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

Maeve Madigan

UNIVERSITY OF CAMBRIDGE

PBSP

European Research Council
Established by the European Commission

HEFT
16/06/22

Based on

- *arXiv:2104.02723, A.Greljo, S. Iranipour, Z. Kassabov, MM, J. Moore, J. Rojo, M. Ubiali, C. Voisey*

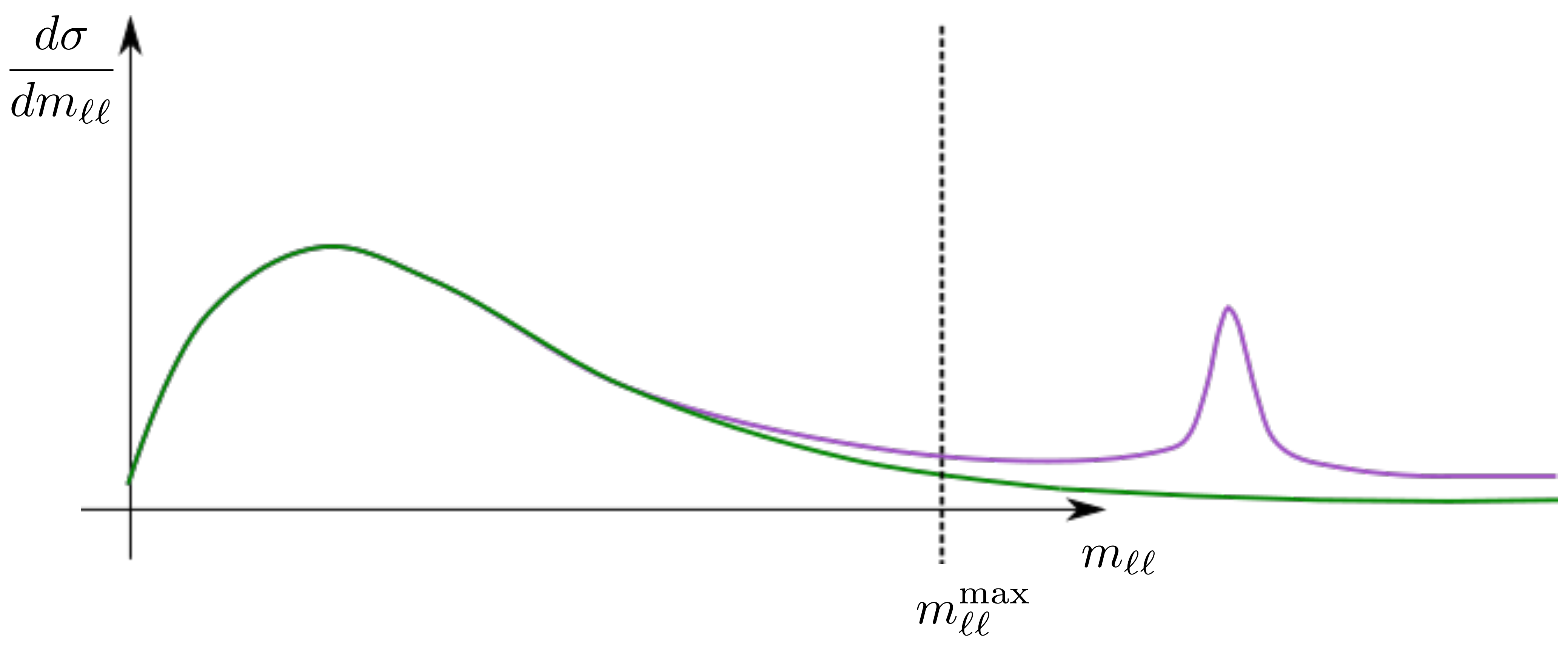
- ongoing work by:
    - *Maria Ubiali, Elie Hammou, Zahari Kassabov, MM, Luca Mantani, James Moore, Manuel Morales*

PBSP

Fitting the SMEFT

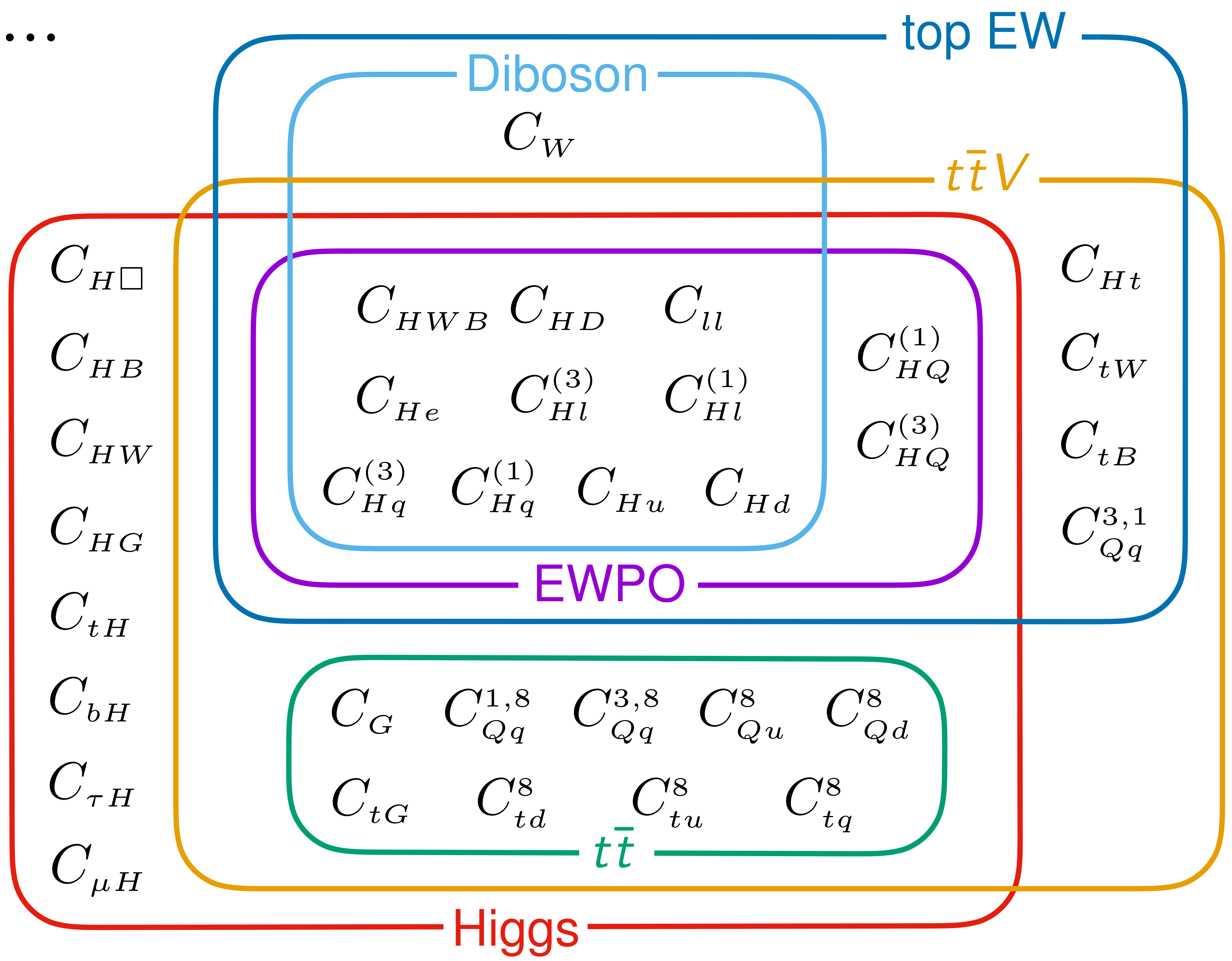
The SMEFT: a powerful framework for capturing deviations from the SM:

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \frac{C^{(5)}}{\Lambda}\mathcal{O}^{(5)} + \sum_i \frac{C_i^{(6)}}{\Lambda^2}\mathcal{O}_i^{(6)} + \ldots$$

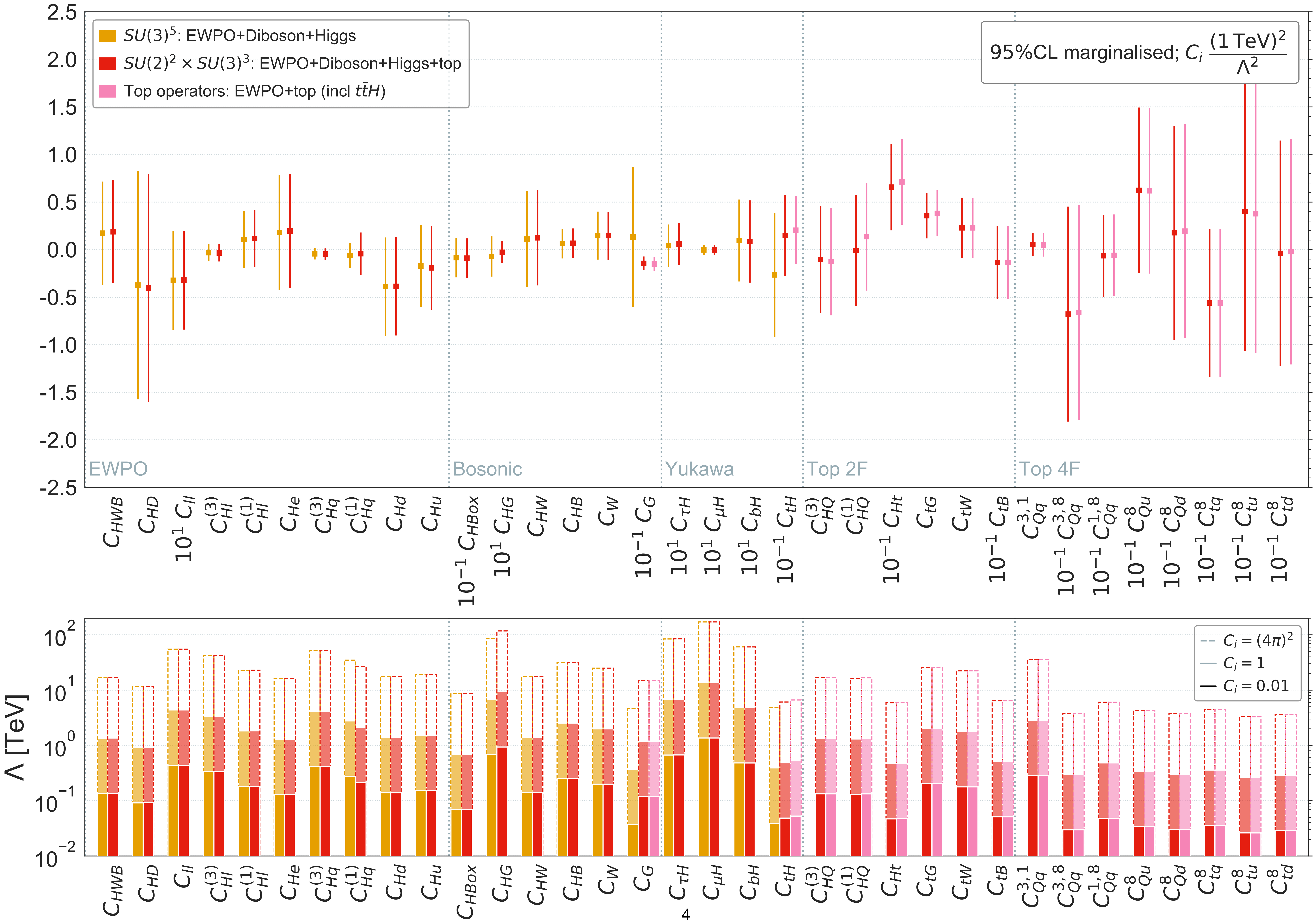


The SMEFT framework connects different sectors of observables measured at the LHC.

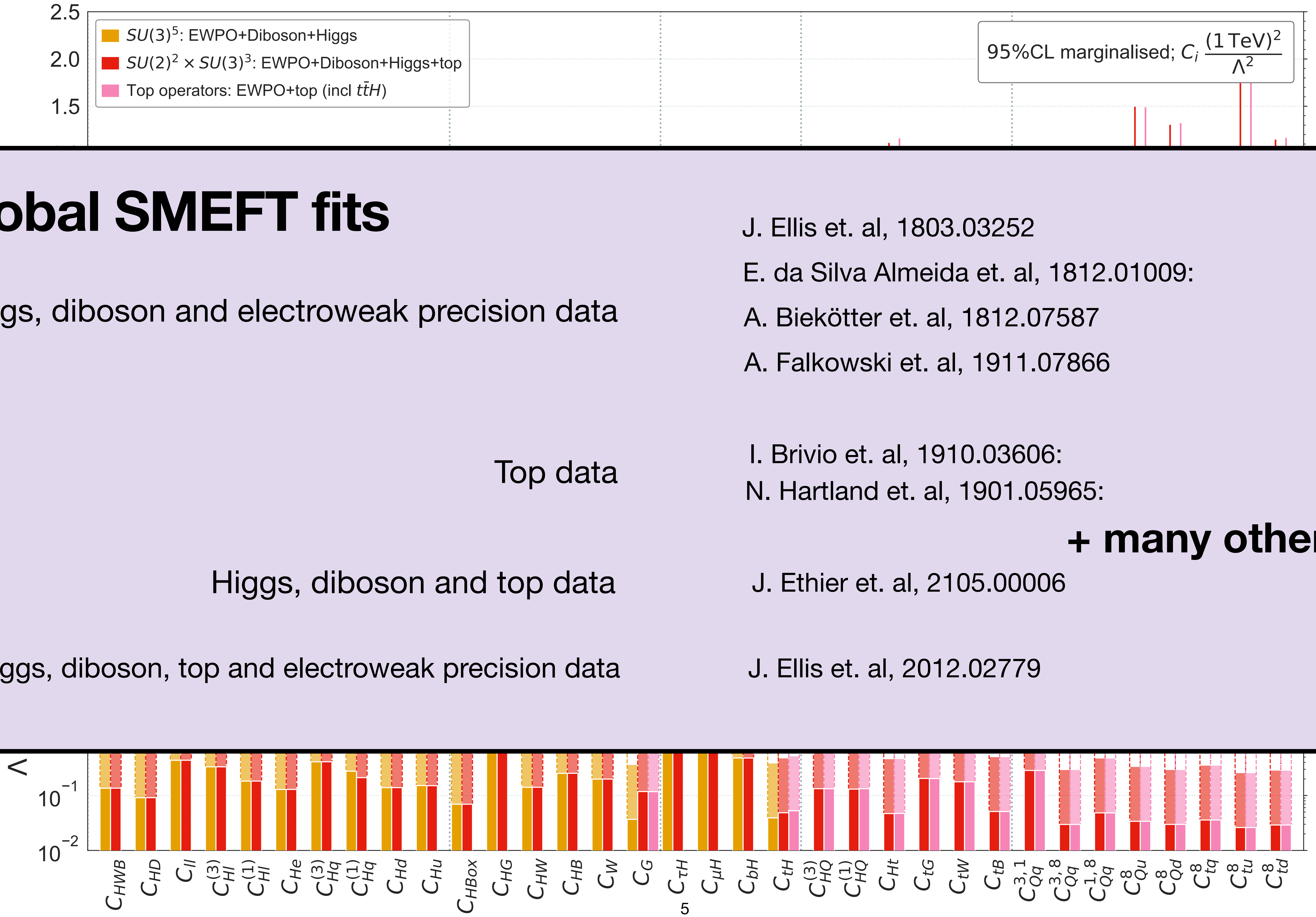
We can probe the SMEFT by taking a **global approach,** including as many datasets as possible.



2012.02779, J. Ellis et. al



2012.02779,
J. Ellis, MM,
K. Mimasu,
V. Sanz, T. You



Global SMEFT fits

Higgs, diboson and electroweak precision data

J. Ellis et. al, 1803.03252

E. da Silva Almeida et. al, 1812.01009:

A. Biekötter et. al, 1812.07587

A. Falkowski et. al, 1911.07866

Top data

I. Brivio et. al, 1910.03606:

N. Hartland et. al, 1901.05965:

+ many others….

