Parton distributions in the SMEFT, and beyond

James Moore, University of Cambridge
(on behalf of the PBSP group)

PBSP

European Research Council
Established by the European Commission

PBSP: Physics Beyond the Standard Proton

- The **PBSP group** is based at the **University of Cambridge**, and is headed by **Maria Ubiali**; the project is **ERC-funded.**
- The aim is to **investigate interplay between BSM physics and proton structure** - the subject of the rest of this talk!
- The team members are:
  - *Postdocs:* Zahari Kassabov, Maeve Madigan, Luca Mantani
  - *PhD students:* Shayan Iranipour (*former*), Elie Hammou, **James Moore**, Manuel Morales, Cameron Voisey (*former*)

European Research Council
Established by the European Commission

Introduction:
Fitting PDFs and physical parameters

Fitting PDFs and physical parameters

- **Theory predictions** for collider experiments are obtained from the standard **factorisation formula**; schematically, we have:

$$T(c, \theta) = \mathrm{FK}(c) \otimes \mathrm{PDF}(\theta)$$

observable

PDF evolution kernel + partonic cross-section

initial-scale PDF contribution

- Predictions are functions of:

(i) **'physics' parameters** c, e.g. $\alpha_S(m_Z)$, m_W, **Wilson coefficients** if we use the **SMEFT**;

(ii) **PDF parameters** θ, e.g. the weights of a neural network parametrising the initial-scale PDFs (in the NNPDF framework).

Fitting PDFs and physical parameters

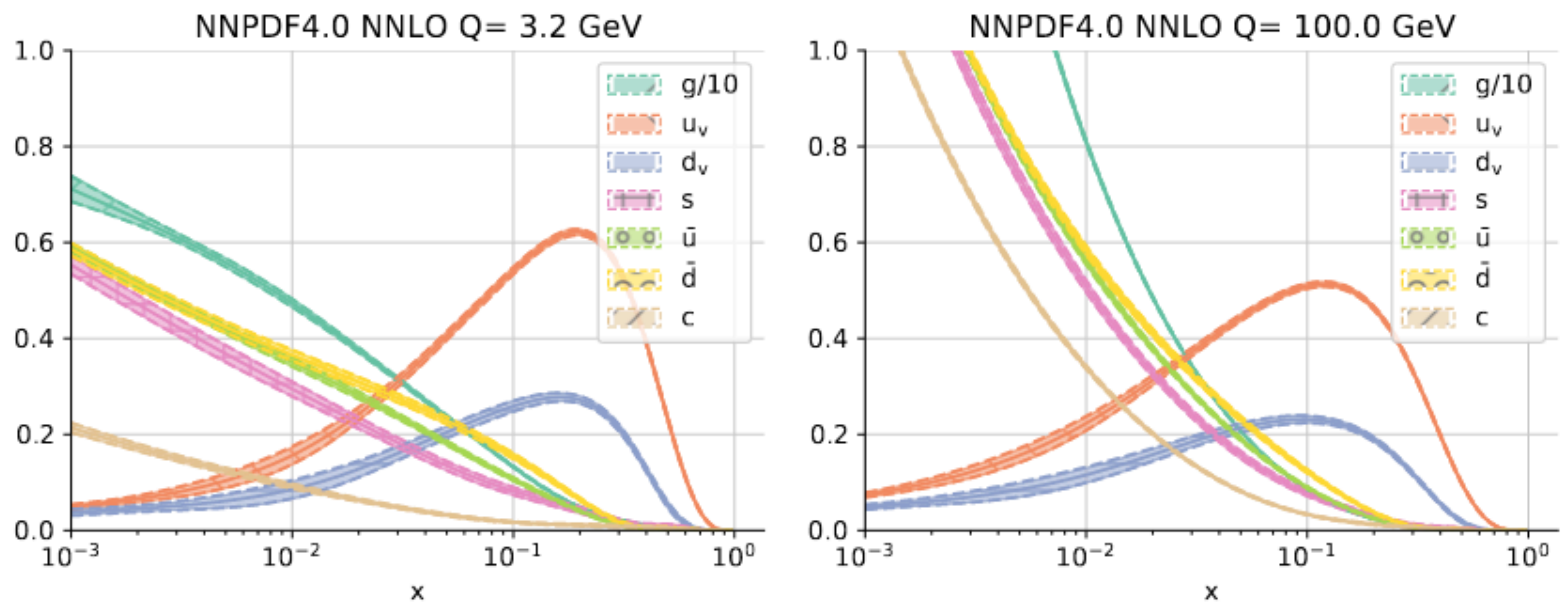
- Typically, the ‘physics’ parameter fits and PDF parameter fits **don’t talk.**

PDF parameter fits

- Fix physics parameters $c = \bar{c}$:

$$T(\bar{c}, \theta) = \text{FK}(\bar{c}) \otimes \text{PDF}(\theta)$$

- Optimal PDF parameters θ^* then have an **implicit dependence** on initial physics parameter choice: $\text{PDF}(\theta^*) \equiv \text{PDF}(\theta^*(\bar{c}))$.
- E.g. NNPDF4.0 fit, Ball et al., 2109.02653.

