

European Research Council

Established by the European Commission



MARIA UBIALI
UNIVERSITY OF CAMBRIDGE

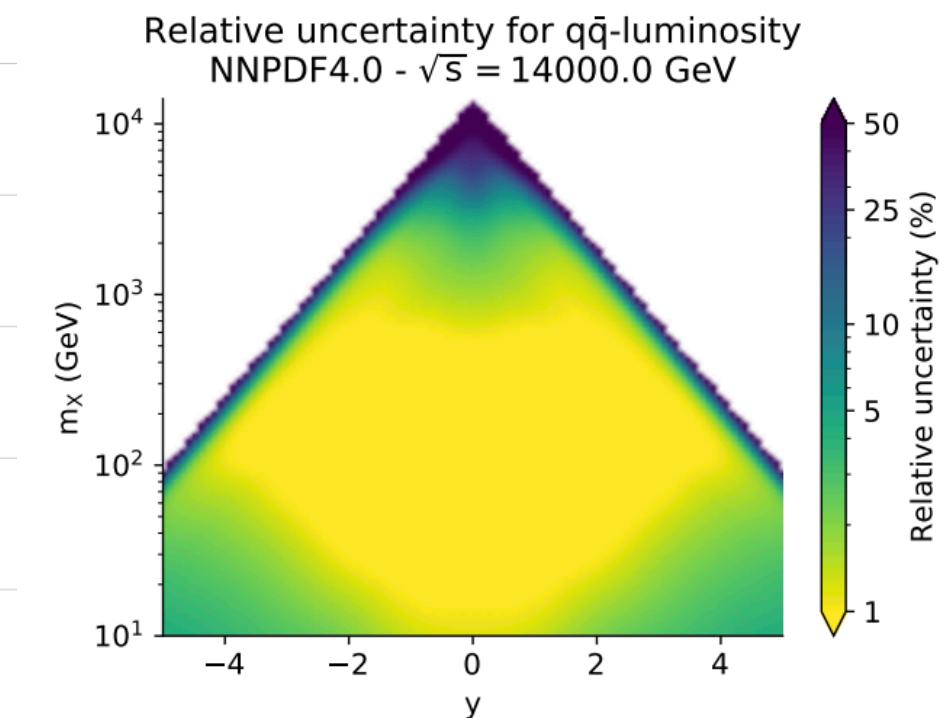
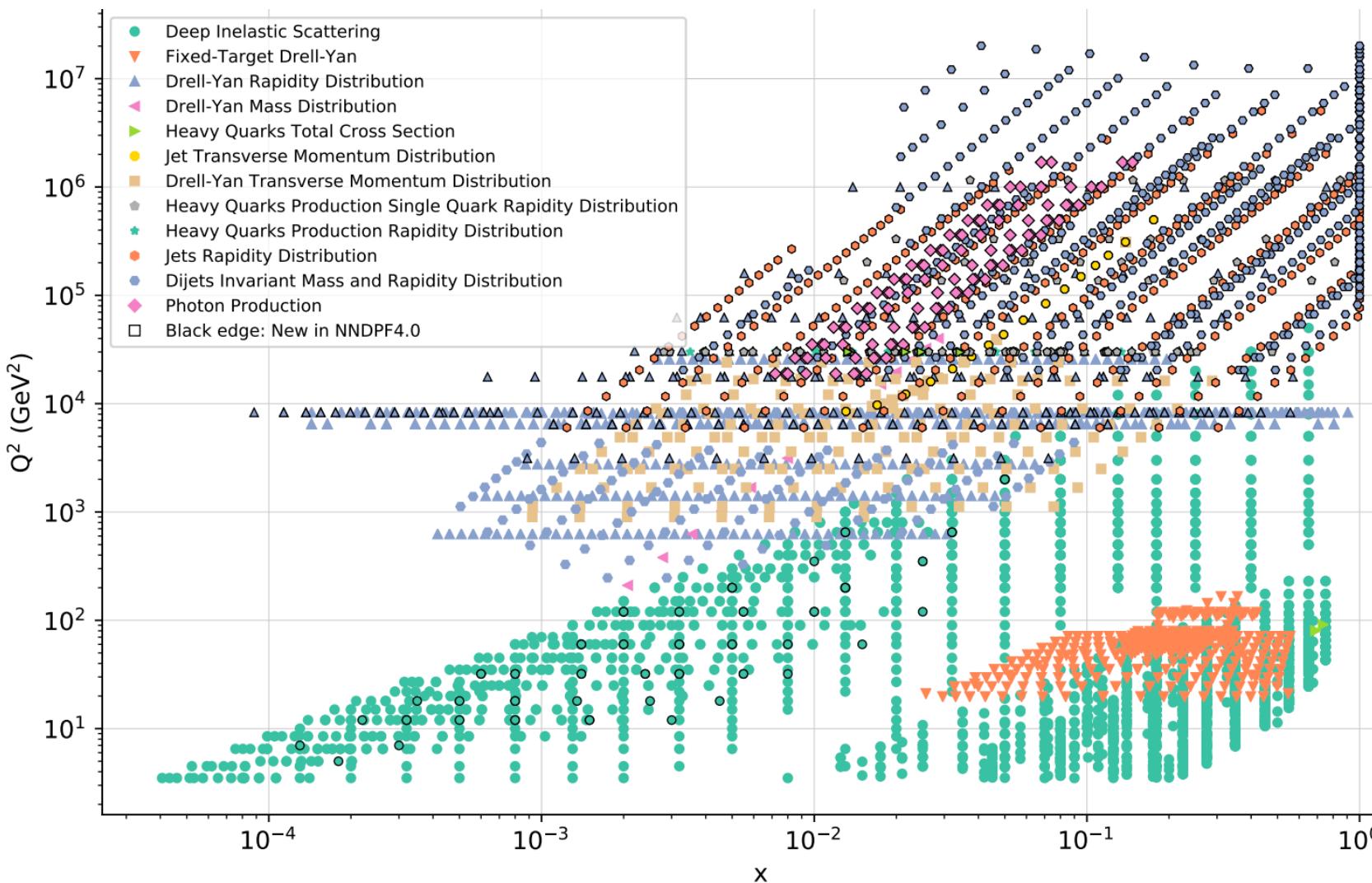
NEW AVENUES IN PDF DETERMINATION

HIGH PRECISION FOR HARD PROCESSES AT THE LHC - NEWCASTLE, UK

22ND SEPTEMBER 2022

PDFS AND LHC PRECISION PROGRAM

✓ Abundance of precise LHC data allows to extract PDFs with unprecedented precision.



At current level of precision, issues that were considered unimportant become crucial, and a new level of rigour and accuracy is needed.

OUTLINE

- Precision and accuracy: the NNPDF4.0 case
 - ➡ Prior probability in PDF fits
 - ➡ Response to the “hopscotch” study
- Missing higher orders uncertainties
 - ➡ Theory Covariance Matrix approach
 - ➡ MCscales approach
- Beyond the standard proton
 - ➡ SMEFT fits and PDF fits interplay
- Conclusions and outlook

PRECISION AND ACCURACY: THE NNPDF4.0 CASE

PRIOR PROBABILITY IN PDF FITS

- ✓ PDF fitting example of inverse problem: aim to find a posterior probability of \mathbf{f} given the data \mathcal{D} .

$$p(f|D) \propto p(D|f) p(f)$$

- ✓ Parametrization of PDFs: finite-dimensional problem.

$$f(x) \approx \tilde{f}(x, \theta) \in \mathcal{F}$$

- ✓ The posterior probability for the parametrization depends on both the figure of merit that maximises the data likelihood given the parameters and on prior probability \mathcal{H} .

$$p(\theta|D, \mathcal{H}) \propto p(D|\theta, \mathcal{H}) p(\theta|\mathcal{H})$$

$$= \exp(-\mathcal{L}(\theta, D)) p(\theta|\mathcal{H})$$



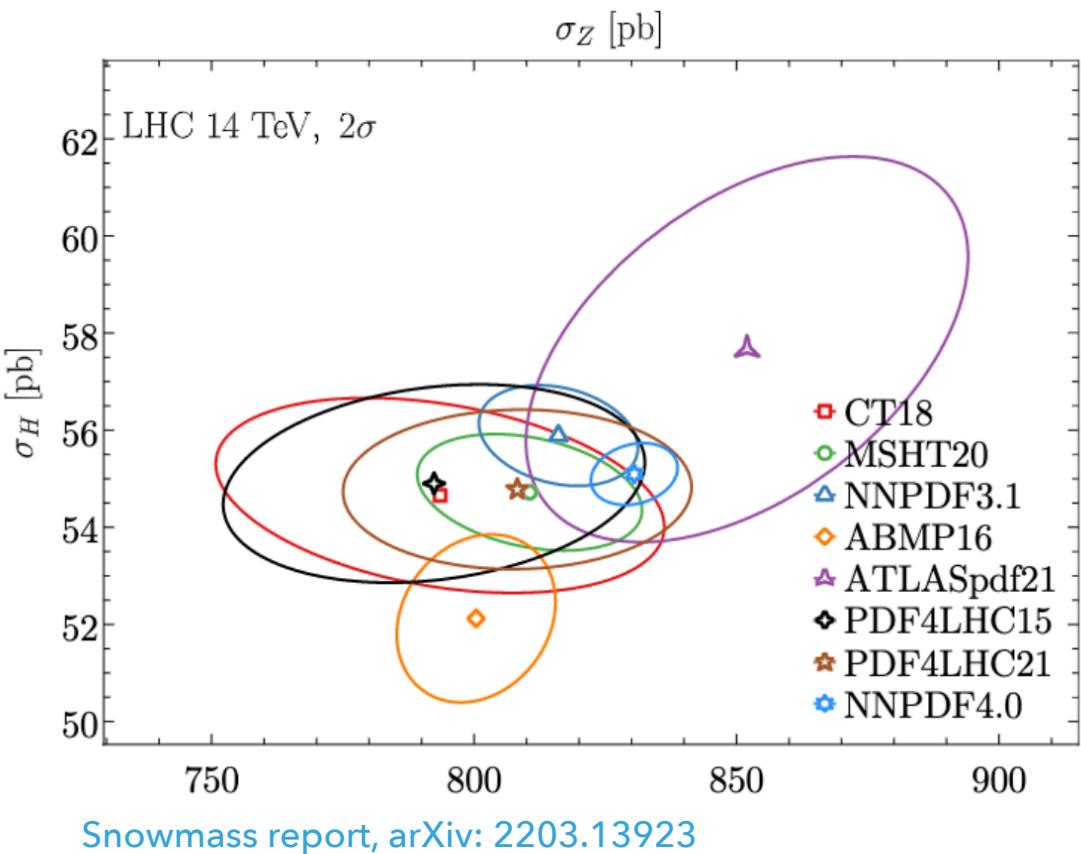
$$\chi^2 = \frac{1}{N_{\text{dat}}} \sum_{i=1}^{N_{\text{dat}}} (T_i(\{\theta\}, \{c\}) - D_i) \text{cov}_{ij}^{-1} (T_j(\{\theta\}, \{c\}) - D_j)$$

$$\text{cov}_{ij} \equiv \text{cov}_{ij}^{t_0} = \left(\sum_{l=1}^{N-N_{\text{norm}}} \sigma_{i,l} \sigma_{j,l} \right) T_i T_j + \left(\sum_{m=1}^{N_{\text{norm}}} \sigma_{i,m} \sigma_{j,m} \right) T_i^{(0)} T_j^{(0)}$$

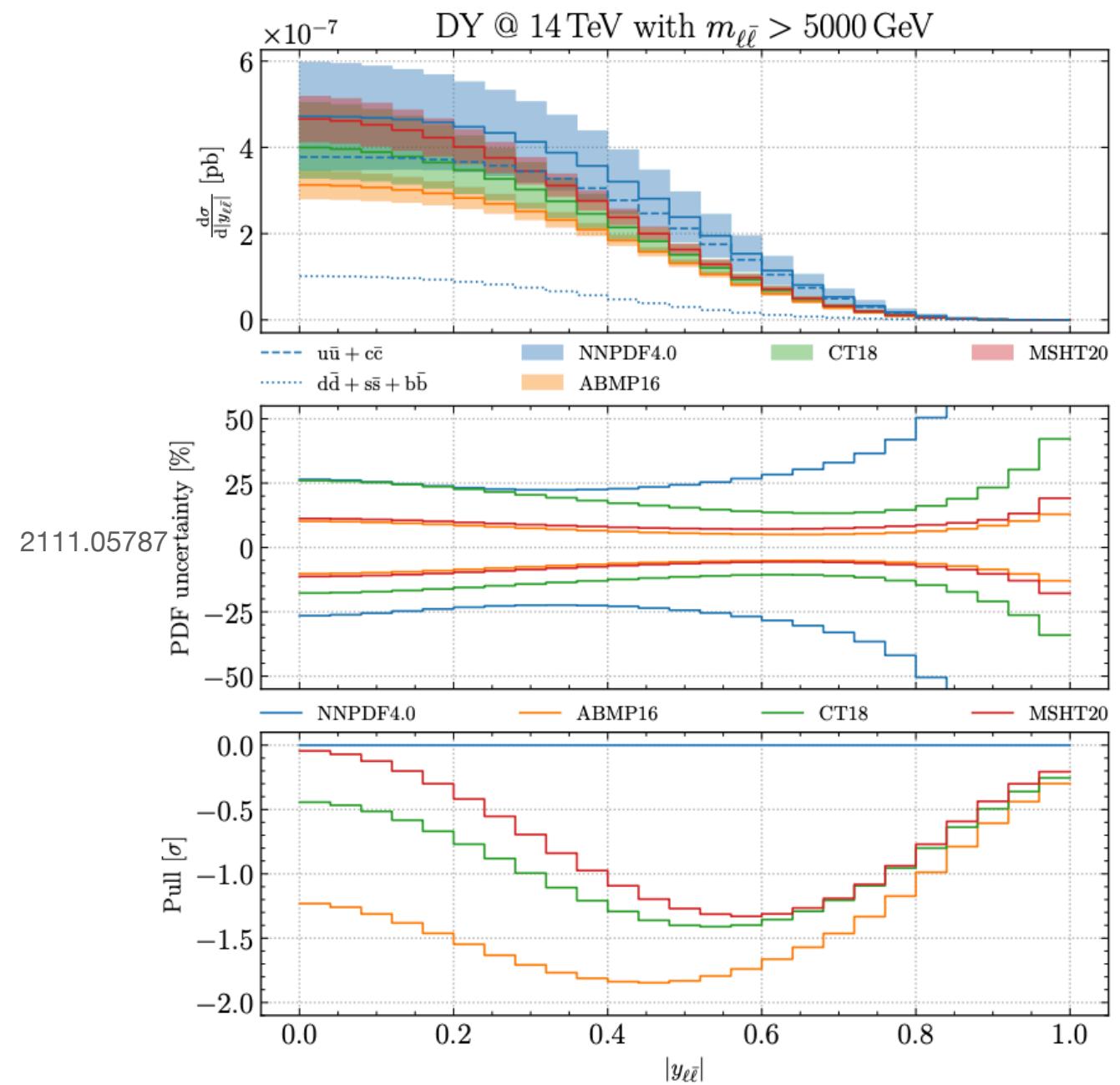
Ball et al, arXiv:0912.2276

Prior: functional form, integrability, positivity, sum rules, behaviour at small-x and large-x...

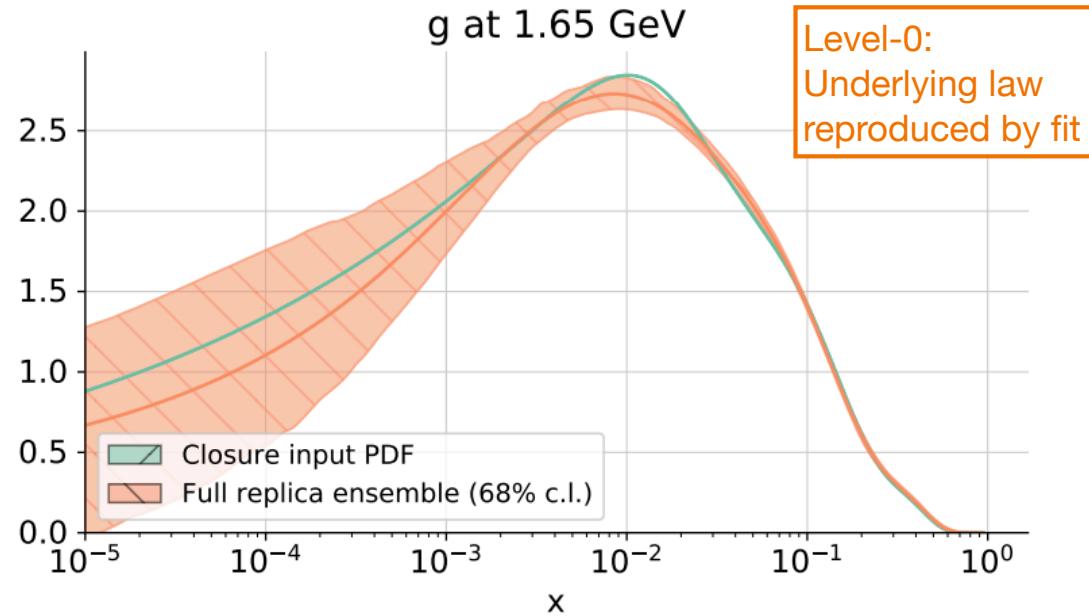
THE NNPDF4.0 PDFS



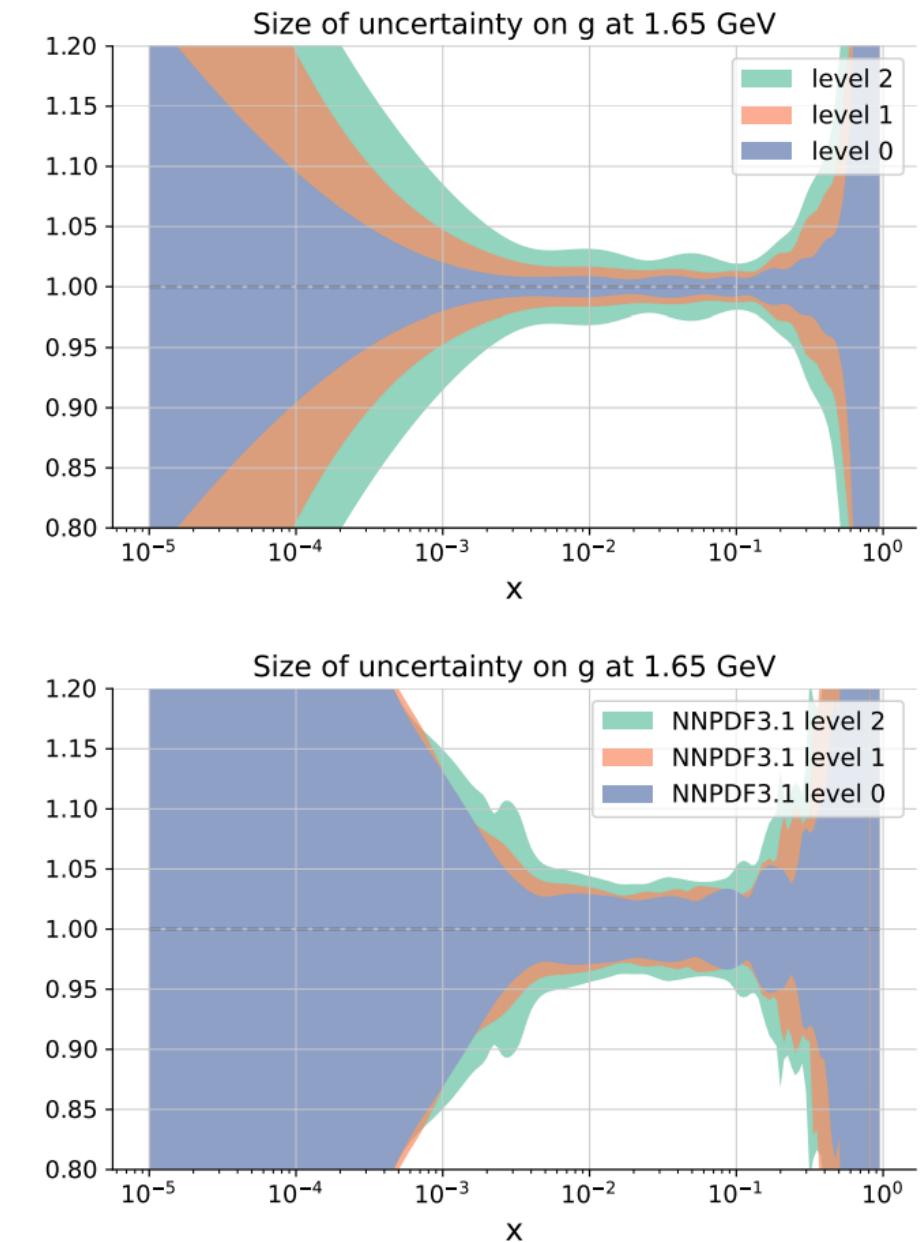
- ✓ NNPDF4.0: Small uncertainty in data region (due to large number of experimental data included and hyper-optimised methodology)
- ✓ Larger uncertainty in the extrapolation region (due to flexibility of NN parametrization)



TESTING THE NNPDF4.0 METHODOLOGY

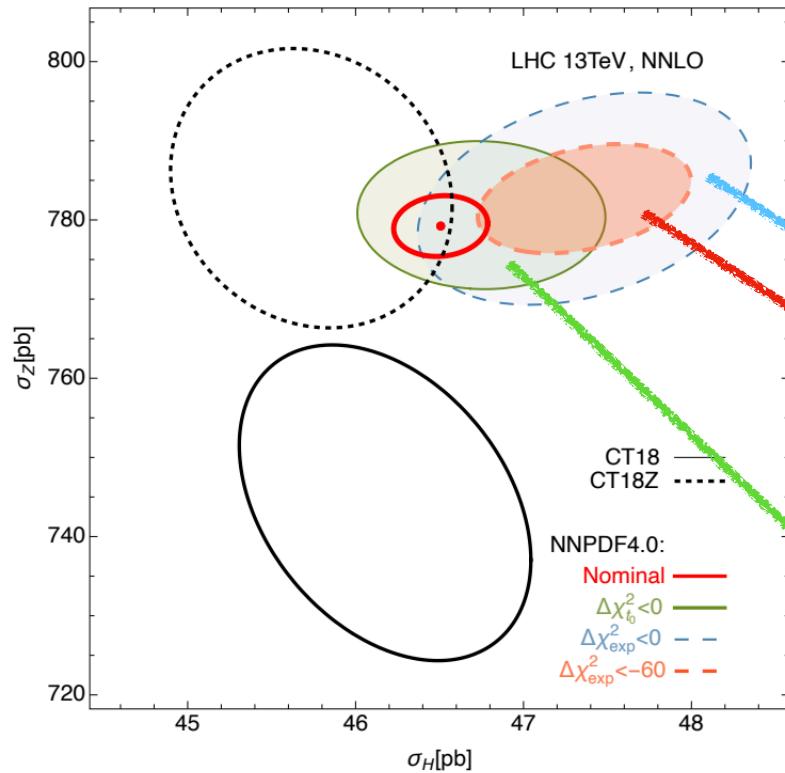


- ✓ Methodology efficiency and PDF uncertainties tested via closure test (in the data region) [[Del Debbio, Giani, Wilson, 2111.05787](#)] and future test (in the extrapolation region) [[Cruz-Martinez, Forte, Nocera, 2103.08606](#)].
- ✓ Closure tests assess faithfulness of uncertainty estimate. They do not account for theory uncertainties nor for data inconsistencies.



THE HOPSCOTCH STUDY: QUESTIONS

- ✓ A "hopscotch" scan made to search for solutions with equal or better χ^2 by building ad-hoc linear combinations of NNPDF4.0 Hessian e-vectors [Courtoy et al, arXiv: 2205.10444].
- ✓ Is the sampling of the PDF uncertainty of an experimental observable truly representative of all acceptable solutions?

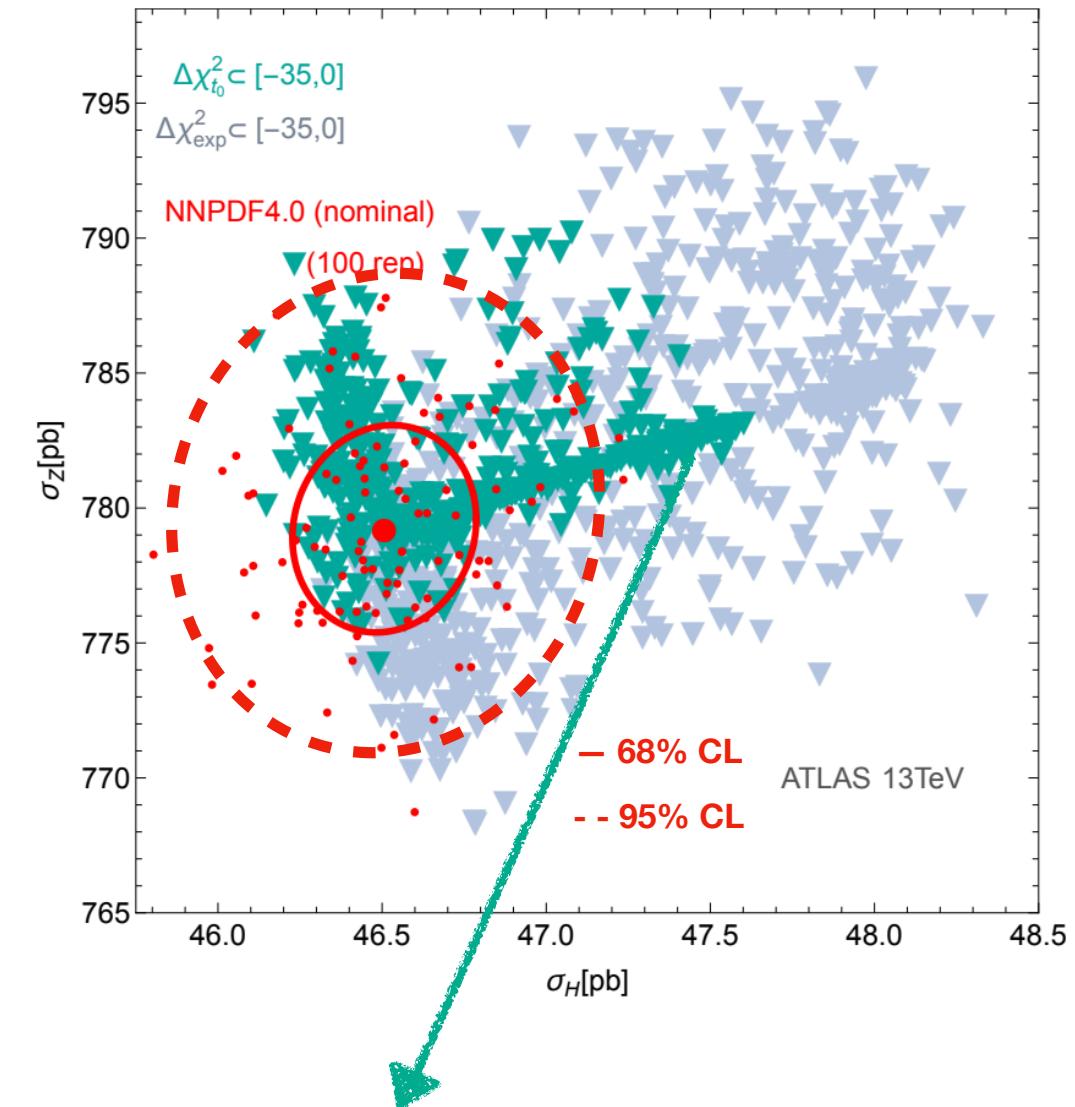


✓ Note: only black and red ellipses have statistical meaning

Irrelevant: not the χ^2 used to fit NNPDF replicas

Relevant: are there NNPDF replicas that look like the hopscotch (HS) PDFs? If so, why do they have low probability?