Higgs, diboson and top data

J. Ethier et. al, 2105.00006

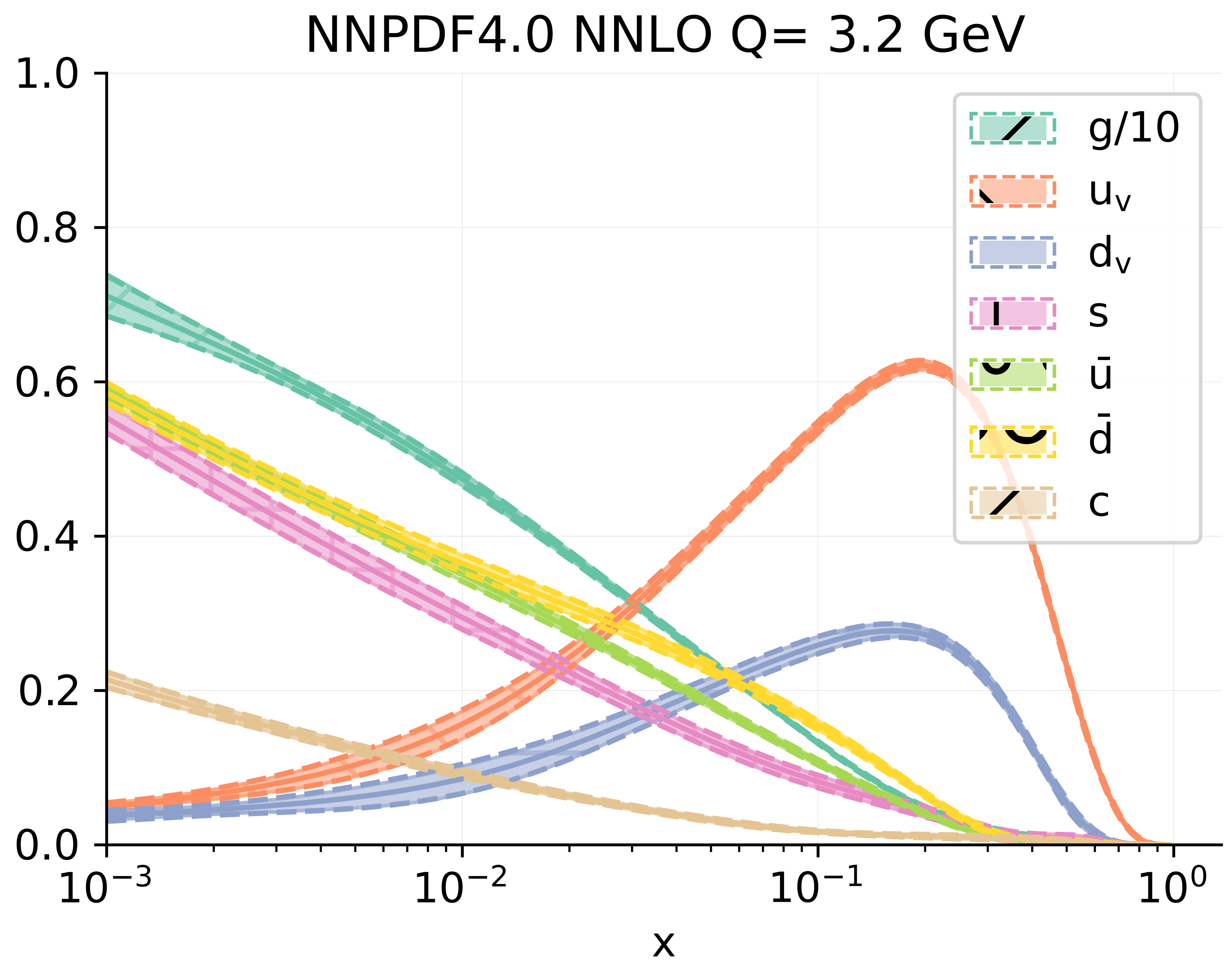
Higgs, diboson, top and electroweak precision data

J. Ellis et. al, 2012.02779

Inputs and assumptions

Global SMEFT fits are dependent on many inputs and assumptions:

- SMEFT flavour symmetry
- Electroweak input scheme: $\{\alpha_{\rm EW}, m_Z, G_F\}$
- Inclusion of $\mathcal{O}(\Lambda^{-4})$ contributions
- Choice of likelihood
- …..
- **Parton distribution functions**



Ball et. al, NNPDF4.0, 2109.02653

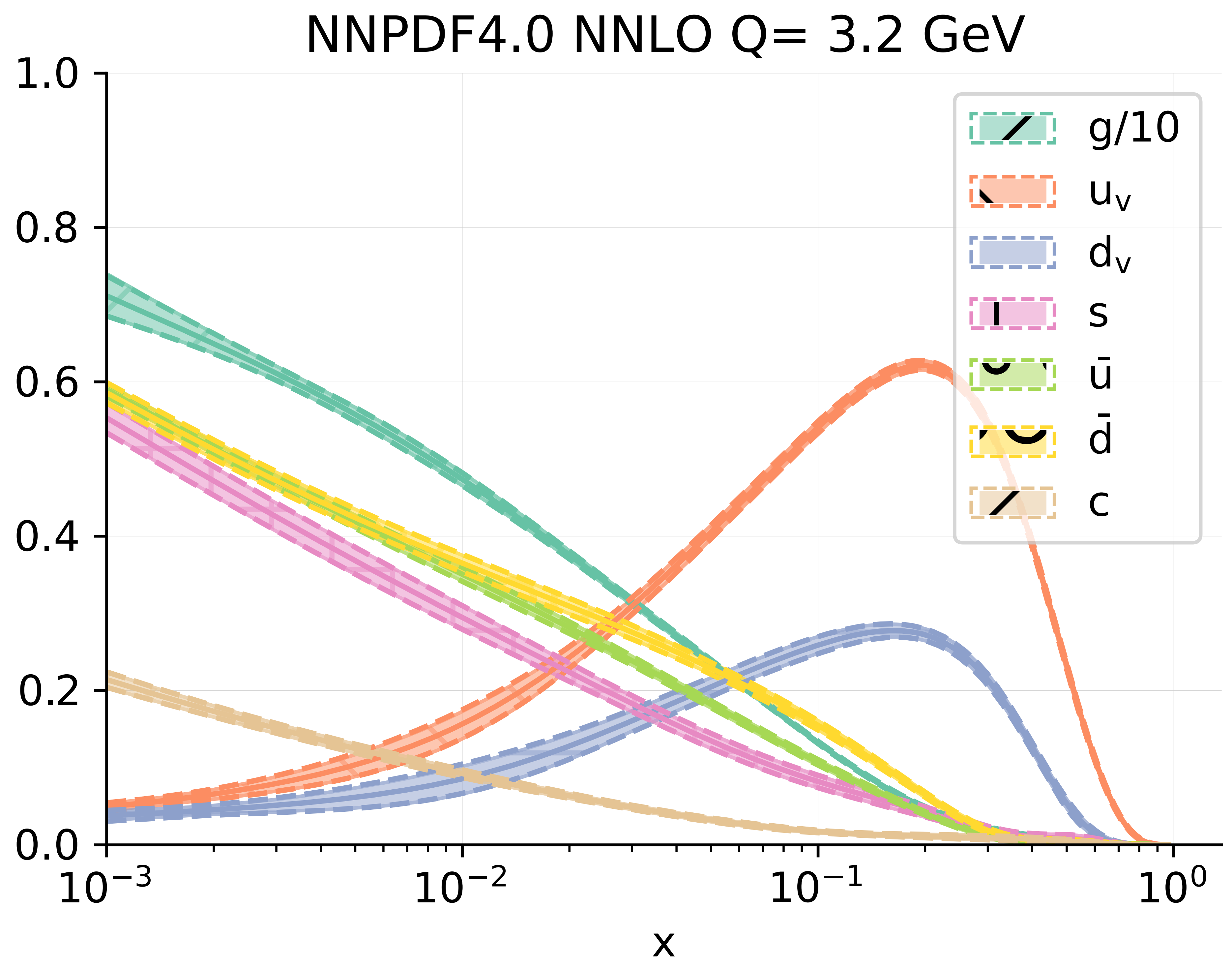
Parton distribution functions

- Describe the quark and gluon constituents of the proton
- Parametrised by Bjorken *x* and energy scale *Q*

$$f_q(x, Q^2)$$

$$\sigma = \int_0^1 dx_1 \int_0^1 dx_2 \sum_{q_1, q_2} f_{q_1}(x_1) f_{q_2}(x_2) \hat{\sigma}(x_1, x_2)$$

- Q^2 dependence determined by DGLAP evolution
- x dependence determined from fits to data



Ball et. al, NNPDF4.0, 2109.02653

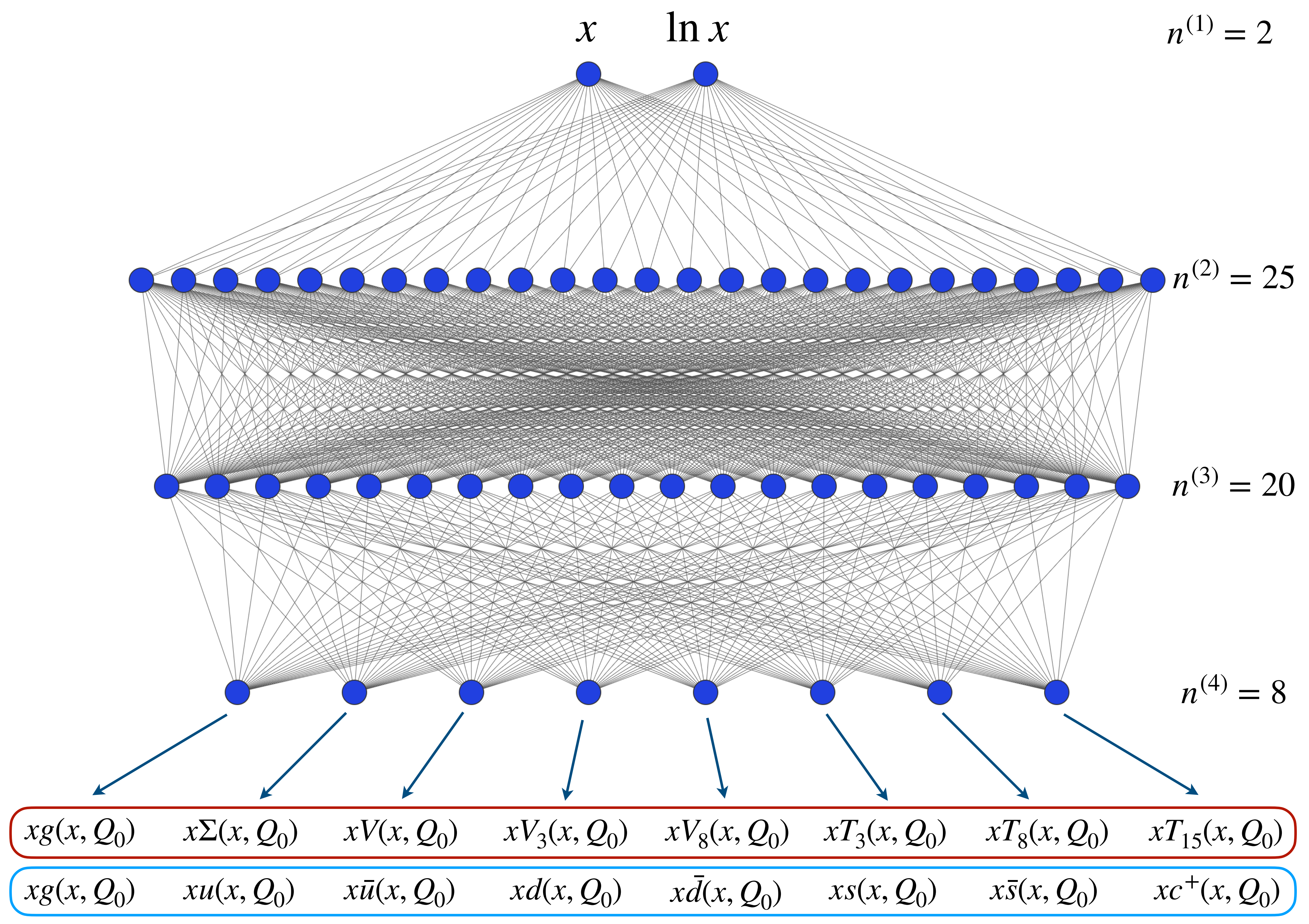
Parton distribution functions

- Describe the quark and gluon constituents of the proton
- Parametrised by Bjorken *x* and energy scale *Q*

$$f_q(x, Q^2)$$

$$\sigma = \int_0^1 dx_1 \int_0^1 dx_2 \sum_{q_1, q_2} f_{q_1}(x_1) f_{q_2}(x_2) \hat{\sigma}(x_1, x_2)$$

- Q^2 dependence determined by DGLAP evolution
- x dependence determined from fits to data
- **in NNPDF: parametrised by neural networks**



