# Prospects for PDF benchmarks and combination

Ultimate Precision at Hadron Colliders


Emanuele R. Nocera

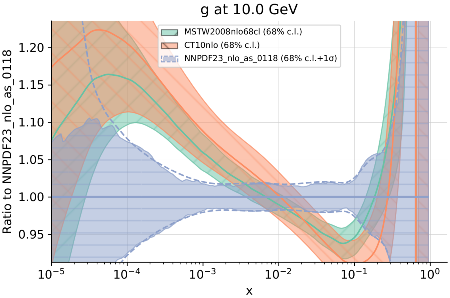
Nikhef - Amsterdam


Institut Pascal, Orsay – 26th November 2019

MARIE CURIE ACTIONS

# Foreword: paths to PDF benchmarks

circa 2012

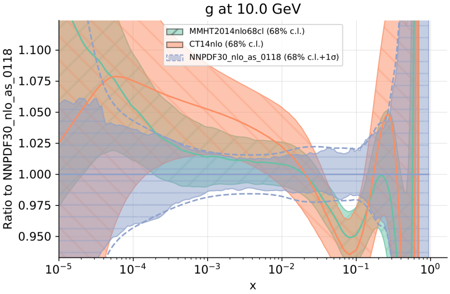



incompatible results from different groups

benchmarking exercise largely inconclusive

recommendation (PDF4LHC11): [1101.0538]
ignore individual group uncertainties
take the envelope of individual determinations

circa 2015




compatible results from different groups

PDF uncertainties become meaningful

recommendation (PDF4LHC15): [1510.03865]
combine individual group uncertainties
into a statistically meaningful set

Several benchmarking exercises between 2011 and 2015

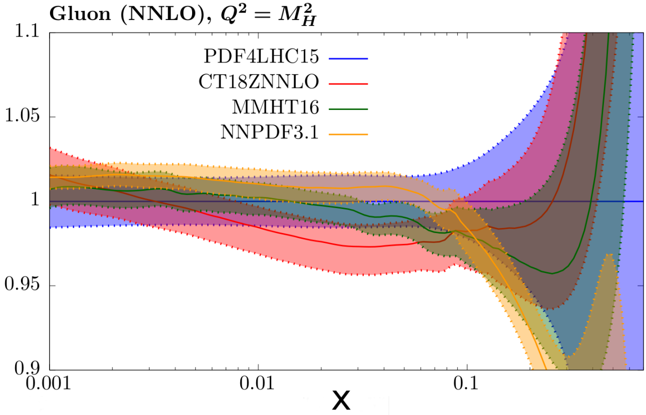
HXSWG benchmarking: PDF correlations [1201.3084]

Global PDF set benchmarking: codes, statistical methods, standard candles [1211.5142]

LH 2013 benchmarking: HQ scheme, EW corrections, cuts, scale choices, data [1405.1067]

# Foreword: paths to PDF benchmarks

circa 2019




[See L. Harland-Lang's talk]

Can residual differences among groups be explained in terms of differences in the data set, details of the QCD analysis and methodology? [PRD 86 (2012) 074017]

Progress in **data**, **theory** and **methodology** led to past benchmarking exercises

# 1. Data

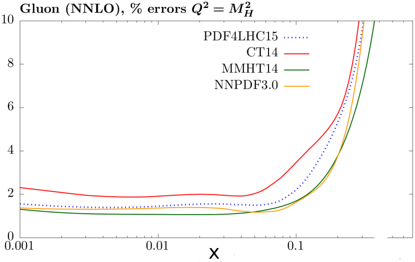
# Overview of current PDF determinations

