European Research Council

Established by the European Commission

PBSP

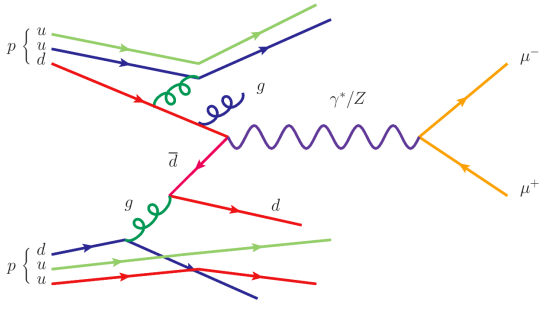
MARIA UBIALI
UNIVERSITY OF CAMBRIDGE

PDFS AND BSM (SMEFT) INTERPLAY

WORKSHOP ON TOOLS FOR HIGH PRECISION LHC SIMULATIONS - RINGBERG CASTLE, GERMANY

4TH NOVEMBER 2022

PDFS AND LHC PRECISION PROGRAM



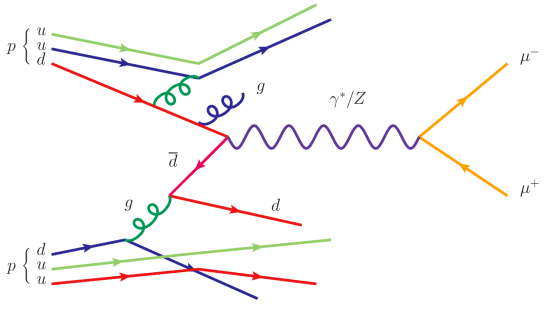
$$\boxed{d\sigma^{pp\to ab}} = \sum_{i,j} f_i \otimes f_j \otimes d\hat{\sigma}^{ij\to ab} + \ldots$$

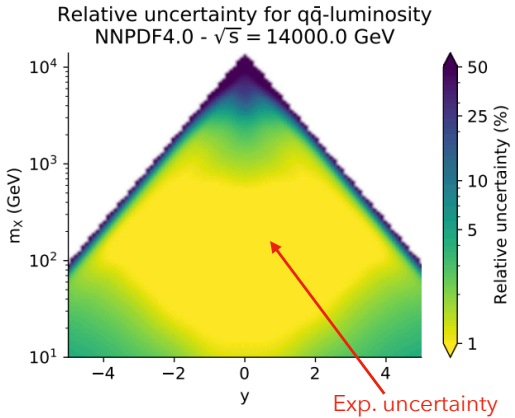
T

Partonic cross sections computed in pQCD (+EW)

PDFs: non-perturbative functions that indicate the fraction of proton's momentum carried by partons i and j

PDFS AND LHC PRECISION PROGRAM





$$\boxed{d\sigma^{pp\to ab}} = \sum_{i,j} f_i \otimes f_j \otimes d\hat{\sigma}^{ij\to ab} + \ldots$$

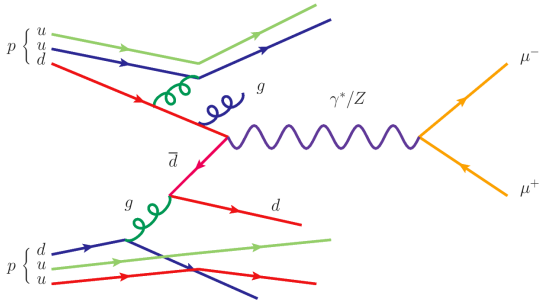
T

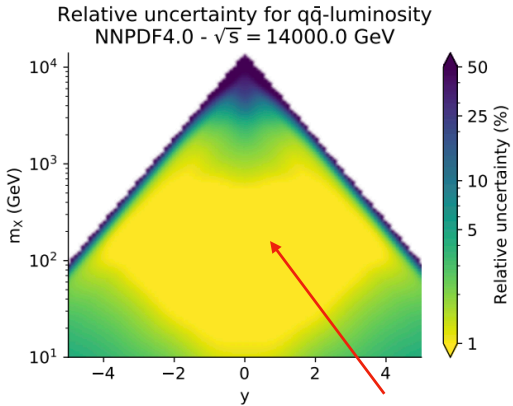
Partonic cross sections computed in pQCD (+EW)

PDFs: non-perturbative functions that indicate the fraction of proton's momentum carried by partons i and j

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

PDFS AND LHC PRECISION PROGRAM





$$\boxed{d\sigma^{pp\to ab}} = \sum_{i,j} f_i \otimes f_j \otimes d\hat{\sigma}^{ij\to ab} + \ldots$$

T

Partonic cross sections computed in pQCD (+EW)

PDFs: non-perturbative functions that indicate the fraction of proton's momentum carried by partons i and j

PBSP

When we extract PDFs from current experimental data, we assume the SM both in the parton content of the proton and in the matrix elements. Is this a good enough assumption at the level of precision that we are aiming for?

E. Hammou, Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales, MU

OUTLINE

- Introduction
  - ➡ The statement of the problem

- PDFs and SMEFT fits interplay
  - ➡ The simuNET methodology
  - ➡ Simultaneous fits: the Drell-Yan sector and the top sector
  - ➡ Can PDFs absorb new physics?

- BSM partons
  - ➡ A dark photon example

- Conclusions and outlook

INTRODUCTION

EXTRACTING PHYSICS PARAMETERS FROM LHC DATA

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

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \sum_i^{N_{d6}} \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_j^{N_{d8}} \frac{b_j}{\Lambda^4}\mathcal{O}_j^{(8)} + \ldots$$

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

SM parameters [$\alpha_s(M_z)$, M_w, G_F, α, m_t,…] & **SMEFT WCs**

Parameters determining PDFs at initial scale

✓ In a PDF fit typically

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

$\bar{c} = 0$ in the case of SMEFT WCs

EXTRACTING PHYSICS PARAMETERS FROM LHC DATA

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

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \sum_i^{N_{d6}} \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_j^{N_{d8}} \frac{b_j}{\Lambda^4}\mathcal{O}_j^{(8)} + \ldots$$

