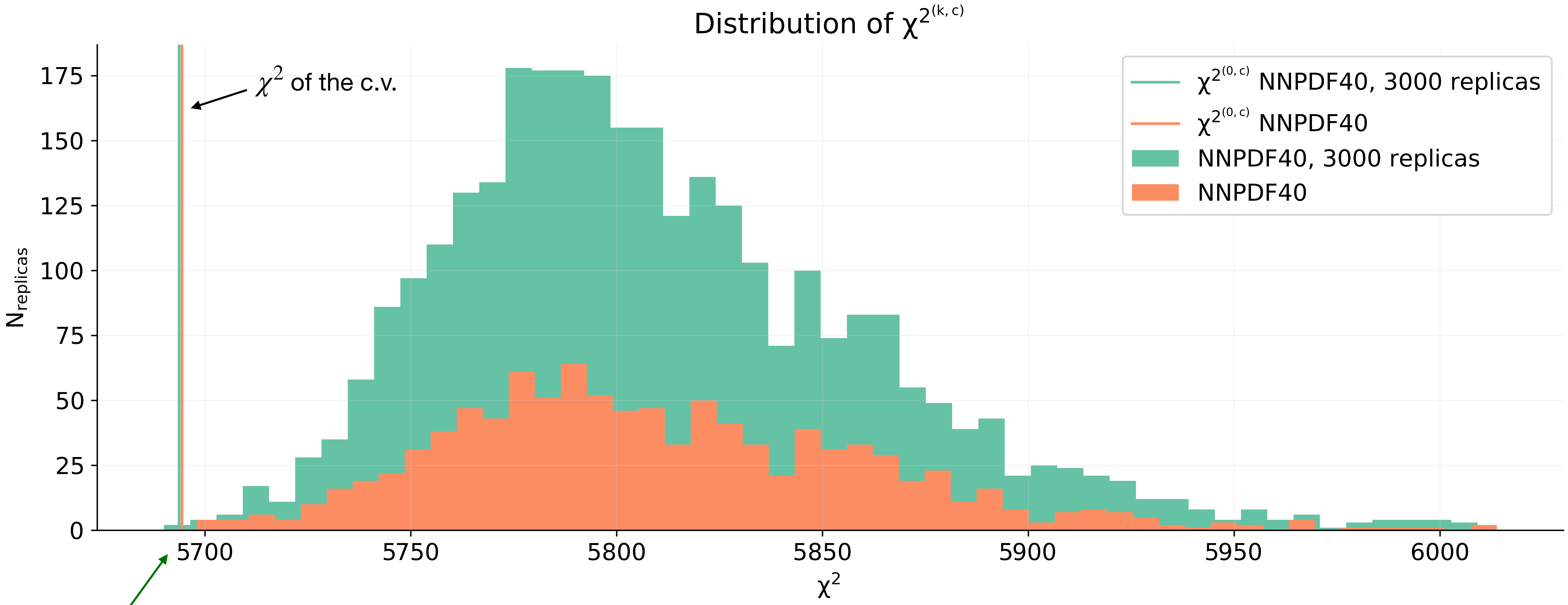


Regardless of the training algorithm or frameworks used the fitting method consist on

1. Reducing the loss function
2. Check the constraints are fulfilled
3. Continue until the validation metric stops improving.

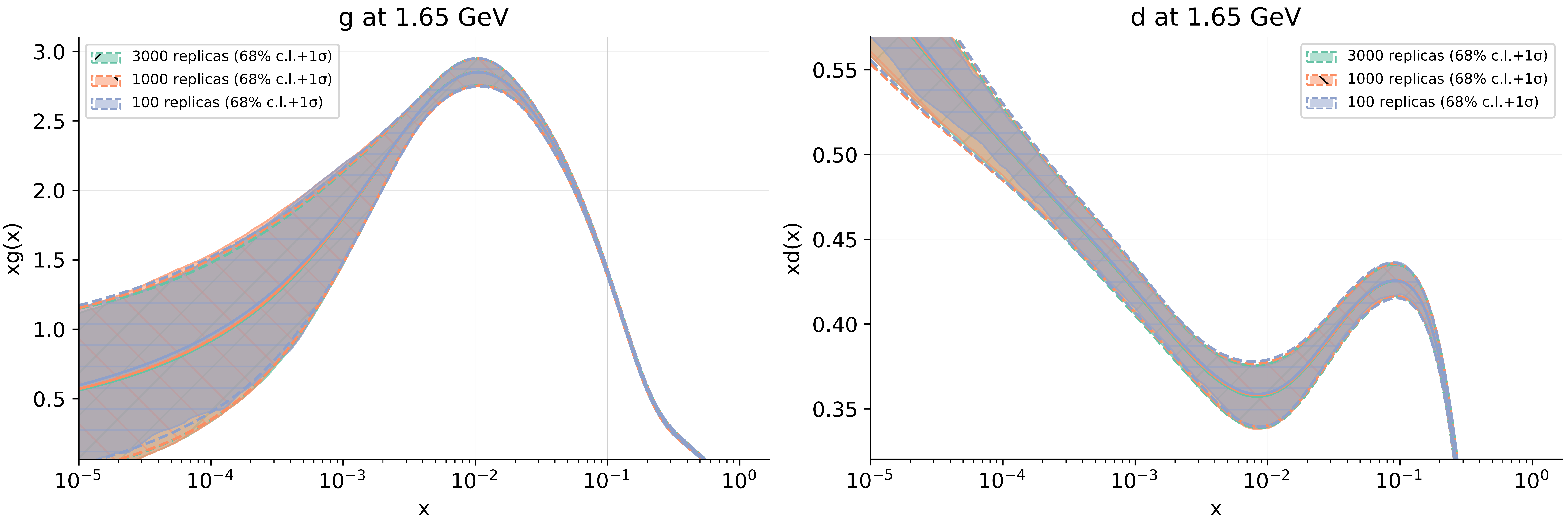
χ^2_0 distribution of the replicas

A common misconception is that the central PDF corresponds to the best fit to the data.