| | NNPDF3.1 | MMHT2014 | CT18 | HERAPDF2.0 | CJ15 | ABMP16 |
|---|---|---|---|---|---|---|
| Fixed target DIS | ☑ | ☑ | ☑ | ☒ | ☑ | ☑ |
| JLAB | ☒ | ☒ | ☒ | ☒ | ☑ | ☒ |
| HERA I+II | ☑ | (☑) | ☑ | ☑ | (☑) | ☑ |
| HERA jets | ☒ | ☑ | ☒ | ☒ | ☒ | ☒ |
| Fixed target DY | ☑ | ☑ | ☑ | ☒ | ☑ | ☑ |
| Tevatron $W$, $Z$ | ☑ | ☑ | ☑ | ☒ | ☑ | ☑ |
| Tevatron jets | ☑ | ☑ | ☑ | ☒ | ☑ | ☒ |
| LHC jets | ☑ | ☑ | ☑ | ☒ | ☒ | ☒ |
| LHC vector boson | ☑ | ☑ | ☑ | ☒ | ☒ | ☑ |
| LHC top (incl.) | ☑ | (☑) | ☑ | ☒ | ☒ | ☑ |
| LHC top (diff.) | ☑ | (☑) | ☑ | ☒ | ☒ | ☒ |
| LHC single top | ☒ | ☒ | ☒ | ☒ | ☒ | ☑ |
| statistical treatment | Monte Carlo | Hessian $\Delta\chi^2$ dynamical | Hessian $\Delta\chi^2$ dynamical | Hessian $\Delta\chi^2 = 1$ | Hessian $\Delta\chi^2 = 1.645$ | Hessian $\Delta\chi^2 = 1$ |
| parametrisation | Neural Network (259 pars) | Chebyschev pol. (37 pars) | Bernstein pol. (30-35 pars) | polynomial (14 pars) | polynomial (24 pars) | polynomial (15 pars) |
| HQ scheme | FONLL | TR$'$ | ACOT-$\chi$ | TR$'$ | ACOT-$\chi$ | FFN |
| latest update | EPJ C77 (2017) 663 | EPJ C75 (2015) 204 | arXiv:1908.11394 | EPJ C75 (2015) 580 | PRD 93 (2016) 114017 | PRD 96 (2017) 014011 |

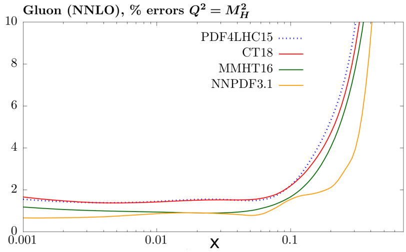
An increasingly significant amount of LHC data

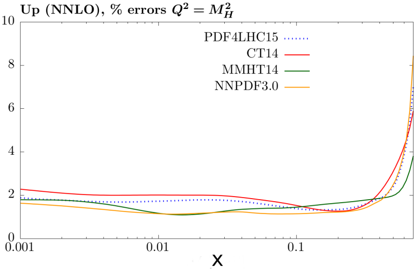
# PDF uncertainties

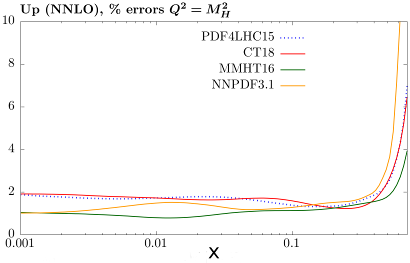
## 2015




## 2019

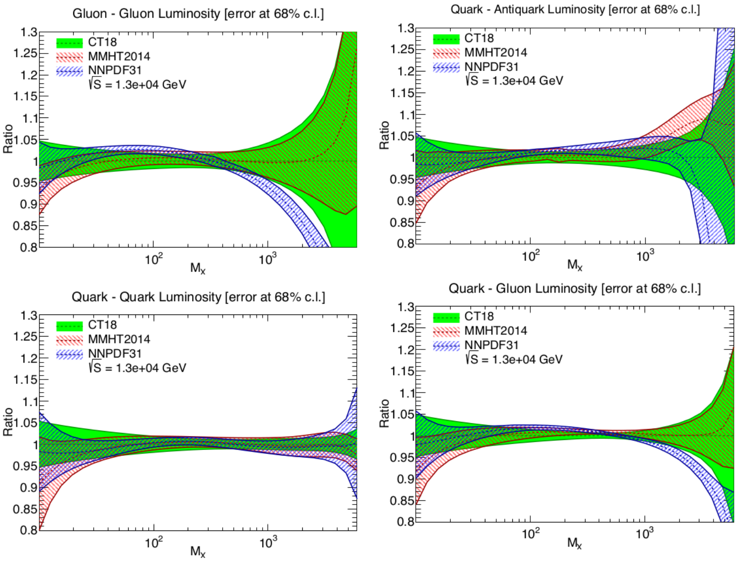



Clear reduction of PDF uncertainties, down to few %, mostly led by LHC data

# Parton Luminosities




Accompanied by some spread across PDF sets

Cracks starting to appear in data/theory comparison: benchmark exercise(s)