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

SM parameters [$\alpha_s(M_z)$, M_w, G_F, α, m_t,…] & **SMEFT WCs**

Parameters determining PDFs at initial scale

✓ In a PDF fit typically

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

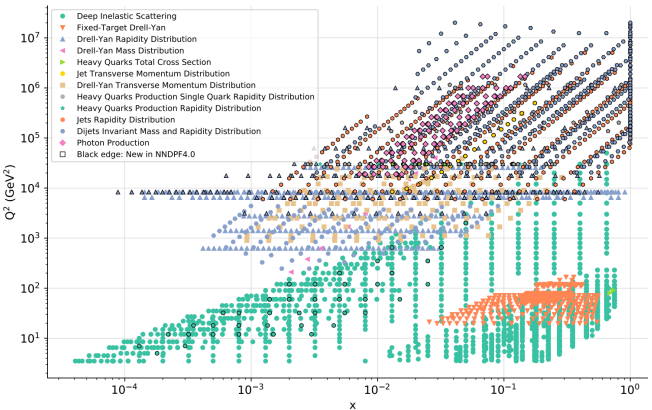
$\bar{c} = 0$ in the case of SMEFT WCs

✓ In a SMEFT fit

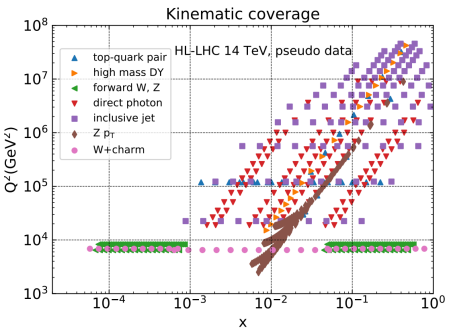
$$T_i(\{c\}) = {\rm PDFs}(\{\bar{\theta}\},\{\bar{c}\}) \otimes \hat{\sigma}_i(\{c\})$$

INTERPLAY BETWEEN PDF FITS AND SMEFT FITS

- 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.
- In principle low-scale physics is separable from high-scale physics, BUT the complexity of the LHC environment might well intertwine them (e.g. due to data overlap).



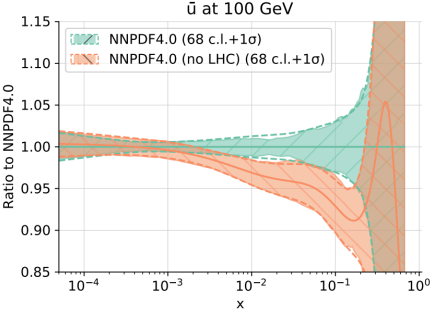
Ball et al, arXiv:2109.02653



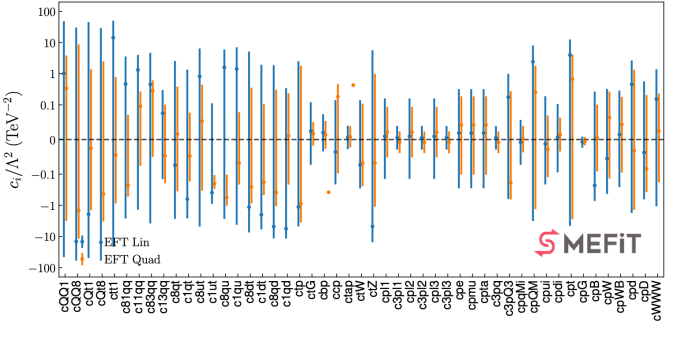
Abdul Khalek et al, arXiv:1810.03639

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 theory predictions computed adding the same operators that I am fitting?



Ball et al, arXiv:2109.02653



Ethier et al, arXiv: 2105.00006

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 theory predictions computed adding the same operators that I am fitting?

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

Simultaneous fits

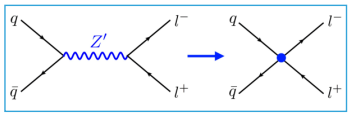
$$D_i(\{\bar{\theta}\},\{\bar{c}\})$$

Contaminated fits

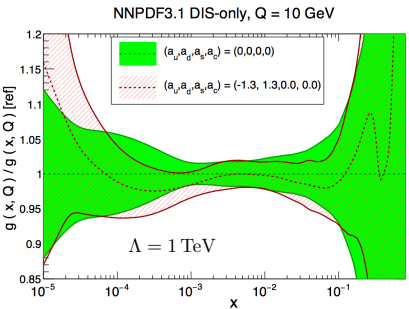
INTERPLAY BETWEEN SMEFT FITS AND PDF FITS

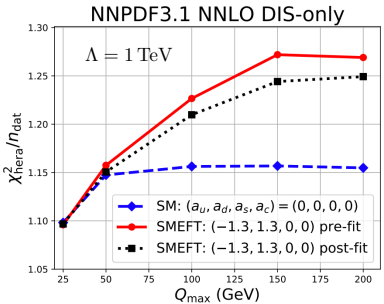
DEEP-INELASTIC SCATTERING

- First study of interplay in case of DIS data [Carrazza, Degrande, Iranipour, Rojo, MU, Phys.Rev.Lett. 123 (2019) 13, 132001]
- Simple scenario, only right-handed 4F operators, lepton flavour blind, quark flavours split to evade LEP constraints
- PDF fits based on DIS only data (Q ≲ 200 GeV for HERA data)



$$\mathcal{O}_{lu} = \left(\bar{l}_R \gamma^\mu l_R\right)\left(\bar{u}_R \gamma_\mu u_R\right)$$
$$\mathcal{O}_{lc} = \left(\bar{l}_R \gamma^\mu l_R\right)\left(\bar{c}_R \gamma_\mu c_R\right)$$
$$\mathcal{O}_{ld} = \left(\bar{l}_R \gamma^\mu l_R\right)\left(\bar{d}_R \gamma_\mu d_R\right)$$
$$\mathcal{O}_{ls} = \left(\bar{l}_R \gamma^\mu l_R\right)\left(\bar{s}_R \gamma_\mu s_R\right)$$





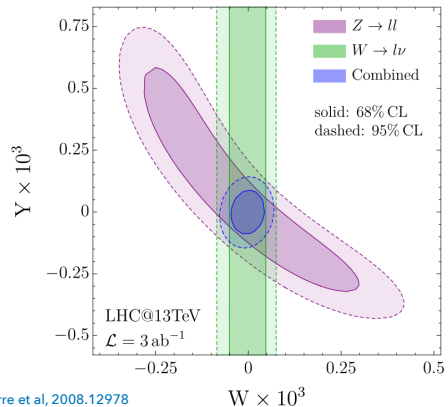
Only gluon affected by the presence of non-zero a_R coefficients, but distortion of PDFs leads to a deterioration of data-theory agreement that scales with energy => A fit based on DIS data is basically unaffected by interplay and the effects of new physics can be disentangled

DRELL-YAN INVARIANT MASS DISTRIBUTIONS

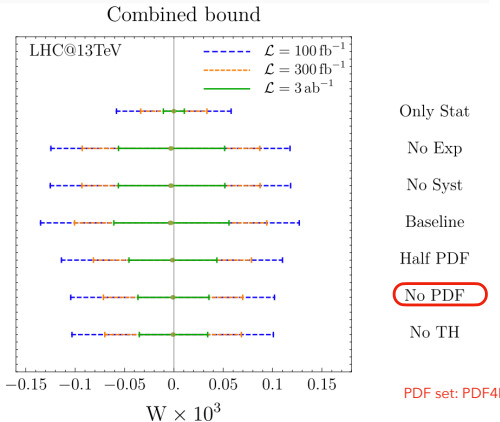
- Case study at higher energy: EW oblique corrections in high-mass NC and CC Drell-Yan tails.
- W and Y parametrise the self-energy of gauge bosons and are powerful probes of quark-lepton contact interactions that produce effects that grow with energy [Torre et al, 2008.12978]
- Interplay first studied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

$$\mathscr{L}_{\rm SMEFT} \supset -\frac{\hat{W}}{4m_W^2}(D_\rho W^a_{\mu\nu})^2 - \frac{\hat{Y}}{4m_W^2}(\partial_\rho B_{\mu\nu})^2$$

$$\hat{W} \propto (\bar{l}\sigma^a\gamma^\mu l)(\bar{q}\sigma^a\gamma_\mu q)$$
$$\hat{Y} \propto (\bar{l}_i\gamma^\mu l_i)(\bar{q}_j\gamma_\mu q_j)$$