Courtoy et al, arXiv: 2205.10444



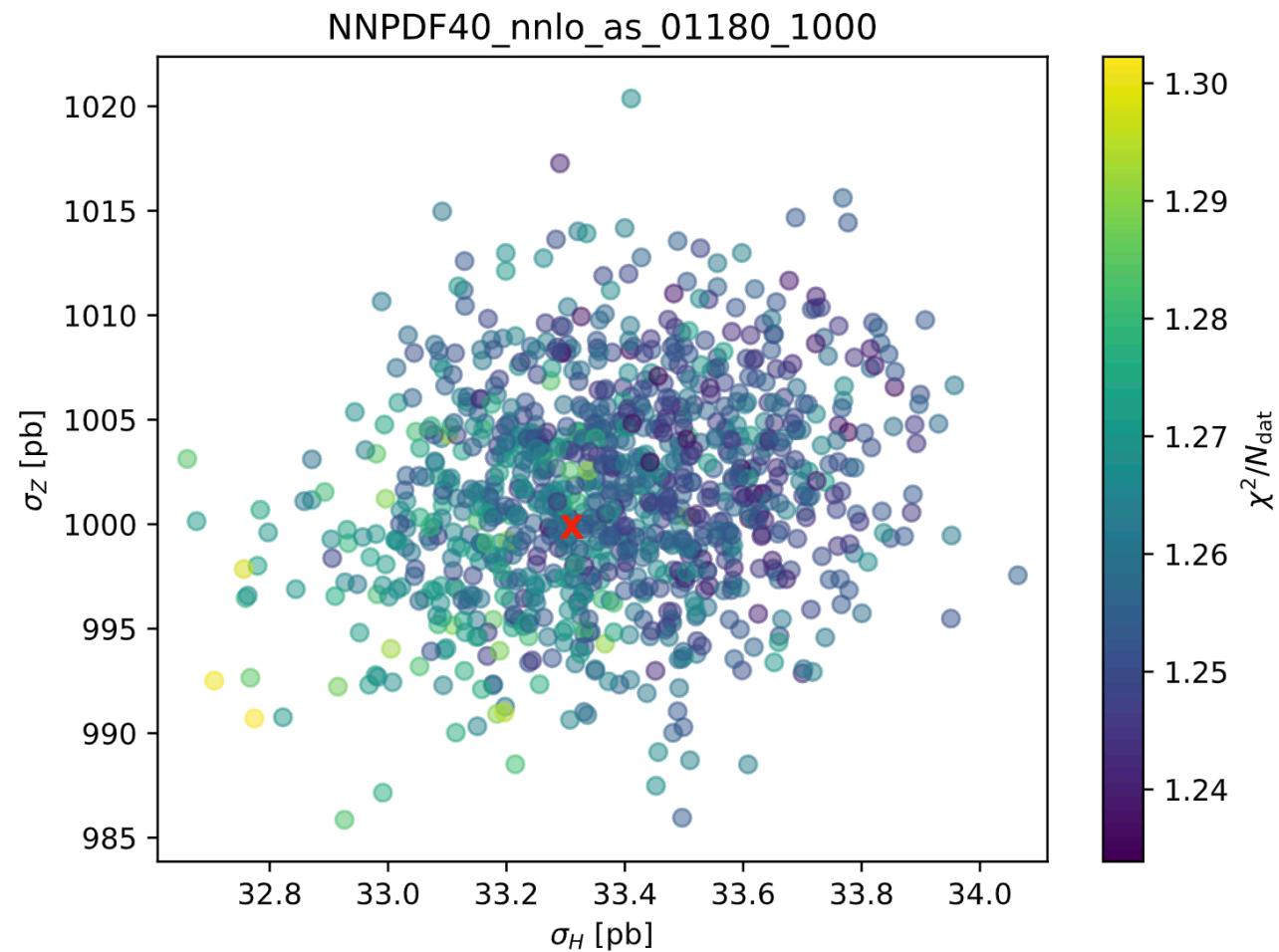
THE HOPSCOTCH STUDY: ANSWERS

Preliminary observations on MC sampling

- ✓ The χ^2 of the central replica (average of PDF reps) is not necessarily the min χ^2 of the replicas
- ✓ Given that each PDF replica is the best fit of a different MC pseudo-data, the $\chi^{2(k)}$ does not increase monotonically moving away from the $\chi^{2(0)}$ of the central replica.

$$D^{(k)} = D^{(0)} + \eta + \epsilon^{(k)}$$

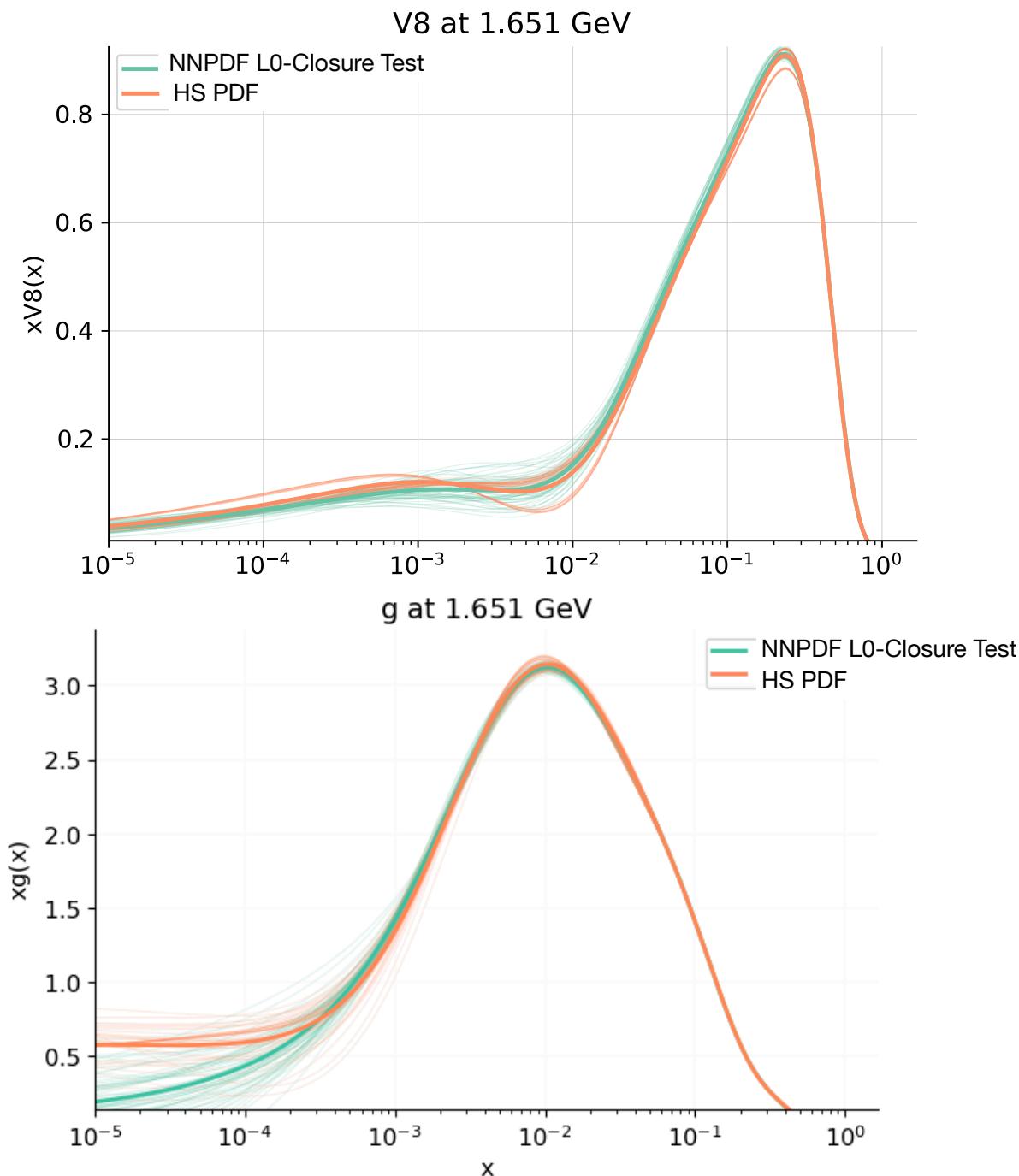
$$\eta, \epsilon \sim \mathcal{N}(0, \text{cov})$$



THE HOPSCOTCH STUDY: ANSWERS

Q1: Are there PDF replicas that look like the PDFs picked in the HS scan? – Yes

- ✓ If HS PDFs are used as input PDF set in level-0 closure test (underlying law = one of the HS PDFs outside 95% C.L. of NNPDF4.0) the shape is reproduced in the data region (not the wiggles), not in the extrapolation region and get $\chi^2=0$
- ✓ NNPDF4.0 parametrization is flexible enough to reproduce the HS parametrization in the data region if HS was the underlying law.
- ✓ HS PDFs not forbidden by the NNPDF methodology

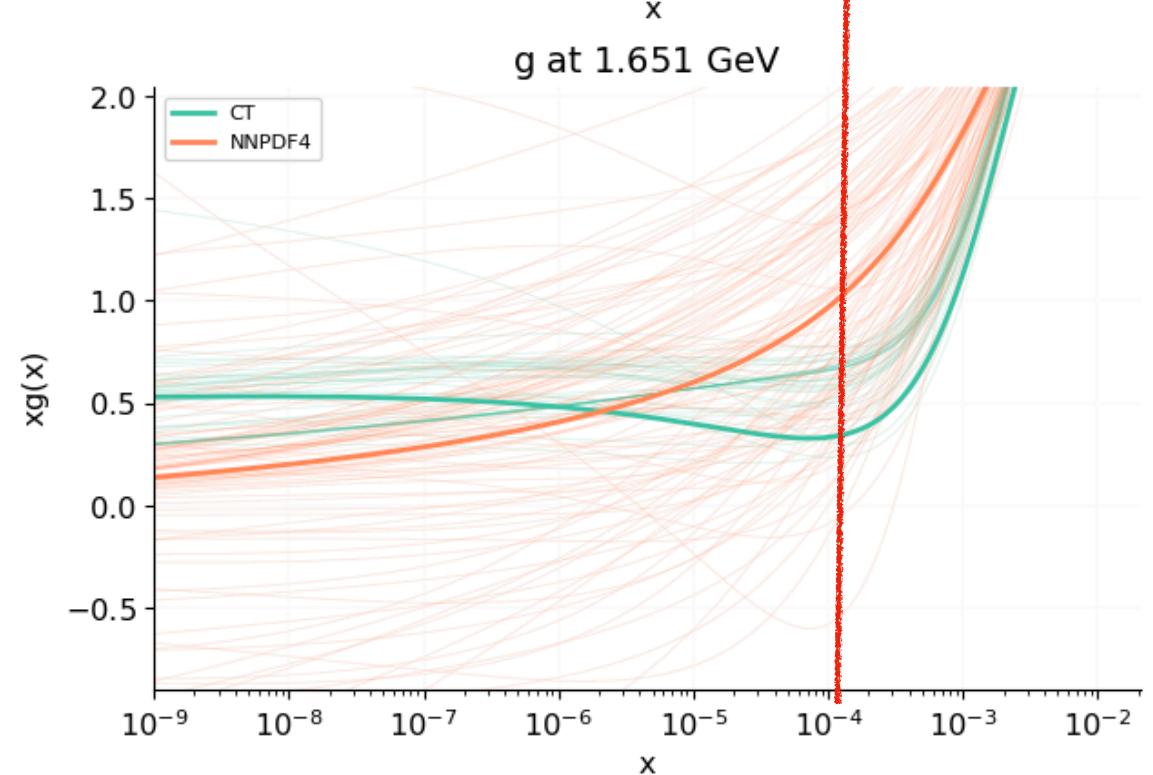
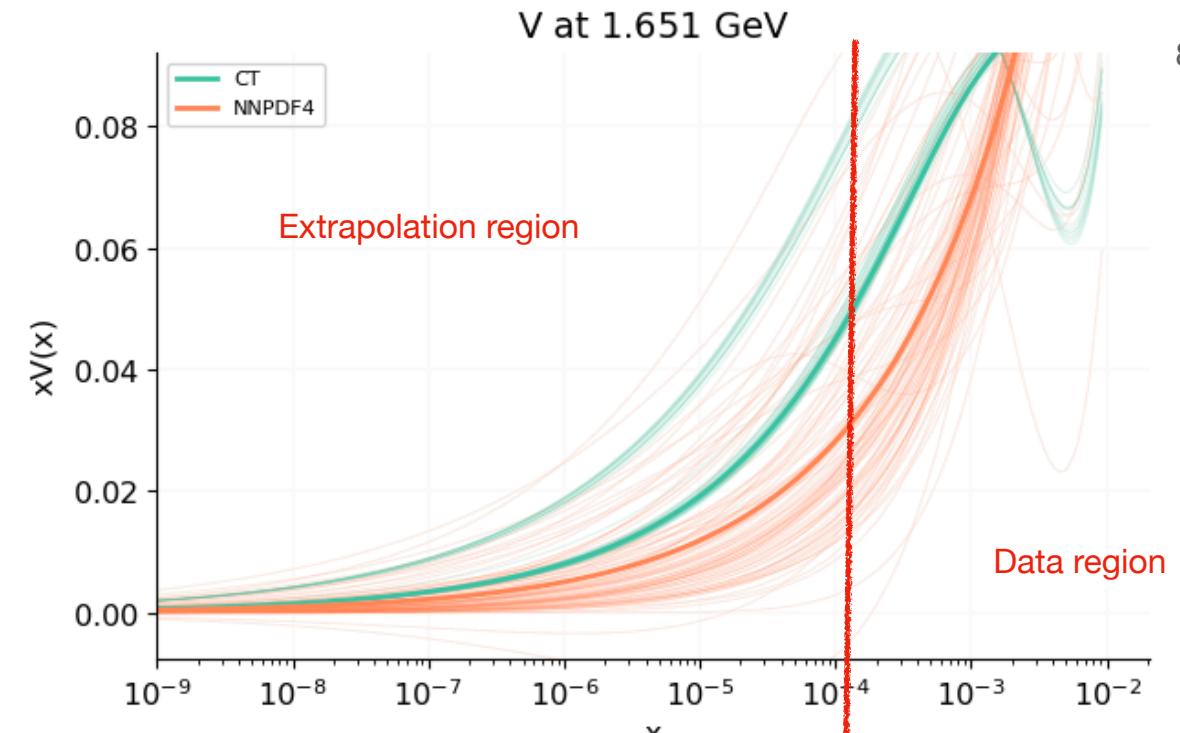


THE HOPSCOTCH STUDY: ANSWERS

8/22

Q2: Why are the HS PDFs unlikely in the NNPDF4.0 probability distribution?

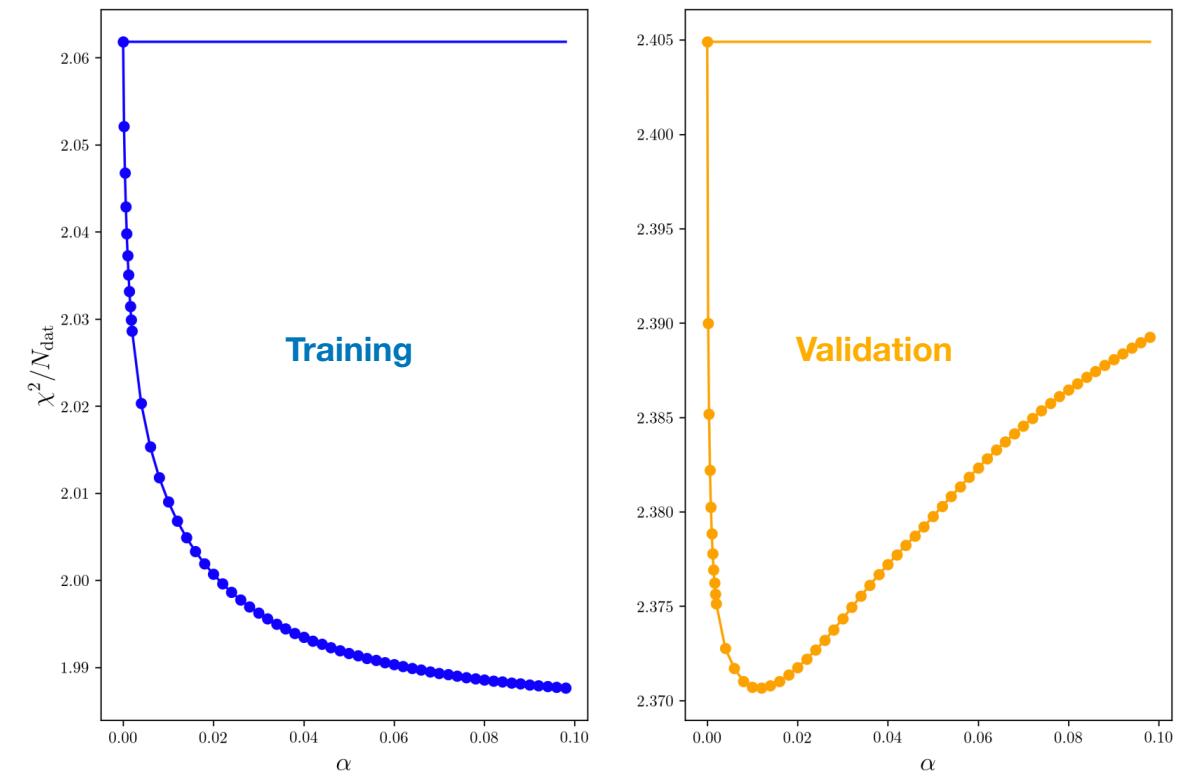
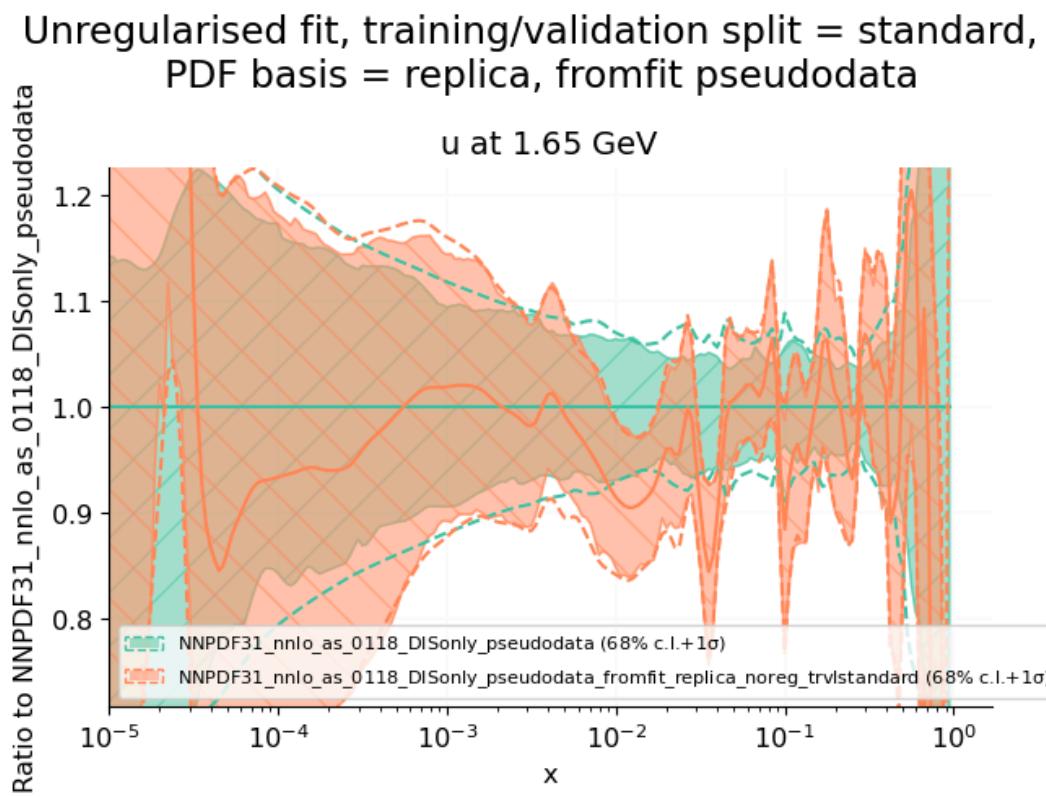
- ✓ The HS replicas that fall outside the NNPDF4.0 95% C.L. error bands (in green) display kink in gluon in the extrapolation region (below 10^{-4}).
- ✓ The prior \mathbf{H} disfavours this behaviour as it does not assume any structure in the extrapolation region.
- ✓ Also, notice signs of overfit in the data region (possibly to compensate kink in the extrapolation region).



THE HOPSCOTCH STUDY: ANSWERS

- More generally: if one fits (rather than scan) the coefficients of a linear combination of NNPDF replicas, can find a better χ^2 , but one must be careful with overfitting

[Kassabov, Moore, MU in preparation].



$$f^{(p)} = \bar{f}^{(p)} + \sum_{\substack{k=1 \\ k \neq p}}^{N_{\text{rep}}} w_k^{(p)} [\bar{f}^{(k)} - \bar{f}^{(p)}],$$

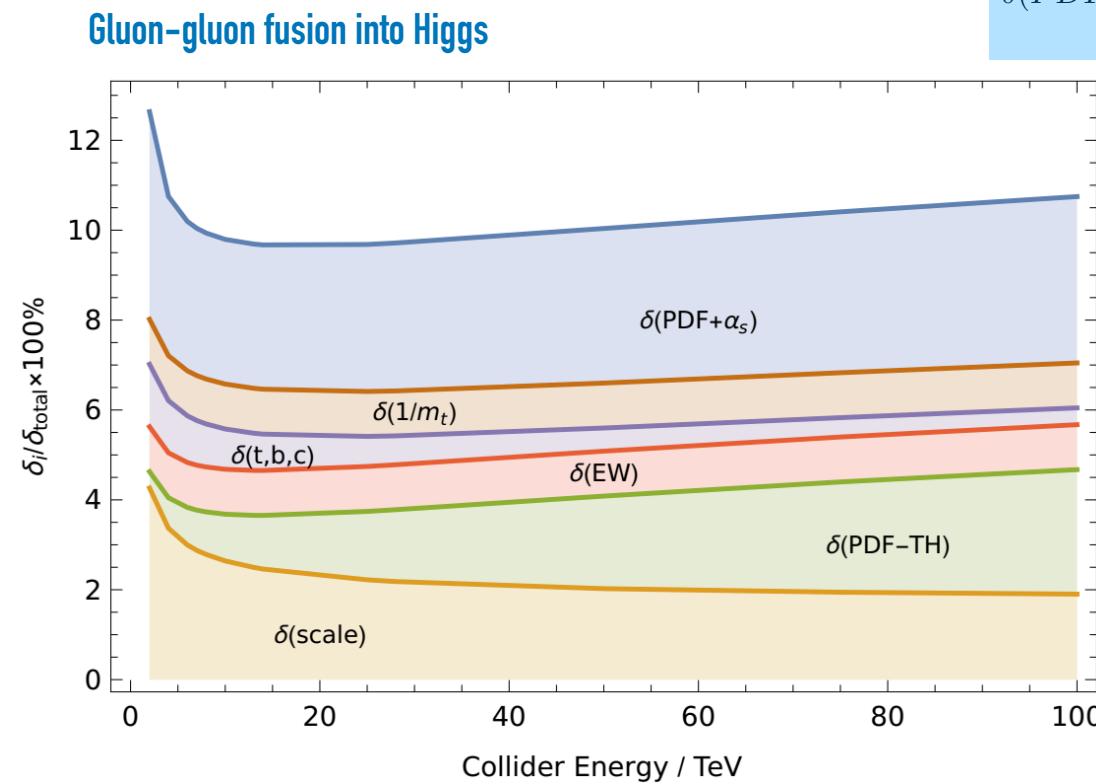
$$\bar{\chi}^2 = \chi^2 + \frac{1}{\alpha} \sum_k |w_k^{(p)}|^2$$

MISSING HIGHER ORDER UNCERTAINTIES

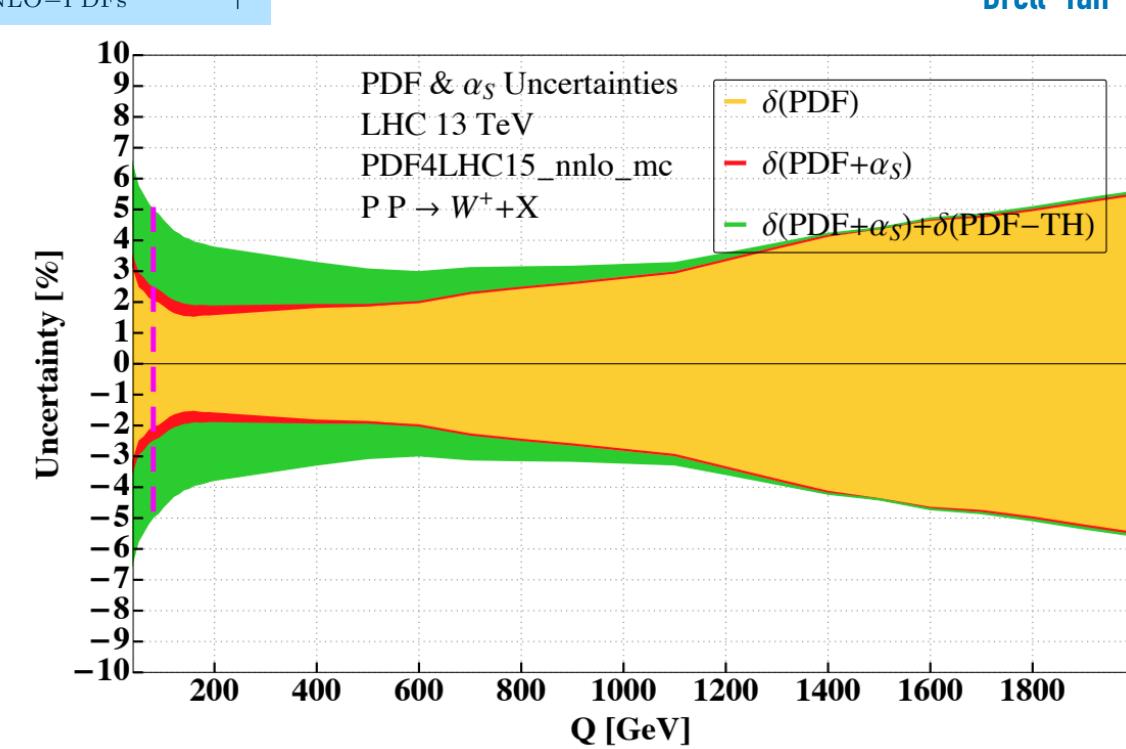
THEORY UNCERTAINTIES IN PDF FITS

$$\sigma = \alpha_s^p \sigma_0 + \alpha_s^{p+1} \sigma_1 + \alpha_s^{p+2} \sigma_2 + \mathcal{O}(\alpha_s^{p+3})$$