# Dealing with highly correlated data sets

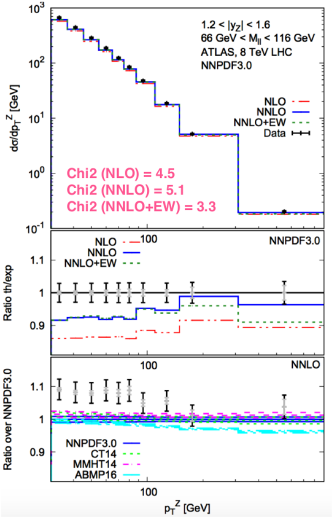
$Z\,p_T$ distributions [JHEP 1707 (2017) 130]
an uncorrelated uncertainty
should be included to achieve a good fit

Single-jet distributions [EPJ C78 (2018) 248]
default correlations: terrible $\chi^2$
(correlations across rapidity bins);
decorrelation model: imporves the fit a lot;
no significant effect on the extracted gluon;
similar gluon irrespective of the rapidity bin

$t\bar{t}$ distributions [arXiv:1909.10541]
default correlations: terrible $\chi^2$
(correlations across distributions)
loosening correlations: improves the fit a lot;
BUT large effect on the extracted gluon PDF

Can we establish as a fact
that these inconsistencies are originated by a
ill-defined experimental covariance matrix?

$Z\,p_T$ distributions




Can we devise a procedure
to deal with ill-defined experimental covariance matrices? [See talk by Z. Kassabov]

# Dealing with highly correlated data sets

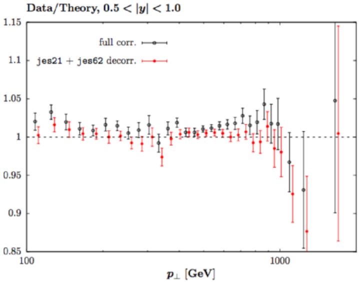
$Z\,p_T$ distributions [JHEP 1707 (2017) 130]
an uncorrelated uncertainty
should be included to achieve a good fit

Single-jet distributions [EPJ C78 (2018) 248]
default correlations: terrible $\chi^2$
(correlations across rapidity bins);
decorrelation model: imporves the fit a lot;
no significant effect on the extracted gluon;
similar gluon irrespective of the rapidity bin

$t\bar{t}$ distributions [arXiv:1909.10541]
default correlations: terrible $\chi^2$
(correlations across distributions)
loosening correlations: improves the fit a lot;
BUT large effect on the extracted gluon PDF

Can we establish as a fact
that these inconsistencies are originated by a
ill-defined experimental covariance matrix?

single-inclusive jet distributions




| Experiment | $N_{dat}$ | NLO (EW) | NNLO (EW) |
|---|---|---|---|
| CMS_1JET_8TEV | 185 | 6.67 (6.33) | 8.60 (8.37) |
| CMS_2JET_7TEV | 54 | 2.16 (1.92) | 2.26 (2.09) |
| CMS_2JET_3D_8TEV | 122 | 8.27 (8.46) | 3.96 (3.75) |
| ATLAS_1JET_8TEV_R04 | 171 | 4.37 (3.78) | 8.34 (7.88) |
| ATLAS_1JET_8TEV_R06 | 171 | 3.12 (2.81) | 4.25 (4.09) |
| ATLAS_2JET_7TEV_R04 | 90 | 1.77 (1.50) | 2.94 (2.63) |
| ATLAS_2JET_7TEV_R06 | 90 | 1.60 (1.35) | 2.31 (1.96) |

Can we devise a procedure
to deal with ill-defined experimental covariance matrices? [See talk by Z. Kassabov]

# Dealing with highly correlated data sets

$Z\,p_T$ distributions [JHEP 1707 (2017) 130]
an uncorrelated uncertainty
should be included to achieve a good fit

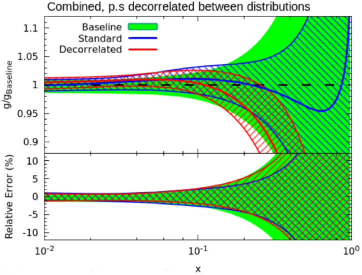
Single-jet distributions [EPJ C78 (2018) 248]
default correlations: terrible $\chi^2$
(correlations across rapidity bins);
decorrelation model: imporves the fit a lot;
no significant effect on the extracted gluon;
similar gluon irrespective of the rapidity bin

$t\bar{t}$ distributions [arXiv:1909.10541]
default correlations: terrible $\chi^2$
(correlations across distributions)
loosening correlations: improves the fit a lot;
BUT large effect on the extracted gluon PDF

Can we establish as a fact
that these inconsistencies are originated by a
ill-defined experimental covariance matrix?