Torre et al, 2008.12978



PDF set: PDF4LHC15_nlo_30_pdfas

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]

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

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

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

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

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

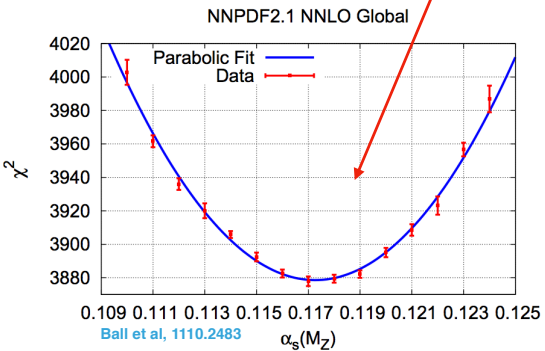
SMEFT PDFs / Simultaneous fit

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

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

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



Ball et al, 1110.2483

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

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

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

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

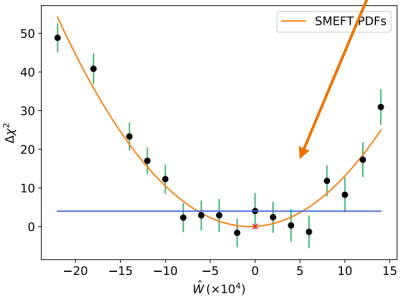
SMEFT PDFs / Simultaneous fit

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

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

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



Greljo et al, 2104.02723

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

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

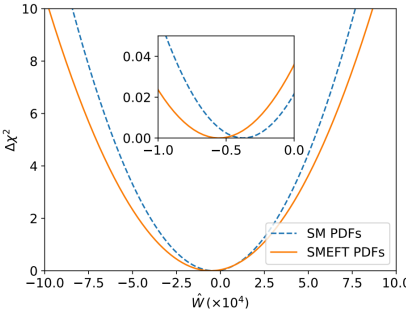
SM PDFs

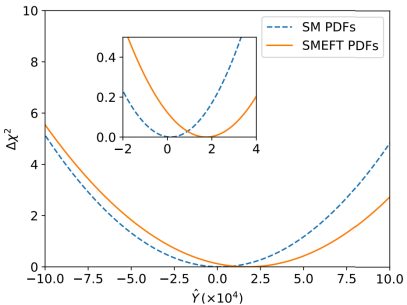
1. Take data, make theoretical predictions accounting for operator in partonic cross section and PDFs.
2. Compute chi2 as a function of WCs (Wilson Coefficients)
3. Minimise chi2 and find best-fit and C.L.s of WCs
4. Extract bounds

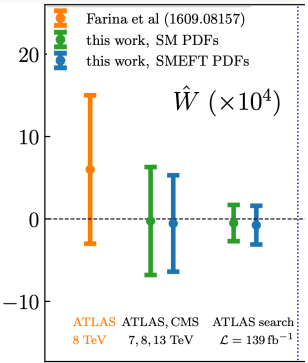
$$T = f_{1,\rm BSM} \otimes f_{2,\rm BSM} \otimes \hat{\sigma}_{\rm BSM}$$

SMEFT PDFs / Simultaneous fit

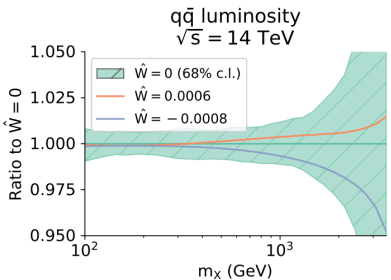
INTERPLAY @ RUN I AND RUN II







- 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)

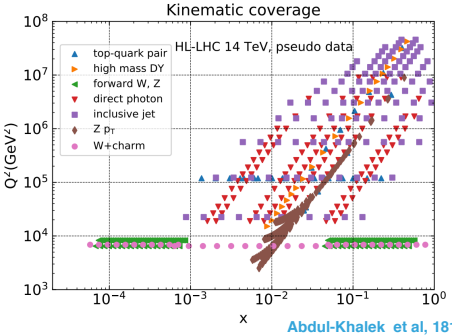


RESULTS: DRELL-YAN DATA @HL-LHC

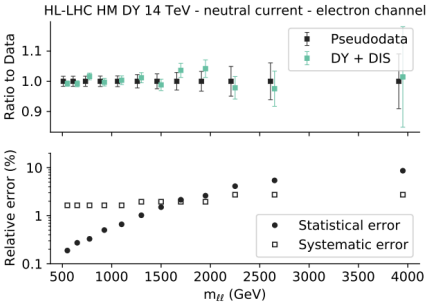
- Add HL-LHC projections for both NC and CC in PDF fit

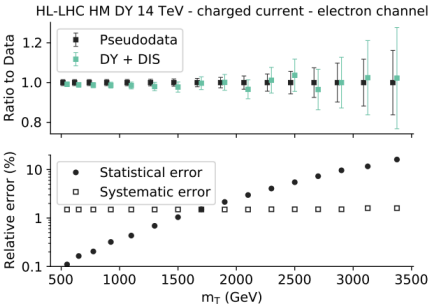
$$\sigma_i^{\rm hllhc} \equiv \sigma_i^{\rm th}\left(1+\lambda\delta_{\mathcal{L}}^{\rm exp}+r_i\delta_{{\rm tot},i}^{\rm exp}\right), \qquad i=1,\ldots,n_{\rm bin}$$

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



Abdul-Khalek et al, 1810.03639

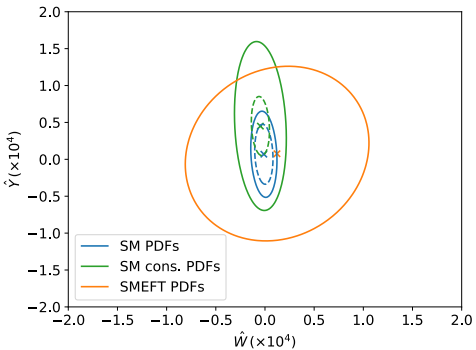


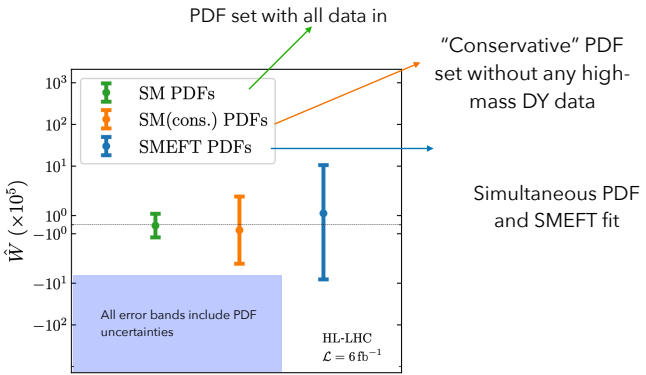


+ muon channel

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!



