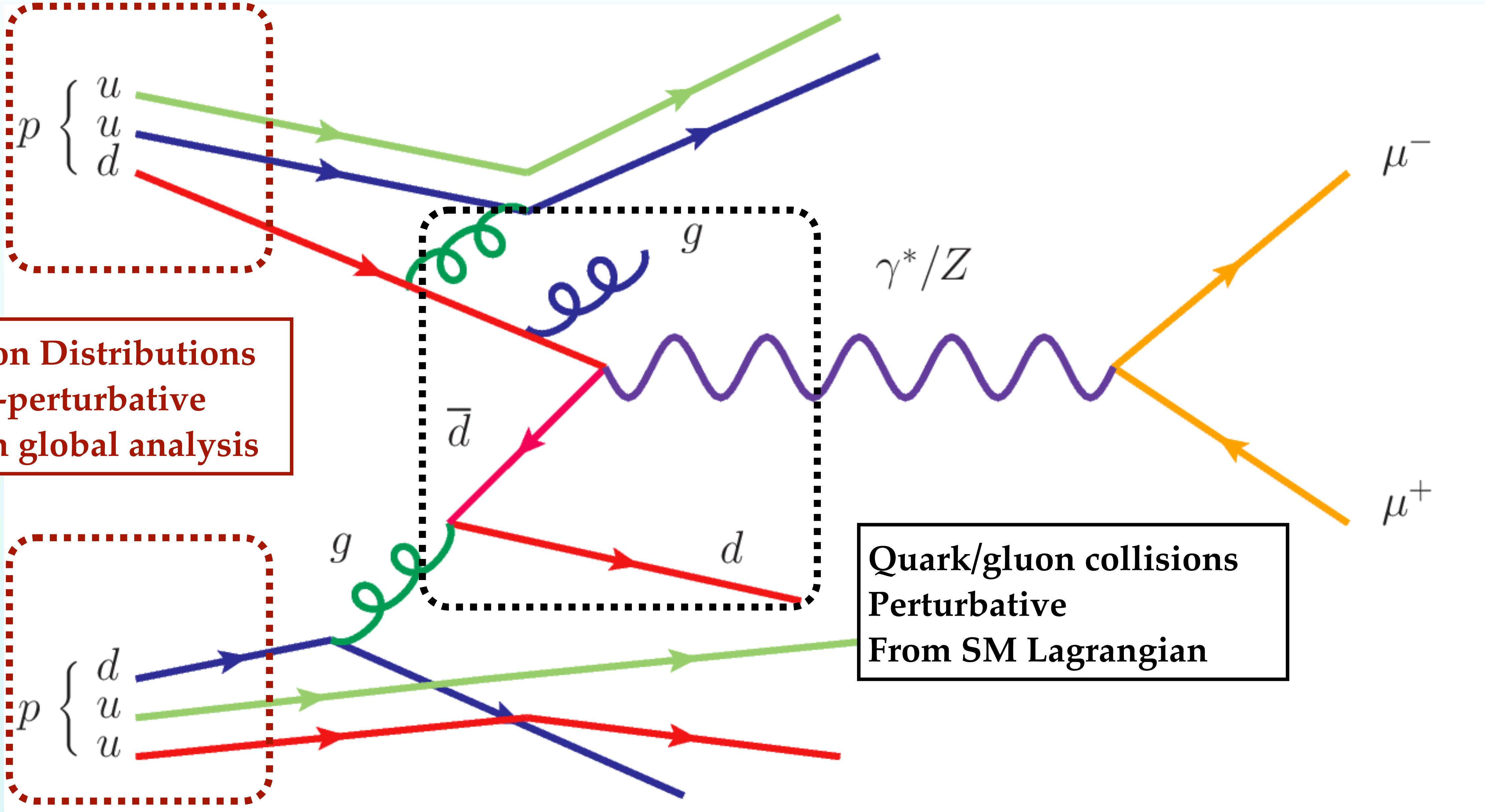
The quark and gluon structure of the proton in the high-precision LHC era

Juan Rojo
VU Amsterdam & Theory group, Nikhef

Anatomy of hadronic collisions

In high-energy **hadron colliders** the collisions involve **composite particles** (protons) with internal substructure (quarks and gluons): the LHC is actually a quark/gluon collider!

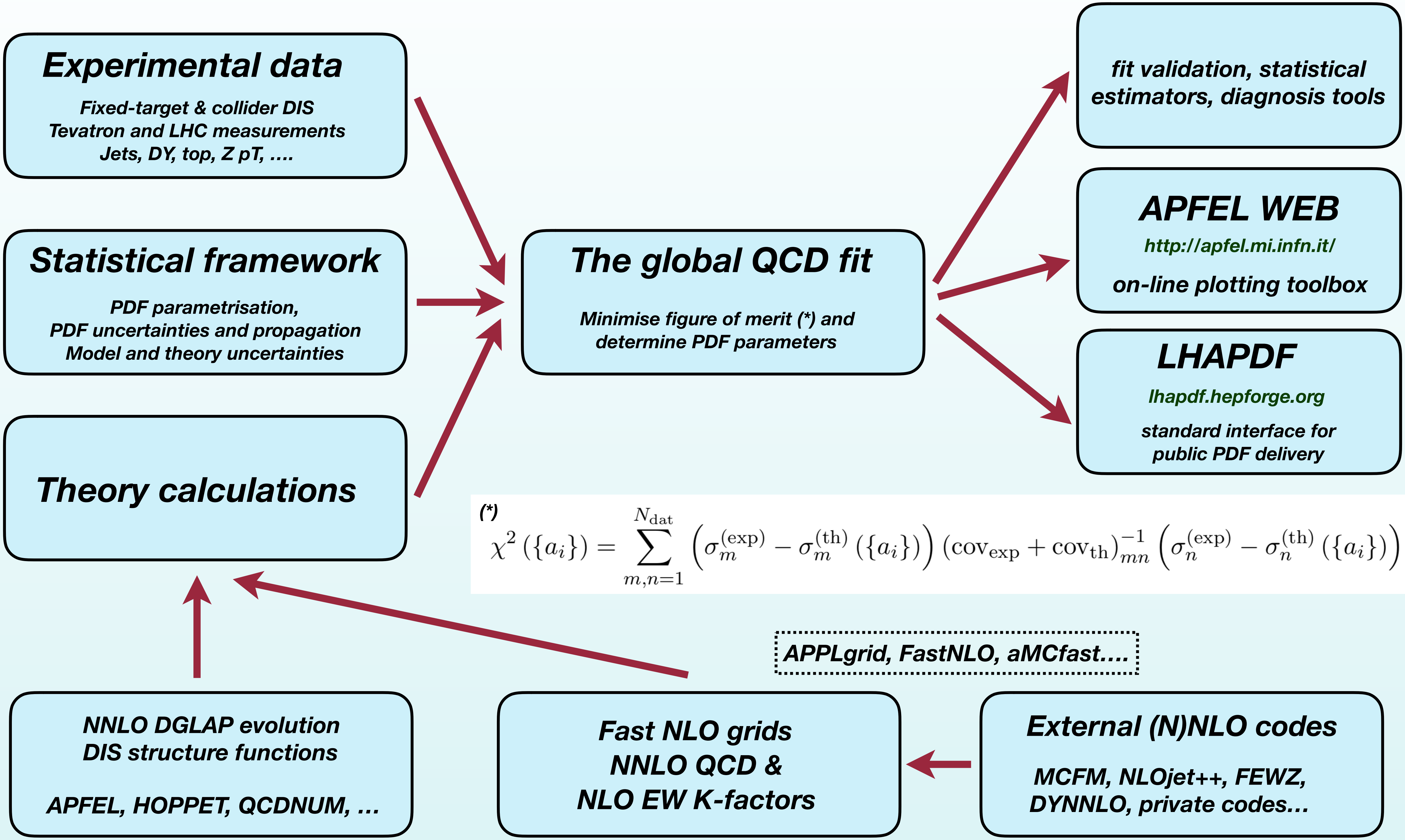


Parton Distributions
Non-perturbative
From global analysis

Quark/gluon collisions
Perturbative
From SM Lagrangian

Calculations of **cross-sections** in hadron collisions require the combination of **perturbative cross-sections** with **non-perturbative parton distribution functions (PDFs)**

The global QCD fit machinery



Experimental data

Fixed-target & collider DIS
Tevatron and LHC measurements
Jets, DY, top, Z pT, ….

Statistical framework

PDF parametrisation,
PDF uncertainties and propagation
Model and theory uncertainties

Theory calculations

The global QCD fit

Minimise figure of merit (*) and determine PDF parameters

fit validation, statistical estimators, diagnosis tools

APFEL WEB

http://apfel.mi.infn.it/

on-line plotting toolbox

LHAPDF

lhapdf.hepforge.org

standard interface for public PDF delivery

(*)

$$\chi^2\left(\{a_i\}\right)=\sum_{m,n=1}^{N_{\rm dat}}\left(\sigma_m^{(\rm exp)}-\sigma_m^{(\rm th)}\left(\{a_i\}\right)\right)\left({\rm cov}_{\rm exp}+{\rm cov}_{\rm th}\right)^{-1}_{mn}\left(\sigma_n^{(\rm exp)}-\sigma_n^{(\rm th)}\left(\{a_i\}\right)\right)$$

APPLgrid, FastNLO, aMCfast….

NNLO DGLAP evolution
DIS structure functions

APFEL, HOPPET, QCDNUM, …

Fast NLO grids
NNLO QCD &
NLO EW K-factors

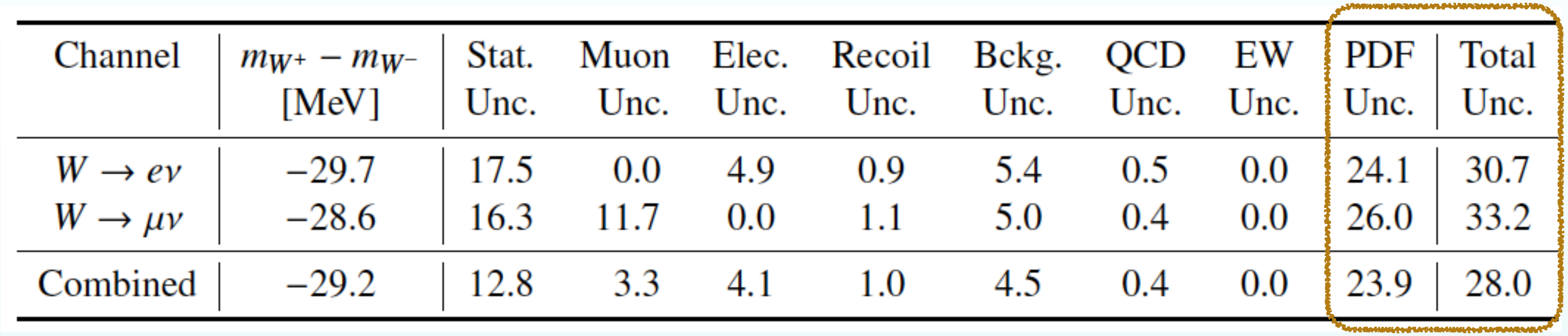
External (N)NLO codes

MCFM, NLOjet++, FEWZ,
DYNNLO, private codes…

Why better PDFs?

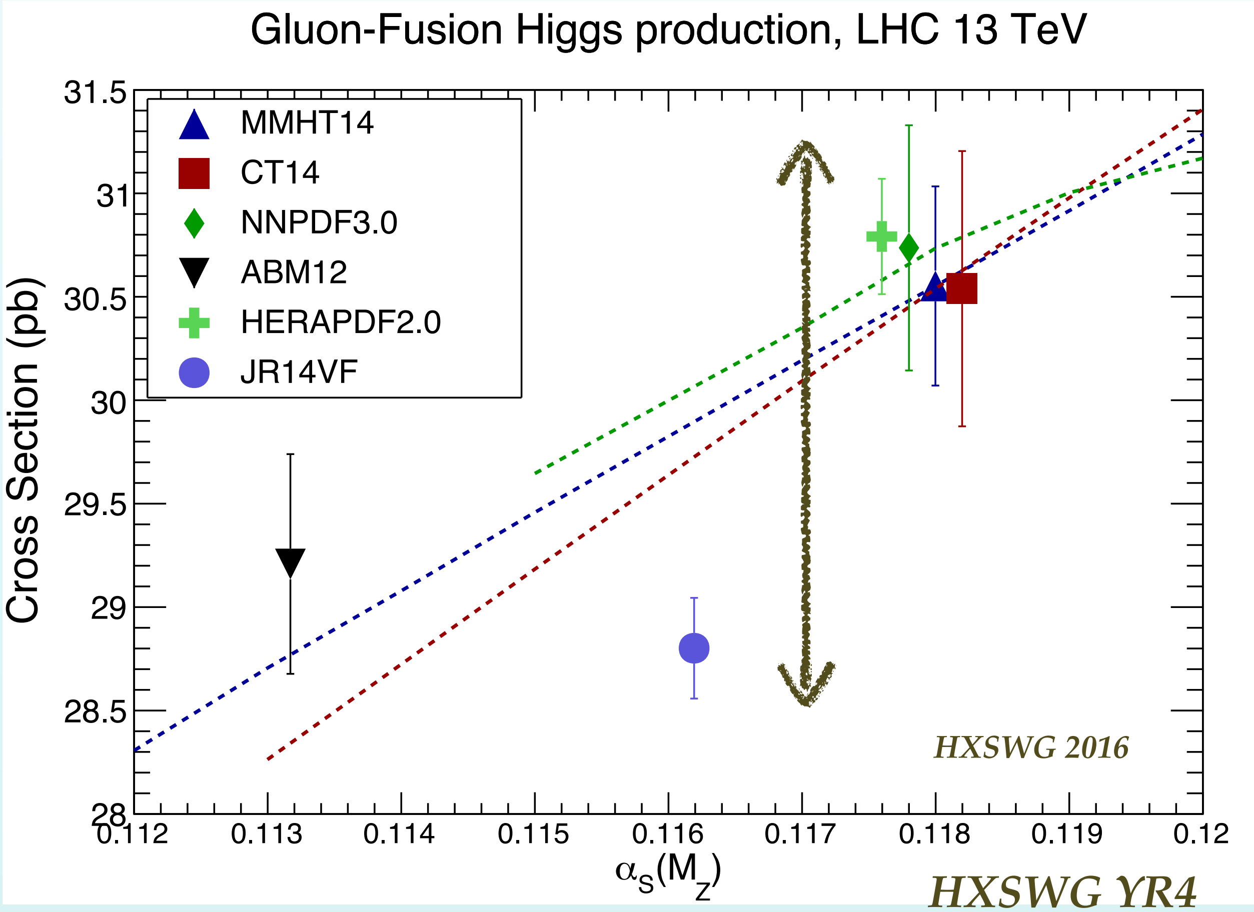
Dominant TH unc for M_W measurements at LHC

ATLAS 2017

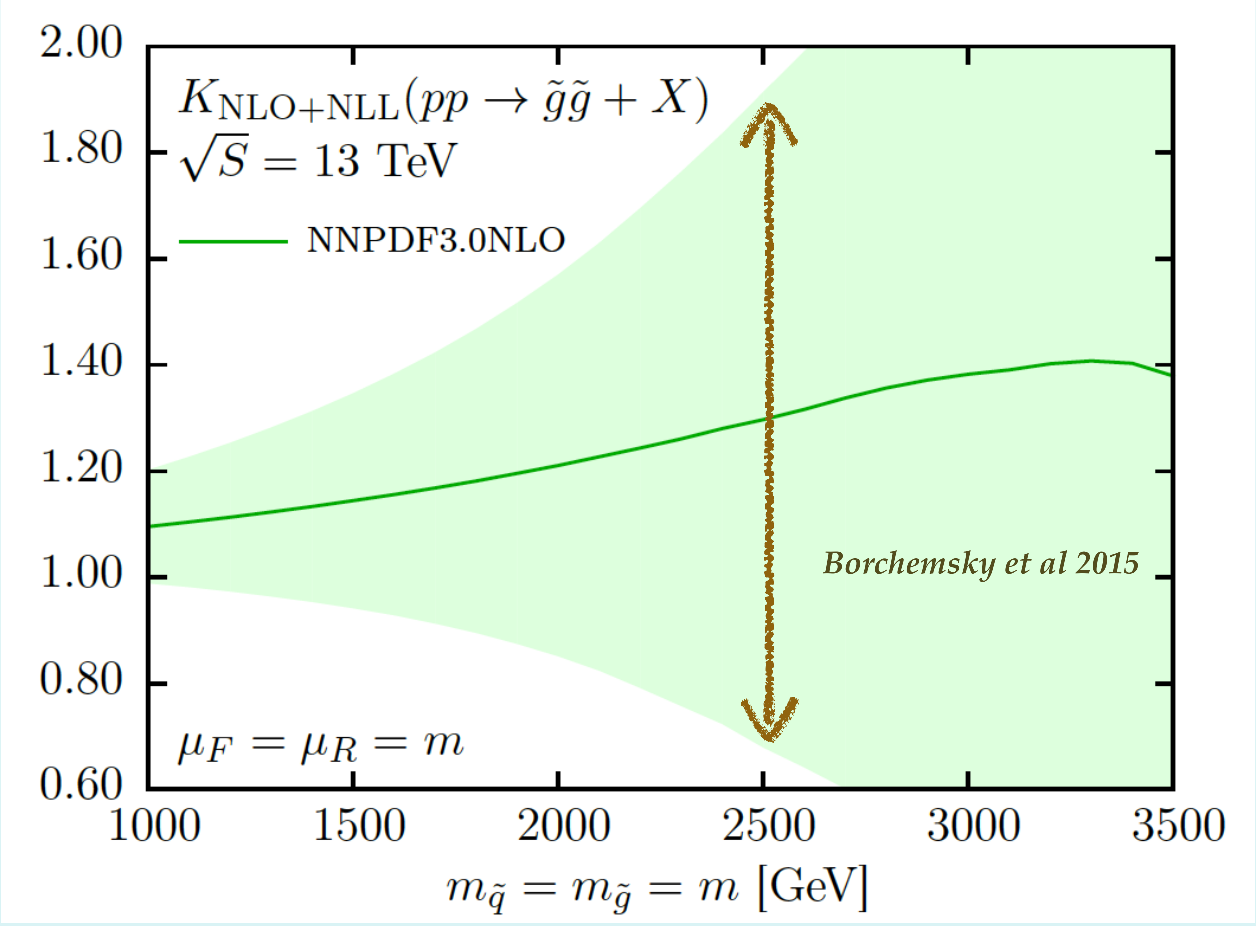


Channel	$m_{W^+} - m_{W^-}$ [MeV]	Stat. Unc.	Muon Unc.	Elec. Unc.	Recoil Unc.	Bckg. Unc.	QCD Unc.	EW Unc.	PDF Unc.	Total Unc.
$W \to e\nu$	−29.7	17.5	0.0	4.9	0.9	5.4	0.5	0.0	24.1	30.7
$W \to \mu\nu$	−28.6	16.3	11.7	0.0	1.1	5.0	0.4	0.0	26.0	33.2
Combined	−29.2	12.8	3.3	4.1	1.0	4.5	0.4	0.0	23.9	28.0

Higgs coupling measurements

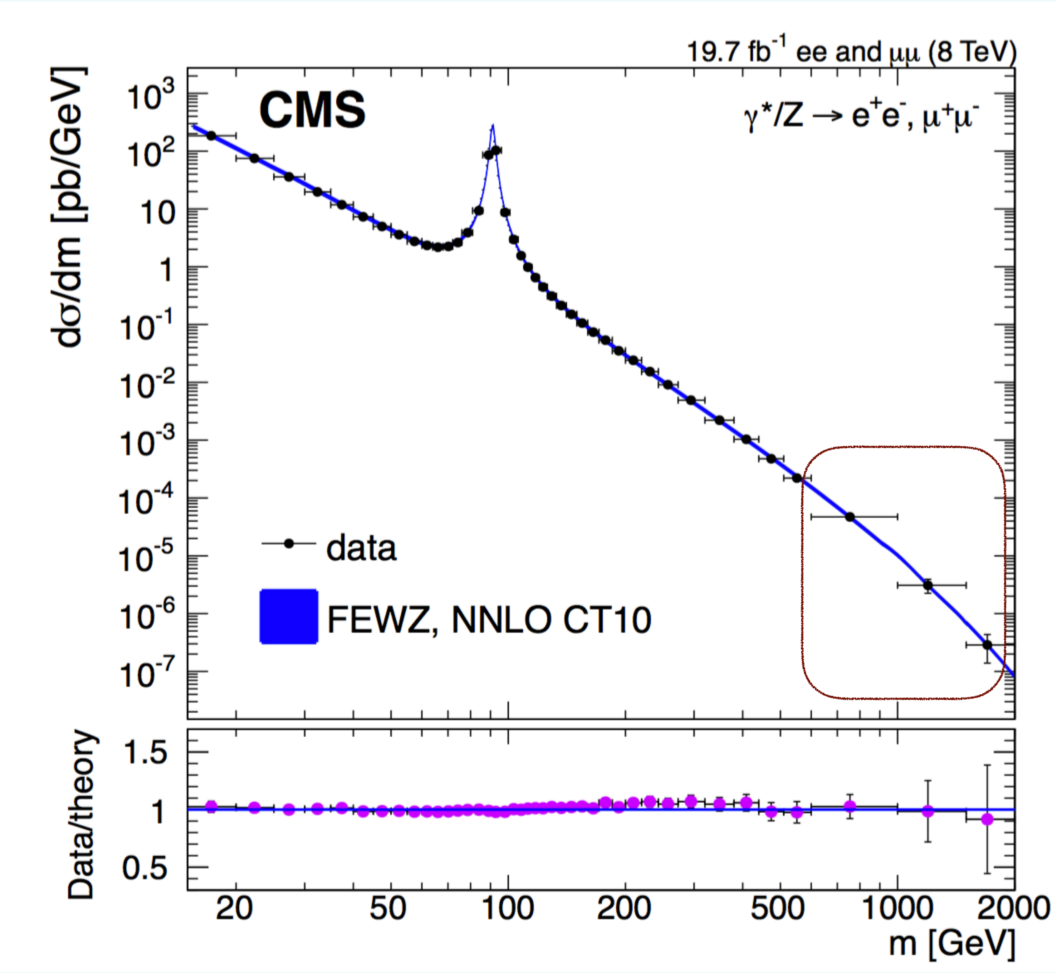


High-mass BSM cross-sections



Why better PDFs?