$t\bar{t}$ distributions
$\chi^2/N_{\rm pts}$ ($N_{\rm pts}^{\rm tot}=25$)

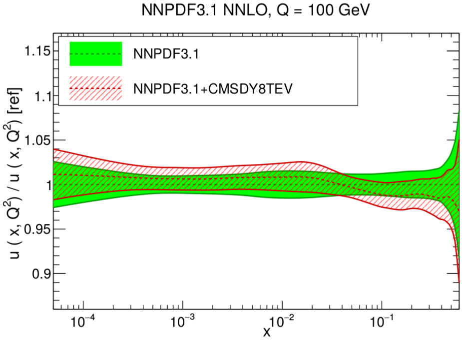
| | |
|---|---|
| $p_T$ | 0.53 |
| $y_t$ | 3.12 |
| $y_{tt}$ | 3.51 |
| $M_{tt}$ | 0.70 |
| $p_T+M_{tt}$ | 5.73 |
| Combined | 7.00 |

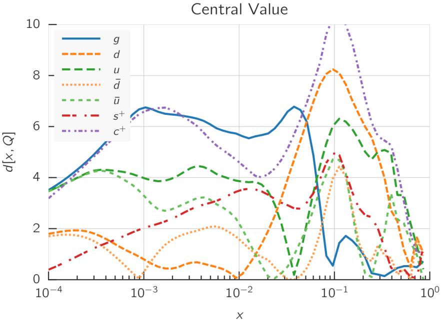



Can we devise a procedure
to deal with ill-defined experimental covariance matrices? [See talk by Z. Kassabov]

# Dealing with inconsistent data sets

Example: the CMS 8 TeV double-differential Drell-Yan distributions




⋆ If the measurements do not have clearly defined systematic errors, it is justified not to use them in a global PDF fit

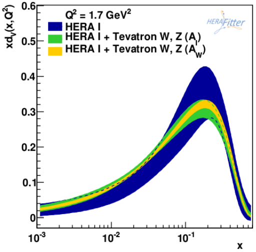
⋆ If the data sets are in strong tension with the other data sets used in a global fit, then they can be excluded; this happens on a case-by-case basis

Is the same pattern of inconsistencies/tensions seen across PDF determinations by different groups?

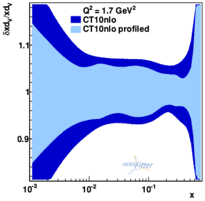
# Shall we use a reduced data set?

## IMPACT OF THE TEVATRON $W$ ASYMMETRY

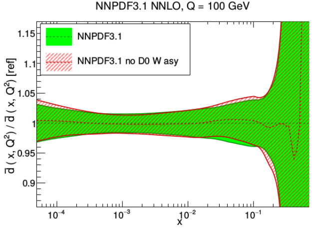
XFITTER: IMPACT ON HERA

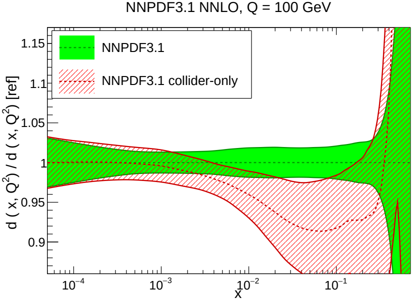



XFITTER: IMPACT ON CT10




NNPDF3.1: IMPACT ON GLOBAL FIT




⋆ Impact of a data set may be exaggerated if added to a PDF set determined from a reduced data set

⋆ PDFs are maximally constrained only in a global fit

A benchmark study can be feasible with a reduced data set, but the resulting PDFs will be (far) less accurate/precise than the PDFs obtained from a global data set

