Beyond the Standard Proton

for 'Parton Distributions and Nucleon Structure, 2022' ,
at the Institute for Nuclear Theory, Seattle

James Moore, University of Cambridge
(on behalf of the 'Physics Beyond the Standard Proton' group)

PBSP

erc
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*)

PBSP

erc
European Research Council
Established by the European Commission

Talk overview

1. Introduction: Joint fits of PDFs and BSM parameters

2. Simultaneous fits of PDFs and SMEFT Wilson coefficients

3. The dark side of the proton

4. Conclusions

1. - Introduction: Joint fits of PDFs and BSM 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**, masses and couplings of new particles in specific **BSM models**;

(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

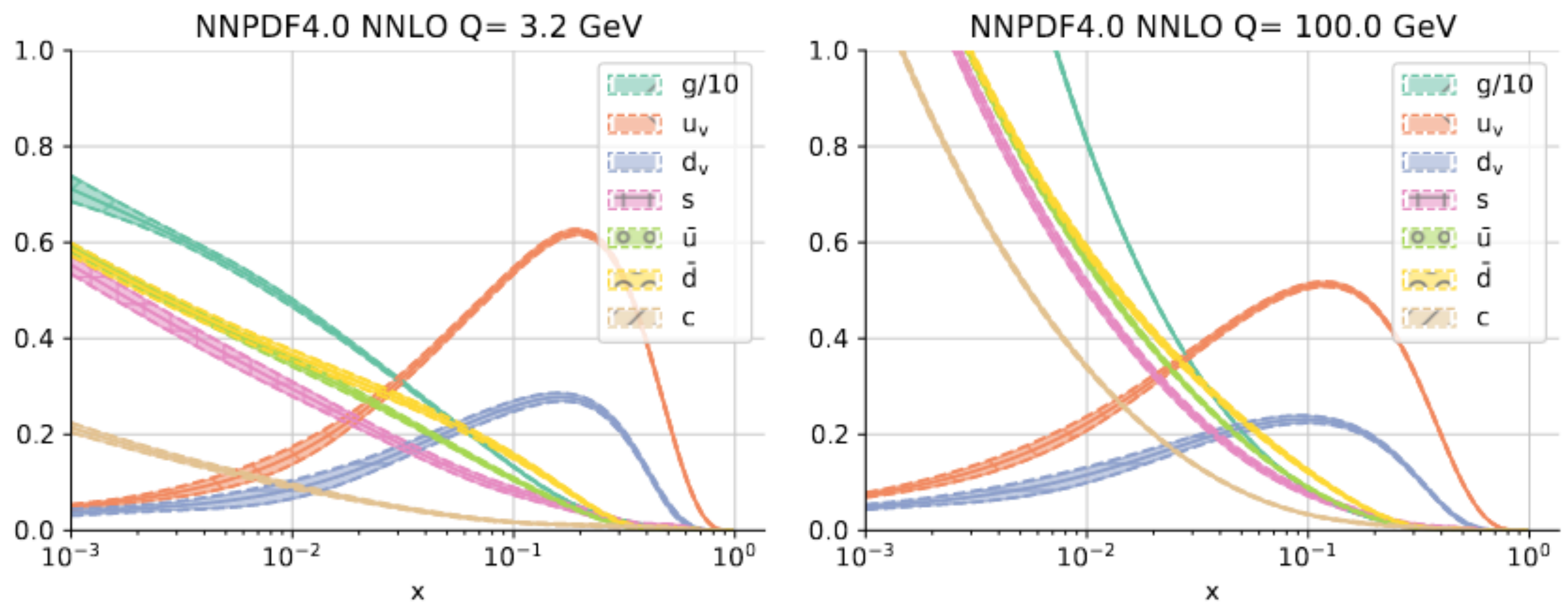
- Typically, the ‘physics’ parameter fits and PDF parameter fits **don’t talk.**

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.

