Parton Distributions with Theory Uncertainties:
General Theory and First Phenomenological Studies

Rabah Abdul Khalek, Richard D. Ball, Stefano Carrazza, Stefano Forte, Tommaso Giani, Zahari Kassabov, Rosalyn L. Pearson, Emanuele R. Nocera, Juan Rojo, Luca Rottoli, Maria Ubiali, Cameron Voisey, Michael Wilson

Cameron Voisey
University of Cambridge

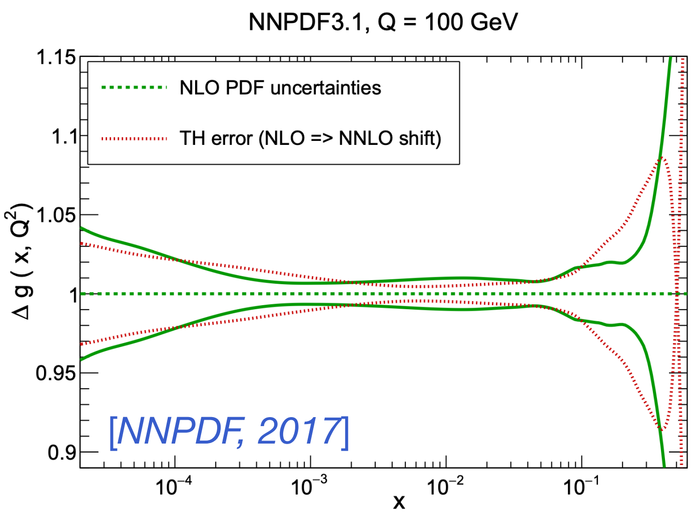
DIS 2019
9 April 2019

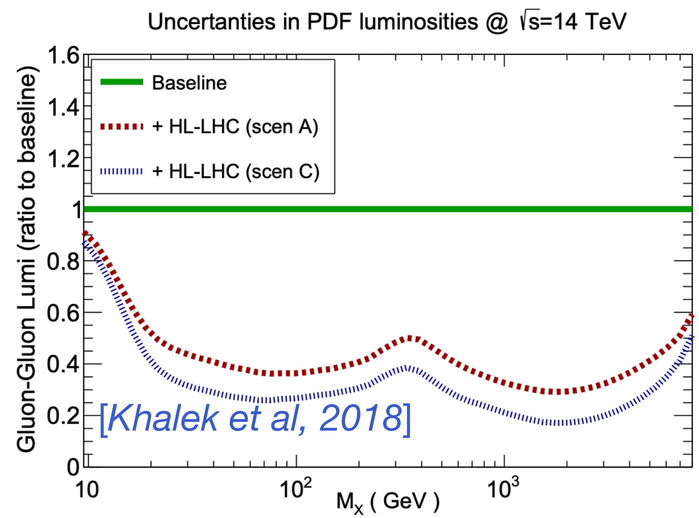
UNIVERSITY OF CAMBRIDGE NNPDF Science & Technology Facilities Council

Missing higher order uncertainties & PDFs

- Standard PDF fits use **fixed-order** hard cross sections, e.g. LO, NLO, …
- Uncertainty due to truncation of these perturbative expansions: **MHOUs**

What is the **potential impact** of MHOUs in PDF fits?





- PDFs now **high precision** → **NNLO-NLO PDF shift** now of **same order** or **larger** than PDF uncertainties
- Should we worry about MHOUs on NNLO PDFs? Looking forward: yes

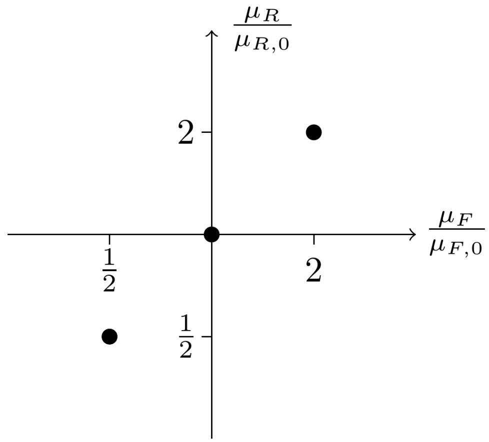
Estimating MHOUs

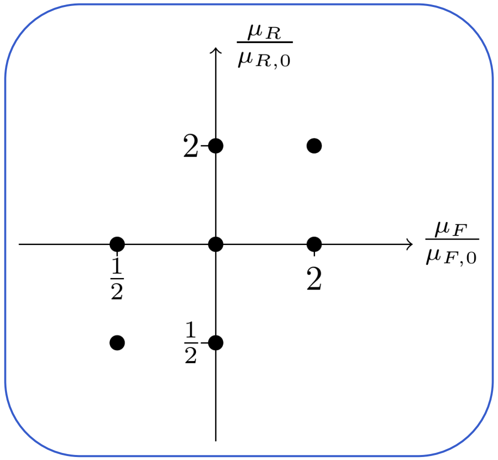
Standard technique: **scale variations**

- Convention (for hadronic processes): vary μ_R in hard cross section and μ_F in PDF, where

$$\mu_R, \mu_F \in \left[\frac{1}{2}, 2\right]$$

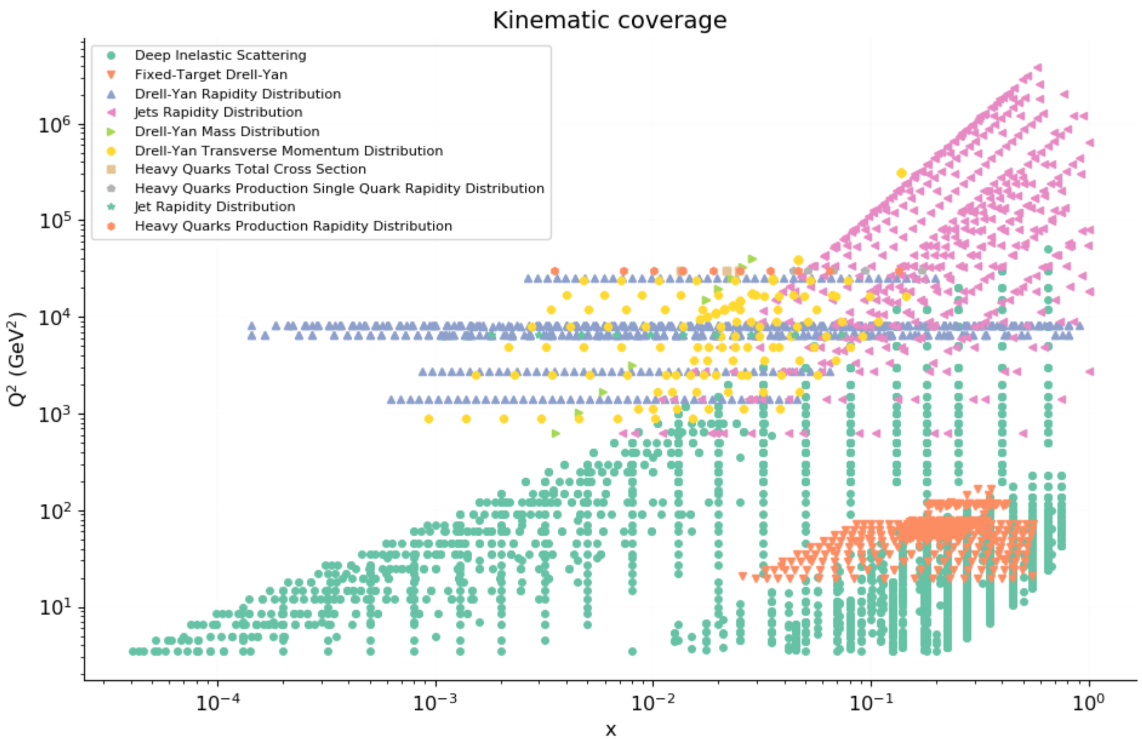
- Compute observable for different scale combinations and take **envelope**