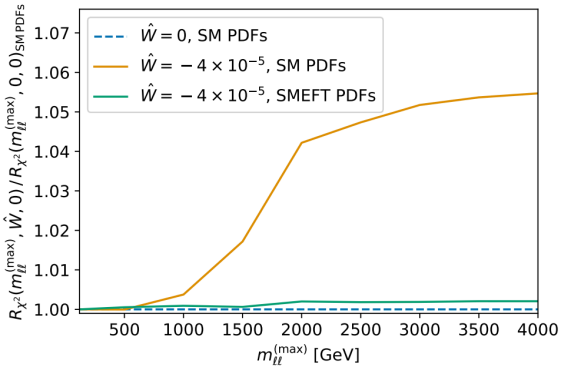


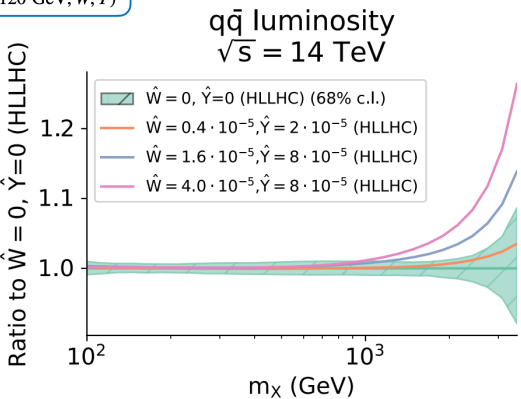
Greljo et al, 2104.02723

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!

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





Greljo et al, 2104.02723

SIMUNET: A DEEP-LEARNING BASED SIMULTANEOUS FIT

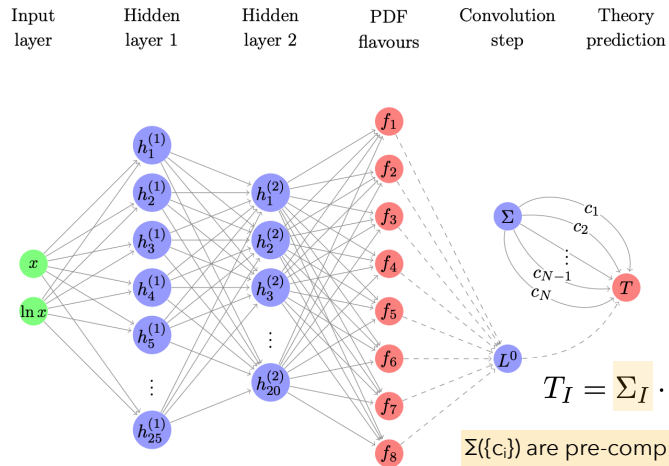
- The idea: take a PDF fit based on NNPDF4.0 methodology and make dependence of observables on physics parameters $\{c_i\}$ explicit via fast interface before computing the loss function (e.g. adding SMEFT corrections, expanding observables in terms of SM precision parameters)

- Perform minimisation of loss function over

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

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.



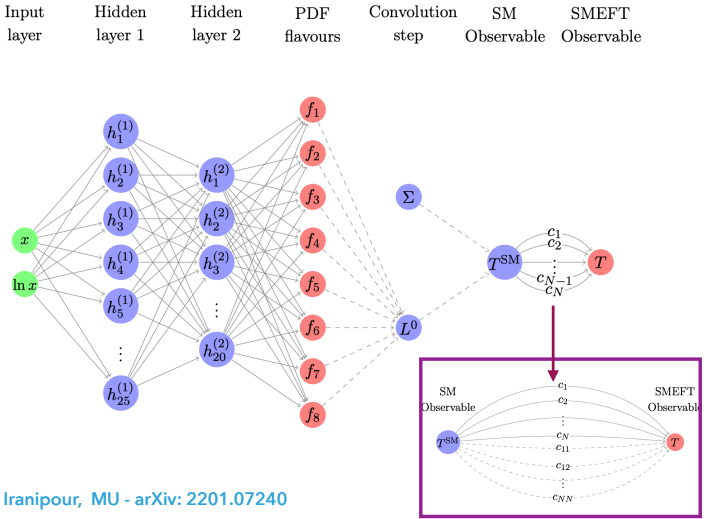
$$T_I = \Sigma_I \cdot L^0$$

$\Sigma(\{c_i\})$ are pre-computed tables for fast interface accounting for PDF evolution and part. xsec

S. Iranipour, MU - arXiv: 2201.07240

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!



Linear dim-6 operator

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

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

Quadratic dim-6 operator

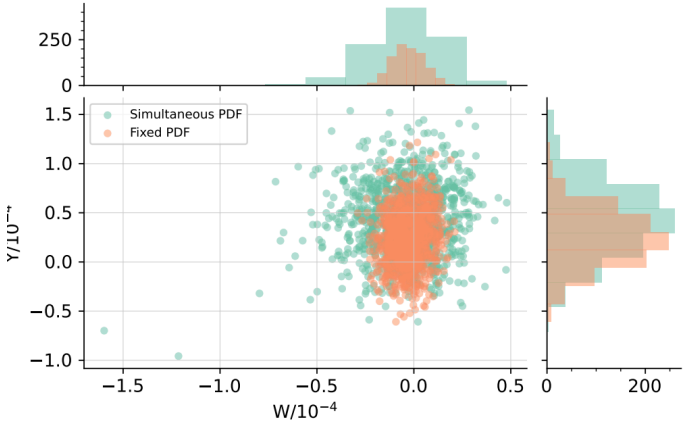
$$T(\hat{\theta}) = T^{\rm SM}(\theta) \cdot \left(1 + \sum_{n=1}^{N} c_n R^{(n)}_{\rm SMEFT} + \sum_{1 \le n \le m \le N} c_{nm} R^{(n,m)}_{\rm SMEFT}\,,\right)$$