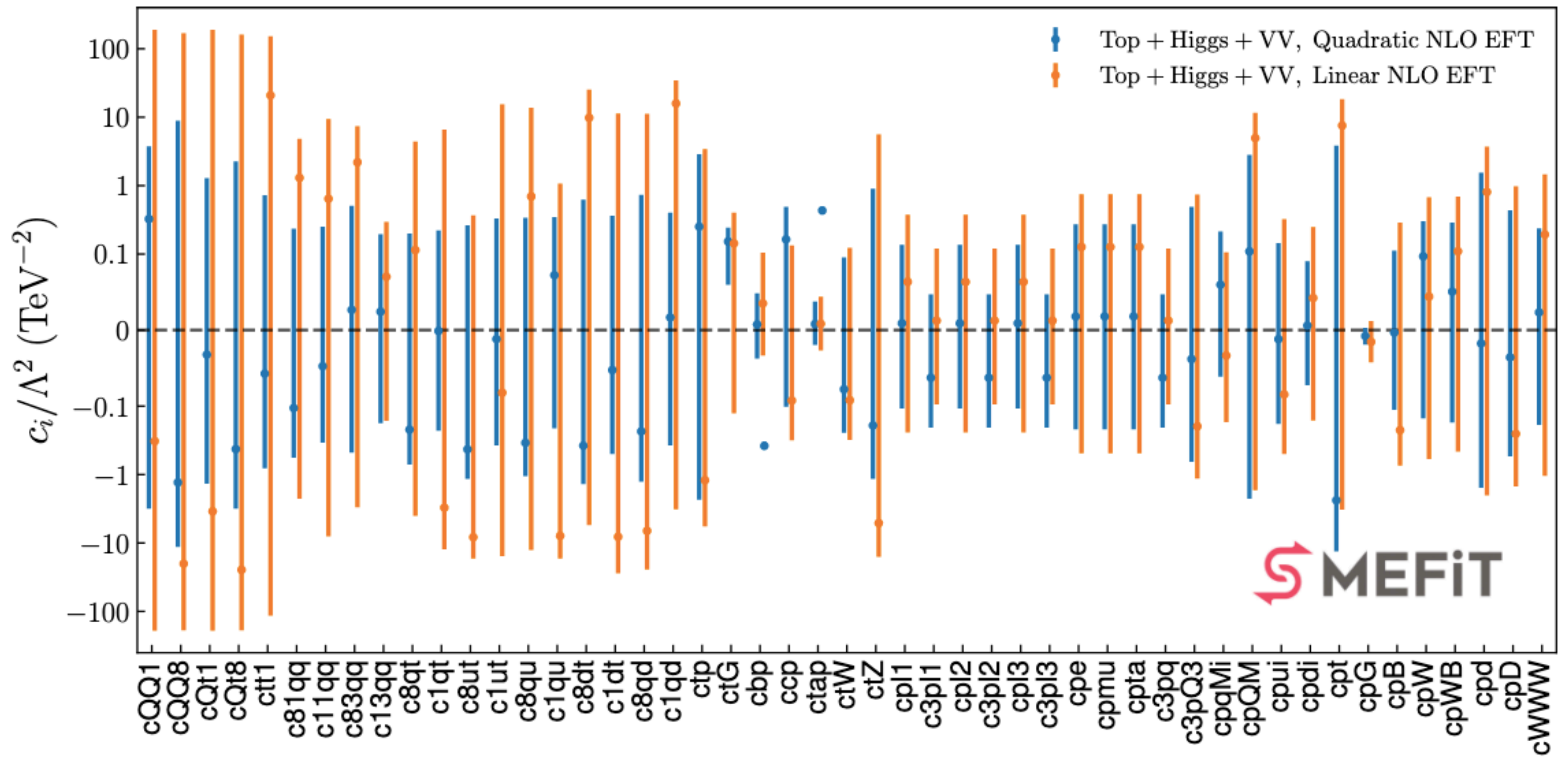


‘Physics’ parameter fits

- Fix PDF parameters $\theta = \bar{\theta}$:

$$T(c, \bar{\theta}) = \text{FK}(c) \otimes \text{PDF}(\bar{\theta})$$

- Optimal ‘physics’ parameters c^* then have an **implicit dependence** on PDF choice: $c^* = c^*(\bar{\theta})$.
- E.g. SMEFiT, Ethier et al., 2105.00006.



Fitting PDFs and physical parameters

- **This could lead to inconsistencies.**

PDF parameter fits

$$\text{PDF}(\theta^*) \equiv \text{PDF}(\theta^*(\bar{c}))$$

- Fitted PDFs can depend implicitly on fixed physical parameters used in the fit.

'Physics' parameter fits

$$c^* \equiv c^*(\bar{\theta})$$

- Bounds on physical parameters can depend implicitly on the fixed PDF set used in the fit.

- For example, if we fit PDFs **assuming all Wilson coefficients in the SMEFT are zero**, but then **use those PDFs in a fit of SMEFT Wilson coefficients**, our resulting bounds **might be misleading**. The same applies to SM parameters.
- In the case of SMEFT, we could even **miss New Physics**, or **see New Physics that isn't really there**!

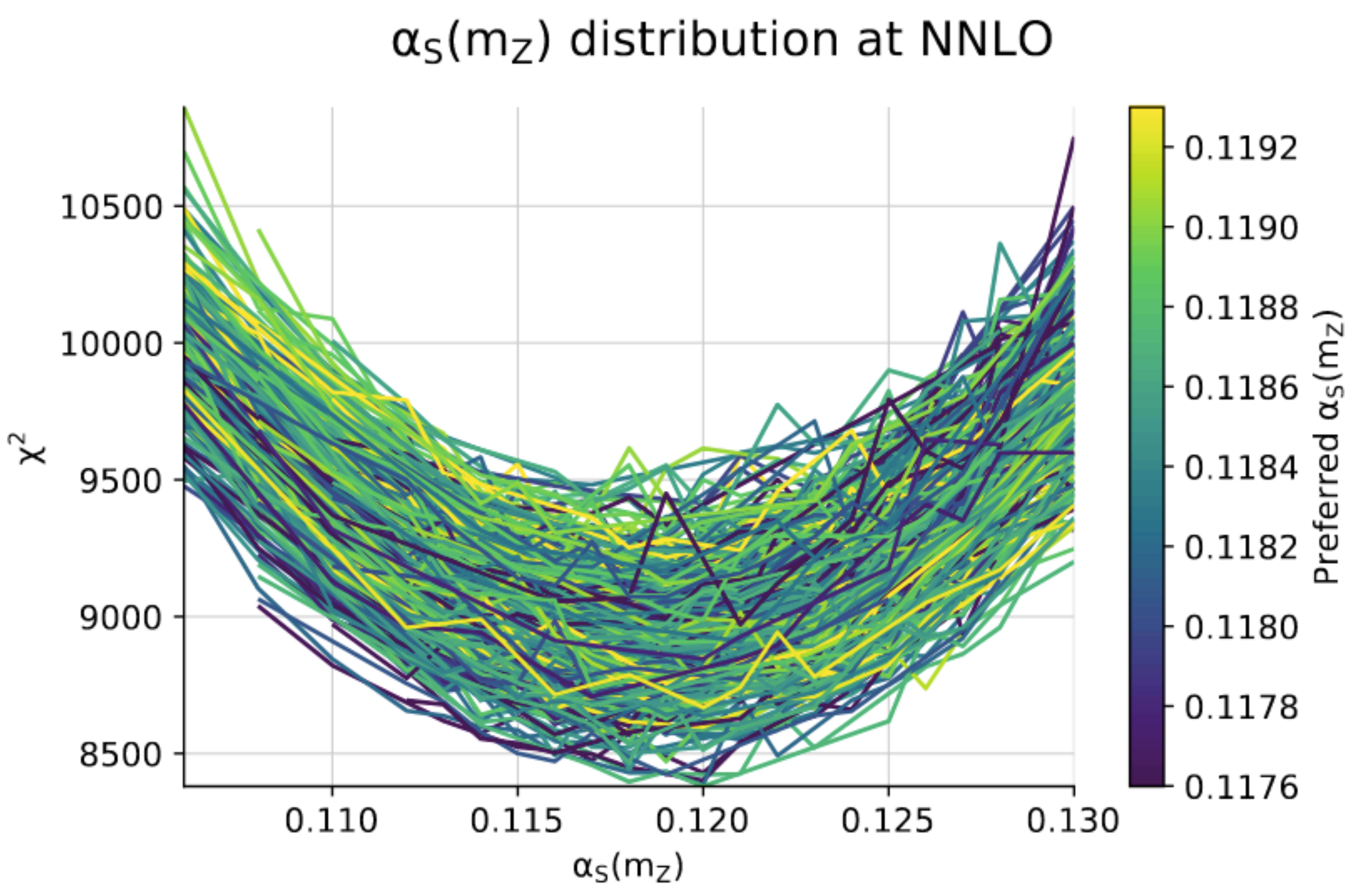
Key question for this talk:

To what extent do bounds on physical parameters change if they are fitted simultaneously with PDF parameters? Is a consistent treatment important?