HXSWG recommendation

Estimating MHOUs on PDFs



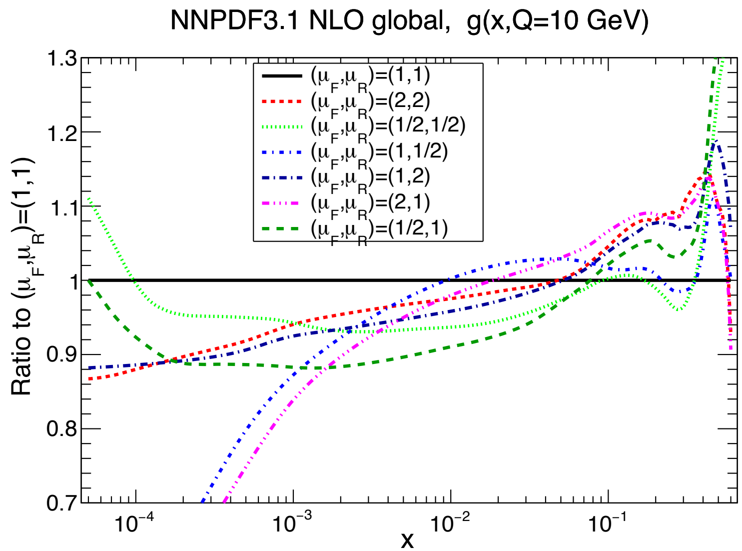
How to **extend** this to **global PDF fits**?

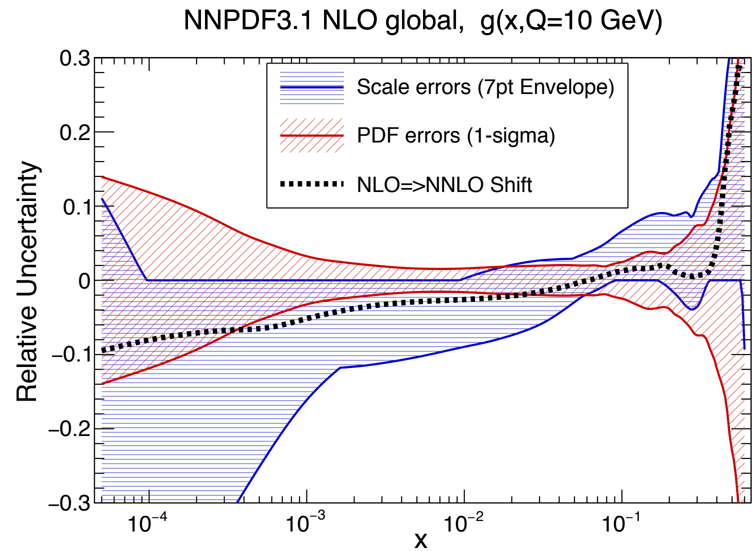
- **O(4000)** data points from **different processes**
- How to **correlate**? Common DGLAP evolution, different α_s dependence in coefficient functions

PDF fits with varied scales

Starting point for estimating MHOUs:

- Produce PDF fits for **range of scale combinations**
- Define MHOUs band as **envelope of central values**

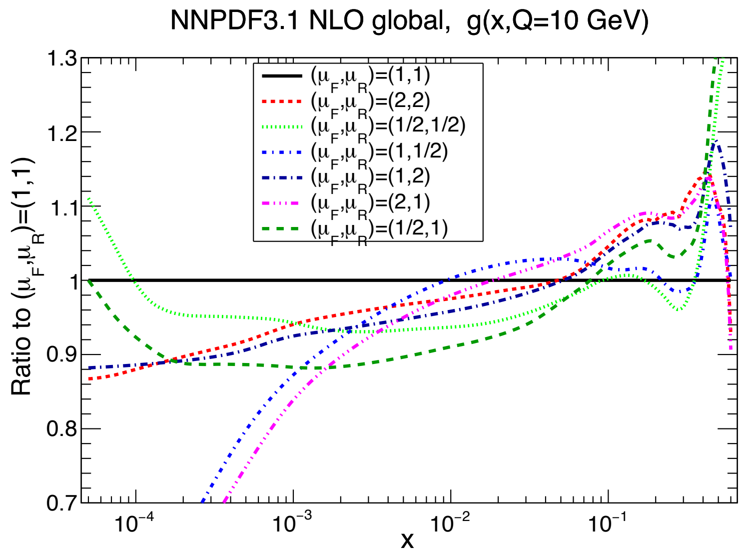


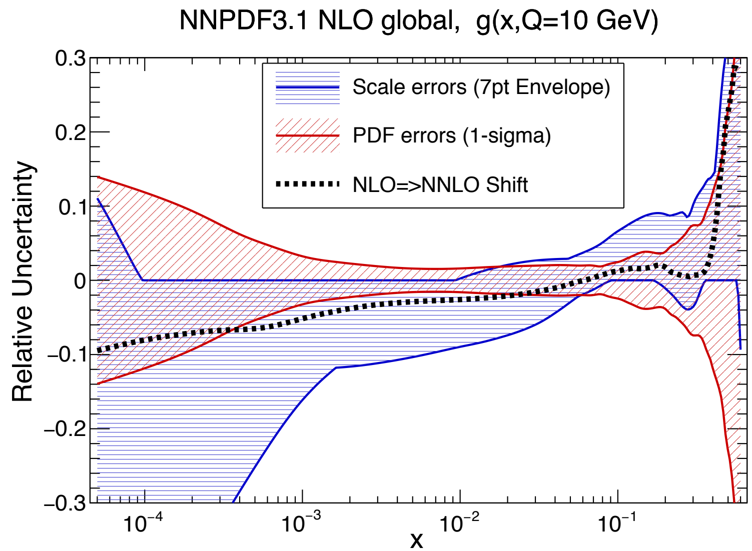


PDF fits with varied scales

Starting point for estimating MHOUs:

- Produce PDF fits for **range of scale combinations**
- Define MHOUs band as **envelope of central values**





- **Neglects correlations** in scale variations
- MHOUs only **estimated**, not **included** in PDF uncertainties

Can we include MHOUs and their correlations in PDF uncertainties by accounting for them in **fitting methodology**?

The theoretical covariance matrix

Experimental uncertainties propagated to PDFs via minimisation of figure of merit:

$$\chi^2 = (\text{data} - \text{theory})^T (\text{cov}_{\text{exp}})^{-1} (\text{data} - \text{theory})$$

Modify this to account for theory errors: *[R. D. Ball & A. Deshpande, 2018]*

$$\chi^2_{\text{tot}} = (\text{data} - \text{theory})^T (\text{cov}_{\text{exp}} + \text{cov}_{\text{th}})^{-1} (\text{data} - \text{theory})$$

Assumptions:

1. Theoretical uncertainties **independent** from experimental uncertainties
   → we are adding exp. and th. uncertainties in quadrature
2. Theoretical uncertainties are **Gaussianly distributed**

Applicable to other types of theoretical uncertainty, e.g. Monte Carlo, nuclear uncertainties, …

The theoretical covariance matrix

See talk later by R. L. Pearson

Edinburgh 2018/4
Nikhef 2018-062

Nuclear Uncertainties in the Determination of Proton PDFs

The NNPDF Collaboration:

Richard D. Ball[1], Emanuele R. Nocera[1,2] and Rosalyn L. Pearson[1]

[1] *The Higgs Centre for Theoretical Physics, University of Edinburgh, JCMB, KB, Mayfield Rd, Edinburgh EH9 3FD, Scotland*
[2] *Nikhef Theory Group, Science Park 105, 1098 XG Amsterdam, The Netherlands*

A theoretical covariance matrix for MHOUs

Construct $\mathrm{cov}_{\mathrm{th}}$ from **scale variations** to estimate:

1. MHOU on each point
2. Correlations between points

i,j: data points

k: scale combinations

$$\mathrm{cov}_{\mathrm{th},ij} = \frac{1}{N}\sum_k \Delta_i^{(k)}\Delta_j^{(k)} \qquad \Delta_i^{(k)} = t_i(\mu_R, \mu_F) - t_i(\mu_{R,0}, \mu_{F,0})$$

Choices:

0. Definition of covariance matrix
1. Range of scale variation
2. Number of scale combinations (3, 7, …)
3. Correlation between scales (same process, different processes)
4. Process categorisation

A theoretical covariance matrix for MHOUs

Construct cov_{th} from **scale variations** to estimate:

1. MHOU on each point
2. Correlations between points

i,j: data points

k: scale combinations

$$\mathrm{cov}_{\mathrm{th},ij} = \frac{1}{N}\sum_k \Delta_i^{(k)}\Delta_j^{(k)} \qquad \Delta_i^{(k)} = t_i(\mu_R, \mu_F) - t_i(\mu_{R,0}, \mu_{F,0})$$

Choices:

0. Definition of covariance matrix
1. Range of scale variation → $\frac{1}{2} \leq \frac{\mu_F}{\mu_{F,0}}, \frac{\mu_R}{\mu_{R,0}} \leq 2$
2. Number of scale combinations (3, 7, …)
3. Correlation between scales (same process, different processes)
4. Process categorisation

A theoretical covariance matrix for MHOUs

Construct $\mathrm{cov}_{\mathrm{th}}$ from **scale variations** to estimate:

1. MHOU on each point
2. Correlations between points

i,j: data points

k: scale combinations

$$\mathrm{cov}_{\mathrm{th},ij} = \frac{1}{N} \sum_k \Delta_i^{(k)} \Delta_j^{(k)} \qquad \Delta_i^{(k)} = t_i(\mu_R, \mu_F) - t_i(\mu_{R,0}, \mu_{F,0})$$

Choices:

0. Definition of covariance matrix
1. Range of scale variation
2. Number of scale combinations (3, 7, …)
3. Correlation between scales (same process, different processes)
4. Process categorisation

How do we correlate scales in this multi-scale problem?

See next slides

A theoretical covariance matrix for MHOUs

Construct $\mathrm{cov}_{\mathrm{th}}$ from **scale variations** to estimate:

1. MHOU on each point
2. Correlations between points

i,j: data points

k: scale combinations

$$\mathrm{cov}_{\mathrm{th},ij} = \frac{1}{N}\sum_k \Delta_i^{(k)}\Delta_j^{(k)} \qquad \Delta_i^{(k)} = t_i(\mu_R, \mu_F) - t_i(\mu_{R,0}, \mu_{F,0})$$

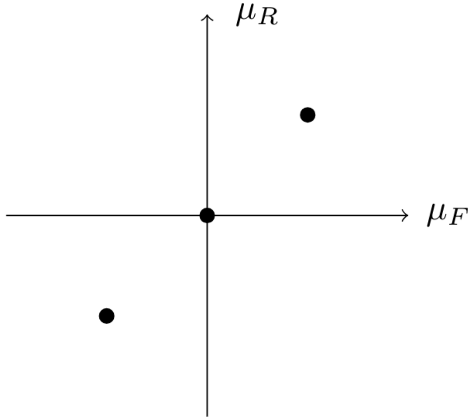
Choices:

0. Definition of covariance matrix
1. Range of scale variation
2. Number of scale combinations (3, 7, …)
3. Correlation between scales (same process, different processes)
4. Process categorisation →

- DIS neutral current
- DIS charged current
- Drell-Yan
- Jets
- Top

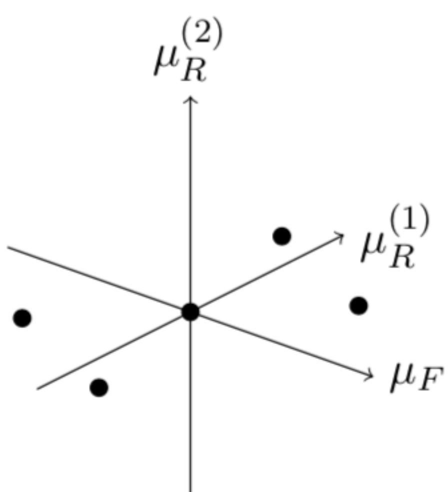
Example: 3-pt theoretical covariance matrix

i, j from ***same process***



$$\text{cov}_{\text{th,ij}} = \frac{1}{2}\{\Delta_i(+,+)\Delta_j(+,+) + \Delta_i(-,-)\Delta_j(-,-)\}$$

i, j from ***different processes***



$$\text{cov}_{\text{th,ij}} = \frac{1}{4}\{(\Delta_i(+,+) + \Delta_i(-,-))(\Delta_j(+,+) + \Delta_j(-,-))\}$$

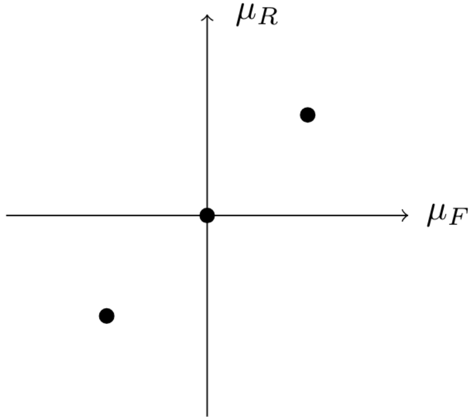
where

$$\Delta_i(+,+) = t_i(\mu_R = 2Q, \mu_F = 2Q) - t_i(\mu_R = Q, \mu_F = Q)$$

$$\Delta_i(-,-) = t_i\left(\mu_R = \frac{Q}{2}, \mu_F = \frac{Q}{2}\right) - t_i(\mu_R = Q, \mu_F = Q)$$

Example: 3-pt theoretical covariance matrix

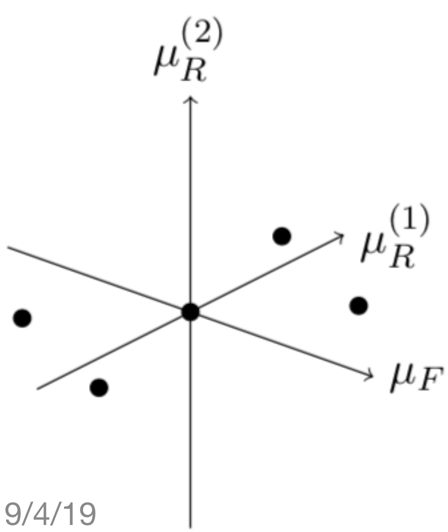
i, j from ***same process***



$$\mathrm{cov}_{\mathrm{th,ij}} = \frac{1}{2}\{\Delta_i(+,+)\Delta_j(+,+) + \Delta_i(-,-)\Delta_j(-,-)\}$$

μ_R, μ_F fully correlated

i, j from ***different processes***



$$\mathrm{cov}_{\mathrm{th,ij}} = \frac{1}{4}\{(\Delta_i(+,+) + \Delta_i(-,-))(\Delta_j(+,+) + \Delta_j(-,-))\}$$