- ▶ Standard global PDF fits based on fixed-order NNLO QCD calculations (using fast interpolation grid for NLO predictions accompanied by local K-factors for NNLO). PDF uncertainty reflects experimental uncertainty.
 - ▶ N3LO is now the precision frontier for partonic cross sections (N3LO splitting functions partially known [\[R. Thorne's talk\]](#))
 - ▶ Mismatch between perturbative order of partonic cross section and PDFs becoming significant source of uncertainty



$$\delta(PDF - TH) = \frac{1}{2} \left| \frac{\sigma_{\text{NNLO-PDFs}}^{(2)} - \sigma_{\text{NLO-PDFs}}^{(2)}}{\sigma_{\text{NNLO-PDFs}}^{(2)}} \right|$$

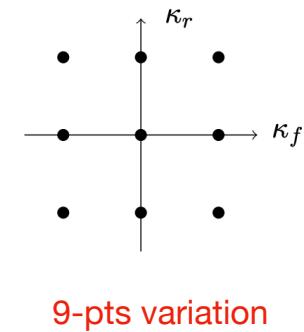
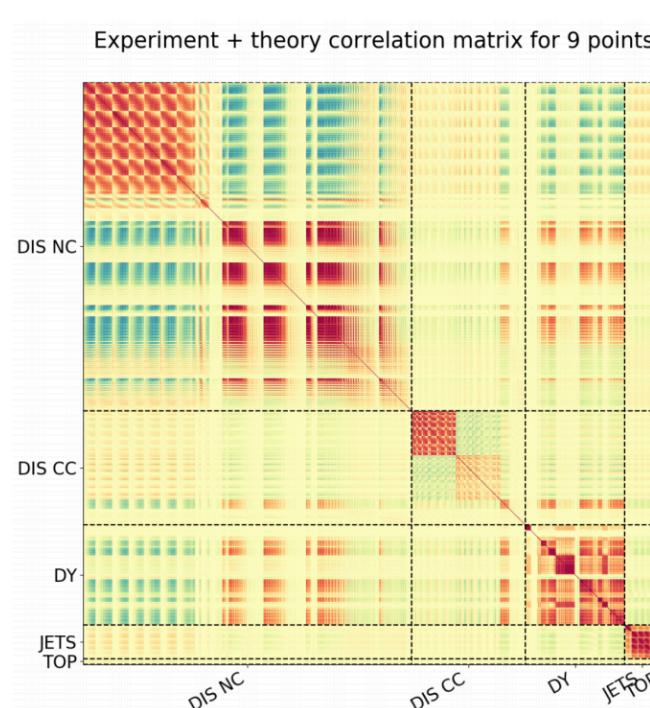
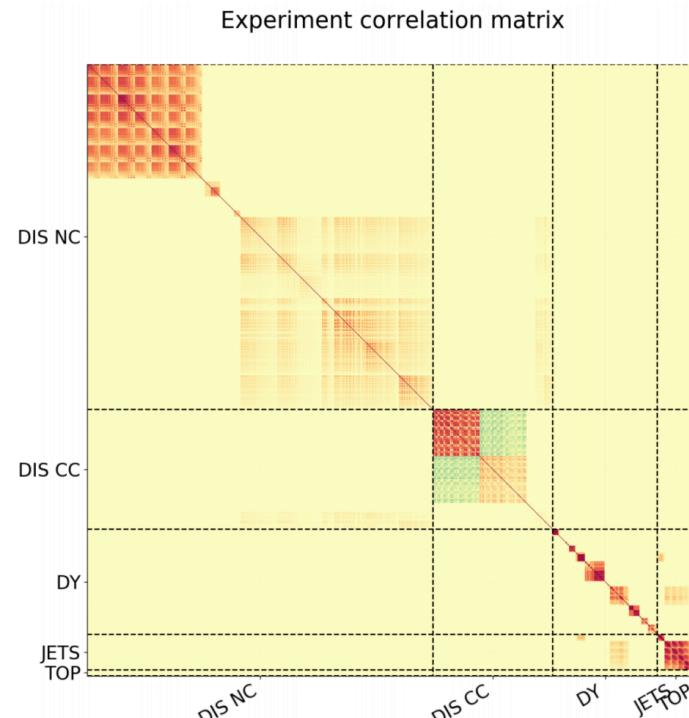


THE THEORY COVARIANCE MATRIX APPROACH

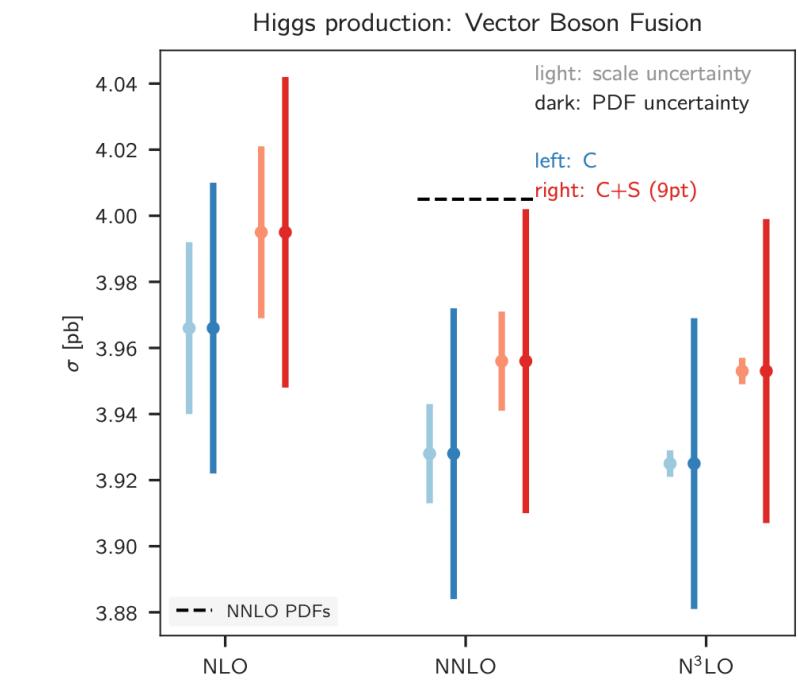
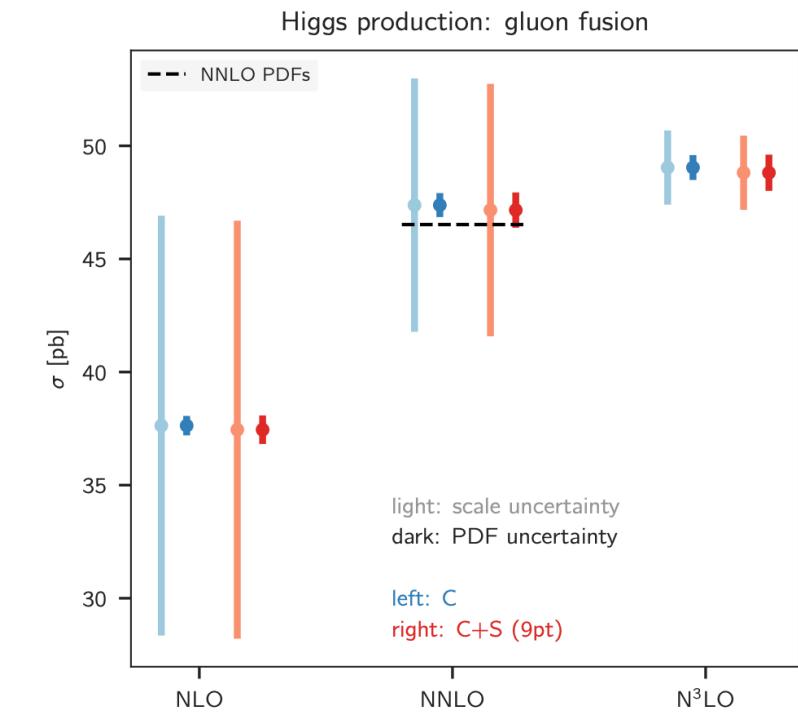
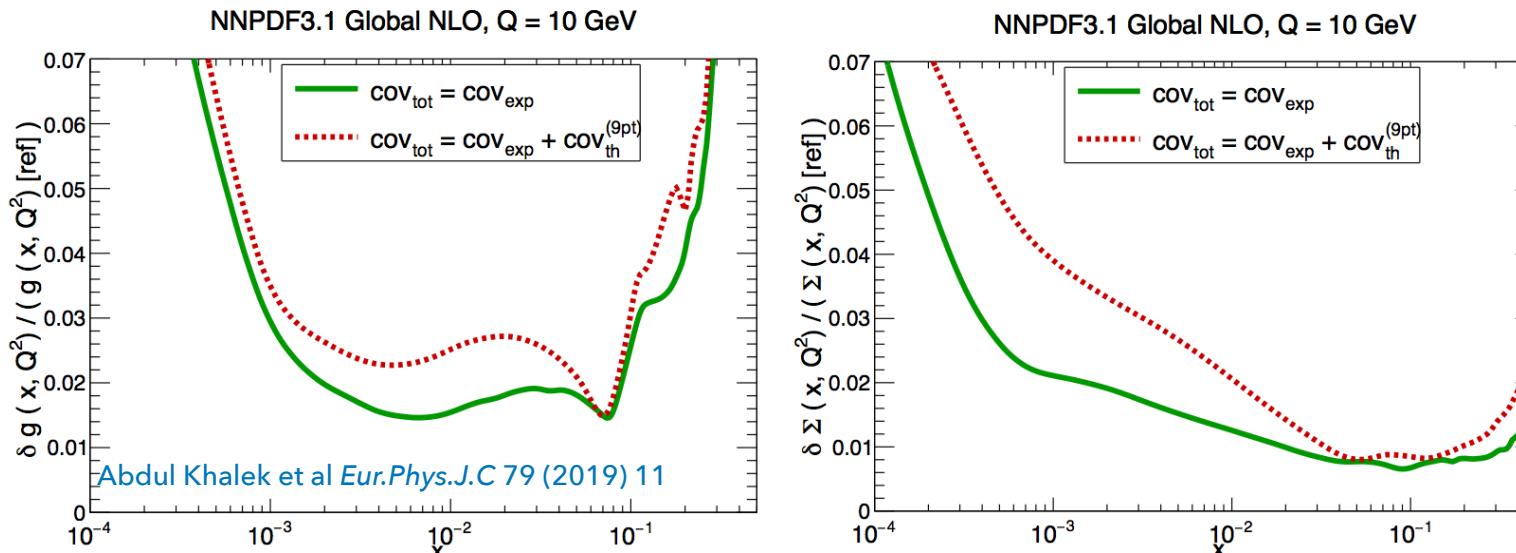
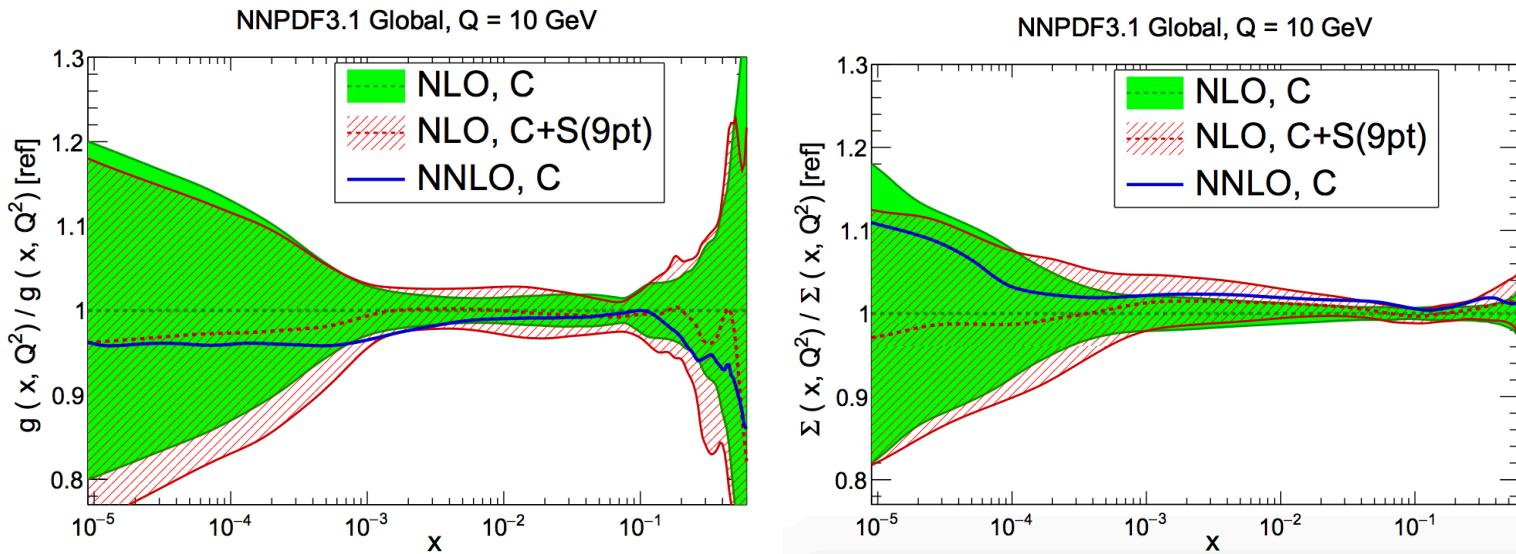
- How to include Missing Higher Order Uncertainties in a PDF fit? So far tested at NLO but soon there for NNLO & N3LO
- Construct a theory covariance matrix from scale-varied cross sections and combine it with the experimental covariance matrix [R.D. Ball et al, Eur.Phys.J.C 79 (2019) 11, 931, Eur.Phys.J. C (2019) 79:838]

$$\chi^2 = \sum_{m,n=1}^N (d_m - t_m)(\text{cov}_{\text{exp}} + \text{cov}_{\text{th}})^{-1}_{mn}(d_n - t_n)$$

- Assumptions: experimental and theoretical errors independent and Gaussian
- Accounting for the theory covariance matrix will modify the relative weight that each of the datasets carries in the global fit: processes with higher MHOUs will be “de-weighted”
- Assumptions on correlation of scales and scale ratio will determine the specific form of the covariance matrix

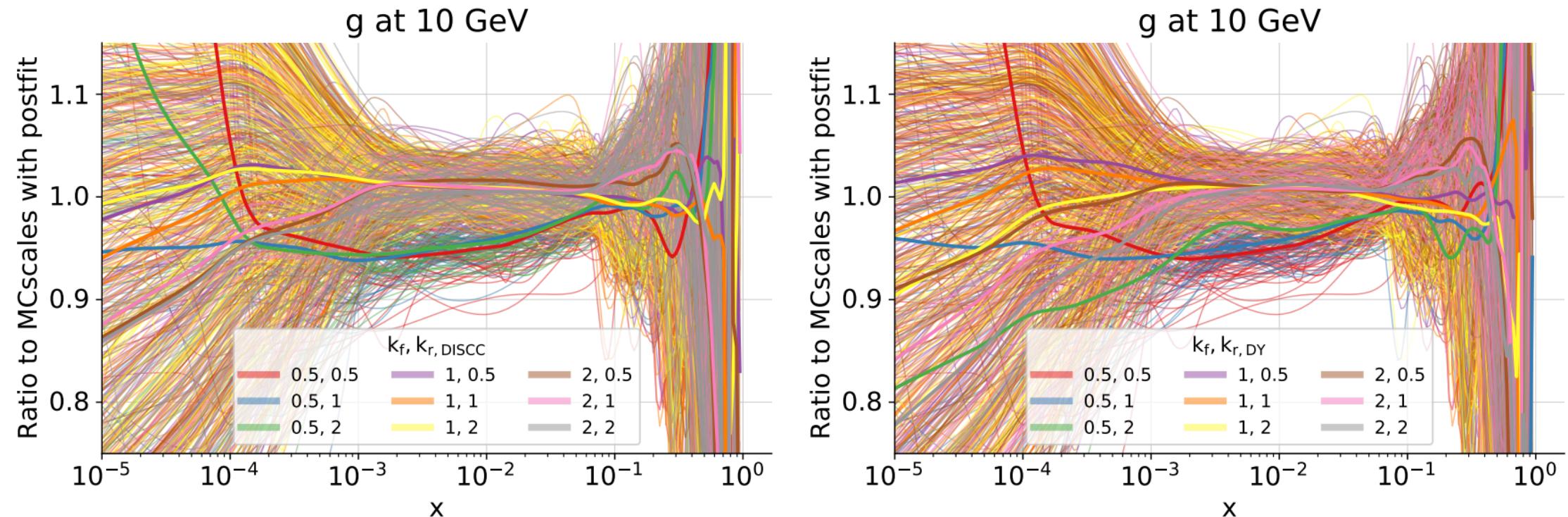


THEORY UNCERTAINTIES IN PDF FITS



→ Missing correlation between scale variation in PDF fits and hard cross sections
[Harland-Lang, Thorne Eur.Phys.J.C 79 (2019) 3, 225]

THE MCSCALES APPROACH



- ✓ Main idea of MCscales: the renormalisation and factorisation scales are free parameters of the fixed-order theory, that induce an uncertainty on the theory predictions included in a PDF fit & need to be propagated
- ✓ Joint sampling of experimental uncertainty (propagated to PDF uncertainty by MC sampling) by specifying a suitable prior probability distribution of all possible scale choices & a-posteriori criterion based on agreement with the data.

$$P\left(k_f = \xi_f, k_{r_1} = \xi_1, \dots, k_{r_{N_p}} = \xi_{N_p}\right) = P(\omega)$$

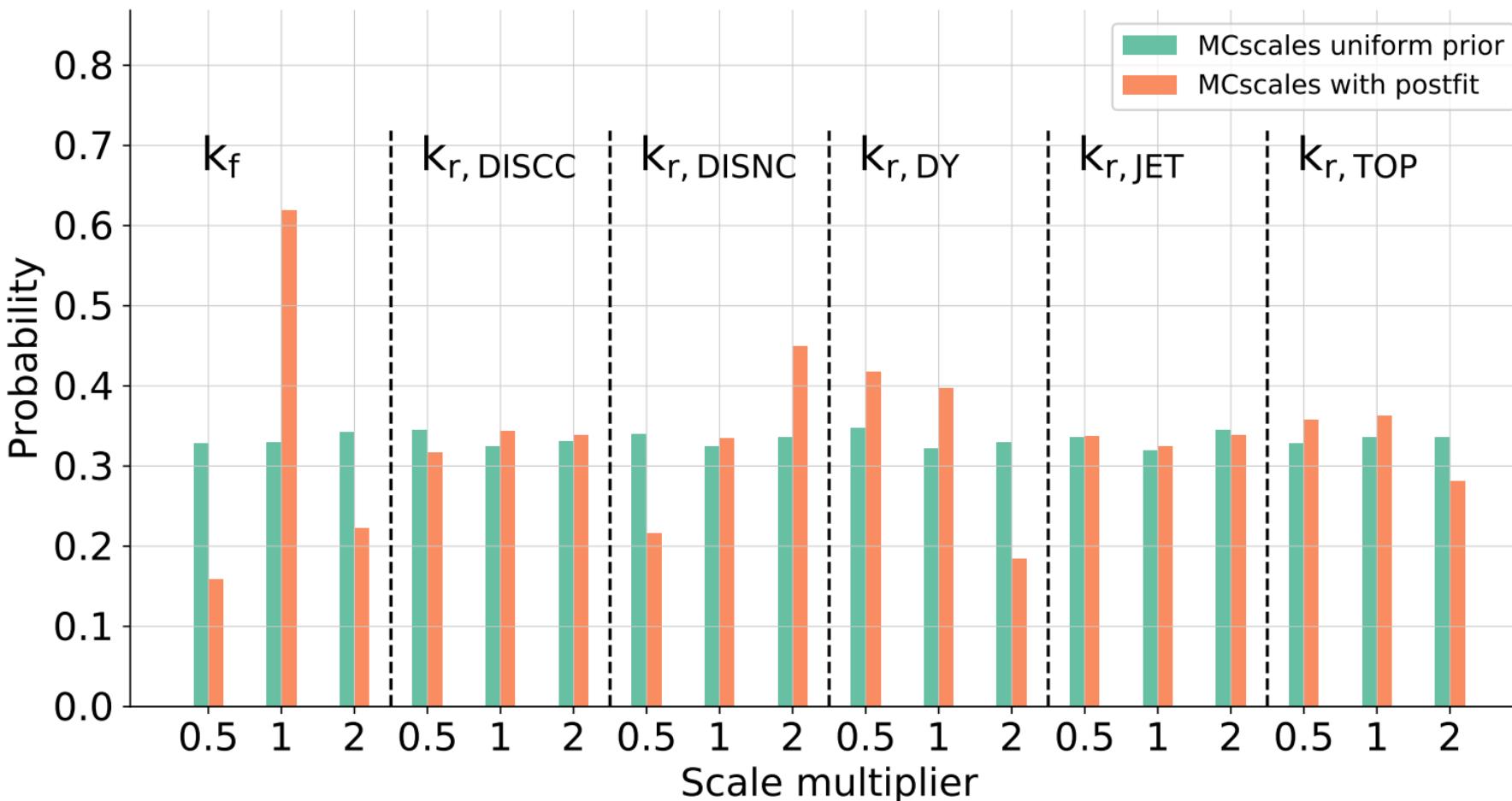
with $\omega \in \Omega = \{(\xi_f, \xi_1, \dots, \xi_{N_p}) \forall \xi_f, \xi_1, \dots, \xi_{N_p} \in \Xi\}$

3^{1+N_p} elements, with $N_p = 5$, p=DIS NC, DIS CC, DY, JET, TOP

Choose prior = choose $P(\omega)$
Posterior

$$\chi_n^2 > \langle \chi^2 \rangle_{n| \omega^{(n)}=\{1, \dots, 1\}} + 4 \text{std}(\chi^2)_{n| \omega^{(n)}=\{1, \dots, 1\}}$$

THE MCSCALES APPROACH



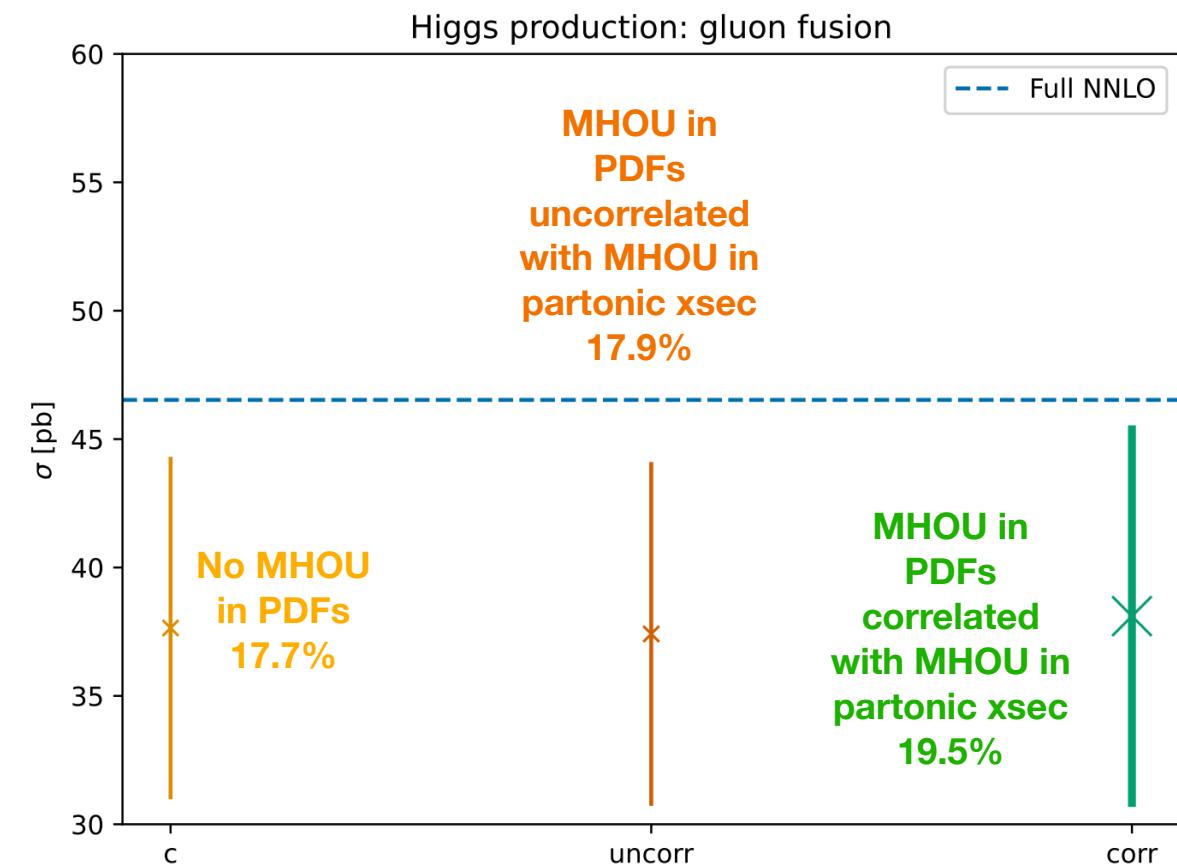
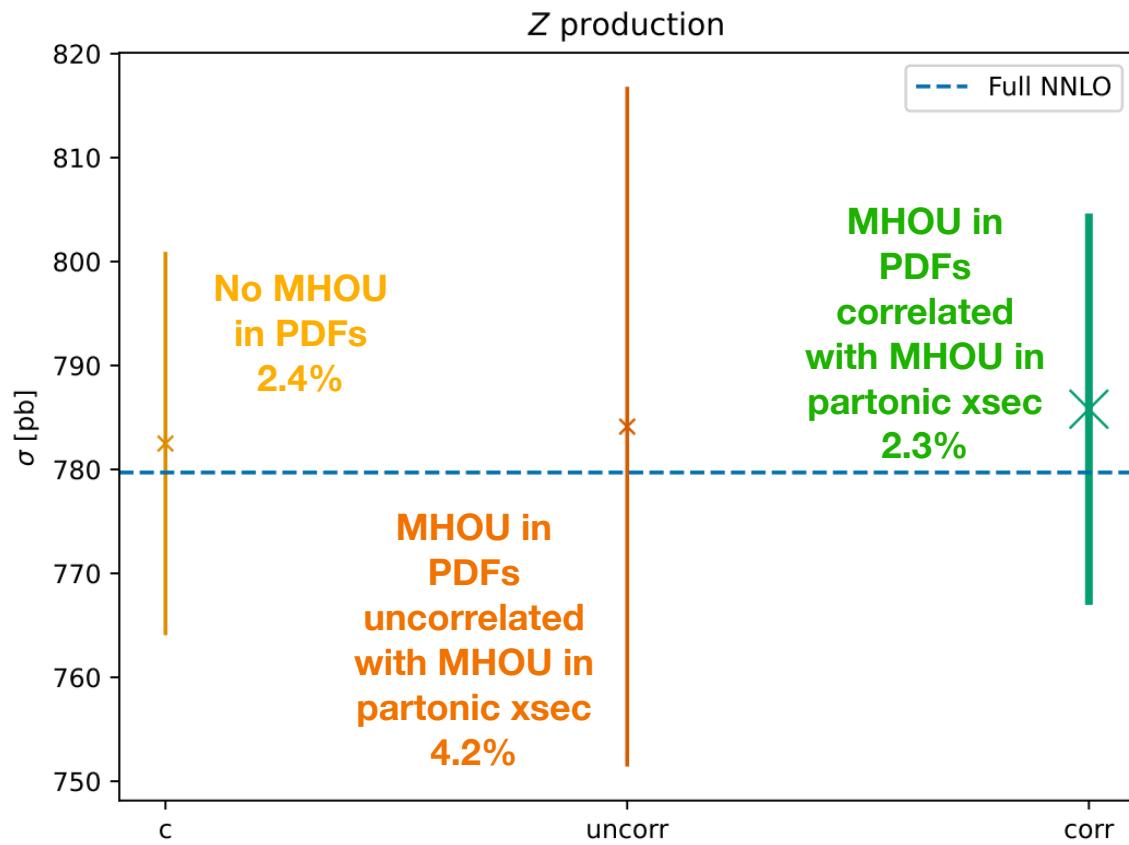
- ✓ Can look at the distribution of each of the scales over replicas.
- ✓ Flat distribution for the MCscales uniform prior.
- ✓ After applying postfit observe preference for central factorisation scale.
- ✓ Each process affected in a different way.