Simultaneous fits of SM parameters, SMEFT parameters and PDFs

Simultaneous SM fits

- **This is not a new problem!** It's been known for a while that **simultaneous fits** of **SM parameters** alongside PDFs can be **important** in many cases. In particular, PDF parameters have a **strong correlation** with the value of $\alpha_S(m_Z)$ (see Forte, Kassabov, 2001.04986).



- The standard method for simultaneous extraction of $\alpha_S(m_Z)$ and PDFs is the **correlated replica method**, 1802.03398. In a nutshell:

  1. A grid of benchmark $\alpha_S(m_Z)$ points is selected.
  2. A **PDF fit** is performed at each benchmark point, with $\alpha_S(m_Z)$ set to the appropriate value for both **PDF evolution** and **convolution with the partonic cross-section**. The PDF replicas are correlated appropriately so as to be comparable for different values of $\alpha_S(m_Z)$.
  3. χ^2 parabolas for each set of correlated replicas are produced, and hence bounds on $\alpha_S(m_Z)$ are found.

Simultaneous SMEFT fits

- More recently, however, it has been shown that there can be a **non-negligible** interplay between **PDFs** and **Wilson coefficients in the SMEFT**.
- There are **four main works** in this direction:

1. **Carrazza et al., 1905.05215**. *Can New Physics Hide Inside the Proton?*

   A proof-of-concept study, performing a simultaneous extraction of 4 four-fermion SMEFT operators together with PDFs, using DIS-only data.

2. **Liu, Sun, Gao, 2201.06586**. *Machine learning of log-likelihood functions in global analysis of parton distributions.*

   A methodological study; simultaneous SMEFT/PDF extraction is noted as a possible application, and one SMEFT four-fermion operator is fitted using DIS-only data.

3. **PBSP team + Greljo and Rojo, 2104.02723**. *Parton distributions in the SMEFT from high-energy Drell-Yan tails.*

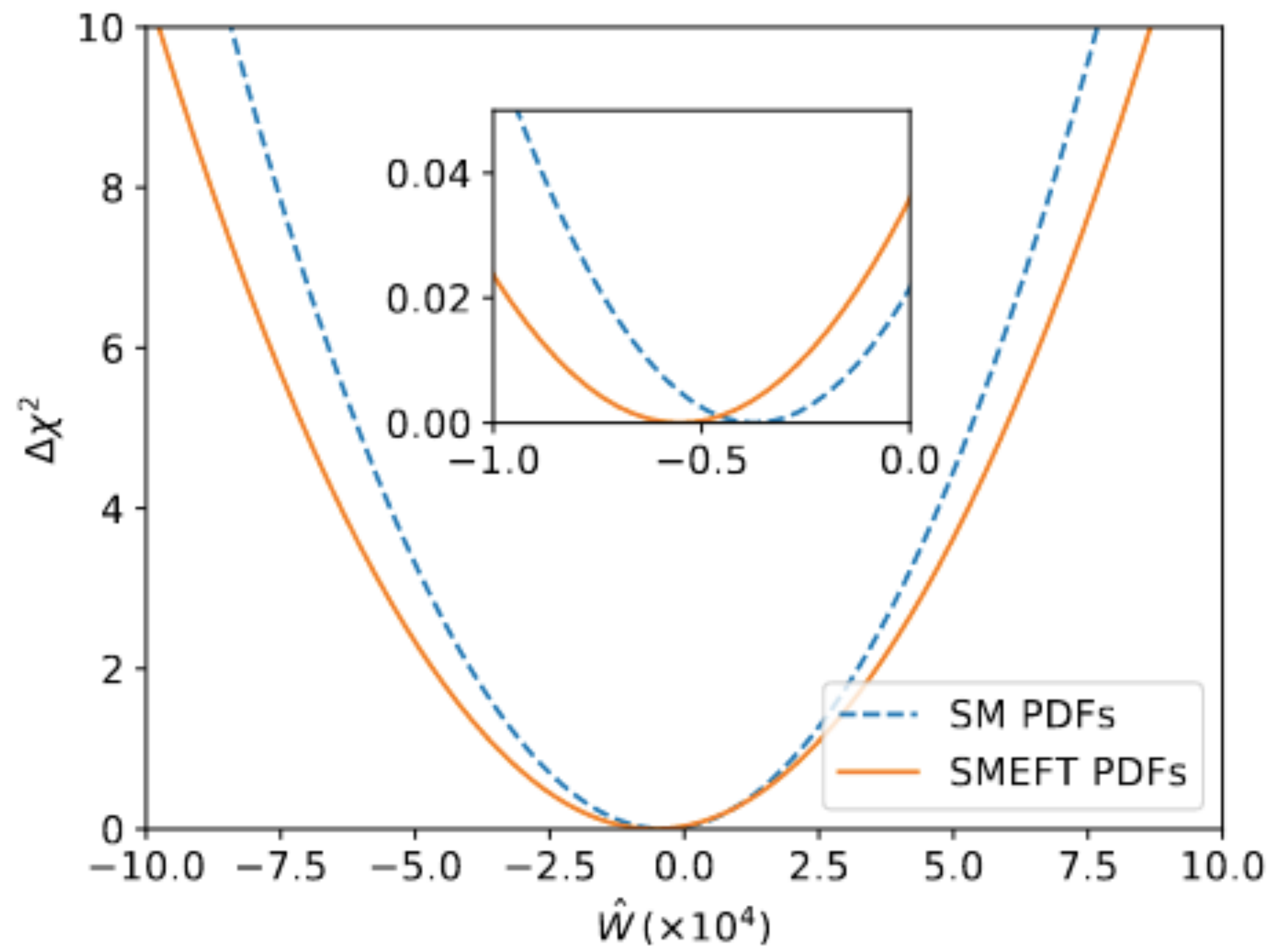
   A phenomenological study, demonstrating the impact of a simultaneous SMEFT/PDF fit in the context of the oblique W, Y parameters using current and projected Drell-Yan data.