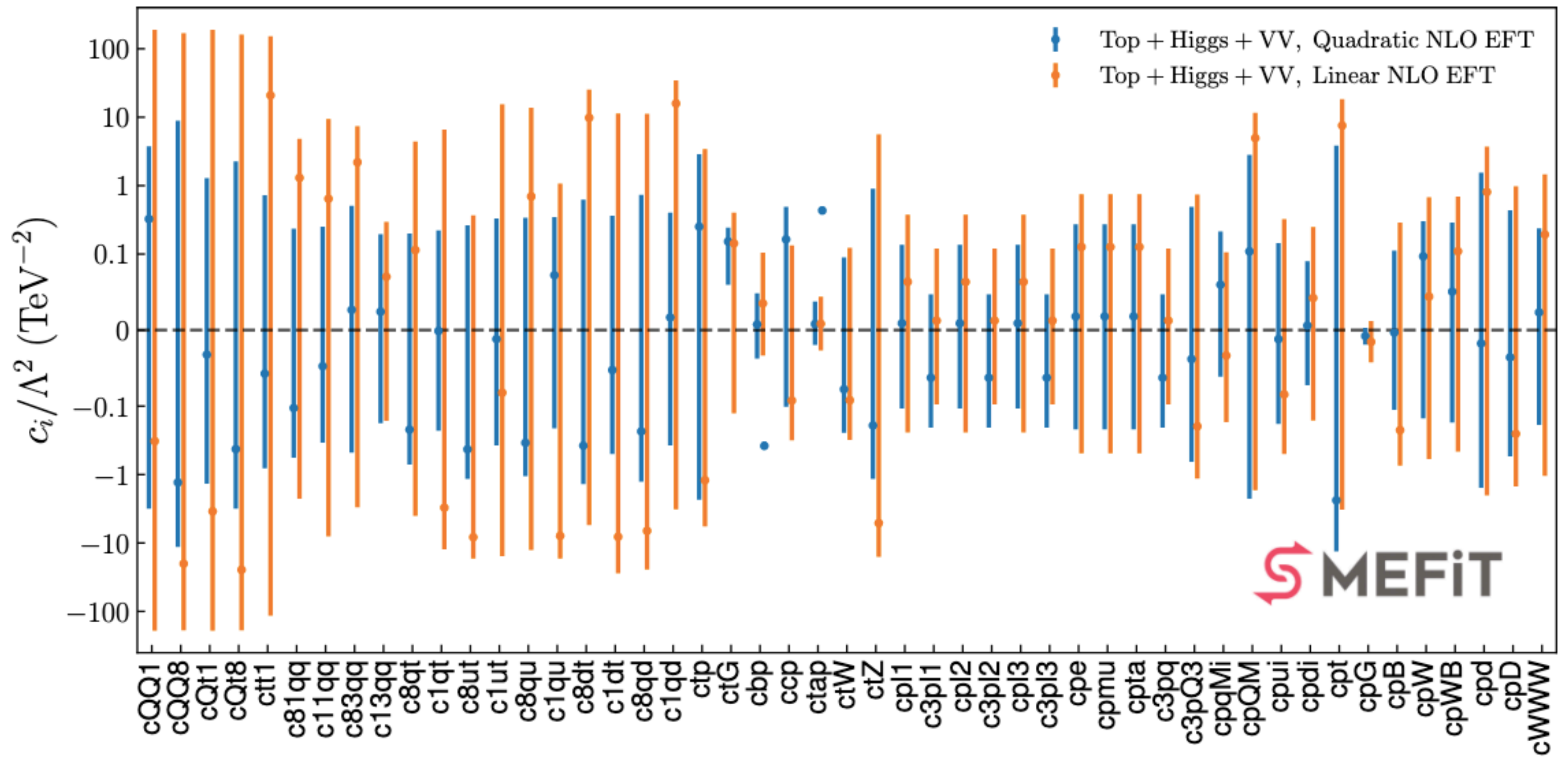


‘Physics’ parameter fits

- 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.

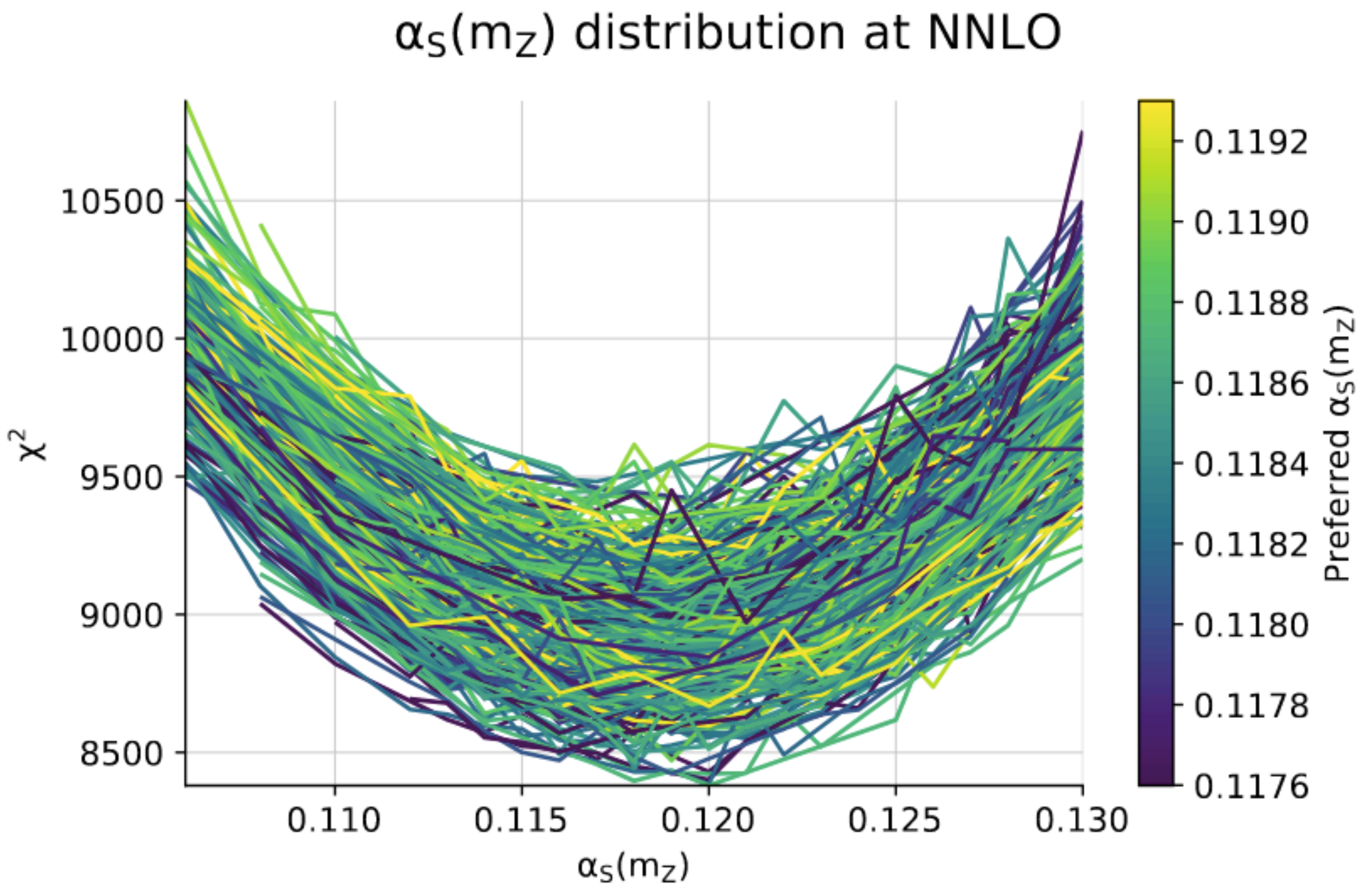
- For example, if we fit PDFs **assuming all Wilson coefficients in the SMEFT are zero**, but then **use those PDFs in a fit of SMEFT Wilson coefficients**, our resulting bounds **might be misleading**. The same applies to SM parameters.
- In the case of BSM models, we could even **miss New Physics**,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

Key question for this talk:

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

2. - Simultaneous fits of PDFs and SMEFT Wilson coefficients

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 e.g. 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.*