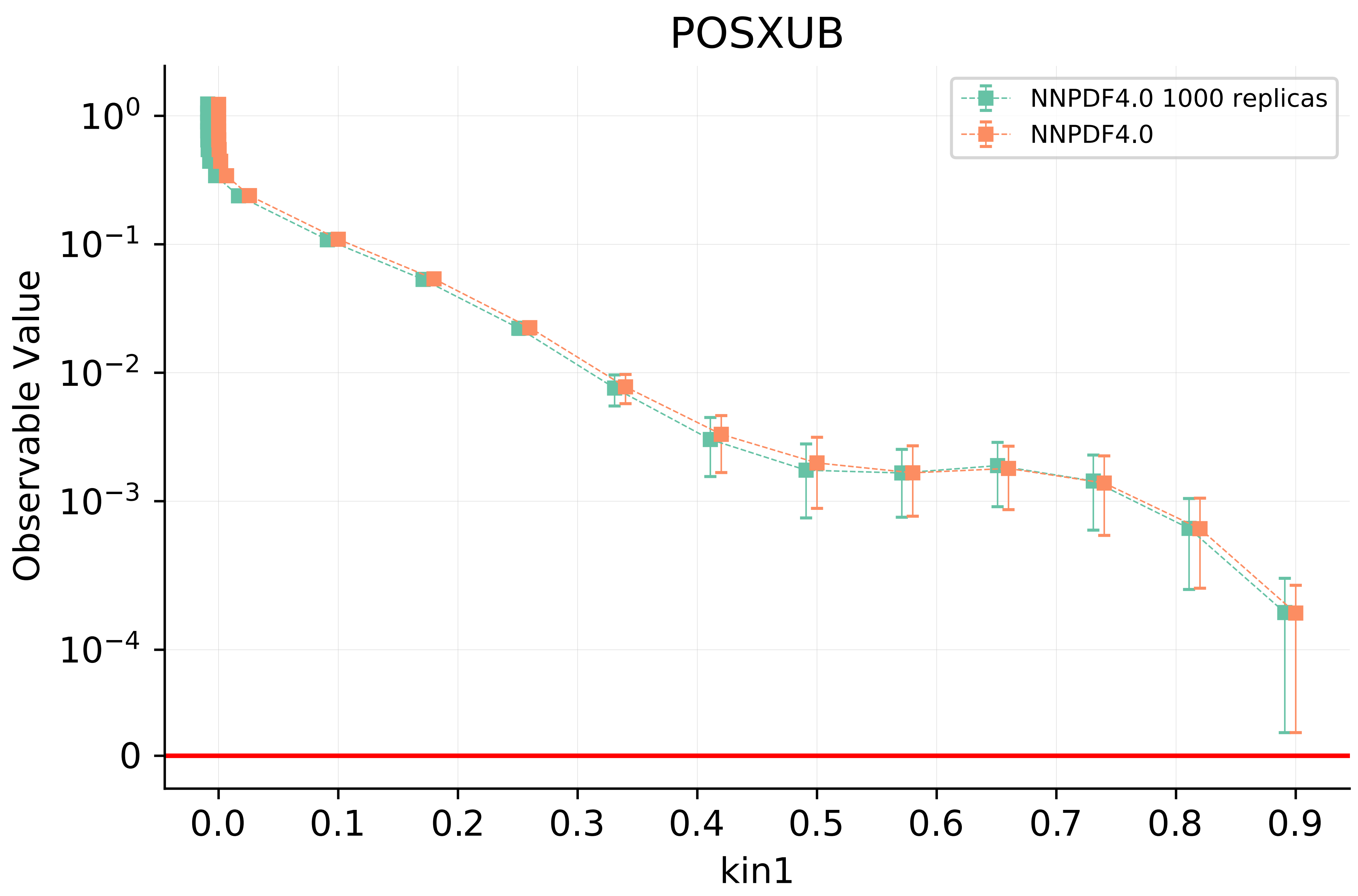
The χ^2 of the CV is not necessarily the minimum χ^2

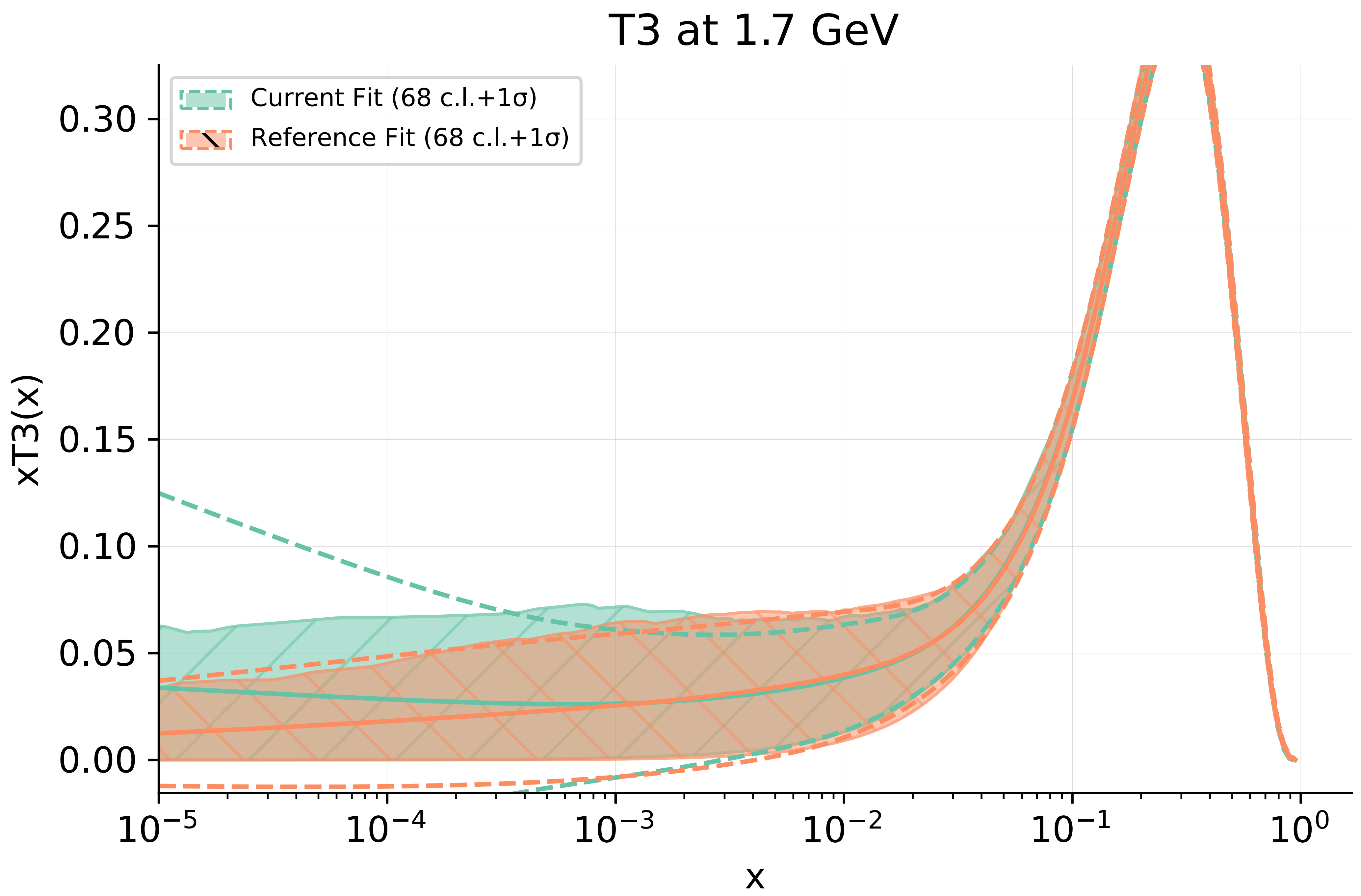
How do the PDFs change with the number of replicas?