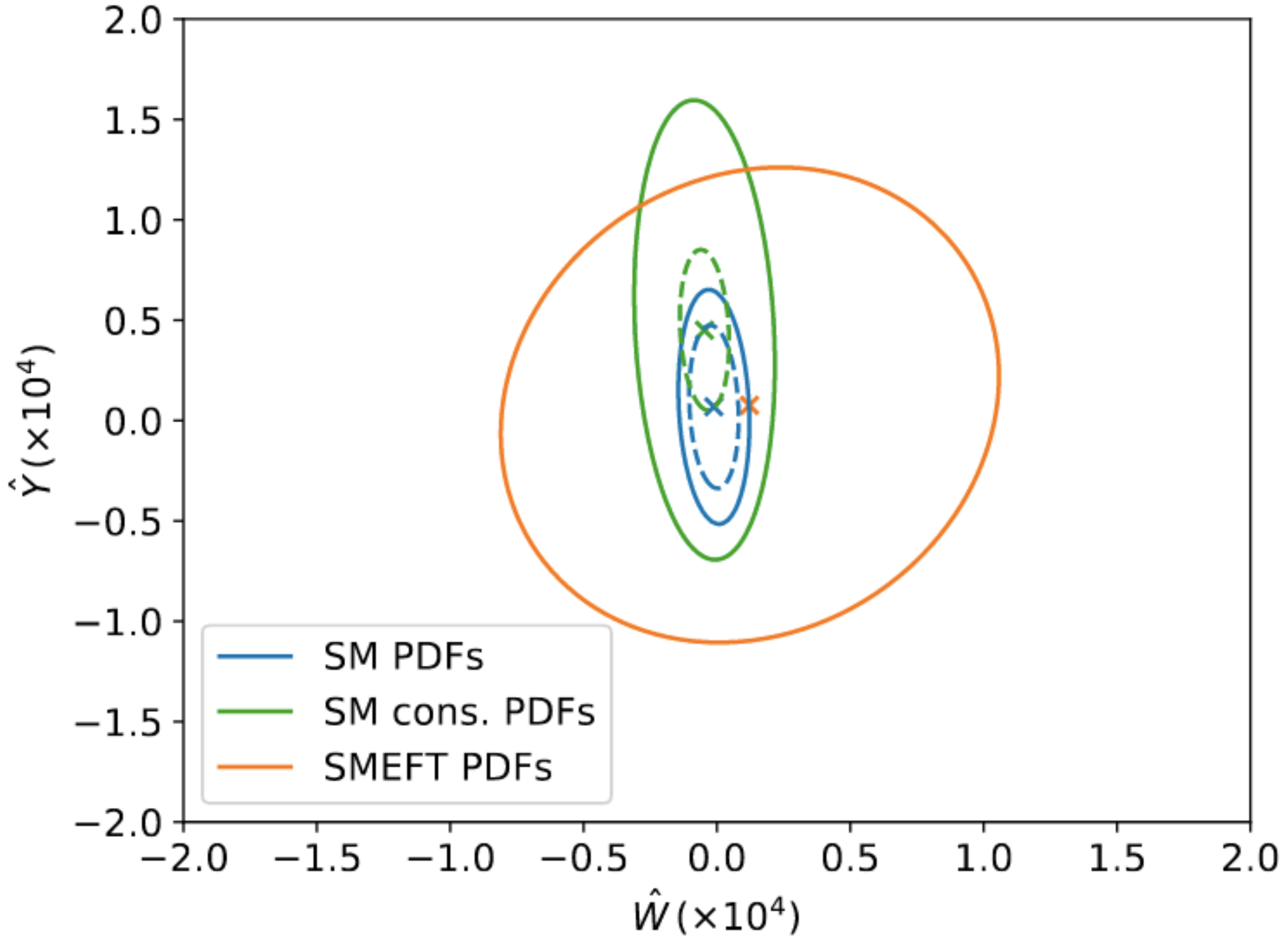
4. **CMS, 2111.10431**. *Measurement and QCD analysis of double-differential inclusive jet cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV.*

   A proof-of-concept study in the SMEFT case, involving a simultaneous extraction of PDFs, $\alpha_S(m_Z)$, the top pole mass and one SMEFT Wilson coefficient.

Parton distributions in the SMEFT from high-energy Drell-Yan tails

- In particular, in the paper 2104.02723 from the PBSP team (+ Greljo, Rojo), we find that in the context of the **oblique** W, Y **parameters**, a simultaneous fit of PDFs and the SMEFT parameters using **current data** has a **small impact on the bounds**.
- Furthermore, when we use **projected HL-LHC data**, the impact of a simultaneous fit versus a fixed PDF fit becomes **enormous**!
- The methodology used is similar to the **'scan' methodology** described for the $\alpha_S(m_Z)$ fit, but replicas are not correlated, we simply take the χ^2 of a PDF fit at each **benchmark point** in Wilson coefficient space to **construct bounds**.