- BSM physics could manifest as **subtle deviations** wrt to the Standard Model predictions
- Even for high-mass resonances, **PDF uncertainties degrade or limit many BSM searches**
- The robustness of **global stress-tests of the SM** (electroweak fit, SM Effective Field Theory analysis) relies crucially in high-precision theoretical calculations



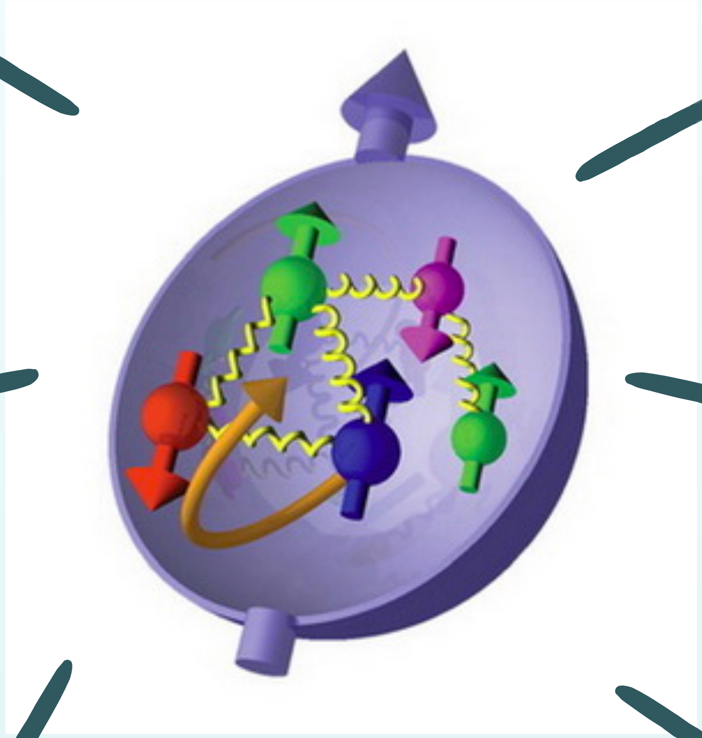
SMEFT expansion

$$\sigma(E) = \sigma_{SM}(E)\left(1 + \epsilon\frac{m_{SM}^2}{m_W^2} + \boxed{\epsilon\frac{E^2}{m_W^2}} + \dots\right)$$

For $\mathbf{E} \simeq \mathbf{1}$ **TeV**, a measurement with $\delta\sigma/\sigma \simeq \mathbf{10}\%$ is sensitive to $\epsilon \simeq \mathcal{O}(\mathbf{0.1}\%)$!

BSM physics might very well hiding itself in the tails of LHC distributions, but need to make sure first that PDF uncertainties are under control

The inner life of protons



Gluon

Strong coupling

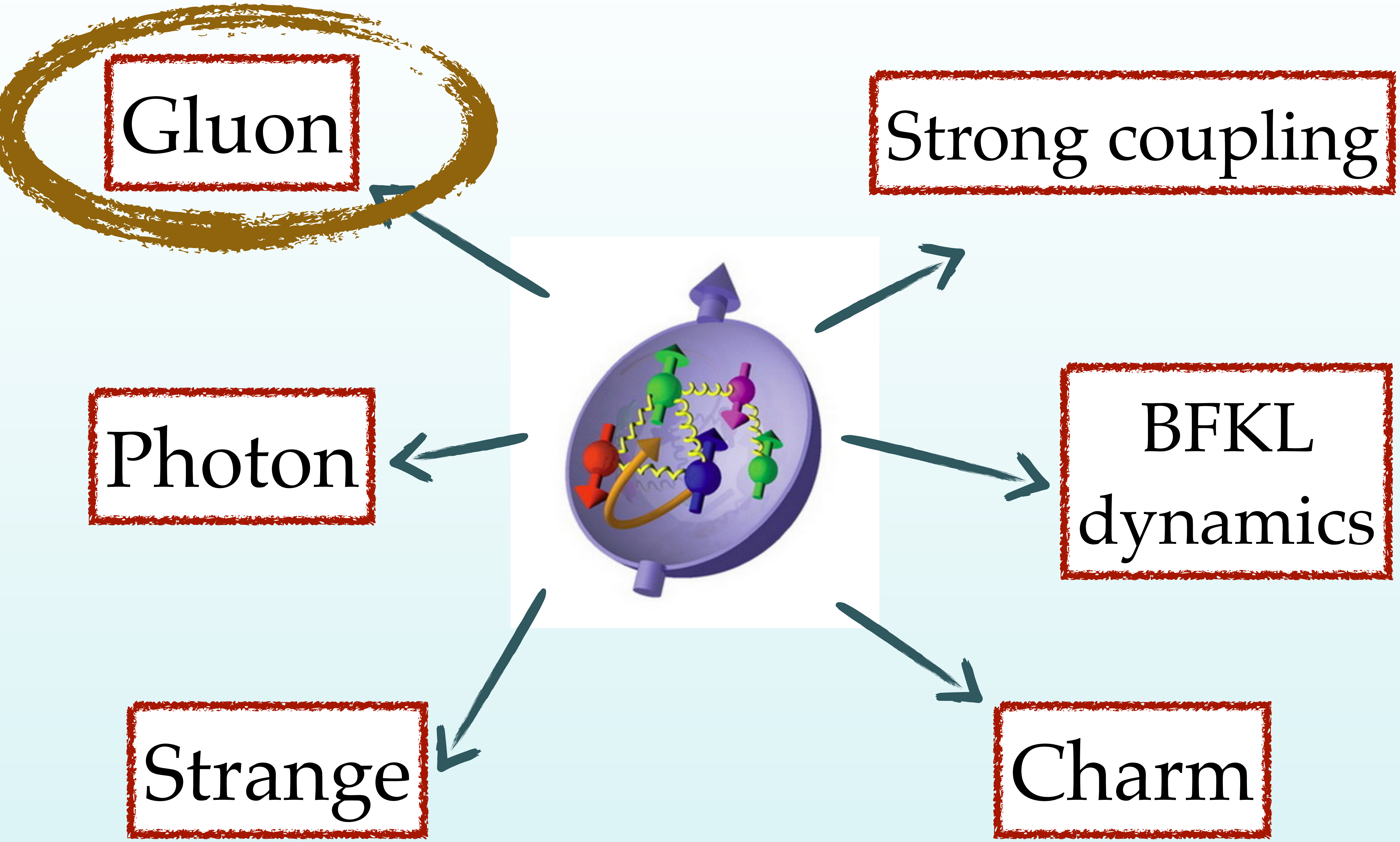
Photon

BFKL dynamics

Strange

Charm

The inner life of protons



Gluon

Strong coupling

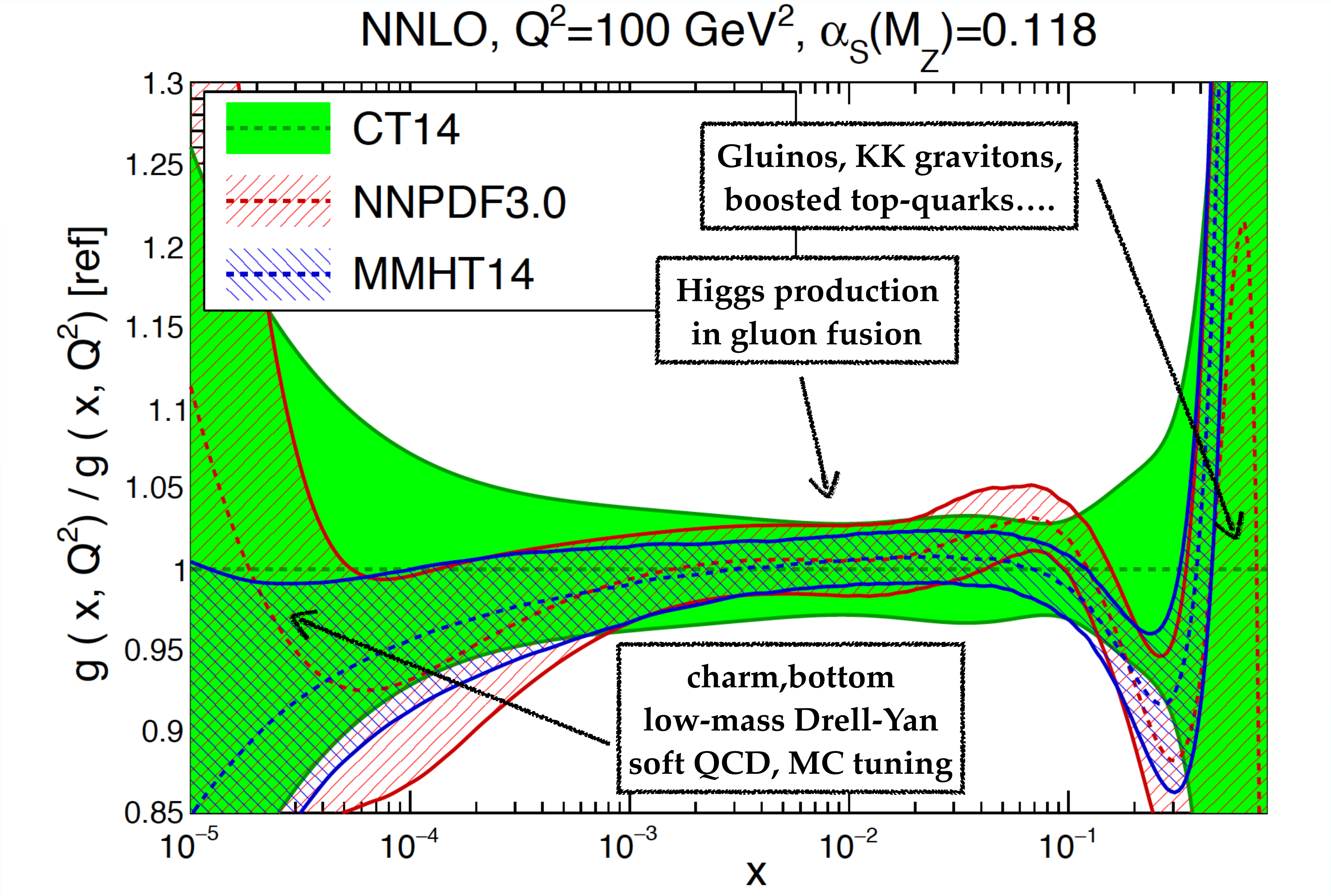
Photon

BFKL dynamics

Strange

Charm

One glue to bind them all



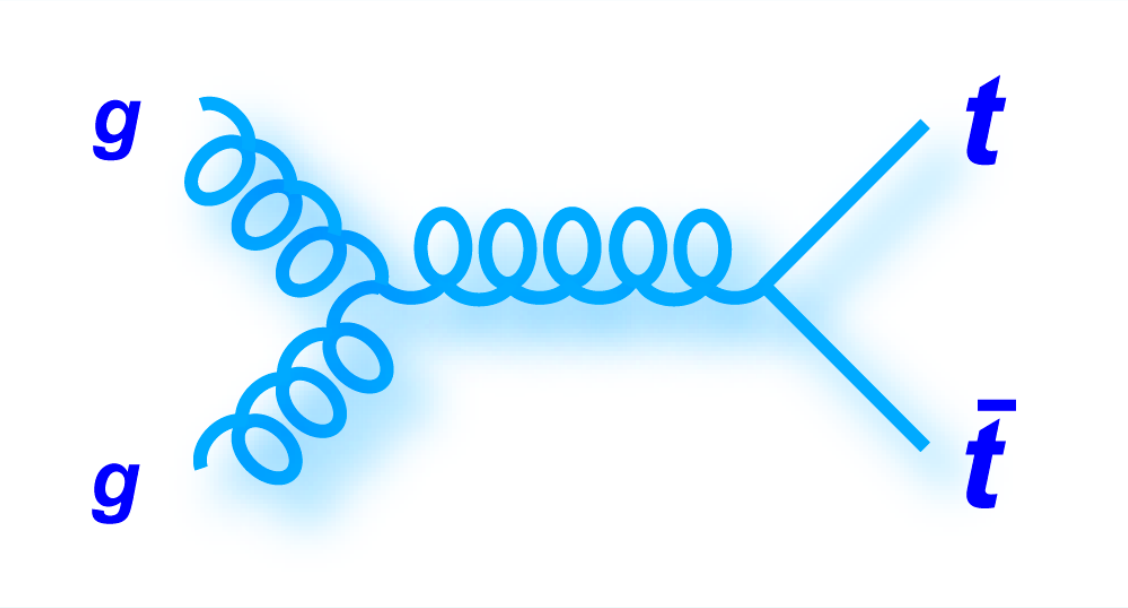
Gluinos, KK gravitons, boosted top-quarks….

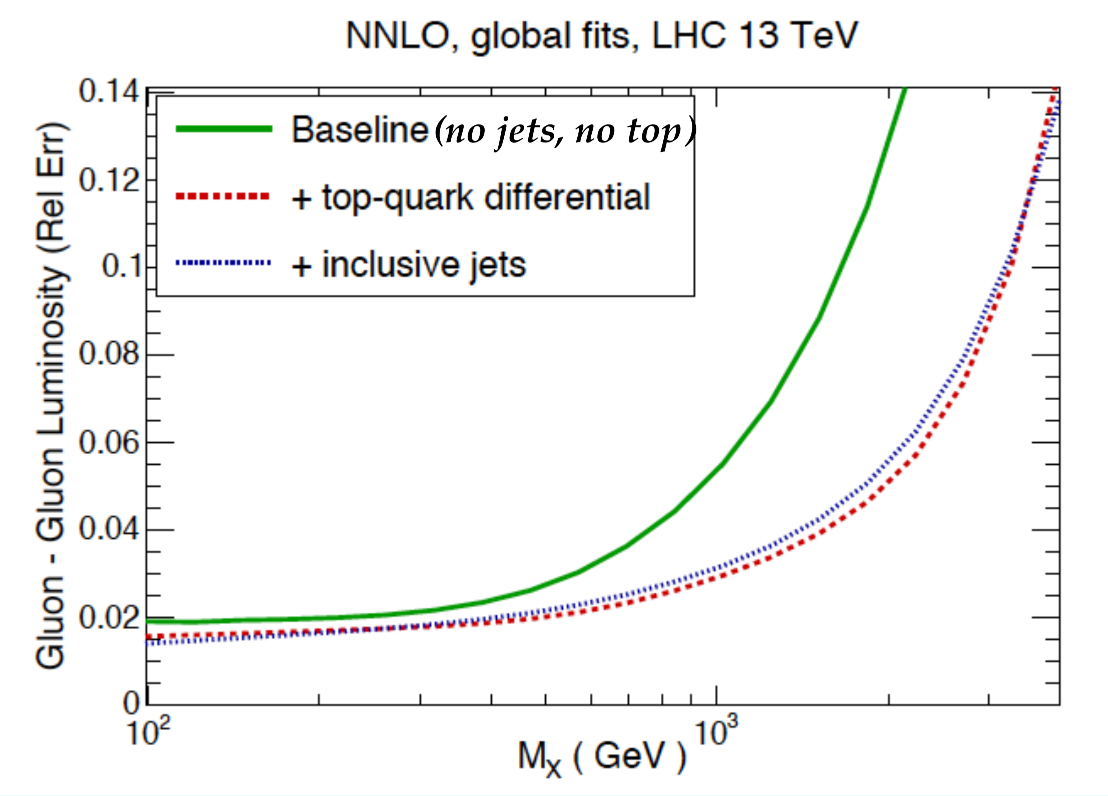
Higgs production in gluon fusion

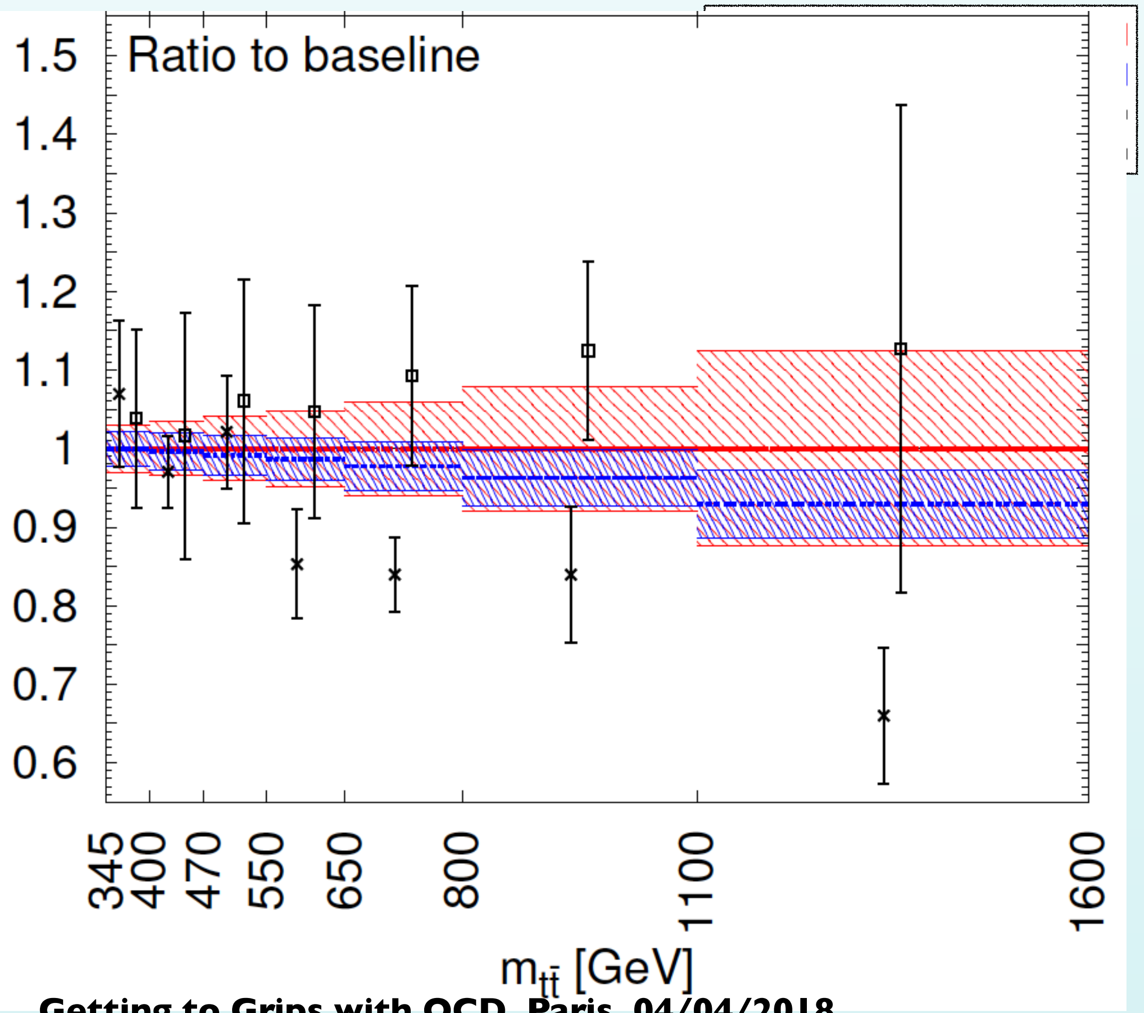
charm,bottom
low-mass Drell-Yan
soft QCD, MC tuning

At the LHC, precise knowledge of the gluon is required **from small-x to large-x**

The large-x gluon from differential top quarks







- Top-quark production driven by **gluon-gluon scattering**

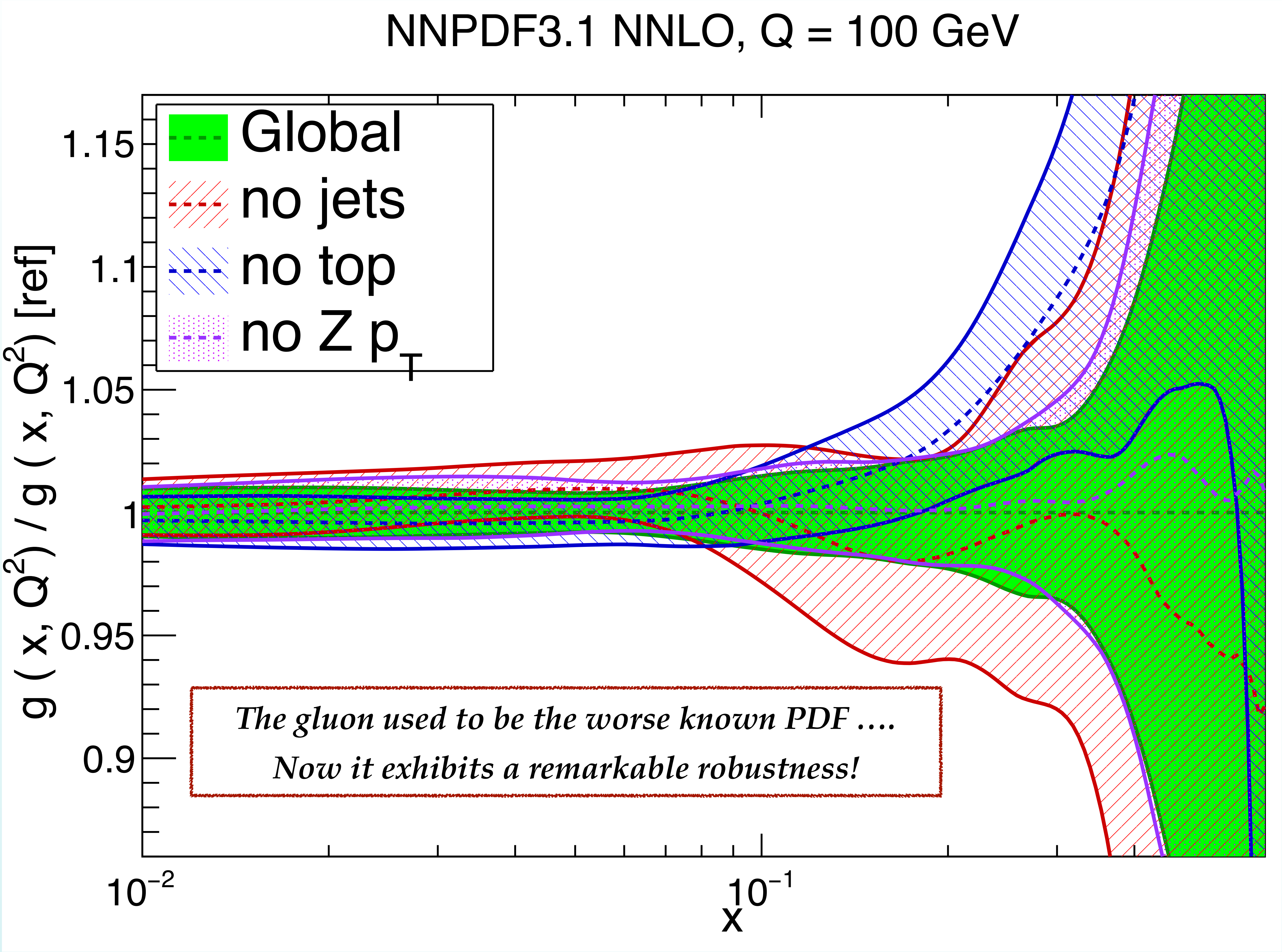
- **NNLO calculations for stable top quarks** available *Czakon, Mitov et al 2015-2017*