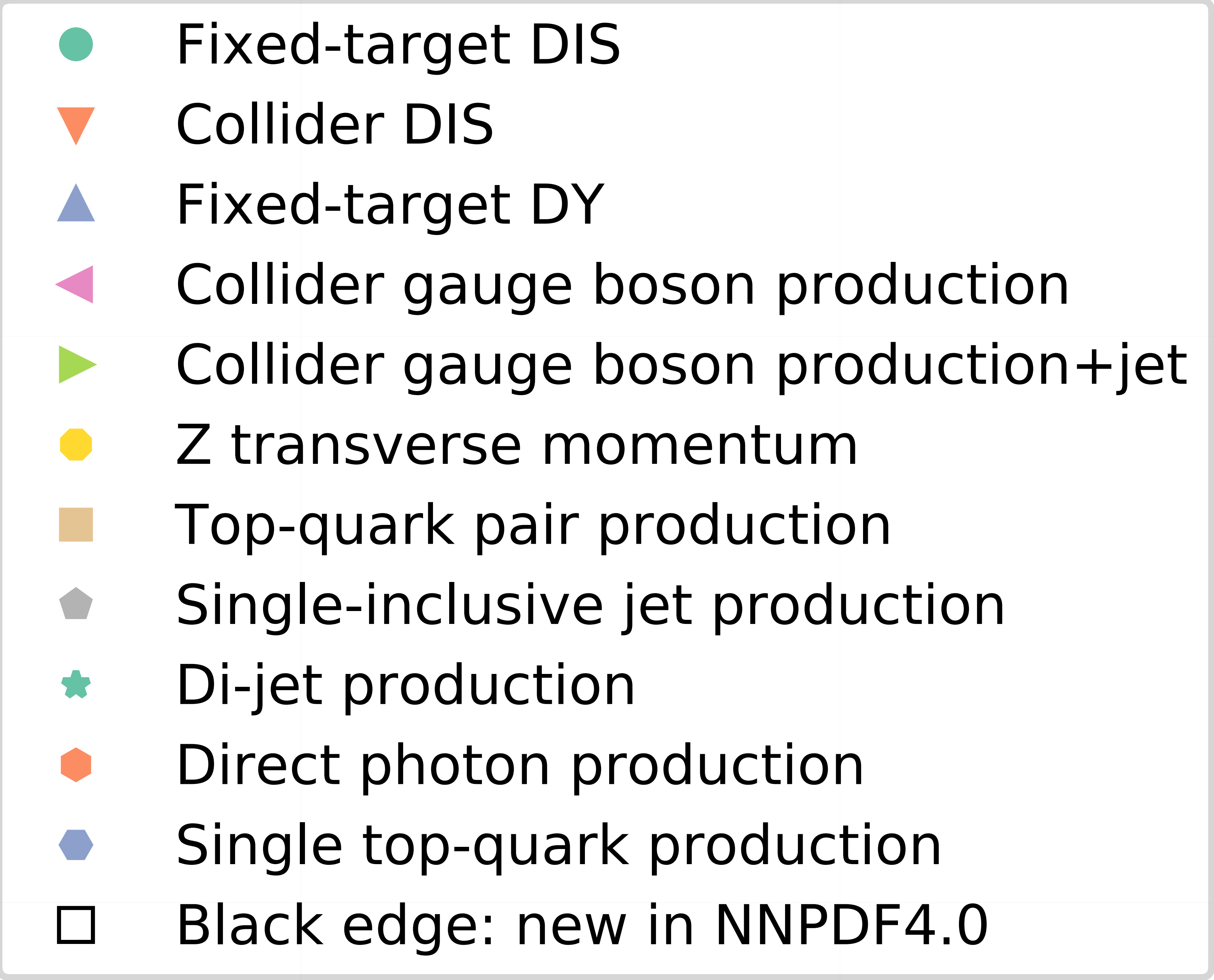
Ball et. al, NNPDF4.0, 2109.02653

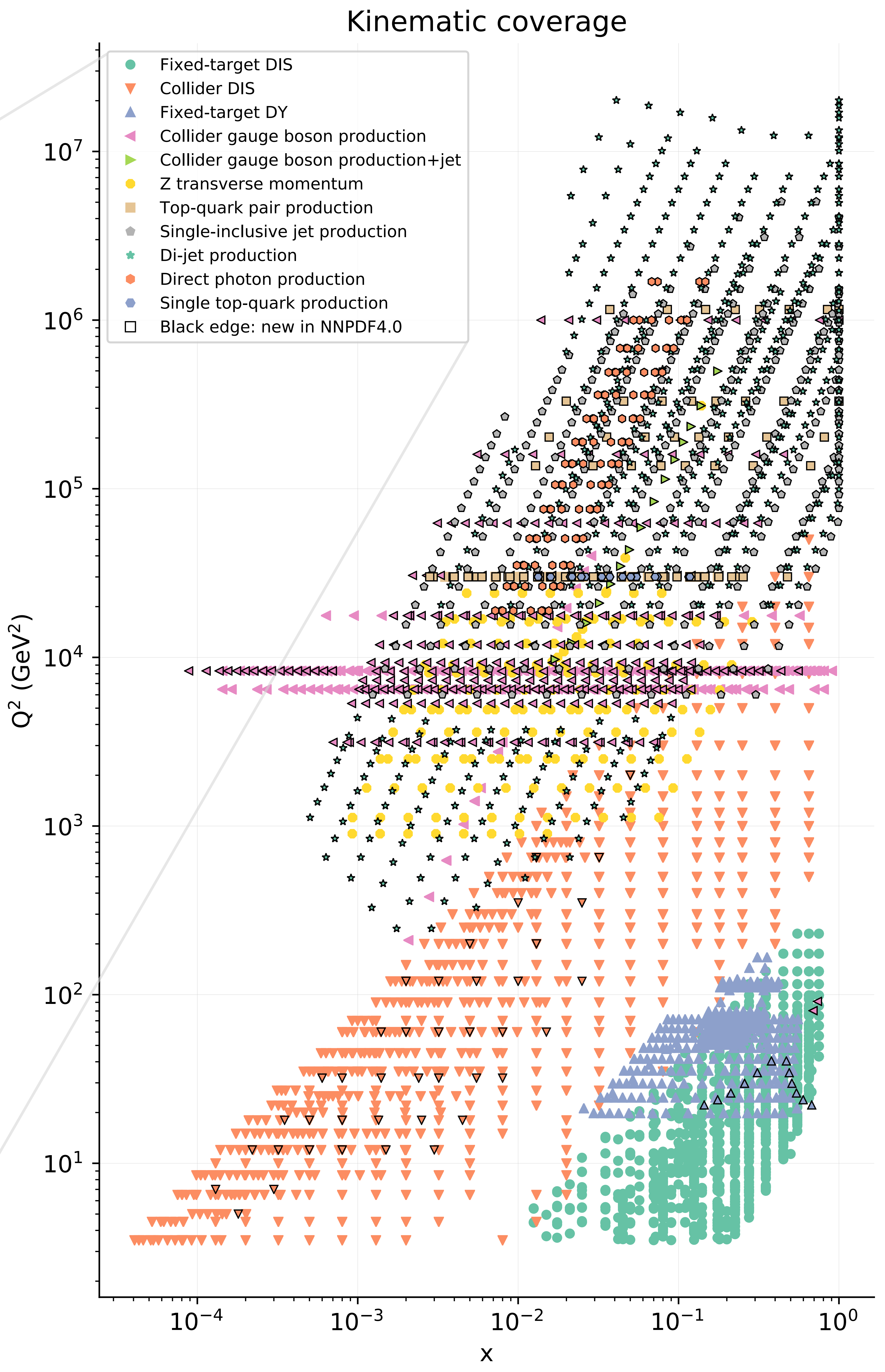
Data overlap

Often the data used in PDF fits are also used in EFT fits.

This overlap will grow as we continue to take a global approach to constraining the SMEFT.

Data included in NNPDF4.0, [2109.02653]:



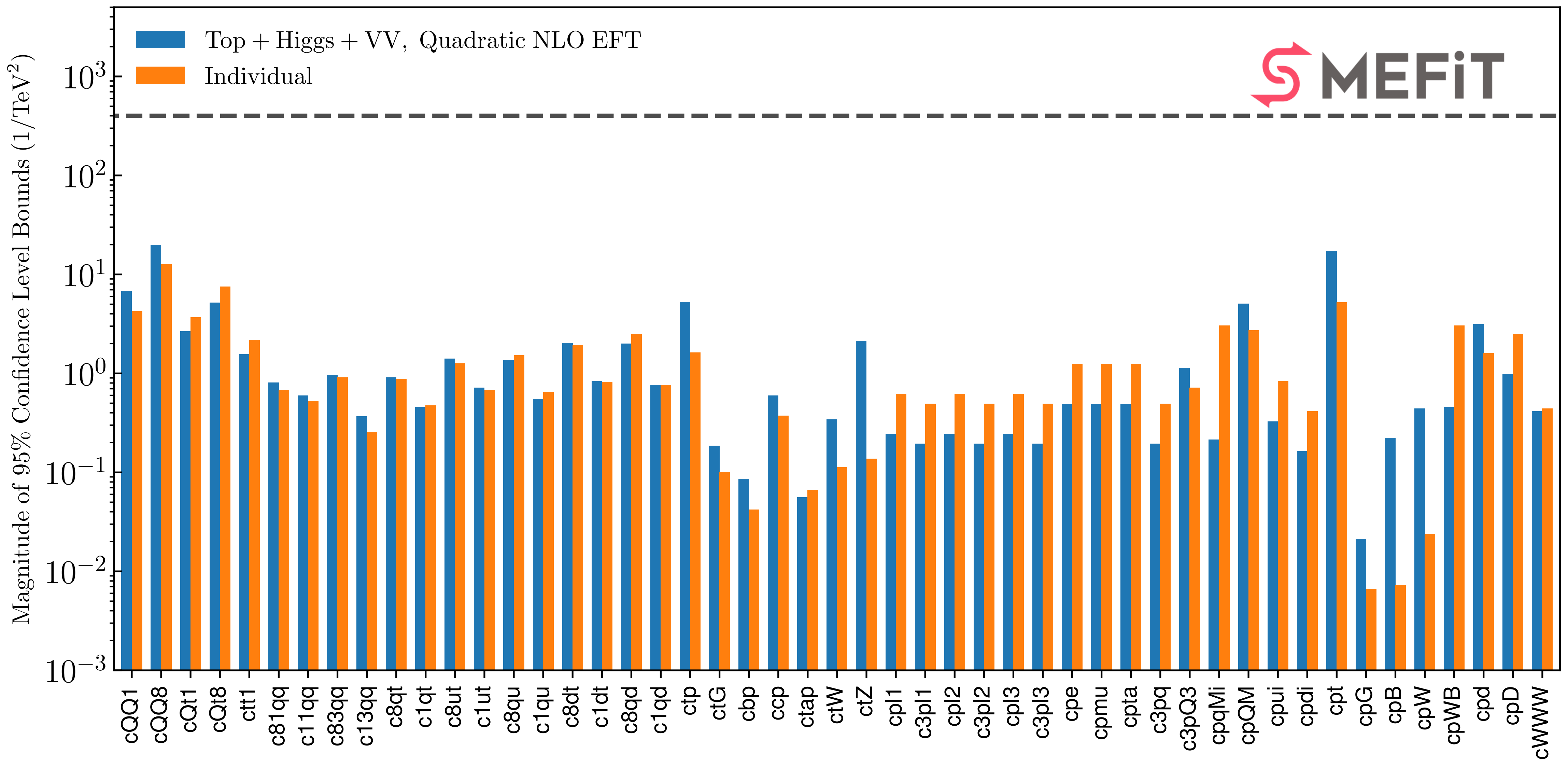


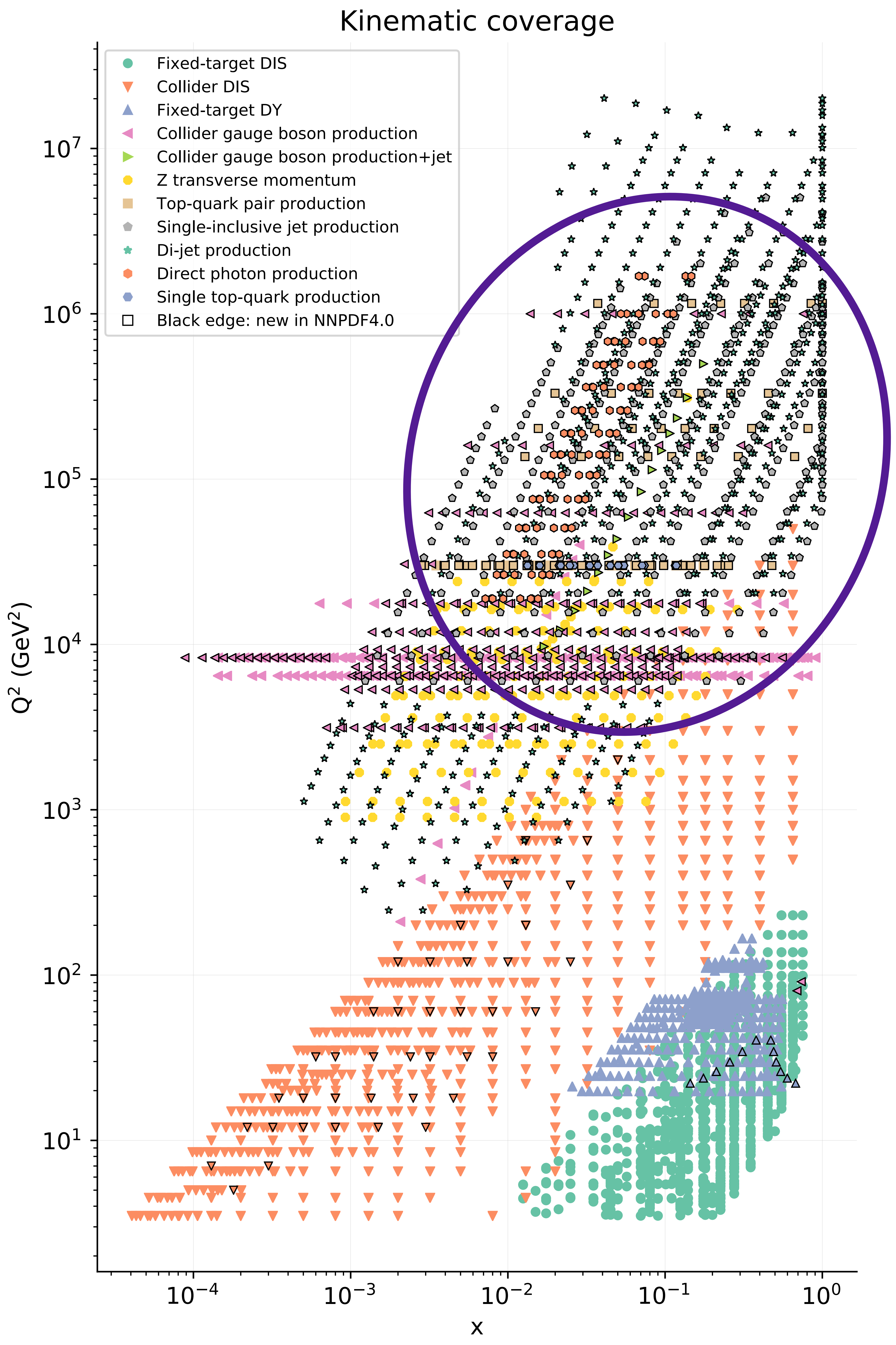
Data overlap

Often the data used in PDF fits are also used in EFT fits.

This overlap will grow as we continue to take a global approach to constraining the SMEFT.

- e.g. Top quark data used to fit the SMEFT in the global fit of *J. Ethier et. al, 2105.00006, SMEFiT*



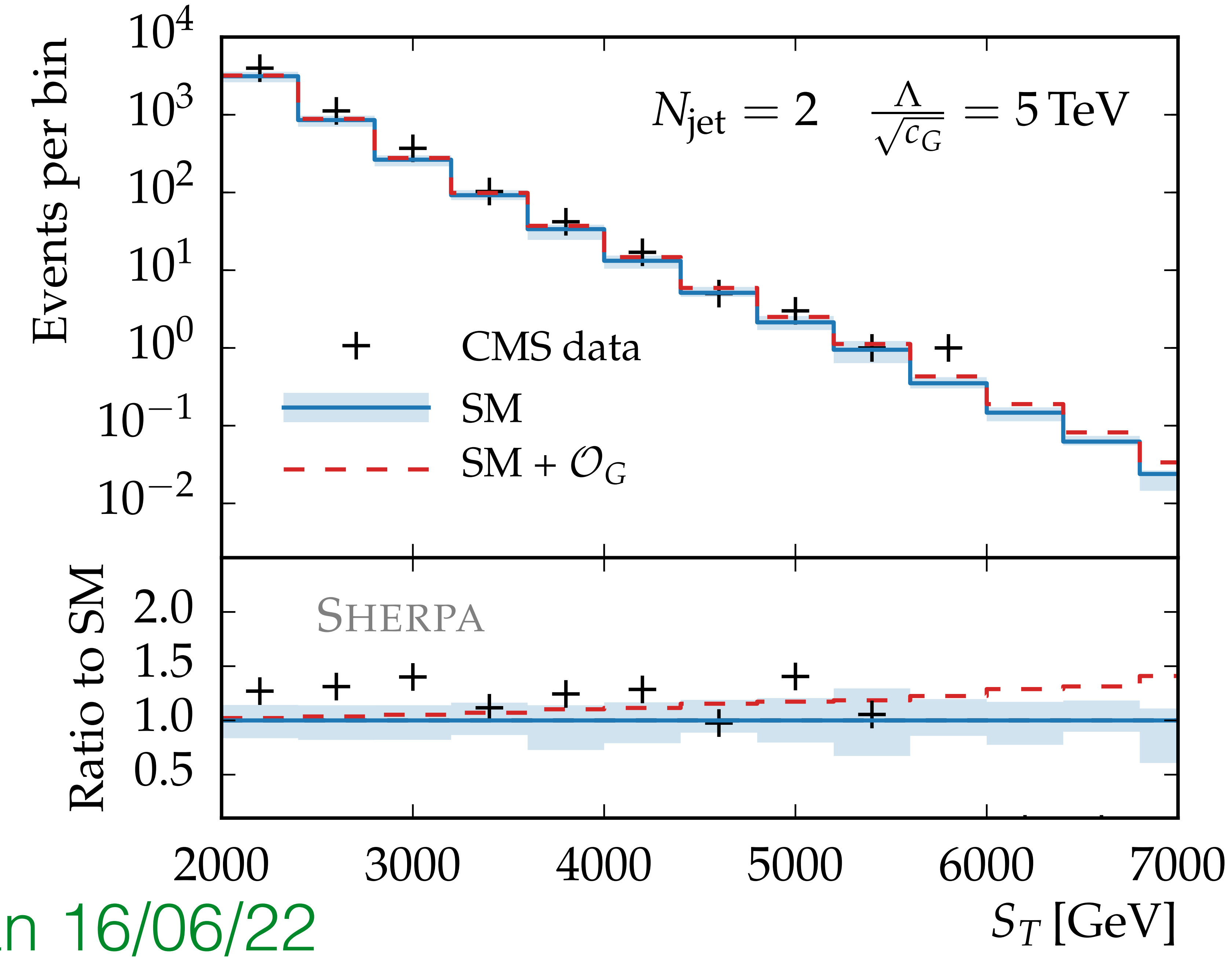


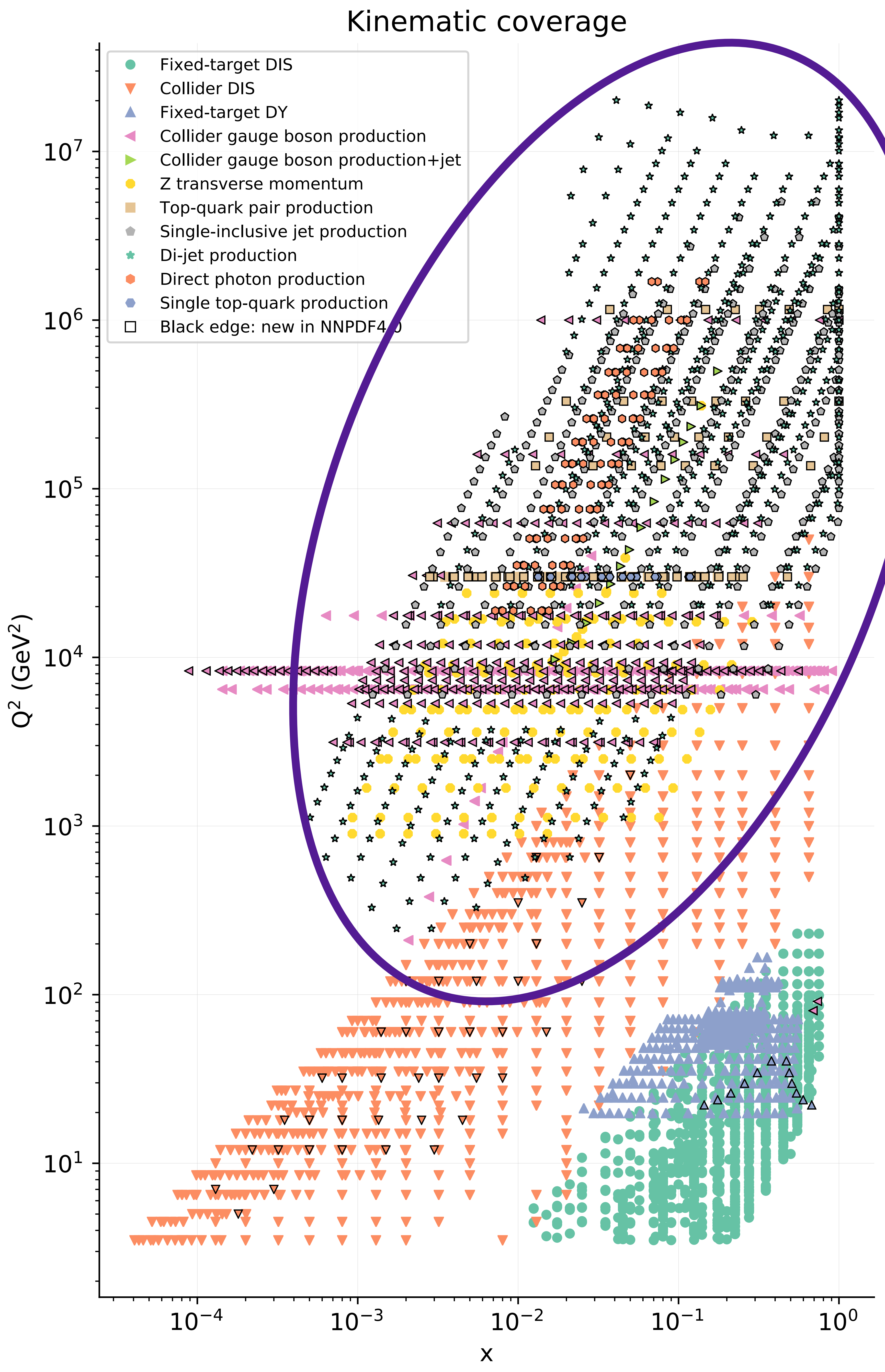
Data overlap

Often the data used in PDF fits are also used in EFT fits.