Scale multipliers	Process	Preferred values
(k_f, k_r)	DIS CC	$(1, 1)$
	DIS NC	$(1, 2)$
	DY	$(1, 1)$
	Jets	$(1, \frac{1}{2})$
	Top	$(1, 1)$

THE MCSCALES APPROACH

- ✓ Can compute full PDF+SCALE uncertainty in cross sections at NLO by matching the scales in the hard cross section computation with the scales in the MCscale PDF set: correlation fully taken into account

$$\left\{ \sigma_n = \hat{\sigma}_p(k_f^{(n)}, k_{r_p}^{(n)}) \otimes f_n(k_f^{(n)}, k_{r_p}^{(n)}) \quad \forall n = 1, \dots, N \right\}$$

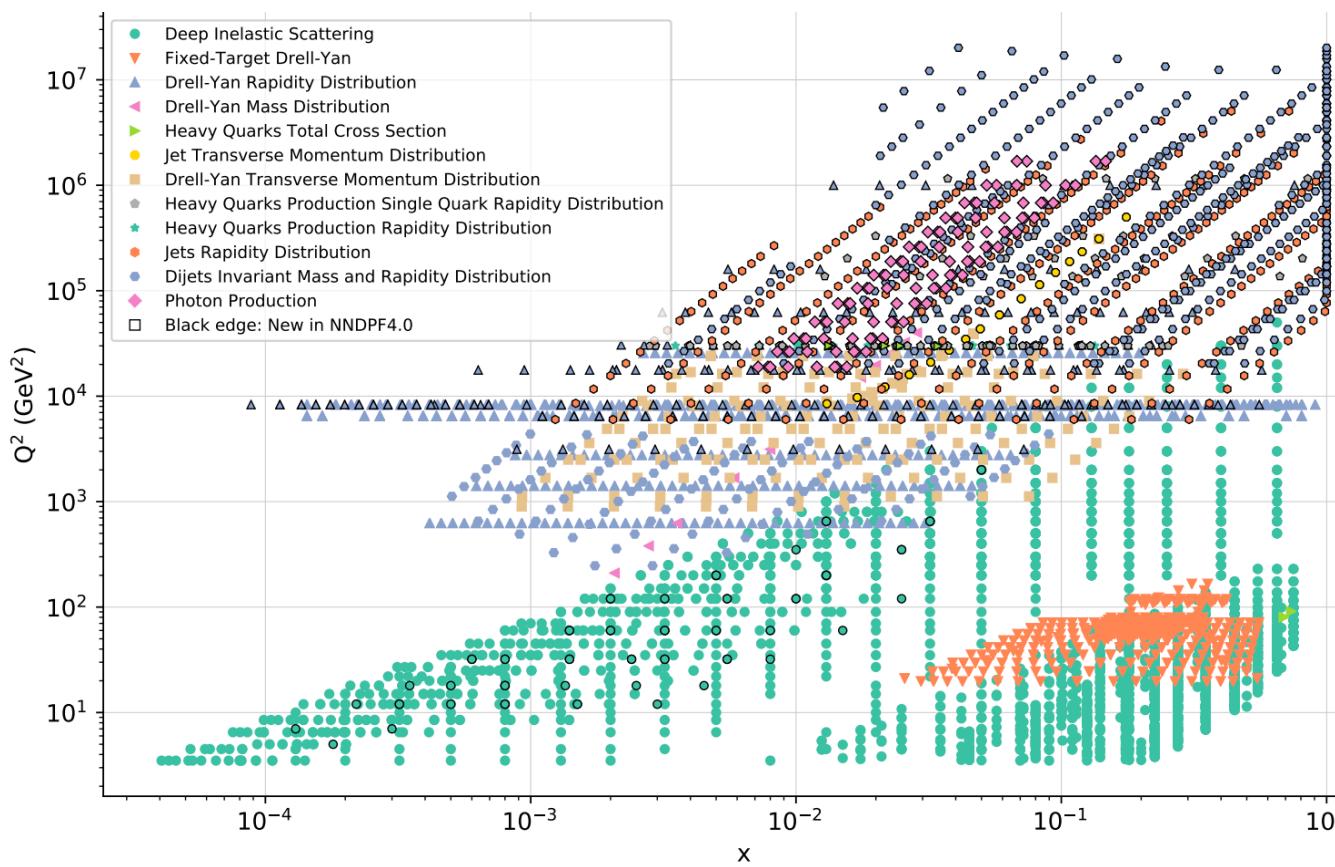


BEYOND THE STANDARD PROTON

EXTRACTING PHYSICS PARAMETERS FROM LHC DATA

- ✓ Abundance of precise LHC data allows to extract information on SM and BSM parameters

$$\chi^2 = \frac{1}{N_{\text{dat}}} \sum_{i=1}^{N_{\text{dat}}} (T_i(\{\theta\}, \{c\}) - D_i) \text{cov}_{ij}^{-1} (T_j(\{\theta\}, \{c\}) - D_j)$$



$$T_i(\{\theta\}, \{c\}) = \text{PDFs}(\{\theta\}, \{c\}) \otimes \hat{\sigma}_i(\{c\})$$

(B)SM parameters: $\alpha_s(M_z)$, M_w , θ_w , SMEFT WCs.....

Parameters determining PDFs at initial scale

- ✓ In a PDF fit typically

$$T_i(\{\theta\}) = \text{PDFs}(\{\theta\}, \{c = \bar{c}\}) \otimes \hat{\sigma}_i(\{c = \bar{c}\})$$

EXTRACTING PHYSICS PARAMETERS FROM LHC DATA

- ✓ Abundance of precise LHC data allows to extract information on SM and BSM parameters

$$\chi^2 = \frac{1}{N_{\text{dat}}} \sum_{i=1}^{N_{\text{dat}}} (T_i(\{\theta\}, \{c\}) - D_i) \text{cov}_{ij}^{-1} (T_j(\{\theta\}, \{c\}) - D_j)$$

$$T_i(\{\theta\}, \{c\}) = \text{PDFs}(\{\theta\}, \{c\}) \otimes \hat{\sigma}_i(\{c\})$$



 (B)SM parameters: $\alpha_s(M_z)$, M_w , θ_w , SMEFT WCs.....

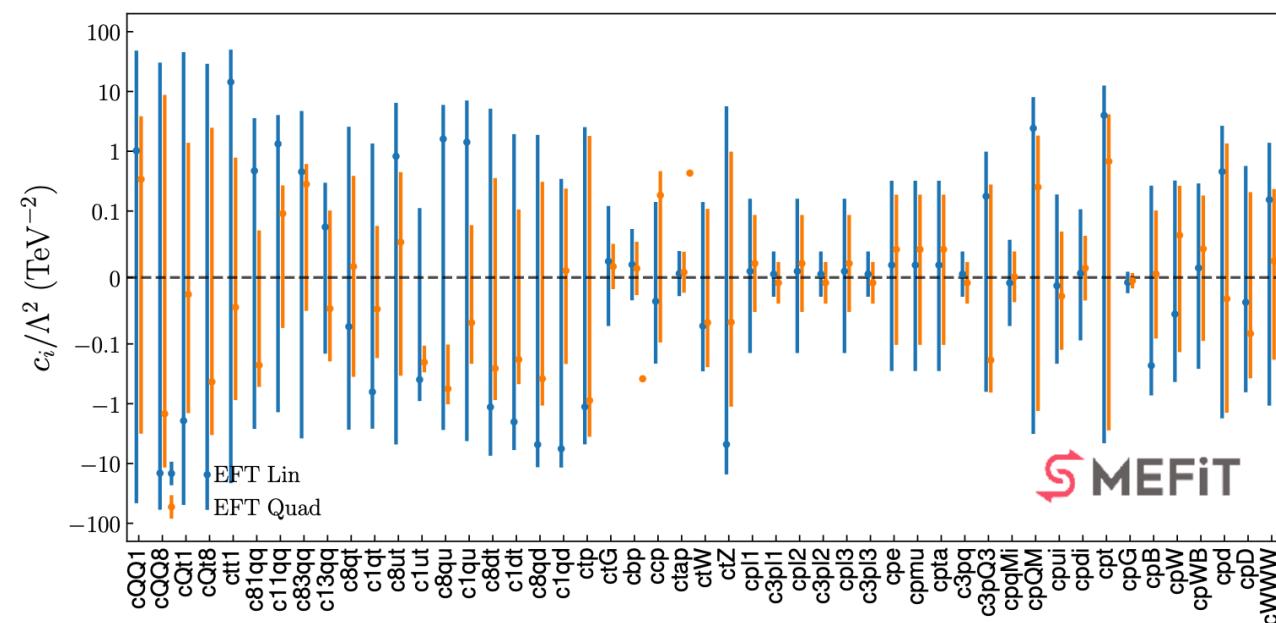
Parameters determining PDFs at initial scale

- ✓ In a PDF fit typically

$$T_i(\{\theta\}) = \text{PDFs}(\{\theta\}, \{c = \bar{c}\}) \otimes \hat{\sigma}_i(\{c = \bar{c}\})$$

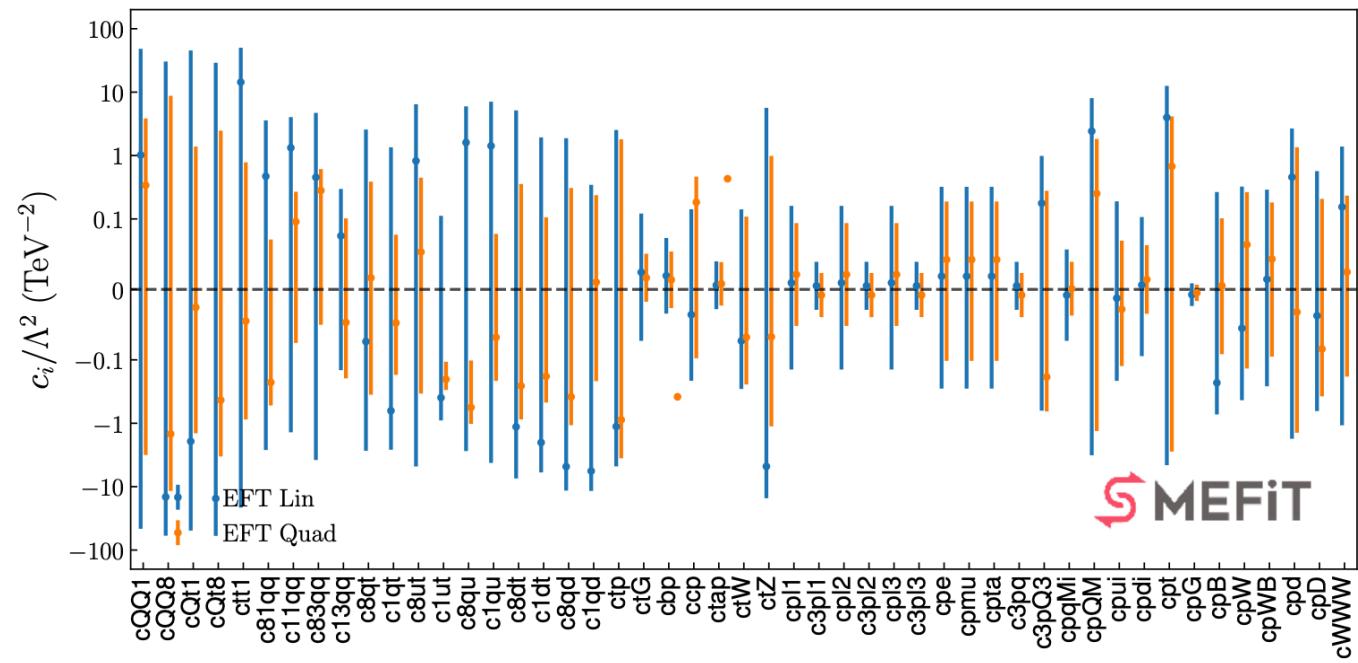
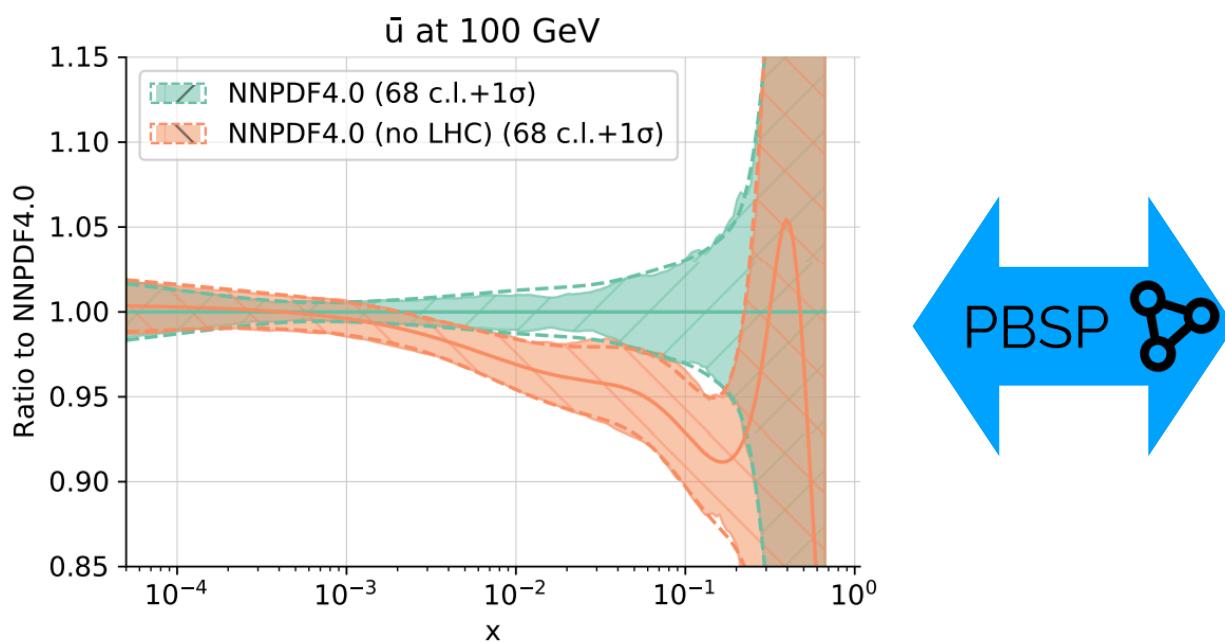
- ✓ In a fit of (B)SM parameters

$$T_i(\{c\}) = \text{PDFs}(\{\bar{\theta}\}, \{\bar{c}\}) \otimes \hat{\sigma}_i(\{c\})$$



INTERPLAY BETWEEN PDF FITS AND SMEFT FITS

- In principle low-scale physics is separable from high-scale physics, BUT the complexity of the LHC environment might well intertwine them.
 - PDFs are low-scale quantities extracted from experimental data at all scales, without considering any potential contamination due to new physics in high-energy data.
 - (SM)EFT fits are performed by assuming a priori that PDFs are SM-like.



INTERPLAY BETWEEN PDF FITS AND SMEFT FITS

- From the point of view of PDF fits:
 - How to make sure that new physics effects are not inadvertently fitted away in a PDF fit?
- From the point of view of SMEFT fits:
 - Should I make sure I am using a clean set of PDFs in a SMEFT analysis? How to define it? Is it enough?
 - How would the bounds change if I was consistently using PDFs that include in the fit the same operators that I am fitting?

T

$$d\sigma^{pp \rightarrow ab} = \sum_{i,j} [f_i \otimes f_j \otimes d\hat{\sigma}^{ij \rightarrow ab}] + \dots$$

↑

↑

$f(\{\theta_k\})$

$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{c_i}{\Lambda^2} \mathcal{O}_i^{(6)} + \dots$

Simultaneous fits
can shed light on
their interplay

$T(\{\theta_k\}, \{c_i\})$

SIMUNET: A DEEP-LEARNING BASED SIMULTANEOUS FIT

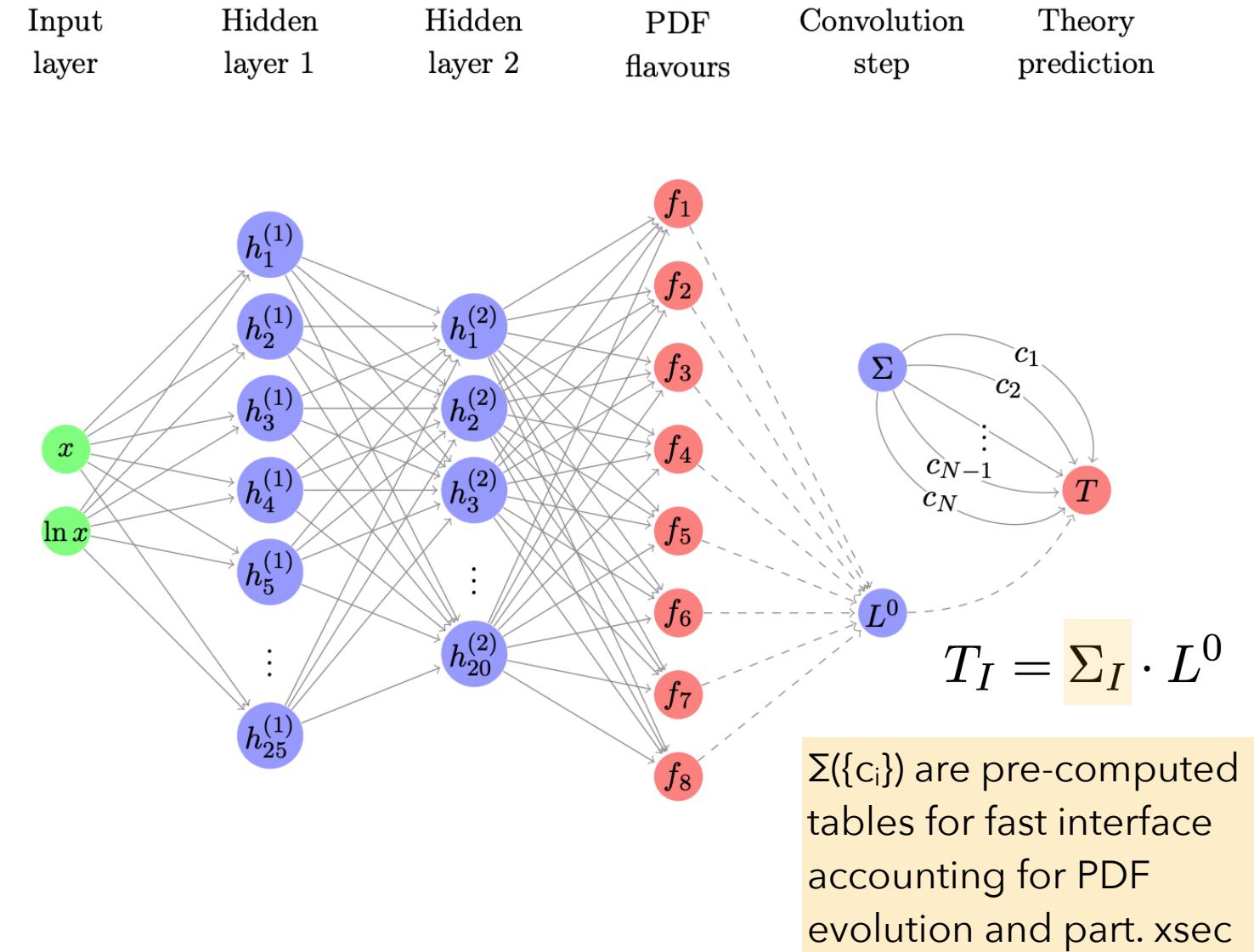
- The idea: take a PDF fit based on NNPDF4.0 methodology and make dependence of observables on physics parameters $\{c_i\}$ explicit via fast interface before computing the loss function (e.g. adding SMEFT corrections, expanding observables in terms of SM precision parameters)

- Perform minimisation of loss function over

$$\hat{\theta} = \theta \cup \{c_i\}$$

by adding new layer to the deep neural network used in NNPDF4.0

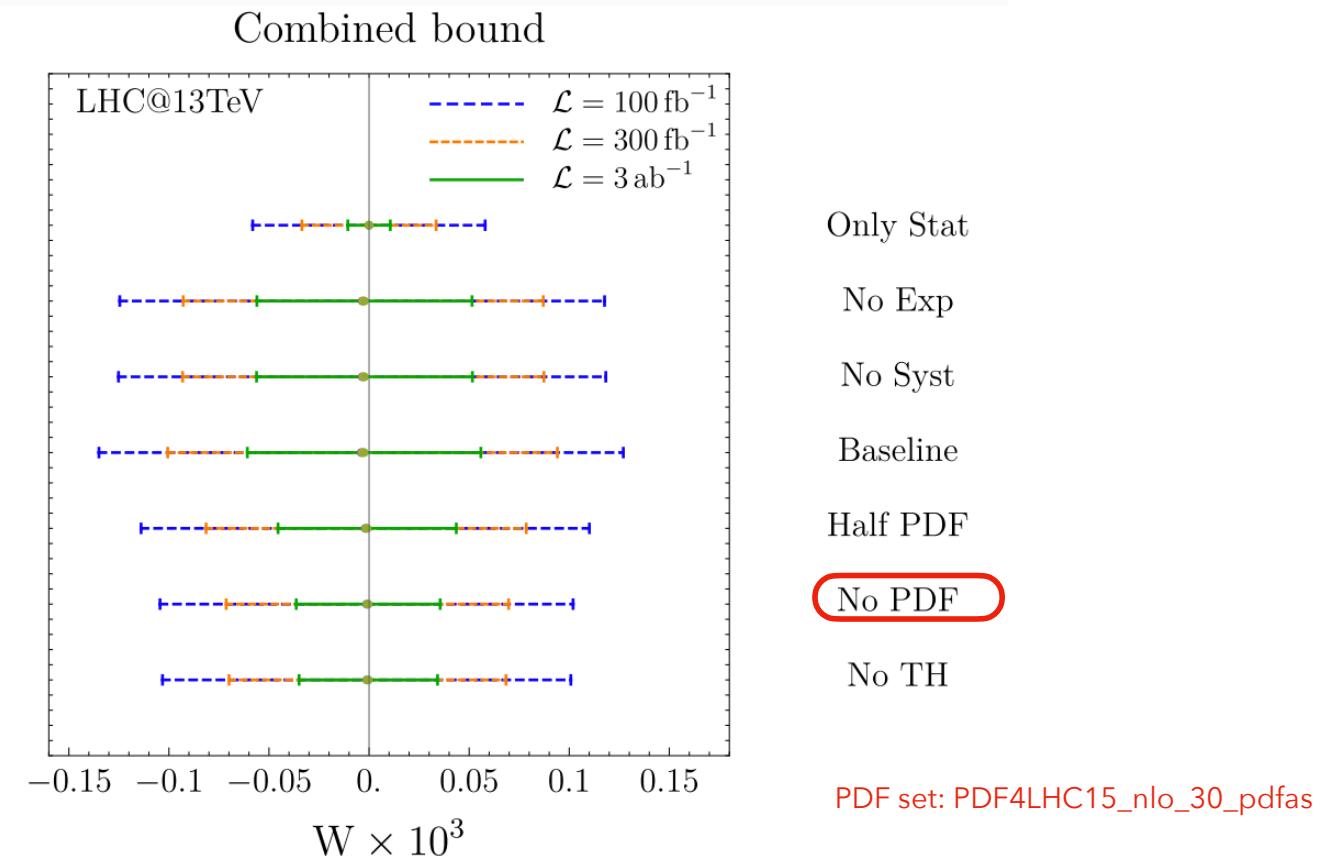
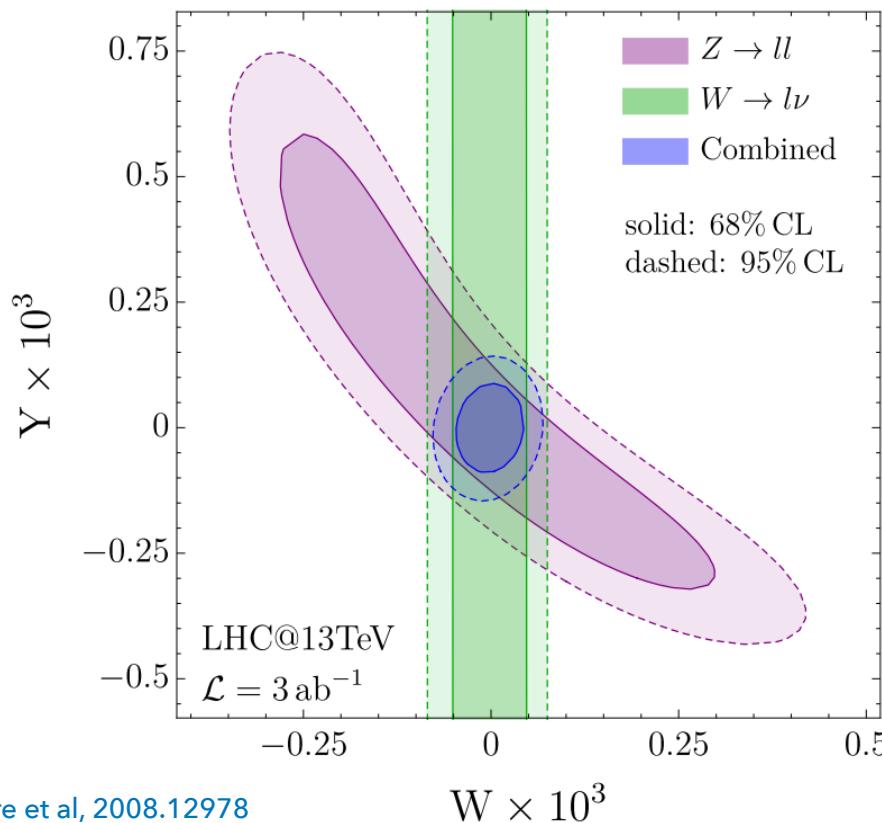
- Can expand dependence on c_i beyond linear terms in T (up to generic power in polynomial expansion) by adding non-trainable edges
- Can be done both for SM parameters and SMEFT coefficients.