- **Data from ATLAS and CMS at 8 TeV** available with breakdown of systematic uncertainties

- Included **differential top data into NNPDF3.0**: constraints on the **large-*x* gluon** comparable to those of inclusive jet production *Czakon et al 2017*

- Improved theory uncertainties in **regions crucial for BSM searches**, *i.e.*, $\mathbf{m_{tt} > 1}$ **TeV** (while fitting only $\mathbf{y_t}$ and $\mathbf{y_{tt}}$)

One (upgraded) glue to bind them all



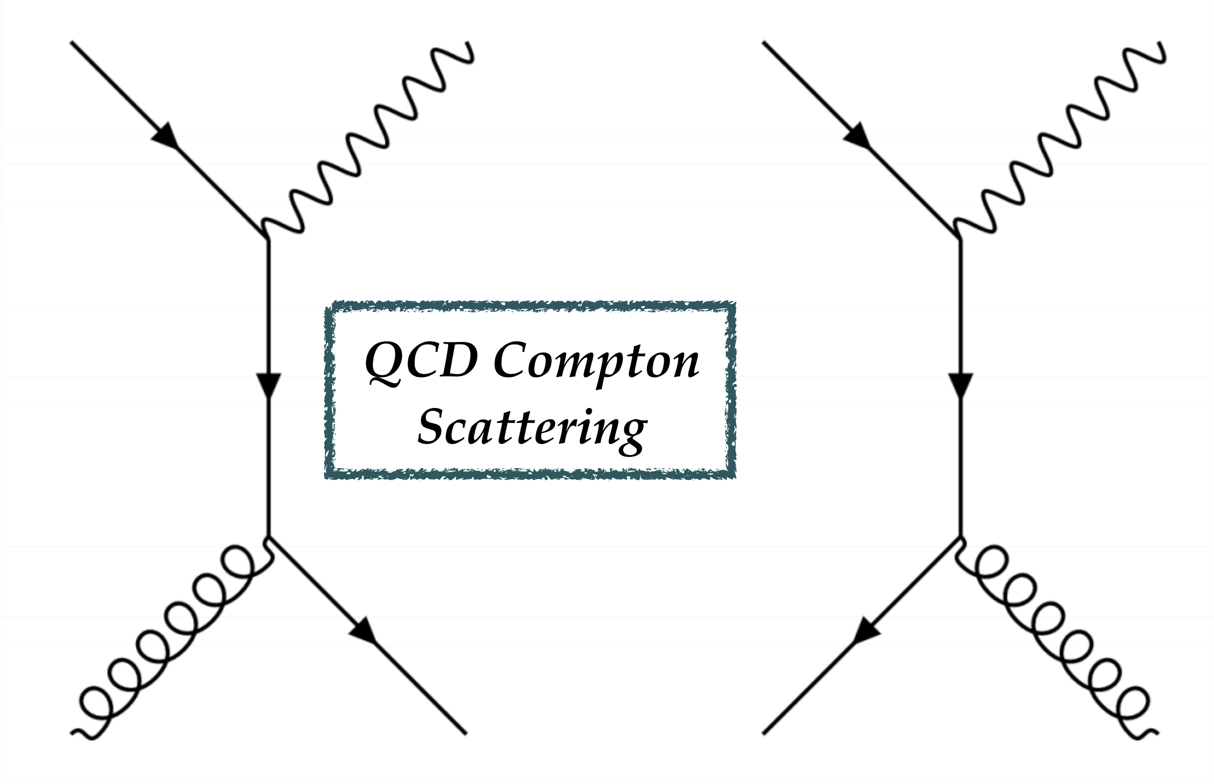
The gluon used to be the worse known PDF ….

Now it exhibits a remarkable robustness!

Direct photon production and PDF fits

- Photon production in hadronic collisions is directly sensitive to the **gluon PDF**

- Photon production data from **fixed-target experiments** was used in the very early global PDF fits to constrain the gluon PDF, but the apparent tension with some data lead to its **replacement by jets**



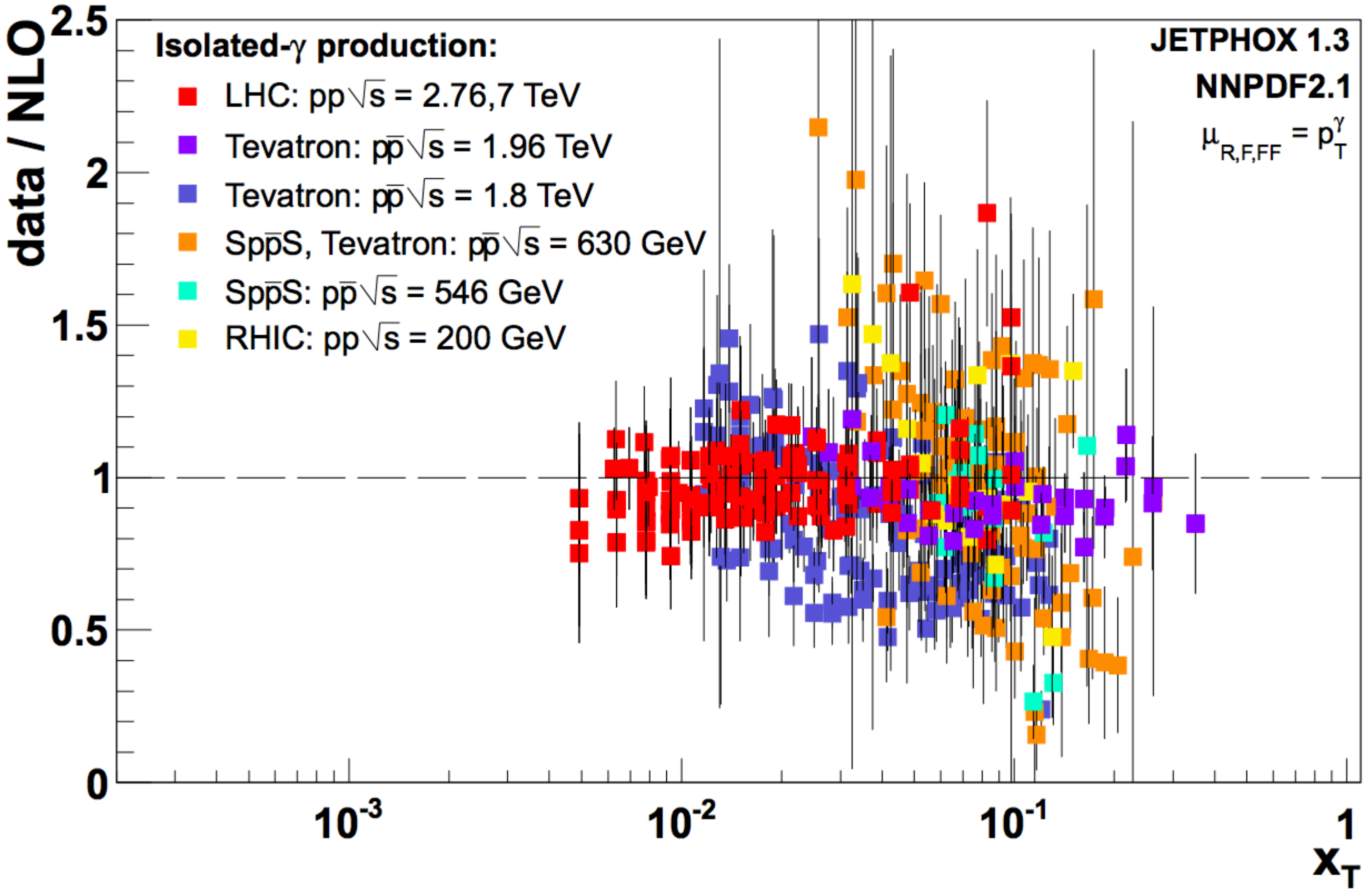
QCD Compton Scattering

- In 2012 we showed that all **available isolated photon production data** was consistent with NLO QCD calculations

D'Enterria, Rojo 12

- However the precision of most recent LHC data required using NNLO QCD theory, which only recently became available

Campbell, Ellis, Williams 16



Direct photon production and PDF fits

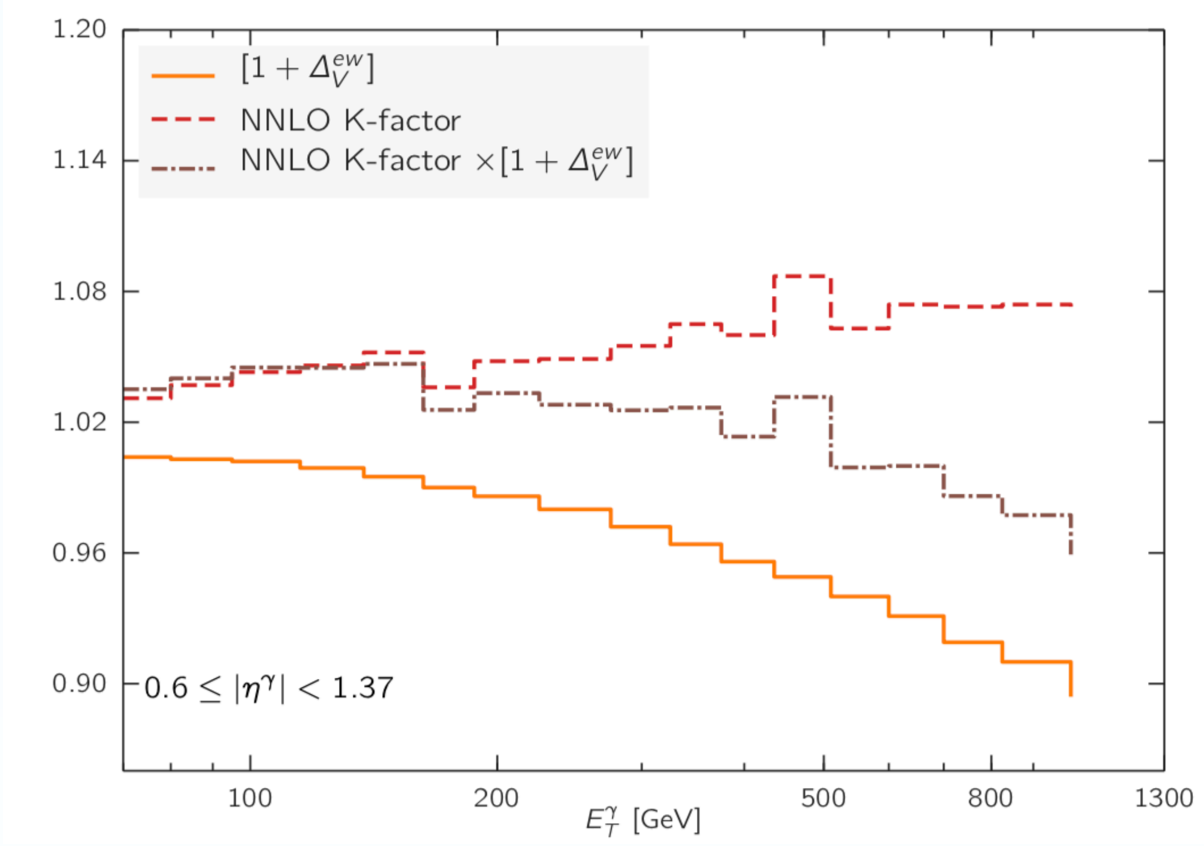
- Recently we have revisited the impact of LHC photon data into the global PDF fit, specifically the ATLAS 8 TeV data

  Campbell, Rojo, Slade, Williams 18

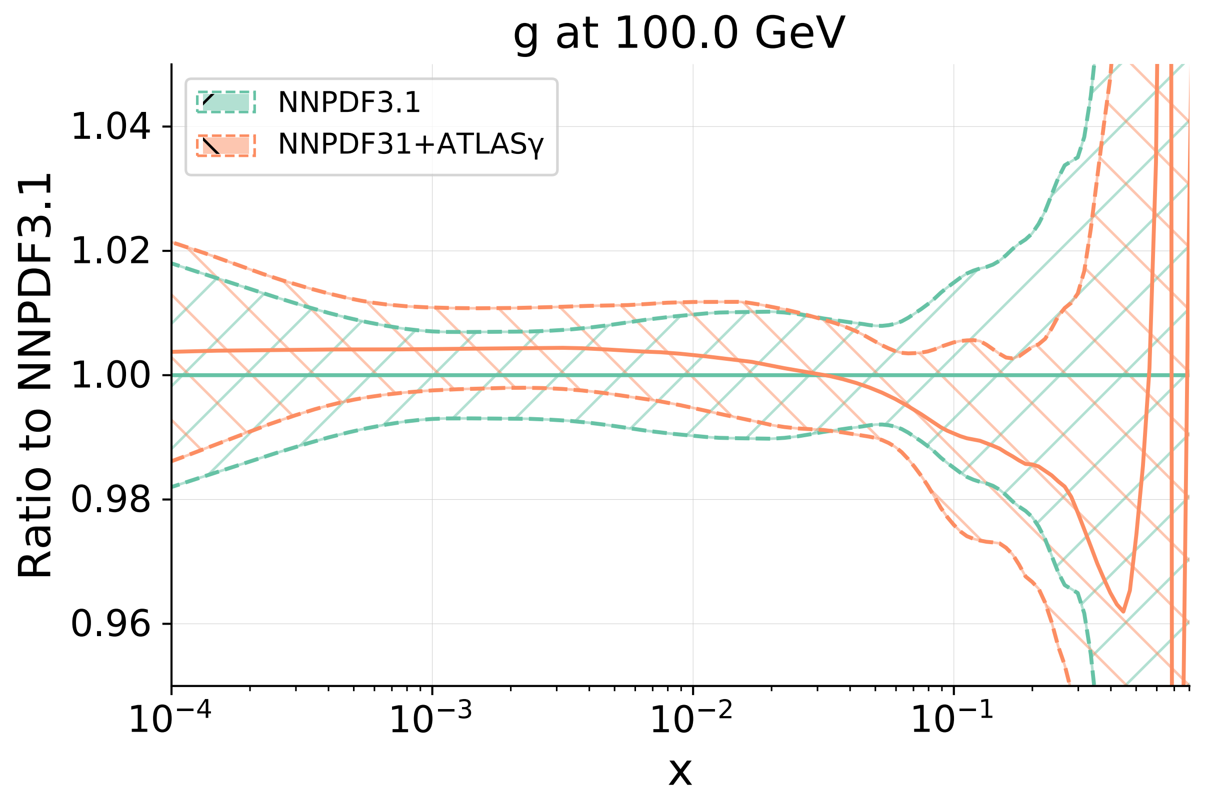
- Theory based on **NNLO QCD** and **LL electroweak** calculations

- Moderate impact on the medium-x gluon, consistent with previous studies at NLO

- Good consistency with the rest of gluon-sensitive experiments in NNPDF3.1



	NNPDF3.1	NNPDF3.1+ATLASγ
Fixed-target lepton DIS	1.207	1.203
Fixed-target neutrino DIS	1.081	1.087
HERA	1.166	1.169
Fixed-target Drell-Yan	1.241	1.242
Collider Drell-Yan	1.356	1.346
Top-quark pair production	1.065	1.049
Inclusive jets	0.939	0.915
Z p_T	0.997	0.980
Total dataset	**1.148**	**1.146**



The small-x gluon from forward charm production

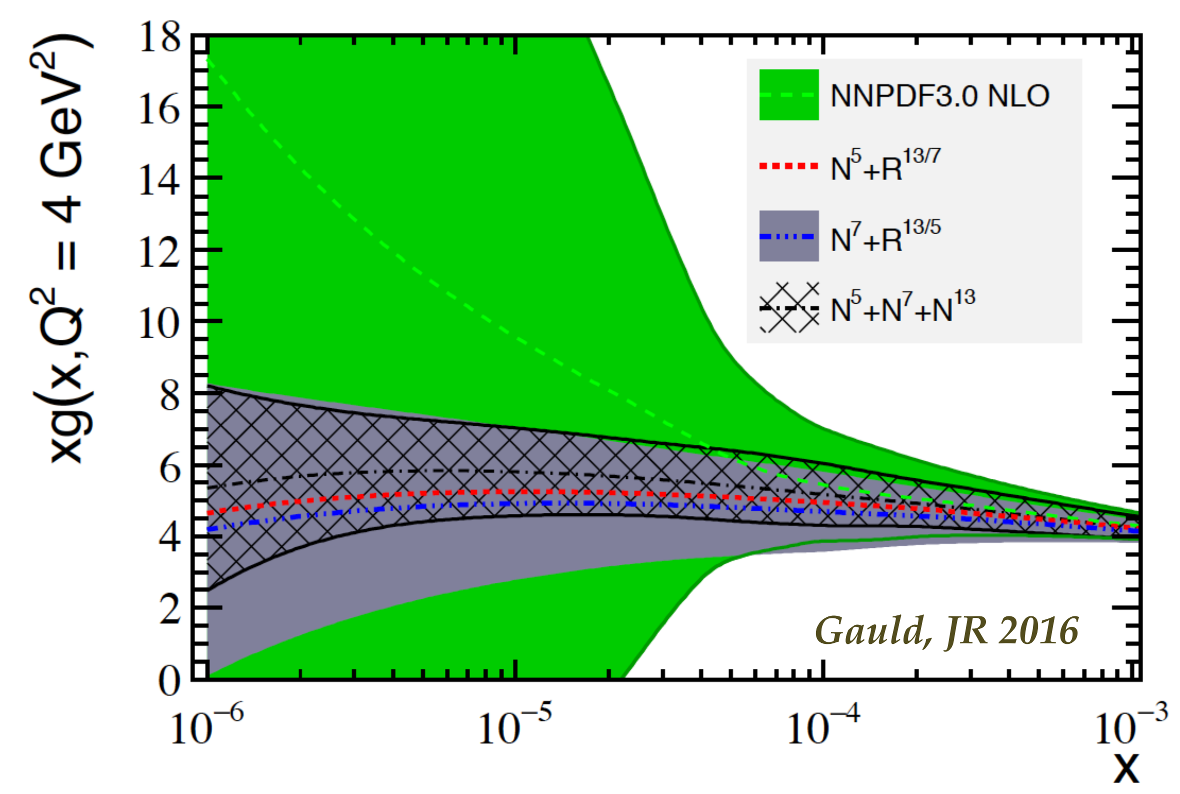
- **D and B meson production from LHCb** allow accessing the **gluon down to x=10^{-6}**, well below the HERA coverage

PROSA 2015, Gauld et al 2015

- Gluon PDF errors **reduced by up to a factor 10!**

- Allows robust estimate for the ***prompt neutrino flux***, the main background for astrophysical neutrinos at IceCube

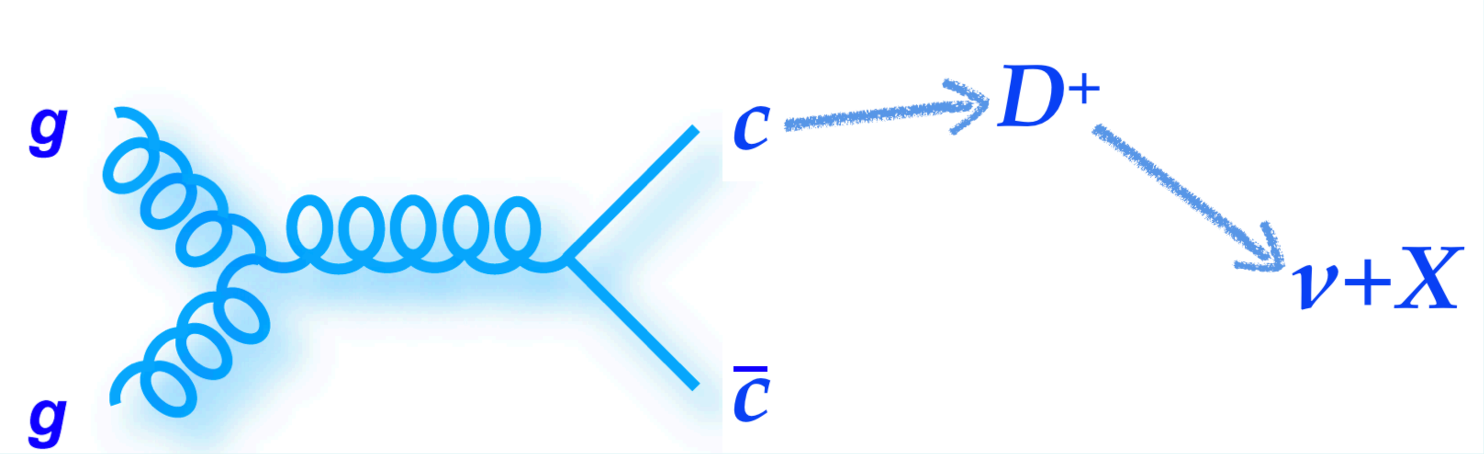
- Precision calculation of the **UHE neutrino-nucleus cross-section**, with few-percent TH errors up to $E_\nu=10^{12}$ GeV

