PROCEEDINGS OF SCIENCE

Can PDFs Absorb New Physics?

Maeve Madigan*

Institut für Theoretische Physik, Universität Heidelberg, Germany

E-mail: madigan@thphys.uni-heidelberg.de

The role of the Parton Distribution Functions (PDFs) is crucial not only in the precise determination of Standard Model (SM) parameters, but also in the interpretation of new physics searches at the LHC. In this talk we show the potential of global PDF analyses to inadvertently 'fit away' signs of new physics, by identifying specific scenarios in which the PDFs may completely absorb signs of new physics, thus biassing theoretical predictions. At the same time, we discuss several strategies to single out and disentangle such effects.

The European Physical Society Conference on High Energy Physics (EPS-HEP2023)
21-25 August 2023
Hamburg, Germany

*Speaker

https://pos.sissa.it/

PoS(EPS-HEP2023)479

1. Introduction

Parton distribution functions (PDFs) form a crucial input into the calculation of LHC observables. In particular, searches for new physics (NP) at the LHC often rely on a precise knowledge of the PDFs at large-x. However, PDFs are determined by fitting their x-dependence from data, a vast proportion of which consists of LHC data, see for example [1–3]. By determining the PDFs under the assumption of the SM, a potential for interplay with the effects of new physics is introduced; the PDFs have the potential to absorb or 'fit away' signs of new physics in the data, thus biassing the resulting theoretical predictions.

The interplay between PDFs and new physics has been widely studied in recent years. Several studies investigated this interplay in deep inelastic scattering (DIS) observables [4–6]. In particular, a simultaneous fit of the PDFs and Standard Model Effective Field theory (SMEFT) was performed using DIS data in [6], and extended to include high-mass Drell-Yan distributions in [7, 8]. The interplay between PDFs and new physics has been further studied in top quark data [9, 10] and jet data [11]. The level of interplay varies across these studies depending on the measurement and associated uncertainties. It was found, however, that in high-mass Drell-Yan tails there is a significant potential for interplay at the High-Luminosity LHC [7, 8]: neglecting this interplay may lead to overly-precise uncertainties on the SMEFT and on the PDFs.

In this study, we further the investigation of the interplay between PDFs and new physics by questioning the ability of a PDF fit to absorb signs of new physics. In particular, we study the behaviour of a PDF fit in the presence of new physics. We introduce two interesting beyond the Standard Model (BSM) scenarios which impact the observables entering into our PDF fit. Firstly, we investigate the consequences of performing a PDF fit in the presence of this new physics, and ask the question: would this new physics go undetected? Secondly, we analyse the impact of using such 'contaminated' PDF sets in the calculation of theoretical predictions for LHC observables. For a more detailed discussion we refer to our main work [12].

1.1 Data and methodology

We begin by briefly outlining the data and methodology used in this study, and refer to the main work, Ref. [12], for further details. The dataset considered in this study encompasses the data entering into the determination of the `NNPDF4.0` PDF set [1]. We add to this a number of high-mass Drell-Yan (DY) datasets including high-luminosity LHC projections for both neutral current DY and charged current DY, as used in Ref. [7].

The methodology closely follows the closure test methodology developed by the NNPDF collaboration [13]. The PDFs are fit to pseudodata generated under the assumption of a known underlying law,

$$D \sim \mathcal{N}(T(\theta_{\rm SM}, \theta_{\rm NP}), \Sigma)\,, \qquad (1)$$

where Σ denotes the experimental systematic and statistical uncertainties, and T denotes the underlying theory. This theory consists of an input PDF set, parametrised here by $\theta_{\rm SM}$, and a known BSM model, parametrised by $\theta_{\rm NP}$. A PDF fit to this pseudodata is then performed using the NNPDF4.0 methodology under the assumption of the SM.

The quality of this fit is assessed by studying the resulting χ^2: in particular, a χ^2 of 1 indicates excellent agreement between data and the resulting PDF. We analyse the deviation of the χ^2 from this expected value of 1, and identify two possible scenarios. Firstly, the PDF may not be able to accommodate the impact of new physics on the data. If we find a deviation of the χ^2 from 1 of more than 2 standard deviations we observe that the PDF has not absorbed signs of new physics. Alternatively, the PDF may be able to shift to compensate for the effects of new physics, resulting in a value of $\chi^2 \sim 1$: in this case we claim that new physics has been absorbed by the PDF.