APPLICATION TO SIMULTANEOUS DETERMINATION OF PDFS AND SMEFT WCs

- Case study at higher energy: EW oblique corrections in high-mass NC and CC Drell-Yan tails.
- W and Y parametrise the self-energy of gauge bosons and are powerful probes of quark-lepton contact interactions that produce effects that grow with energy [Torre et al, 2008.12978]
- Interplay first studied with scan method over W - Y parameters in [Greljo et al, 2104.02723]

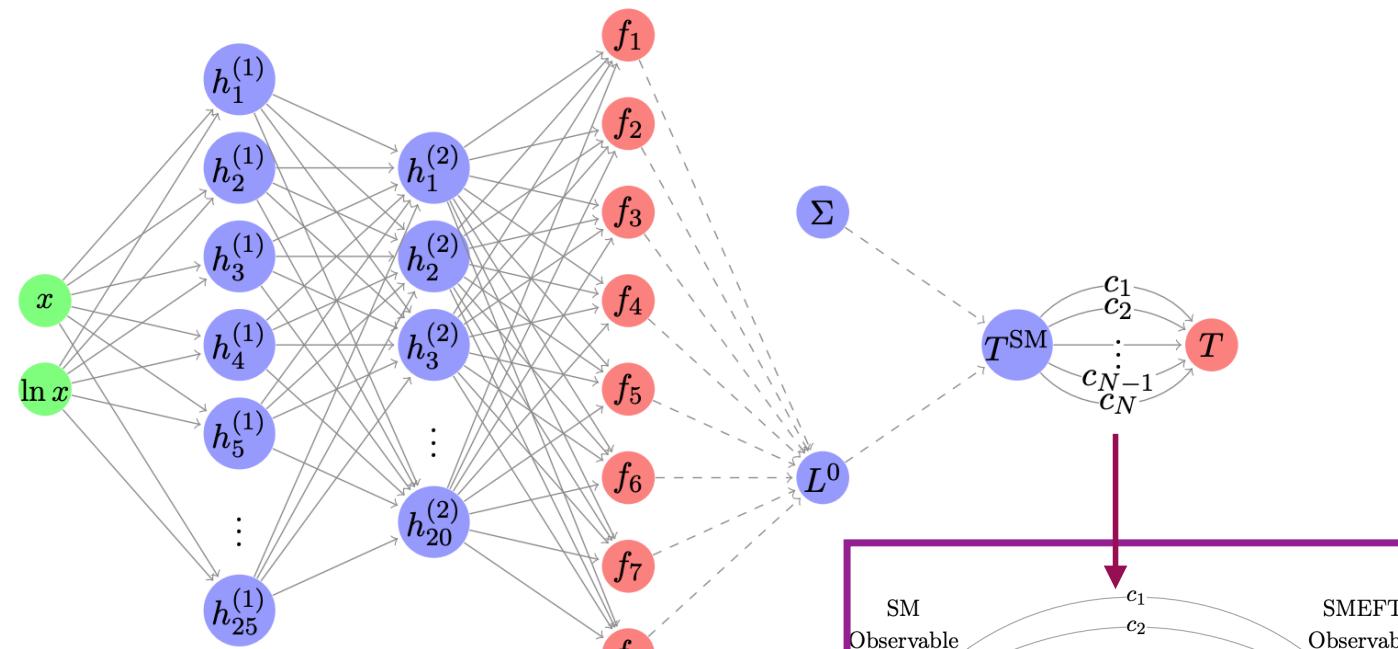
$$\mathcal{L}_{\text{SMEFT}} \supset -\frac{\hat{W}}{4m_W^2}(D_\rho W_{\mu\nu}^a)^2 - \frac{\hat{Y}}{4m_W^2}(\partial_\rho B_{\mu\nu})^2$$



THE SIMUNET ANALYSIS

- SimuNET yields a truly simultaneous fit, rather than a scan in benchmark point in W-Y parameter space and it does not have limit in number of parameters that can be fitted alongside PDFs at the initial scale!

Input layer	Hidden layer 1	Hidden layer 2	PDF flavours	Convolution step	SM Observable	SMEFT Observable
-------------	----------------	----------------	--------------	------------------	---------------	------------------



Linear dim-6 operator

$$T(\hat{\theta}) = \Sigma(\{c_n\}) \cdot L^0(\theta) = T^{\text{SM}}(\theta) \cdot \left(1 + \sum_{n=1}^N c_n R_{\text{SMEFT}}^{(n)} \right)$$

$$T^{\text{SM}}(\theta) = \Sigma^{\text{SM}} \cdot L^0(\theta)$$

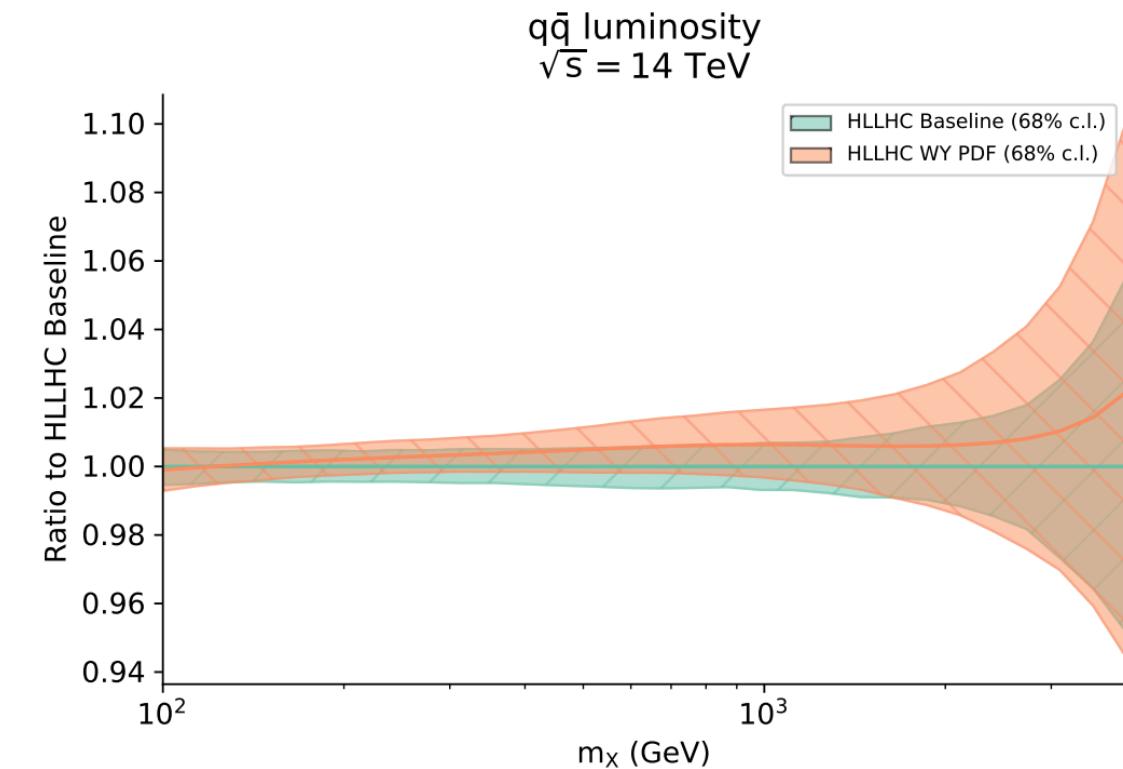
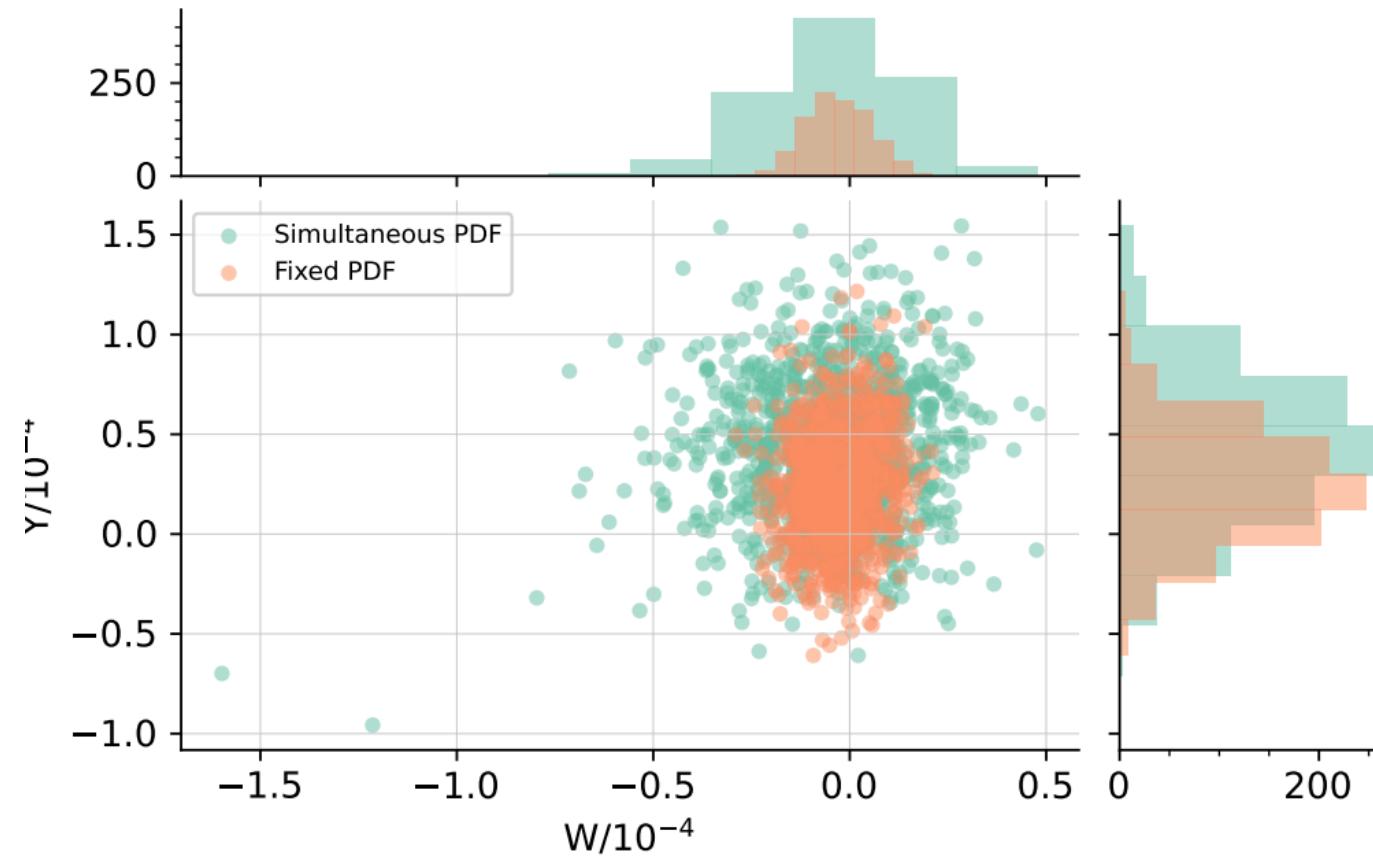
Quadratic dim-6 operator

$$T(\hat{\theta}) = T^{\text{SM}}(\theta) \cdot \left(1 + \sum_{n=1}^N c_n R_{\text{SMEFT}}^{(n)} + \sum_{1 \leq n \leq m \leq N} c_{nm} R_{\text{SMEFT}}^{(n,m)} \right)$$

$$c_n c_m$$

RESULTS: DRELL-YAN DATA @HL-LHC

S. Iranipour, MU - arXiv: 2201.07240



- ✓ Simultaneous analysis of PDFs and W&Y SMEFT coefficient of DIS + DY (including HL-LHC projections) using simuNET method shows that at HL-LHC the effect of interplay becomes important as WCs bounds broaden and PDFs change significantly once SMEFT effects allowed in theory predictions entering PDF fit
- ✓ Stress-tested and shown robustness with closure tests

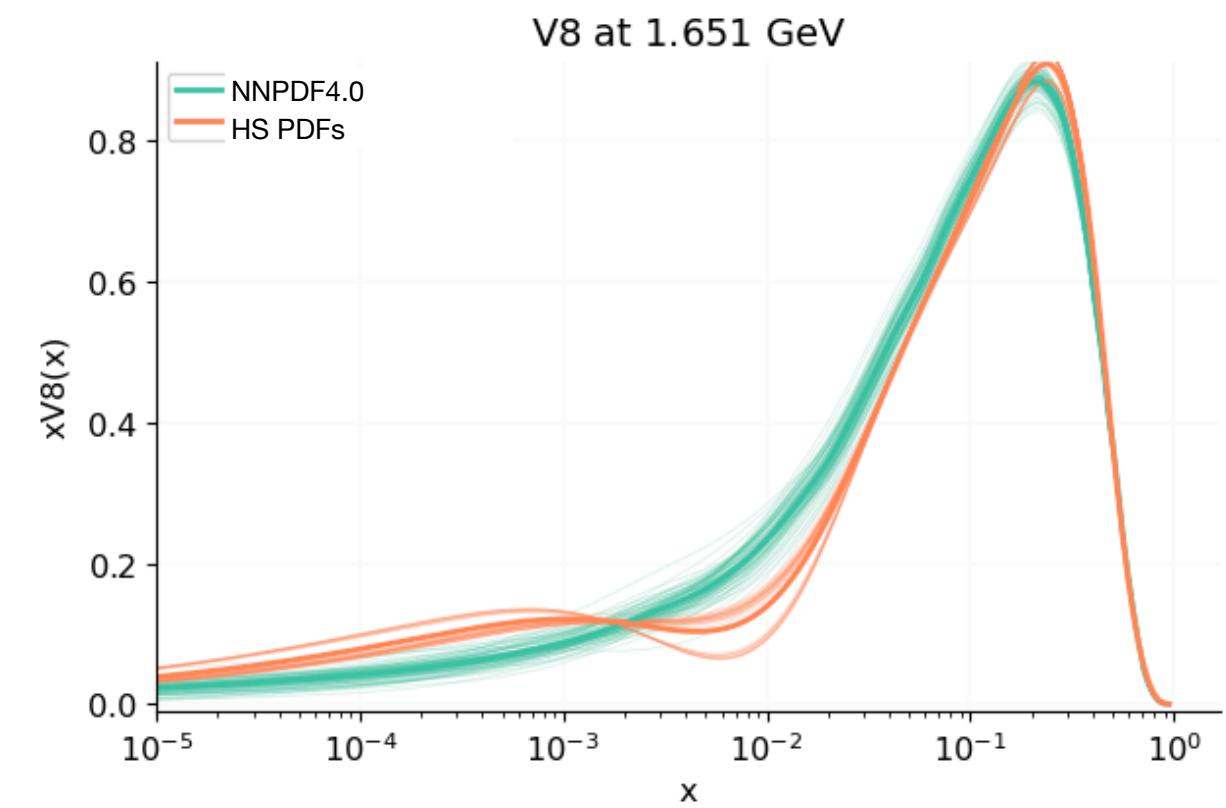
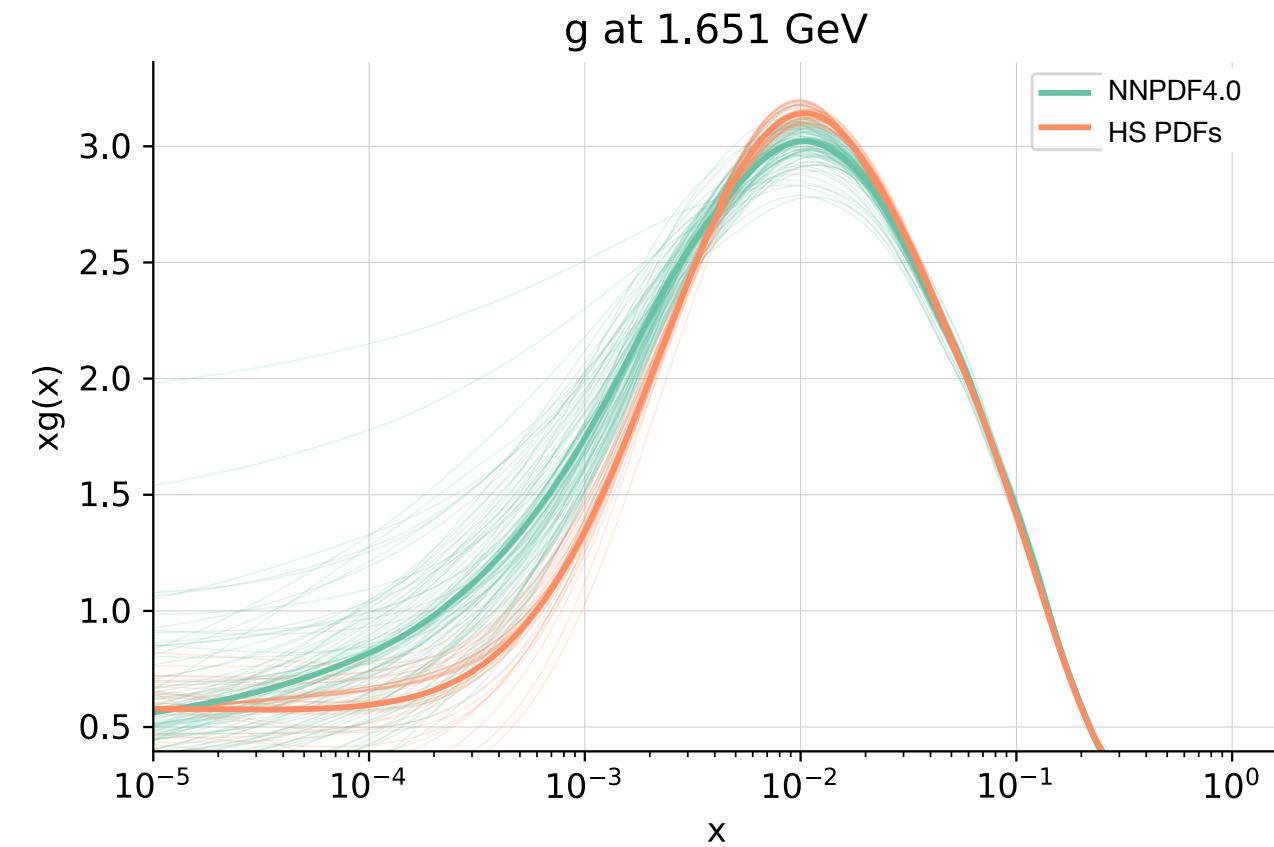
CONCLUSIONS AND OUTLOOK

- In an era of precision, need careful assessment of PDF uncertainties and rigorous claims about faithfulness of PDF uncertainties.
- PDF uncertainties in standard fit underestimated due to lack of inclusion of MHOU, several methods developed to include it in PDF uncertainties under scrutiny, and progress towards approx N3LO PDFs.
- While huge progress made in determining key ingredients of theoretical predictions from the data, PDFs, α_s , SMEFT WCs coefficients, not yet evident how to combine all these partial fits into a global interpretation of the LHC data
- SimuNET methodology based on an extension of the NNPDF4.0 NN architecture, allows the addition of an extra layer to simultaneously determine PDFs alongside an arbitrary number of physics parameters that enter predictions.
- No time to mention other interesting interplay between PDFs and new physics due to the presence of new partons (dark photon, see [M. McCullough, J. Moore, MU, arXiv:2203.12628](#))
- Lots of new exciting avenues being explored:
 - Determination of PDFs and α_s ([Stegeman et al](#))
 - Determination of PDFs and SMEFT coefficients in the top sector ([Kassabov, Madigan, Mantani, Moore, Morales, Rojo, MU](#))
 - Systematic study of new physics contamination in PDF fits
 - Determination of PDFs and electroweak parameters

EXTRA MATERIAL

THE HOPSCOTCH STUDY: ANSWERS

Pick HS PDFs that are outside 95% C.L. NNPDF4.0 band



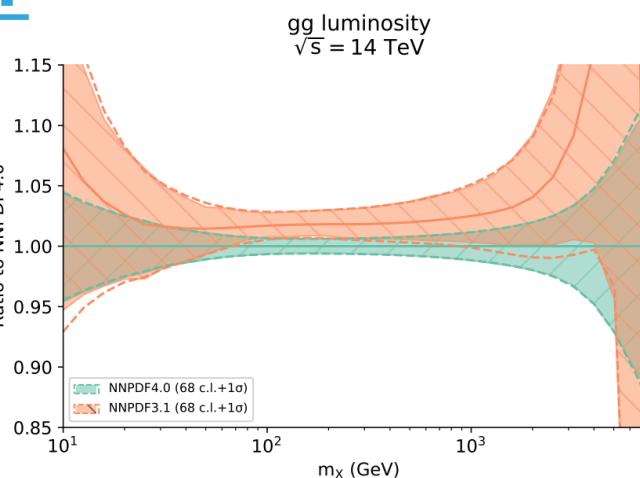
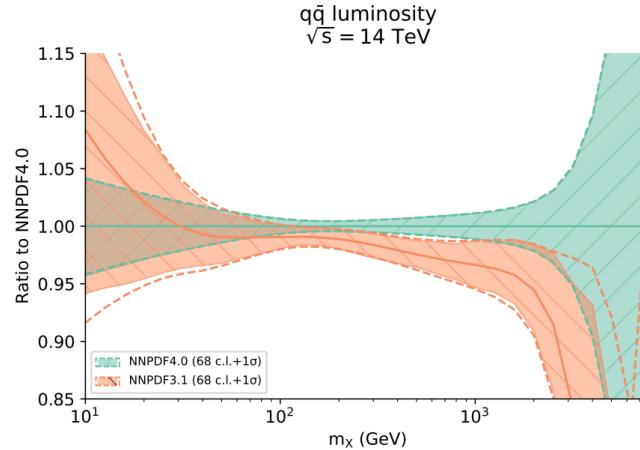
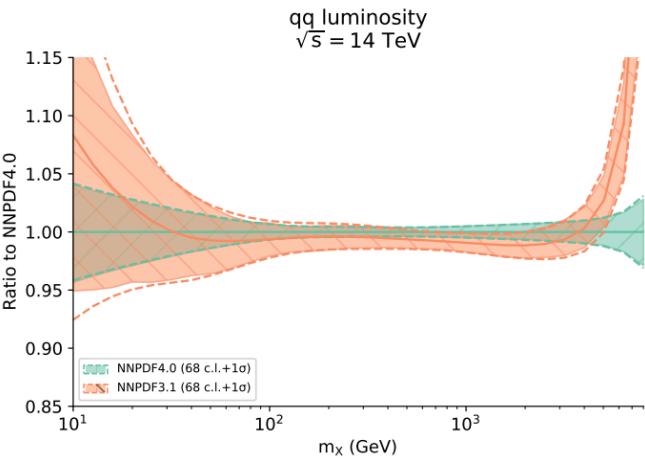
NNPDF4.0 & LHC DATA

Data set	NNPDF4.0	NNPDF3.1	ABMP16	CT18	MSHT20
ATLAS W, Z 7 TeV (2010)	✓	✓	✓	✓	✓
ATLAS W, Z 7 TeV (2011)	✓	✓	✗	✓	✓
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 8 TeV	✓	✓	✗	✓	✓
ATLAS σ_{tt}^{tot} 7, 8 TeV	✓	✓	✓	✗	✗
ATLAS σ_{tt}^{tot} 13 TeV	✓	✓	✓	✗	✗
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 13 TeV	✓	✗	✗	✗	✗
ATLAS single top R_t 7, 8, 13 TeV	✓	✗	✓	✗	✗
ATLAS single top diff. 7, 8 TeV	✓	✗	✗	✗	✗
ATLAS single top diff. 8 TeV	✓	✗	✗	✗	✗