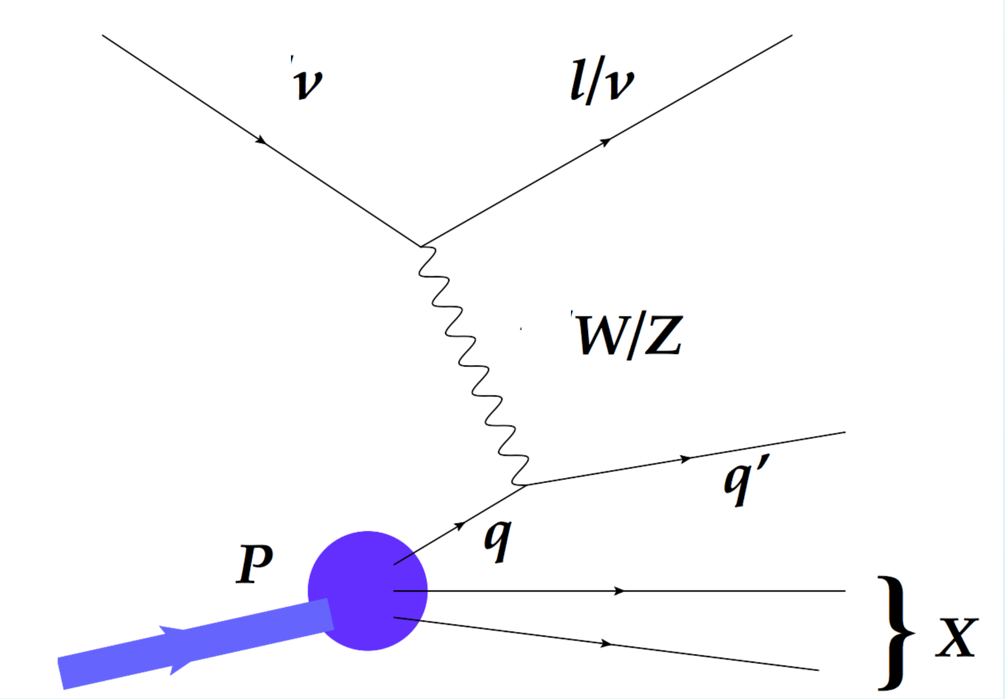


Gauld, JR 2016

Prompt neutrino flux at IceCube

UHE neutrino-nucleus xsecs





The small-x gluon from forward charm production

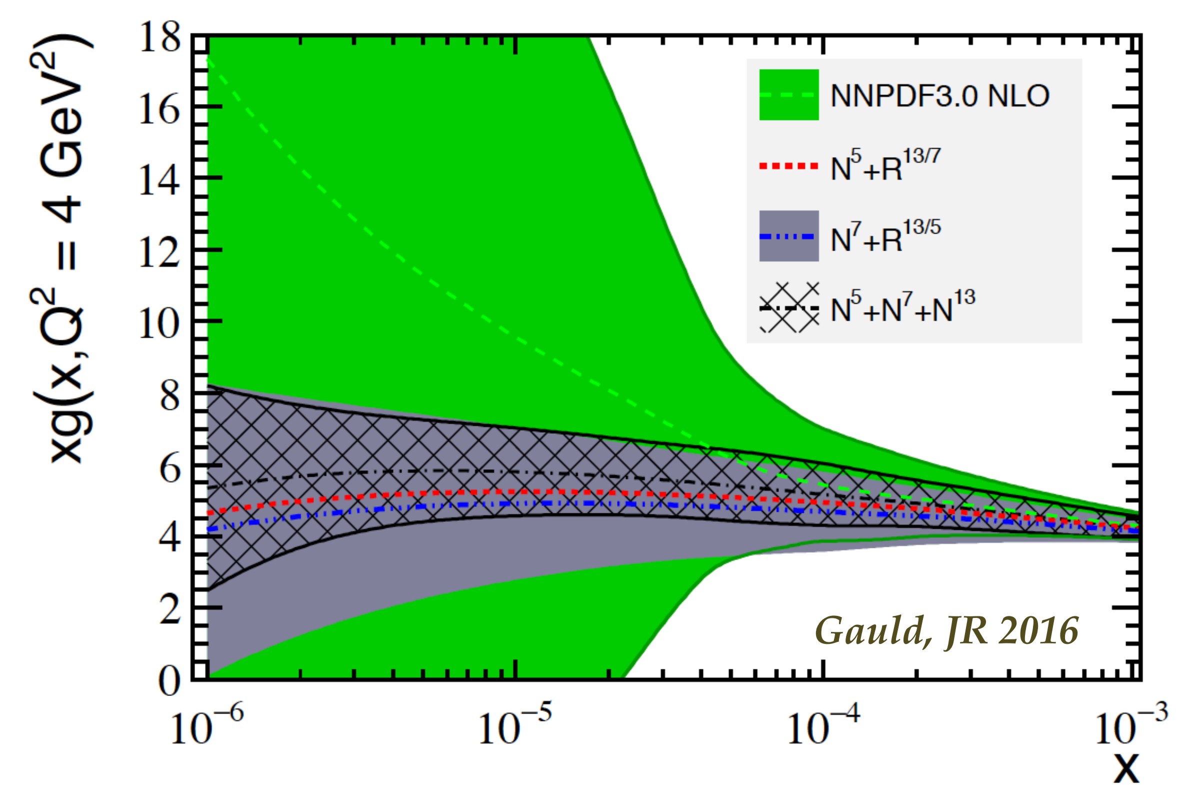
- **D and B meson production from LHCb** allow accessing the **gluon down to x≈10^{-6}**, well below the HERA coverage

PROSA 2015, Gauld et al 2015

- Gluon PDF errors **reduced by up to a factor 10!**

- Allows robust estimate for the ***prompt neutrino flux***, the main background for astrophysical neutrinos at IceCube

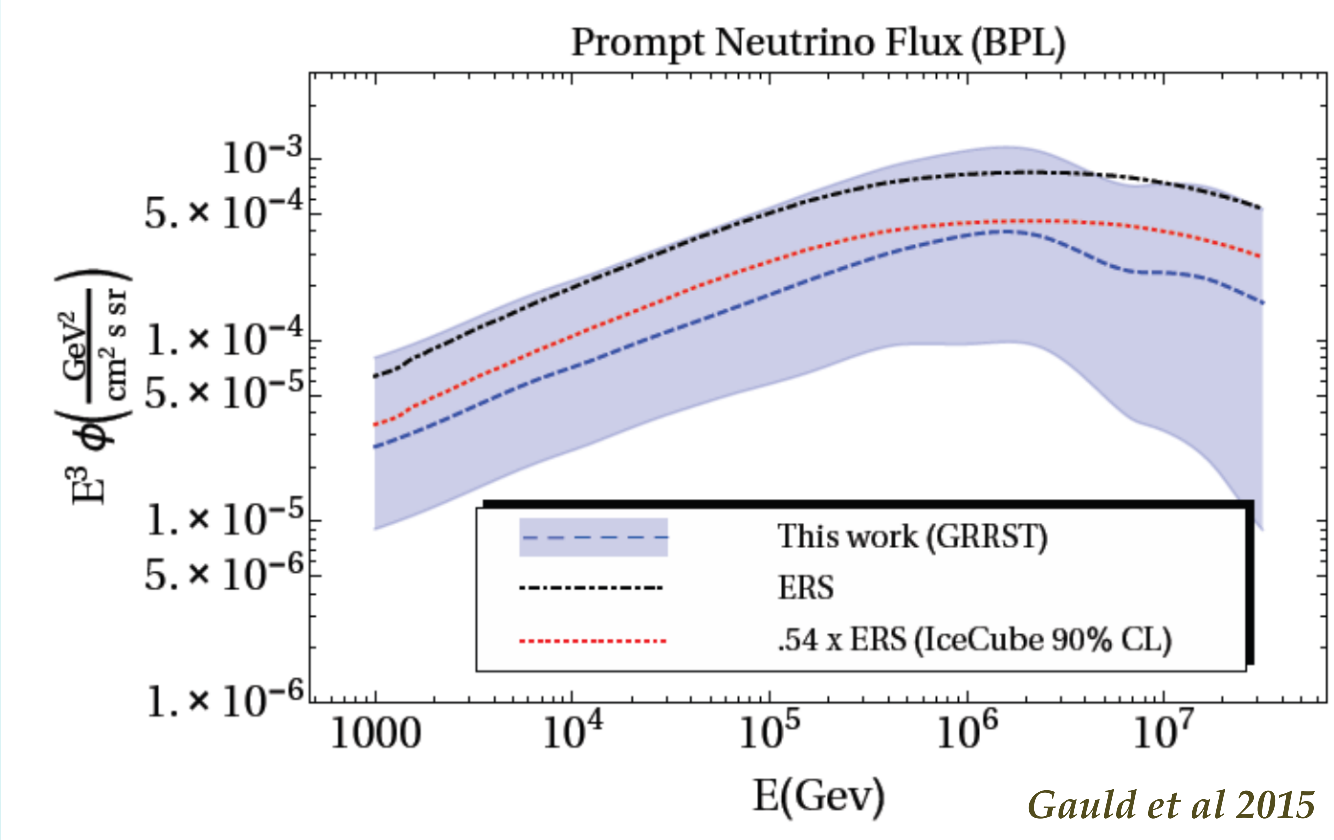
- Precision calculation of the **UHE neutrino-nucleus cross-section**, with few-percent TH errors up to $E_\nu=10^{12}$ GeV



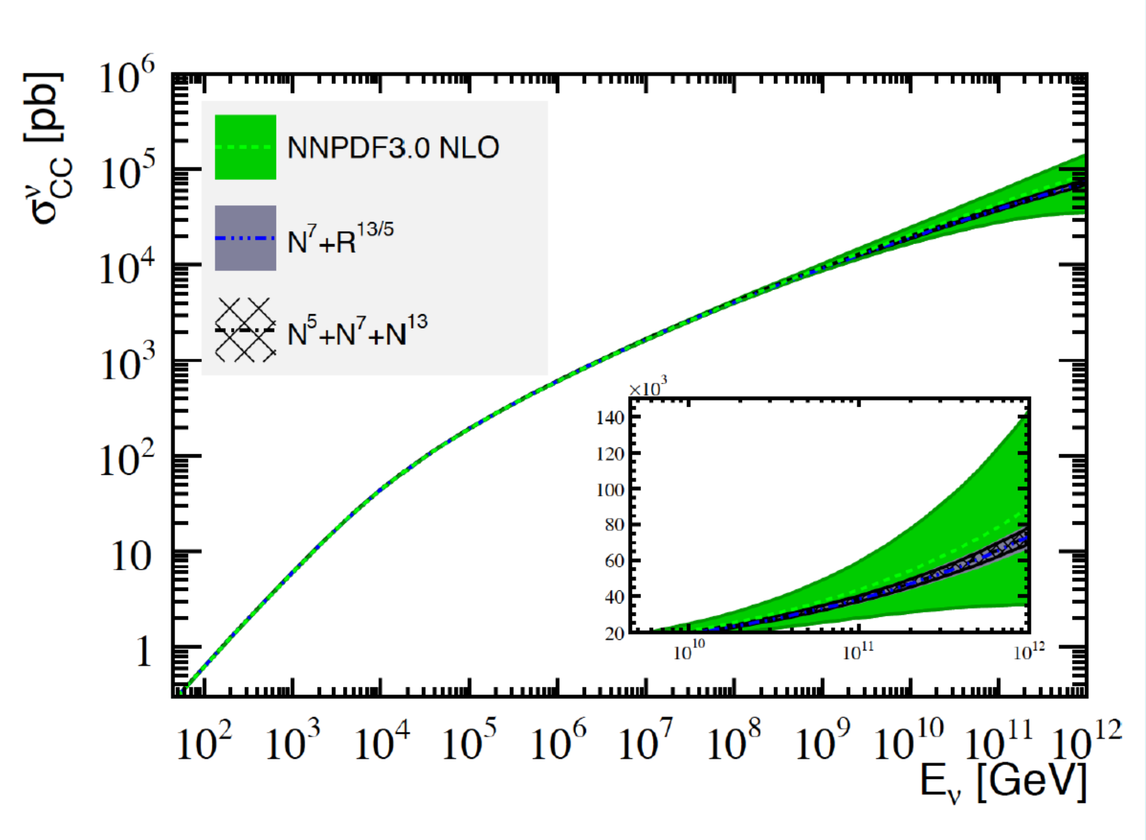
Gauld, JR 2016

Prompt neutrino flux at IceCube

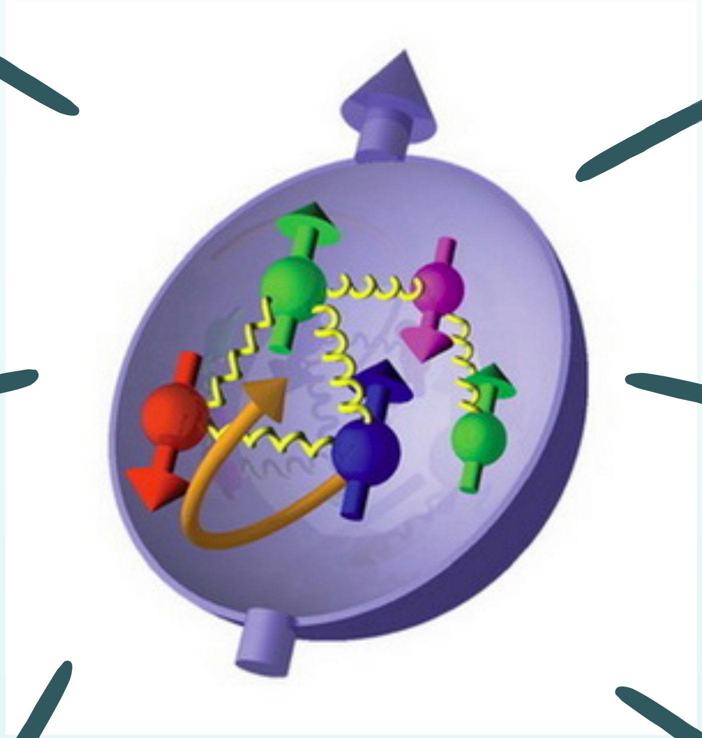
UHE neutrino-nucleus xsecs



Gauld et al 2015



The inner life of protons



Gluon

Strong coupling

Photon

BFKL dynamics

Strange

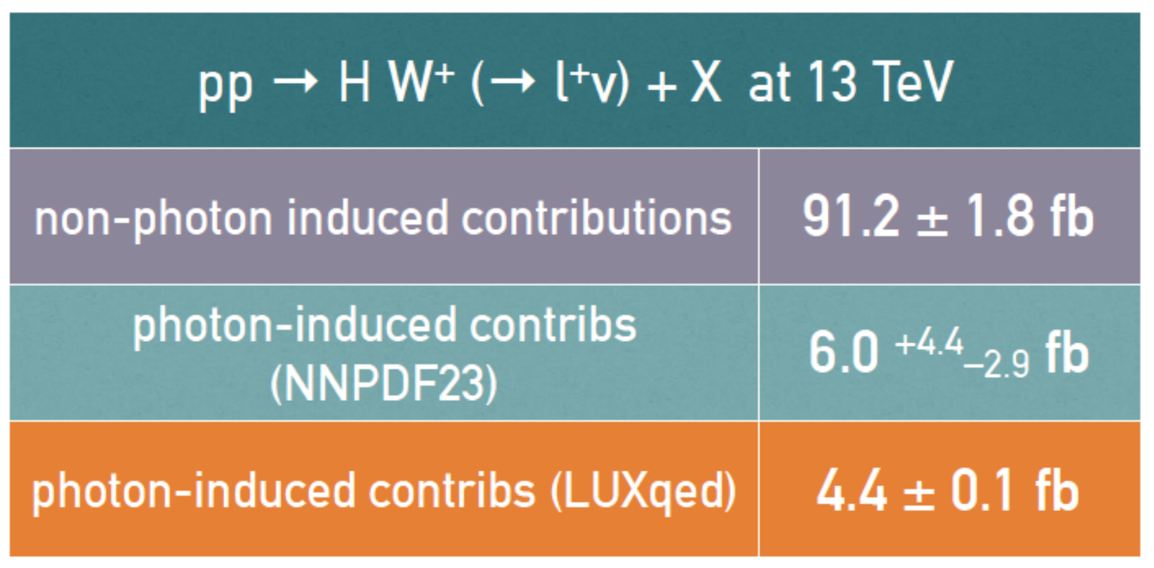
Charm

How bright is the proton?

- The calculation of **QED and electroweak corrections to hadron collider processes** requires by consistency to introduce the PDF of the photon in the proton, **$\gamma(x,Q)$**
- The first model-independent determination of **$\gamma(x,Q)$** from LHC *W,Z* data was **NNPDF2.3QED**, which however affected by **large uncertainties**, due to the limited experimental information
- Recently, **$\gamma(x,Q)$ computed** in terms of the well-known **inclusive structure functions F_2 and F_L**: the resulting photon PDF, exhibits now **few-percent uncertainties**

$$x\gamma(x,\mu) = \frac{1}{2\pi\alpha(\mu)} \int_x^1 \frac{dz}{z} \left\{ \int_{Q^2_{\rm min}}^{\mu^2/(1-z)} \frac{dQ^2}{Q^2} \alpha^2(Q^2) \left[-z^2 F_L(x/z,Q^2) + \left(zP_{\gamma q}(z) + \frac{2x^2 m_p^2}{Q^2} \right) F_2(x/z,Q^2) \right] - \alpha^2(\mu) z^2 F_2(x/z,\mu^2) \right\} + \mathcal{O}\left(\alpha\alpha_s, \alpha^2\right)$$

Manohar, Nason, Salam, Zanderighi, 16-17

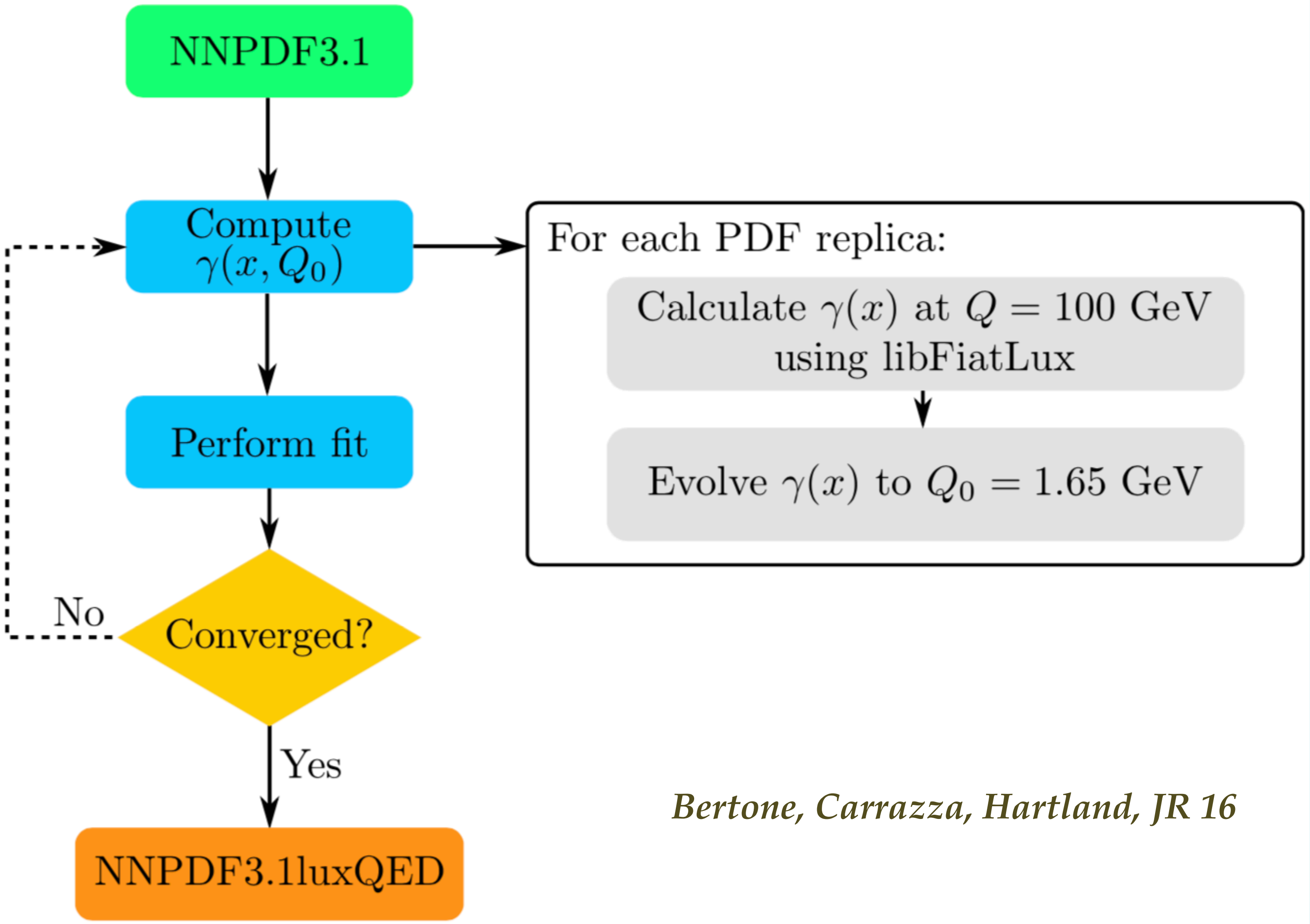


pp → H W^+ (→ $l^+\nu$) + X at 13 TeV	
non-photon induced contributions	91.2 ± 1.8 fb
photon-induced contribs (NNPDF23)	$6.0^{+4.4}_{-2.9}$ fb
photon-induced contribs (LUXqed)	4.4 ± 0.1 fb

Crucial implications for LHC pheno: high-precision determination of photon-initiated (PI) contributions

Illuminating the photon content of the proton

- **NNPDF3.1luxQED**: variant of the NNPDF3.1 global analysis supplemented by
  - the LUXqed theoretical constraints,
  - NLO QED corrections to DGLAP evolution, and
  - NLO QED coefficients functions in DIS
- Iterative procedure: photon PDF recomputed at each iteration until convergence is achieved



NNPDF3.1

Compute $\gamma(x, Q_0)$

For each PDF replica:

Calculate $\gamma(x)$ at $Q = 100$ GeV using libFiatLux

Evolve $\gamma(x)$ to $Q_0 = 1.65$ GeV

Perform fit

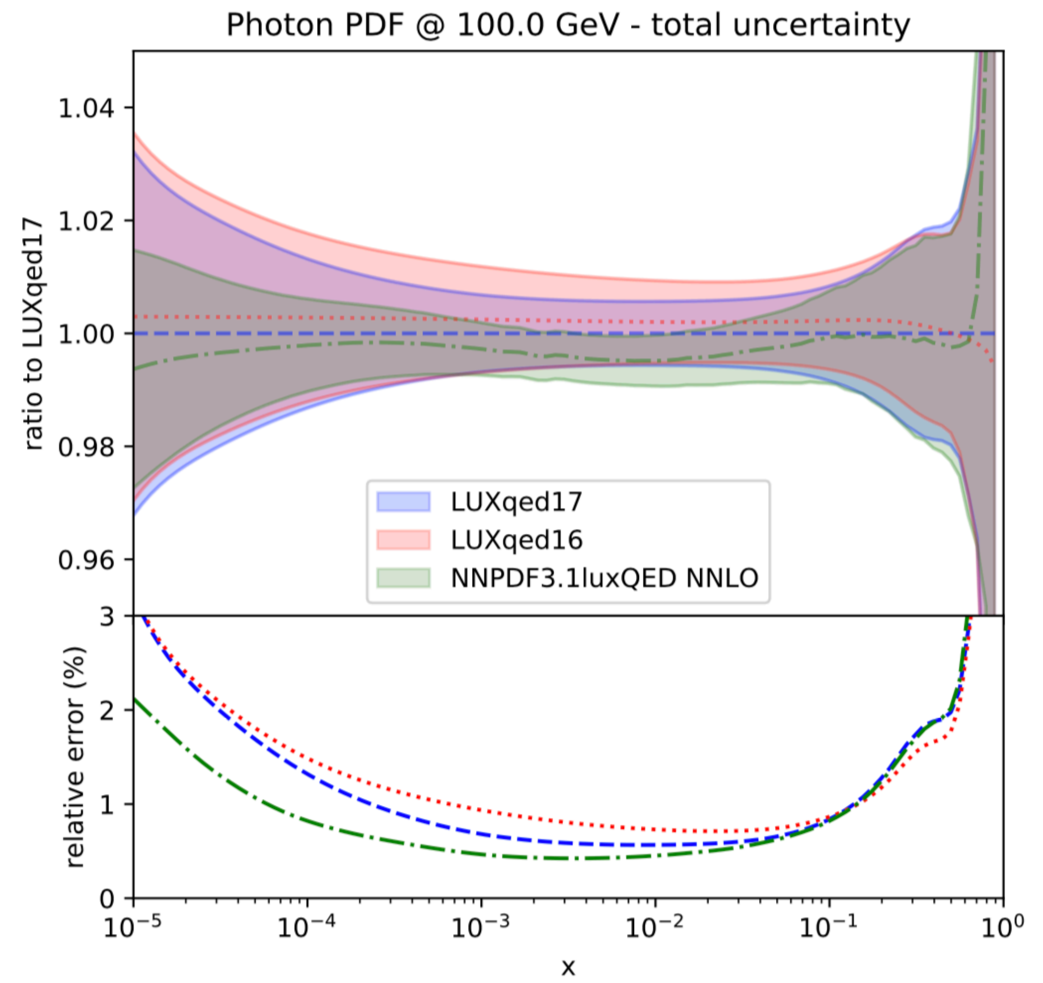
Converged?