# Should everybody use the same data set?

⋆ The wider the data set, the better, BUT not all PDF groups on the same page

## New data for NNPDF4.0

ELECTROWEAK
* ATLAS high-mass Drell-Yan double-differential distributions at 8 TeV
* ATLAS W/Z total xsec at 13 TeV (81pb-1)
* ATLAS triple-differential Z production at 8 TeV (20.2 fb-1)
* ATLAS W+jets differential distributions at 8 TeV
* CMS differential distributions in Z production at 13 TeV
* LHCb W -> e nu rapidity dist, 8 TeV (2 fb-1)
* LHCb Z rapidity distribution, 13 TeV
* CMS W pt distribution, 8 TeV (18.4 fb-1)
* CMS Z+charm at 8 TeV, 19.7 fb-1
* CMS W+charm differential distributions at 13 TeV

JETS and PHOTONS
* ATLAS isolated photon production 8 TeV, 20 fb-1
* ATLAS isolated photon production, 13 TeV, 3.2 fb-1
* ATLAS dijet cross-sections at 7 TeV
* ATLAS inclusive jet cross-sections at 8 TeV from the 2012 dataset
* CMS dijet cross-sections at 7 TeV
* CMS inclusive jet production at 8 TeV, 19.6 fb-1
* CMS triple differential dijet cross-sections at 8 TeV (19.6 fb-1 )
* CMS double-differential dijet distributions at 5 TeV
* Inclusive jet and di-jet production in neutral-current DIS from H1 and ZEUS (HERA DIS jets)

prompt photons (at NNLO)

Dijets (at NNLO)

DIS jets (at NNLO)

TOP QUARK
* CMS total xsec of top-pair production at 5.02 TeV, 27.4 pb-1
* CMS double differential distributions top-quark production 8 TeV, 19.7 fb-1
* CMS single differential distributions in top-pair production (lepton+jets) at 13 TeV, L=35.8 fb-1(2016)
* CMS single differential distributions in top-pair production (dilepton) at 13 TeV, 35.8 fb-1(2016)
* CMS single top t-channel total cross section ratio at 7 TeV
* CMS single top t-channel total cross section ratio at 8 TeV
* CMS single top t-channel total cross section ratio at 13 TeV
* ATLAS single top t-channel total cross section ratio and diff. distributions at 7 TeV
* ATLAS single top t-channel total cross section ratio at 8 TeV
* ATLAS single top t-channel total cross section ratio at 13 TeV

single top (at NNLO)

Cutoff date for new data:
end of 2019

## Upgrades

* ATLAS W/Z production, 7 TeV (4.6 fb-1) => added the off-peak and forward Z prod bins
* Final combination of charm and beauty str fns from HERA (Runs I+II): replaces HERA-I charm comb and H1, ZEUS structure functions

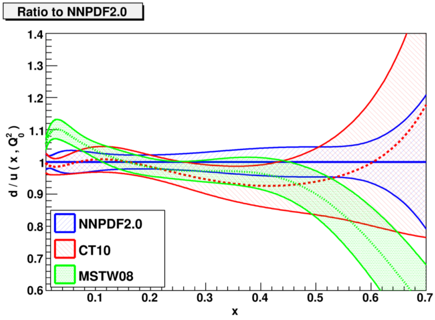
⋆ Advisable to have updated CT/MMHT releases

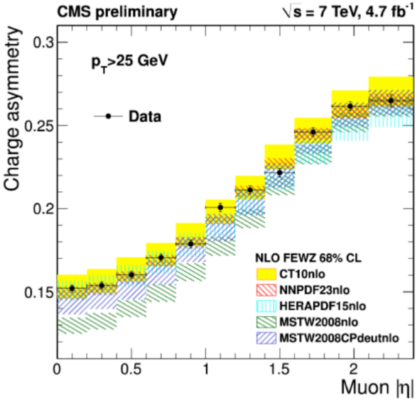
⋆ If needed, may consider a NNPDF3.2 release with a partial data set update

# 2. Methodology

# Should everybody use the same data set?

Example 1: the $d/u$ ratio (in 2011)