μ_R, μ_F fully uncorrelated

$\rightarrow$ missing μ_F correlation

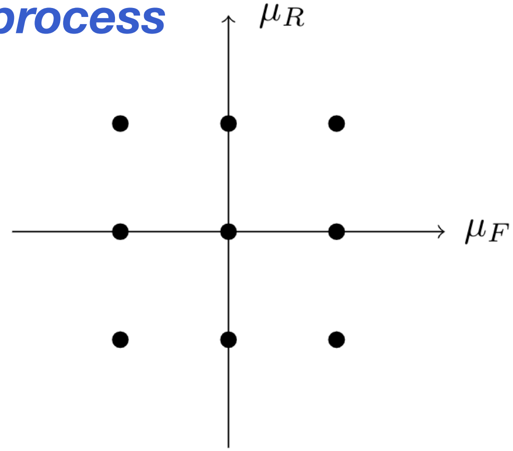
where

$$\Delta_i(+,+) = t_i(\mu_R = 2Q, \mu_F = 2Q) - t_i(\mu_R = Q, \mu_F = Q)$$

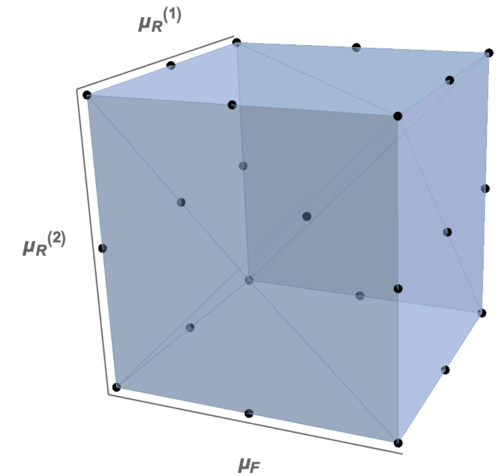
$$\Delta_i(-,-) = t_i\left(\mu_R = \frac{Q}{2}, \mu_F = \frac{Q}{2}\right) - t_i(\mu_R = Q, \mu_F = Q)$$

More complex scale combinations: 9-pt

i, j from ***same process***



i, j from ***different processes***

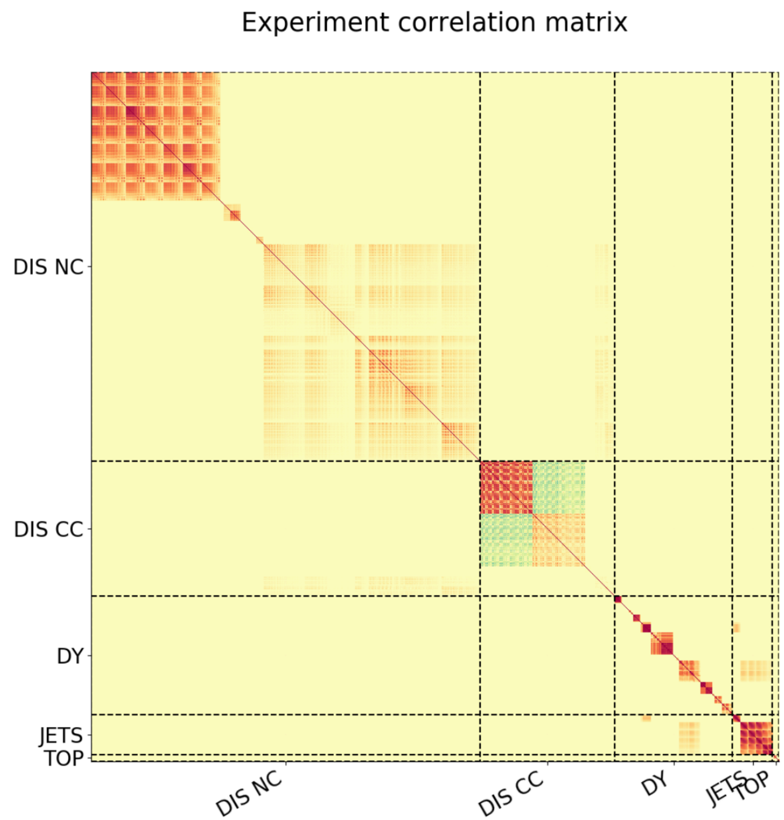


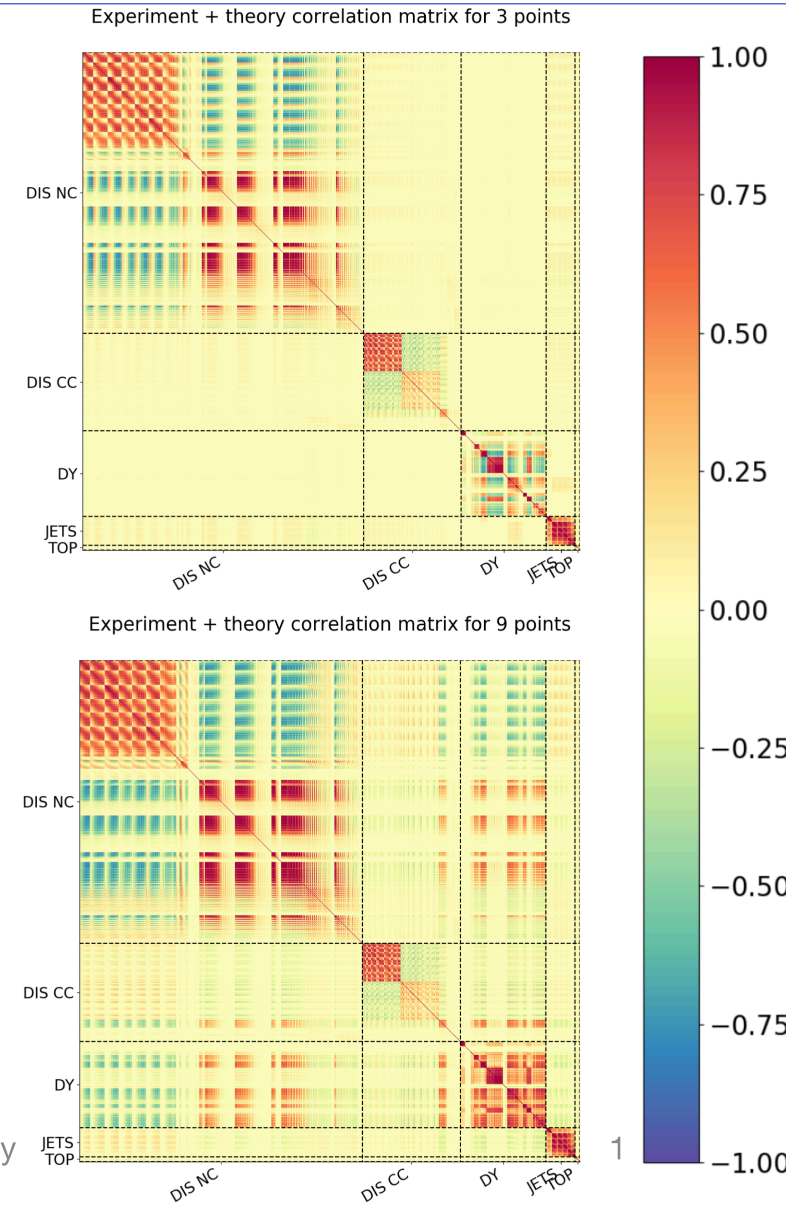
The more complex scale combination allows us to define **more complex correlation structure**:

- same process: μ_R, μ_F fully correlated
- different processes: μ_F fully correlated, μ_R fully uncorrelated

We expect this to produce a more **accurate** correlation structure, since we account for common DGLAP evolution, and different α_s dependence in coefficient functions

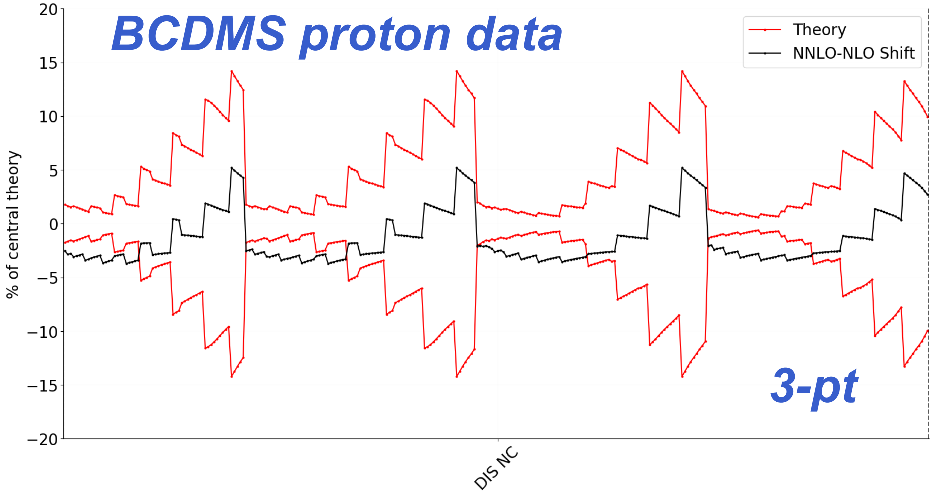
A theoretical covariance matrix for MHOUs