2. Contaminated PDFs

2.1 New physics scenarios

Two new physics scenarios are considered as benchmarks in this study. The first is a flavour universal Z' boson charged under a $U(1)_Y$ symmetry. The second scenario consists of a flavour universal W' boson charged under $SU(2)_L$. In the effective field theory approximation, the effect of these UV models on our observables can be parametrised by the electroweak oblique parameters $\hat{Y}$ and $\hat{W}$ respectively, for example:

$$\mathcal{L}^{Z'}_{\rm SMEFT} = \mathcal{L}_{\rm SM} - \frac{g'^2 \hat{Y}}{2m_W^2} J_Y^{\mu} J_{Y,\mu}, \qquad \hat{Y} = \frac{g_{Z'}^2}{M_{Z'}^2} \frac{m_W^2}{g'^2}\,, \tag{2}$$

where we note that the $\hat{Y}$ parameter is inversely proportional to the Z' mass $m_{Z'}$. The dominant impact of both Z' and W' scenarios is on Drell-Yan at high invariant mass. However, while the neutral Z' boson has an effect only on neutral current DY, the W' model impacts both neutral and charged current DY. As we will see in the next subsection, this difference has a significant effect on the results of the PDF fits.

2.2 PDF fit results

PDF fits are performed to pseudodata generated under the assumption of new physics at benchmark values of $\hat{Y} = 5, 15, 25 \times 10^{-5}$ and $\hat{W} = 3, 8, 15 \times 10^{-5}$ in the Z' and W' scenarios respectively. We observe that for values of $\hat{Y}$ with non-negligible impact on the observables in our fit, the Z' scenario leads to a poor fit quality and relatively large deviations in the χ^2. This indicates that, if such a Z' model were present in the data, this would not go undetected in a PDF fit; this scenario would instead lead to a noticeably large χ^2. Conversely, the W' scenario leads to a value of χ^2 close to 1, even in the case of a W' of mass $m_{W'} \approx 14$ TeV parametrised by $\hat{W} = 8 \times 10^{-5}$. In this scenario, we find excellent agreement between the NP-modified pseudodata and the PDF, indicating that such a W' could be absorbed by the PDFs.

The origin of the difference between these two scenarios is that while the Z' impacts only neutral current DY, the W' impacts both neutral and charged current DY. In Figure 1 we show the resulting u, $\bar{u}$, d and $\bar{d}$ PDFs convolved to produce the luminosity which enters into the calculation of charged current DY. In both plots, the green PDF indicates the baseline of our study: a PDF determined from SM pseudodata. We compare this baseline to the PDFs determined from the Z' and W' scenarios. We can see that in the case of the Z' scenario, the charged current DY distribution constrains the PDFs to remain aligned with the baseline and therefore consistent with the SM, due to the lack of new physics in this observable. This limits the flexibility of the PDF fit to absorb new physics in neutral current DY. In the W' scenario, however, the PDFs shift significantly in order to