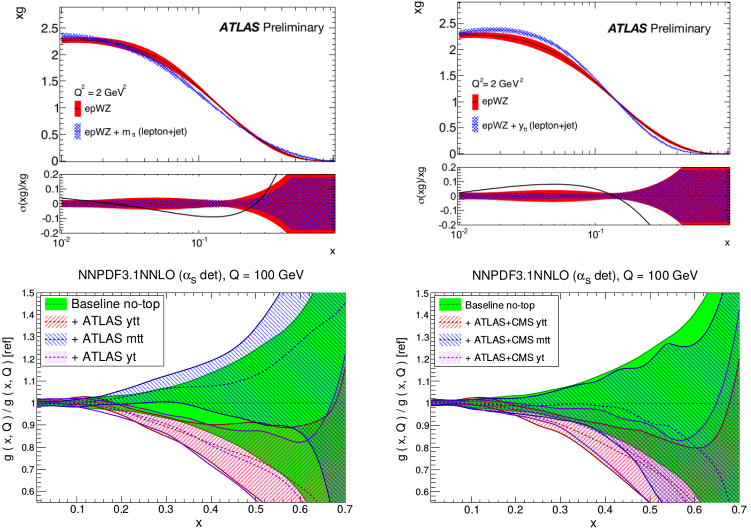




MSTW08 discrepancy traced to a parameterisation issue, now solved

- ⋆ The wider the data set, the better, BUT not all methodologies can accommodate it
- ⋆ If the PDF sets include the data, but do not agree with the data, and the other PDF sets do, then it is crucial to understand the source of the disagreement

# Should everybody use the same data set?

Example 2: $g$ from ATLAS differential $t\bar{t}$ data



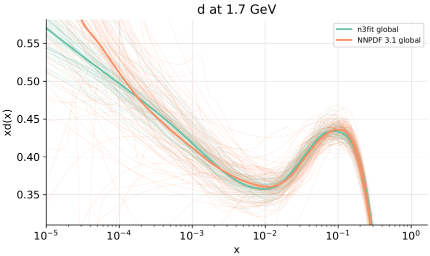

⋆ There is an irreducible relationship between the data set and methodology

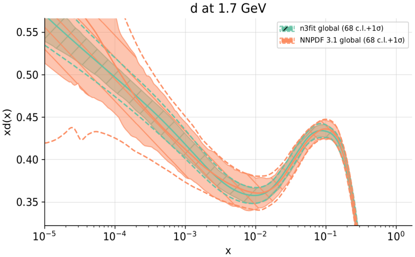
# Should everybody use the same methodology?

⋆ Differences in the methodology should be part of the benchmarking exercise (along with differences in the details of the QCD analysis)

⋆ This does not imply that better methodologies shouldn't be pursued

Improvements in the NNPDF methodology [EPJ C79 (2019) 676]




Gradient descent techniques implemented with Keras + TensorFlow

Performance increased by a factor ∼20; allows for removing a lot of legacy code

Central values and fit quality remarkably stable; PDF uncertainties somewhat affected

comparable in the data region significantly reduced outside

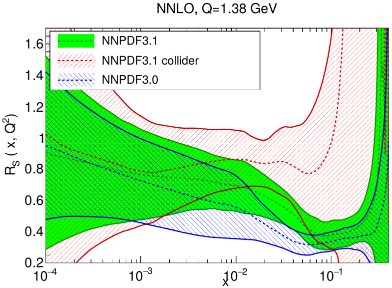
Fewer replicas for equal accuracy; completely new classes of studies open up

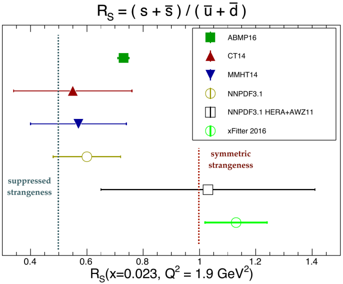
Methodological improvements are likewise pursued in the MMHT and CT frameworks

What's their impact?

# 3. Theory

# Interplay between Data, Methodology and Theory




In most PDF fits the strange PDF is suppressed w.r.t up and down sea quark PDFs

effect mostly driven by neutrino dimuon data

A symmetric strange sea PDF is preferred by collider data

in particular by ATLAS $W, Z$ rapidity distributions (2011) [EPJ C77 (2017) 367]

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

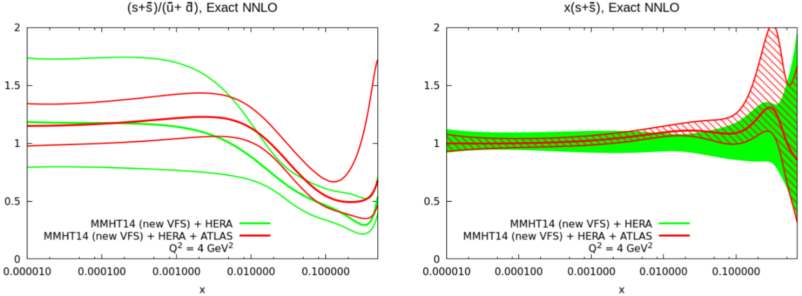
The ATLAS data can be accommodated in the global fit

more flexible methodology: NNPDF3.1 vs XFitter [EPJ C77 (2017) 663]

better theory: massive corrections in CC DIS [JHEP 1802 (2018) 026]

better treatment of data: covariance matrix regularisation [See Z. Kassabov's talk]