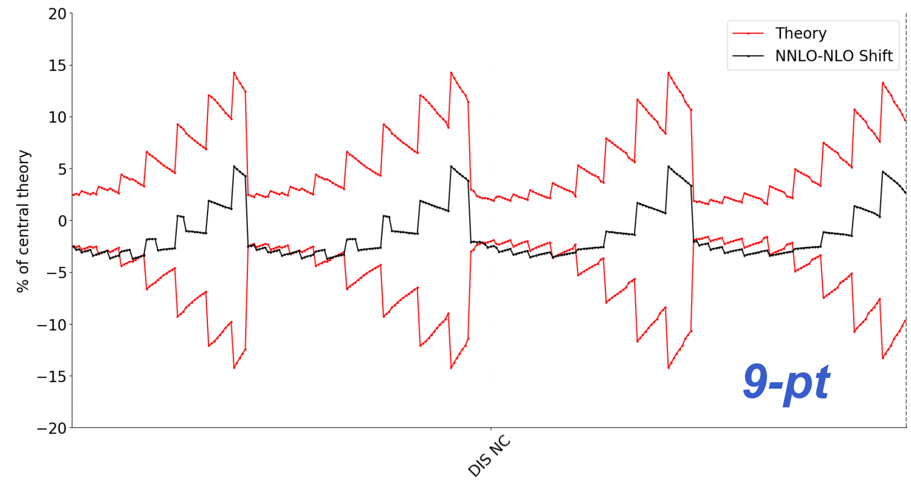




How can we **validate** and **compare** our theory covariance matrices?

Validation





We can compare **MHOU per point**, but this only tests diagonal elements of theoretical covariance matrix

→ We want to test **full covariance matrix**: MHOU per point + correlations

Validation: uncertainties + correlations

- We validate cov_{th} against exact result: **NNLO-NLO shift**

- cov_{th} is **positive semi-definite** (eigenvalues > 0 or 0)
- Eigenvalue of covariance matrix is variance in direction of eigenvector
- Eigenvalue = 0 $\Rightarrow$ no variance/**shift** predicted by cov_{th} in direction of eigenvector
- Define **efficiency**, ε, of matrix as proportion of shift that is contained within **non-zero eigenvectors** (normalised to shift projected into **full** eigenvector basis)

$$0 \leq \varepsilon \leq 1$$

$\varepsilon = 1$: cov_{th} predicts variation in same directions as shift

Validation: uncertainties + correlations

3-pt

Per **data set**: $0.12 \leq \varepsilon \leq 0.99$

Per **process**:

Process	Efficiency, ε
DIS NC	0.20
DIS CC	0.41
DY	0.16
Jets	0.67
Top	0.89

Global: $\varepsilon = 0.19$

Validation: uncertainties + correlations

3-pt

Per **data set**: $0.12 \leq \varepsilon \leq 0.99$

Per **process**:

Process	Efficiency, ε
DIS NC	0.20
DIS CC	0.41
DY	0.16
Jets	0.67
Top	0.89

Global: $\varepsilon = 0.19$

9-pt

Per **data set**: $0.70 \leq \varepsilon \leq 0.99$

Per **process**:

Process	Efficiency, ε
DIS NC	0.48
DIS CC	0.71
DY	0.85
Jets	0.99
Top	0.98

Global: $\varepsilon = 0.54$

Validation: uncertainties + correlations

3-pt

Per **data set**: $0.12 \leq \varepsilon \leq 0.99$

Per **process**:

Process	Efficiency, ε
DIS NC	0.20
DIS CC	0.41
DY	0.16
Jets	0.67
Top	0.89

Global: $\varepsilon = 0.19$

9-pt

Per **data set**: $0.70 \leq \varepsilon \leq 0.99$

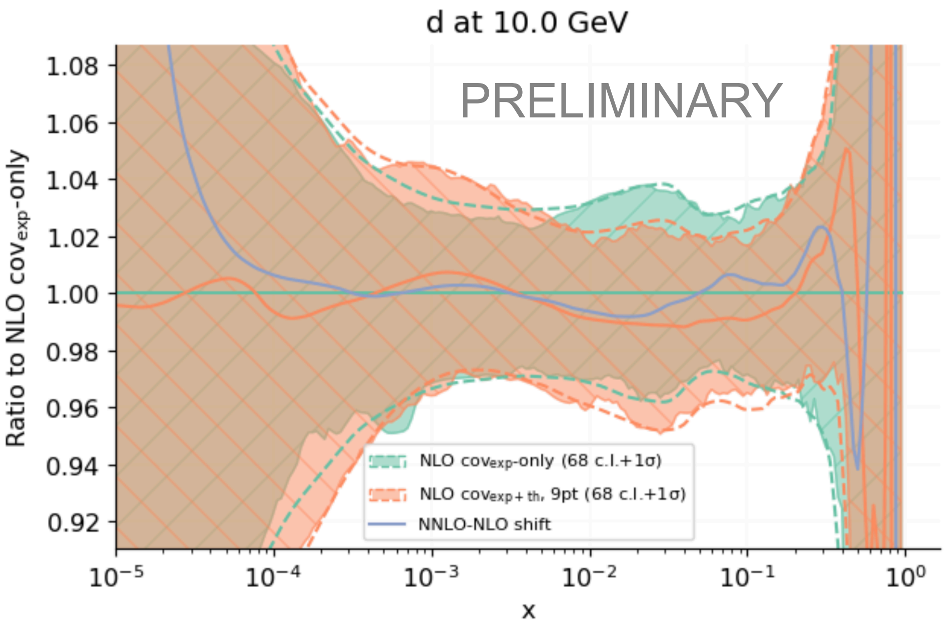
Per **process**:

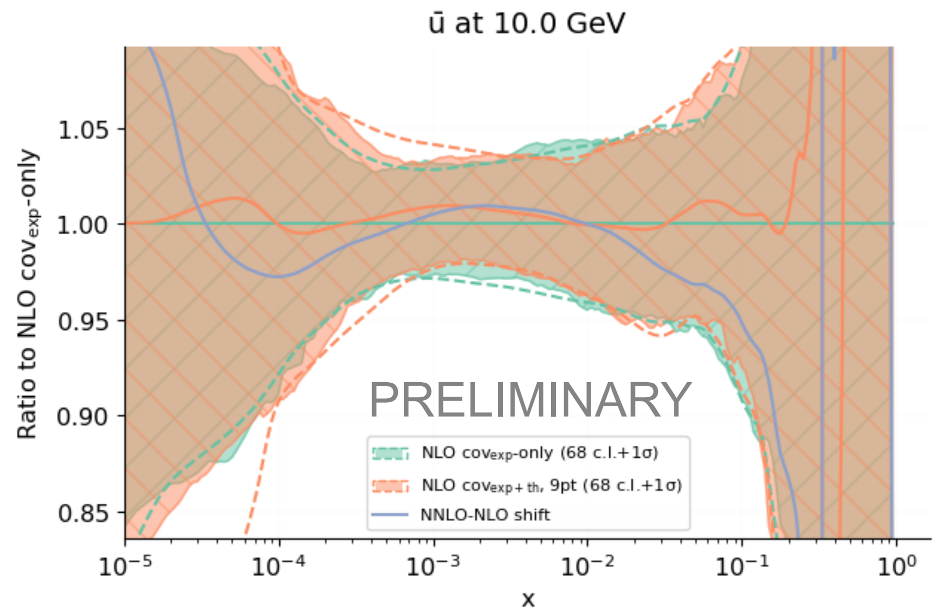
Process	Efficiency, ε
DIS NC	0.48
DIS CC	0.71
DY	0.85
Jets	0.99
Top	0.98