No

Yes

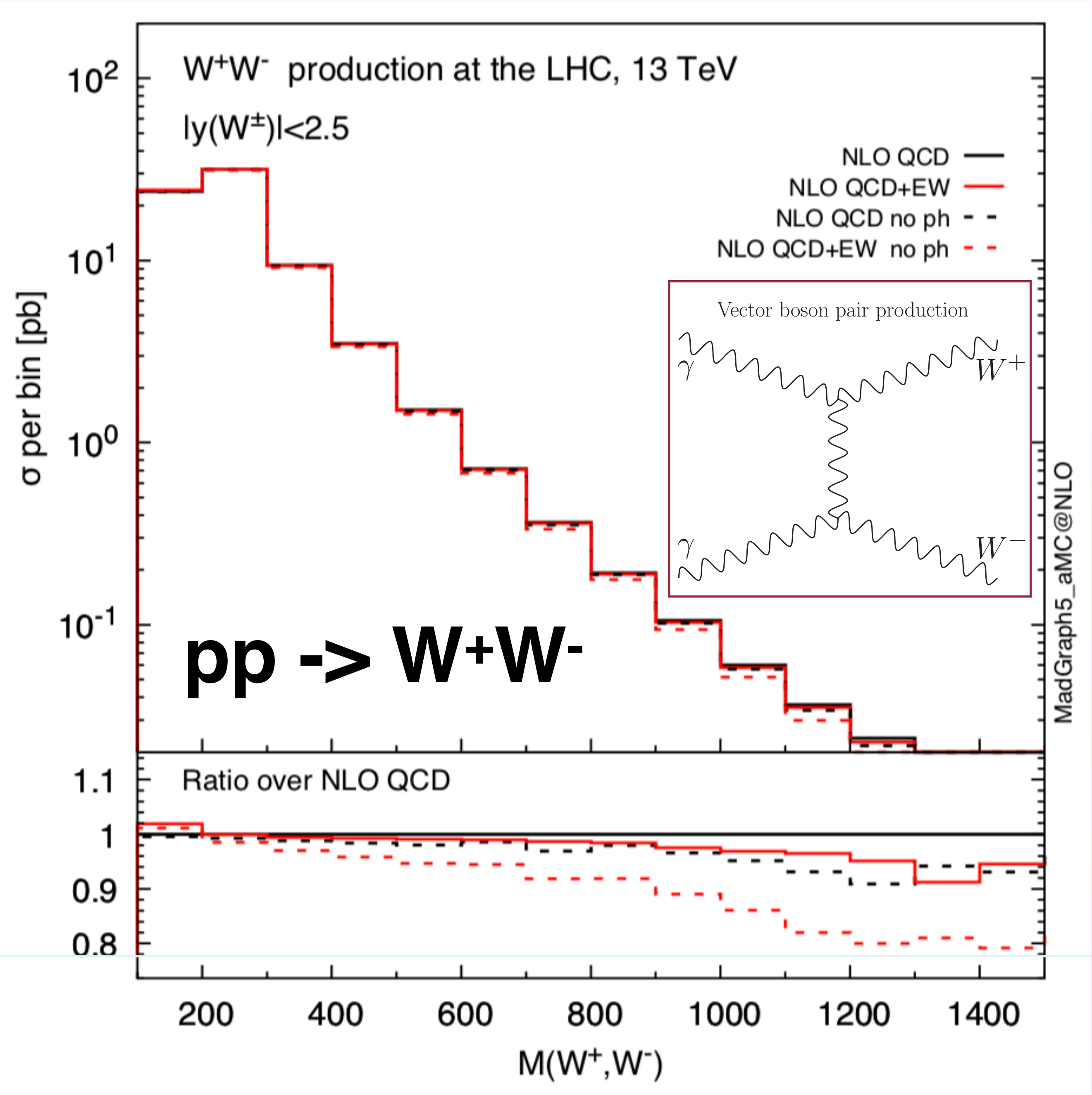
NNPDF3.1luxQED

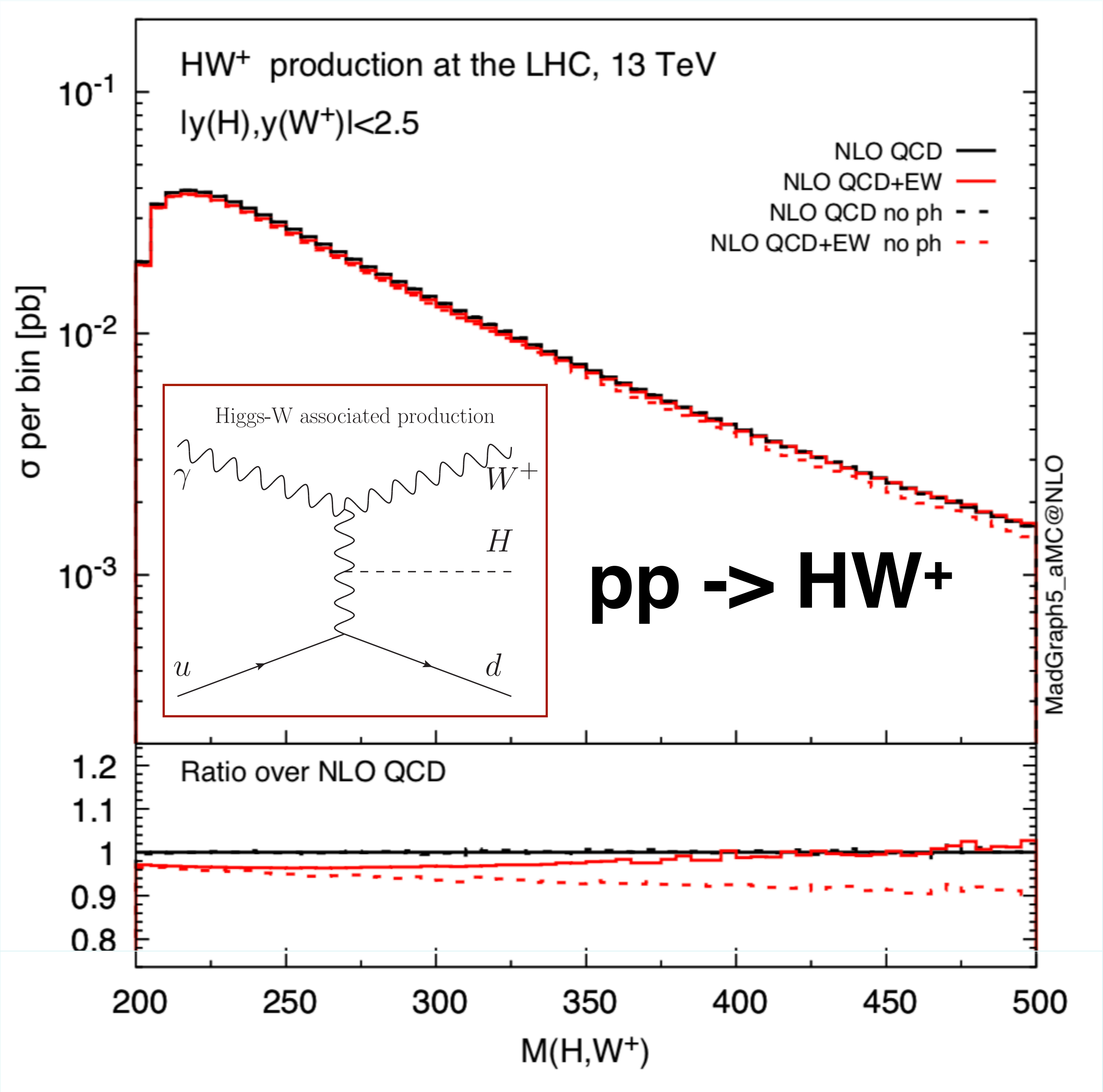
Bertone, Carrazza, Hartland, JR 16



Photon-initiated processes at the LHC

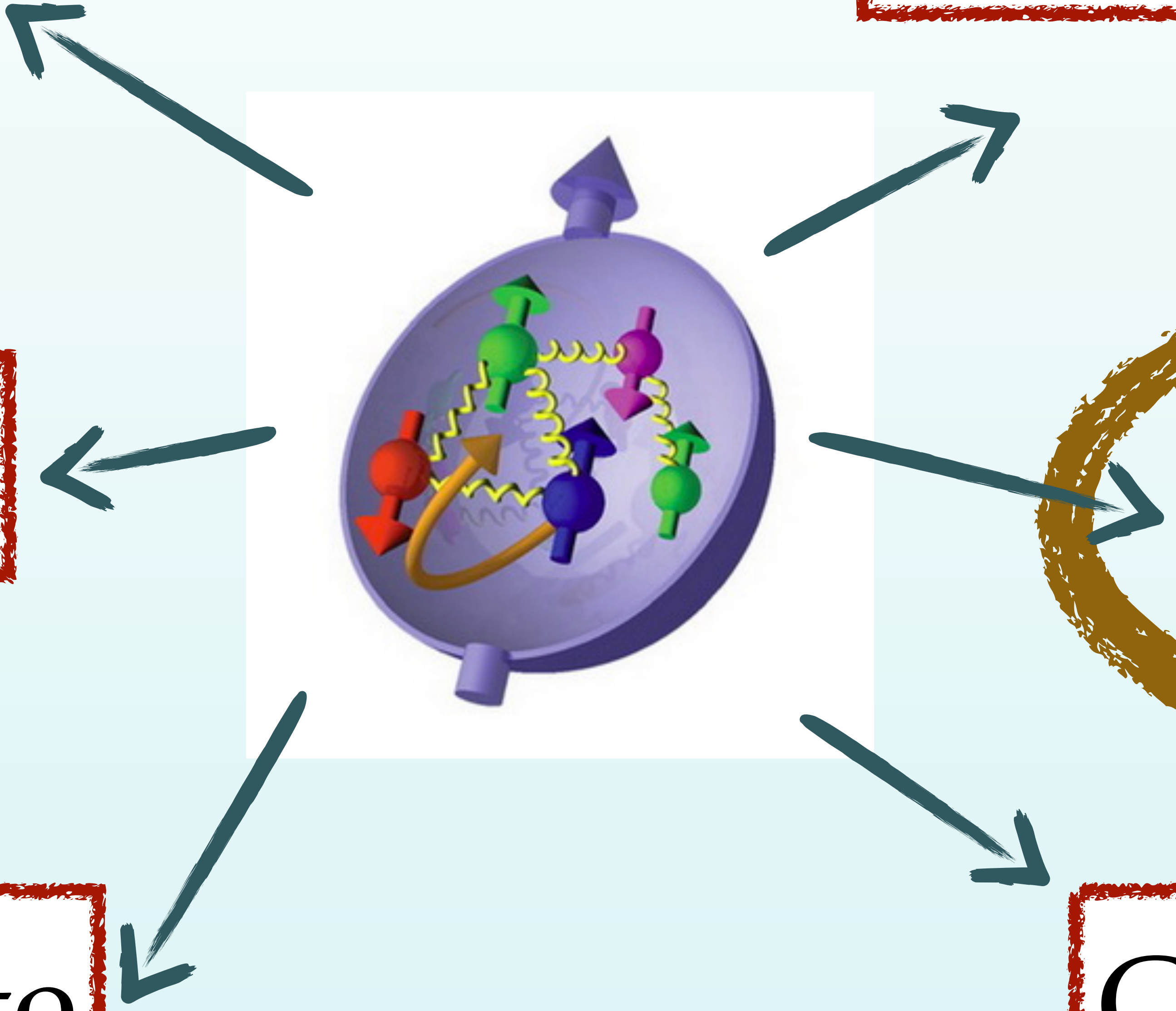
- At high mass **PI contributions and NLO EW corrections have opposite sign**, and thus in general one expects a partial cancellation among them
- This seems to be the case for many processes: **once PI effects included, NLO EW corrections rather small**





Bertone, Carrazza, Pagani, Rojo, Vicini, Zaro (in preparation)

The inner life of protons



Gluon

Strong coupling

Photon

BFKL dynamics

Strange

Charm

Parton distributions with BFKL resummation

- **Perturbative fixed-order QCD calculations** have been extremely successful in describing a wealth of data from proton-proton and electron-proton collisions

- There are theoretical reasons that eventually we need to go beyond DGLAP: at small-x, **logarithmically enhanced terms in $1/x$ become dominant** and need to be resummed to all orders

- ***BFKL/high-energy/small-x resummation*** can be matched to the **DGLAP collinear framework**, and thus be included into a standard PDF analysis

DGLAP Evolution in Q²

$$\mu^2 \frac{\partial}{\partial \mu^2} f_i(x,\mu^2) = \int_x^1 \frac{dz}{z} P_{ij}\left(\frac{x}{z}, \alpha_s(\mu^2)\right) f_j(z,\mu^2),$$

BFKL Evolution in x

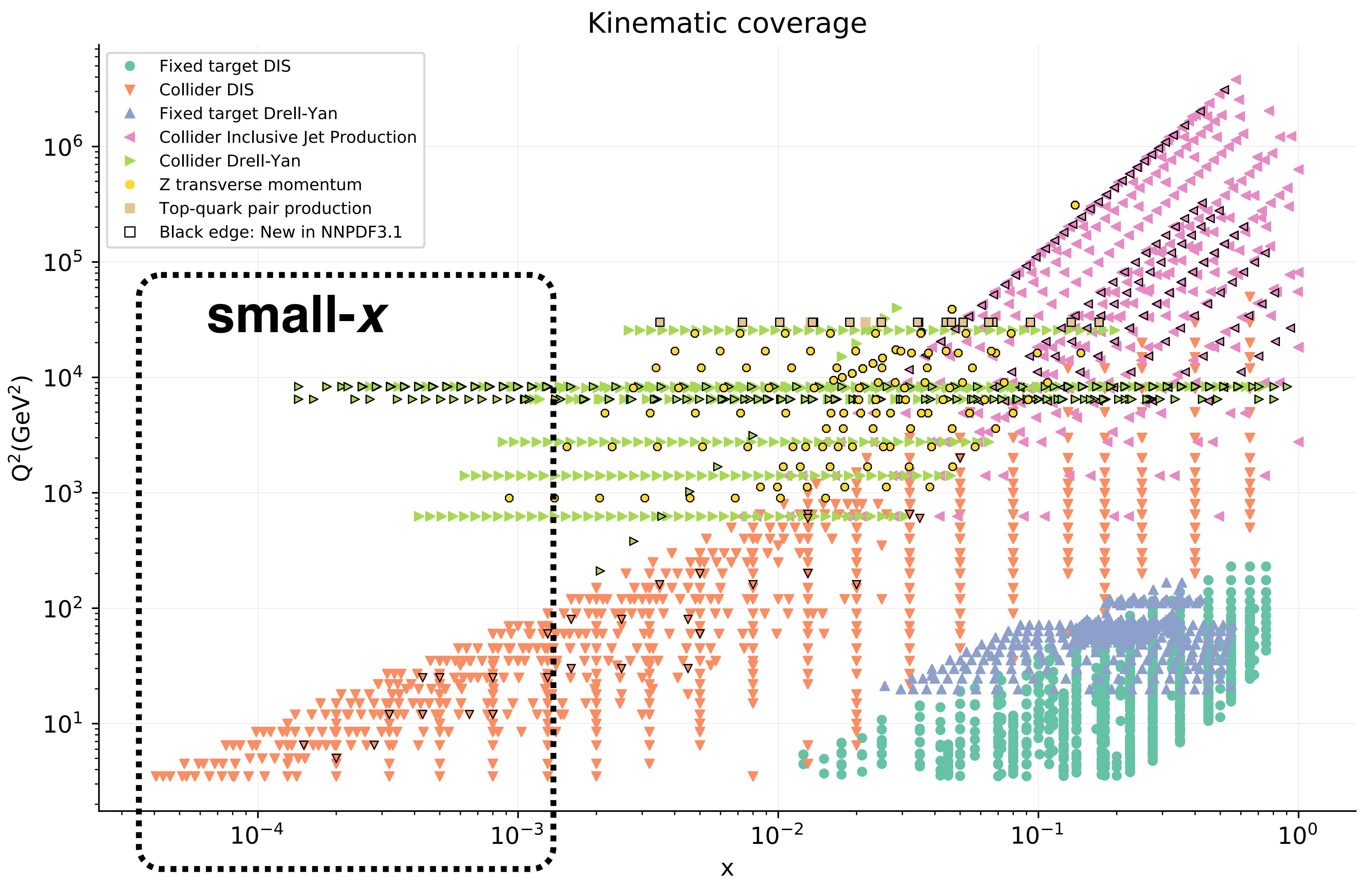
$$-x\frac{d}{dx} f_+(x,\mu^2) = \int_0^\infty \frac{d\nu^2}{\nu^2} K\left(\frac{\mu^2}{\nu^2}, \alpha_s\right) f_+(x,\nu^2)$$

Within small-x resummation, the **N^{k}LO fixed-order DGLAP splitting functions** are complemented with the **N^{h}LLx contributions from BKFL**

ABF, CCSS, TW + others, 94-08

$$P_{ij}^{\mathrm{N}^k\mathrm{LO}+\mathrm{N}^h\mathrm{LL}x}(x) = P_{ij}^{\mathrm{N}^k\mathrm{LO}}(x) + \Delta_k P_{ij}^{\mathrm{N}^h\mathrm{LL}x}(x),$$