This overlap will grow as we continue to take a global approach to constraining the SMEFT.

- e.g. Dijet data used to fit the SMEFT operator $\mathcal{O}_G$ in *F. Krauss et. al, 1611.00767*



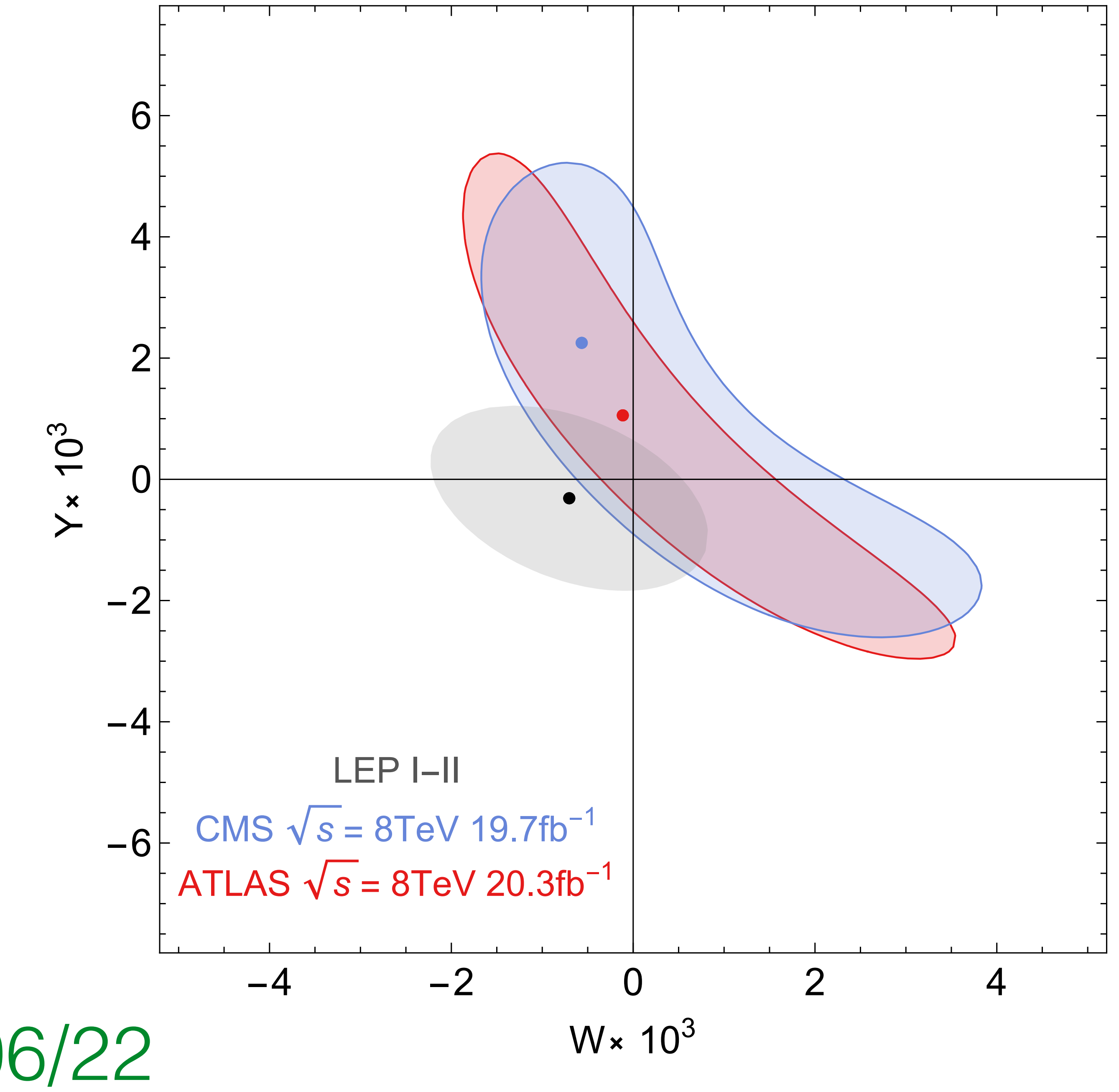


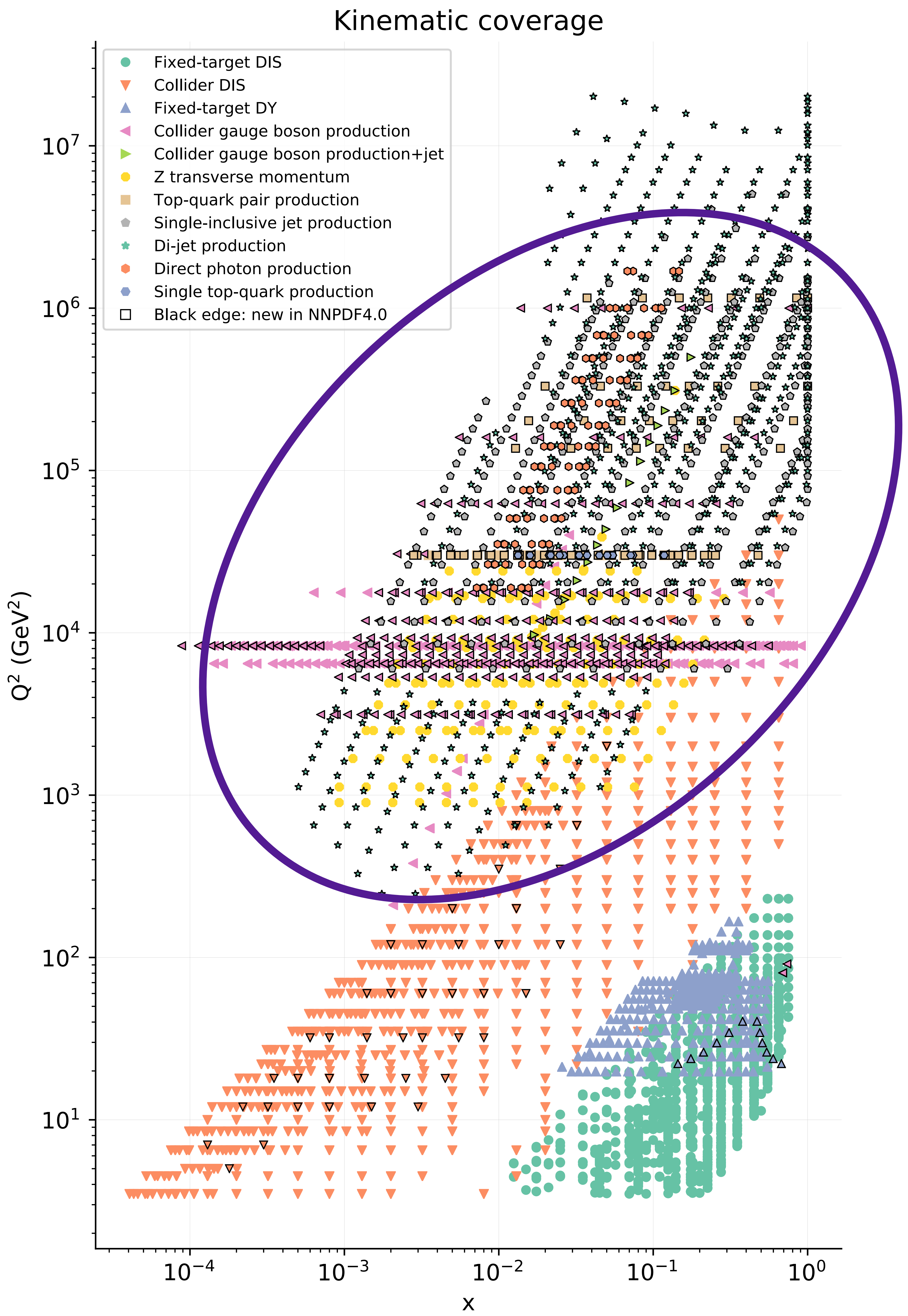
Data overlap

Often the data used in PDF fits are also used in EFT fits.

This overlap will grow as we continue to take a global approach to constraining the SMEFT.

- e.g. Drell-Yan data used to fit the SMEFT 4-fermion operators in *Farina et. al 1609.08157*





Theoretical inconsistencies

PDFs are an input to SMEFT fits: $\sigma_{\mathrm{SMEFT}}(C) = f_1 \otimes f_2 \otimes \hat{\sigma}_{\mathrm{SMEFT}}(C)$

But PDFs are found assuming the SM: $\sigma = f_1 \otimes f_2 \otimes \hat{\sigma}_{SM}$

'Standard Model PDFs'

How do the constraints on the **SMEFT** change if we perform a consistent joint determination of the PDFs and SMEFT?

How do the constraints on the **SMEFT** change if we perform a consistent joint determination of the PDFs and SMEFT?

How do the **PDFs** change if we perform a consistent joint determination of the PDFs and SMEFT?

How do the constraints on the **SMEFT** change if we perform a consistent joint determination of the PDFs and SMEFT?

How do the **PDFs** change if we perform a consistent joint determination of the PDFs and SMEFT?

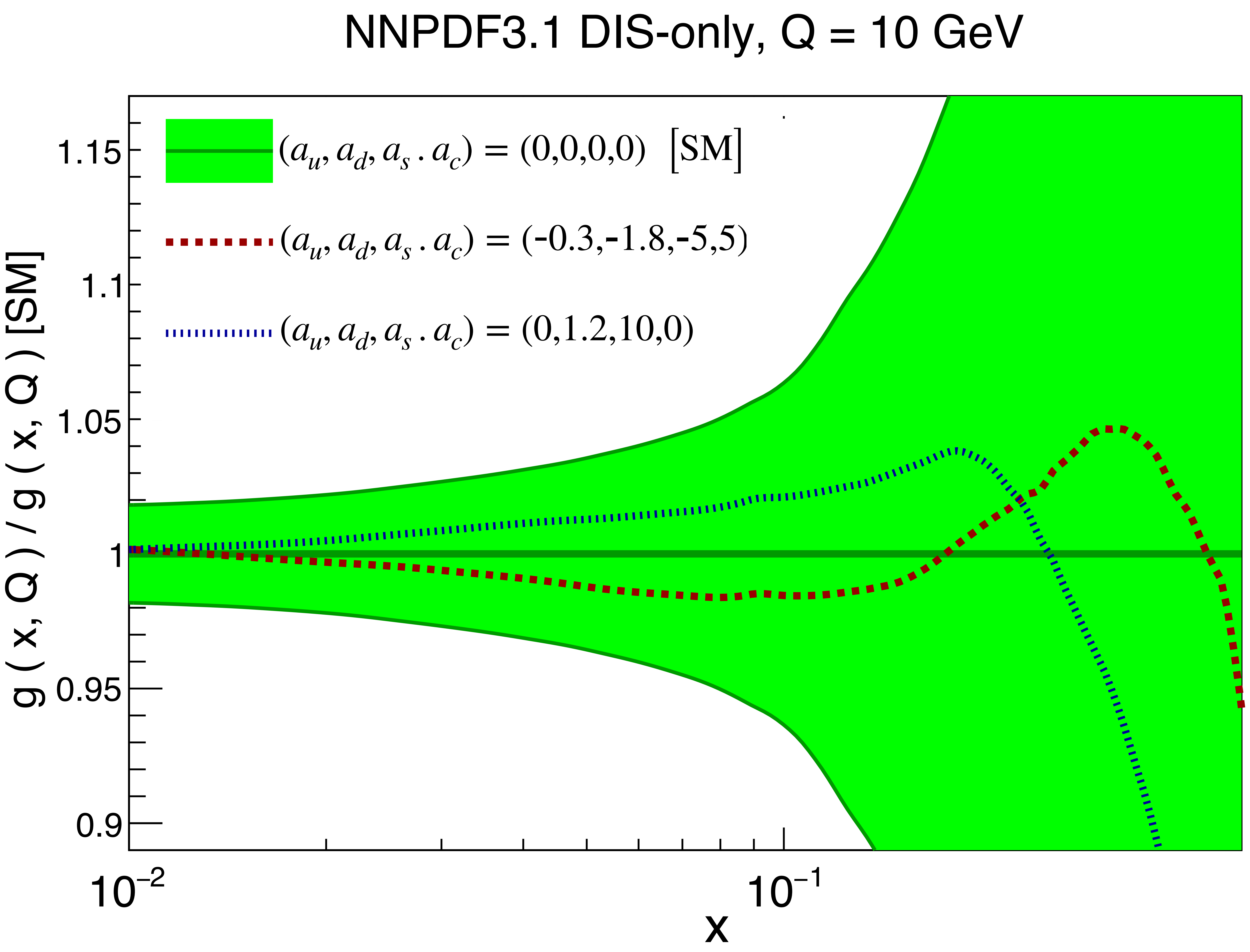
Could we be absorbing signs of new physics into the PDFs?

Can New Physics Hide Inside the Proton?