The loss function

Positivity and integrability





$$\mathscr{L} = \chi_0^2 + \lambda_{pos}\Theta(\sigma < 0) + \lambda_{int}\Theta(\sigma > \text{th})$$

Disentangling intrinsic charm

$$c^{(n_f=4)}(x,Q) \simeq c^{(n_f=4)}_{(\mathrm{pert})}(x,Q) + c^{(n_f=4)}_{(\mathrm{intr})}(x,Q)$$

Extracted phenomenologically from data

from pQCD evolution and matching

from intrinsic component $c^{(n_f=3)}_{(\mathrm{intr})}(x) \neq 0$

4FNS Charm PDF constrained by experimental data for $Q > Q_0$

- NNPDF4.0 dataset
- NNLO QCD calculations

$\uparrow$ **QCD evolution**

starting point: NNPDF 4.0 methodology

$Q_0 =$ 1.65 GeV

4FNS Charm PDF parametrised at Q_0

- Deep-learning parametrisation
- Monte Carlo representation of uncertainties

$\downarrow$ **QCD evolution**

subtract perturbative component

4FNS to 3FNS transformation

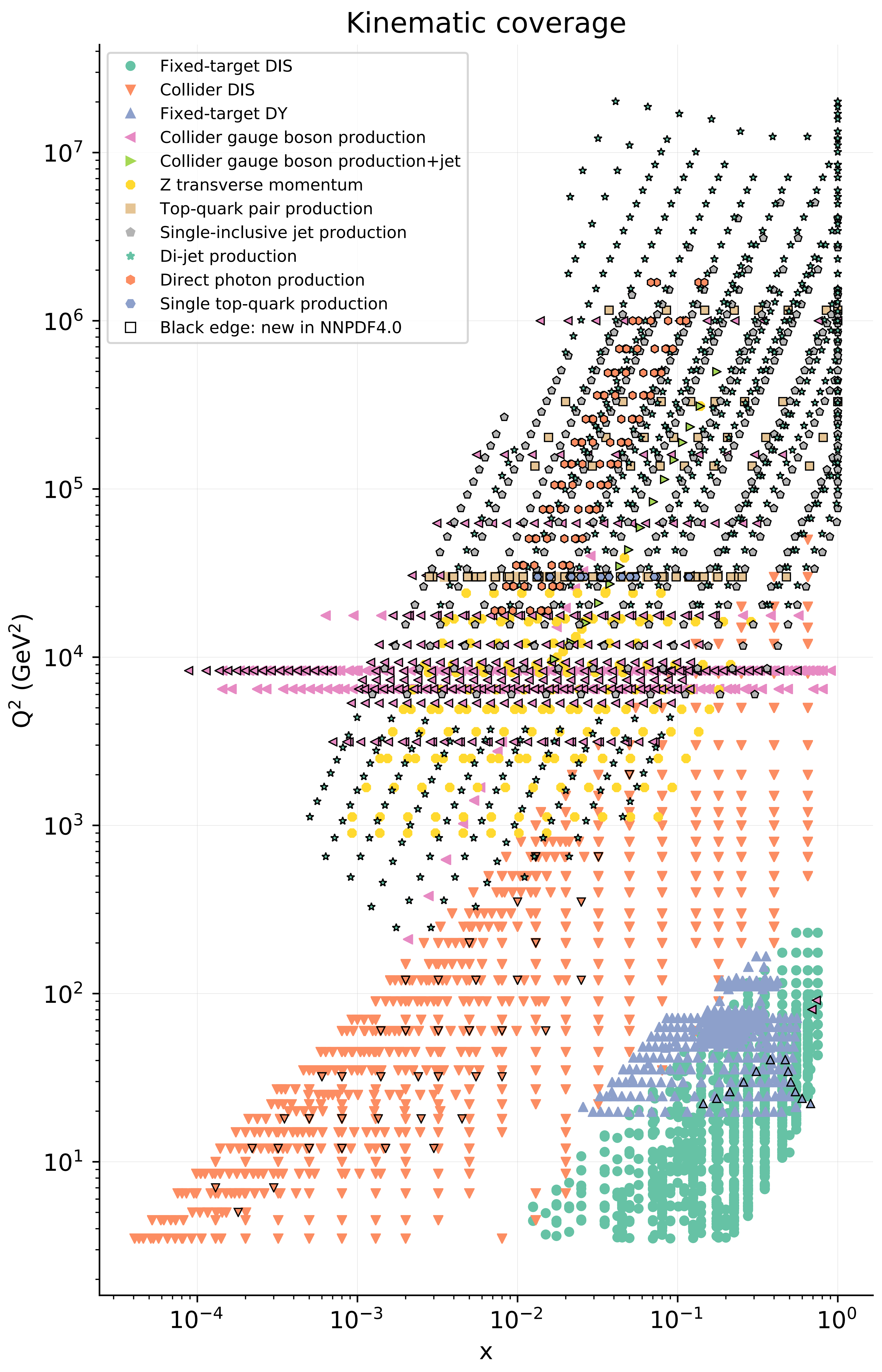
NNLO or N^3LO matching conditions

$$c^{(n_f=3)}(x,Q) = c_{(\mathrm{intr})}(x)$$

Intrinsic (3FNS) Charm

- Scale-independent
- PDF and MHO uncertainties

Breakdown of the NNPDF4.0 datasets by process type.



The full list of datasets with all references and details on the theory predictions used for each of the datasets can be consulted in the NNPDF4.0 paper: link

In the following slides the list of datasets is reproduced with the PDF determinations that use each of them.