Global: $\varepsilon = 0.54$

9-pt does best $\rightarrow$ use this for our PDF fits

Results: PDF fits with cov_{th}



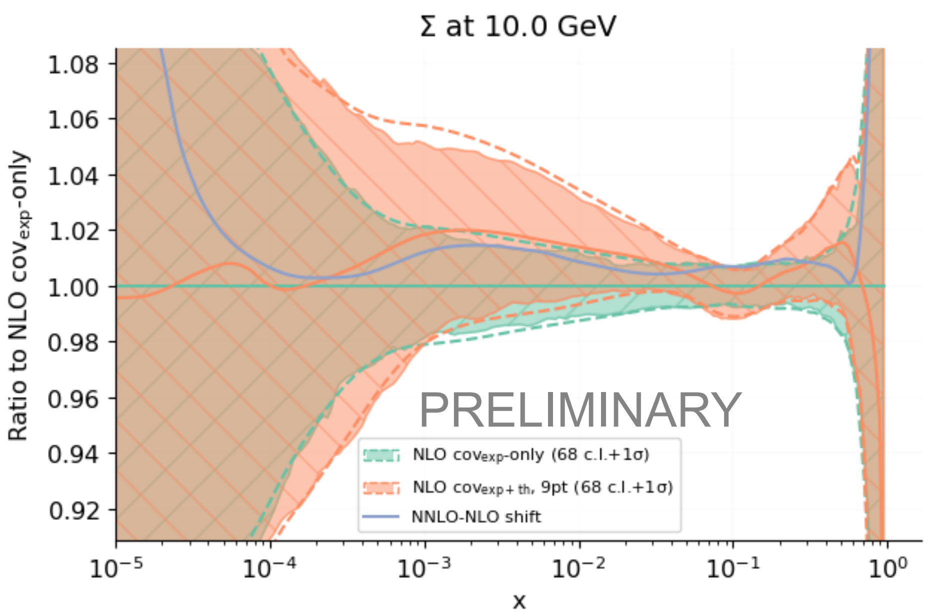


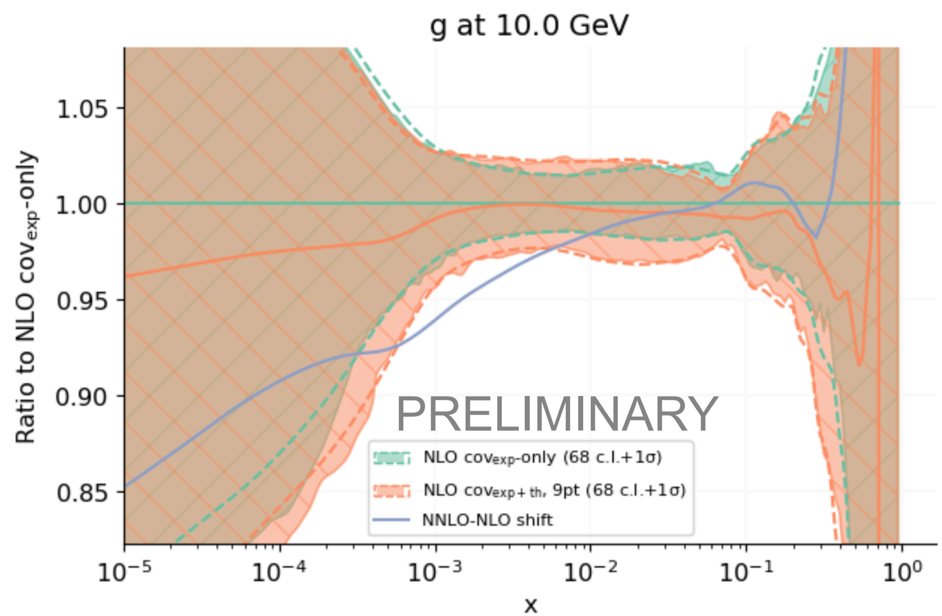
Shape of central value with cov_{th} resembles shift in data regions: **closer to true NNLO PDF**

Overall **small increase** in uncertainties (if at all): **tensions relieved**

- Increase in PDF uncertainties counteracted by change of data set weighting in fit: addition of MHOUs leads to **better fit**

Results: PDF fits with $\mathrm{cov_{th}}$





If NNLO-NLO shift is large while standard NLO PDF uncertainty is small:

- PDF uncertainty increases with addition of $\mathrm{cov_{th}}$
- **More reliable PDF uncertainties**

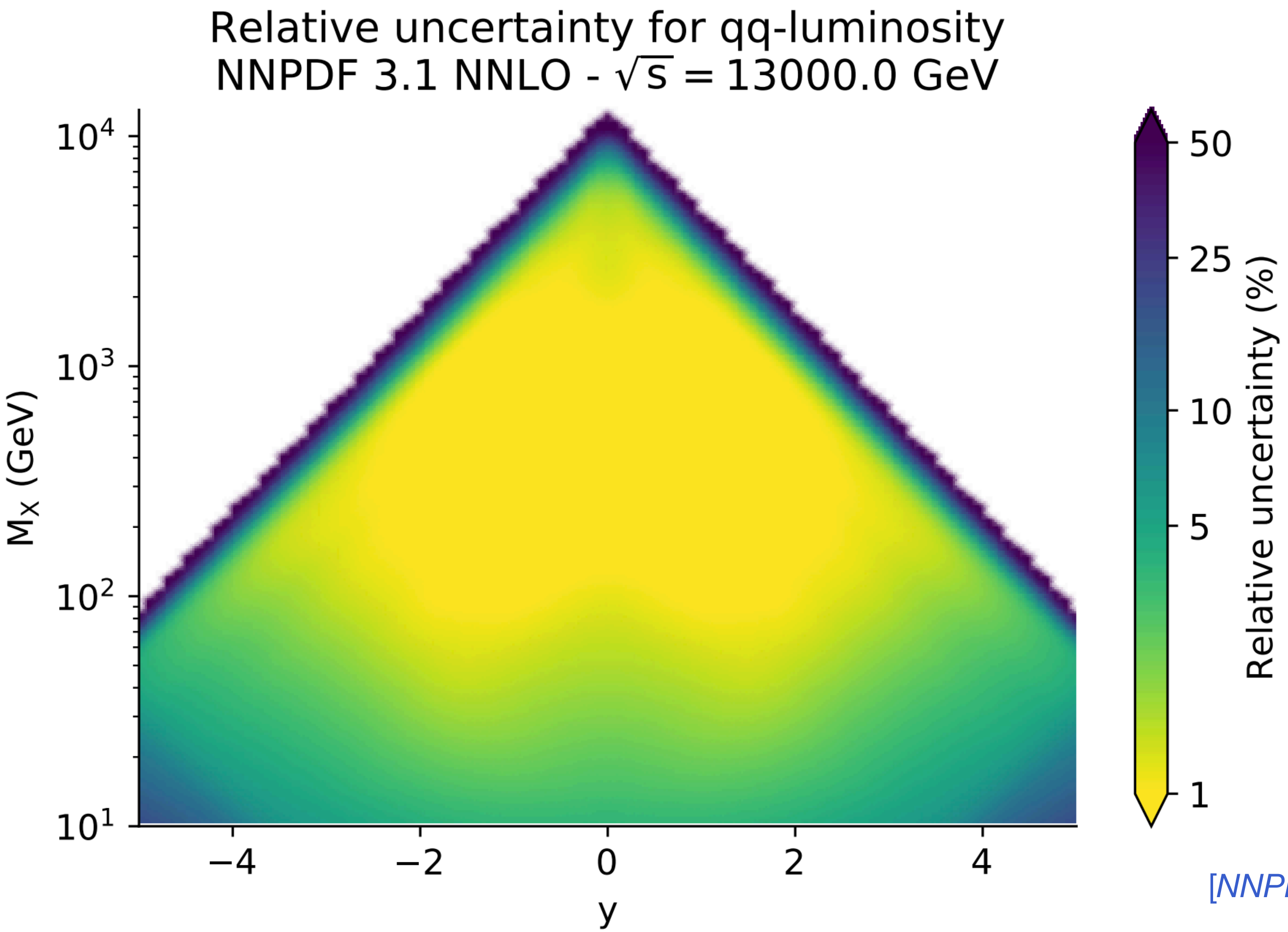
Conclusion and outlook

- Systematically including MHOUs in PDFs is now important, and will become crucial
- A new framework for including MHOUs in PDFs has been developed, based on **fitting with a theory covariance matrix**
- This is **validated** against **NNLO-NLO shift**
- Using this we have produced the **first PDF fits including MHOUs**, which are **more consistent** with NNLO PDFs than standard NLO fits
- Framework is applicable to **other sources of theoretical uncertainty**

Thank you for listening!