The ‘SimuNET’ methodology

The need for *fast* simultaneous fits

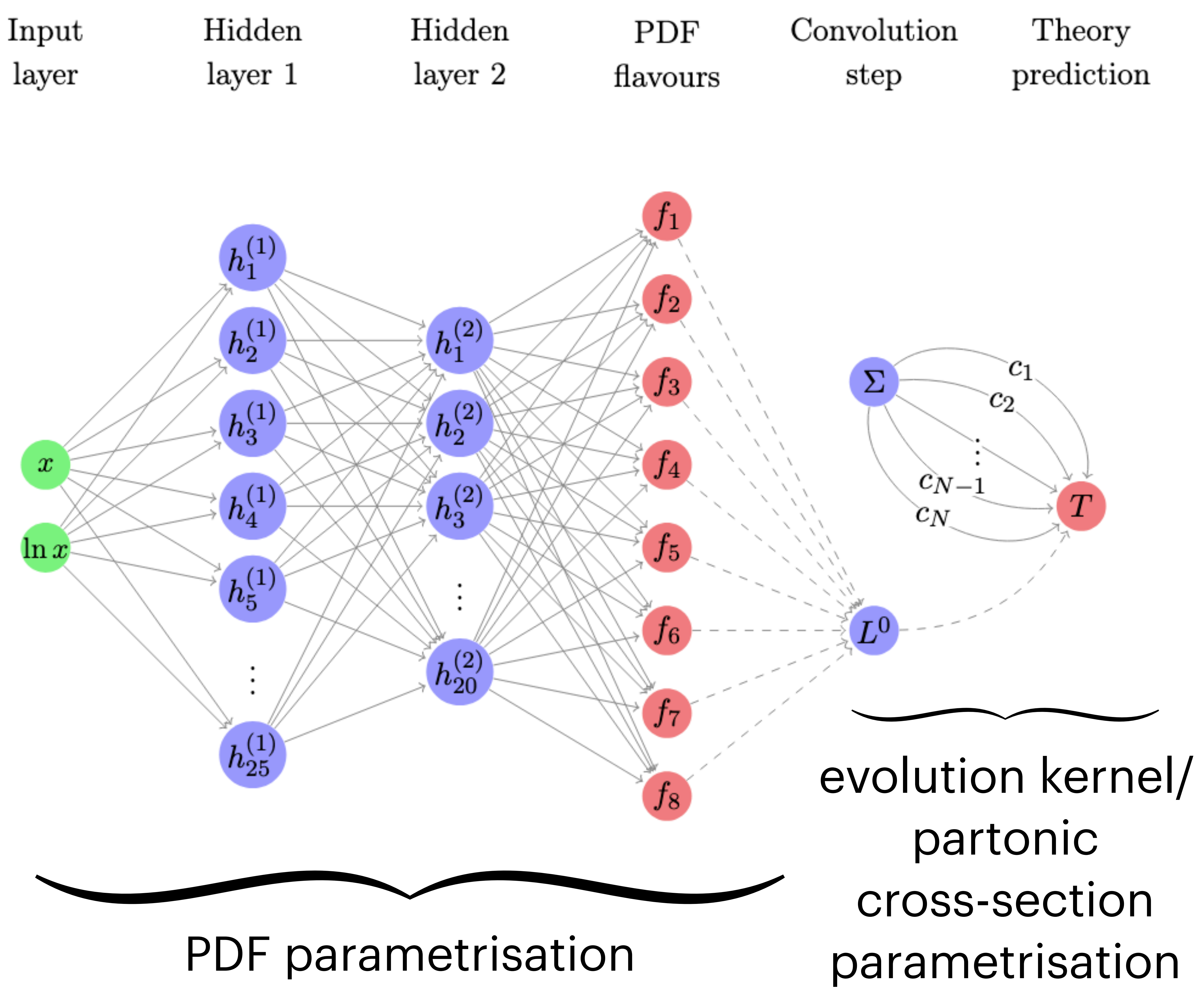
- The 'scan' methodology used for simultaneous fits in the work 2104.02723 becomes **exponentially slower** as more physical parameters are added to the simultaneous fit.

- Hence, we need a **new method** which will **scale well**. One suggestion is given in Liu, Sun, Gao, 2201.06586.

- Two members of the PBSP group have developed another approach based directly on the NNPDF4.0 PDF-fitting framework, which we call the **SimuNET methodology**, presented in Iranipour, Ubiali, 2201.07240.

Fast simultaneous fits through SimuNET

- In the NNPDF4.0 framework, PDFs are modelled by **neural networks**. The neural network PDFs are convolved with the **PDF evolution kernel** and the **partonic cross-section** to produce theory predictions, which are compared to data.
- The idea of **SimuNET** is to **add the convolution step** to the neural network itself, with the physical parameters added as **weights** of **neural network edges**.
- In principle any (non-linear) **polynomial** dependence on physical parameters can be captured through the use of **non-trainable edges**.
- An **arbitrary number** of physical parameters can be fitted at basically **no extra cost**!



SMEFT and beyond through SimuNET

- In the case of the SMEFT, we can use a **K-factor approximation** to write:

$$T_{\text{SMEFT}}(c,\theta) = \left(1 + cK_{\text{SMEFT,lin}} + c^2K_{\text{SMEFT,quad}} + \cdots\right) T_{\text{SM}}(\theta)\,.$$

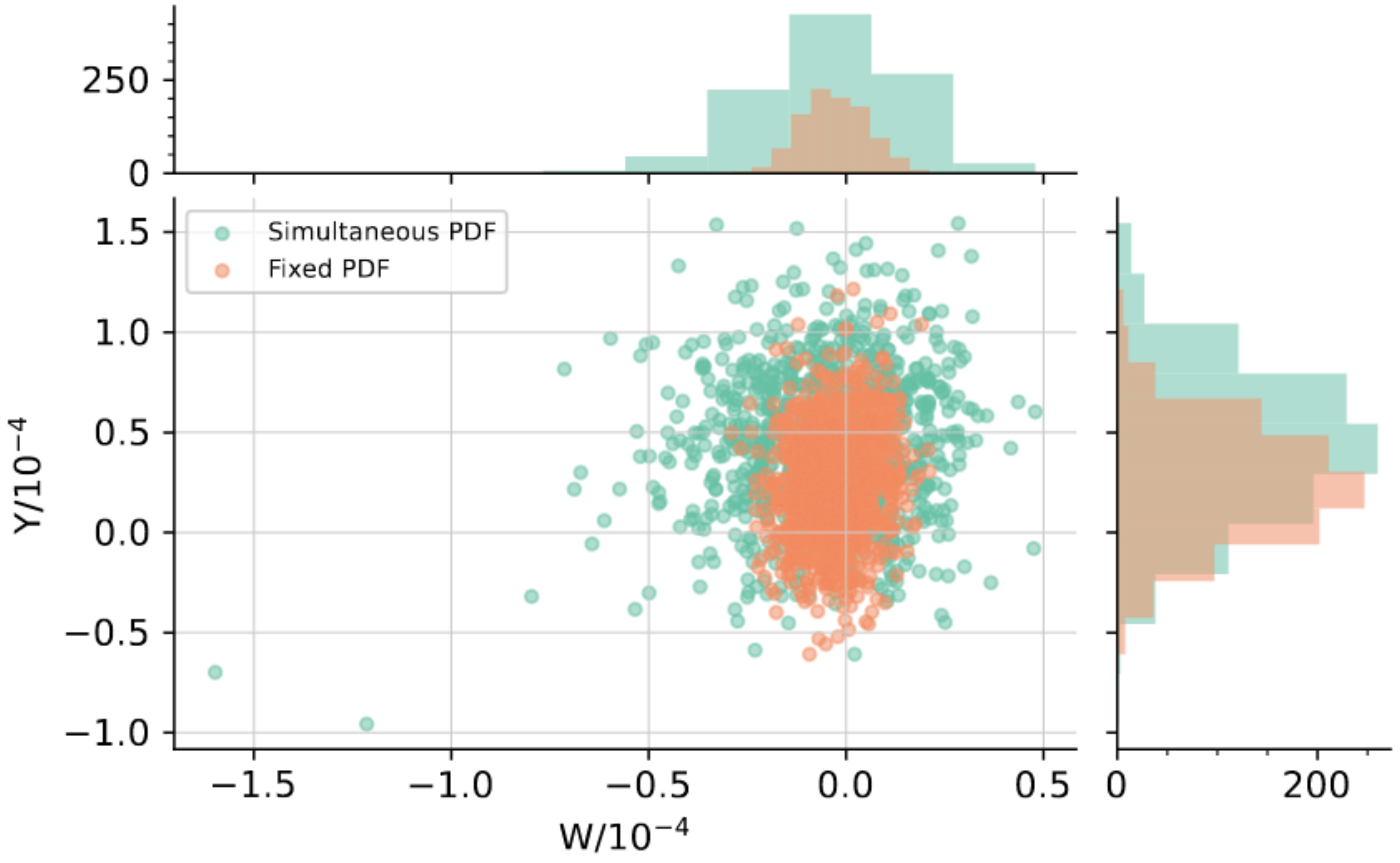
- Hence, the additional layer of the neural network **easily implemented** in the case of the SMEFT.
- In principle, the method can also be applied to **more complicated parameters**, e.g. $\alpha_S(m_Z)$, where the dependence cannot be well-modelled by a K-factor approximation. This can be achieved by constructing a **polynomial interpolant** between PDF evolution kernels and partonic cross-sections evaluated at different values of $\alpha_S(m_Z)$.