Data set	NNPDF3.1	NNPDF4.0	ABMP16	CT18	MSHT20
CMS W asym. 7 TeV ($\mathcal{L} = 36$ pb^{-1})	✗	✗	✗	✗	✓
CMS Z 7 TeV ($\mathcal{L} = 36$ pb^{-1})	✗	✗	✗	✗	✓
CMS W electron asymmetry 7 TeV	✓	✓	✗	✓	✓
CMS W muon asymmetry 7 TeV	✓	✓	✓	✓	✗
CMS Drell-Yan 2D 7 TeV	✓	✓	✗	(✓)	✓
CMS Drell-Yan 2D 8 TeV	(✓)	✗	✗	✗	✗
CMS W rapidity 8 TeV	✓	✓	✓	✓	✓
CMS W, Z p_T 8 TeV ($\mathcal{L} = 18.4$ fb^{-1})	✗	✗	✗	(✓)	✗
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 $\sigma_{tt}^{\rm tot}$ 8 TeV	✗	✗	✗	✗	✓
CMS $\sigma_{tt}^{\rm tot}$ 5, 7, 8, 13 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	✗	✓	✓	✗	✗
CMS single top 13 TeV	✗	✗	✗	✗	(✓)

Data set	NNPDF3.1	NNPDF4.0	ABMP16	CT18	MSHT20
ATLAS W, Z 7 TeV ($\mathcal{L} = 35$ pb^{-1})	✓	✓	✓	✓	✓
ATLAS W, Z 7 TeV ($\mathcal{L} = 4.6$ fb^{-1})	✓	✓	✗	(✓)	✓
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 7 TeV	(✓)	✗	✗	(✓)	✗
ATLAS Z p_T 8 TeV	✓	✓	✗	✓	✓
ATLAS $W + c$ 7 TeV	✗	✓	✗	(✓)	✗
ATLAS $\sigma_{tt}^{\rm tot}$ 7, 8 TeV	✓	✓	✓	✗	✗
ATLAS $\sigma_{tt}^{\rm tot}$ 7, 8 TeV	✗	✗	✓	✗	✗
ATLAS $\sigma_{tt}^{\rm tot}$ 13 TeV ($\mathcal{L} = 3.2$ fb^{-1})	✓	✗	✓	✗	✗
ATLAS $\sigma_{tt}^{\rm tot}$ 13 TeV ($\mathcal{L} = 139$ fb^{-1})	✗	✓	✗	✗	✗
ATLAS $\sigma_{tt}^{\rm tot}$ and Z ratios	✗	✗	✗	✗	(✓)
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 8 TeV	✗	(✓)	✗	✗	✗
ATLAS direct photon production 13 TeV	✗	✓	✗	✗	✗
ATLAS single top R_t 7, 8, 13 TeV	✗	✓	✓	✗	✗
ATLAS single top diff. 7 TeV	✗	✓	✗	✗	✗
ATLAS single top diff. 8 TeV	✗	✓	✗	✗	✗

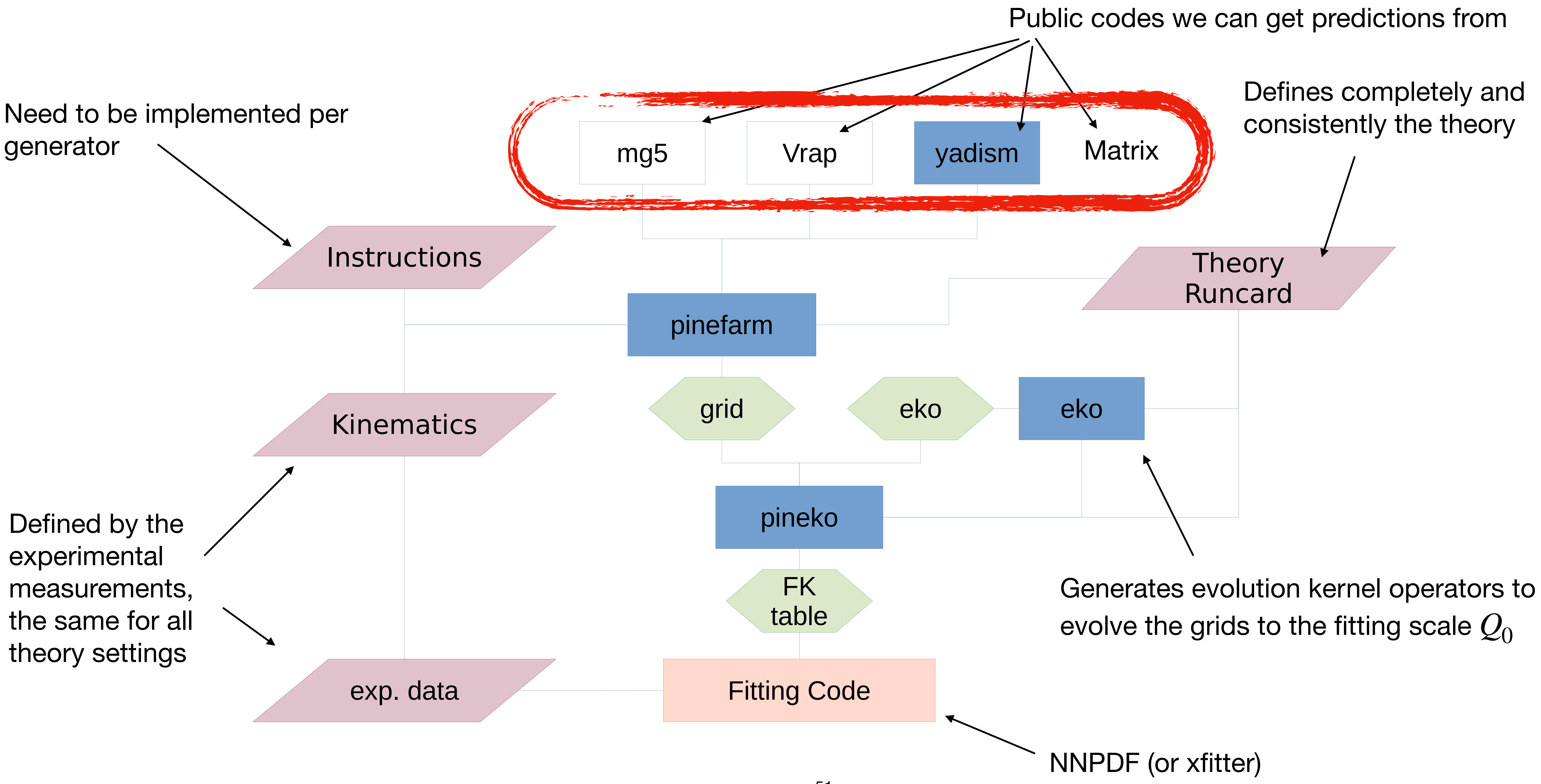
Data set	NNPDF3.1	NNPDF4.0	ABMP16	CT18	MSHT20
CDF Z rapidity	✓	✓	✗	✓	✓
CDF $W \to \ell\nu$ asymmetry (1.8 TeV)	✗	✗	✗	✓	✗
CDF $W \to e\nu$ asymmetry ($\mathcal{L} = 170$ pb^{-1})	✗	✗	✗	✓	✗
CDF $W \to e\nu$ asymmetry ($\mathcal{L} = 1$ fb^{-1})	✗	✗	✗	✗	✓
CDF k_t inclusive jets	✓	✗	✗	✗	✓
CDF cone-based inclusive jets	✗	✗	✗	✓	✗
D0 Z rapidity	✓	✓	✗	✓	✓
D0 $W \to e\nu$ asymmetry ($\mathcal{L} = 0.75$ fb^{-1})	✗	✗	✗	✗	✓
D0 $W \to e\nu$ (prod.) asymmetry ($\mathcal{L} = 9.7$ fb^{-1})	✗	✗	(✓)	✗	✓
D0 $W \to e\nu$ (prod. and decay) asymmetry ($\mathcal{L} = 9.7$ fb^{-1})	✓	(✓)	✓	✓	✗
D0 $W \to \mu\nu$ asymmetry ($\mathcal{L} = 0.3$ fb^{-1})	✗	✗	✗	✓	✗
D0 $W \to \mu\nu$ asymmetry ($\mathcal{L} = 7.3$ fb^{-1})	✓	✓	✓	✗	✓
D0 cone-based inclusive jets	✗	✗	✗	✓	✓
CDF and D0 top-pair production	✗	✗	(✓)	✗	✓
CDF and D0 single-top production	✗	✗	✓	✗	✗