A new world at small-x

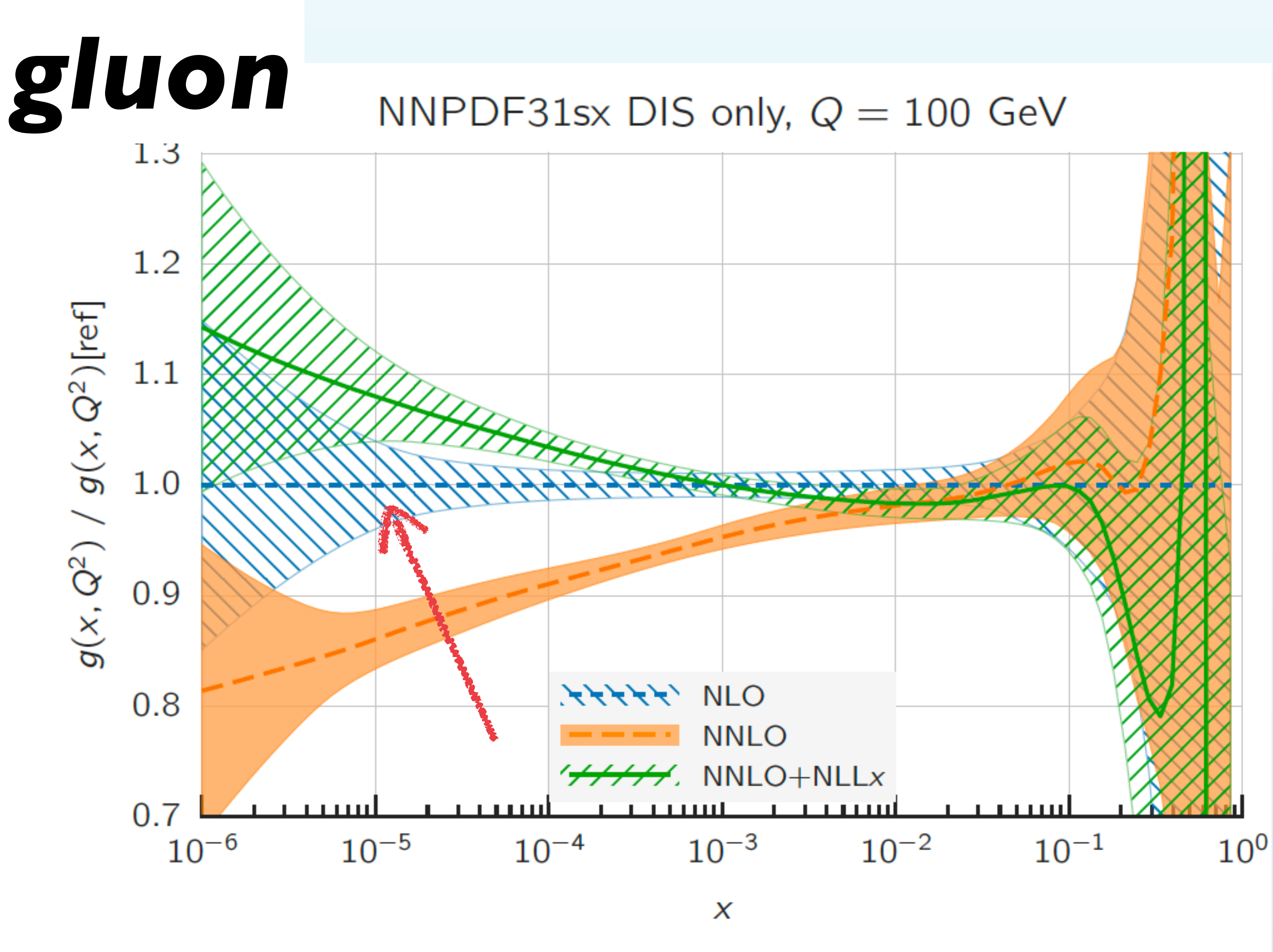


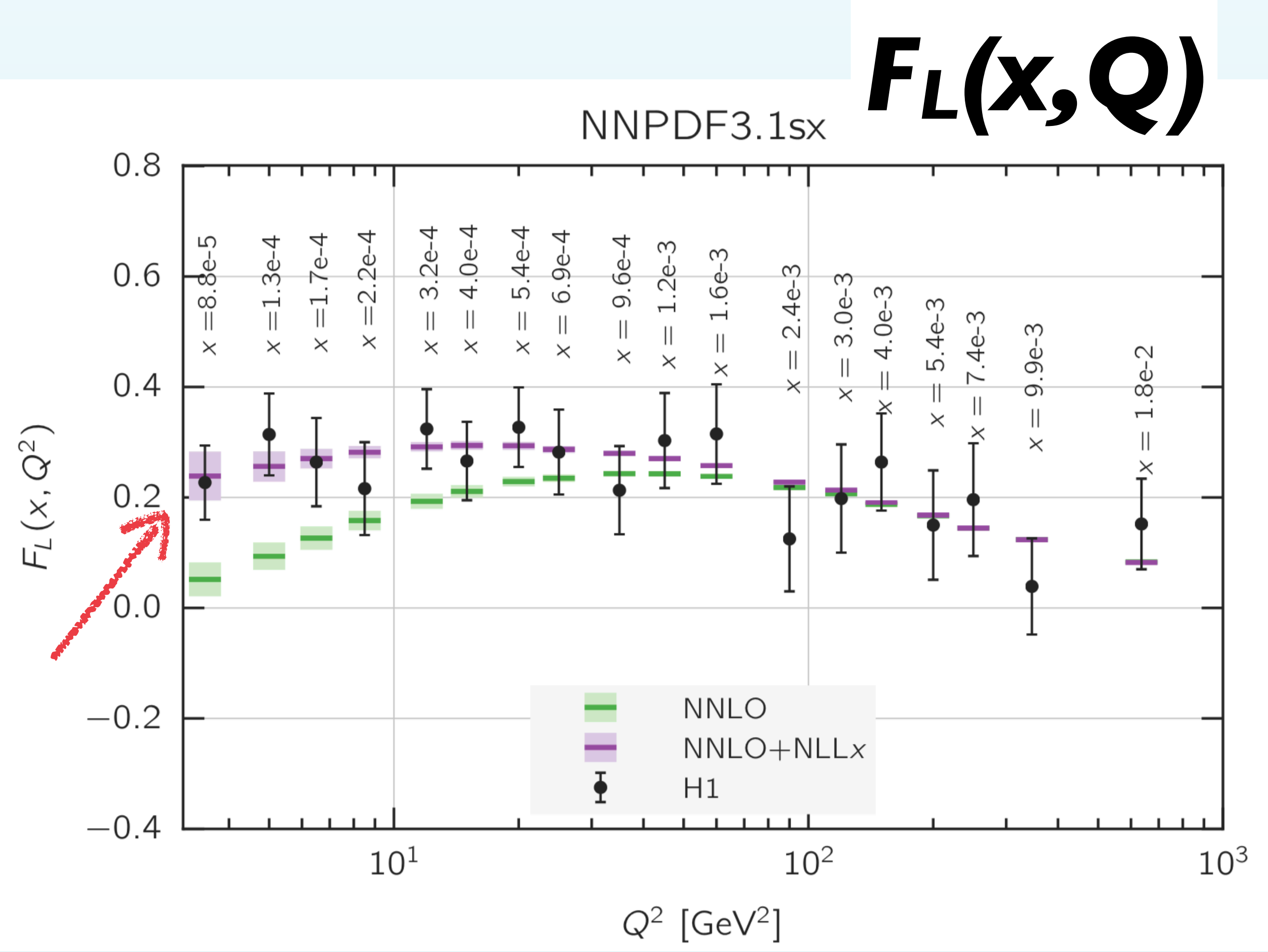
A new world at small-x

- Ultimately, the need for (or lack of) BKFL resummation in **ep and pp collider data** can only be assessed by performing a **global PDF analysis based on (N)NLO+NLLx theory**

- Theoretical tools are now available: **HELL for NLL*x* resummation**, interfaced to **APFEL**

- NNPDF3.1 (N)NLO+NLL fits **stabilize the perturbative PDF expansion at small-*x***, in particular for the gluon, and markedly improve the **fit quality to the small-*x* HERA data**

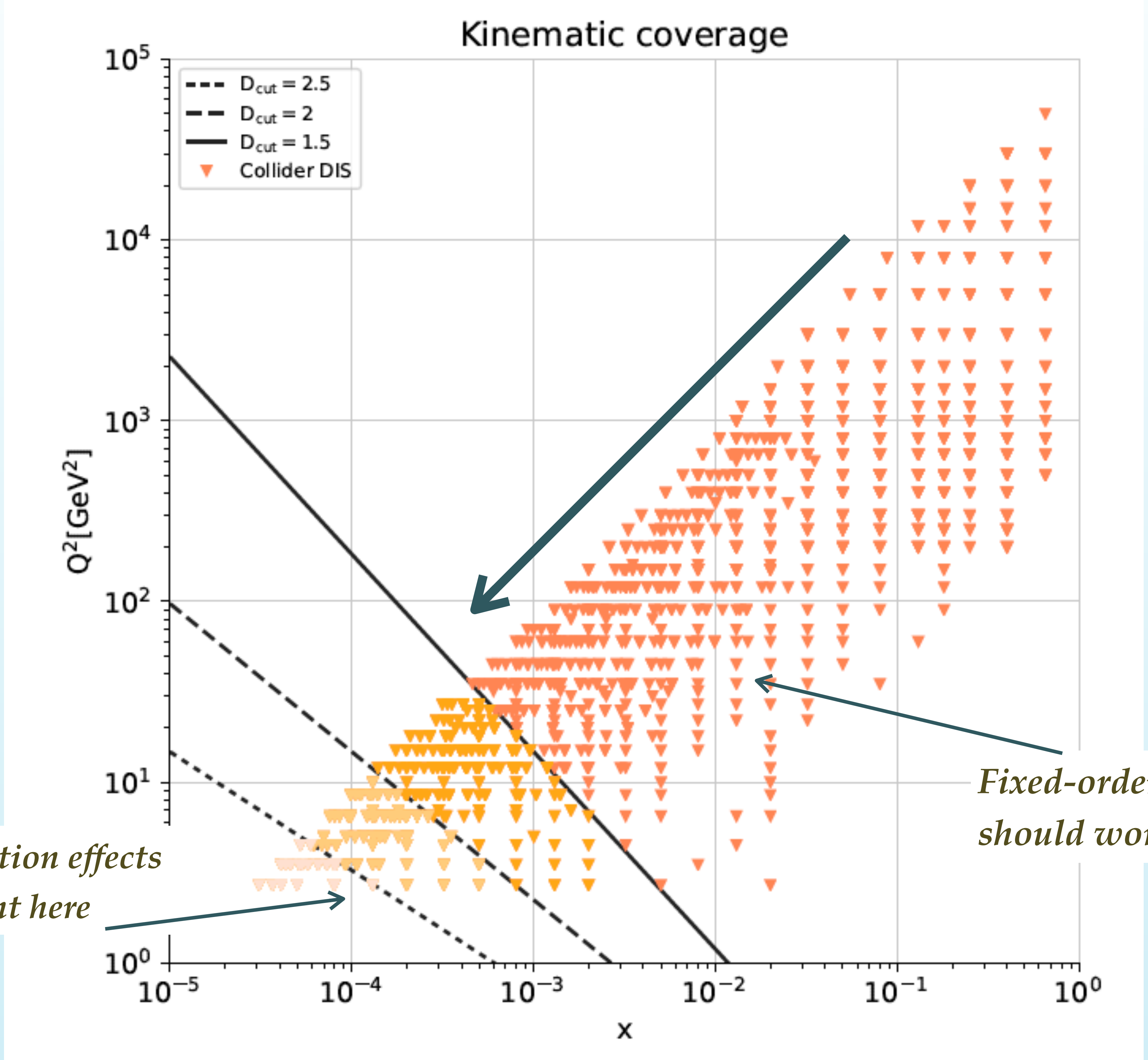




Ball, Bertone, Bonvini, Marzani, JR,Rottoli 16

Evidence for BFKL dynamics in HERA data

In order to assess the impact of small-x resummation for the description of the small-x and Q^2 HERA data, compute the χ^2 **removing data points in the region where resummation effects are expected**



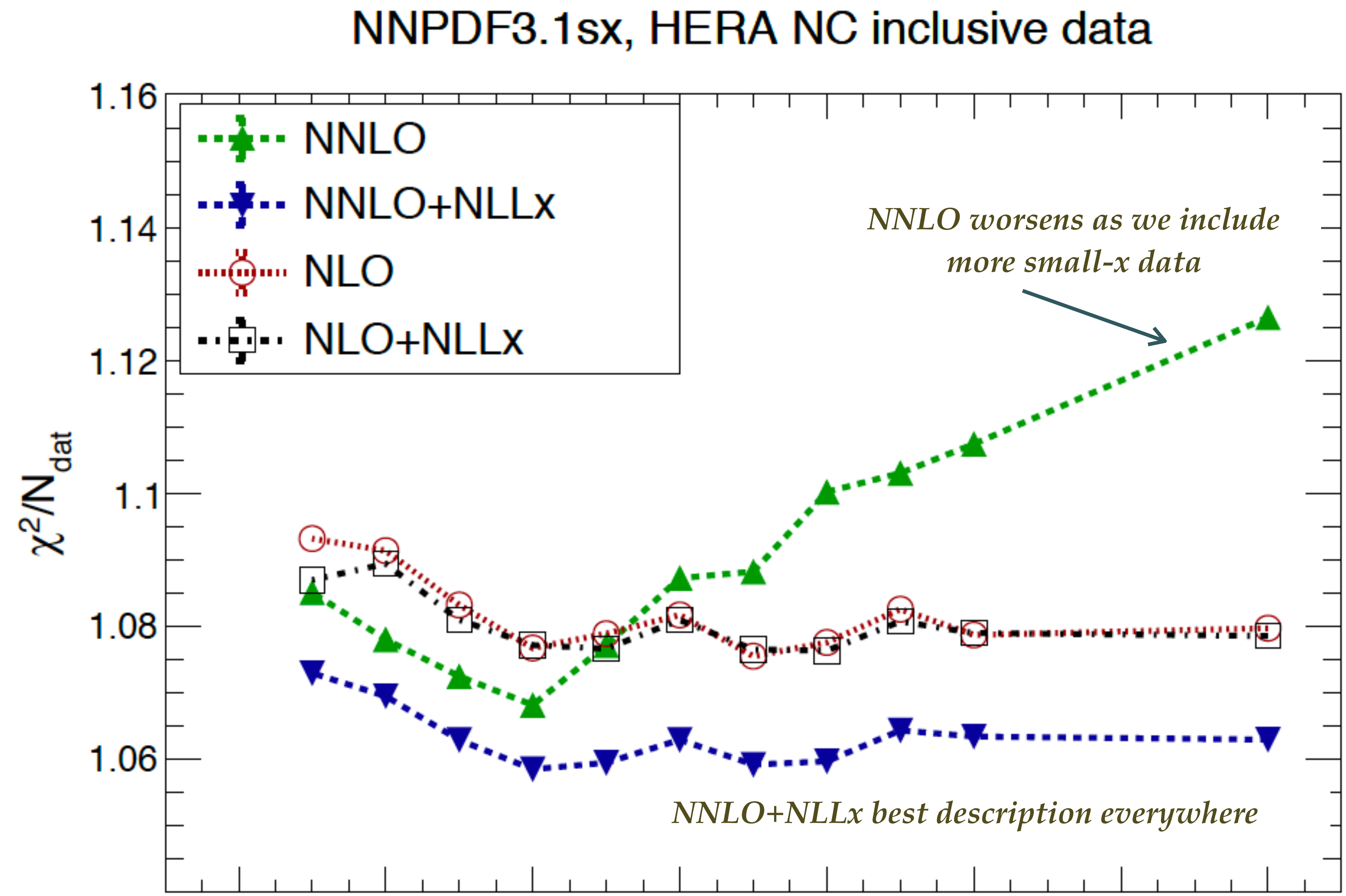
Fixed-order theory should work fine here

Small-x resummation effects could be important here

Evidence for BFKL dynamics in HERA data

Using **NNLO+NLL*x* theory, the NNLO instability at small-*x* of the χ^2 disappears**

Excellent fit quality to **inclusive and charm HERA** data achieved in the **entire (x,Q²) region**



NNLO worsens as we include more small-x data

NNLO+NLLx best description everywhere

Including more low-x, low Q data

Nunca es tarde si la dicha es buena

Science
Life and Physics

After 40 years of studying the strong nuclear force, a revelation

This was the year that analysis of data finally backed up a prediction, made in the mid 1970s, of a surprising emergent behaviour in the strong nuclear force

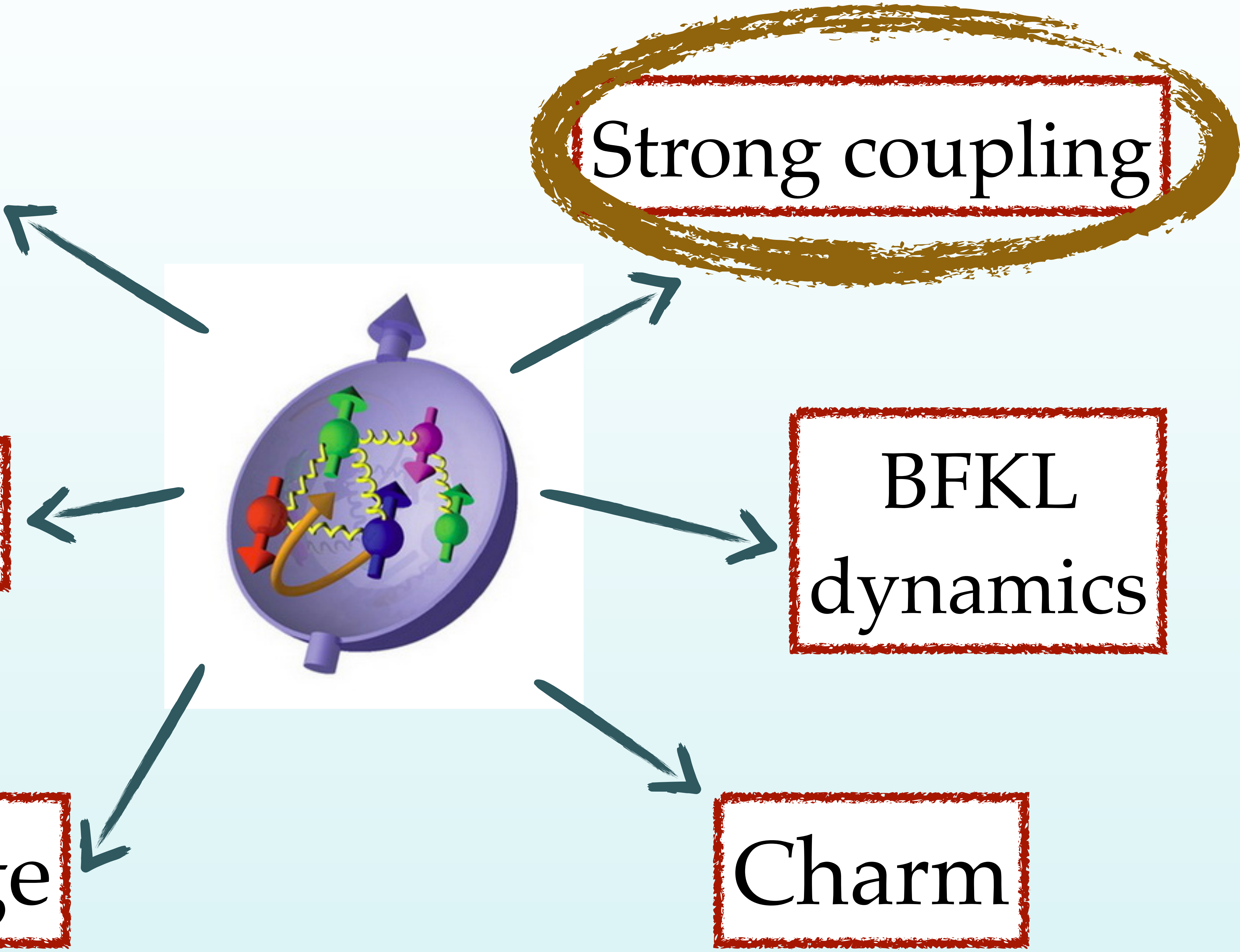
Jon Butterworth
@jonmbutterworth
Thu 28 Dec 2017 17.30 GMT

529 59

In the mid 1970s, four Soviet physicists, Batlisky, Fadin, Kuraev and Lipatov, made some predictions involving the strong nuclear force which would lead to their initials entering the lore. “BFKL” became a shorthand for a difficult-to-

Jon Butterworth,
the Guardian, 28/12/2018

The inner life of protons



Gluon

Strong coupling

Photon

BFKL dynamics

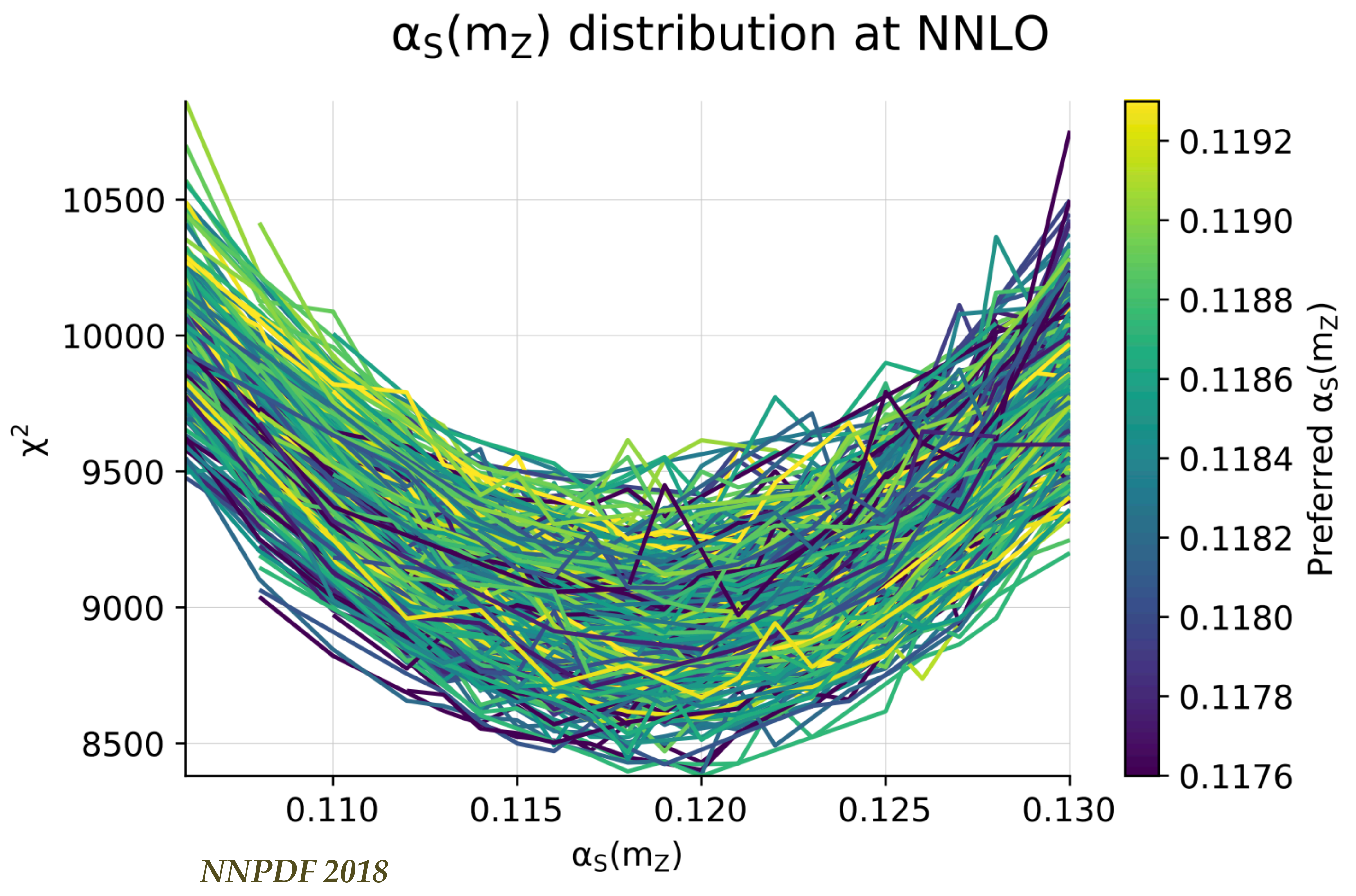
Strange

Charm

Precision determination of $\alpha_S(m_Z)$

- Determination of the strong coupling constant using the new **Monte Carlo correlated replica** method

- In a nutshell, carry out an **independent $\alpha_S(m_Z)$ fit for each MC replica** to determine its complete probability distribution accounting for all correlations with the PDFs