$c_n c_m$

S. Iranipour, MU - arXiv: 2201.07240

RESULTS: DRELL-YAN DATA @HL-LHC

S. Iranipour, MU - arXiv: 2201.07240



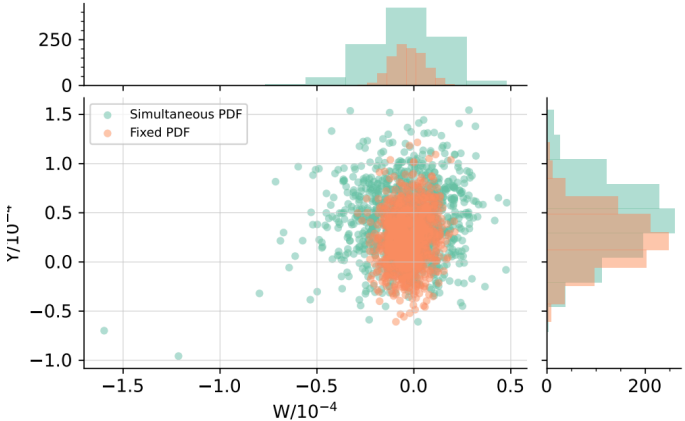
	SM PDFs	SMEFT PDFs
$W \times 10^5$ (68% CL)	$[-1.1, 0.5]$	$[-2.4, 1.5]$
$W \times 10^5$ (95% CL)	$[-2.0, 1.4]$	$[-4.3, 3.4]$
$Y \times 10^5$ (68% CL)	$[-0.4, 5.2]$	$[0.6, 8.0]$
$Y \times 10^5$ (95% CL)	$[-3.2, 8.1]$	$[-3.1, 11.7]$

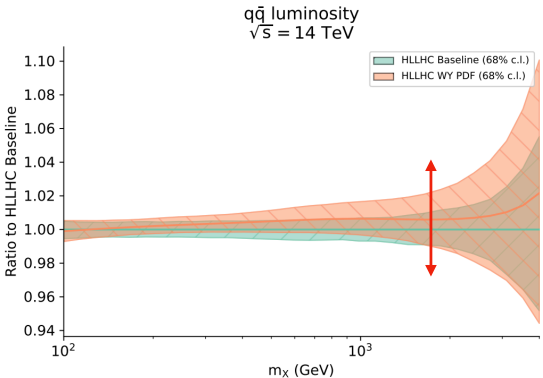
x 2.3 broadening of bounds for W
x 1.3 broadening of bounds for Y

✓ 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

RESULTS: DRELL-YAN DATA @HL-LHC

S. Iranipour, MU - arXiv: 2201.07240

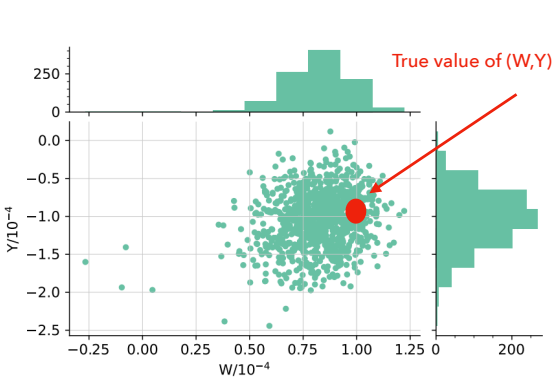


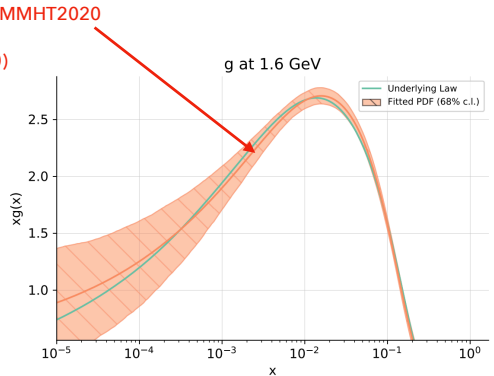


✓ Simultaneous analysis of PDFs and W&Y SMEFT coefficient of DIS + DY (including HL-LHC projections) using simuNET method shows that at HL-LHC the effect of interplay becomes important as WCs bounds broaden and PDFs change significantly once SMEFT effects allowed in theory predictions entering PDF fit

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

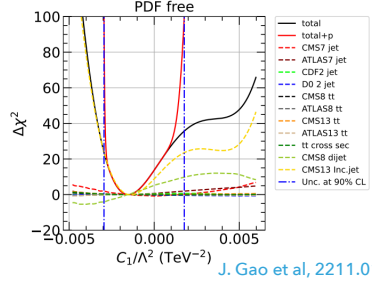
✓ Stress-tested and shown robustness with closure tests

THE TOP SECTOR

- We extend previous analysis to 20 operators in the top sector, including all available unfolded top data that constrain PDFs and/or SMEFT operators: tt~ inclusive and differential/double-differential cross sections at 5, 7, 8 and 13 TeV, tt~ asymmetries at 8 an 13 TeV, W-helicity fractions at 8 and 13 TeV, tt~V associated production at 8 TeV and 13 TeV, tt~tt~ and tb~tb~ production total cross sections, single top t-channel and s-channel production and associated single top and vector boson production.

- Improved simuNET algorithm allows to include both PDF dependent and independent measurements.

- Strategy (Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales, J. Rojo, MU) :
  - ➡ Add all available top-sector data and assess impact on PDFs
  - ➡ Add all available top-sector data and fit SMEFT coefficients
  - ➡ Add all available top-sector data in simultaneous fit of PDFs and SMEFT to assess interplay and correlations
  - ➡ Verify results via closure tests



J. Gao et al, 2211.01094

- Note that this is different from analysis by J. Gao et al both in terms of approach (individual vs marginalised, scan versus fit) and in number of operators (4 versus 20) and datasets & interesting to compare.

CAN PDFS ABSORB NEW PHYSICS?

- Imagine that the "true" UV model is some heavy W' boson with M ~ 10 TeV
- This is mapped down to non-zero Wilson coefficients in the SMEFT expansion
- Generate MC pseudodata assuming "true" law of nature
- Fit PDFs assuming SM
- Can PDFs absorb signs of new physics?

➡ Compare total chi2 of PDF fit to the chi2 once gets by fitting MC pseudo-data generated assuming SM

➡ Compare chi2 of the datasets that are mostly affected by new physics

	SM data	BSM data
χ^2	0.96833	0.95868
$< E_{trn} >$	1.958±0.053	1.944±0.051
$< E_{val} >$	1.986±0.091	1.972±0.089
$< TL >$	11900±2900	11400±3000
$< \chi^2 >$	0.989±0.011	0.979±0.010
ϕ	0.1438±0.0040	0.1437±0.0036

E. Hammou, Z. Kassabov, M. Madigan, M. Mangano, L. Mantani, J. Moore, M. Morales,, MU - in progress

CAN PDFS ABSORB NEW PHYSICS?

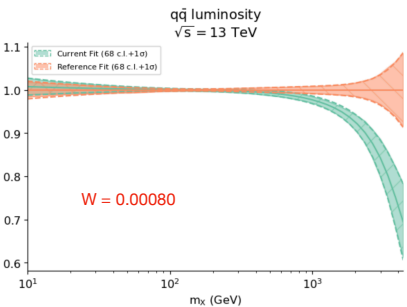
- Imagine that the "true" UV model is some heavy W' boson with M ~ 10 TeV
- This is mapped down to non-zero Wilson coefficients in the SMEFT expansion
- Generate MC pseudodata assuming "true" law of nature
- Fit PDFs assuming SM
- Can PDFs absorb signs of new physics?