Data set	NNPDF3.1	NNPDF4.0	ABMP16	CT18	MSHT20
DY E866 $\sigma^d_{\rm DY}/\sigma^p_{\rm DY}$ (NuSea)	✓	✓	✓	✓	✓
DY E866 $\sigma^p_{\rm DY}$	✓	✓	✗	✓	✓
DY E605 $\sigma^p_{\rm DY}$	✓	✓	✓	✓	✗
DY E906 $\sigma^d_{\rm DY}/\sigma^p_{\rm DY}$ (SeaQuest)	✗	✓	✗	✗	✗

LHCb Z 7 TeV ($\mathcal{L} = 940$ pb^{-1})	✓	✓	✗	✗	✓
LHCb $Z \to ee$ 8 TeV ($\mathcal{L} = 2$ fb^{-1})	✓	✓	✓	✓	✓
LHCb W 7 TeV ($\mathcal{L} = 37$ pb^{-1})	✗	✗	✗	✗	✓
LHCb $W, Z \to \mu$ 7 TeV	✓	✓	✓	✓	✓
LHCb $W, Z \to \mu$ 8 TeV	✓	✓	✓	✓	✓
LHCb $W \to e$ 8 TeV	✗	(✓)	✗	✗	✗
LHCb $Z \to \mu\mu, ee$ 13 TeV	✗	✓	✗	✗	✗

Data set	NNPDF3.1	NNPDF4.0	ABMP16	CT18	MSHT20
NMC F_2^d/F_2^p	✓	✓	✗	✗	✓
NMC $\sigma^{{\rm NC},p}$	✓	✓	✗	✓	✓
SLAC F_2^p, F_2^d	✓	✓	✓	✗	✓
BCDMS F_2^p	✓	✓	✓	✓	✓
BCDMS F_2^d	✓	✓	✗	✓	✓
BCDMS, NMC, SLAC F	✗	✗	✗	✗	✓
CHORUS $\sigma^{\nu}_{CC}, \sigma^{\bar\nu}_{CC}$	✓	✓	✗	✗	✓
CHORUS	✗	✗	✓	✗	✗
NuTeV F_2, F_3	✗	✗	✗	✗	✓
NuTeV/CCFR $\sigma^{\nu}_{CC}, \sigma^{\bar\nu}_{CC}$	✓	✓	✓	✓	✓
EMC F_2^c	(✓)	(✓)	✗	✗	✗
NOMAD	✗	(✓)	✓	✗	✗
CCFR xF_3^p	✗	✗	✗	✓	✗
CCFR F_2^p	✗	✗	✗	✓	✗
CDSHW F_2^p, xF_3^p	✗	✗	✗	✓	✗
E665 F_2^p, F_2^d	✗	✗	✗	✗	✓
HERA NC, CC	✗	✗	✗	✗	✓
HERA I+II $\sigma^p_{\rm NC,CC}$	✓	✓	✓	✓	✗
HERA I+II $\sigma^{\rm red}_{c\bar c}$	✗	✓	✗	(✓)	✓
HERA I+II $\sigma^{\rm red}_{b\bar b}$	✗	✓	✗	(✓)	✗
HERA I+II $\sigma^{\rm red}_{c\bar c}$	✓	✗	✓	✓	✗
H1 $F_2^{c\bar c}$	✗	✗	✗	✓	✗
H1 $F_2^{b\bar b}$	✓	✗	✓	✗	✗
ZEUS $\sigma^{\rm red}_{b\bar b}$	✓	✗	✓	✗	✗
H1 $F_{\rm L}$	✗	✗	✗	✓	✓
H1 and ZEUS $F_{\rm L}$	✗	✗	✗	✗	✓
ZEUS 820 (HQ) (1j)	✗	(✓)	✗	✗	✗
ZEUS 920 (HQ) (1j)	✗	(✓)	✗	✗	✗
H1 (LQ) (1j-2j)	✗	(✓)	✗	✗	✗
H1 (HQ) (1j-2j)	✗	(✓)	✗	✗	✗
ZEUS 920 (HQ) (2j)	✗	(✓)	✗	✗	✗

The NNPDF theory pipeline



Automatic hyperparameter selection

The usage of Neural Networks had as primary goal eliminating the biases associated with the choice of a specific functional form.

However, there are still many choices associated with the optimization:

- Number and width of the layers
- Activation functions and initialization
- Optimization algorithm (and associated parameters)
- Training length, stopping patience, etc.
- Strength of lagrange multipliers (positivity, integrability)

Collectively called “hyperparameters”, we are going to sample them automatically in order to remove any kind of human intervention.

Automatic hyperparameter selection

The usage of Neural Networks had as primary goal eliminating the biases associated with the choice of a specific functional form.