NNPDF 2018

Precision determination of $\alpha_S(m_Z)$

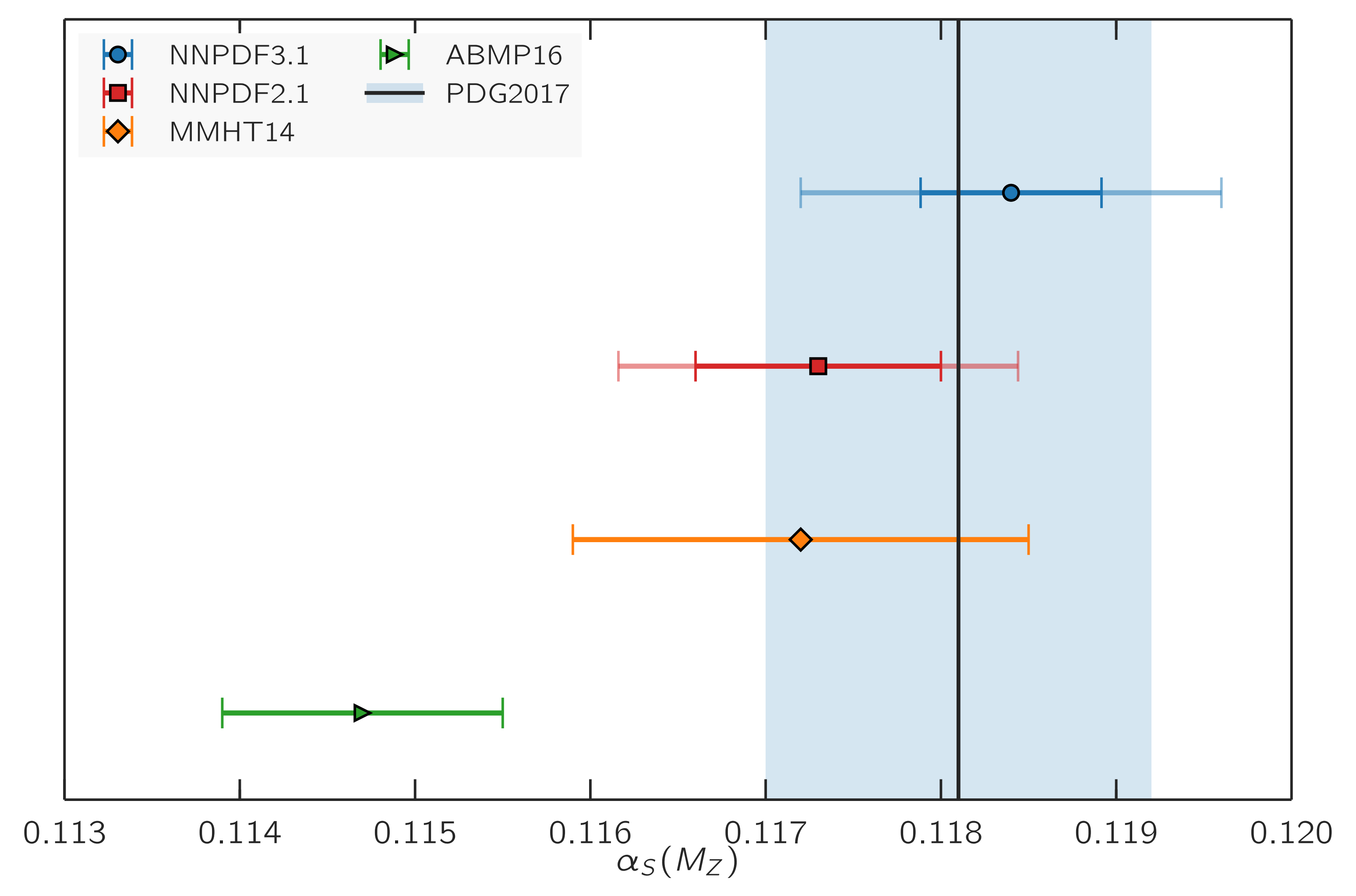
- Fit of $\alpha_S(m_Z)$ based NNPDF3.1NNLO, where **all collider processes are treated using exact NNLO theory**

- First ever determination of the strong coupling based on a **global NNLO analysis** with inclusive jet, top quark pair production, Z transverse momentum distributions: various **complementary handles!**

$$\alpha_s^{\rm NNLO}(m_Z) = 0.1185 \pm 0.0005^{\rm exp} \ (0.4\%)$$

$$\Delta\alpha_s^{\rm pert} \equiv |\alpha_s^{\rm NNLO} - \alpha_s^{\rm NLO}| = 0.0022$$

$$\delta\alpha_s^{\rm mhou} \simeq 0.0011$$

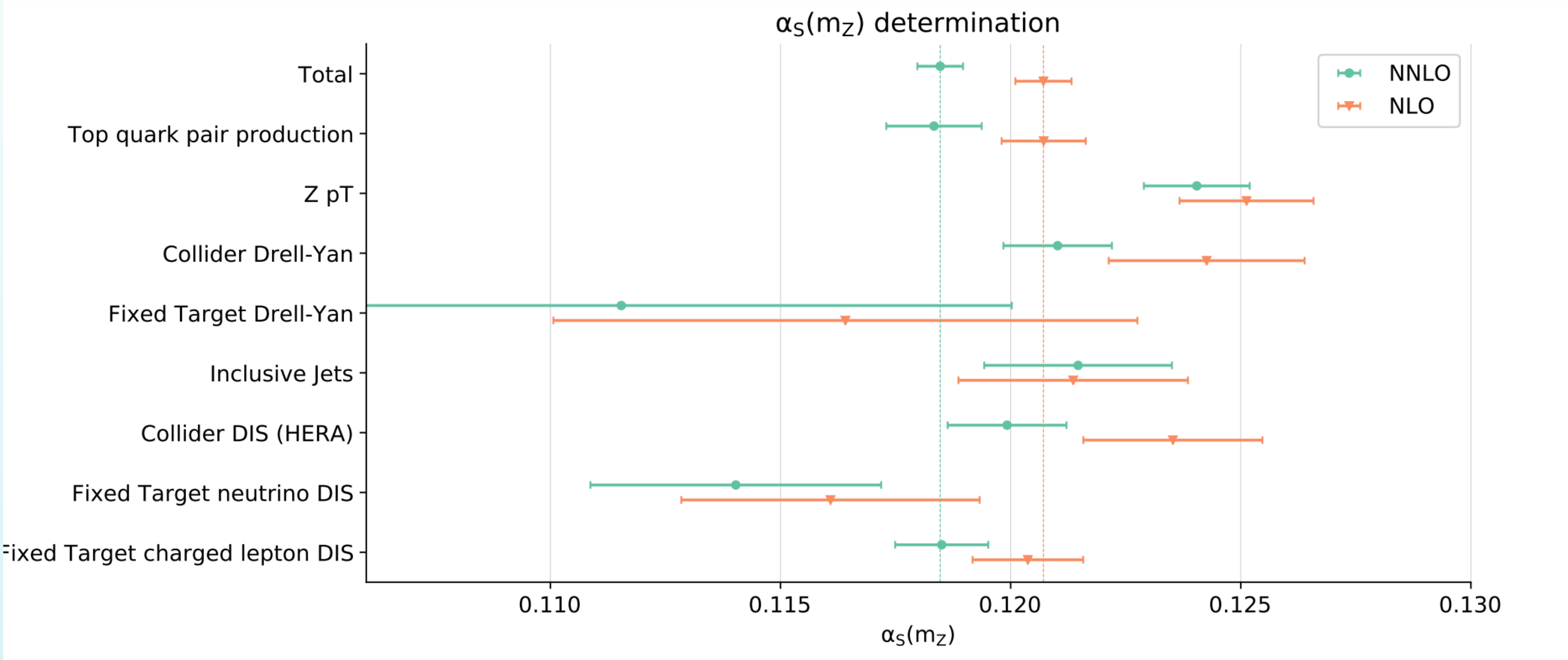


Precision determination of $\alpha_S(m_Z)$

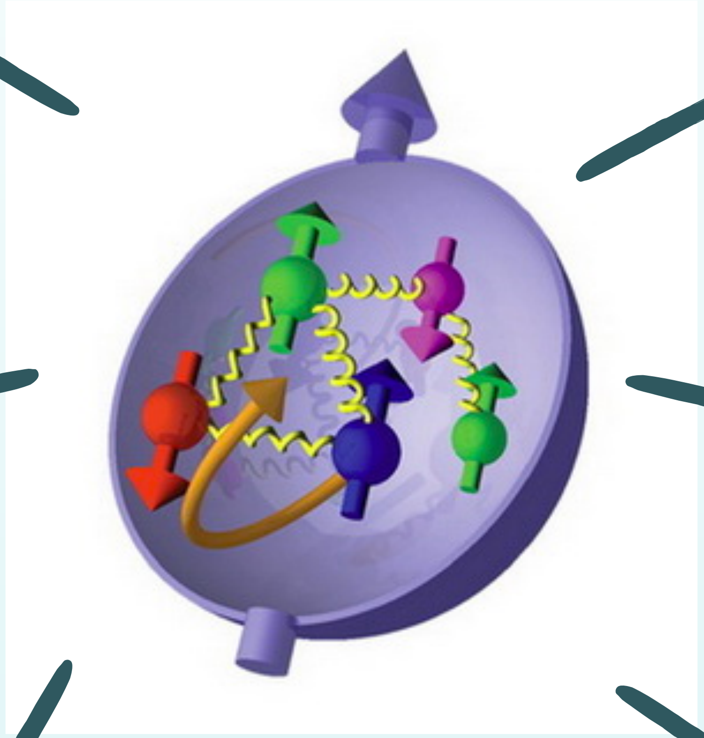
- Significant constraints from top quark pair, *Z pT*, and inclusive jets ….

- But also from fixed-target NC DIS structure functions and from collider DIS and DY data

- Main limitation of the fit is the **poor control over the MHOUs**: once these are accounted for, the global PDF fit would become **one of the most competitive methods** to extract $\alpha_S(m_Z)$



The inner life of protons



Gluon

Strong coupling

Photon

BFKL dynamics

Strange

Charm

A charming story

In global PDF fits there are **two strategies** to treat the **charm PDF:**

- The charm PDF is **generated perturbatively** via collinear splittings **off gluons and light quarks:**

$$c(x,Q)=0\,,\qquad Q<m_c$$
$$c(x,Q)\simeq\alpha_s(Q^2)\ln\frac{Q^2}{m_c^2}\int_x^1\frac{dy}{y}P_{gg}^{(0)}(y)g\left(\frac{y}{x},Q^2\right)\,,\qquad Q\geq m_c$$

- The charm PDF is **fitted from data**, that is, it is treated on an **equal footing to light quarks:**

$$c(x,Q_0)=f(x,\{a_i\})\,,\qquad Q_0>m_c$$

Note that the first option is necessarily an **assumption**, which can only be validated by comparing with the results of a **direct fit of the charm PDF**

Moreover **fitting the charm PDF** offers several advantages:

☑ **Reduce the dependence of high-scale cross-sections** with respect of the value of the **charm mass**

☑ Improve the description of **high-precision collider data** that depend on **quark flavor separation**

☑ Compare with models of the **non-perturbative (``*intrinsic*") charm content of the proton**

***Fitting the charm PDFs** has many advantages even with for a vanishing non-perturbative component*

Some charm really improves things

	NNLO Fitted Charm	NNLO Pert Charm	NLO Fitted Charm	NLO Pert Charm
HERA	1.16	1.21	1.14	1.15
ATLAS	1.09	1.17	1.37	1.45
CMS	1.06	1.09	1.20	1.21
LHCb	1.47	1.48	1.61	1.77
TOTAL	1.148	1.187	1.168	1.197

- In NNPDF3.1, for collider data, **NNLO theory** leads to a markedly better fit quality that than **NLO**

- The new precision data included has small experimental errors, with **NNLO corrections mandatory**

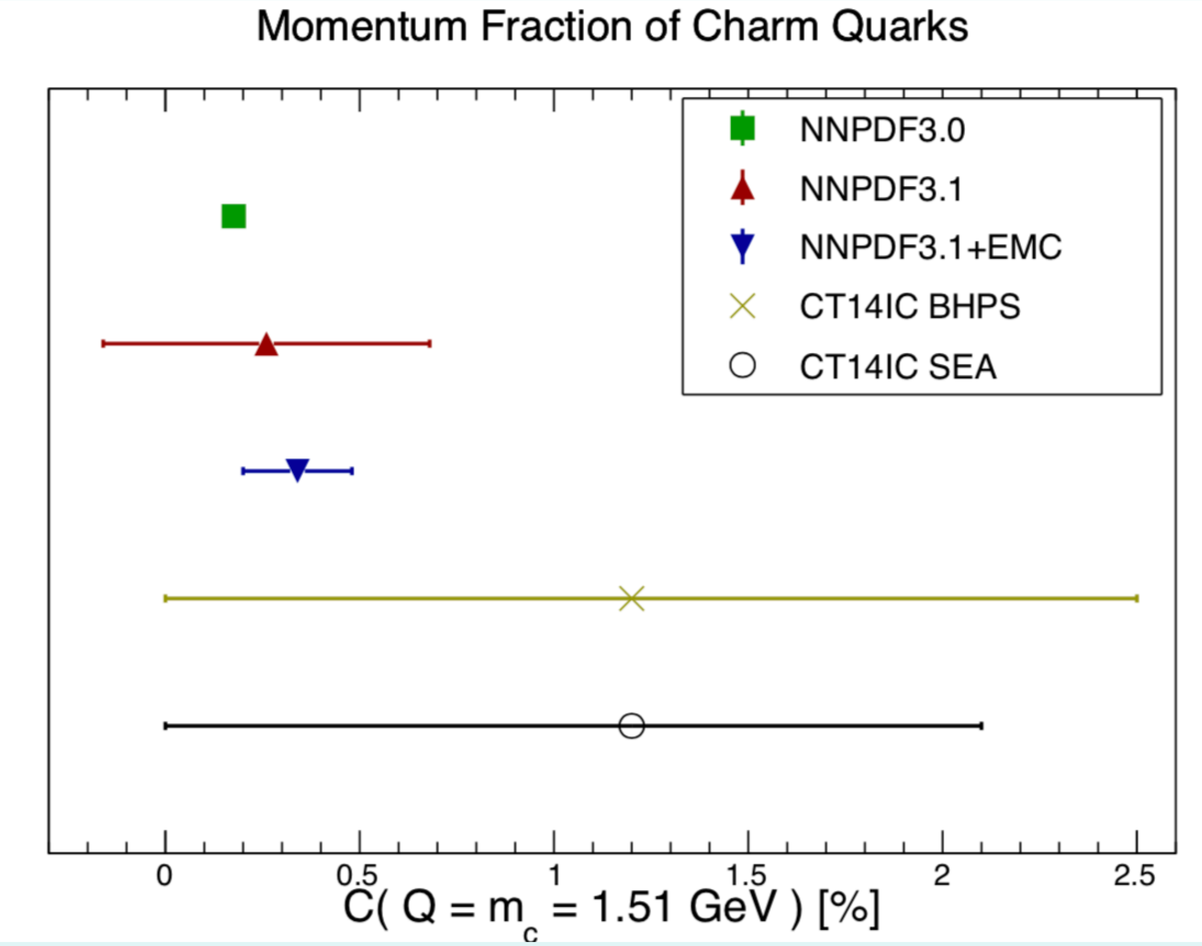
- The global PDF analysis where the charm PDF is fitted leads to a **slightly superior fit quality** than assuming a perturbatively generated charm PDF

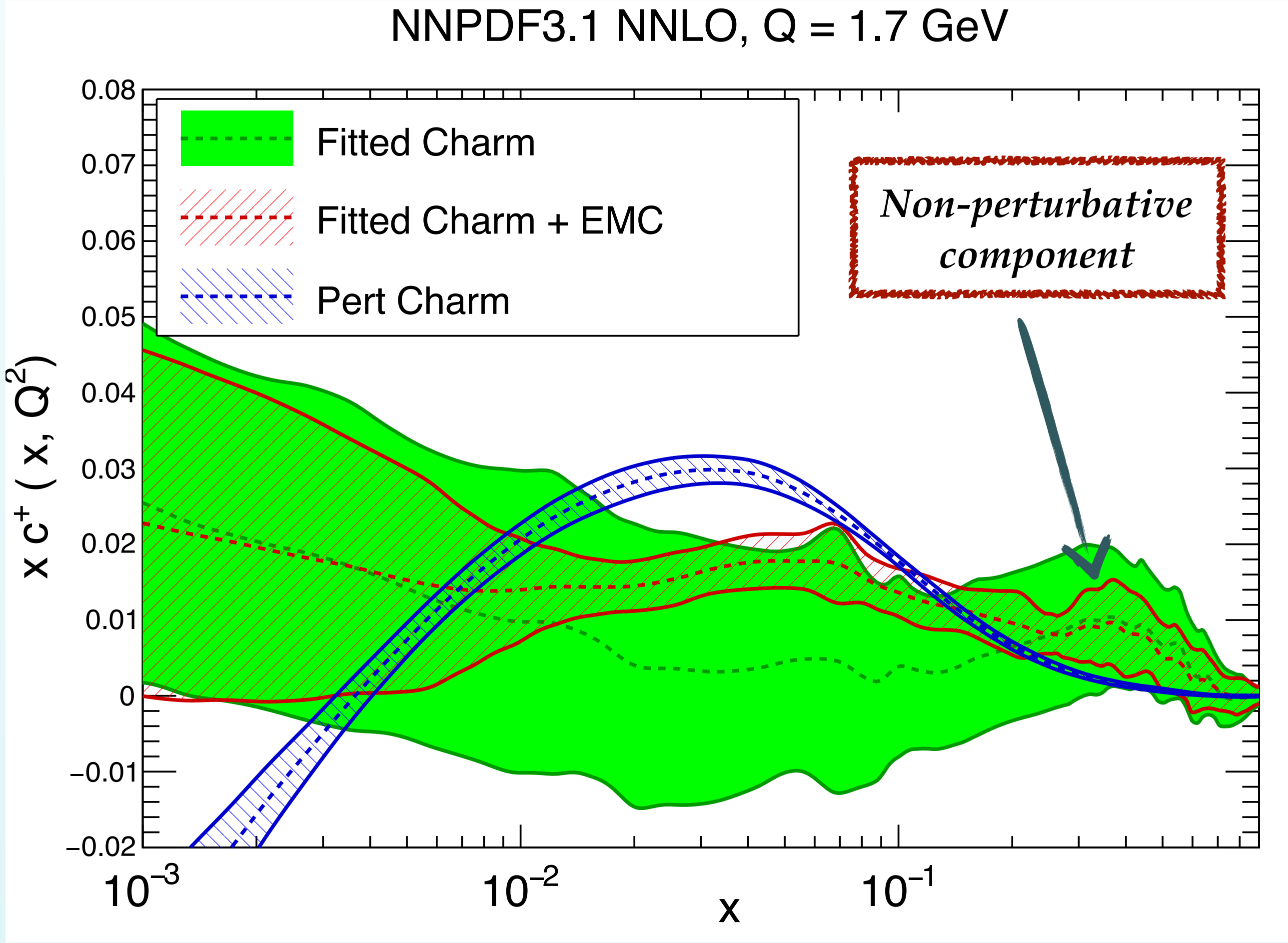
A charming story

The NNPDF3.1 global analysis is performed with a **fitted charm PDF**. The recent LHC *W,Z* data, in particular from LHCb, impose stringent constraints on the **size of the charm PDFs at $Q=m_c$**

$$C(Q^2) \equiv \int_0^1 dx\, x \left(c(c, Q^2) + \bar{c}(x, Q^2) \right)$$

Charm momentum fraction in the proton



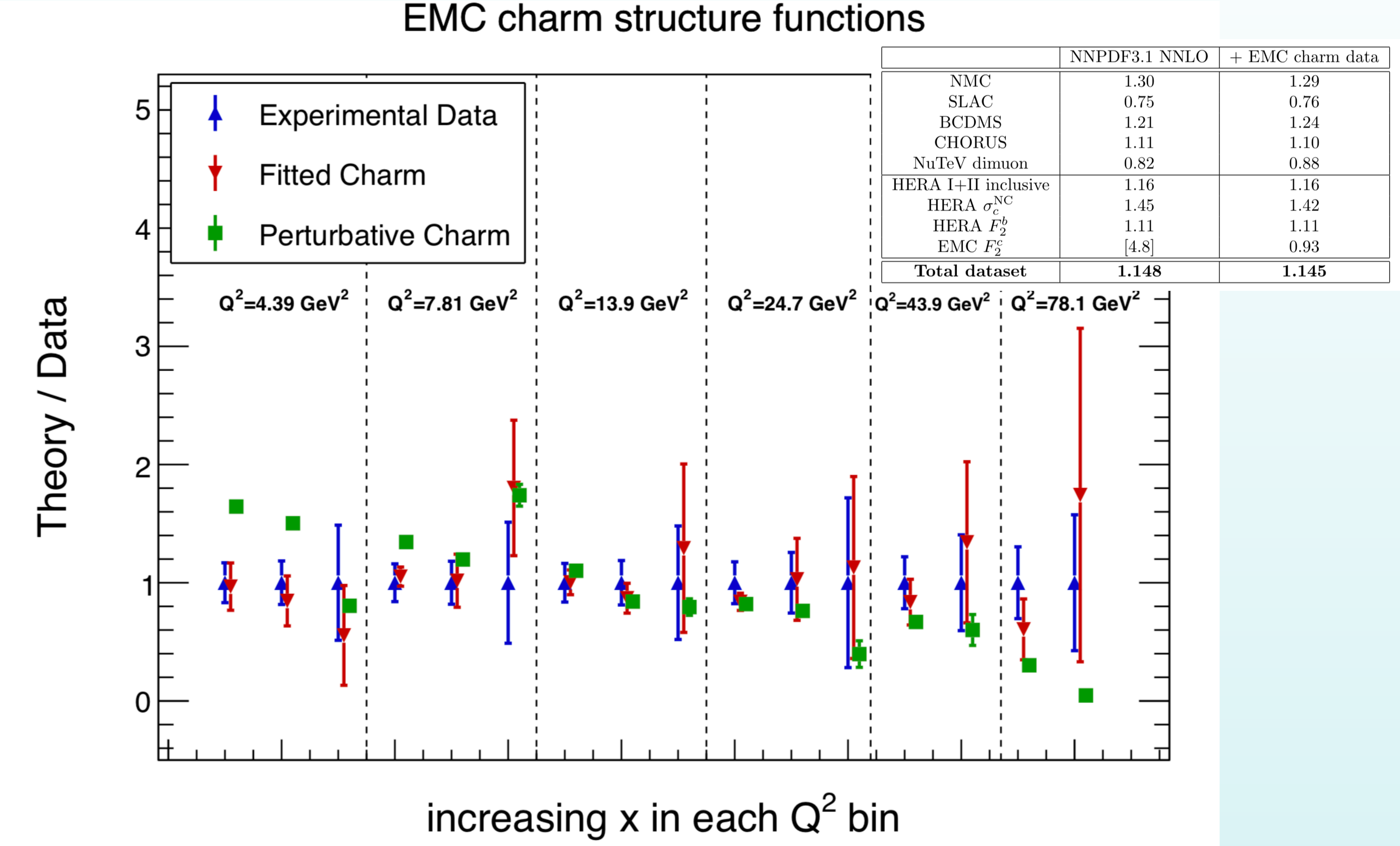


Non-perturbative component

Indications of a **non-perturbative component of the charm PDF localised at large-*x***, though its statistical significance is still limited….