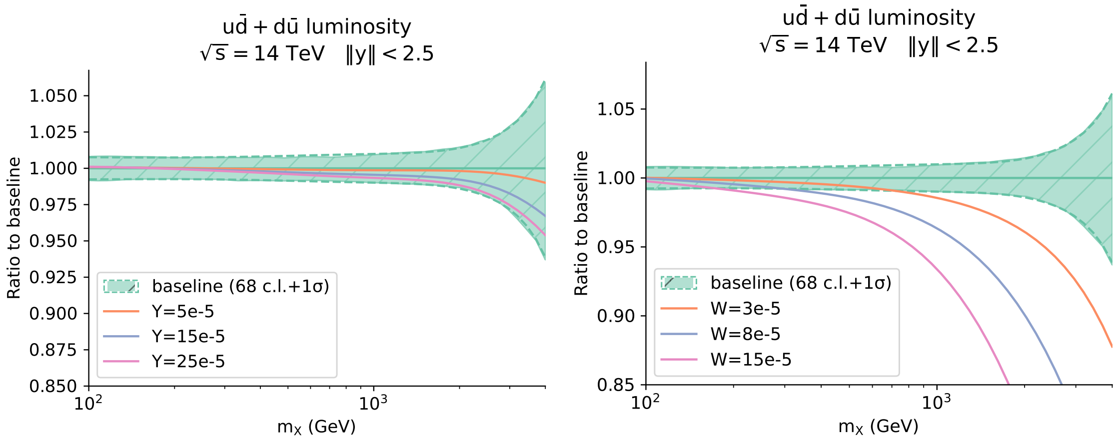

Figure 1: Contaminated versus baseline luminosity at $\sqrt{s}$ = 14 TeV in the central rapidity region. The results are normalised to the baseline SM PDF. Contaminated PDFs are obtained for benchmark values of $\hat{Y} = 5, 15, 25 \times 10^{-5}$ (left) and $\hat{Y} = 3, 8, 15 \times 10^{-5}$ (right).

compensate the effect of the W' on both neutral and charged current categories of high mass DY distributions.

Figure 2 demonstrates the result of using the PDFs obtained in the Z' and W' scenarios to calculate the high-mass neutral and charged current DY distributions that entered into the PDF fit. The red curve indicates the prediction and associated PDF uncertainties from a PDF fit to the W' scenario, assuming a value of $\hat{W} = 8 \times 10^{-5}$. This is compared to the W'-modified pseudodata, as shown in blue. Similarly, the purple curve indicates the predictions obtained from a PDF fit to the Z' scenario, assuming a value of $\hat{Y} = 15 \times 10^{-5}$, and is compared to pseudodata generated under the same $\hat{Y}$ assumption shown in green. In the middle panel we clearly see the agreement between the W'-modified pseudodata and the resulting PDF: the PDF compensates the effect of new physics in the data, and therefore signs of new physics are completely missed.

3. Consequences of new physics contamination

3.1 Impact on electroweak processes

The absorption of new physics into PDFs has a significant impact on the search for new physics. As we saw in Fig. 2, the presence of new physics in the W' scenario is absorbed by the PDF, and consequently signs of new physics in this scenario would be missed. However, the impact of this new physics absorption is not limited to the high-mass Drell-Yan distributions entering into the PDF fit. In Fig. 3 (left) we demonstrate the effect of using the W'-contaminated PDF set, assuming $\hat{W} = 8 \times 10^{-5}$, in calculating theory predictions for diboson production $pp \to W^+W^-$ at the High-Luminosity LHC assuming $\sqrt{s}$ = 14 TeV and $\mathcal{L}$ = 3 ab^{-1}. This high-mass kinematic distribution is often used as a probe of new physics, and the corresponding theory predictions rely on knowledge of the large-x quark and antiquark PDFs.

In Fig. 3 (left) we observe a deviation between data and theory with a statistical significance of over 3 standard deviations. Note that W' model has no impact on the partonic cross section of

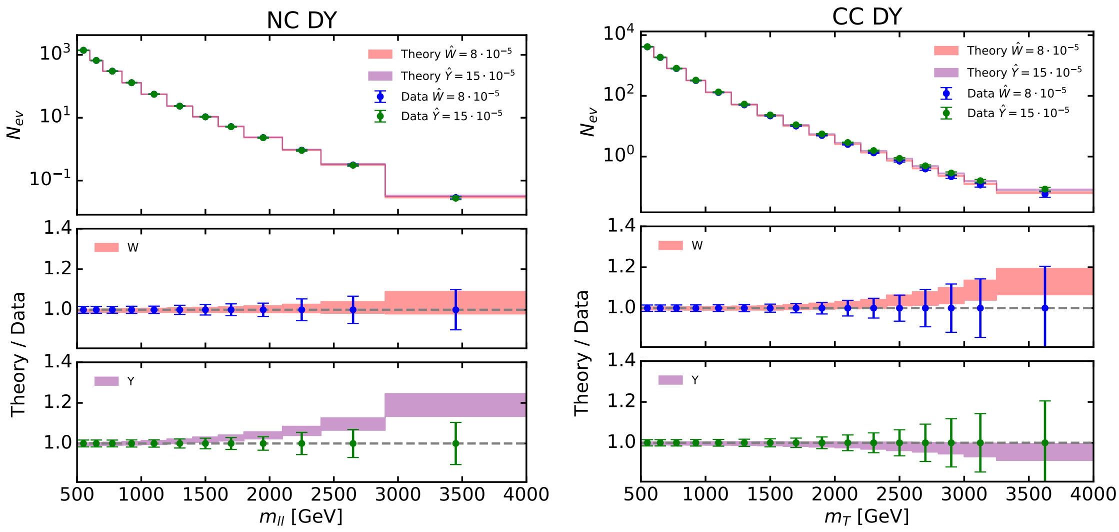

Figure 2: Comparison between the pseudodata calculated in the presence of new physics in the W' and Y' scenarios, and SM theory predictions obtained with the resulting contaminated PDFs . Left panel: neutral current DY $m_{\ell\ell}$ distribution. Right panel: charged current DY m_T distribution.