However, there are still many choices associated with the optimization:

- Number and width of the layers
- Activation functions and initialization
- Optimization algorithm (and associated parameters)
- Training length, stopping patience, etc.
- Strength of lagrange multipliers (positivity, integrability)

Collectively called "hyperparameters", we are going to sample them automatically in order to remove any kind of human intervention.

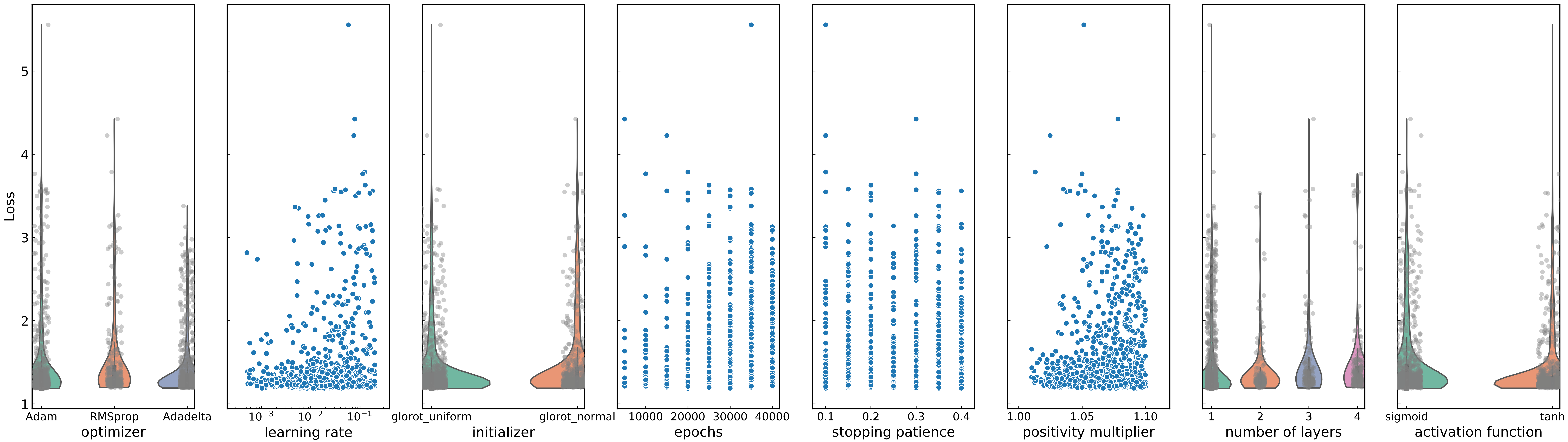
Automatic hyperparameter selection

Model selection

Select a model such that:

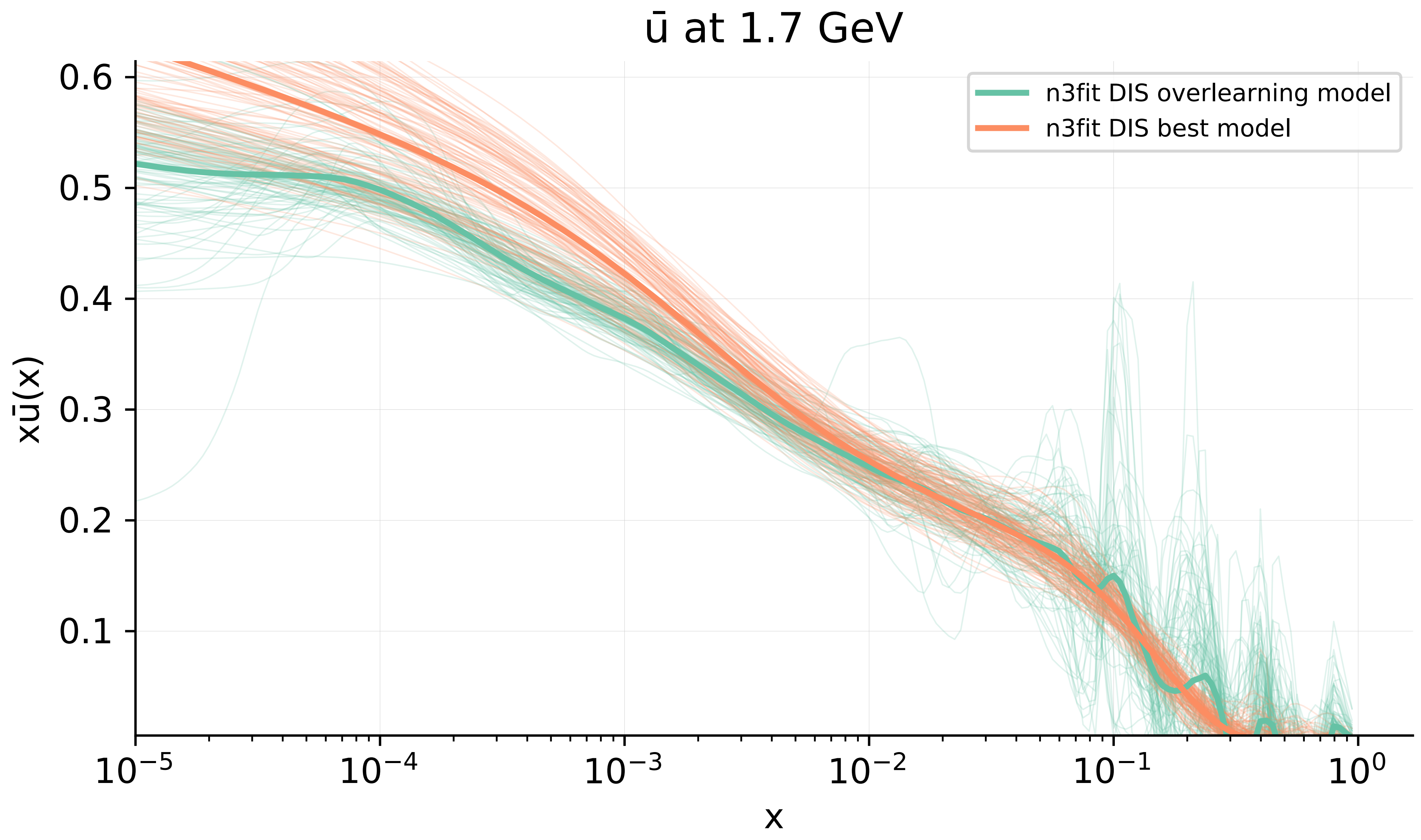
- The χ^2 is minimized (so the data is well described)
- Generalizable
- Fast (choose the faster methodology for the same quality!)
- Stable upon variations (we don't want to redo it too often!)