➡ Compare total chi2 of PDF fit to the chi2 once gets by fitting MC pseudo-data generated assuming SM

➡ Compare chi2 of the datasets that are mostly affected by new physics

➡ If chi2 is comparable and PDFs are different then <u>PDFs can absorb new physics</u>

	SM data	BSM data
χ^2	0.96833	0.95868
$< E_{trn} >$	1.958±0.053	1.944±0.051
$< E_{val} >$	1.986±0.091	1.972±0.089
$< TL >$	11900±2900	11400±3000
$< \chi^2 >$	0.989±0.011	0.979±0.010
ϕ	0.1438±0.0040	0.1437±0.0036



E. Hammou, Z. Kassabov, M. Madigan, M. Mangano, L. Mantani, J. Moore, M. Morales,, MU - in progress

CAN PDFS ABSORB NEW PHYSICS?

- Imagine that the "true" UV model is some heavy W' boson with M ~ 10 TeV
- This is mapped down to non-zero Wilson coefficients in the SMEFT expansion
- Generate MC pseudodata assuming "true" law of nature
- Fit PDFs assuming SM
- Can PDFs absorb signs of new physics?

➡ Compare total chi2 of PDF fit to the chi2 once gets by fitting MC pseudo-data generated assuming SM

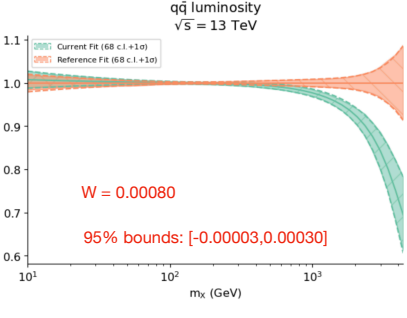
➡ Compare chi2 of the datasets that are mostly affected by new physics

➡ If chi2 is comparable and PDFs are different then <u>PDFs can absorb new physics</u>

➡ If we use contaminated PDFs in a search, are we missing new physics?

➡ How do we avoid this? Work in progress

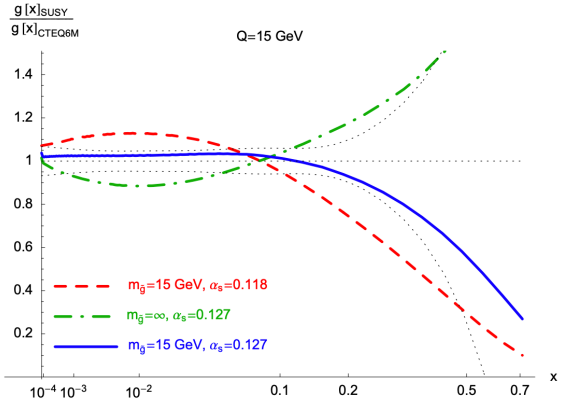
	SM data	BSM data
χ^2	0.96833	0.95868
$< E_{trn} >$	1.958±0.053	1.944±0.051
$< E_{val} >$	1.986±0.091	1.972±0.089
$< TL >$	11900±2900	11400±3000
$< \chi^2 >$	0.989±0.011	0.979±0.010
ϕ	0.1438±0.0040	0.1437±0.0036



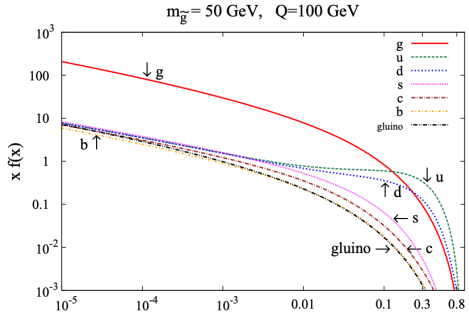
E. Hammou, Z. Kassabov, M. Madigan, M. Mangano, L. Mantani, J. Moore, M. Morales,, MU - in progress

BEYOND THE STANDARD PROTON: NEW PARTONS

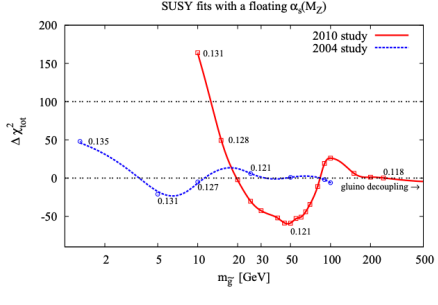
COLOURED PARTONS



Berger et al hep-ph/0406143



Berger et al 1010.4315



- Pre-LHC studies: what is there was a light SUSY coloured partner?
- A light SUSY Parton would modify DGLAP equation and running of αs
- Comparison to data excludes any light coloured parton on increasing mass range as more (and more precise) data are included in the global PDF fit

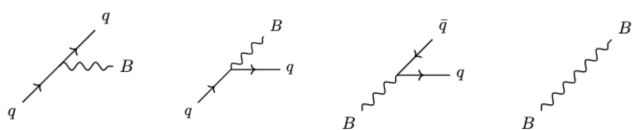
COLOURLESS PARTONS: A DARK PHOTON EXAMPLE

- Idea: now PDFs are known very precisely, and their uncertainties will continue to reduce in the near future with the HL-LHC, could we do the same for a colourless particle too?
- If there was a lepto-phobic dark photon weakly coupled to quarks via effective lagrangian

$$\mathcal{L}_{\text{int}} = \frac{1}{3} g_B \bar{q} \not{B} q \qquad m_B \in [2,80] \text{ GeV}$$

it would appear among the partons of the proton.

- To include the dark photon as a constituent of the proton: compute the dark photon splitting functions, and add them to DGLAP evolution. Starting from an appropriate initial-scale ansatz (dark photon generated dynamically off quarks and antiquarks at threshold) and a reference PDF set, evolve using the modified DGLAP equations