Benchmark of results

- In Iranipour, Ubiali, 2201.07240, the authors repeat the 'scan' study of Greljo et al, 2104.02723, now using the new **SimuNET methodology**.
- **Compatible bounds** in all cases are obtained, with **similar broadenings of the bounds** on the Wilson coefficients compared with **fixed PDFs** in the projected HL-LHC fit.

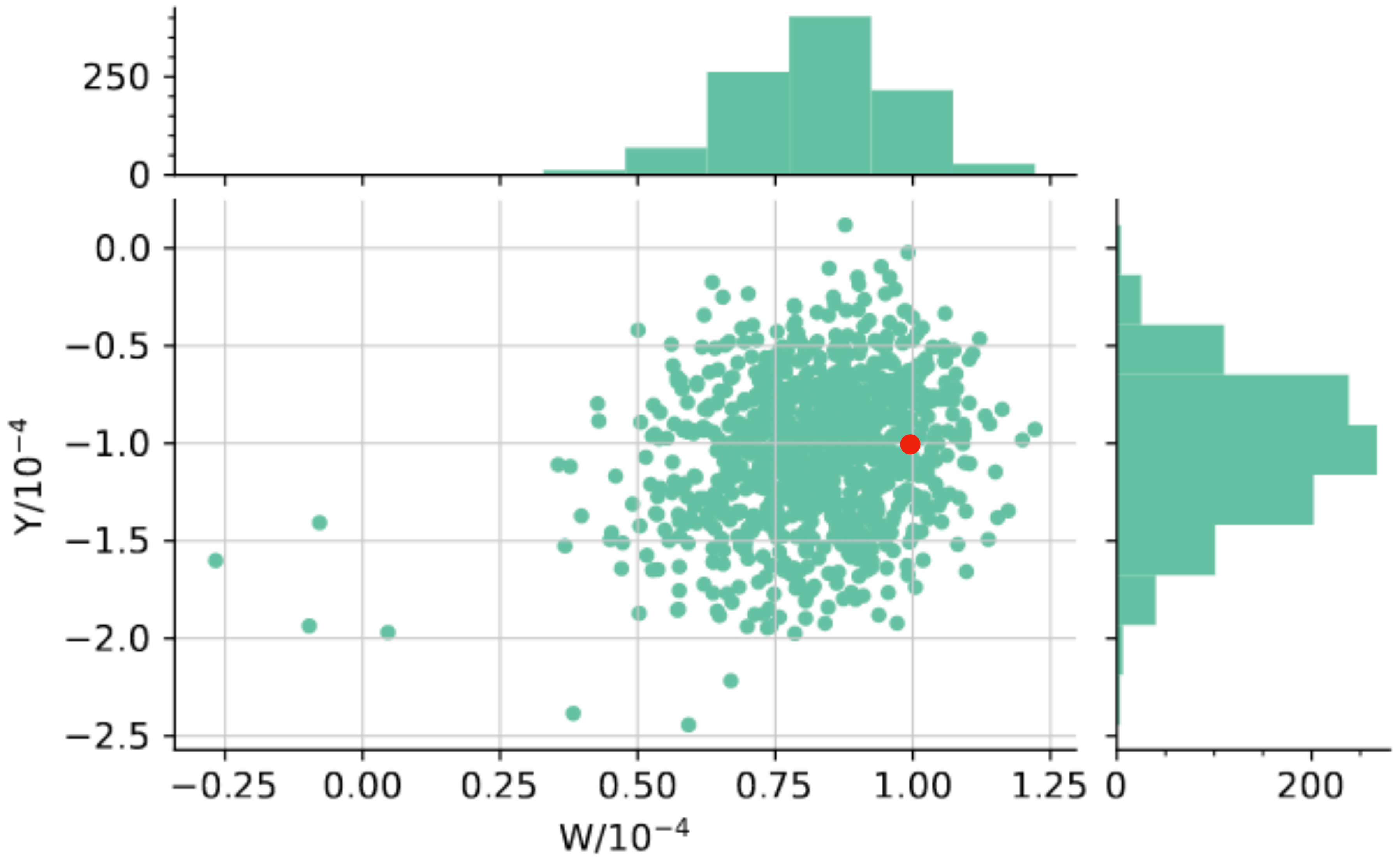
	SM PDFs	SMEFT PDFs	best-fit shift	broadening
$W \times 10^5$ (this work)	$[-2.0, 1.4]$	$[-4.3, 3.4]$	-0.2	$+126\%$
$W \times 10^5$ [17]	$[-1.4, 1.2]$	$[-8.1, 10.6]$	-1.4	$+620\%$
$Y \times 10^5$(this work)	$[-3.2, 8.1]$	$[-3.1, 11.7]$	$+1.9$	$+31\%$
$Y \times 10^5$[17]	$[-5.3, 6.3]$	$[-11.1, 12.6]$	$+0.3$	$+110\%$