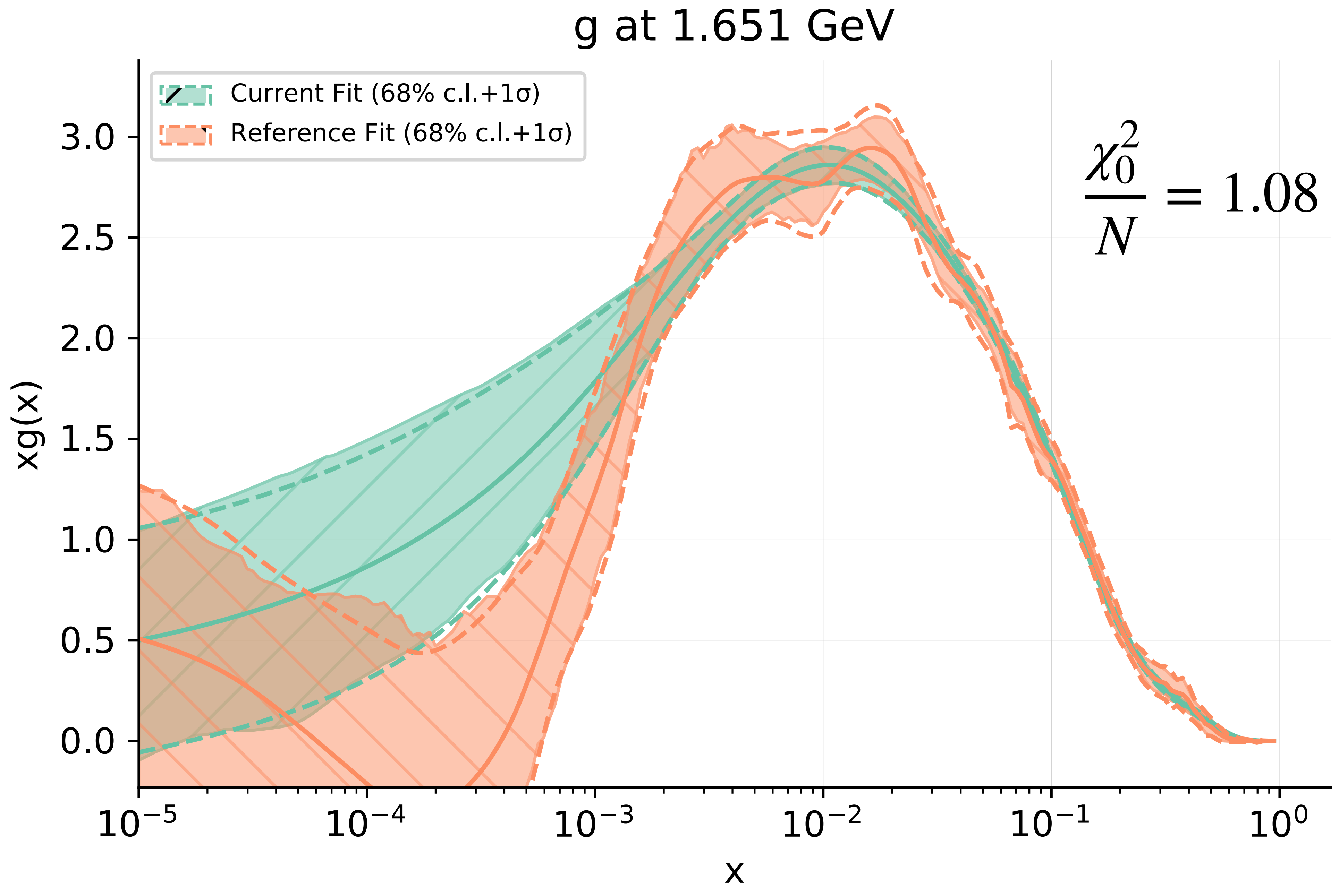
Perform many fits with many different hyperparameter choices and select the absolute best



Automatic hyperparameter selection

With great power comes great responsability





Getting a χ^2 many units below the nominal NNPDF4.0 (~1.16) is relatively “easy”, but that doesn’t mean it is a good fit.

Automatic hyperparameter selection

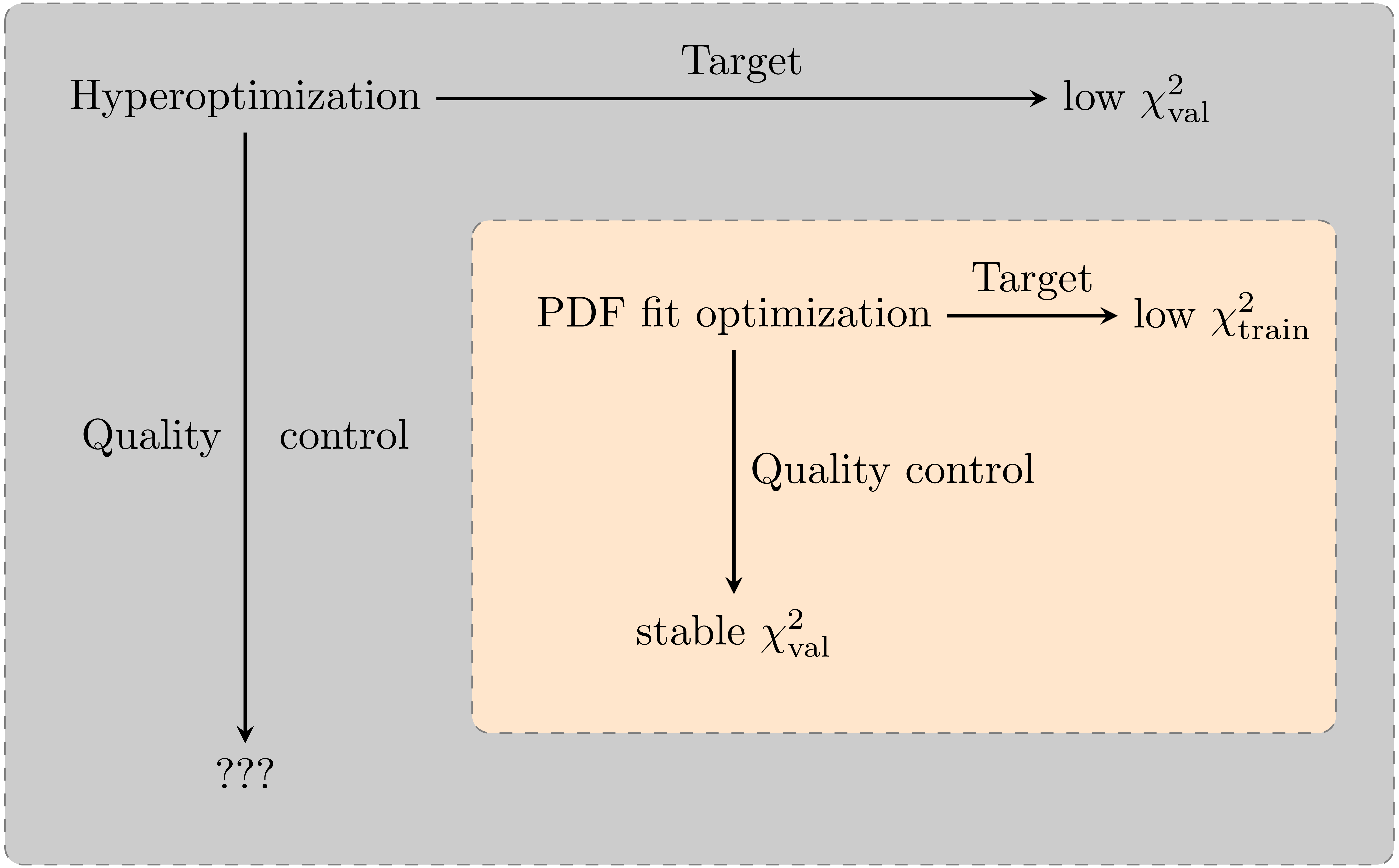
$$\begin{tikzcd} \text{PDF fit optimization} \arrow[r, "\text{Target}"] \arrow[d, "\text{Quality control}"] & \text{low } \chi^2_{\mathrm{train}} \\ \text{stable } \chi^2_{\mathrm{val}} & \end{tikzcd}$$