First studied with deep inelastic scattering data by Carrazza et al.: PRL 123 (2019) 13, 132001

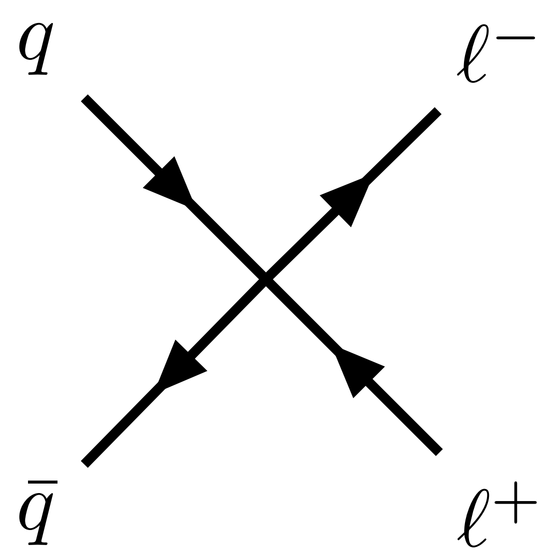
A *proof of concept* that disentangling PDF and SMEFT effects is possible

Used a simple χ^2 methodology, sampling across SMEFT parameter space



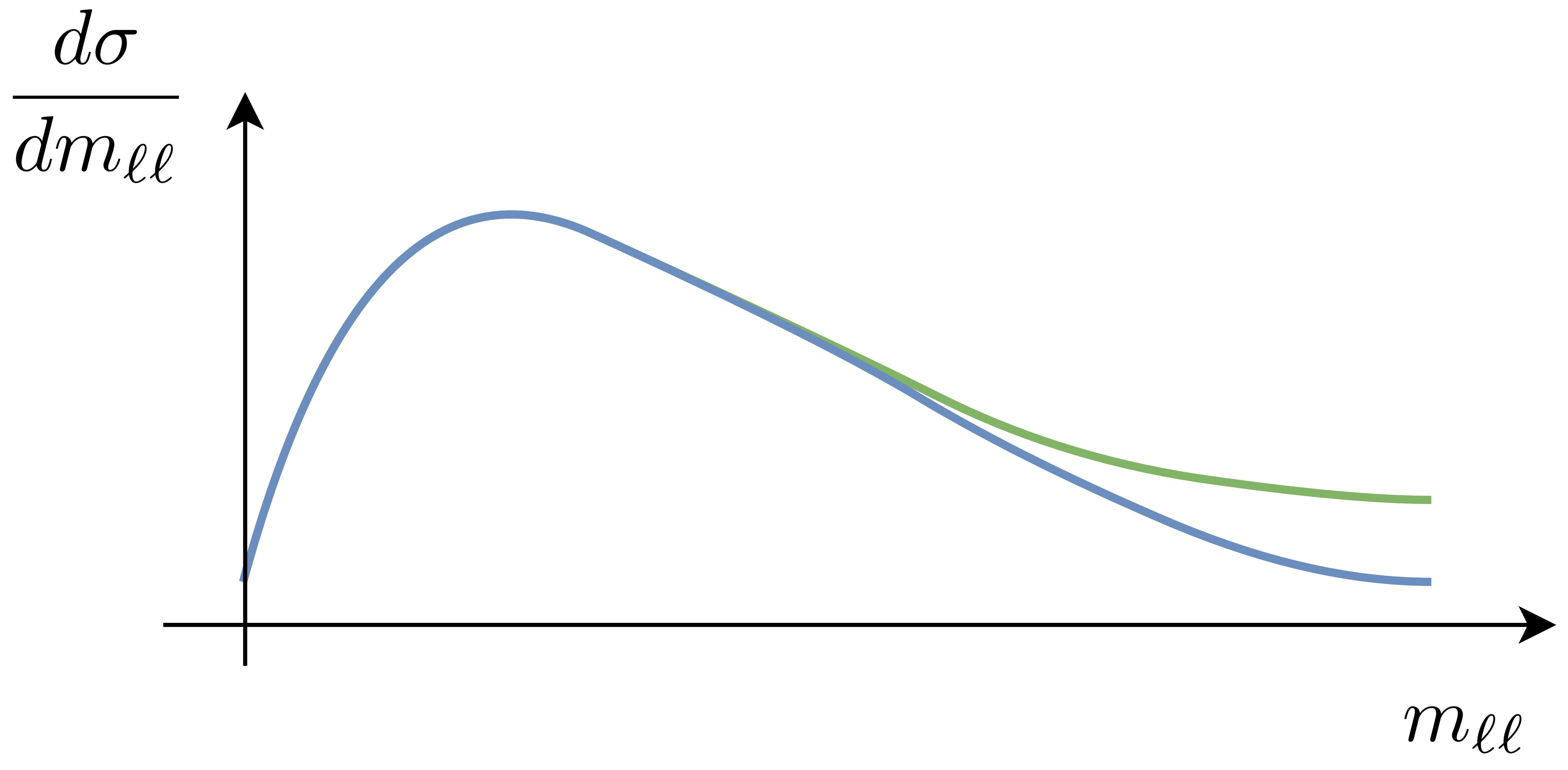
High-mass Drell-Yan tails

High-mass Drell-Yan measurements are used to probe 4-fermion operators.



$$\mathcal{A} \sim \mathcal{A}_{\rm SM} + C\frac{E^2}{\Lambda^2}$$

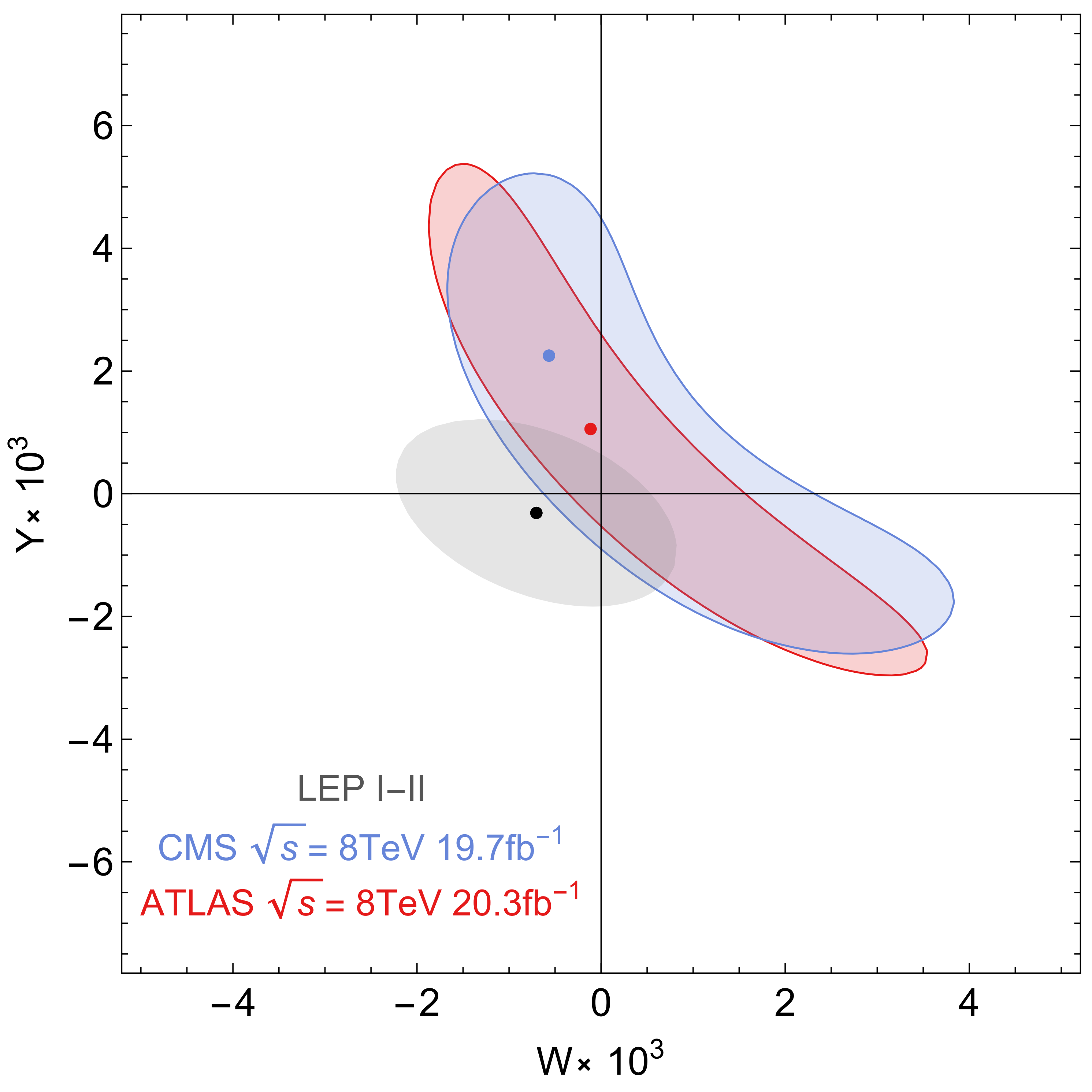
New physics in 4-fermion operators will manifest as a smooth distortion of the high-mass tail



High-mass Drell-Yan tails

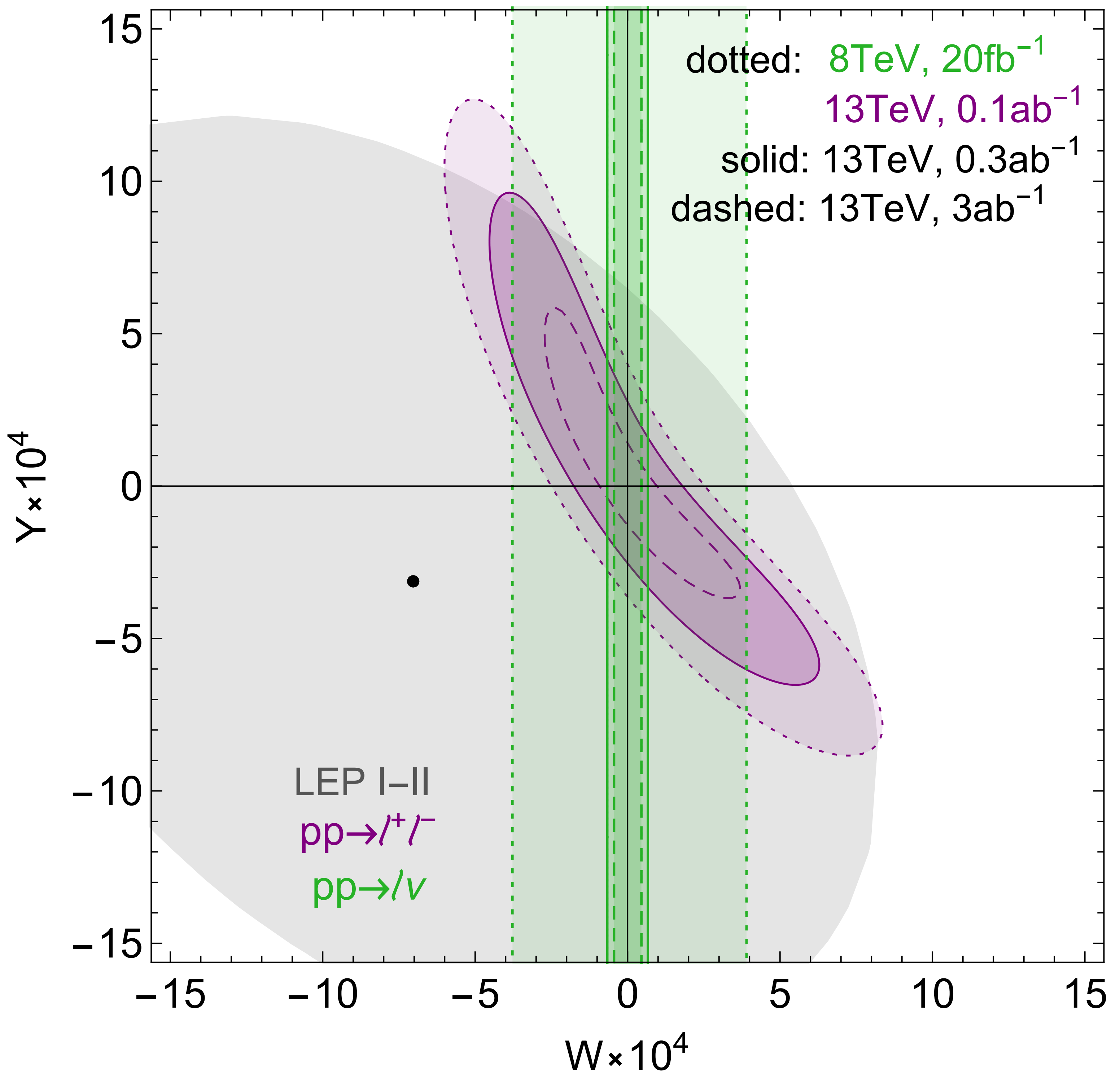
High-mass Drell-Yan measurements provide important constraints, despite statistical uncertainties in the tail:

Energy helps accuracy: *Farina et. al 1609.08157*



Constraints from Run I DY are almost competitive with LEP precision data

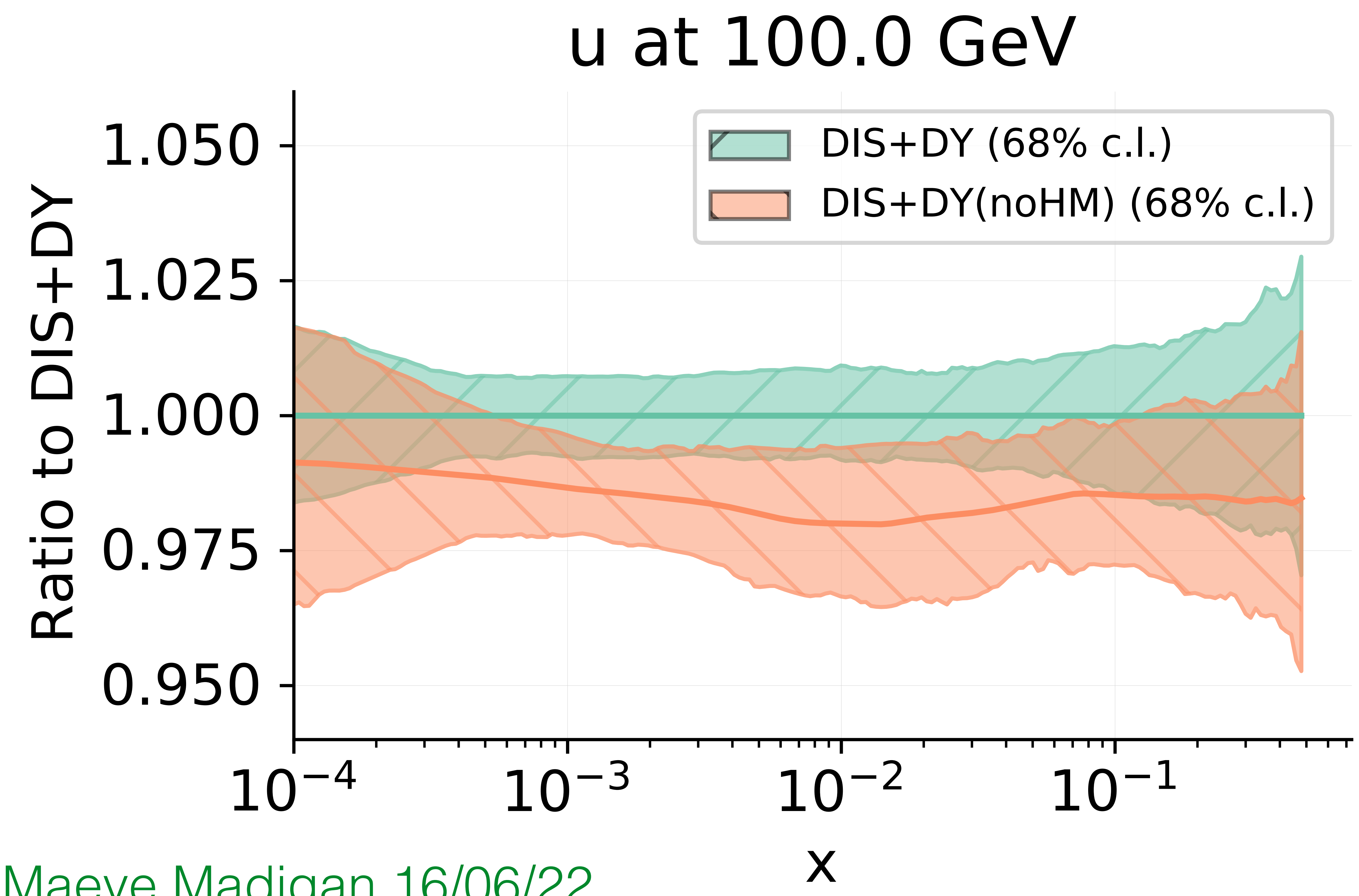
Projected constraints from the LHC at 8 and 13 TeV

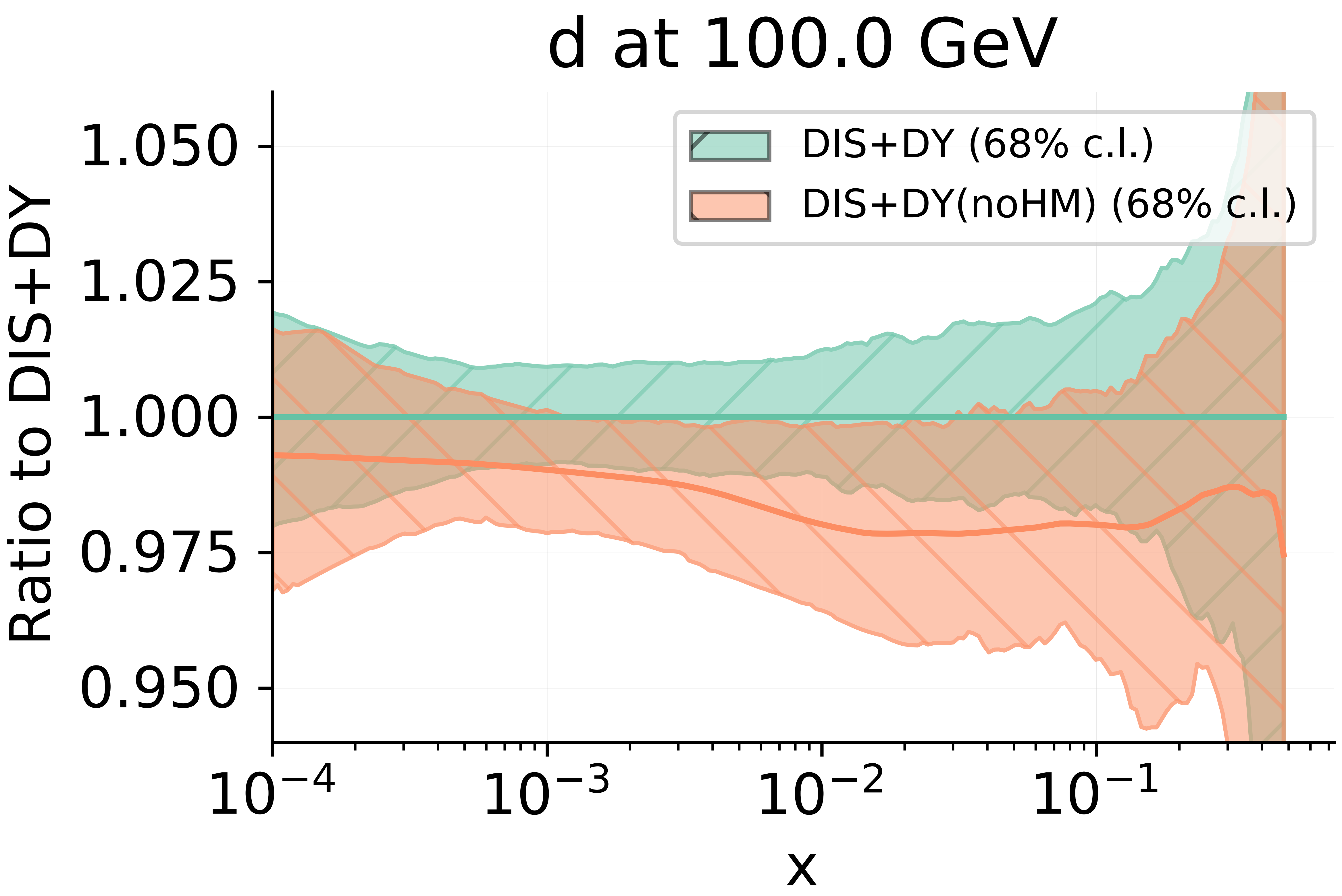


High-mass Drell-Yan tails

High-mass Drell-Yan measurements provide important constraints on the SMEFT, despite statistical uncertainties in the tail.