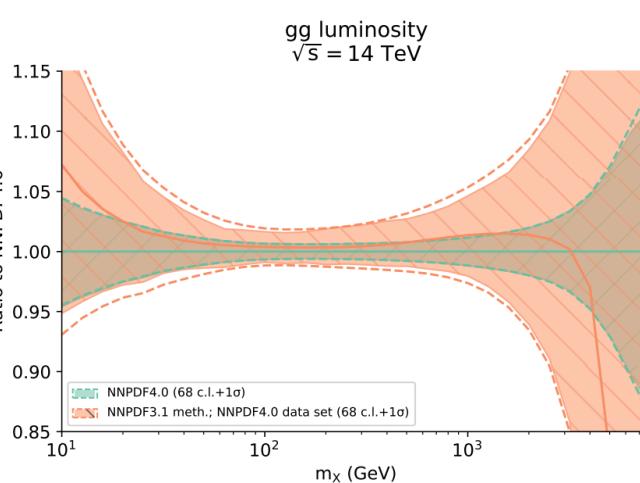
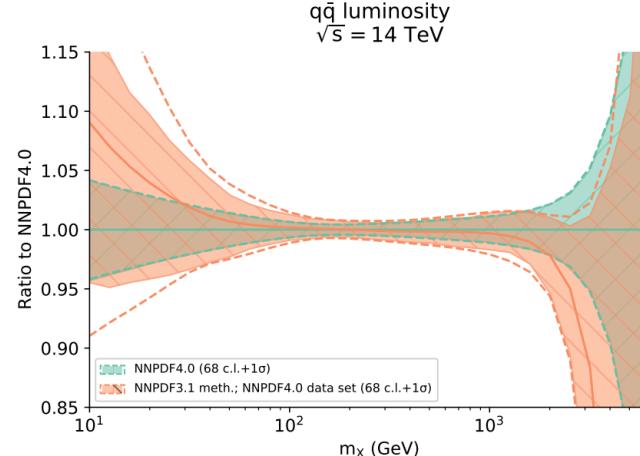
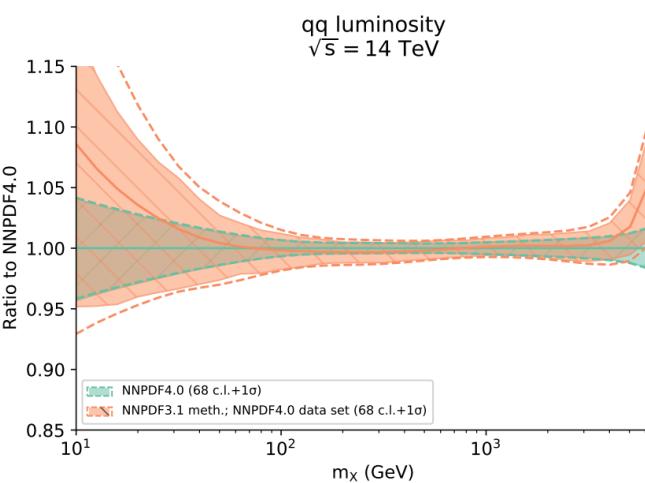
Data set	NNPDF4.0	NNPDF3.1	ABMP16	CT18	MSHT20
LHCb Z 940 pb	✓	✓	✗	✗	✓
LHCb $Z \rightarrow ee$ 2 fb	✓	✓	✓	✓	✓
LHCb $W, Z \rightarrow \mu$ 7 TeV	✓	✓	✓	✓	✓
LHCb $W, Z \rightarrow \mu$ 8 TeV	✓	✓	✓	✓	✓
LHCb $Z \rightarrow \mu\mu, ee$ 13 TeV	✓	✗	✗	✗	✗

Data set	NNPDF4.0	NNPDF3.1	ABMP16	CT18	MSHT20
CMS W electron asymmetry 7 TeV	✓	✓	✗	✓	✓
CMS W muon asymmetry 7 TeV	✓	✓	✓	✓	✗
CMS Drell-Yan 2D 7 TeV	✓	✓	✗	✗	✓
CMS W rapidity 8 TeV	✓	✓	✓	✓	✓
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 σ_{tt}^{tot} 5 TeV	✓	✗	✓	✗	✗
CMS σ_{tt}^{tot} 7, 8 TeV	✓	✓	✓	✓	✓
CMS σ_{tt}^{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	✓	✗	✓	✗	✗

NNPDF4.0: THE ROLE OF METHODOLOGY



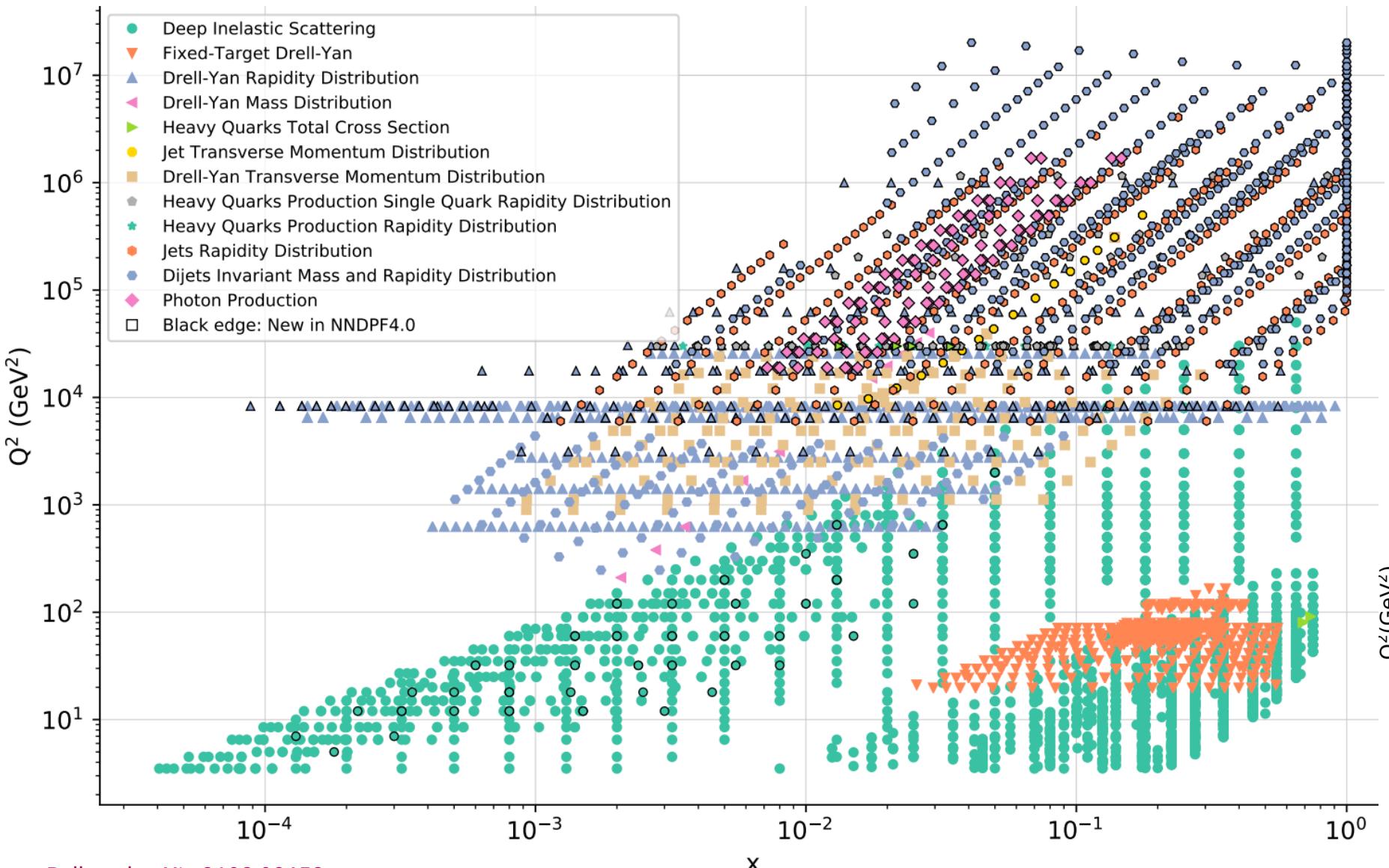
NNPDF3.1 vs NNPDF4.0	
METHODOLOGY	DATA
NNPDF3.1 (4093)	1.19
NNPDF4.0 (4491)	1.12



NNPDF3.1 vs NNPDF4.0	
METHODOLOGY	DATA
NNPDF3.1 (4093)	1.19
NNPDF4.0 (4491)	1.12

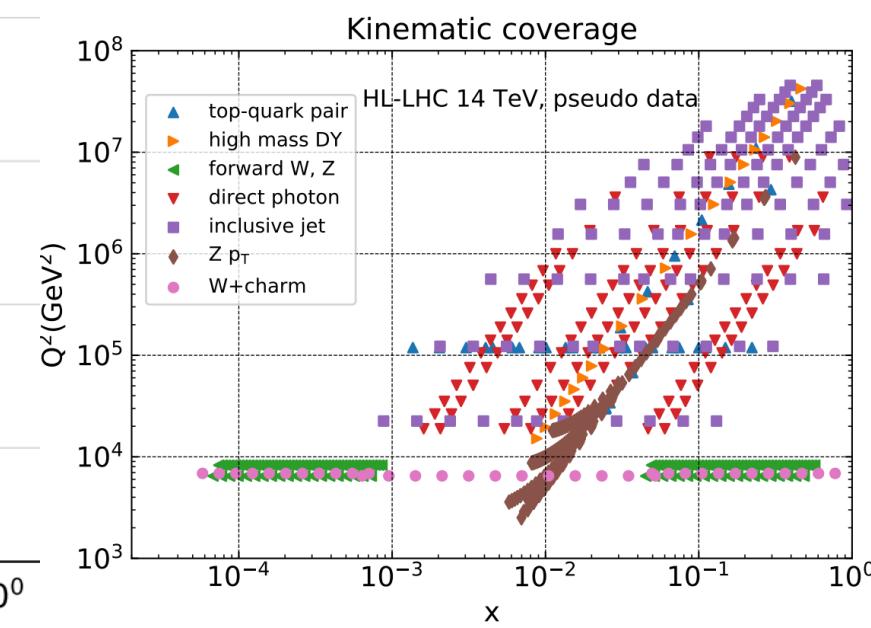
- Shift in parton luminosities mostly due to inclusion of $O(500)$ more data points
- Parton luminosities based on same dataset are consistent with each other but 4.0 methodology displays smaller uncertainty than 3.1 methodology: NNPDF4.0 more accurate and superior to previous methodology

PDFS AND NEW PHYSICS



Ball et al, arXiv:2109.02653

- Top pair production and single top data included in SMEFT analysis [Hartland et al 1901.05965] [Ellis et al 2012.02779]
- Dijets data in [Bordone et al 2103.10332] [Alioli et al 1706.03068]
- Drell-Yan data in [Farina et al 1609.08157] [Torre et al 2008.12978]
- Inclusive jets in [Alte et al 1711.07484]
- Overlap enhanced in HL-LHC projections [Abdul Khalek et al, 1810.03639]



Abdul Khalek et al, arXiv:1810.03639

ANALYSIS METHODOLOGY

- We performed a similar analysis as in Torre et al, now with emphasis on PDF and their interplay with bounds on oblique operators
[Greljo, Iranipour, Kassabov, Madigan, Moore, Rojo, MU, Voisey: 2104.02723]

$$\chi^2 = \frac{1}{n_{\text{dat}}} \sum_{i,j=1}^{n_{\text{dat}}} (D_i - T_i) (\text{cov}^{-1})_{ij} (D_j - T_j)$$

1. Take data, make theoretical predictions accounting for operator in partonic cross section **with fixed SM PDFs**.
2. Compute chi2 as a function of WCs (Wilson Coefficients)
3. Minimise chi2 and find best-fit and C.L.s of WCs
4. Extract bounds

$$T = f_{1,\text{SM}} \otimes f_{2,\text{SM}} \otimes \hat{\sigma}_{\text{BSM}}$$

SM PDFs

1. Take data, make theoretical predictions accounting for operator **in partonic cross section and PDFs**.
2. Compute chi2 as a function of WCs (Wilson Coefficients)
3. Minimise chi2 and find best-fit and C.L.s of WCs
4. Extract bounds

$$T = f_{1,\text{BSM}} \otimes f_{2,\text{BSM}} \otimes \hat{\sigma}_{\text{BSM}}$$

SMEFT PDFs / Simultaneous fit

ANALYSIS METHODOLOGY

- We performed a similar analysis as in Torre et al, now with emphasis on PDF and their interplay with bounds on oblique operators
[Greljo, Iranipour, Kassabov, Madigan, Moore, Rojo, MU, Voisey: 2104.02723]
- Methodology for simultaneous fit is similar to the one adopted in fits of α_s from a global fit of PDFs

$$\chi^2 = \frac{1}{n_{\text{dat}}} \sum_{i,j=1}^{n_{\text{dat}}} (D_i - T_i)(\text{cov}^{-1})_{ij}(D_j - T_j)$$

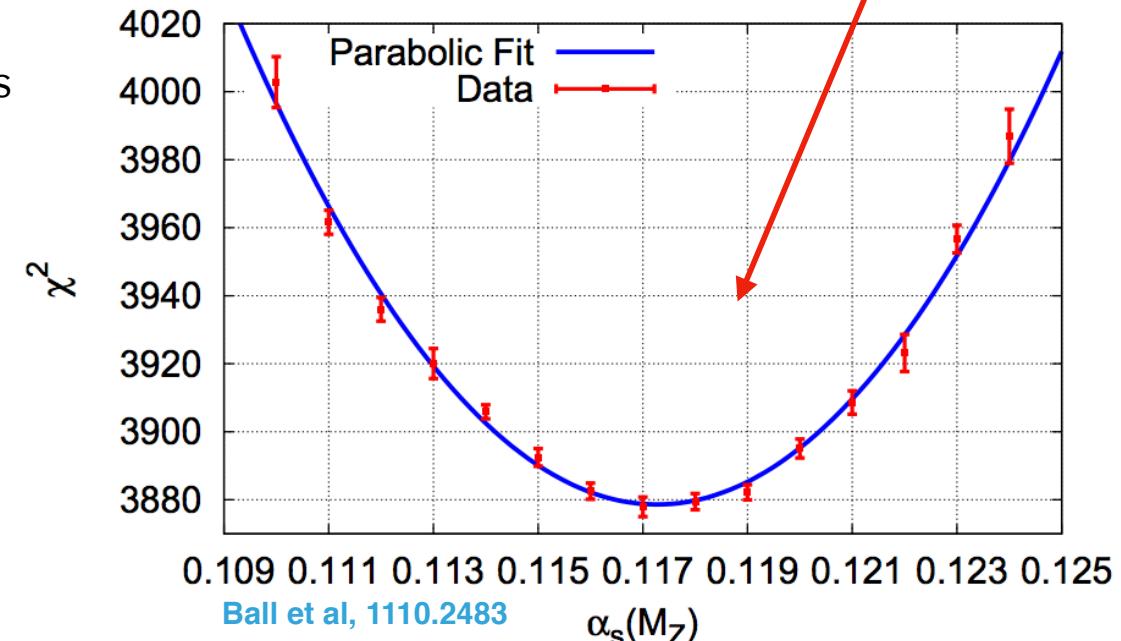
- Take data, make theoretical predictions accounting for operator in partonic cross section **with fixed SM PDFs**.
- Compute chi2 as a function of WCs (Wilson Coefficients)
- Minimise chi2 and find best-fit and C.L.s of WCs
- Extract bounds

$$T = f_{1,\text{SM}} \otimes f_{2,\text{SM}} \otimes \hat{\sigma}_{\text{BSM}}$$

SM PDFs

$$T = f_1(\alpha_s) \otimes f_2(\alpha_s) \otimes \hat{\sigma}(\alpha_s)$$

NNPDF2.1 NNLO Global



Ball et al, 1110.2483

$\alpha_s(M_Z)$

- Take data, make theoretical predictions accounting for operator **in partonic cross section and PDFs**.
- Compute chi2 as a function of WCs (Wilson Coefficients)
- Minimise chi2 and find best-fit and C.L.s of WCs
- Extract bounds

$$T = f_{1,\text{BSM}} \otimes f_{2,\text{BSM}} \otimes \hat{\sigma}_{\text{BSM}}$$

SMEFT PDFs / Simultaneous fit

ANALYSIS METHODOLOGY

- We performed a similar analysis as in Torre et al, now with emphasis on PDF and their interplay with bounds on oblique operators
[\[Greljo, Iranipour, Kassabov, Madigan, Moore, Rojo, MU, Voisey: 2104.02723\]](#)
- Methodology for simultaneous fit is similar to the one adopted in fits of α_S from a global fit of PDFs

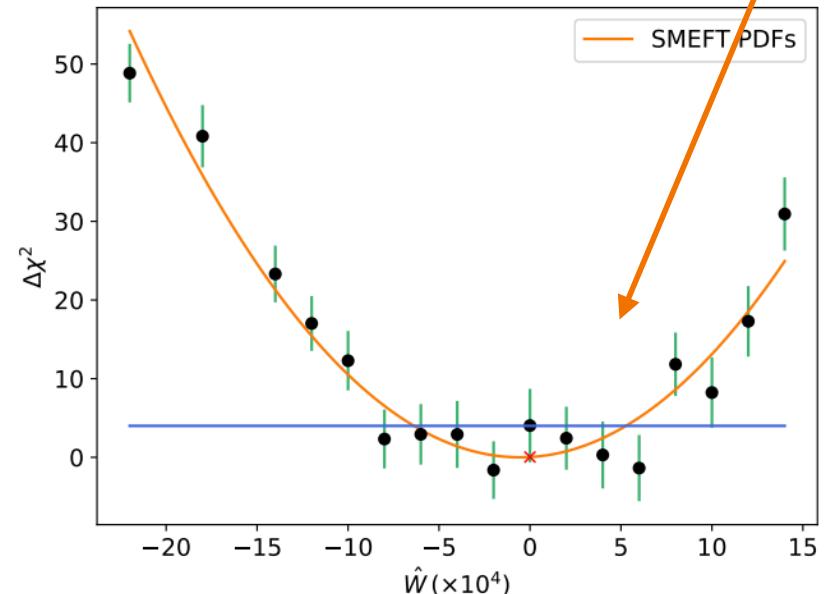
$$\chi^2 = \frac{1}{n_{\text{dat}}} \sum_{i,j=1}^{n_{\text{dat}}} (D_i - T_i)(\text{cov}^{-1})_{ij}(D_j - T_j)$$

- Take data, make theoretical predictions accounting for operator in partonic cross section with fixed SM PDFs.
- Compute chi2 as a function of WCs (Wilson Coefficients)
- Minimise chi2 and find best-fit and C.L.s of WCs
- Extract bounds

$$T = f_{1,\text{SM}} \otimes f_{2,\text{SM}} \otimes \hat{\sigma}_{\text{BSM}}$$

SM PDFs

$$T = f_1(\hat{W}) \otimes f_2(\hat{W}) \otimes \hat{\sigma}(\hat{W})$$



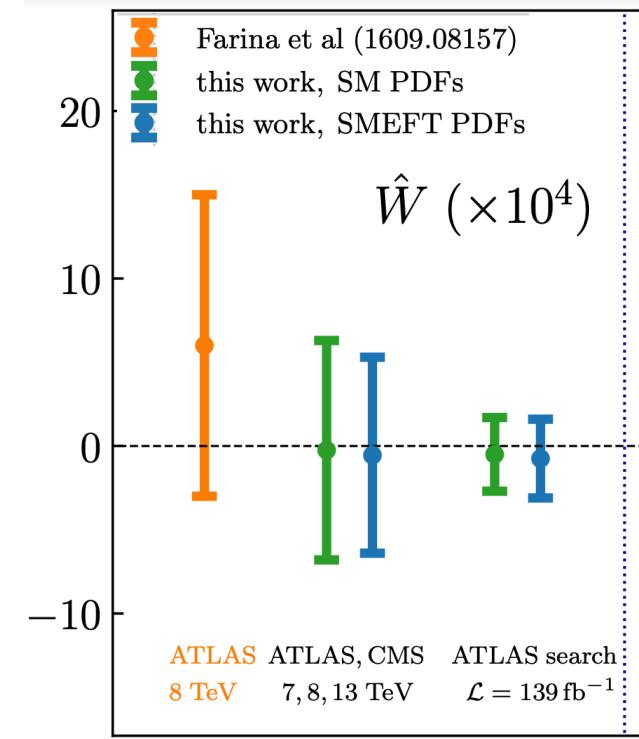
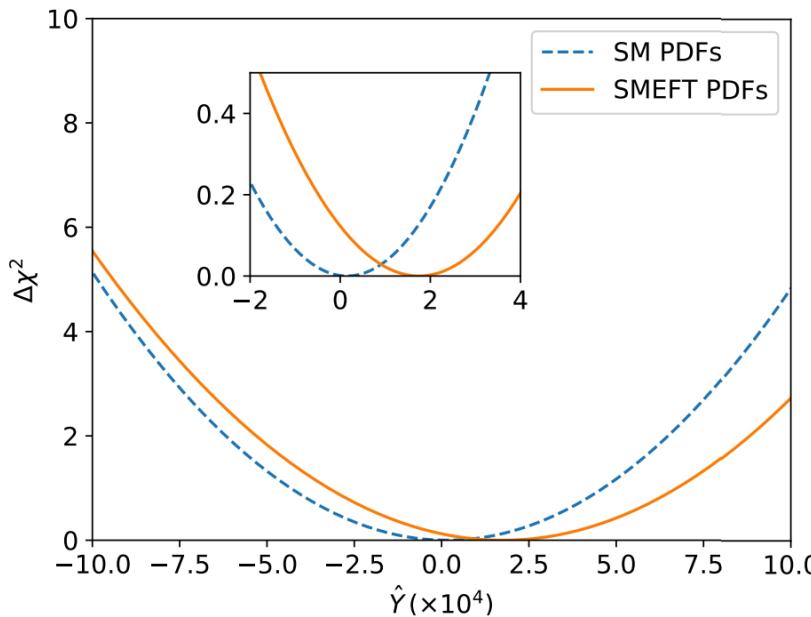
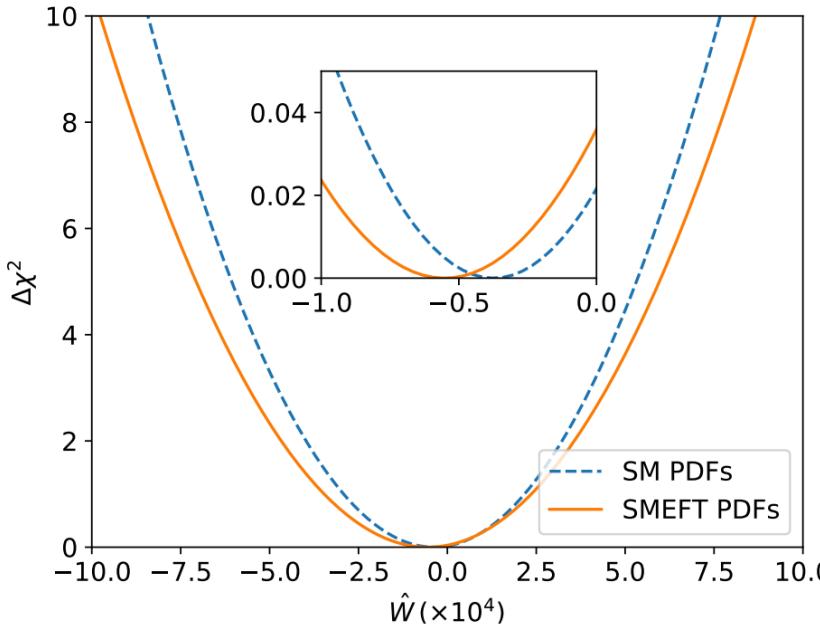
Greljo et al, 2104.02723

- Take data, make theoretical predictions accounting for operator in partonic cross section and PDFs.
- Compute chi2 as a function of WCs (Wilson Coefficients)
- Minimise chi2 and find best-fit and C.L.s of WCs
- Extract bounds

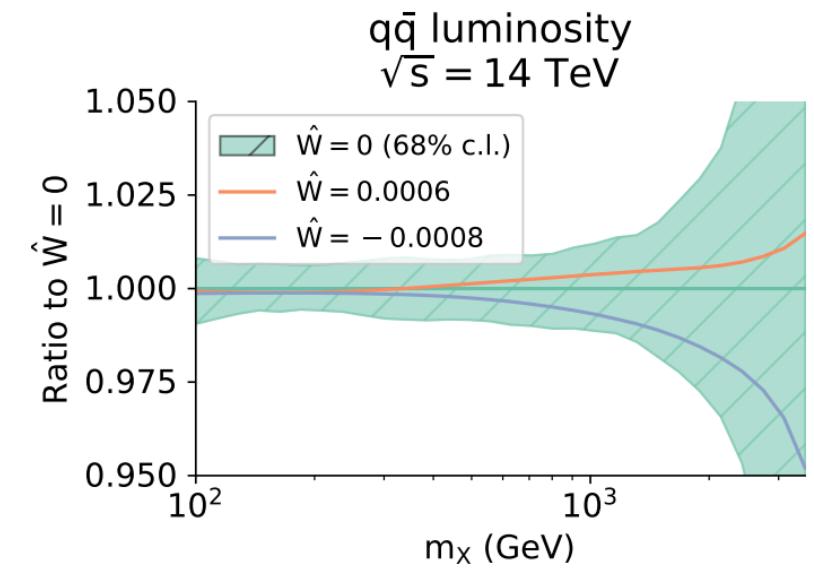
$$T = f_{1,\text{BSM}} \otimes f_{2,\text{BSM}} \otimes \hat{\sigma}_{\text{BSM}}$$

SMEFT PDFs / Simultaneous fit

INTERPLAY @ RUN I AND RUN II



- With current data, PDFs are moderately affected by inclusion of non-zero W and Y coefficients in the fit, mostly quark-antiquark luminosity within uncertainties
- Broadening of individual bounds on W and Y once SMEFT PDFs are used (i.e. PDFs that have been fitted with consistent values of W and Y) is not negligible, but still within PDF uncertainties
- If SMEFT PDFs are used in determining bounds from ATLAS search same mild broadening (larger than PDF uncertainties)