K-folding cross validation:

1. Divide data into k sets
2. Leave one out and fit using the union of the k-1 sets that are still in
3. Compute a reward/loss function on the datasets that are left out

$$\mathscr{L}(\text{parameters}) = \frac{1}{k}\sum_{k}^{i}\frac{\chi_i^2}{N_i}$$

Automatic hyperparameter selection

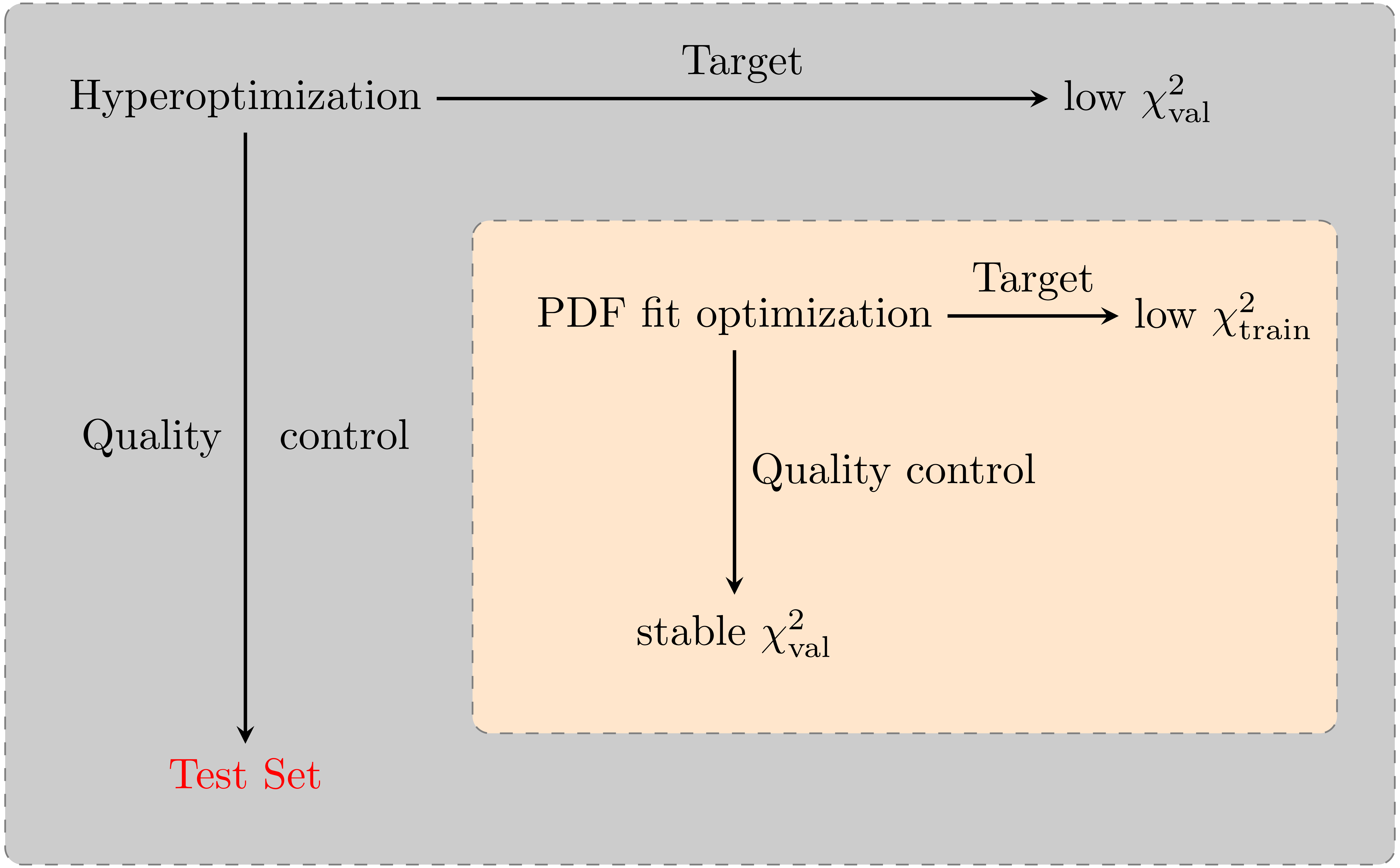


K-folding cross validation:

1. Divide data into k sets
2. Leave one out and fit using the union of the k-1 sets that are still in
3. Compute a reward/loss function on the datasets that are left out

$$\mathscr{L}(\text{parameters}) = \frac{1}{k}\sum_{k}^{i}\frac{\chi_i^2}{N_i}$$

Automatic hyperparameter selection



K-folding cross validation:

1. Divide data into k sets
2. Leave one out and fit using the union of the k-1 sets that are still in
3. Compute a reward/loss function on the datasets that are left out

$$\mathscr{L}(\text{parameters}) = \frac{1}{k}\sum_{k}^{i}\frac{\chi_i^2}{N_i}$$