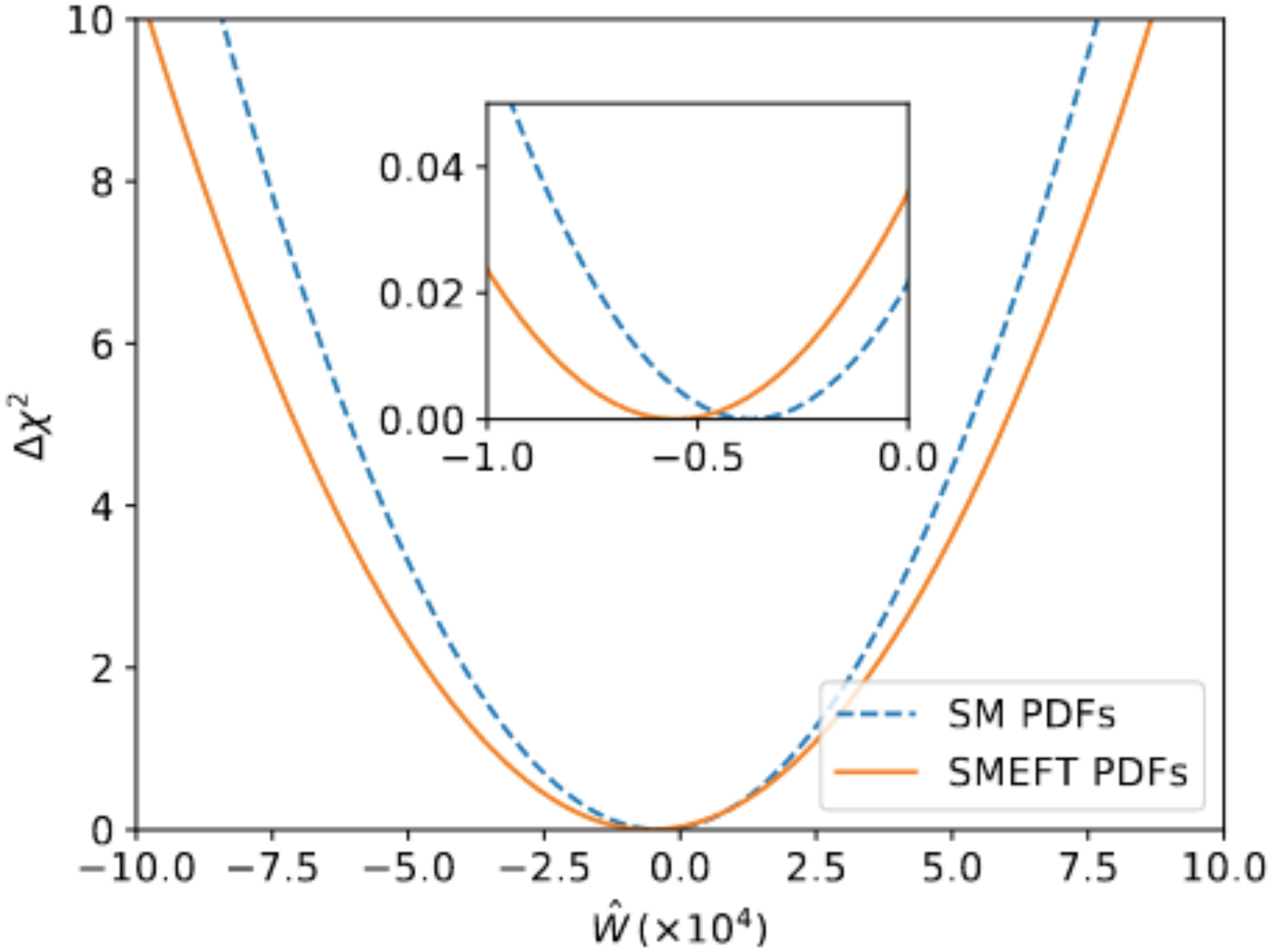
   A phenomenological study, demonstrating the impact of a simultaneous SMEFT/PDF fit in the context of the oblique W, Y parameters using current and projected Drell-Yan data.

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.

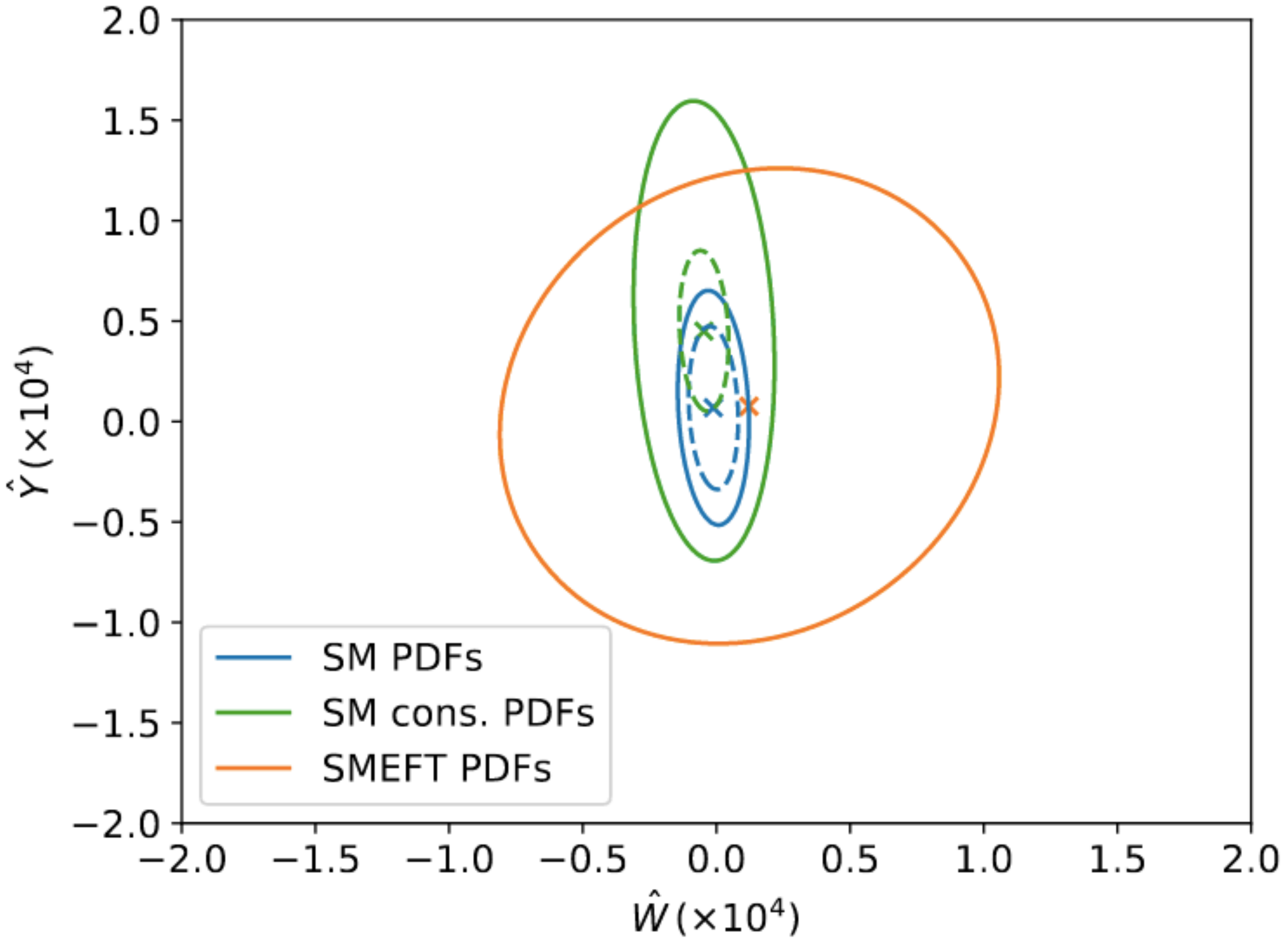
Parton distributions in the SMEFT from high-energy Drell-Yan tails

- 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**.
- 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**.