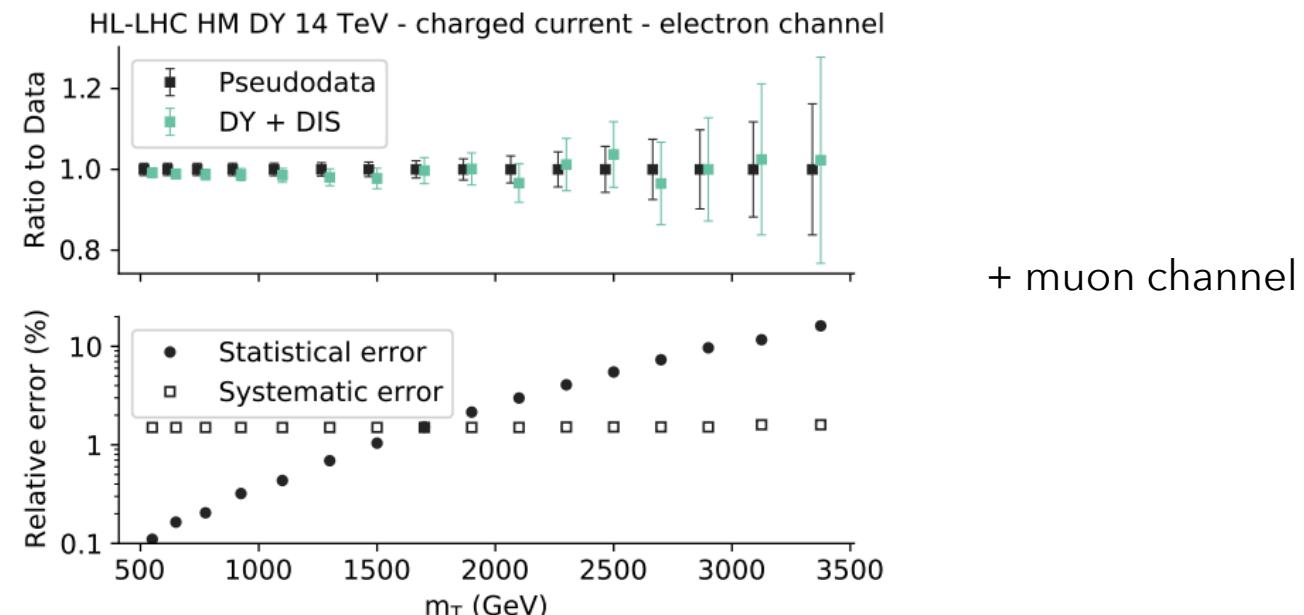
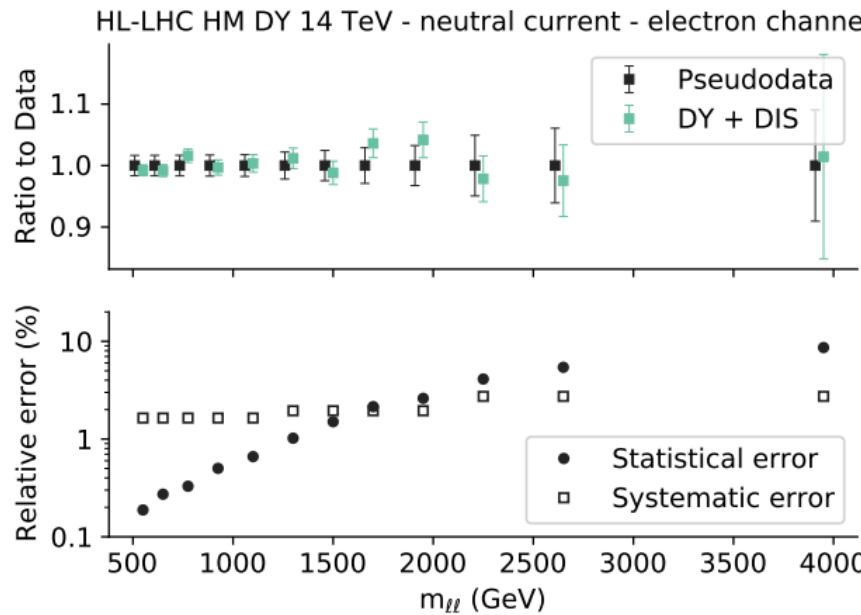
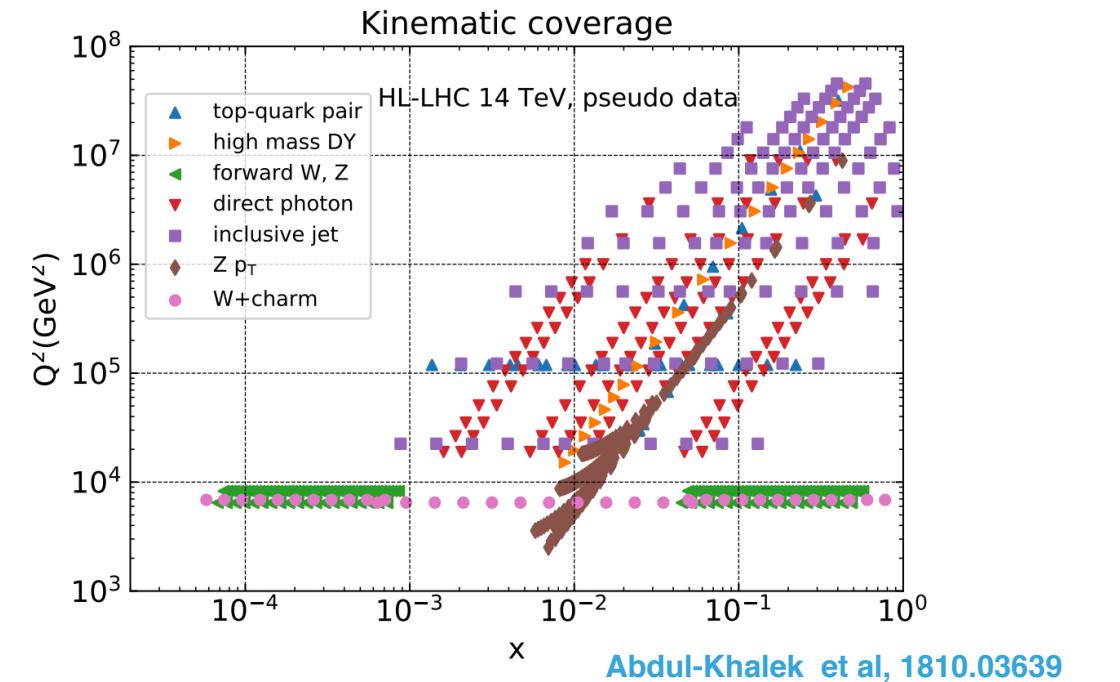


RESULTS: DRELL-YAN DATA @HL-LHC

- Add HL-LHC projections for both NC and CC in PDF fit

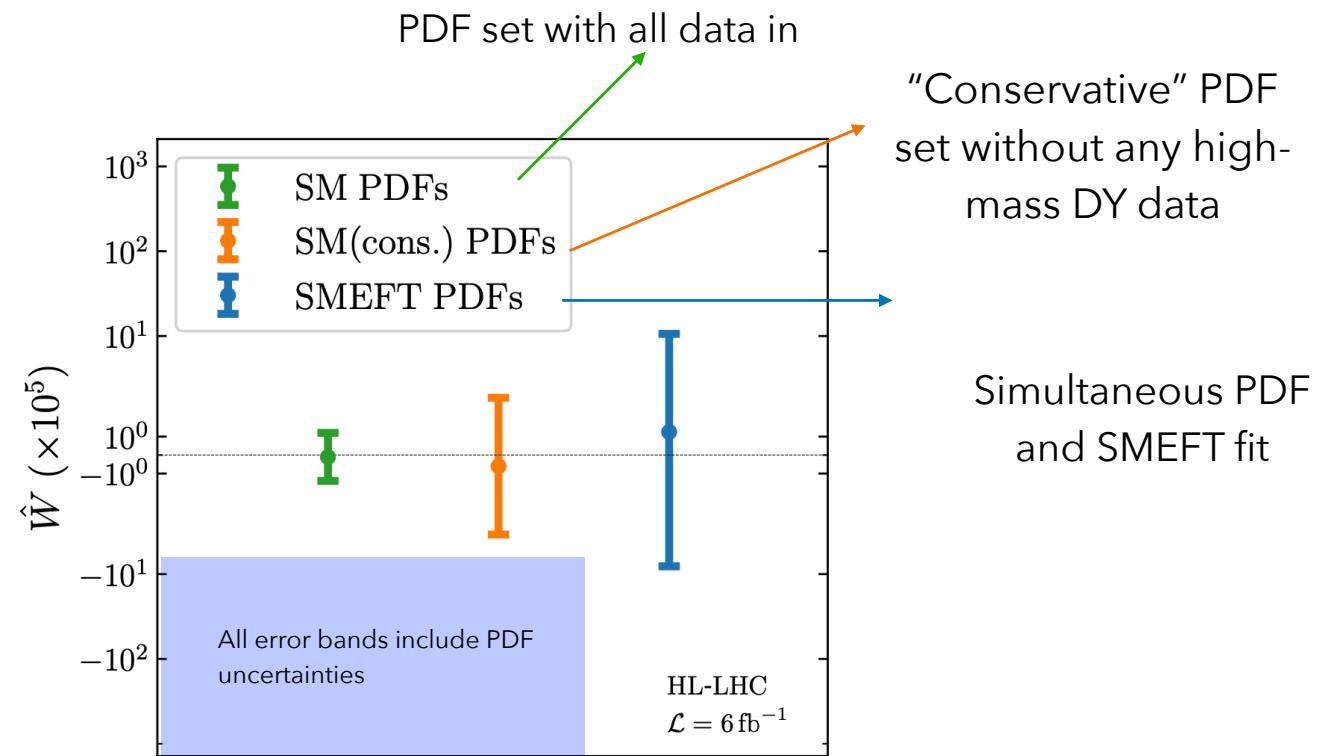
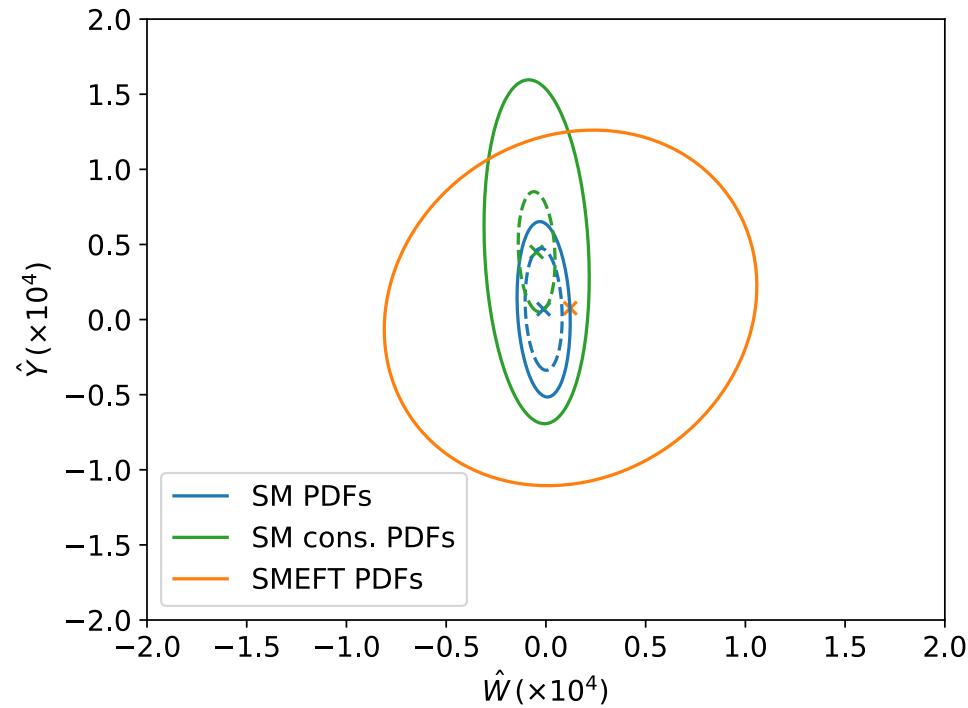
$$\sigma_i^{\text{hlhc}} \equiv \sigma_i^{\text{th}} \left(1 + \lambda \delta_{\mathcal{L}}^{\text{exp}} + r_i \delta_{\text{tot},i}^{\text{exp}} \right), \quad i = 1, \dots, n_{\text{bin}}$$

$$\delta_{\text{tot},i}^{\text{exp}} = \left((\delta_i^{\text{stat}})^2 + \sum_{j=1}^{n_{\text{sys}}} (f_{\text{red},j} \delta_{i,j}^{\text{sys}})^2 \right)^{1/2}$$



INTERPLAY @ HL-LHC

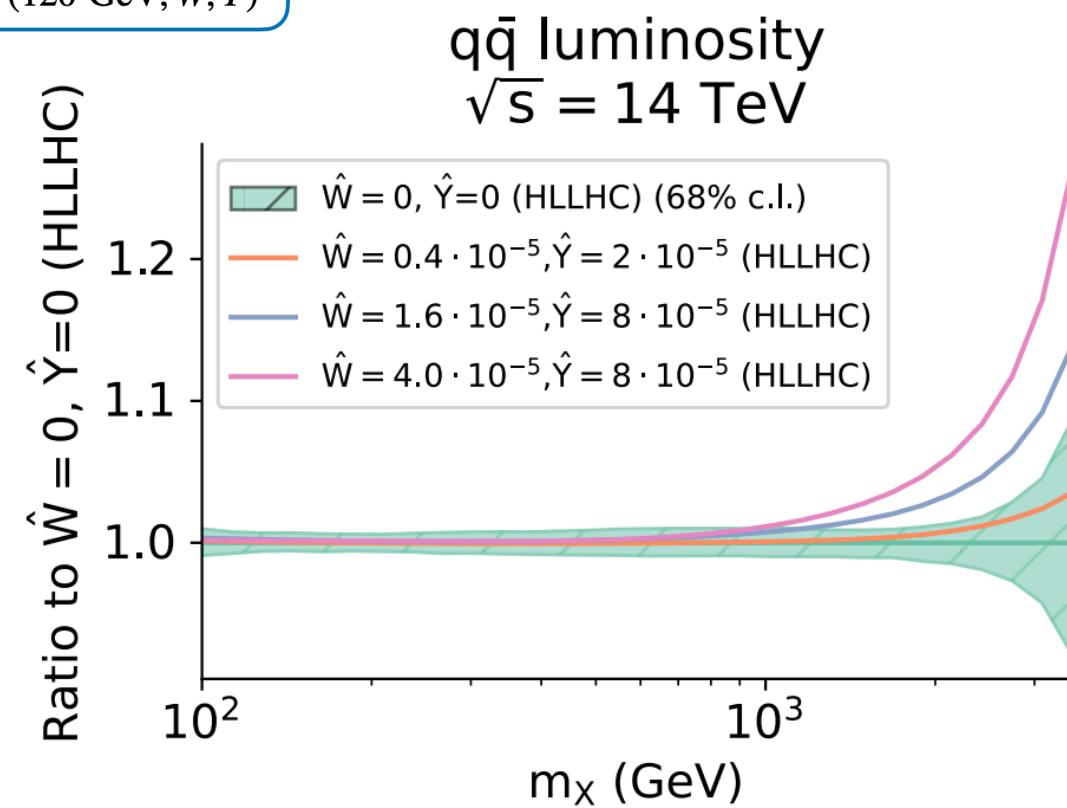
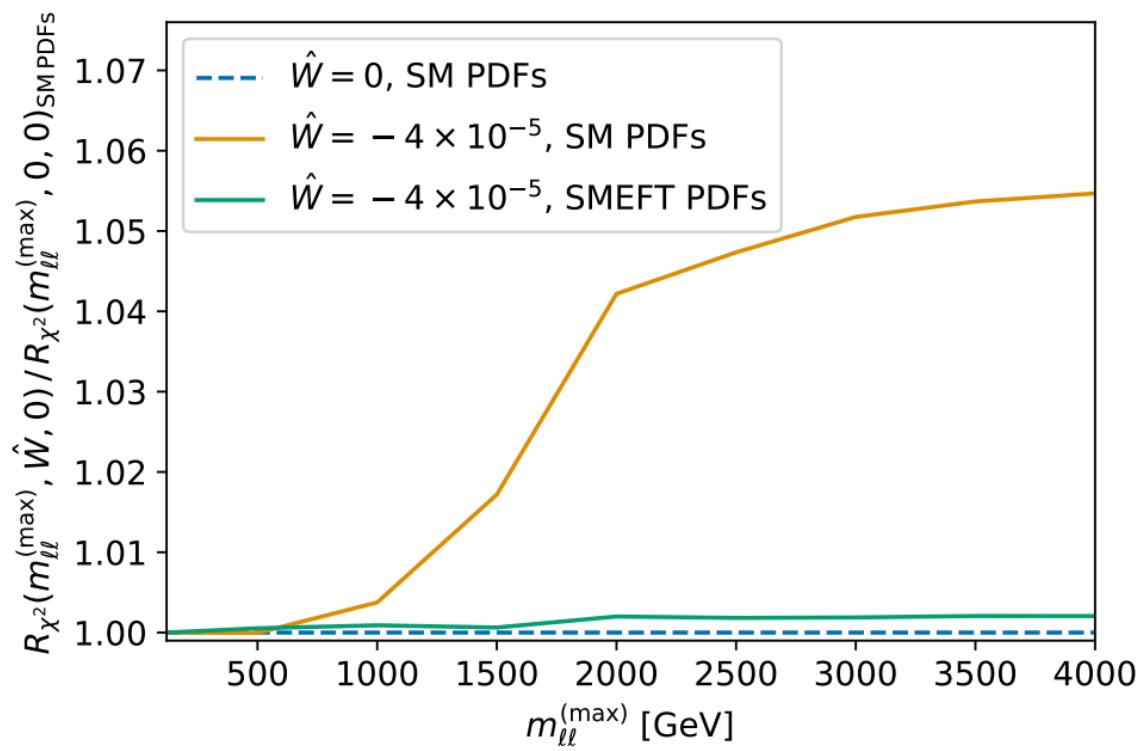
- Compare Wilson coefficients bounds from HL-LHC projections assuming SM PDFs (that include NC+CC data) to the bounds on the same Wilson coefficients obtained from a simultaneous fit of PDFs and Wilson coefficients
- Not accounting for interplay (using PDFs as a black box) leads to over-constrained bounds
- PDFs do absorb effect of new physics in this case!



INTERPLAY @ HL-LHC

- Compare Wilson coefficients bounds from HL-LHC projections assuming SM PDFs (that include NC+CC data) to the bounds on the same Wilson coefficients obtained from a simultaneous fit of PDFs and Wilson coefficients
- Not accounting for interplay (using PDFs as a black box) leads to over-constrained bounds
- PDFs do absorb effect of new physics in this case!

$$R_{\chi^2}(m_{\ell\ell}^{(\max)}, \hat{W}, \hat{Y}) \equiv \frac{\chi^2(m_{\ell\ell}^{(\max)}, \hat{W}, \hat{Y})}{\chi^2(120 \text{ GeV}, \hat{W}, \hat{Y})}$$



THE SIMUNET ANALYSIS

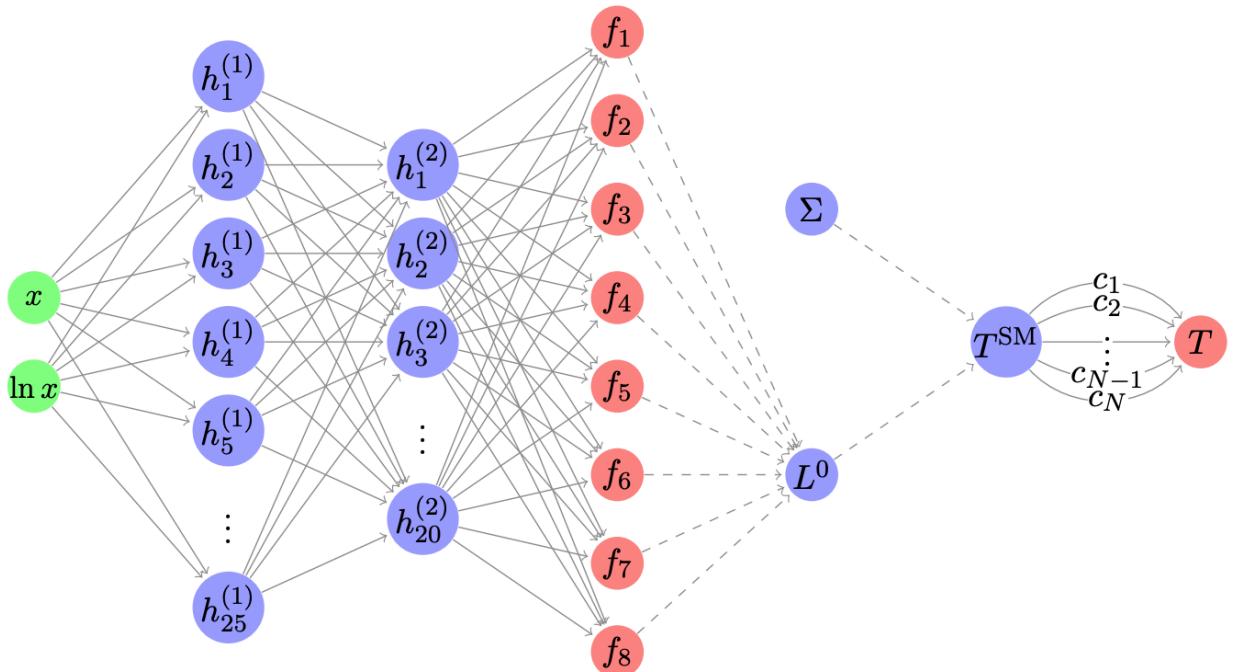
- With SimuNET we can do a truly simultaneous fit, rather than a scan in benchmark point and it does not have limit in number of parameters that can be fitted alongside PDFs at the initial scale!

[Iranipour, MU, Voisey: 2201.07240]

Input layer	Hidden layer 1	Hidden layer 2	PDF flavours	Convolution step	SM Observable	SMEFT Observable
-------------	----------------	----------------	--------------	------------------	---------------	------------------

Linear dim-6 operator

$$T(\hat{\theta}) = \Sigma(\{c_n\}) \cdot L^0(\theta) = T^{\text{SM}}(\theta) \cdot \left(1 + \sum_{n=1}^N c_n R_{\text{SMEFT}}^{(n)} \right)$$



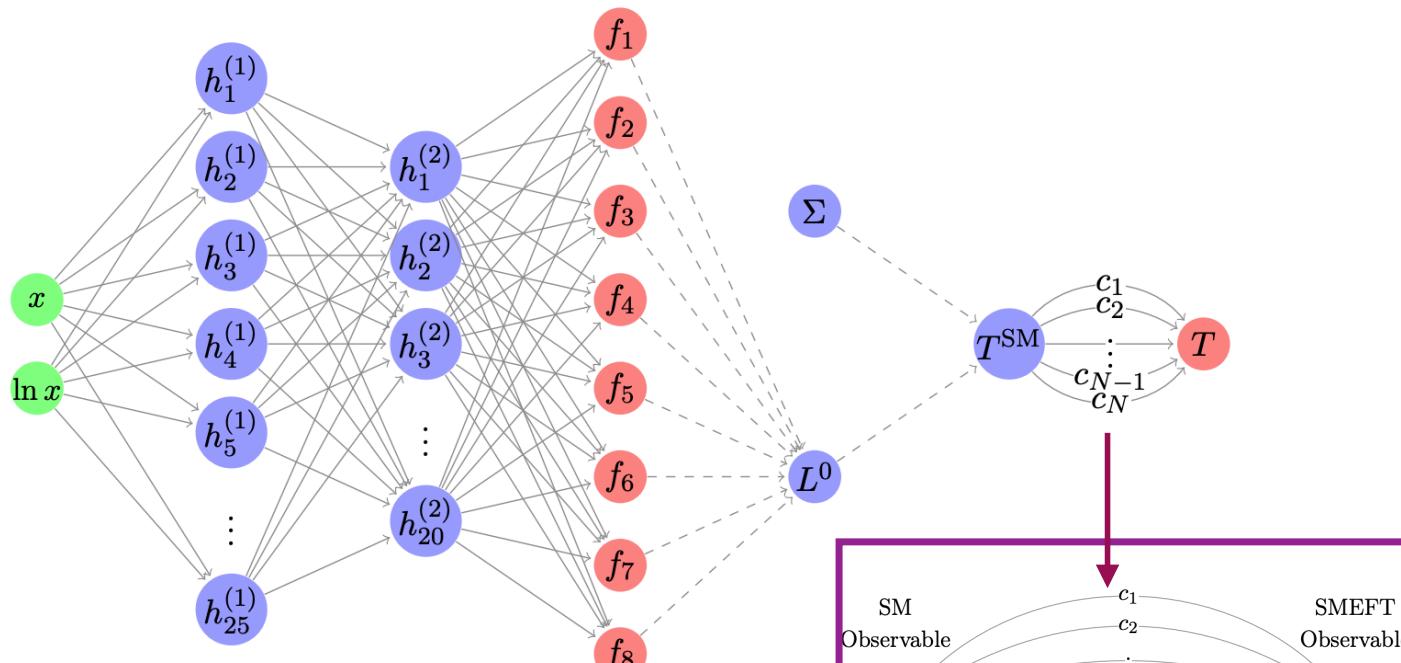
$$T^{\text{SM}}(\theta) = \Sigma^{\text{SM}} \cdot L^0(\theta)$$

THE SIMUNET ANALYSIS

- With SimuNET we can do a truly simultaneous fit, rather than a scan in benchmark point and it does not have limit in number of parameters that can be fitted alongside PDFs at the initial scale!