What about the PDFs?





SMEFT benchmarks

Operators to which DY is sensitive *include:*

$\mathcal{O}_{ld}$ $\mathcal{O}_{lu}$ $\mathcal{O}_{qe}$ $\mathcal{O}_{ed}$ $\mathcal{O}_{eu}$ $\mathcal{O}_{lq}^{(1)}$ $\mathcal{O}_{lq}^{(3)}$ energy-growing 4-fermion operators

$\mathcal{O}_{Hl}^{(3)}$ $\mathcal{O}_{Hl}^{(1)}$ $\mathcal{O}_{Hq}^{(1)}$ $\mathcal{O}_{Hq}^{(3)}$ $\mathcal{O}_{He}$ $\mathcal{O}_{Hu}$ $\mathcal{O}_{Hd}$ corrections to the $Z\bar{f}f$ vertex

+ others……

Our methodology limits us to 1-2 SMEFT parameters.

SMEFT benchmarks

Operators to which DY is sensitive *include:*

$\mathcal{O}_{ld}$ $\mathcal{O}_{lu}$ $\mathcal{O}_{qe}$ $\mathcal{O}_{ed}$ $\mathcal{O}_{eu}$ $\mathcal{O}_{lq}^{(1)}$ $\mathcal{O}_{lq}^{(3)}$ energy-growing 4-fermion operators

$\mathcal{O}_{Hl}^{(3)}$ $\mathcal{O}_{Hl}^{(1)}$ $\mathcal{O}_{Hq}^{(1)}$ $\mathcal{O}_{Hq}^{(3)}$ $\mathcal{O}_{He}$ $\mathcal{O}_{Hu}$ $\mathcal{O}_{Hd}$ corrections to the $Z\bar{f}f$ vertex

+ others……

Our methodology limits us to 1-2 SMEFT parameters.

SMEFT benchmarks

Our methodology limits us to 1-2 SMEFT parameters.

Electroweak oblique parameters $\hat{W}, \hat{Y}$

- generated in **flavour-universal** BSM models
- their effect on DY is equivalent to the effect of a **linear combination of 4-fermion operators**:

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} - \frac{g^2 \hat{W}}{4 m_W^2} \mathcal{O}_{lq}^{(3)} - \frac{g_Y^2 \hat{Y}}{m_W^2} \Big(Y_l Y_d \, \mathcal{O}_{ld} + Y_l Y_u \, \mathcal{O}_{lu} + Y_l Y_q \, \mathcal{O}_{lq}^{(1)} + Y_e Y_d \, \mathcal{O}_{ed} + Y_e Y_u \, \mathcal{O}_{eu} + Y_e Y_q \, \mathcal{O}_{qe} \Big)$$

See 2104.02723 for a flavourful benchmark

Data

- Deep inelastic scattering data (DIS) from NNPDF3.1
- Low-mass and on-shell Drell-Yan datasets from NNPDF3.1

Exp.	$\sqrt{s}$ (TeV)	Ref.	Observable	$n_{\rm dat}$
E886	0.8	[98]	$d\sigma^d_{\rm DY}/d\sigma^p_{\rm DY}$	15
E886	0.8	[99, 100]	$d\sigma^p_{\rm DY}/(dy\,dm_{\ell\ell})$	89
E605	0.04	[101]	$\sigma^p_{\rm DY}/(dx_F\,dm_{\ell\ell})$	85
CDF	1.96	[102]	$d\sigma_Z/dy_Z$	29
D0	1.96	[103]	$d\sigma_Z/dy_Z$	28
D0	1.96	[104]	$d\sigma_{W\to\mu\nu}/d\eta_\mu$ asy.	9
ATLAS	7	[105]	$d\sigma_W/d\eta_l, d\sigma_Z/dy_z$	30
ATLAS	7	[106]	$d\sigma_{Z\to e^+e^-}/dm_{e^+e^-}$	6
ATLAS	7	[107]	$d\sigma_W/d\eta_l, d\sigma_Z/dy_z$	61
ATLAS	7	[108]	$d\sigma_{W+c}/dy_c$	22
ATLAS	8	[109]	$d\sigma_Z/dp_T$	82
ATLAS	8	[110]	$d\sigma_{W+j}/dp_T$	32
CMS	7	[111]	$d\sigma_{W\to l\nu}/d\eta_\ell$ asy.	22
CMS	7	[112]	$d\sigma_{W+c}/dy_c$	5
CMS	7	[112]	$d\sigma_{W^++c}/d\sigma_{W^-+c}$	5
CMS	8	[113]	$d\sigma_Z/dp_T$	28
CMS	8	[114]	$d\sigma_{W\to\mu\nu}/d\eta_\mu$	22
CMS	13	[115]	$d\sigma_{W+c}/dy_c$	5
LHCb	7	[116]	$d\sigma_{Z\to\mu^+\mu^-}/dy_{\mu^+\mu^-}$	9
LHCb	7	[117]	$d\sigma_{W,Z}/d\eta$	29
LHCb	8	[118]	$d\sigma_{Z\to e^+e^-}/dy_{e^+e^-}$	17
LHCb	8	[119]	$d\sigma_{W,Z}/d\eta$	30
Total				**659**

See 2104.02723 for references

Data

- Deep inelastic scattering data (DIS) from NNPDF3.1
- Low-mass and on-shell Drell-Yan datasets from NNPDF3.1
- 5 additional high-mass Drell-Yan datasets

Exp.	$\sqrt{s}$ (TeV)	Ref.	$\mathcal{L}$ (fb^{-1})	Channel	1D/2D	$n_{\rm dat}$	$m_{\ell\ell}^{\rm max}$ (TeV)
ATLAS	7	[120]	4.9	e^-e^+	1D	13	[1.0, 1.5]
ATLAS (*)	8	[86]	20.3	$\ell^-\ell^+$	2D	46	[0.5, 1.5]
CMS	7	[121]	9.3	$\mu^-\mu^+$	2D	127	[0.2, 1.5]
CMS (*)	8	[87]	19.7	$\ell^-\ell^+$	1D	41	[1.5, 2.0]
CMS (*)	13	[122]	5.1	e^-e^+, $\mu^-\mu^+$ $\ell^-\ell^+$	1D	43, 43 43	[1.5, 3.0]
Total						**270 (313)**	

See 2104.02723 for references

Theoretical predictions

SM predictions are calculated at **NNLO QCD** and **NLO EW.**

- we use `MadGraph5_aMC@NLO` interfaced with `APPLgrids` via `aMCfast` to generate predictions at NLO in QCD

- we apply k-factors:

$$K_{\mathrm{QCD}} = \left(\mathcal{L}_{ij}^{\mathrm{NNLO}} \otimes d\widehat{\sigma}_{ij}\big|_{\mathrm{NNLO\,QCD}}\right) / \left(\mathcal{L}_{ij}^{\mathrm{NNLO}} \otimes d\widehat{\sigma}_{ij}\big|_{\mathrm{NLO\,QCD}}\right)$$

$$K_{\mathrm{EW}} = \left(\mathcal{L}_{ij}^{\mathrm{NNLO}} \otimes d\widehat{\sigma}_{ij}\big|_{\mathrm{NLO\ \ QCD+EW}}\right) / \left(\mathcal{L}_{ij}^{\mathrm{NNLO}} \otimes d\widehat{\sigma}_{ij}\big|_{\mathrm{NLO\,QCD}}\right)$$

Theoretical predictions

SM predictions are calculated at **NNLO QCD** and **NLO EW.**

- we use `MadGraph5_aMC@NLO` interfaced with `APPLgrids` via `aMCfast` to generate predictions at NLO in QCD

- we apply k-factors:

$$K_{\rm QCD} = \left(\mathcal{L}_{ij}^{\rm NNLO} \otimes d\widehat{\sigma}_{ij}\big|_{\rm NNLO\,QCD}\right) / \left(\mathcal{L}_{ij}^{\rm NNLO} \otimes d\widehat{\sigma}_{ij}\big|_{\rm NLO\,QCD}\right)$$

$$K_{\rm EW} = \left(\mathcal{L}_{ij}^{\rm NNLO} \otimes d\widehat{\sigma}_{ij}\big|_{\rm NLO\;\;QCD+EW}\right) / \left(\mathcal{L}_{ij}^{\rm NNLO} \otimes d\widehat{\sigma}_{ij}\big|_{\rm NLO\,QCD}\right)$$

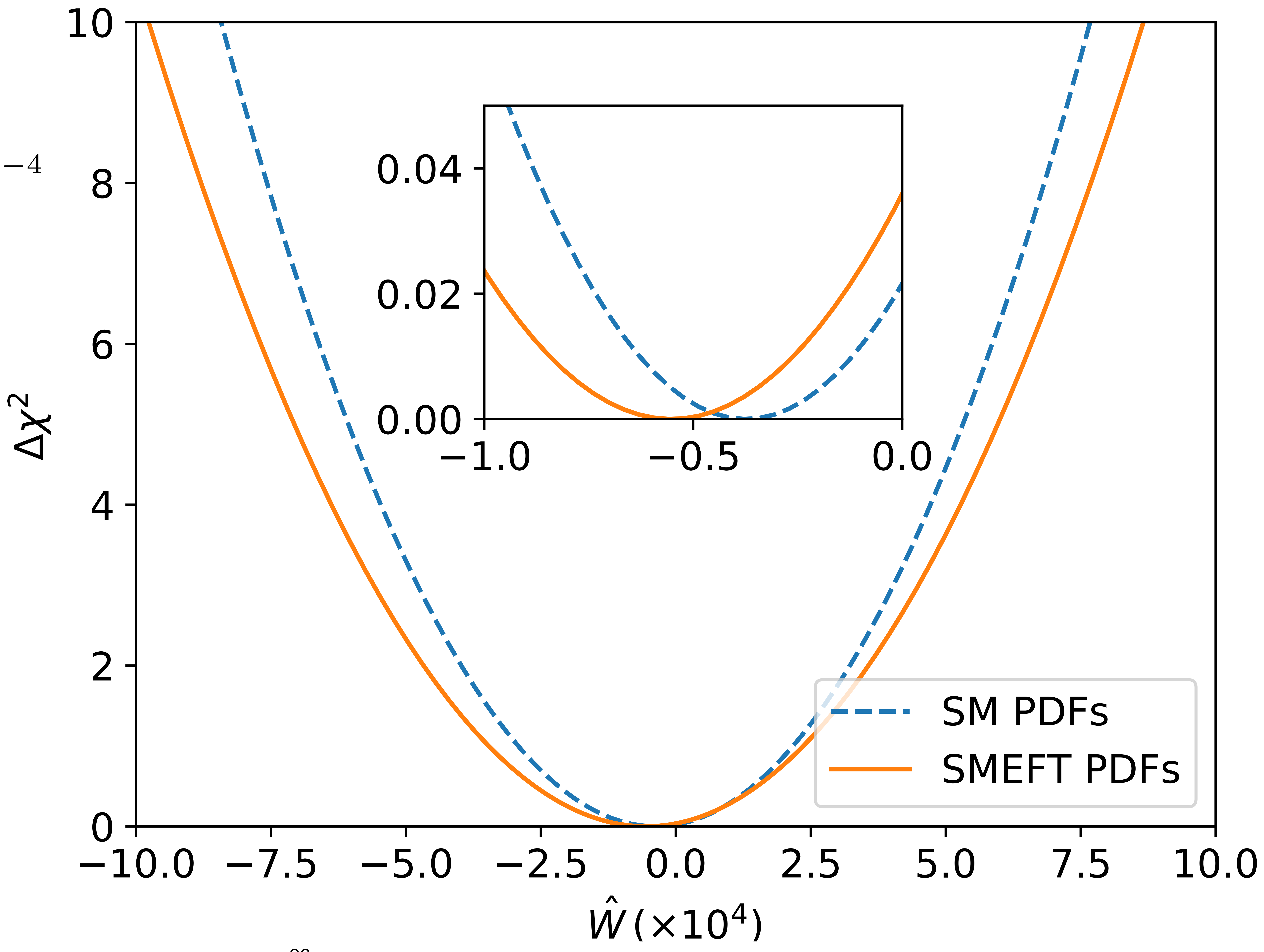
We calculate SMEFT predictions by applying k-factors to the SM predictions calculated at LO in QCD, EW:

$$K_{\rm SMEFT} = 1 + \hat{W} K_{\hat{W}} + \hat{Y} K_{\hat{Y}}$$

Results: $\hat{W}$

Best-fit shifts by $\delta\hat{W} = -0.2 \times 10^{-4}$

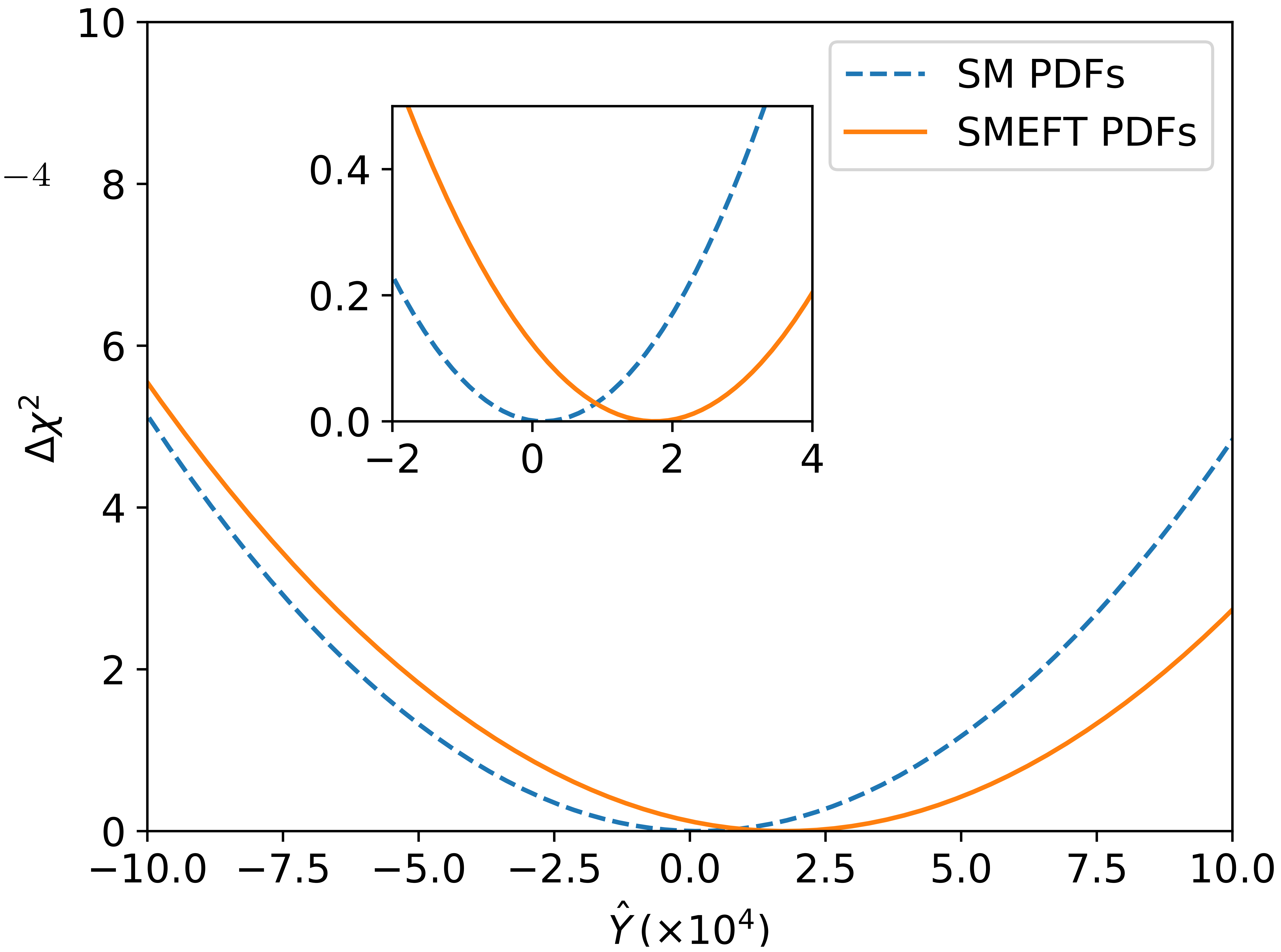
Parabola broadens by 15%



Results: $\hat{Y}$

Best-fit shifts by $\delta\hat{Y} = 1.6 \times 10^{-4}$

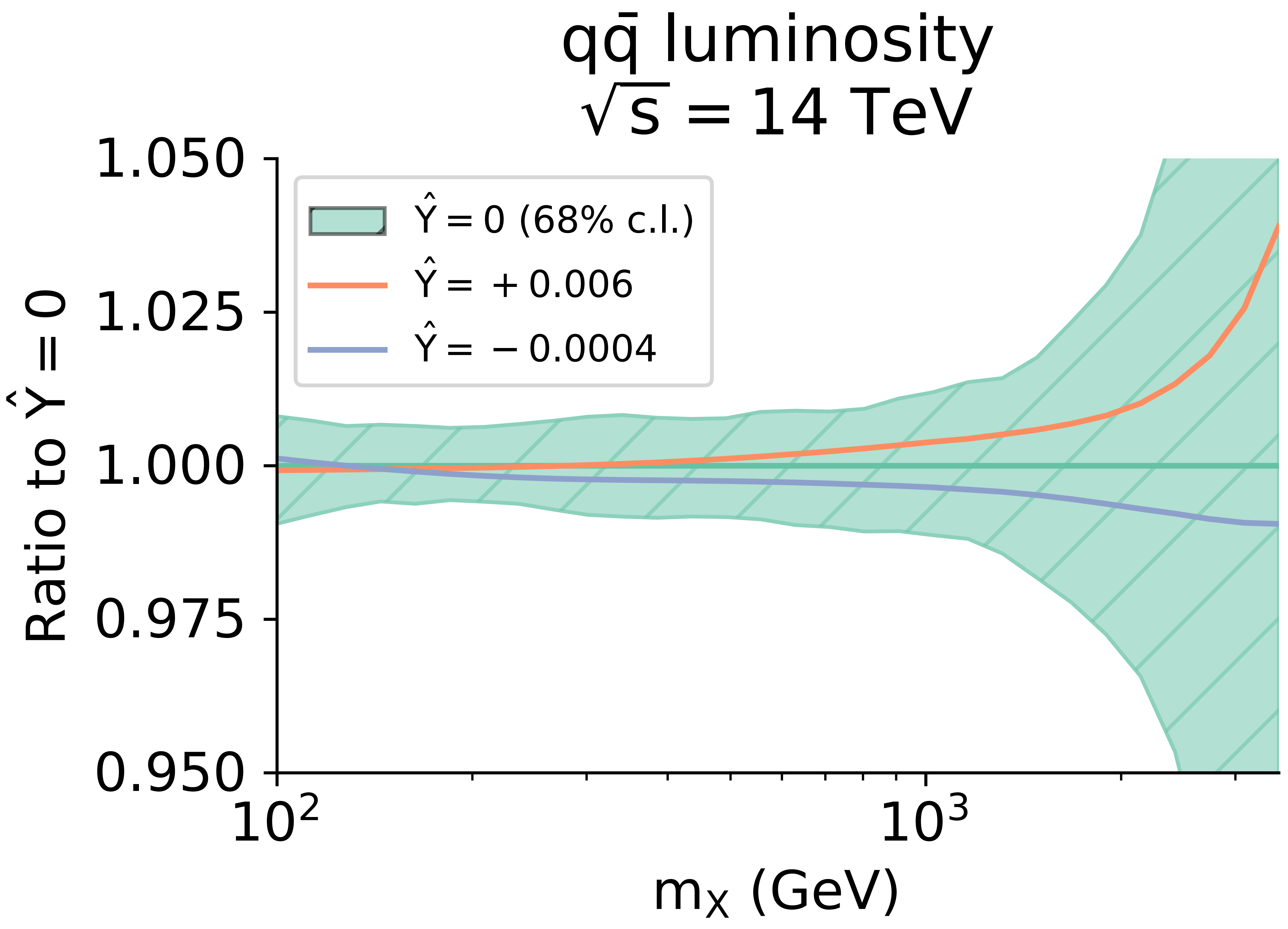
Parabola broadens by 12%

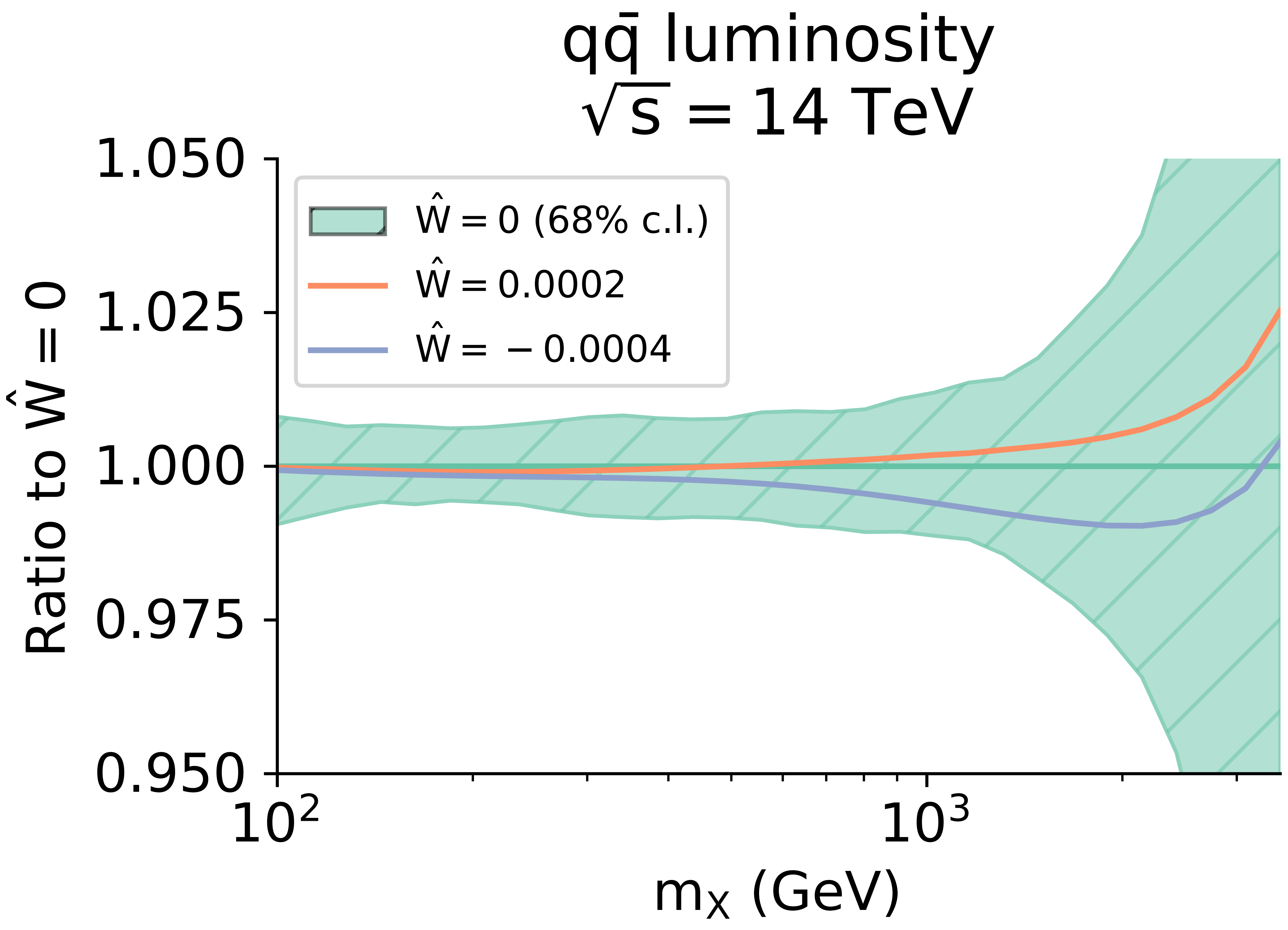


SMEFT PDFs

$$\mathcal{L}_{q\bar{q}} = \sum_{q,\bar{q}} \int_\tau^1 \frac{dx}{x} f_q(x) f_{\bar{q}}(\tau/x)$$

We see a moderate shift of the PDF central values, and no change to the PDF uncertainties.





How do the **SMEFT constraints** and **PDFs** change if we perform a consistent joint determination of the PDFs and SMEFT **at the HL-LHC?**

We expect HL-LHC measurement of high-mass DY to offer much **higher precision.**

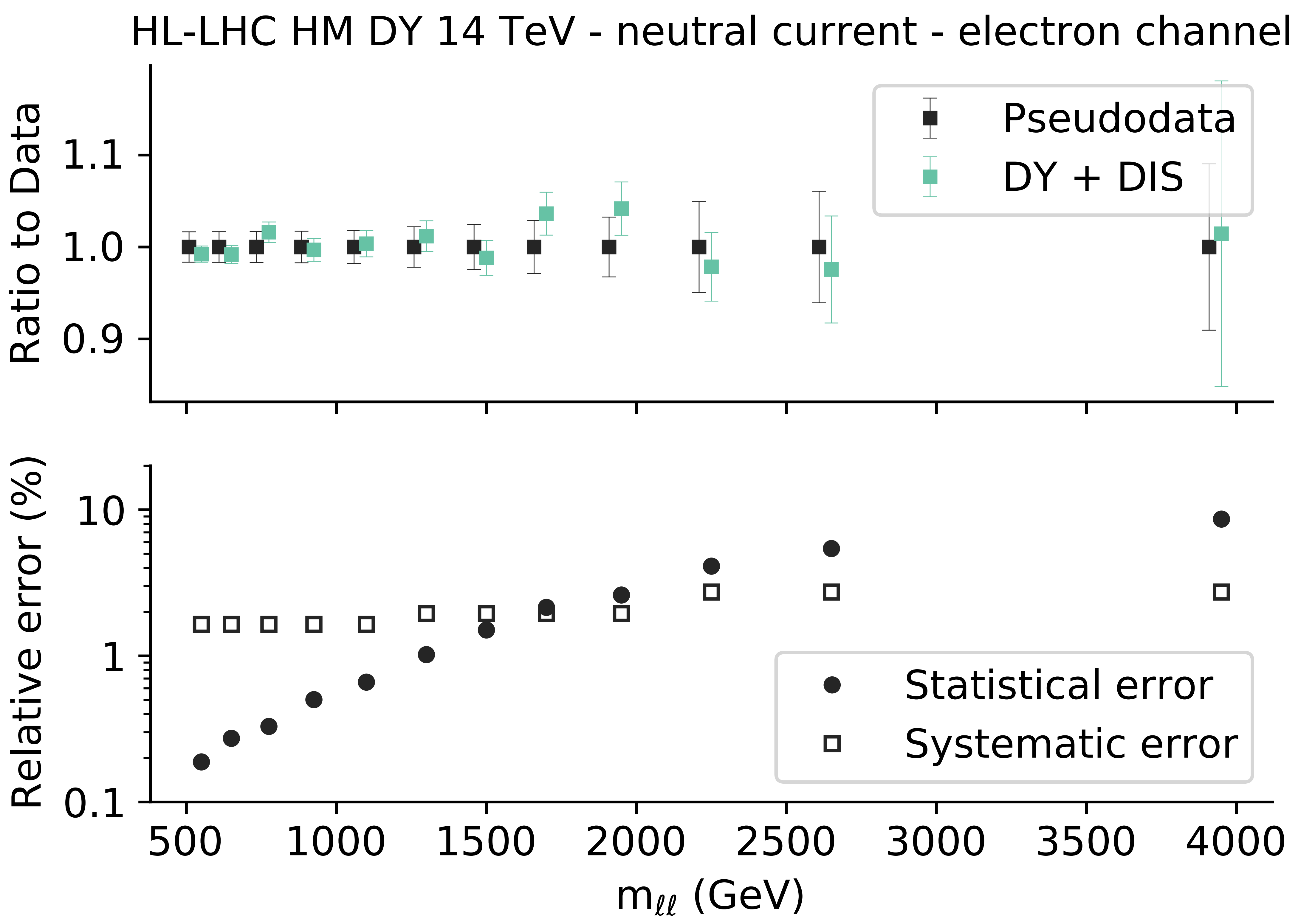
- reduction in PDF uncertainties *(see projections in 1810.03639)*
- reduction in SMEFT constraints

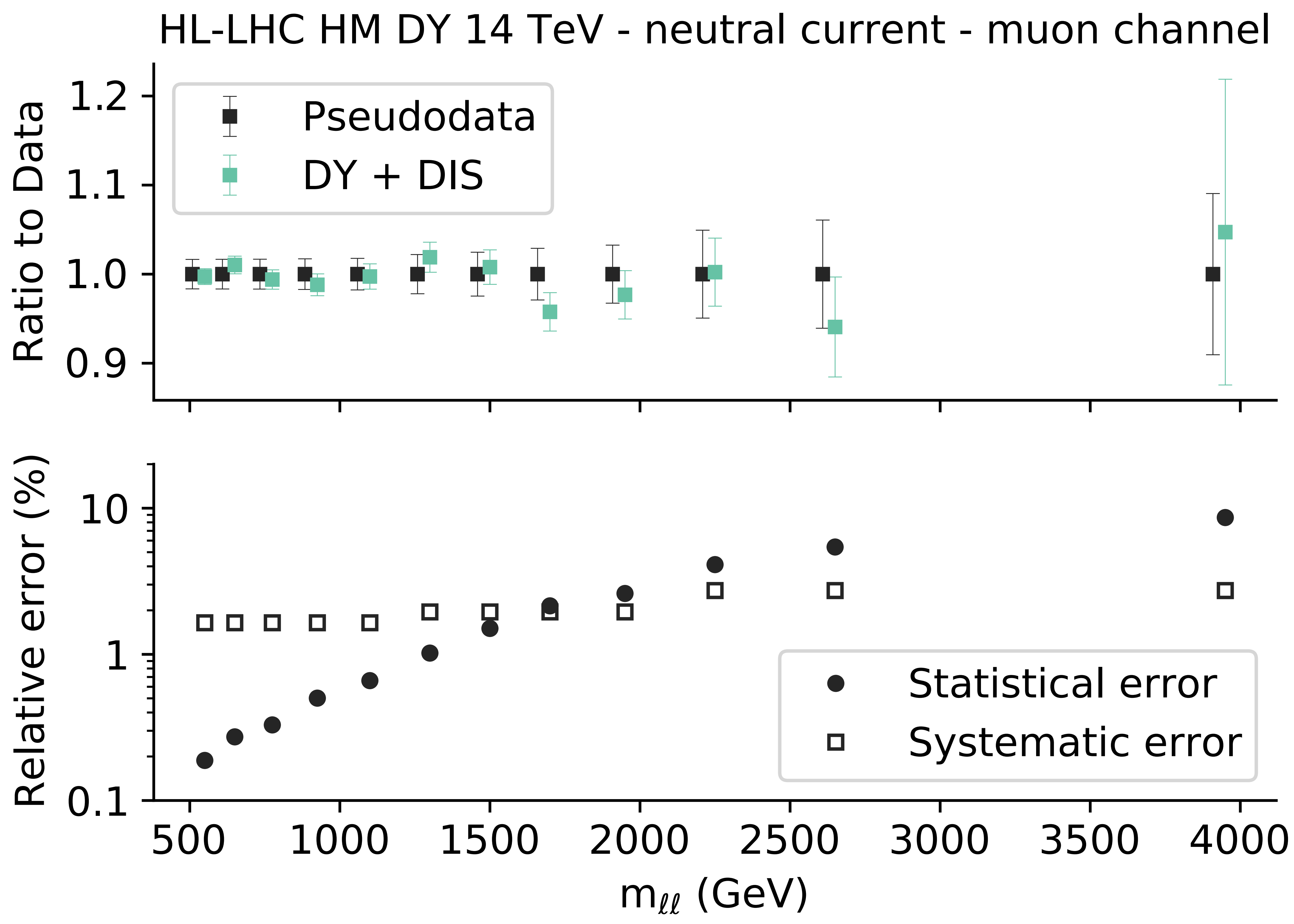
Next, we create projections for the interplay between PDFs and the SMEFT at the HL-LHC.