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

- On the other hand, when we use **projected HL-LHC data**, the impact of a simultaneous fit versus a fixed PDF fit becomes **enormous**!
- Without a simultaneous fit, we find that the size of the bounds is **significantly underestimated** - this could lead to claims of discovering New Physics when it **isn't necessarily there**.



New simultaneous SMEFT fits in the top sector

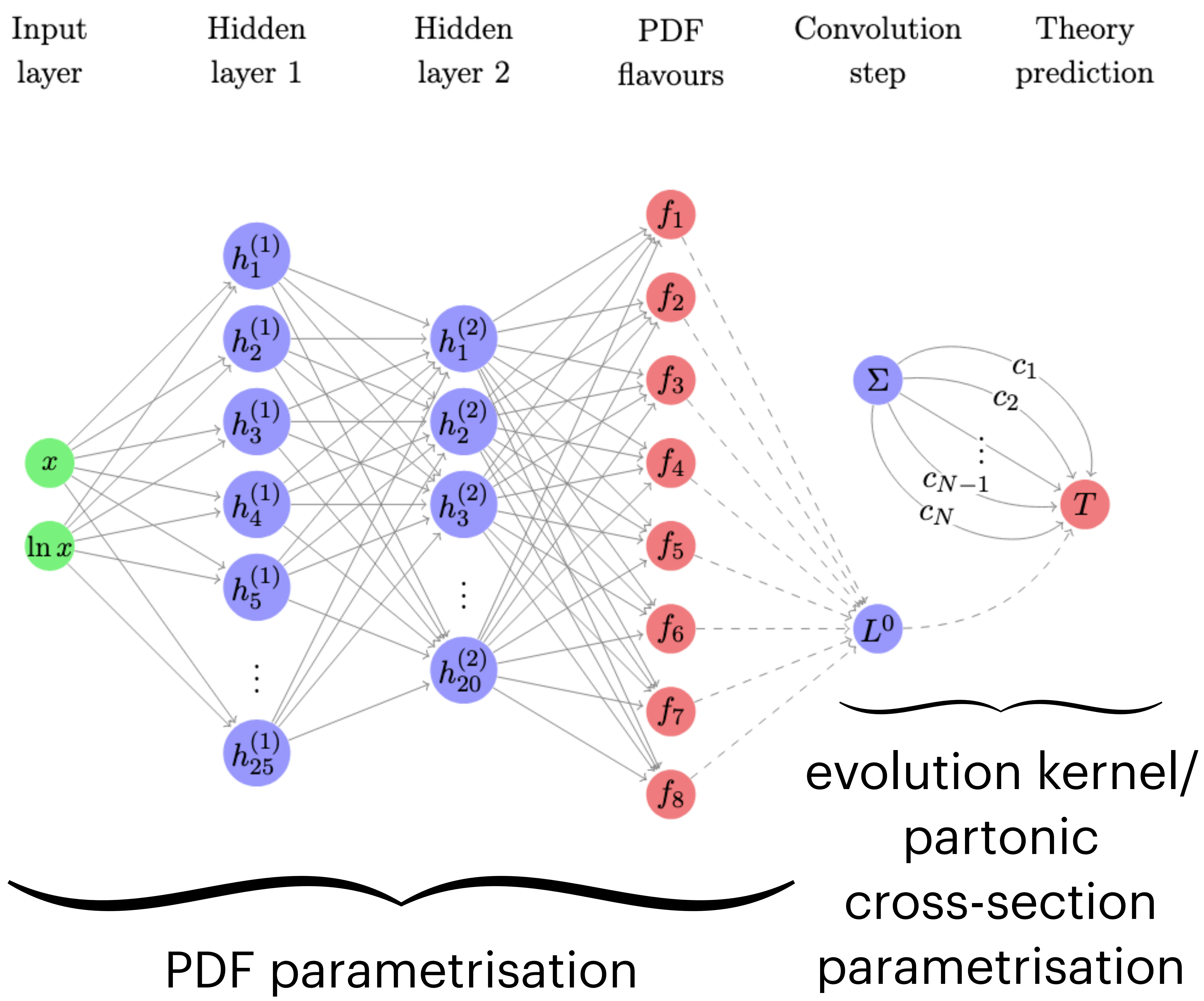
- Our DY study motivates more serious studies of PDF-SMEFT interplay; our next target is the **top sector.** The top sector provides a great playground in the search for New Physics, and has been used in multiple EFT analyses, including **SMEFiT** (2105.00006) and **FitMaker** (2012.02779), for example.

- Unfortunately, in a serious top fit, we need to include many more SMEFT operators: a total of **16 SMEFT operators** contribute to the observables of interest. The ‘scan’ methodology scales terribly though...

- Therefore, in our top study, we use a **new, efficient methodology** proposed in Iranipour, Ubiali, 2201.07240.

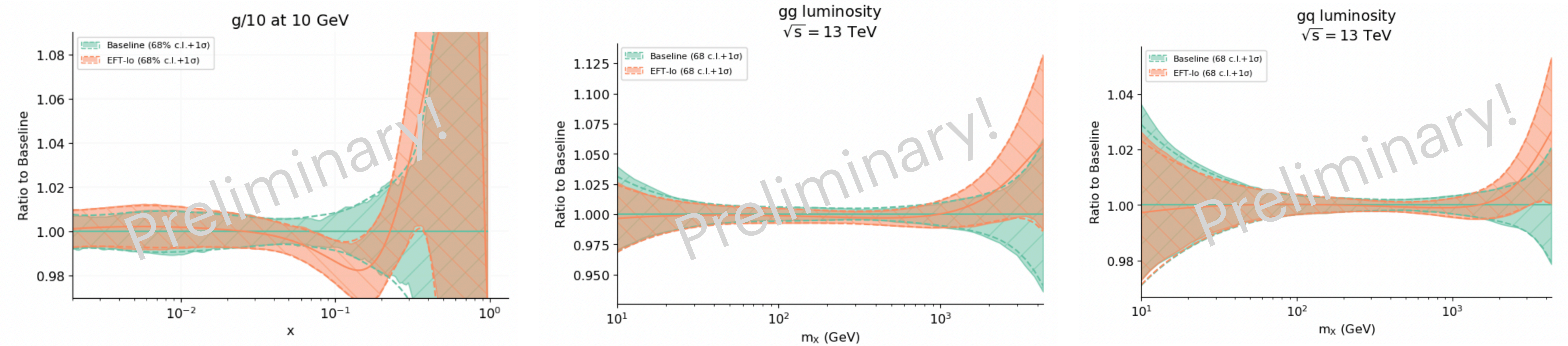
New simultaneous SMEFT fits in the top sector

- 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.
- In brief, the idea of the new method 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 **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**!