[Iranipour, MU, Voisey: 2201.07240]

Input layer	Hidden layer 1	Hidden layer 2	PDF flavours	Convolution step	SM Observable	SMEFT Observable
-------------	----------------	----------------	--------------	------------------	---------------	------------------



Linear dim-6 operator

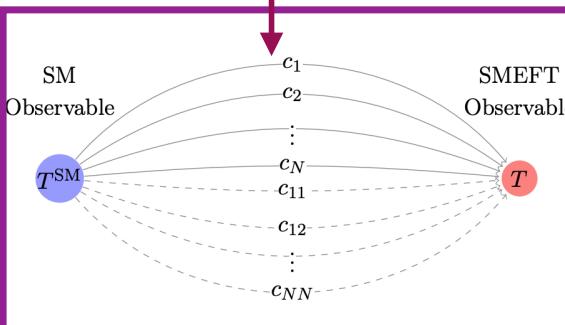
$$T(\hat{\theta}) = \Sigma(\{c_n\}) \cdot L^0(\theta) = T^{\text{SM}}(\theta) \cdot \left(1 + \sum_{n=1}^N c_n R_{\text{SMEFT}}^{(n)} \right)$$

$$T^{\text{SM}}(\theta) = \Sigma^{\text{SM}} \cdot L^0(\theta)$$

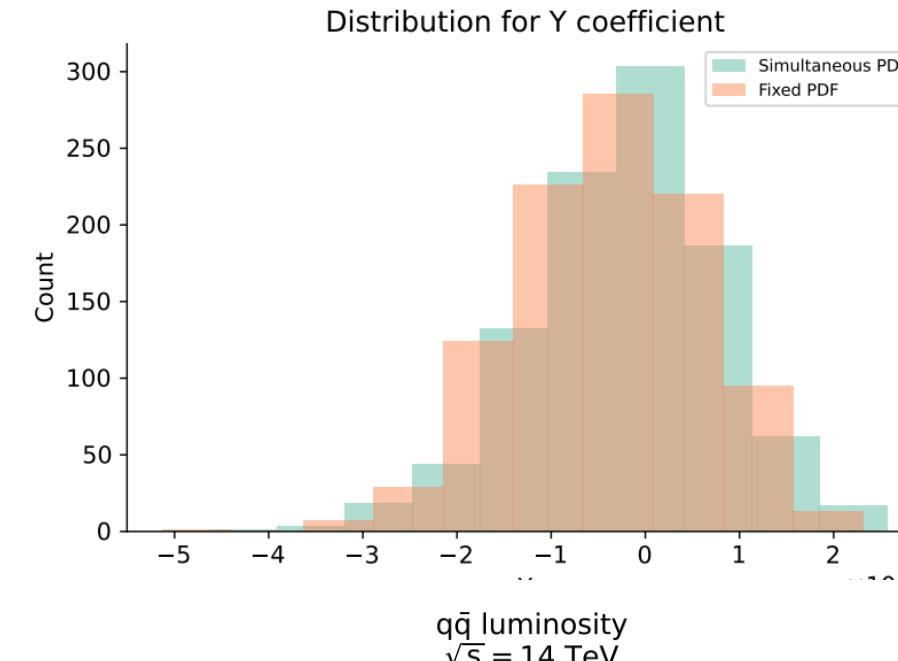
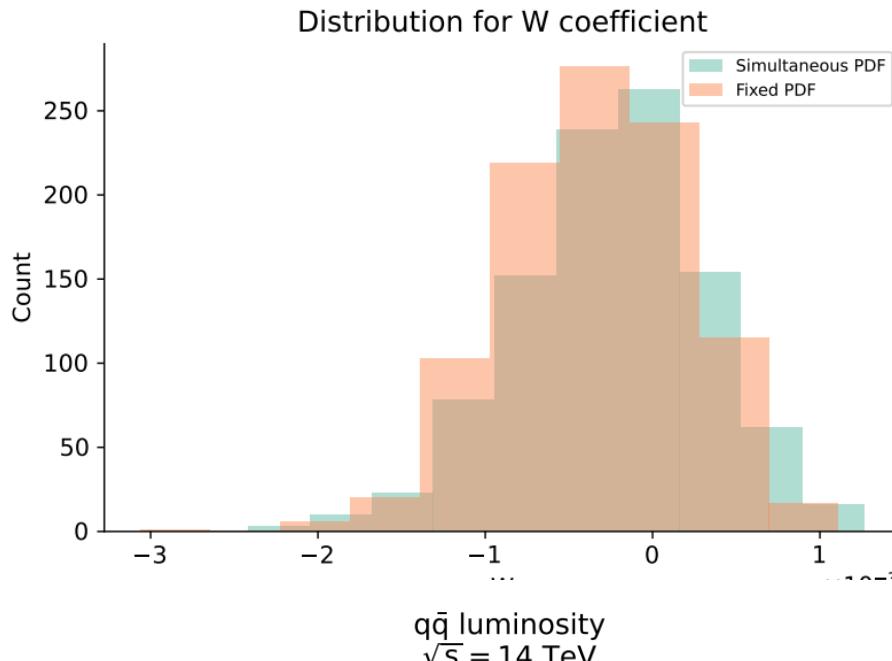
Quadratic dim-6 operator

$$T(\hat{\theta}) = T^{\text{SM}}(\theta) \cdot \left(1 + \sum_{n=1}^N c_n R_{\text{SMEFT}}^{(n)} + \sum_{1 \leq n \leq m \leq N} c_{nm} R_{\text{SMEFT}}^{(n,m)} \right)$$

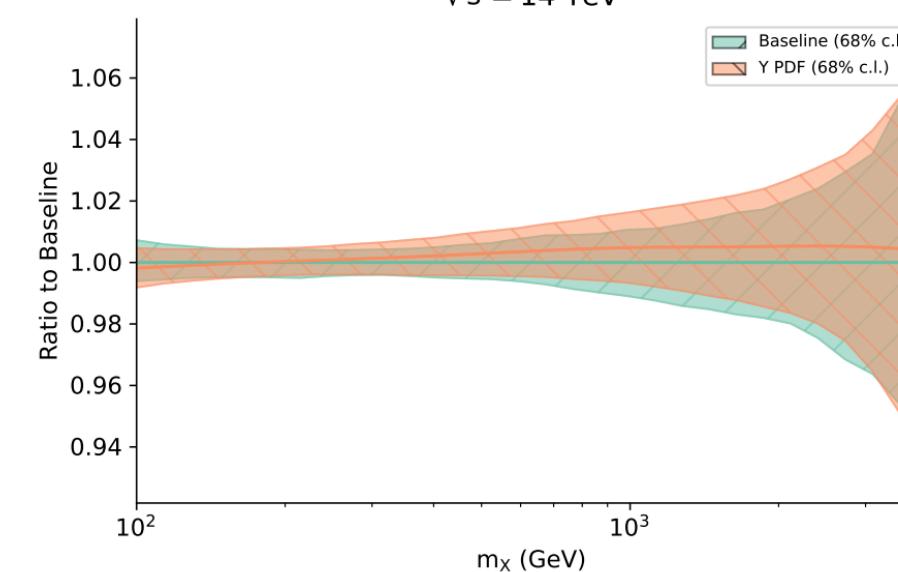
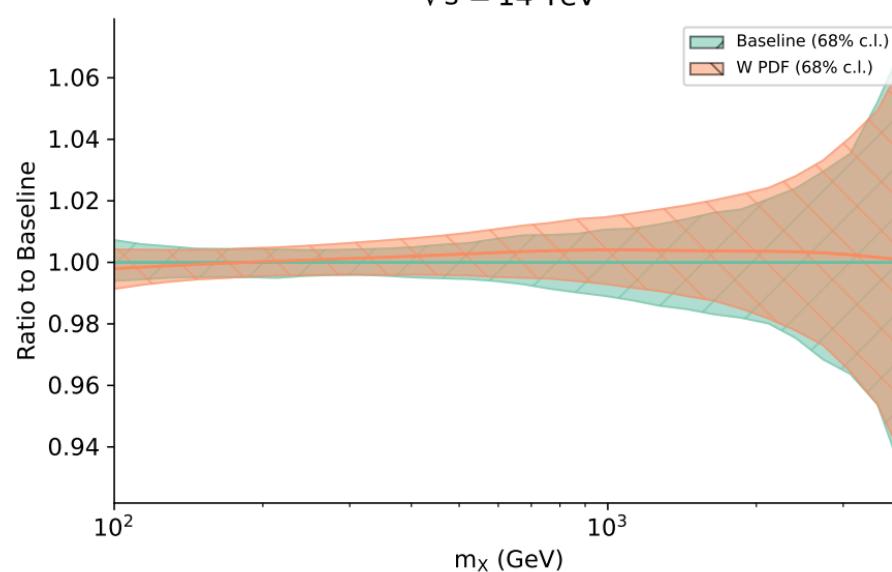
$c_n c_m$



RESULTS: DRELL-YAN DATA @RUN1 AND RUN2



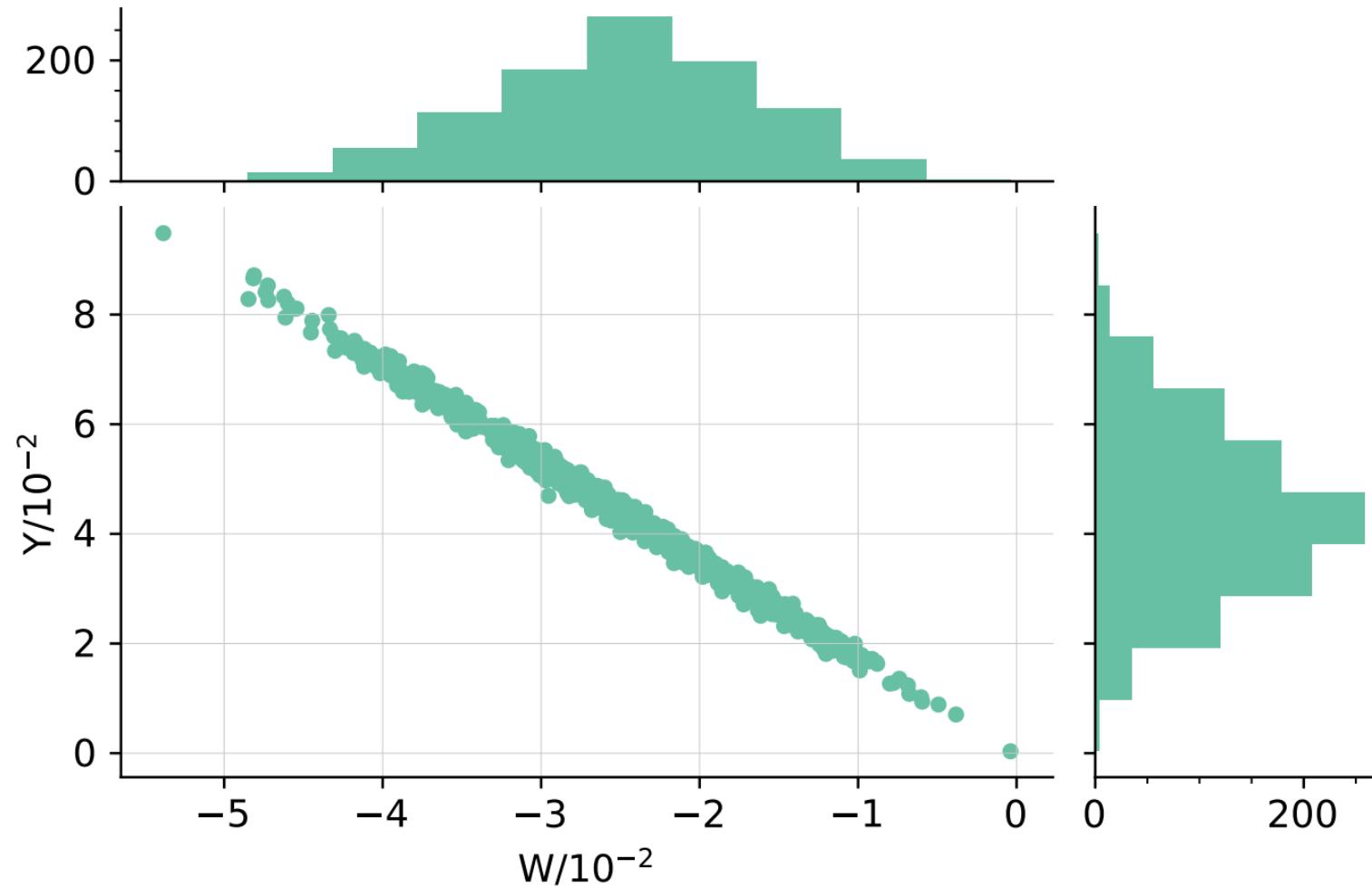
Distribution of W & Y best fits over MC reps with fixed SM PDFs (baseline)



Distribution of W & Y best fits over MC reps with PDFs fitted alongside them

Same comparison for quark-antiquark luminosity

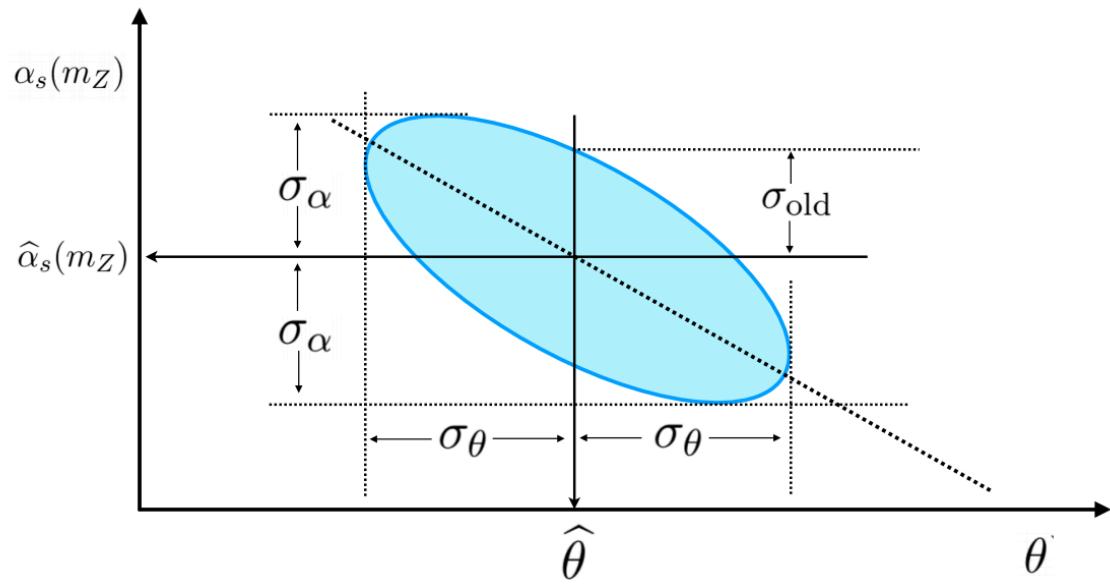
RESULTS: DRELL-YAN DATA @RUN1 AND RUN2



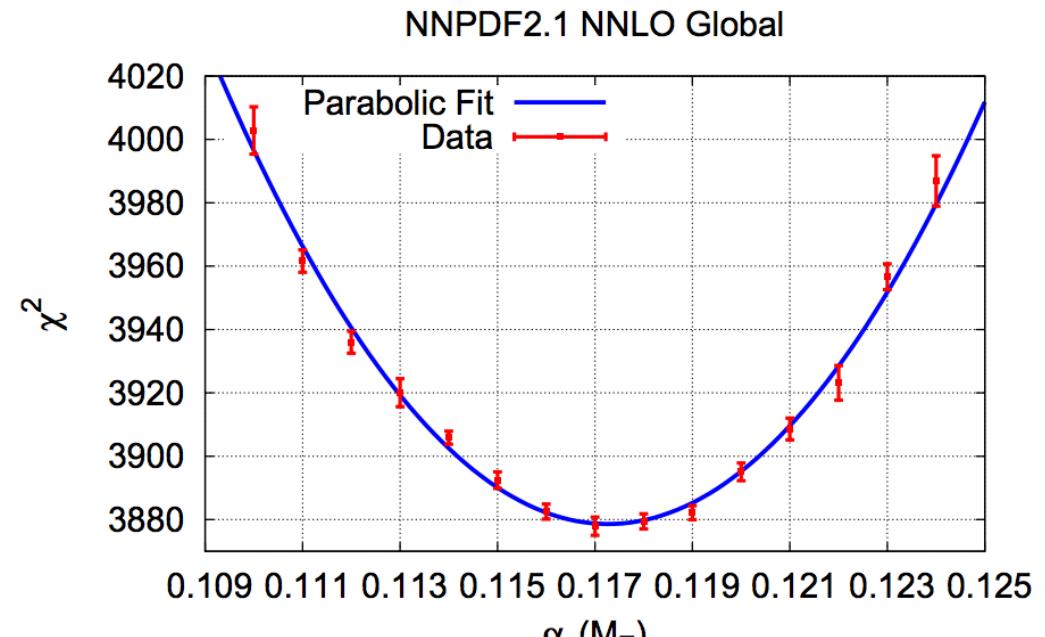
- ✓ Simultaneous analysis confirms results of previous study based on scan on benchmark points in the SMEFT space: with current data effect is not-negligible but small compared to PDF uncertainties
- ✓ Methodology able to find flat direction in W-Y parameter space
- ✓ To eliminate it, need Drell-Yan charged current data

PDFs AND α_s

- PDFs and α_s strongly correlated (PDF evolution with the scale and hard cross sections)
- Cleanest determinations of α_s from processes that do not require knowledge of the PDFs
- A determination of α_s jointly with the PDFs has advantage that it is driven by the combination of many experimental measurements from several different processes.



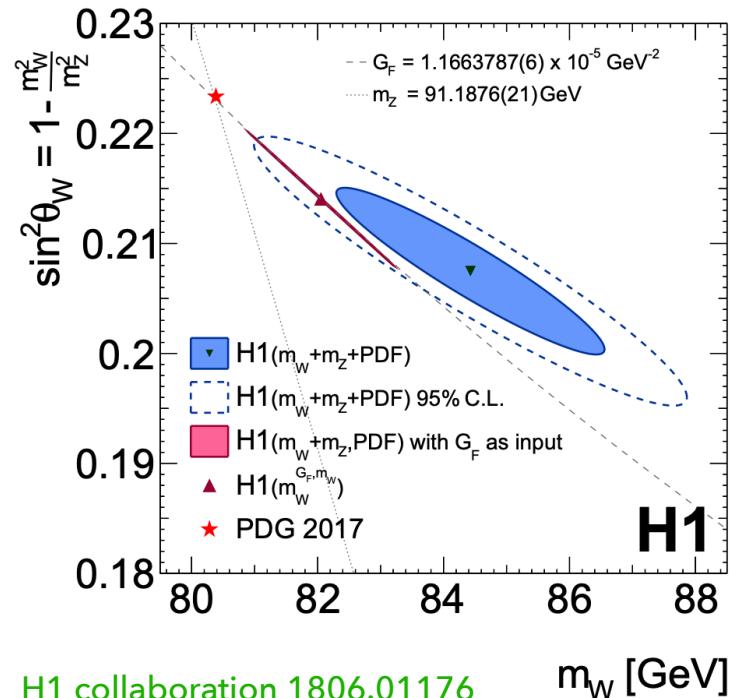
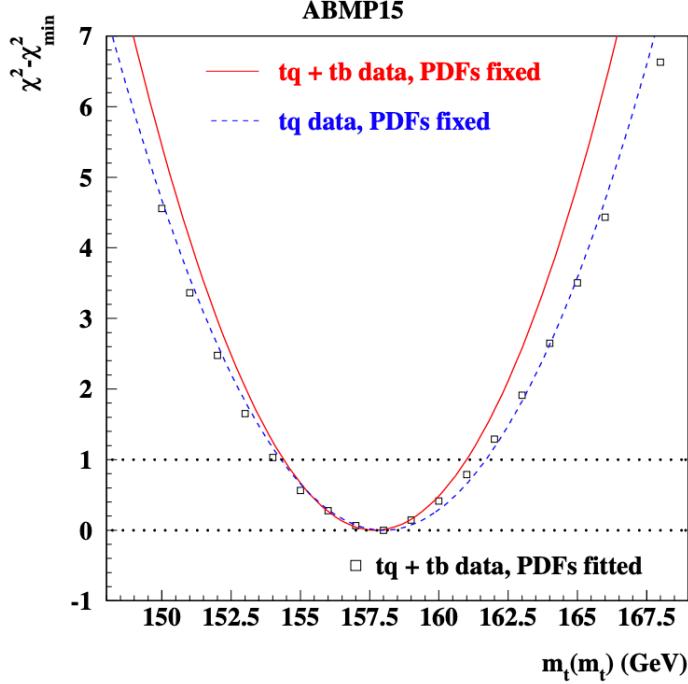
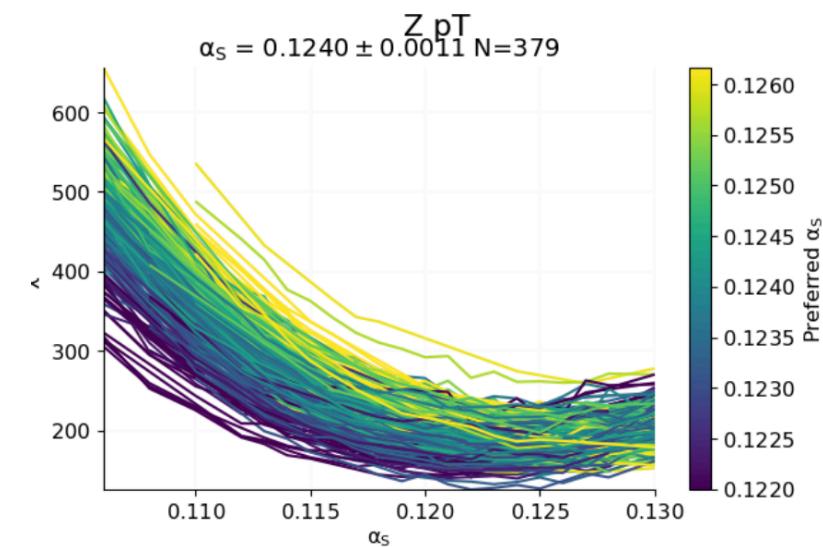
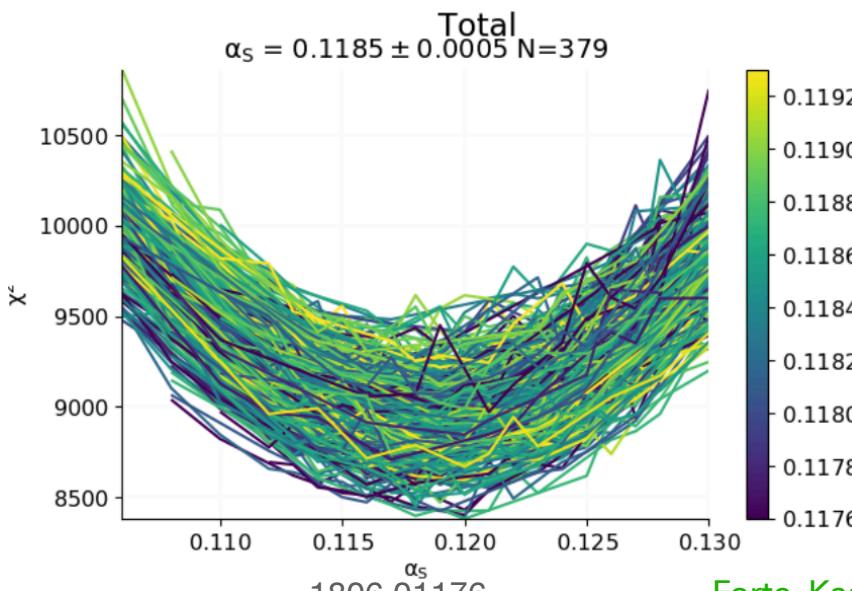
Ball, Carrazza, Del Debbio, Forte, Kassabov, Rojo, Slade, MU 1802.03398



- Early determinations involve a scan over α_s and ignored PDF and α_s correlation in the fit
- Recent simultaneous determination of PDF and α_s using correlated replica method
- Many determination of α_s from analyses of specific LHC processes have been published recently (from $t\bar{t}$, Z and W production, jets)
- How reliable are such partial determination of α_s ?

SIMULTANEOUS FITS FOR SM PARAMETERS

- Given the strong correlation between PDFs of the proton and α_s , a non simultaneous determination of α_s along with the PDFs from LHC processes might yield misleading results

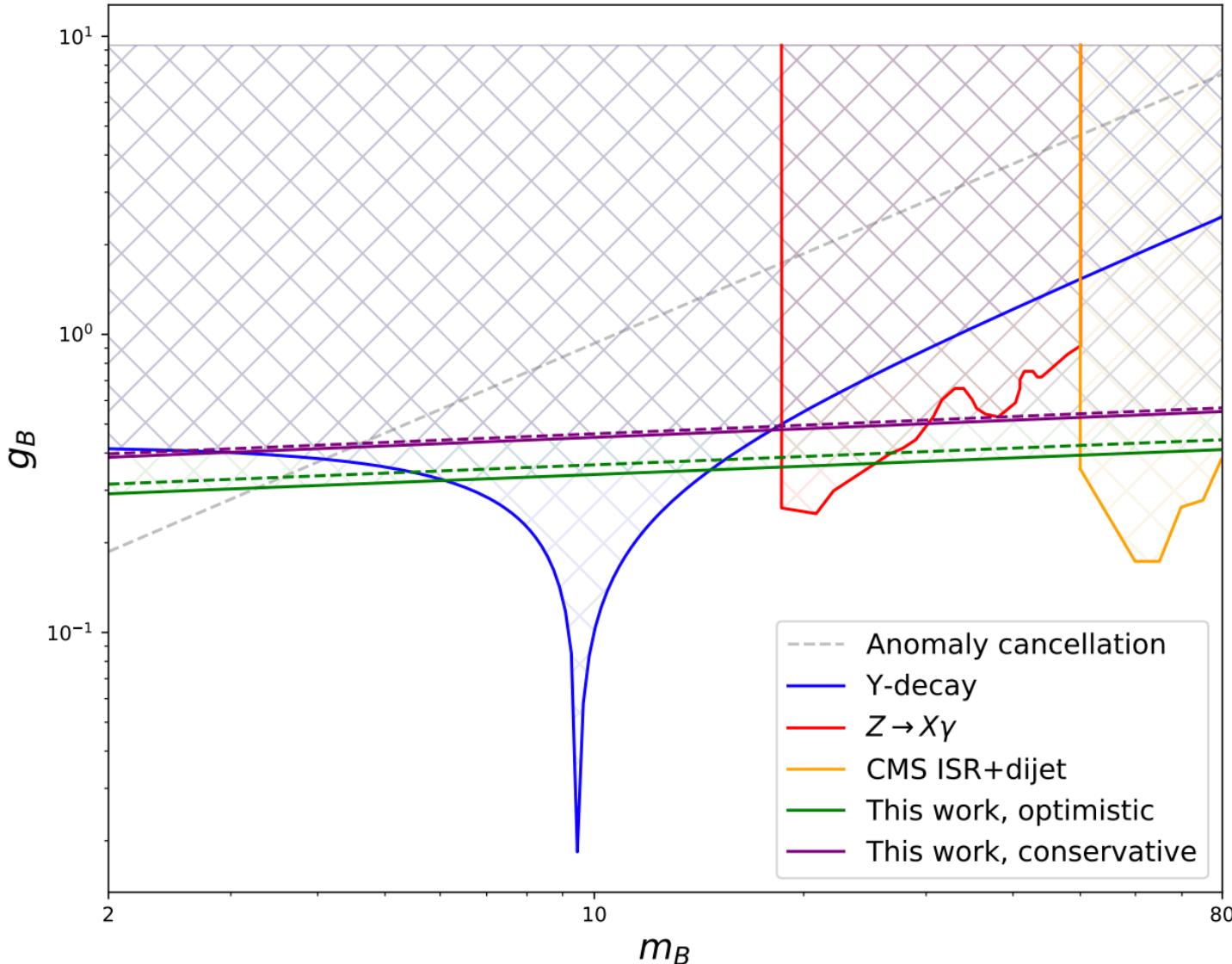


Alekhin, Moch, Their 1608.05212

- Correlation of PDFs and the EW parameters or m_t weaker than in the case of α_s , but the very high accuracy which is sought suggests that the effect of simultaneous determination is not negligible
- Similar considerations for fits of polarised/unpolarised PDFs, proton/nuclear PDFs or PDFs and FFs (universal fits)

DARK PHOTON

M. McCullough, J. Moore, MU, arXiv:2203.12628



- If there was a lepto-phobic dark photon weakly coupled to quarks, it would appear among the partons of the proton.
$$\mathcal{L}_{\text{int}} = \frac{1}{3} g_B \bar{q} \not{B} q$$
- The presence of the dark Parton would modify the evolution of standard quarks and gluon.
- Precise LHC data can indirectly constrain parameter space of the dark photon in a competitive way compared to direct searches