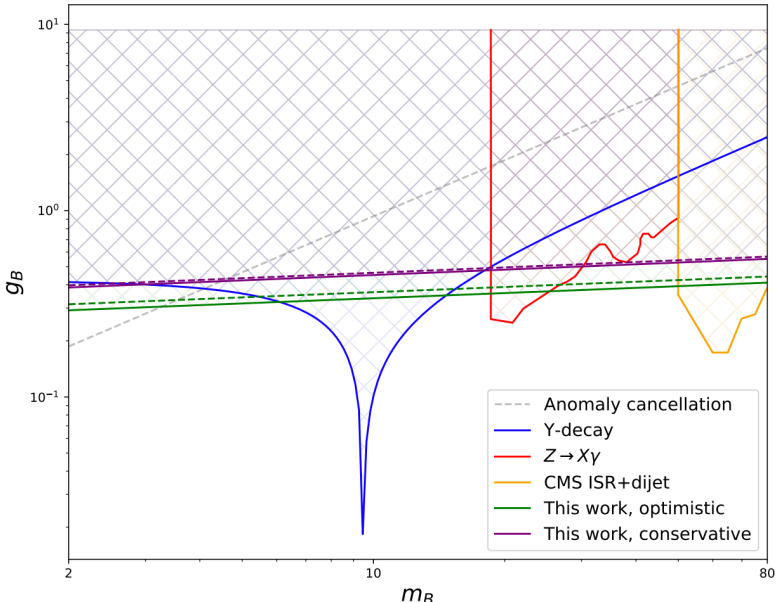


$$P_{ij} = \left(\frac{\alpha_s}{2\pi}\right) P_{ij}^{(1,0,0)} + \left(\frac{\alpha_s}{2\pi}\right)^2 P_{ij}^{(2,0,0)} + \left(\frac{\alpha_s}{2\pi}\right)^3 P_{ij}^{(3,0,0)} + \left(\frac{\alpha}{2\pi}\right) P_{ij}^{(0,1,0)} + \left(\frac{\alpha_s}{2\pi}\right)\left(\frac{\alpha}{2\pi}\right) P_{ij}^{(1,1,0)} + \left(\frac{\alpha}{2\pi}\right)^2 P_{ij}^{(0,2,0)} + \left(\frac{\alpha_B}{2\pi}\right) P_{ij}^{(0,0,1)} + \cdots ,$$

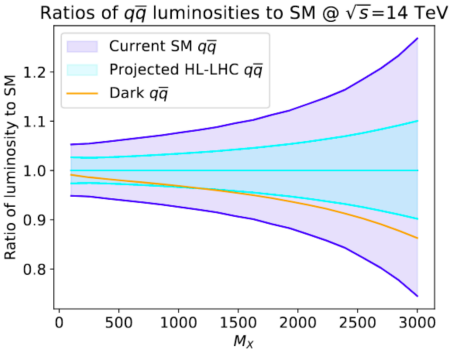
$$\alpha_B \sim 0.001$$

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

COLOURLESS PARTONS: A DARK PHOTON EXAMPLE



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



- The presence of the dark Parton would modify the evolution of standard quarks and gluon.
- Precise LHC data (DY distributions) can indirectly constrain parameter space of the dark photon in a competitive way compared to direct searches

CONCLUSIONS AND OUTLOOK

- In an era of precision, need careful assessment of PDF uncertainties
- 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.
- Another interesting interplay between PDFs and new physics due to the presence of new partons
- Lots of new exciting avenues being explored:
  - ➡ Determination of PDFs and α_s
  - ➡ Determination of PDFs and electroweak parameters
  - ➡ Determination of PDFs and SMEFT coefficients in the top sector
  - ➡ Systematic study of new physics contamination in PDF fits

EXTRA MATERIAL

PRIOR PROBABILITY IN PDF FITS

✓ PDF fitting example of inverse problem: aim to find a posterior probability of ***f*** given the data ***D***.

$$\boxed{p(f|D) \propto p(D|f)\, p(f)}$$

✓ 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 ***H.***

$$\begin{aligned} p(\theta|D,\mathcal{H}) &\propto p(D|\theta,\mathcal{H})\, p(\theta|\mathcal{H}) \\ &= \exp(-\mathcal{L}(\theta,D))\, p(\theta|\mathcal{H}) \end{aligned}$$

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

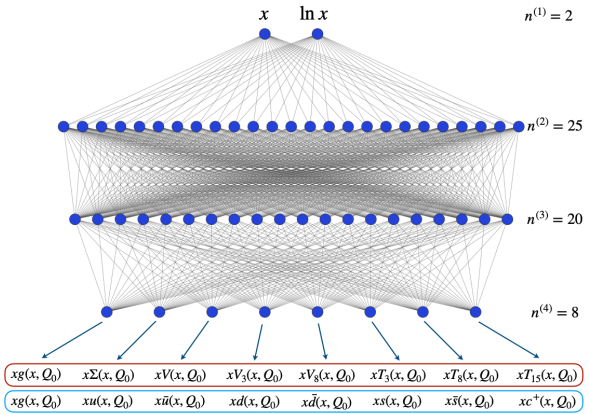
$$\boxed{\chi^2 = \frac{1}{N_{\rm dat}} \sum_{i=1}^{N_{\rm dat}} (T_i(\{\theta\},\{c\}) - D_i)\, {\rm cov}_{ij}^{-1}\, (T_j(\{\theta\},\{c\}) - D_j)}$$

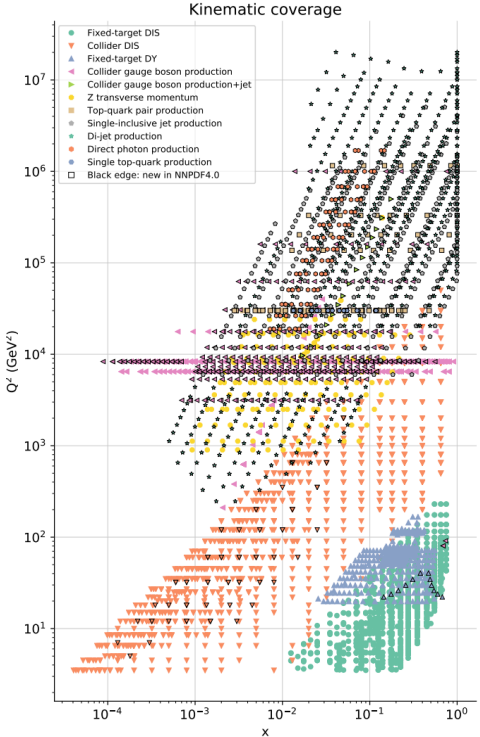
Ball et al, arXiv:0912.2276

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

A RECENT PDF FIT: NNPDF4.0

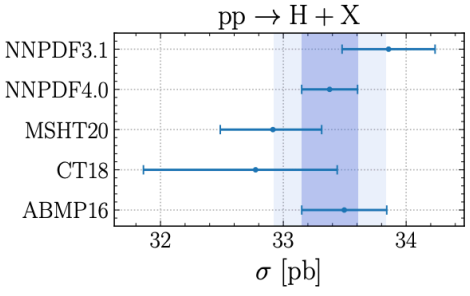
- NNPDF4.0: most recent global PDF set from NNPDF based on very large set of data from LHC: O(5000) data points.
- New methodology based on hyper-parameter optimisation validated by closure tests and future tests. [Ball et al, arXiv:2109.02653]
- Open-source code [Ball et al. arXiv:2109.02671]
- Parton luminosity uncertainties down to 1-2% in many regions