New simultaneous SMEFT fits in the top sector

- **Preliminary results** show that a simultaneous top fit versus a fixed PDF top fit can have significant impact on the **gluon PDF**, and in turn the gg and gq **luminosities** - **watch this space for more results**!



3. - The dark side of the proton

Light new physics and PDFs

- So far, we've focussed on **joint PDF-SMEFT determinations**. However, whilst the SMEFT is a great tool in searching for New Physics, it does not capture **new weakly-coupled, light particles**. Proton structure could also be affected by these new degrees of freedom!
- In this case, we could **still see the impact on proton structure** by including the new particles as **constituents of the proton**.
- The idea is not too far-fetched! The inclusion of new **coloured** particles, e.g. **gluinos**, has already been studied by Berger et al. in 0406143 (from 2005) and 1010.4315 (from 2010). **Strong constraints** can be derived assuming that new coloured particles alter our SM view of proton structure.

Light new physics and PDFs

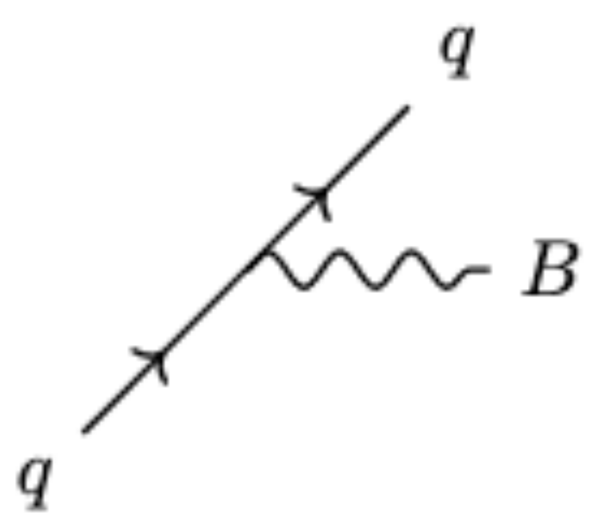
- *Idea:* now PDFs are known **very precisely**, and their uncertainties **will continue to reduce in the near future with the HL-LHC**, could we do the same for a **colourless** particle too?
- In McCullough, **Moore**, Ubiali, 2203.12628, we studied the impact of using a **toy dark matter candidate**, namely a **light leptophobic dark photon** B which couples to quarks via the effective interaction Lagrangian:

$$\mathscr{L}_{\text{int}} = \frac{1}{3} g_B \bar{q} \gamma^\mu B_\mu q$$

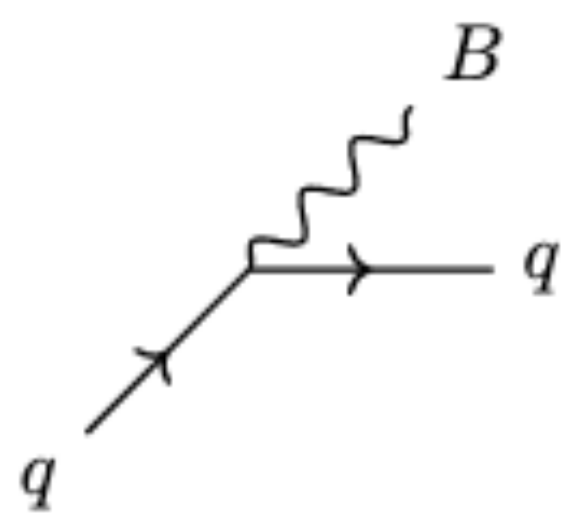
- Low-energy experimental probes already strongly constrain $m_B < 2$ GeV.
- We also treat this as an effective theory, valid up to the mass of the Z, where **kinetic mixing** effects become important; so for us: $m_B \in [2,80]$ GeV.

DGLAP in the presence of dark photons

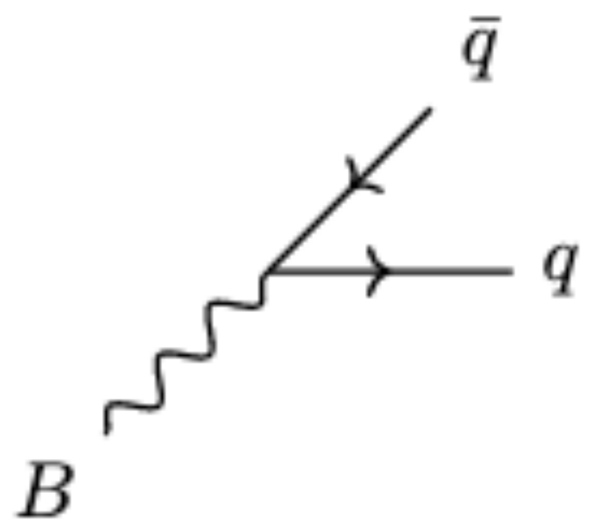
- Now, to include the dark photon as a constituent of the proton, we mimic the earliest studies into **photon PDFs** (namely MRST 0411040, from 2004), using the following procedure:
  1. Compute the **dark photon splitting functions**, and add them to **DGLAP evolution**.
  2. Starting from an **appropriate initial-scale ansatz**, and a **reference PDF set**, evolve using the **modified DGLAP equations**. Since we assume $m_B > 2$ GeV, greater than the standard initial scale 1.65 GeV, we **always generate the dark photon from zero** similar to a **heavy quark**. We choose the **state-of-the-art NNPDF3.1 LUXQED set** as our reference set (this will soon be replaced by NNPDF4.0 LUXQED).
  3. Compare resulting PDF set predictions with reference SM predictions to see **impact of inclusion of a dark photon**.