Extra slides

PDF uncertainties



[NNPDF, 2017]

Data set and cuts

The following datasets are included in both `NNPDF31_nlo_as_0118_1000` and `190302_ern_nlo_central_163_global`:

- HERA I+II inclusive NC $e^{+}p$ 920 GeV
- NMC p
- LHCb Z 940 pb
- CMS W rapidity 8 TeV
- D0 Z rapidity
- HERA I+II inclusive CC $e^{+}p$
- CDF Z rapidity
- ATLAS low-mass DY 2011
- CMS $\sigma_{tt}^{\rm tot}$
- HERA I+II inclusive NC $e^{+}p$ 820 GeV
- CHORUS $\sigma_{CC}^{\bar{\nu}}$
- ATLAS W, Z 7 TeV 2011
- ATLAS HM DY 7 TeV
- ATLAS $\sigma_{tt}^{\rm tot}$
- BCDMS d
- BCDMS p
- LHCb $W, Z \to \mu$ 8 TeV
- CMS W asymmetry 840 pb
- HERA I+II inclusive NC $e^{+}p$ 575 GeV
- NuTeV $\sigma_c^{\bar{\nu}}$
- HERA I+II inclusive NC $e^{+}p$ 460 GeV
- D0 $W \to e\nu$ asymmetry
- HERA I+II inclusive CC $e^{-}p$
- D0 $W \to \mu\nu$ asymmetry
- NMC d/p
- HERA $\sigma_c^{\rm NC}$
- SLAC d
- CMS Drell-Yan 2D 7 TeV 2011
- LHCb $W, Z \to \mu$ 7 TeV
- LHCb $Z \to ee$ 2 fb
- ATLAS $t\bar{t}$ rapidity y_t
- NuTeV σ_c^{ν}
- SLAC p
- ATLAS Z p_T 8 TeV (p_T^{ll}, M_{ll})
- CHORUS σ_{CC}^{ν}
- ATLAS Z p_T 8 TeV (p_T^{ll}, y_{ll})
- CMS jets 7 TeV 2011
- CMS $t\bar{t}$ rapidity $y_{t\bar{t}}$
- HERA I+II inclusive NC $e^{-}p$
- CMS Z p_T 8 TeV (p_T^{ll}, y_{ll})
- CMS W asymmetry 4.7 fb
- ATLAS W, Z 7 TeV 2010
- ATLAS jets 2011 7 TeV

Changes to cuts:

$$Q^2_{\rm min} = 3.49 \to 13.96\ {\rm GeV}^2$$

Intersection of NLO, NNLO cuts

The following datasets are included in `NNPDF31_nlo_as_0118_1000` but not in `190302_ern_nlo_central_163_global`:

- ATLAS jets 2.76 TeV
- CMS $W + c$ ratio
- DY E886 $\sigma^p_{\rm DY}$
- ATLAS jets 2010 7 TeV
- CMS jets 2.76 TeV
- HERA H1 F_2^b
- DYE 866 $\sigma^d_{\rm DY}/\sigma^p_{\rm DY}$
- CMS $W + c$ total
- DY E605 $\sigma^p_{\rm DY}$
- CDF Run II k_t jets
- HERA ZEUS F_2^b

Data removed:

- Fixed target Drell-Yan
- Bottom structure function
- Jets without exact NNLO theory
- W+charm

THEORY COVARIANCE MATRICES
SUBTLETIES I: DEFINITION

"STANDARD" DEFINITION OF SCALE VARIATION:
USE RG INVARIANCE OF PHYSICAL OBSERVABLE

- HADRONIC (HXSWG...): $\sigma(Q^2) = \sum_{ij} \hat\sigma_{ij}\left(\frac{Q^2}{\mu_F^2}, \frac{Q^2}{\mu_R^2}, \alpha_s(\mu_R^2)\right) f_i(\mu_F^2) f_j(\mu_F^2)$
  - FACTORIZATION: $f_i({\mu'_F}^2) = \left(1 + P_0 \ln \frac{{\mu'_F}^2}{{\mu'_F}^2}\right) f_i(\mu_F^2)$
  - RENORMALIZATION: $\alpha({\mu'}_r^2)\left(1 - \beta_0 \alpha \mu_R^2 \ln \frac{{\mu'_R}^2}{{\mu'_R}^2}\right)$
  - μ_F DEP IN PDF, μ_R DEP IN $\hat\sigma$
- DIS (Virchaux-Milsztajn, MRS, PEGASUS, APFEL,...):
  $F(Q^2) = \sum_i C_i\left(\frac{Q^2}{\mu_F^2}, \frac{Q^2}{\mu_R^2}, \alpha_s(Q^2)\right) f_i(\mu_F^2, \mu_R^2)$
  - FACTORIZATION: AS ABOVE
  - RENORMALIZATION: LET $\alpha(\mu_F^2) \to \alpha(\mu_R^2)$ IN EVOLUTION EQUATION
  - BOTH μ_R, μ_F VARIED IN PDF
- DIFFERENCE DIFFERENT NNLO TERMS GENERATED AT NLO
  "ADDITIVE" VS. "MULTIPLICATIVE"
  - DIS NLO $\ln \frac{\mu_R}{\mu_F}$, HADRONIC $\ln \frac{\mu_R}{Q} \ln \frac{\mu_F}{Q}$
  - DIS NLO $\beta_0 P_1$ TERMS, HADRONIC $\beta_0 + P_1$

$\Rightarrow$ ADOPT A COMMON PRESCRIPTION

[credit: *S. Forte, 2018*]

Correlating scale variations between PDFs and predictions

How to use these PDFs consistently in theoretical predictions?

Consider a situation when all data is at one scale. Let us only have evolution uncertainties, i.e. turn off uncertainties in hard cross sections

We have three scales:

- Q_0 : fitting scale of PDFs
- Q_{data} : scale of data
- $Q_{\text{pred.}}$: scale of prediction

→

We have two evolutions:

$Q_0 \to Q_{\text{data}}$

$Q_0 \to Q_{\text{pred.}}$

1. Q_0 is kept fixed. There is no dependence on Q_0 because for a sufficiently flexible parameterisation changes in Q_0 are absorbed by fit
2. We vary Q_{data} in fits (in a correlated way among data points)
3. One varies $Q_{\text{pred.}}$ when making a prediction for an observable

Correlating scale variations between PDFs and predictions

How are Q_{data} and $Q_{\text{pred.}}$ correlated?