this process at leading order, and as a result, the pseudodata is generated under the assumption of the SM, while the 'theory' curve is obtained by convolving the W'-contaminated PDF set with SM predictions. This deviation indicates the potential for the use of a contaminated PDF set to lead to spurious signs of new physics in datasets which are, in fact, not affected by new physics at all.

3.2 Opportunities to disentangle PDFs from new physics effects

Before concluding, we briefly discuss opportunities to disentangle the effect of new physics from PDF fits. Ratio observables, see example Ref. [14], provide sensitivity to new physics while reducing the PDF-dependence. Figure 3 (right) shows the ratio of the diboson W^+W^- invariant mass distribution to the neutral current DY invariant mass distribution, both of which are sensitive to the $q\bar{q}$ luminosity. We observe the clear presence of new physics in this ratio. Such ratio observables provide a clear PDF-independent assessment of the presence of new physics in the dataset, although as a result, it is not clear whether the new physics affects DY or diboson production or both processes. Future PDF fits may also benefit from the addition of precise datasets at low-energy. Such measurements would be less sensitive to new physics than the high-mass DY distributions considered here, and could potentially constrain the large-x $\bar{u}$ and $\bar{d}$ PDFs such that their flexibility to shift in the presence of new physics impacting observables at the LHC and HL-LHC is limited. In this endeavor, future experiments such as the EIC programme [15, 16] will be highly valuable.

4. Conclusions

To conclude, we have demonstrated the ability of parton distribution functions to absorb signs of new physics. In the presence of new physics in the form of a flavour universal W' boson, we find that the PDFs shift to compensate the effects of new physics in HL-LHC projections for high-mass

PoS(EPS-HEP2023)479

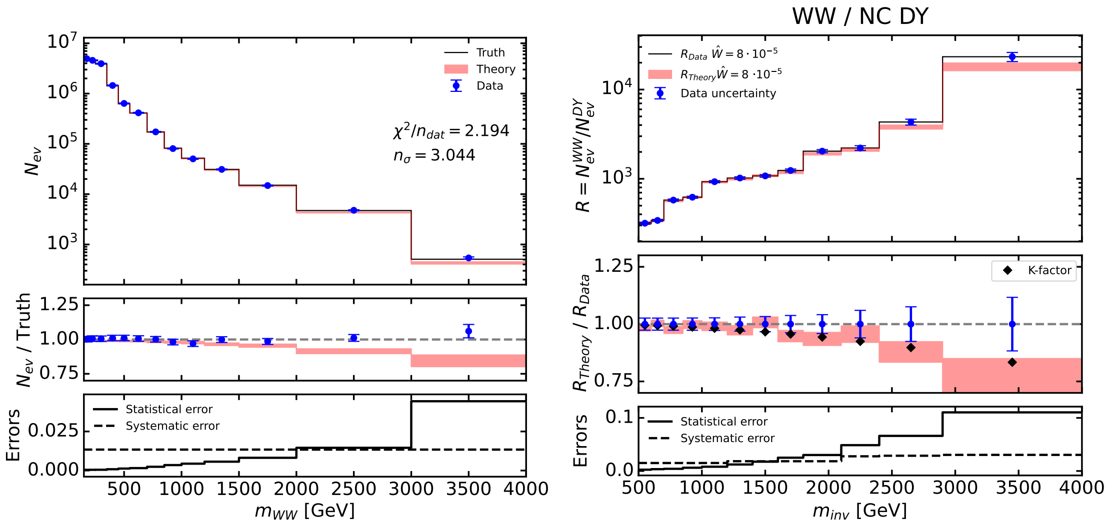

Figure 3: Left: comparison between HL-LHC projections for diboson W^+W^- production and the corresponding theory predictions calculated using the W'-contaminated PDF set. Right: ratio of this diboson W^+W^- production invariant mass distribution to the neutral current DY distribution shown in Fig. 2.