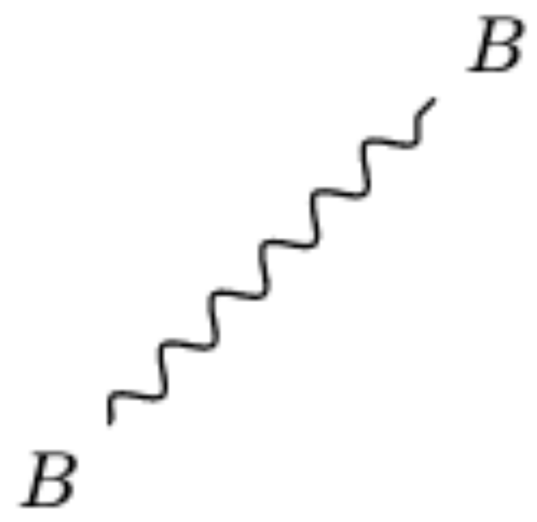
$$P_{qq}(x) = \frac{1+x^2}{9(1-x)_+} + \frac{1}{6}\delta(1-x)$$



$$P_{Bq}(x) = \frac{1}{9}\left(\frac{1+(1-x)^2}{x}\right)$$



$$P_{qB}(x) = \frac{x^2+(1-x)^2}{9}$$



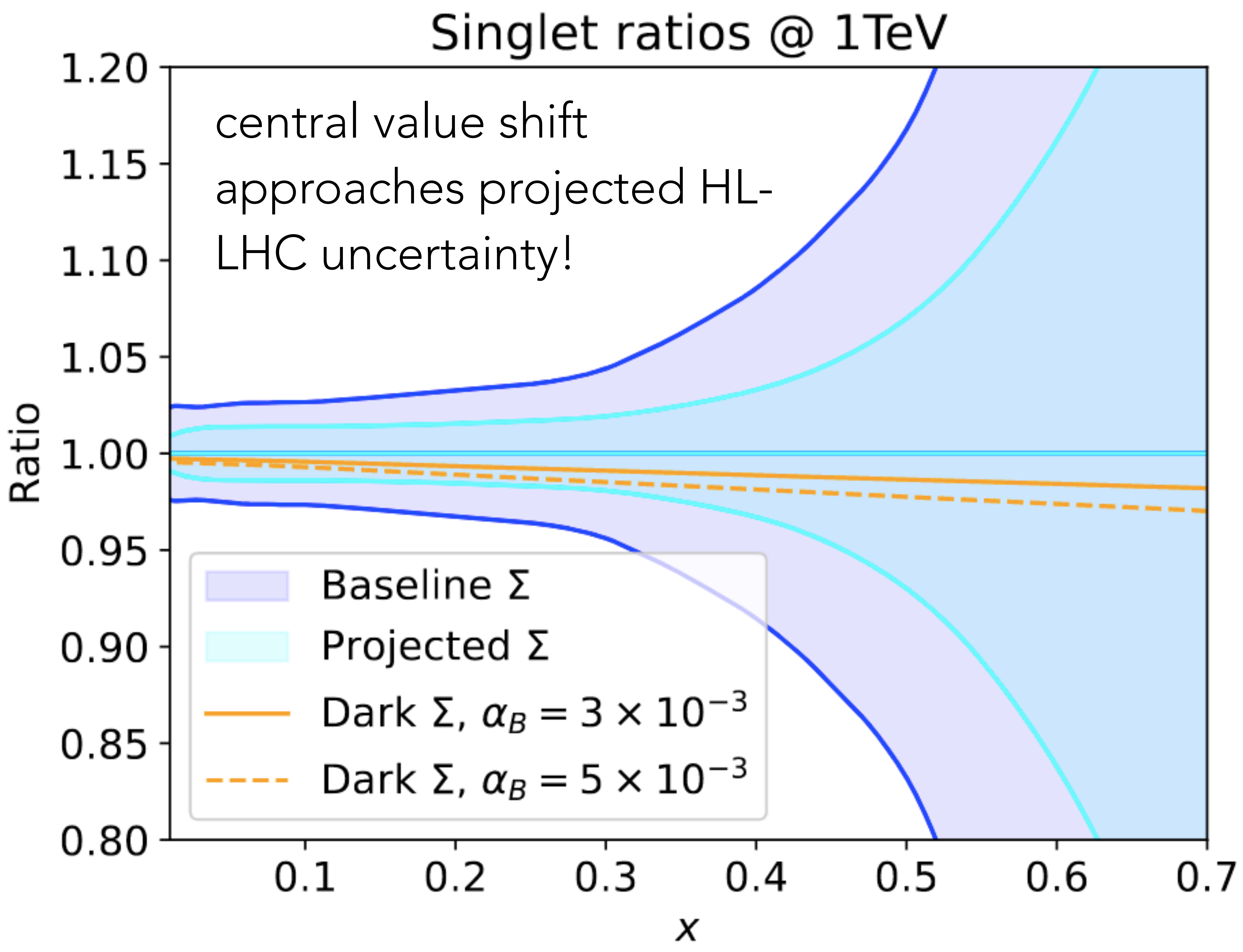
$$P_{BB}(x) = -\frac{2}{27}\delta(1-x)$$

DGLAP in the presence of dark photons

- All four splitting functions are multiplied by $\alpha_B = g_B^2/4\pi$ in the DGLAP equations. Assuming a dark coupling of order $\alpha_B \sim 0.001$ (reasonable in the literature for this model), we see that we must also include:
  - NNLO QCD effects, $\alpha_S^3 \sim 0.001$
  - LO QED effects, $\alpha \sim 0.01$ (this implies that we must use a **photon PDF**; we use the LUXQED PDF from the NNPDF3.1 QED baseline)
  - QED-QCD mixing, $\alpha\alpha_S \sim 0.001$
- These contributions are well-known and already implemented in the **APFEL public evolution code**, which we modify in our work.