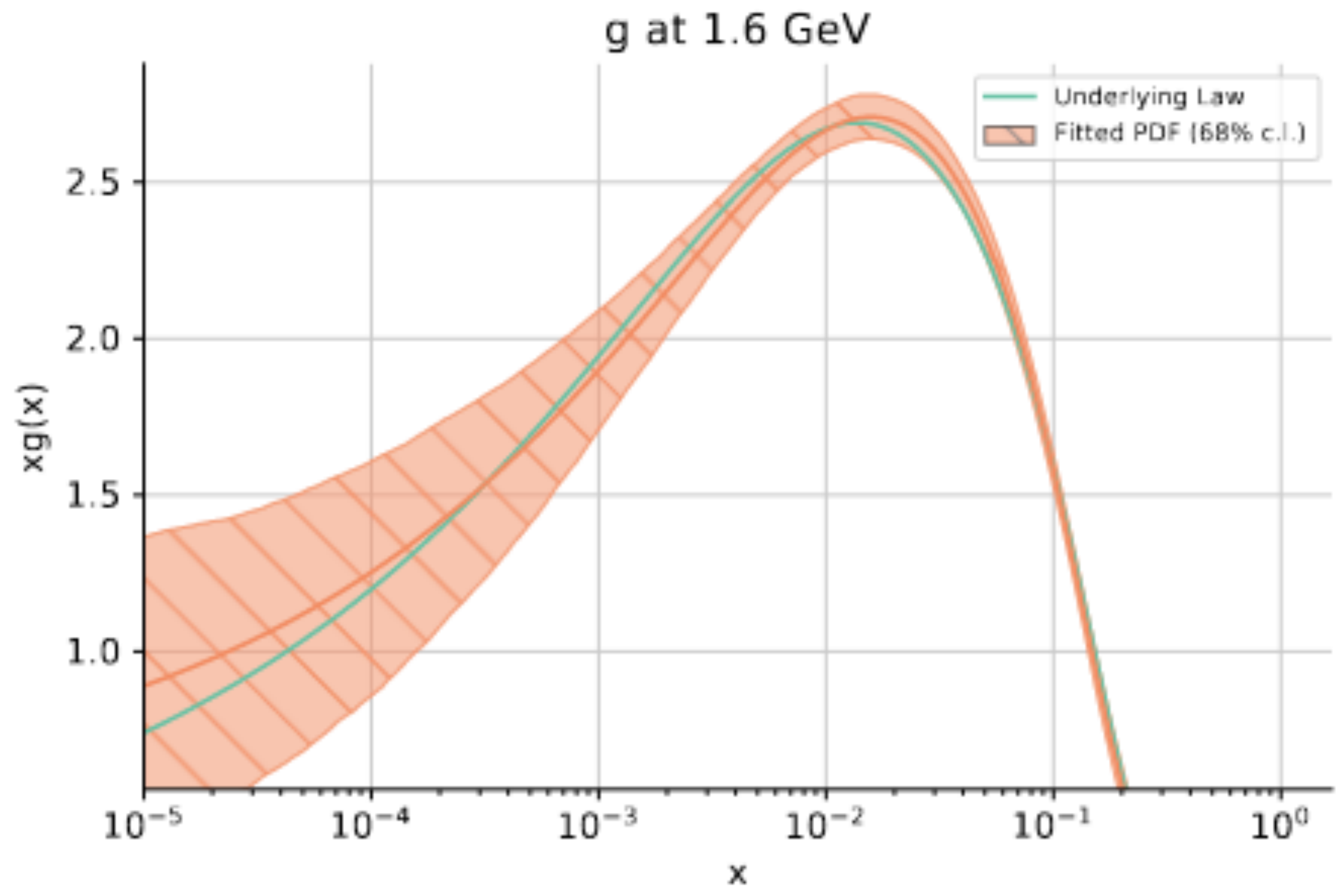
Benchmark of bounds from SimuNET paper against Greljo et al., 2104.02723 ([17] in above)



Closure tests

- The authors also perform **closure tests** to stress-test the methodology; this is standard practice in PDF fitting.
    1. An **arbitrary PDF set** is chosen, and used to generate **fake SM predictions.**
    2. Similarly, **arbitrary values of the Wilson coefficients** are chosen, and the K-factor approximation is used to convert the fake SM predictions to **fake SMEFT predictions**.
    3. The **SimuNET methodology** is deployed to see whether it can determine the underlying PDF and underlying Wilson coefficients.
- **It all works!** Next challenge: use it in a more interesting phenomenological scenario…





Future and related work

New simultaneous SMEFT fits in the top sector

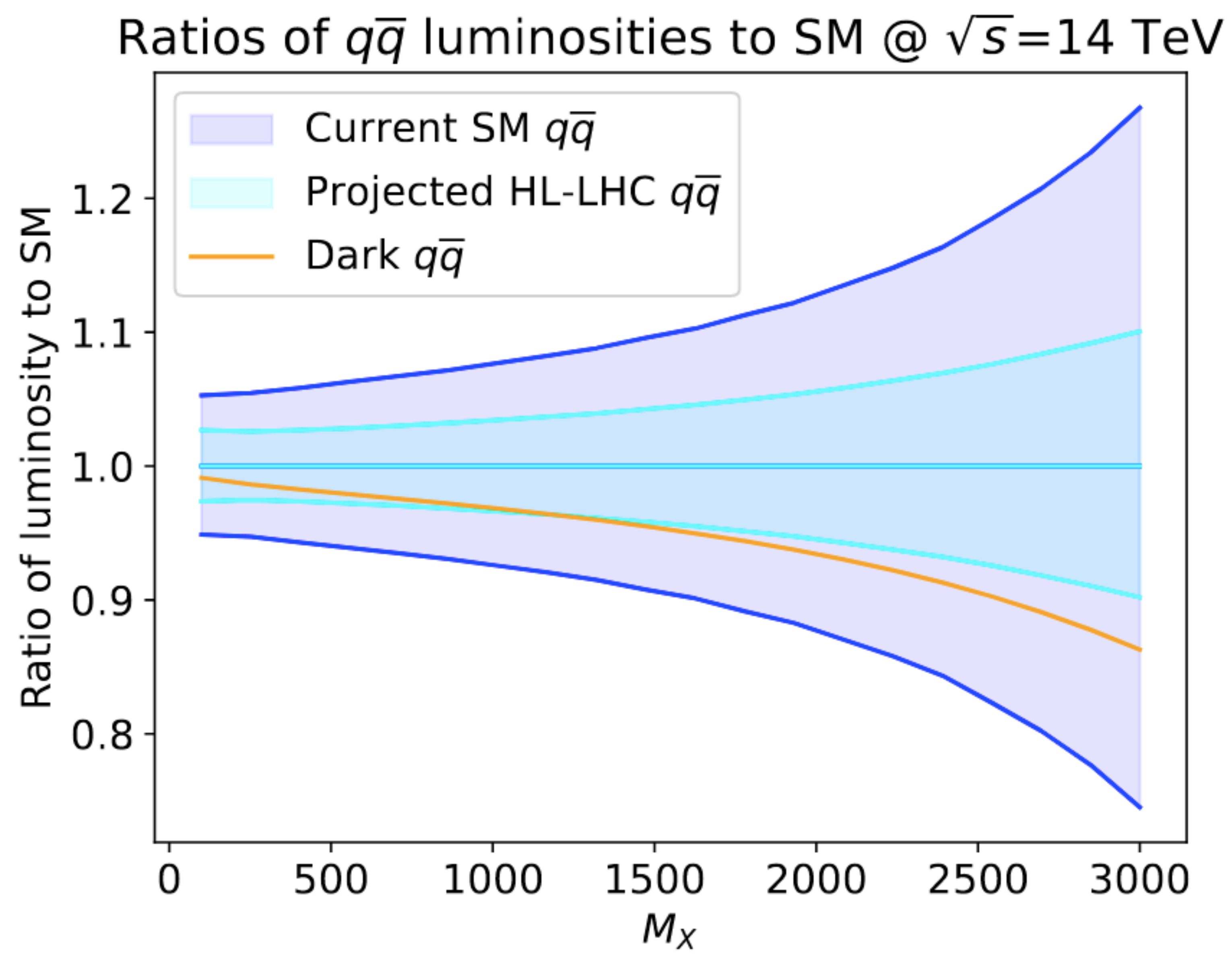
- The **top sector** provides a great playground in the search for New Physics, and has been used in multiple EFT analyses, including **SMEFiT** (2105.00006) and **FitMaker** (2012.02779), for example.