# Interplay between Data, Methodology and Theory




In most PDF fits the strange PDF is suppressed w.r.t up and down sea quark PDFs

effect mostly driven by neutrino dimuon data

A symmetric strange sea PDF is preferred by collider data

in particular by ATLAS $W, Z$ rapidity distributions (2011) [EPJ C77 (2017) 367]

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

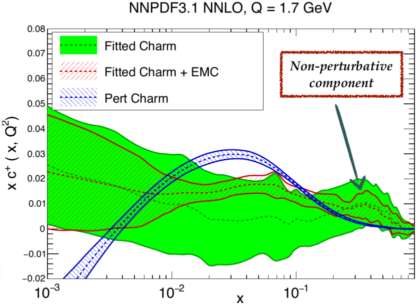
The ATLAS data can be accommodated in the global fit

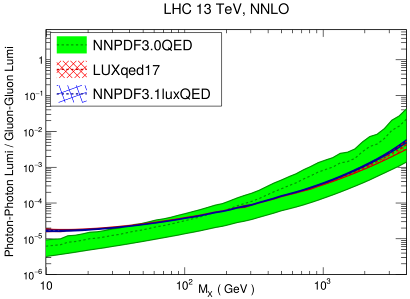
more flexible methodology: NNPDF3.1 vs XFitter [EPJ C77 (2017) 663]

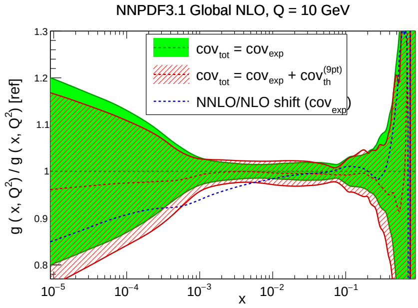
better theory: massive corrections in CC DIS [JHEP 1802 (2018) 026]

better treatment of data: covariance matrix regularisation [See Z. Kassabov's talk]

# Not all theoretical frameworks are created equal




Fitted vs pertubative charm

[EPJ C76 (2016) 647]

The photon PDF (and EW corrections)

[SciPost Phys 5 (2018) 008; see talk by C. Schwan]

Theory uncertainties

[EPJ C79 (2019) 838; EPJ C79 (2019) 931; see talk by C. Voisey]

Not all PDFs are on the same page
Shall we combine PDF sets
determined with different theories?

# 4. Concluding remarks

# Summary

## BENCHMARK(S)

Is the same pattern of data/theory discrepancies
seen across PDF determinations by different groups?

If the PDF sets include the data, but do not agree with the data, and the other PDF sets do,
then it is crucial to understand the source of the disagreement: top pair

Can we establish the origin of the data/theory inconsistencies?

tension across datasets? ill-defined experimental covariance matrix?
incompleteness of the theory? limitations in the methodology?

## COMBINATION

Advisable to have updated CT/MMHT releases

if needed, may consider a NNPDF3.2 release with a partial data set update

NNPDF4.0 too far in the future for a combination?

maybe, anyways default theory choices in NNPDF4.0 will be quite different
from other current PDF sets to justify a combination on an equal footing

# Summary

## BENCHMARK(S)

Is the same pattern of data/theory discrepancies
seen across PDF determinations by different groups?

If the PDF sets include the data, but do not agree with the data, and the other PDF sets do, then it is crucial to understand the source of the disagreement: top pair

Can we establish the origin of the data/theory inconsistencies?

tension across datasets? ill-defined experimental covariance matrix?
incompleteness of the theory? limitations in the methodology?

## COMBINATION

Advisable to have updated CT/MMHT releases

if needed, may consider a NNPDF3.2 release with a partial data set update

NNPDF4.0 too far in the future for a combination?

maybe, anyways default theory choices in NNPDF4.0 will be quite different
from other current PDF sets to justify a combination on an equal footing

**Thank you**