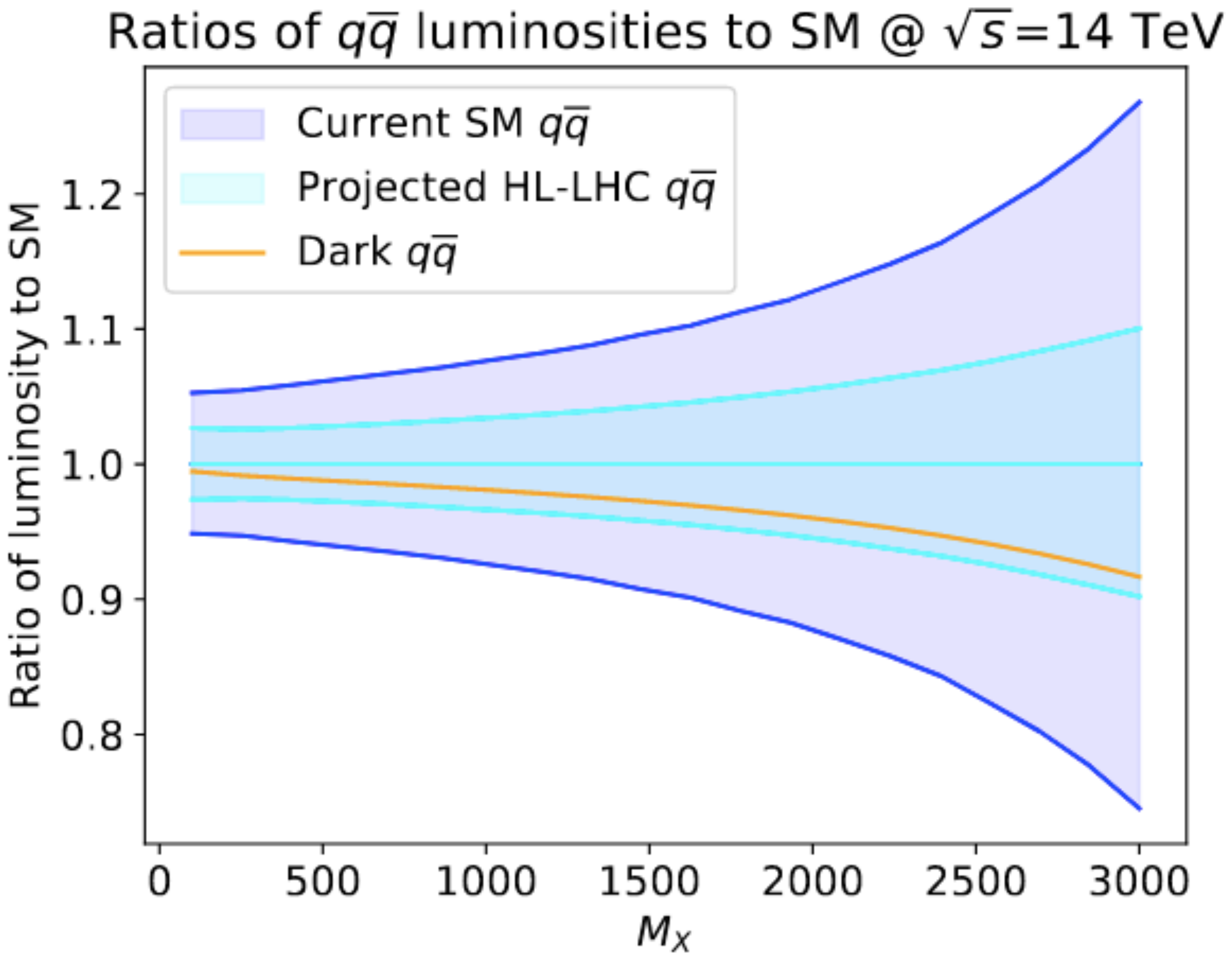
Impact on PDFs and parton luminosities

- We can now study the impact of including a dark photon in DGLAP evolution on **PDFs** and **parton luminosities**, and hence on **theoretical predictions for collider processes**.
- E.g. including a dark photon modifies the **singlet PDF**, as shown on the right. Light blue bands correspond to **projected PDF** uncertainty at the **HL-LHC** (see 1810.03639).
- The region that is most modified suggests that some values of the dark mass and coupling might lead to PDF sets which **perform too poorly** on **Drell-Yan sets**, relative to the baseline.

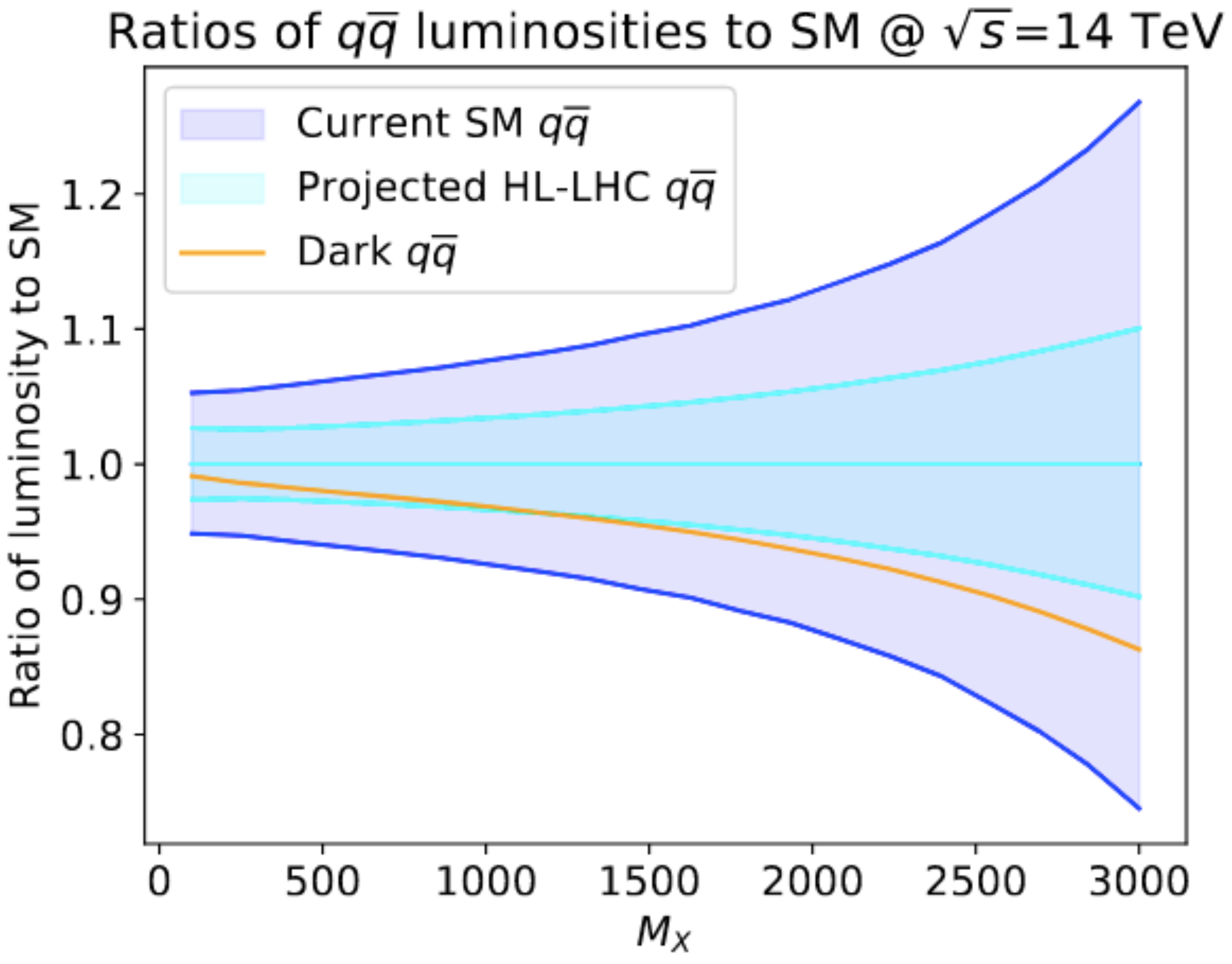


Impact on PDFs and parton luminosities

- The most important luminosity channel for DY is $q\bar{q}$; here, there is **tension with projected HL-LHC uncertainties** for some values of the mass and couplings!