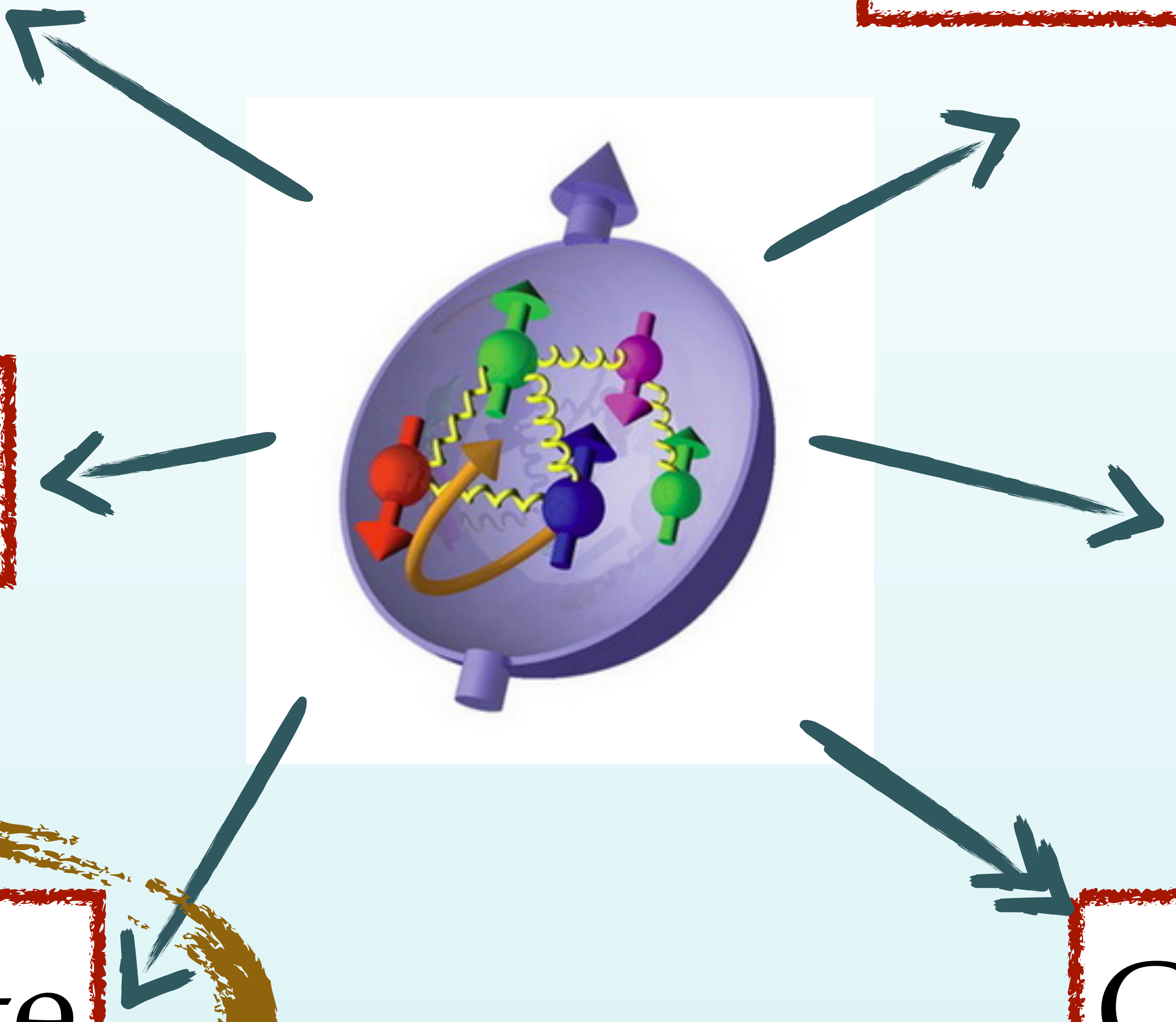
A charming story

Once **charm PDF is fitted**, one achieves excellent description of **EMC charm structure functions**



	NNPDF3.1 NNLO	+ EMC charm data
NMC	1.30	1.29
SLAC	0.75	0.76
BCDMS	1.21	1.24
CHORUS	1.11	1.10
NuTeV dimuon	0.82	0.88
HERA I+II inclusive	1.16	1.16
HERA σ_c^{NC}	1.45	1.42
HERA F_2^b	1.11	1.11
EMC F_2^c	[4.8]	0.93
Total dataset	**1.148**	**1.145**

The inner life of protons



Gluon

Strong coupling

Photon

BFKL dynamics

Strange

Charm

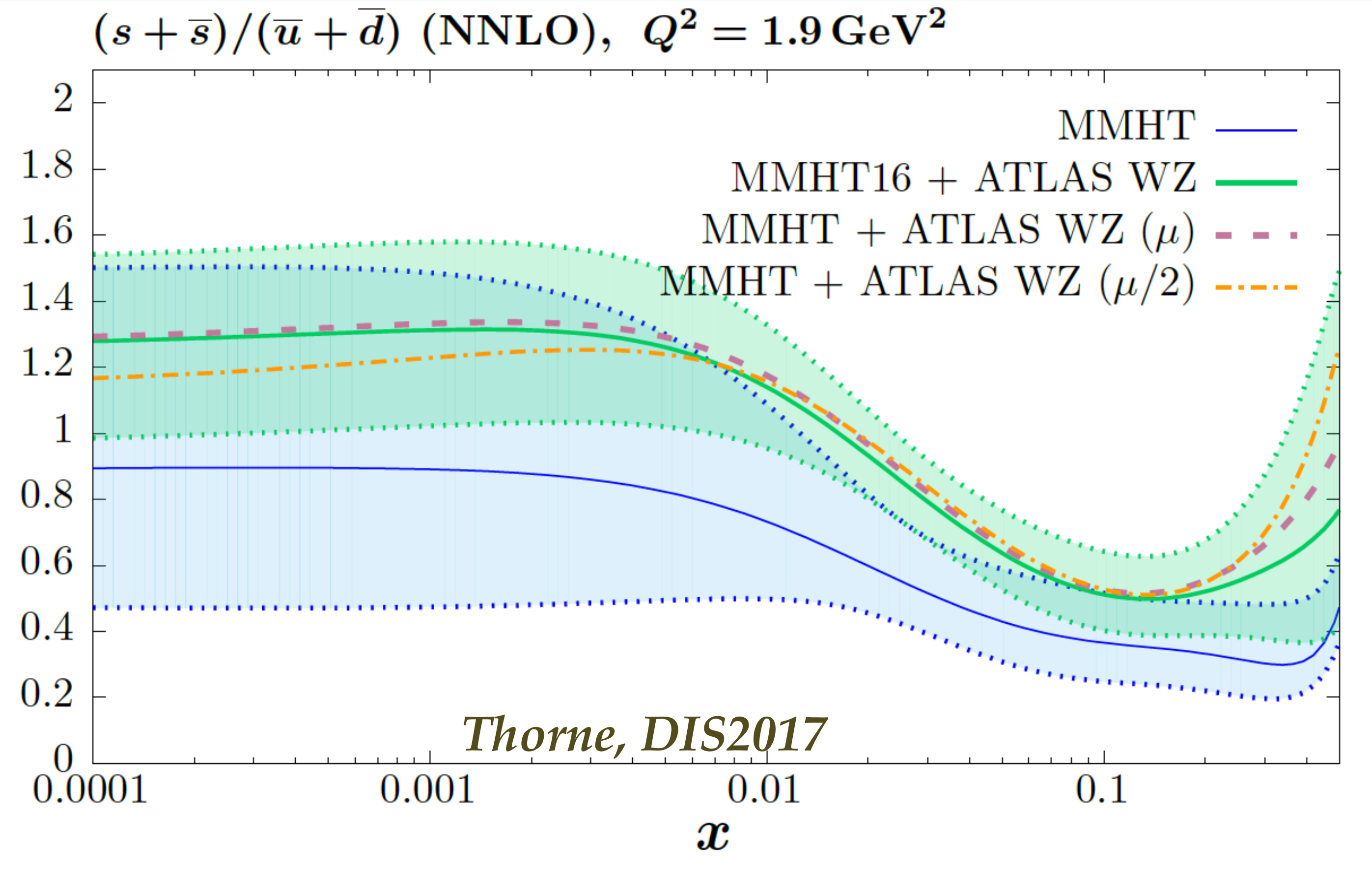
A strange conundrum

- In most PDF fits, strangeness suppressed wrt **up and down quark sea** due to **neutrino dimuon data**

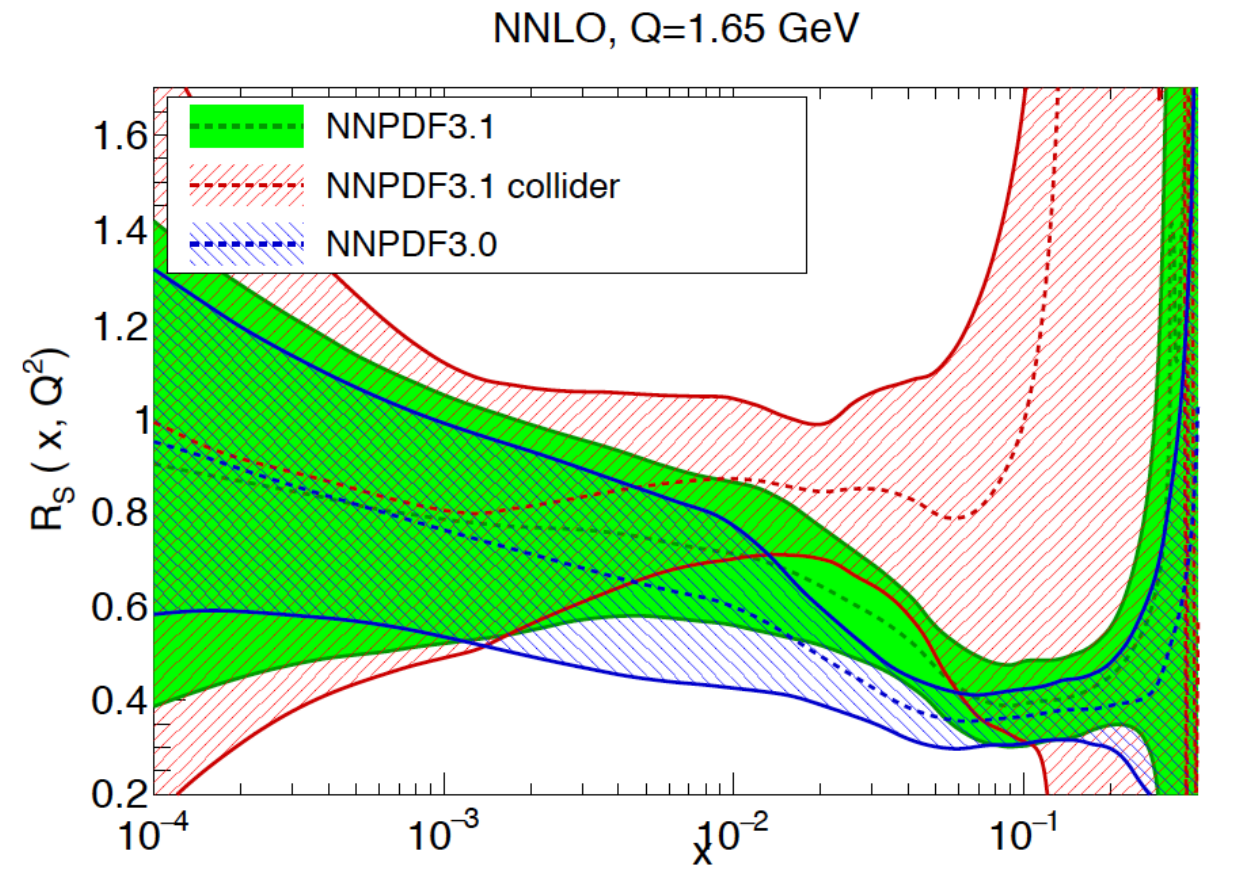
- On the other hand, recent collider data, in particular the ATLAS W,Z 2011 rapidity distributions, prefer instead a **symmetric strange quark sea**

$$R_s(x,Q^2)=\frac{s(x,Q^2)+\bar{s}(x,Q^2)}{\bar{u}(x,Q^2)+\bar{d}(x,Q^2)}\quad\begin{cases}\approx 0.5 \textit{ (from neutrino, CMS W+c)}\\ \approx 1.0 \textit{ (from ATLAS W,Z)}\end{cases}$$

The new ATLAS data can be accommodated in the **global fits**, and *i)* indeed it **increases strangeness**, but not as much as in a collider-only fit, and *ii)* **some tension remain**s between neutrino and collider data

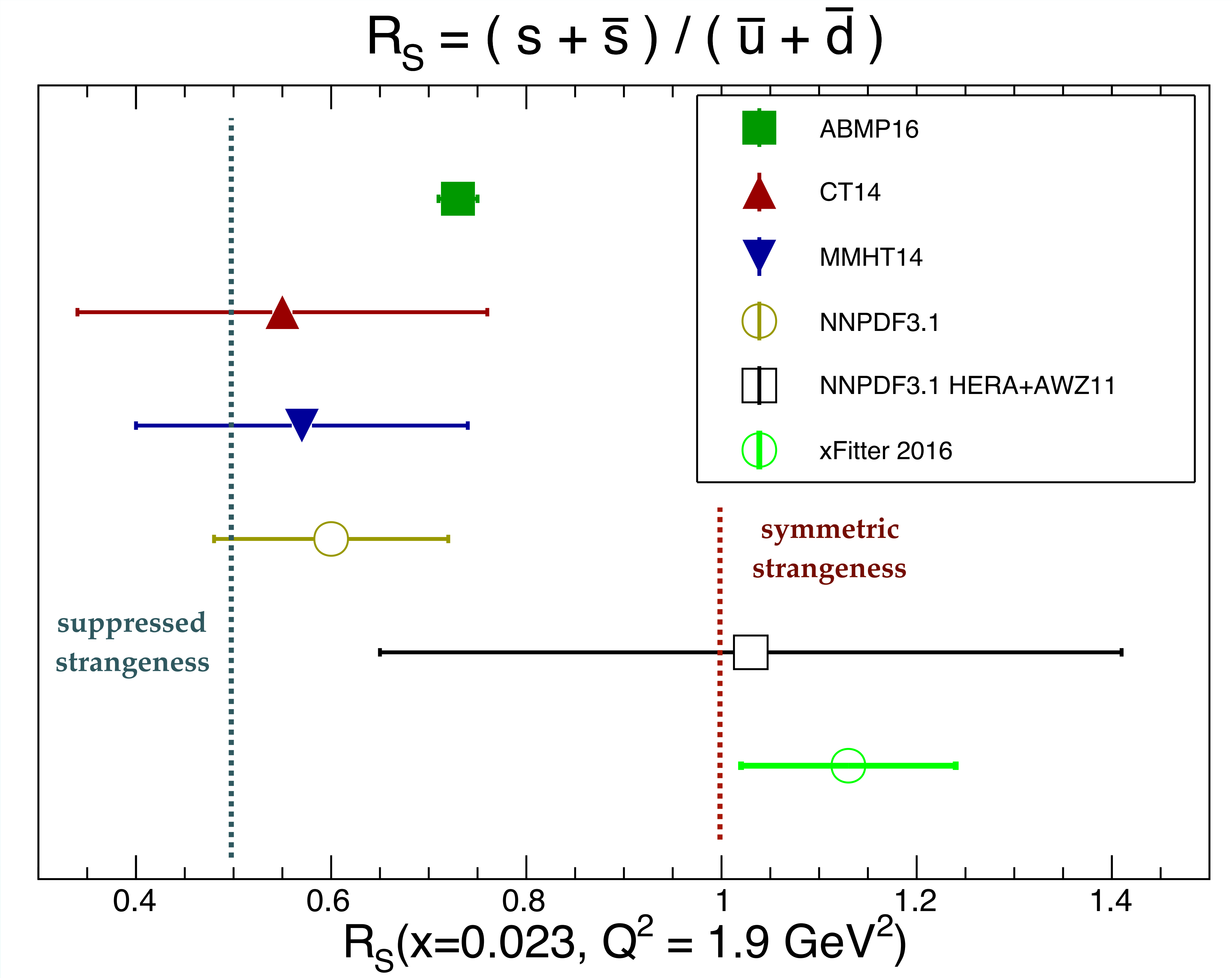


Thorne, DIS2017

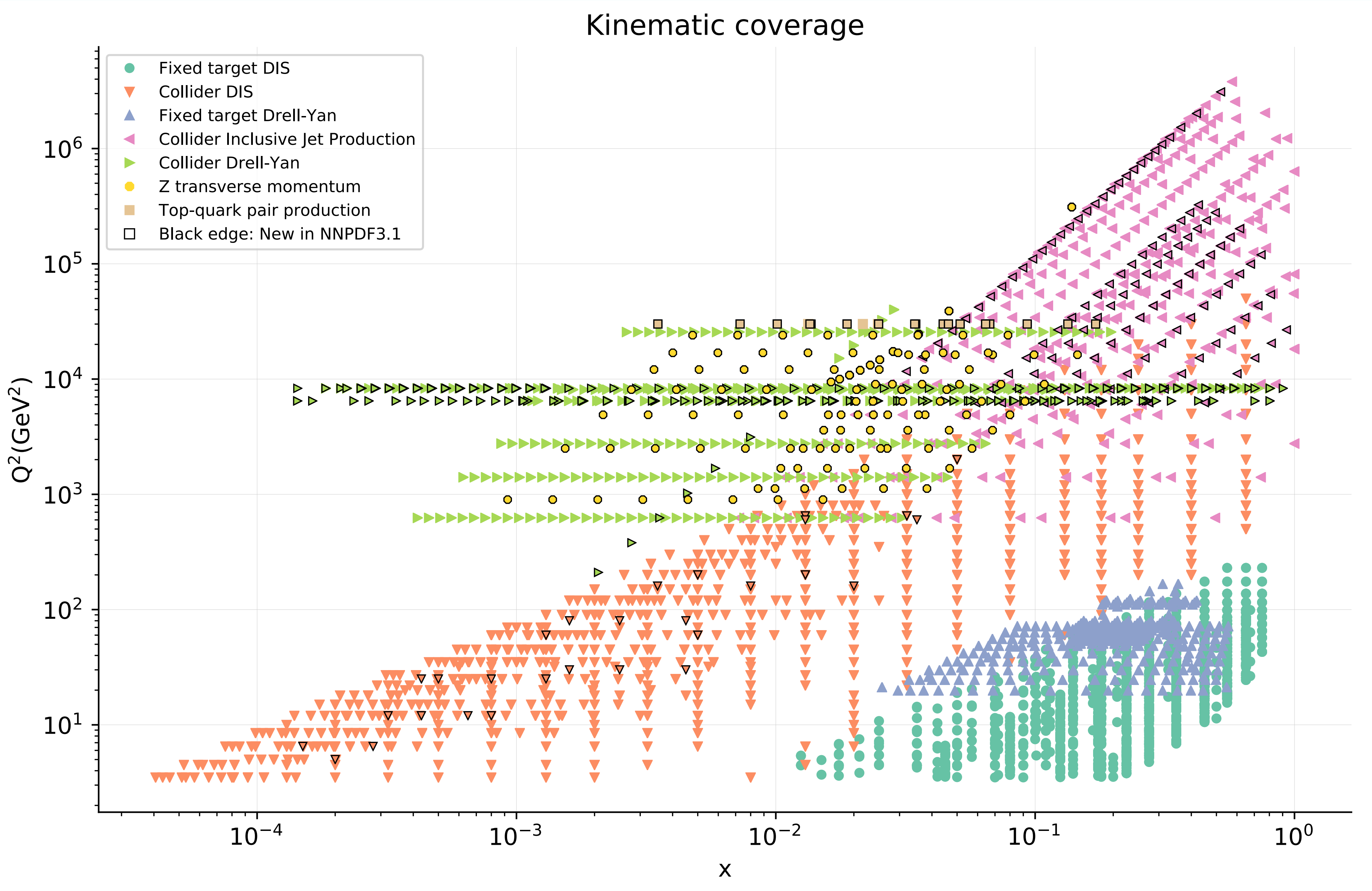


How strange is the proton?

In the global fit, the resulting strange PDF is the best compromise between the **pulls from individual experiments.** More strange-sensitive measurements are needed to shed more light in this **strange mystery!**



The global QCD fit machinery



Global PDF fits: highly non-trivial validation of the **QCD factorisation framework**:

i) *including **O(5000) data** points,*

ii) *from **O(50) experiments**,*

iii) *several of them with **≈1% errors**,*

yet still manage to achieve $\chi^2/N_{dat} \approx 1$!

Summary and outlook

- Recent developments in our understanding of the **quark and gluon structure of the proton** have been driven by a combination of:

  ☑ **Theory**: Progress in NNLO QCD and NLO EW calculations for many collider processes: differential top pairs, inclusive jets, the Z transverse momentum … Also the calculation of the photon PDF in terms of DIS structure functions

  ☑ **Data**: a wealth of high-precision measurements from HERA, Tevatron, ATLAS, CMS and LHCb, in several cases with sub-percent uncertainties.

  ☑ **Methodology**: fitted charm PDF, combination/reduction methods for different PDFs, new software for PDF fits, fast NLO/NNLO interfaces, ….

- Improvements for **many Run II analysis**: Higgs couplings, M_W measurements, heavy BSM particle production, precision SM studies, SMEFT fits, MC validation, …

- **Theory uncertainties** now likely to be a limiting factor in PDF fits and PDF-related studies: urgent need to provide PDF analysis with robust estimates of TH errors, including MHOUs

Summary and outlook

- Recent developments in our understanding of the **quark and gluon structure of the proton** have been driven by a combination of:

  ☑ **Theory**: Progress in NNLO QCD and NLO EW calculatio… der processes: differential top pairs, inclusive jets, the Z tran… he calculation of the photon PDF in terms of DIS stru…

  ☑ **Data**: a wealth of high-p… and LHCb, in severa…

  …s for different PDFs,

- I… : Higgs couplings, M_W measurements, heavy BSM parti… studies, MC validation, …

- **The**… **uncertainties** now likely to be a limiting factor in PDF fits and PDF-related studies: urgent need to provide PDF analysis with robust estimates of TH errors, including MHOUs

PDFs are in excellent shape for the exploitation of the Run II harvest!

Thanks for your attention!