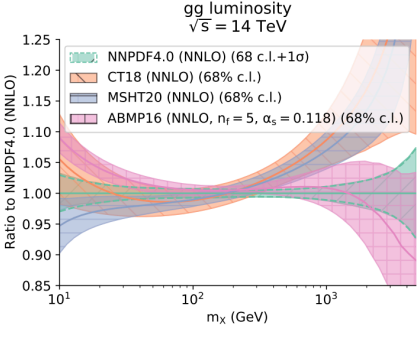


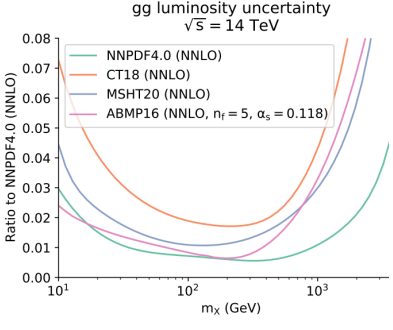


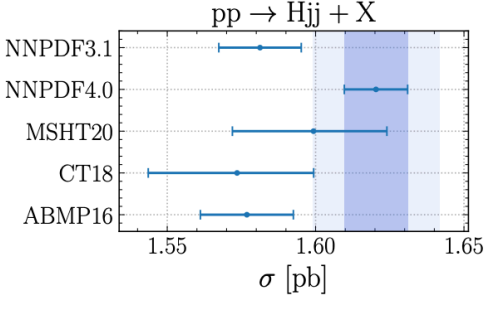
A RECENT PDF FIT: NNPDF4.0

- NNPDF4.0: most recent global PDF set from NNPDF based on very large set of data from LHC Run I: O(5000) data points.
- New methodology based on hyper-parameter optimisation validated by closure tests and future tests. [Ball et al, arXiv:2109.02653]
- Open-source code [Ball et al, arXiv:2109.02671]
- Parton luminosity uncertainties down to 1-2% in many regions.
- At such level of precision crucial to account for all th. uncertainties, starting from missing higher order uncertainty. [Abdul Khalek et al, arXiv:1906.10698]

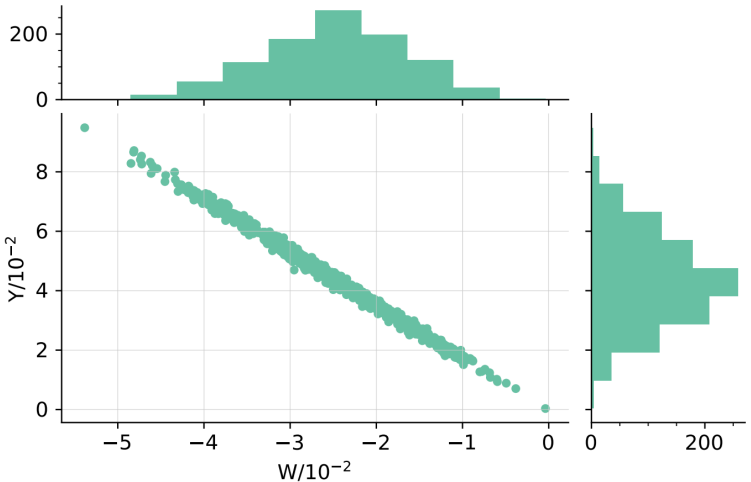








RESULTS: DRELL-YAN DATA @RUN1 AND RUN2



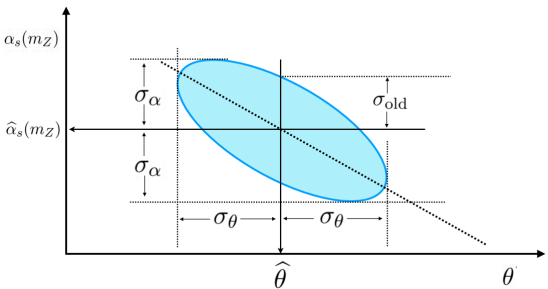
✓ Simultaneous analysis confirms results of previous study based on scan on benchmark points in the SMEFT space: with current data effect is not-negligible but small compared to PDF uncertainties

✓ Methodology able to find flat direction in W-Y parameter space

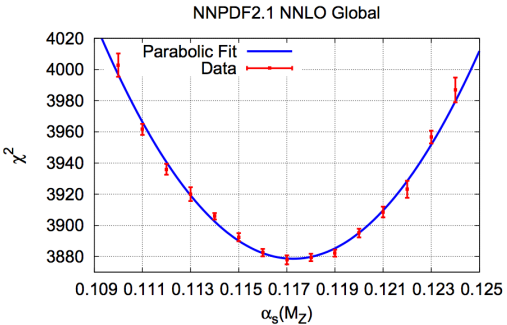
✓ To eliminate it, need Drell-Yan charged current data

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

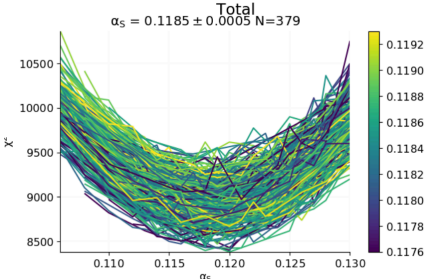


Ball et al, 1110.2483

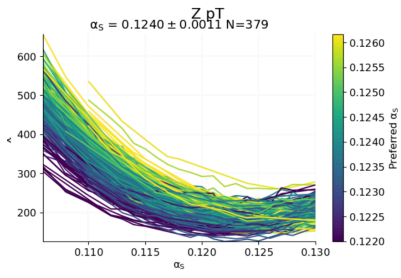
- ➡ Early determinations involve a scan over α_s and ignored PDF and α_s correlation in the fit
- ➡ Recent simultaneous determination of PDF and α_s using correlated replica method
- ➡ Many determination of α_s from analyses of specific LHC processes have been published recently (from tt~, Z and W production,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

SIMULTANEOUS FITS FOR SM PARAMETERS

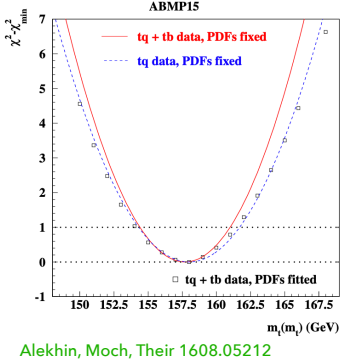
- Given the strong correlation between PDFs of the proton and α_s, a non simultaneous determination of α_s along with the PDFs from LHC processes might yield misleading results



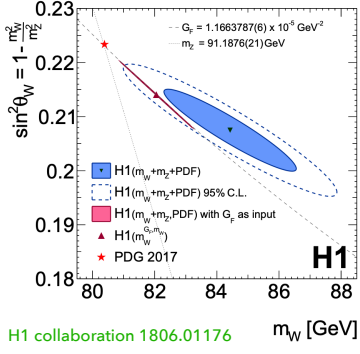
1806.01176



Forte, Kassabov 2001.04986



Alekhin, Moch, Their 1608.05212



H1 collaboration 1806.01176

- 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)