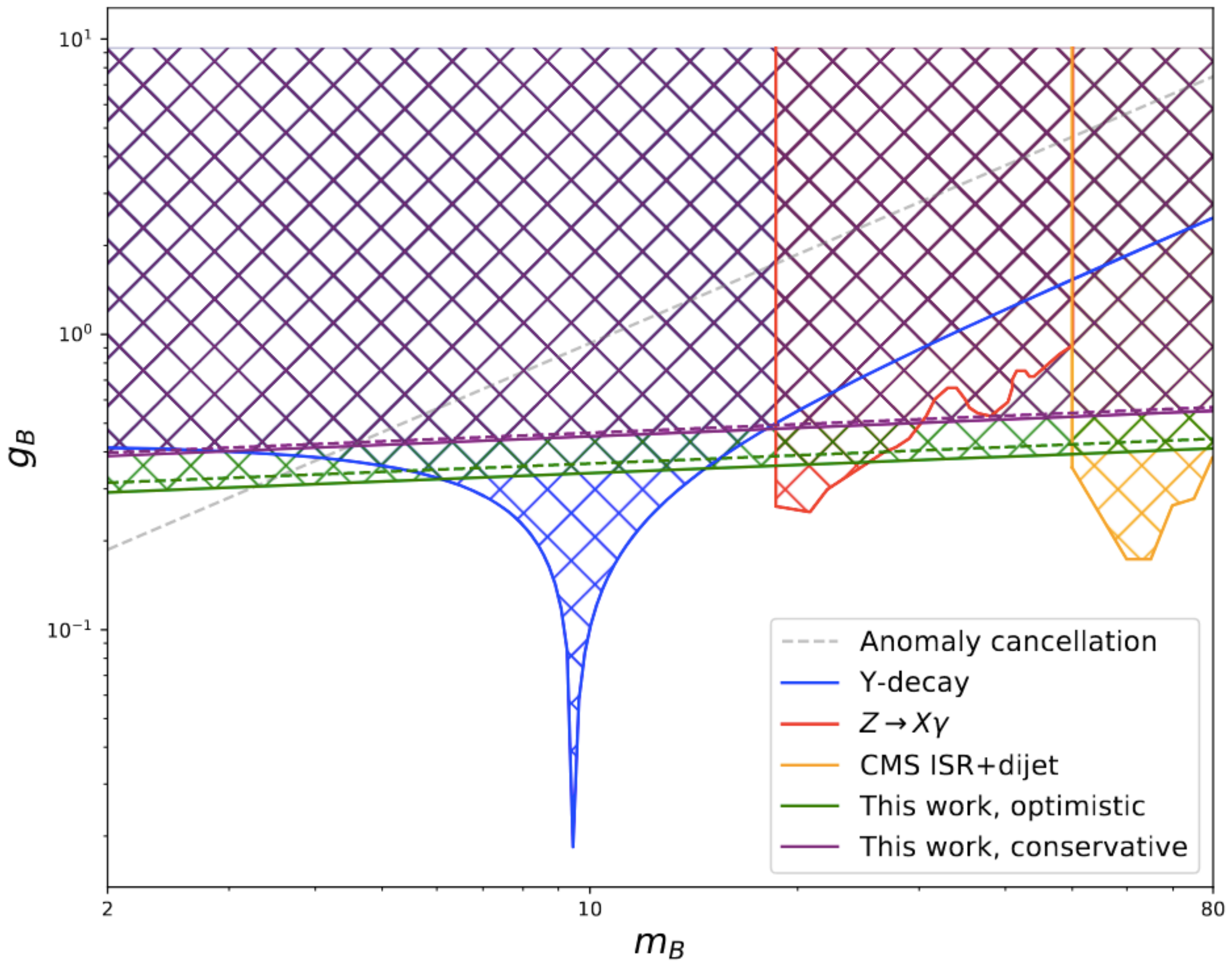
- It is interesting to ask what happens to the bounds from these analyses when **PDF interplay is additionally considered** (especially as more top data is added to PDF fits) - **perfect opportunity for the new SimuNET methodology**!

- We are now preparing a work in this direction, now with **16 SMEFT operators** simultaneously fitted with PDFs (rather than just 2!).

New parton flavours?

- Whilst the SMEFT is a great tool in searching for New Physics, it does not capture **weakly-coupled, light particles**. This does not mean that proton structure is unaffected by these hypothetical particles, though!
- It was shown in McCullough, **Moore**, Ubiali, 2203.12628, that the HL-LHC will be able to **probe proton structure** to such a precise degree as to place **stringent bounds** on certain models of **leptophobic dark photons**, which should appear as **constituents of the proton** (if they exist!).





Conclusions

- **Simultaneous determination of PDFs and physical parameters**, in particular **SMEFT Wilson coefficients**, will be **vital in future analyses**.

- Members of the **PBSP team** have already produced two important works in this direction: (i) a **phenomenological study** 2104.02723 showing the need for simultaneous extraction; (ii) a **methodology** (SimuNET, 2201.07240) capable of **fast simultaneous fitting**.

- The area is extremely **fertile**, with lots of interesting **future directions**, both in simultaneous SMEFT-PDF extraction (e.g. in the **top sector**), and other related projects.

Thanks for listening!
Questions?