.

Drell-Yan distributions. This 'contamination' of the PDFs by new physics has the potential to hinder the discovery of new physics at the HL-LHC, as signs of new physics in high-mass DY distributions are undetected as a result of their absorption into the PDFs. Furthermore, the shift made by the PDF in order to compensate for this new physics leads to the presence of spurious new physics effects in other datasets, for example in the diboson invariant mass distribution. Opportunities to disentangle this interplay between PDFs and new physics include the use of ratio observables, as well as future precision measurements at low-energy which may improve constraints on the large-x PDFs while exhibiting low sensitivity to new physics. Finally, the strategy presented here may be applied to assess the level of contamination of the PDFs due to other BSM scenarios. The tools used in this work are made public alongside detailed instructions at the following link: `https://www.pbsp.org.uk/contamination/`.

References

[1] **NNPDF** Collaboration, R. D. Ball et al., *The path to proton structure at 1% accuracy*, *Eur. Phys. J. C* **82** (2022), no. 5 428, [arXiv:2109.02653].

[2] S. Bailey, T. Cridge, L. A. Harland-Lang, A. D. Martin, and R. S. Thorne, *Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs*, arXiv:2012.04684.

[3] T.-J. Hou et al., *New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC*, *Phys. Rev. D* **103** (2021), no. 1 014013, [arXiv:1912.10053].

[4] **ZEUS** Collaboration, H. Abramowicz et al., *Limits on contact interactions and leptoquarks at HERA*, *Phys. Rev. D* **99** (2019), no. 9 092006, [arXiv:1902.03048].

[5] **H1** Collaboration, F. D. Aaron et al., *Search for Contact Interactions in $e^{\pm}p$ Collisions at HERA*, *Phys. Lett. B* **705** (2011) 52–58, [arXiv:1107.2478].

[6] S. Carrazza, C. Degrande, S. Iranipour, J. Rojo, and M. Ubiali, *Can New Physics hide inside the proton?*, arXiv:1905.05215.

[7] A. Greljo, S. Iranipour, Z. Kassabov, M. Madigan, J. Moore, J. Rojo, M. Ubiali, and C. Voisey, *Parton distributions in the SMEFT from high-energy Drell-Yan tails*, *JHEP* **07** (2021) 122, [arXiv:2104.02723].

[8] S. Iranipour and M. Ubiali, *A new generation of simultaneous fits to LHC data using deep learning*, *JHEP* **05** (2022) 032, [arXiv:2201.07240].

[9] Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. M. Alvarado, J. Rojo, and M. Ubiali, *The top quark legacy of the LHC Run II for PDF and SMEFT analyses*, arXiv:2303.06159.

[10] J. Gao, M. Gao, T. J. Hobbs, D. Liu, and X. Shen, *Simultaneous CTEQ-TEA extraction of PDFs and SMEFT parameters from jet and $t\bar{t}$ data*, arXiv:2211.01094.

[11] **CMS** Collaboration, A. Tumasyan et al., *Measurement and QCD analysis of double-differential inclusive jet cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **02** (2022) 142, [arXiv:2111.10431]. [Addendum: JHEP 12, 035 (2022)].

[12] E. Hammou, Z. Kassabov, M. Madigan, M. L. Mangano, L. Mantani, J. Moore, M. M. Alvarado, and M. Ubiali, *Hide and seek: how PDFs can conceal New Physics*, arXiv:2307.10370.

[13] **NNPDF** Collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040, [arXiv:1410.8849].

[14] M. L. Mangano and J. Rojo, *Cross Section Ratios between different CM energies at the LHC: opportunities for precision measurements and BSM sensitivity*, *JHEP* **08** (2012) 010, [arXiv:1206.3557].

[15] R. A. Khalek, J. J. Ethier, E. R. Nocera, and J. Rojo, *Self-consistent determination of proton and nuclear PDFs at the Electron Ion Collider*, *Phys. Rev. D* **103** (2021), no. 9 096005, [arXiv:2102.00018].

[16] R. Abir et al., *The case for an EIC Theory Alliance: Theoretical Challenges of the EIC*, arXiv:2305.14572.

PoS(EPS-HEP2023)479