(c) $m_B = 5$ GeV, $\alpha_B = 3 \times 10^{-3}$



(d) $m_B = 5$ GeV, $\alpha_B = 5 \times 10^{-3}$

Impact on PDFs and parton luminosities

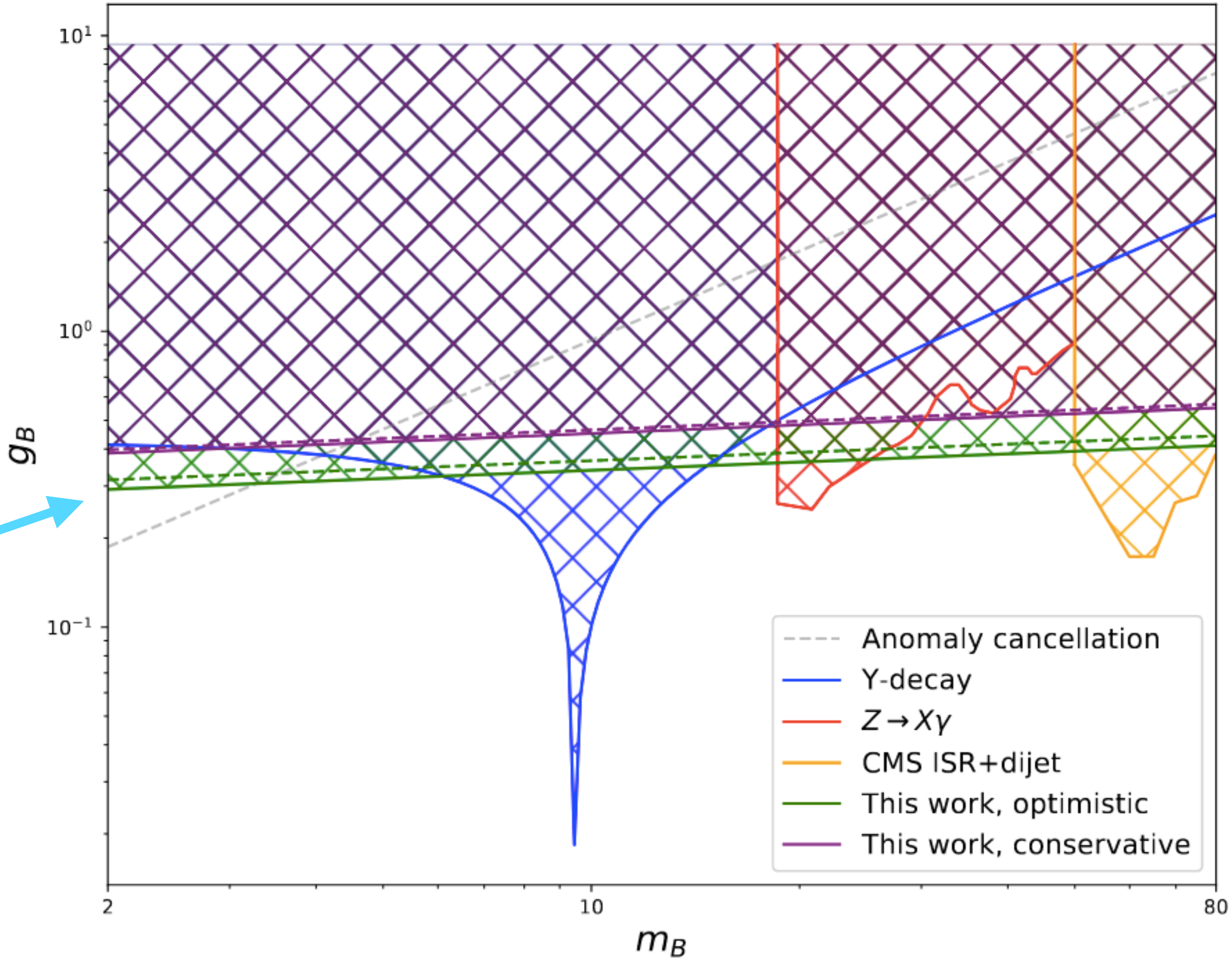
- Results we have seen so far suggest that we can definitely hope to constrain the dark photon's mass and coupling using DY data, **provided** we work with **HL-LHC projections** and **assume that PDF uncertainties will shrink as predicted**.
- We obtain **projected bounds** as follows:

1. Construct a large ensemble of 'dark' PDF sets, one for each point for a grid in dark parameter space (we use 32 points, so 32 PDF sets).
2. Construct predictions for a specific DY observable for each PDF set and compute the χ^2-statistic.
3. Compare to the reference fit's χ^2-statistic, and hence obtain projected bounds.

Impact on PDFs and parton luminosities

- The specific HL-LHC observable we choose to use is **neutral current Drell-Yan** at a centre-of-mass-energy $\sqrt{s} = 14$ TeV, in 12 bins of lepton invariant pair-mass. The projected data we use is a small modification of that produced for **Parton Distributions in the SMEFT from High-Energy Drell-Yan Tails**, 2104.02723.
- Two sets of projected data are used, corresponding to the following two scenarios:
  - *Optimistic*: Total integrated luminosity 6 ab^{-1} (both CMS and ATLAS available), with five-fold reduction in systematics.
  - *Conservative*: Total integrated luminosity 3 ab^{-1} (only CMS or ATLAS is available), with two-fold reduction in systematics.

Comparison of (projected) bounds



dashed lines: including projected HL-LHC PDF uncertainty

4. - Conclusions

Conclusions

- **Simultaneous determination of PDFs and BSM parameters**, will be **very important in future analyses** (especially as we enter Run III).

- Members of the **PBSP team** have already produced two works in the direction of simultaneous PDF-SMEFT fits: (i) a **phenomenological study** 2104.02723 showing the need for simultaneous extraction; (ii) a **methodology** (SimuNET, 2201.07240) capable of **fast simultaneous fitting**. We aim to continue with a more ambitious **top-sector fit**.

- There are interesting directions outside the SMEFT, e.g. studying **light, weakly-coupled particles** inside the proton, like our **dark photon** study.

Thanks for listening!
Quick questions before discussion?