- In our procedure Q_{data} and $Q_{\text{pred.}}$ variations will necessarily be uncorrelated - necessary consequence of delivering universal PDFs
- For points where $Q_{\text{data}} = Q_{\text{pred.}} \neq Q_0$, the variations are fully correlated and we overestimate uncertainty by factor of $\sqrt{2}$
- In global fit overestimate due to missing correlation will be between 1 and $\sqrt{2}$, but likely to be closer to 1
- **Importantly**: if one neglects either variation, one will **in general** underestimate MHOUs
- Better to have a conservative estimate of uncertainties than to underestimate them
- Same for coefficient function: if estimating μ_R uncertainty for process included in fit, we will miss correlations $\Rightarrow$ larger uncertainty than in ideal scenario
- **Not a double counting.** Instead, a problem of **missing correlation**

Theoretical covariance matrix

- Theory is perturbative expansion to some order : $t_p = \sum_{m=0}^{p} c_m$
- Standard case: $P(d|t_p) \propto \exp\Big(-\frac{1}{2}\underbrace{(d-t_p)^T \mathrm{cov}_{\mathrm{exp}}^{-1}(d-t_p)}_{\chi^2_{\mathrm{exp}}}\Big)$
- Bayes' theorem: $P(t_p|d) = \frac{P(d|t_p)P(t_p)}{P(d)} \propto P(d|t_p)P(t_p)$
- Assume Gaussian theory prior:

$$P(t_p) = \prod_{m=0}^{p} P(c_m) \quad \text{where} \quad P(c_m) \propto \exp\Big(-\frac{1}{2}\underbrace{c_m^T \mathrm{cov}_{\mathrm{th},m}^{-1} c_m}_{\chi^2_{\mathrm{th}}}\Big)$$

- Assume MHOUs due to O(α^{p+1}) terms only → marginalise these terms:

$$P(t_p|d) \propto \int dc_{p+1} P(d|c_{p+1})P(t_{p+1})$$

$$\propto \exp\Big(-\frac{1}{2}\underbrace{(d-t_p)^T(\mathrm{cov}_{\mathrm{exp}} + \mathrm{cov}_{\mathrm{th}})^{-1}(d-t_p)}_{\chi^2_{\mathrm{tot}}}\Big)$$

- Include higher order terms by induction

Validation: uncertainties + correlations

- We validate cov_{th} against exact result: **NNLO-NLO shift**
- We use fact that cov_{th} is **positive semi-definite** (eigenvalues > 0 or 0)

Procedure:

1. Find N_s non-zero eigenvectors, e_i^α, and eigenvalues, $\lambda^\alpha = (s^\alpha)^2$, of cov_{th}
2. Compute shift vector: $\delta_i = t_i^{\text{NNLO}} - t_i^{\text{NLO}}$ (fixed NLO PDFs)
3. Project shift vector onto eigenvectors:

$$\delta^\alpha = \sum_{i=1}^{N_{\text{dat}}} \delta_i e_i^\alpha$$

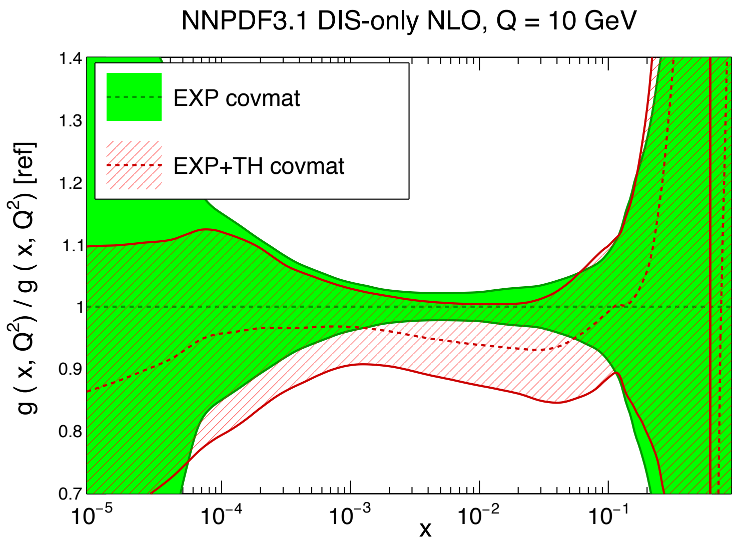
4. Define **efficiency**:

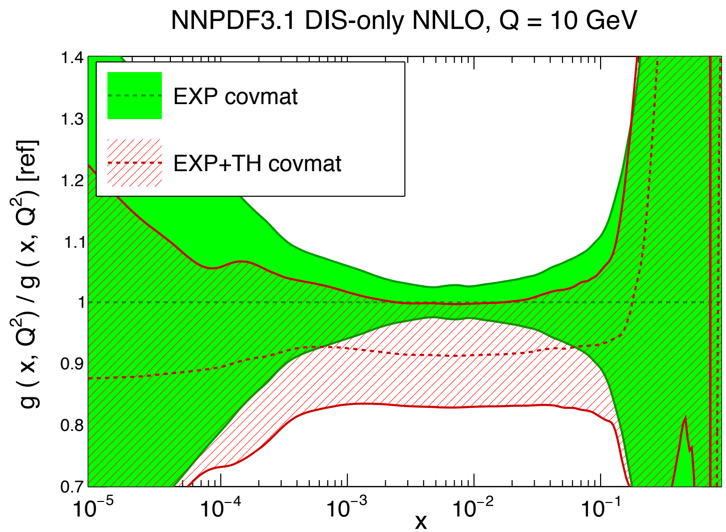
$$\varepsilon = \frac{|\sum_{\alpha=1}^{N_s} \delta^\alpha e^\alpha|}{|\delta|}$$

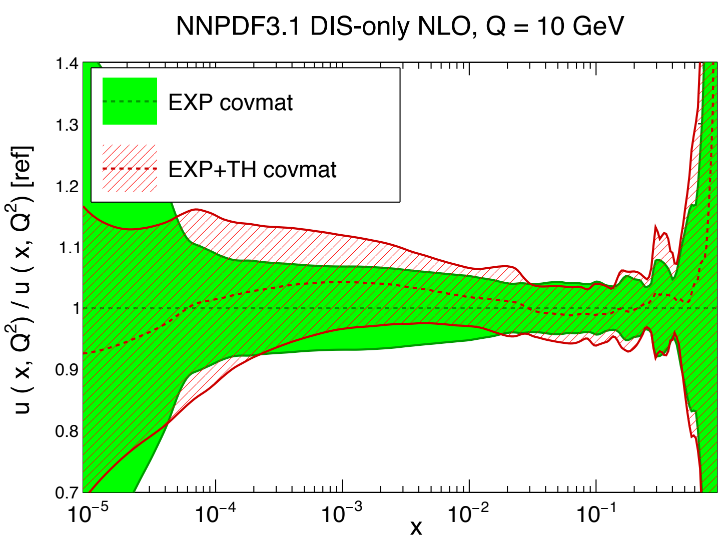
$$0 \le \varepsilon \le 1$$

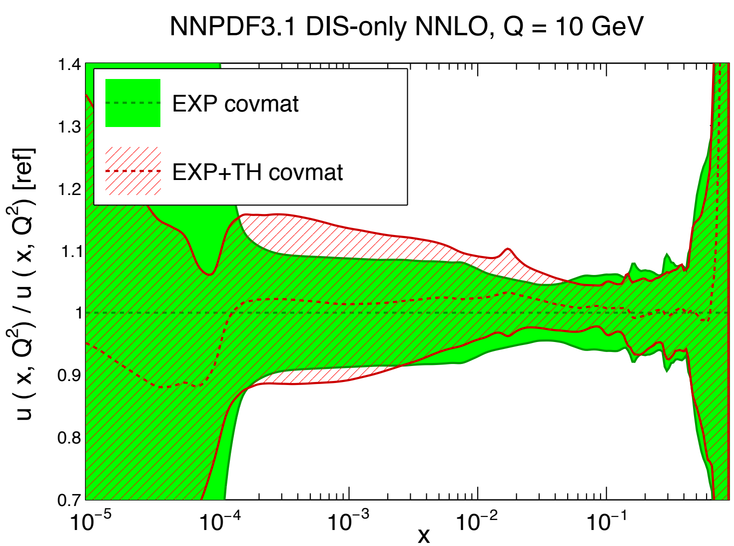
$\varepsilon = 1$: cov_{th} predicts variation in same directions as shift

DIS-only fits with cov_{th}









Impact of theory correlations on fits