HL-LHC projections

- We produce pseudodata under the assumption of the SM for neutral and charged current DY
- We restrict to $m_{\ell\ell} > 500\ \mathrm{GeV}$ to avoid the systematics-dominated region
- Acceptance cuts and systematics uncertainties are modelled on reference measurements from ATLAS and CMS

$$\mathcal{L} = 6\ \mathrm{ab}^{-1} \quad \sqrt{s} = 14\ \mathrm{TeV}$$



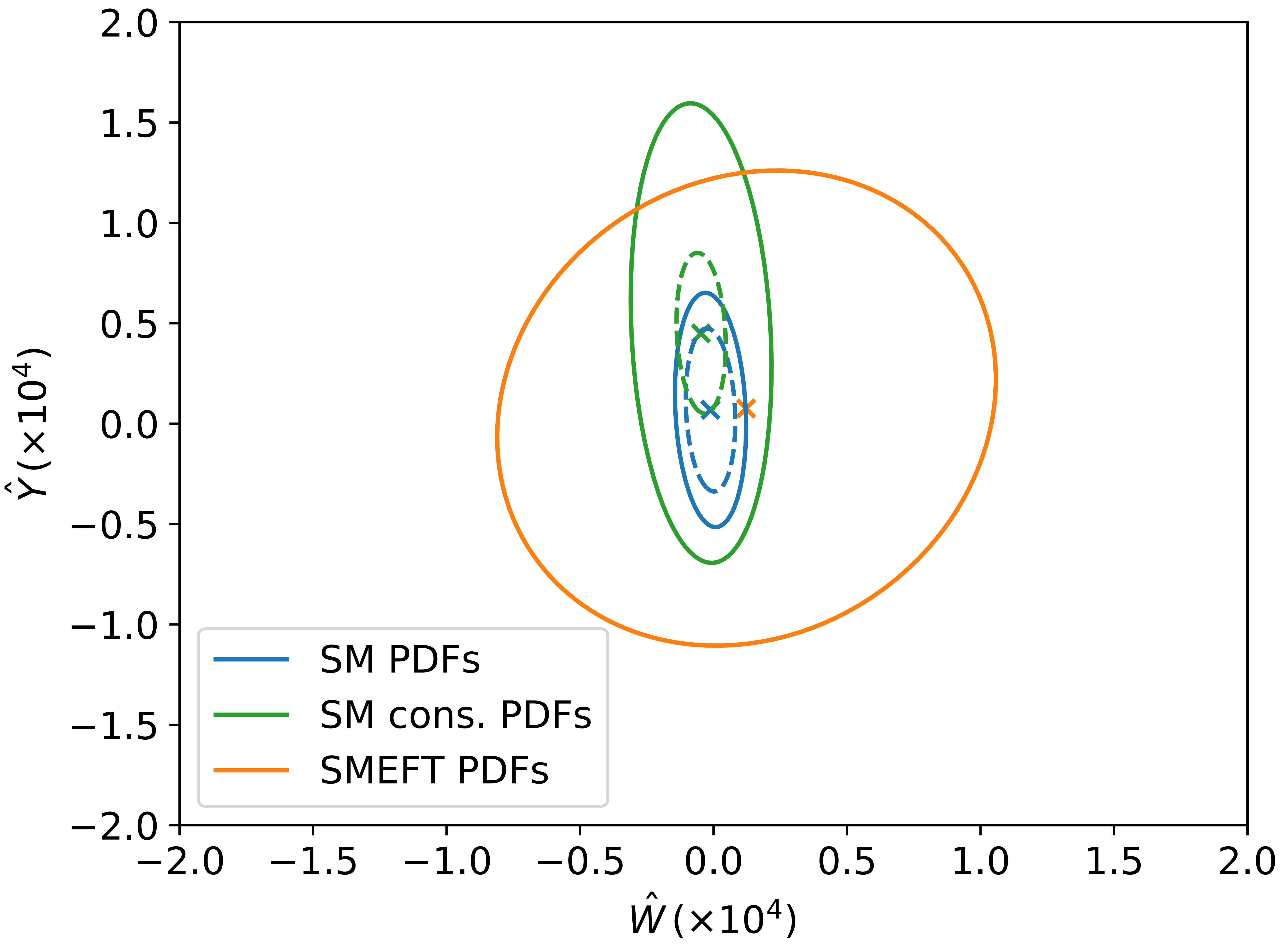


HL-LHC projections

Comparing blue to orange:

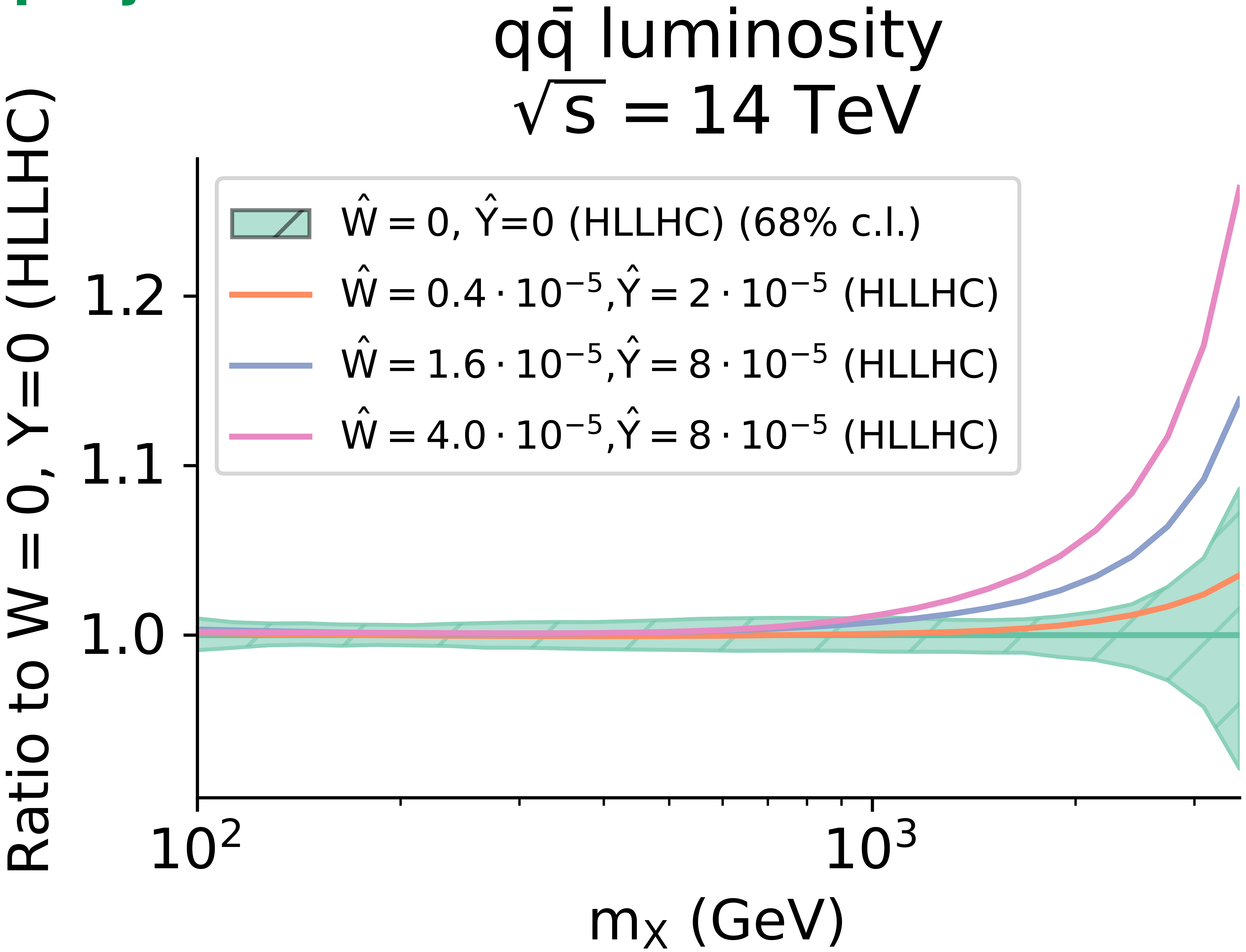
- neglecting the interplay between PDFs and the SMEFT leads to a **significant overestimate of the EFT constraints.**

dashed: not including PDF uncertainties



HL-LHC projections

$$\mathcal{L}_{q\bar{q}} = \sum_{q,\bar{q}} \int_\tau^1 \frac{dx}{x} f_q(x) f_{\bar{q}}(\tau/x)$$

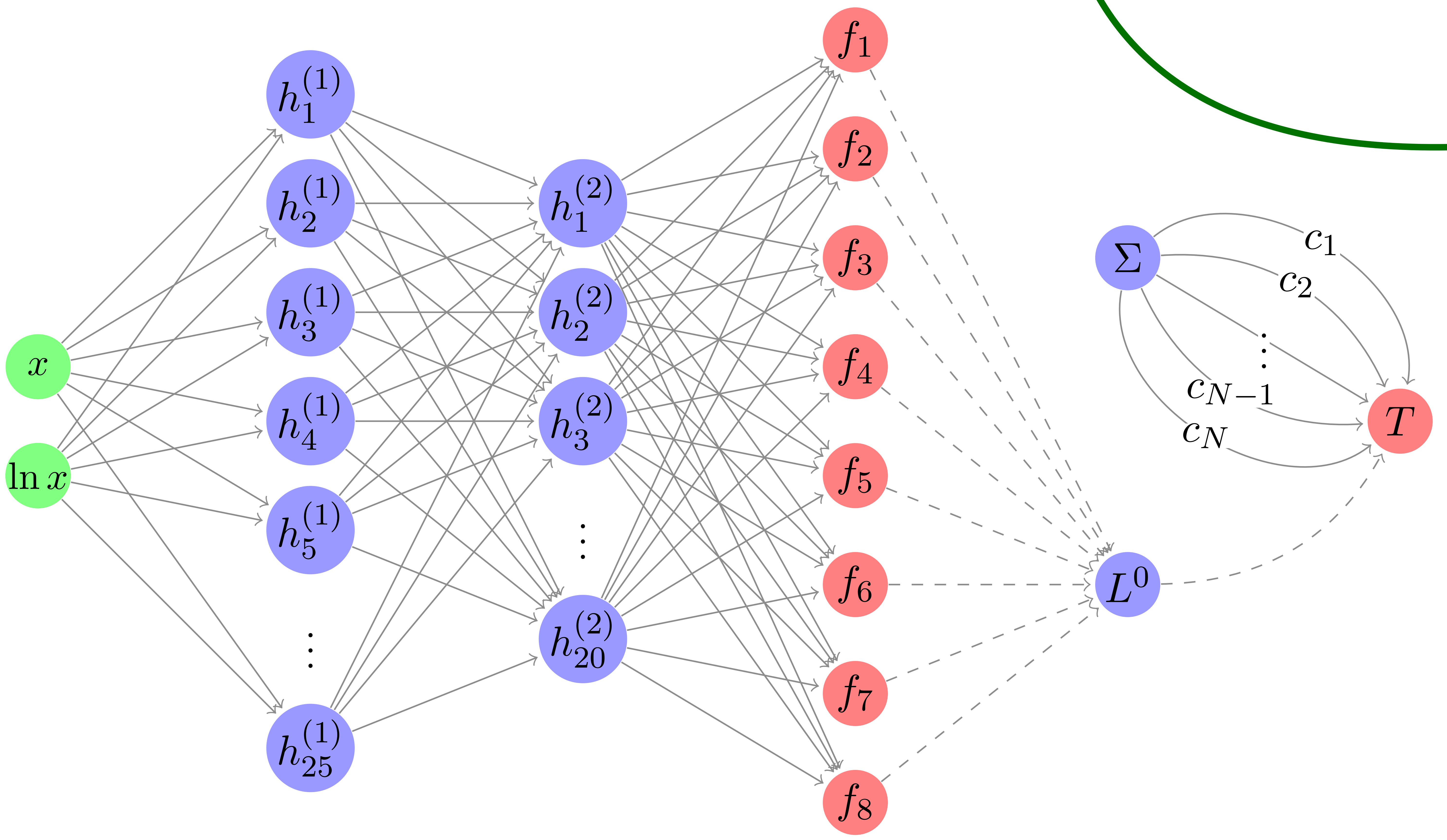


Maeve Madigan 16/06/22

SIMUnet *S. Iranipour, M. Ubiali, 2201.07240*

"A new methodology that is able to yield a simultaneous determination of the PDFs alongside **any set of parameters that determine the theory predictions**"

e.g. Wilson coefficients!

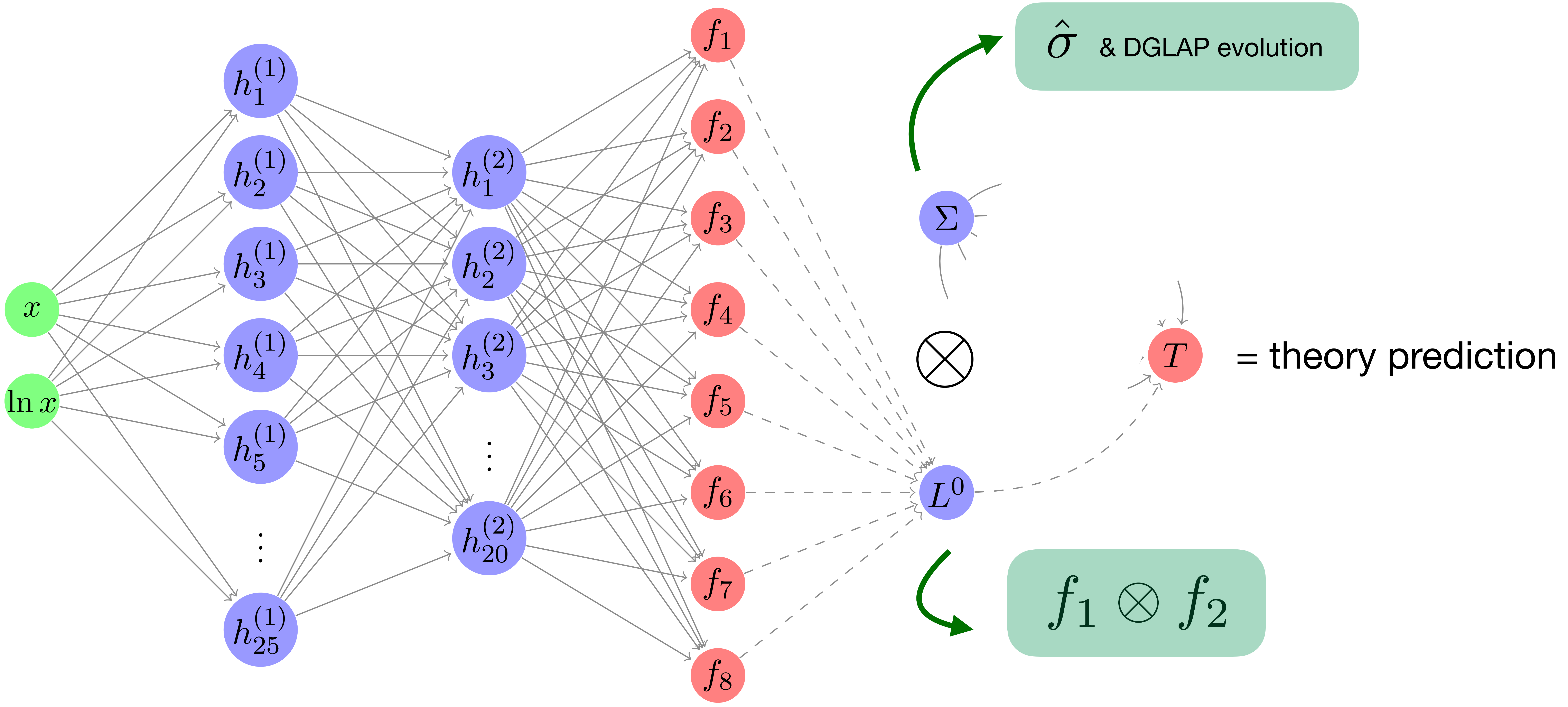


SIMUnet

S. Iranipour, M. Ubiali, 2201.07240

$$\sigma = f_1 \otimes f_2 \otimes \hat{\sigma}_{SM}$$

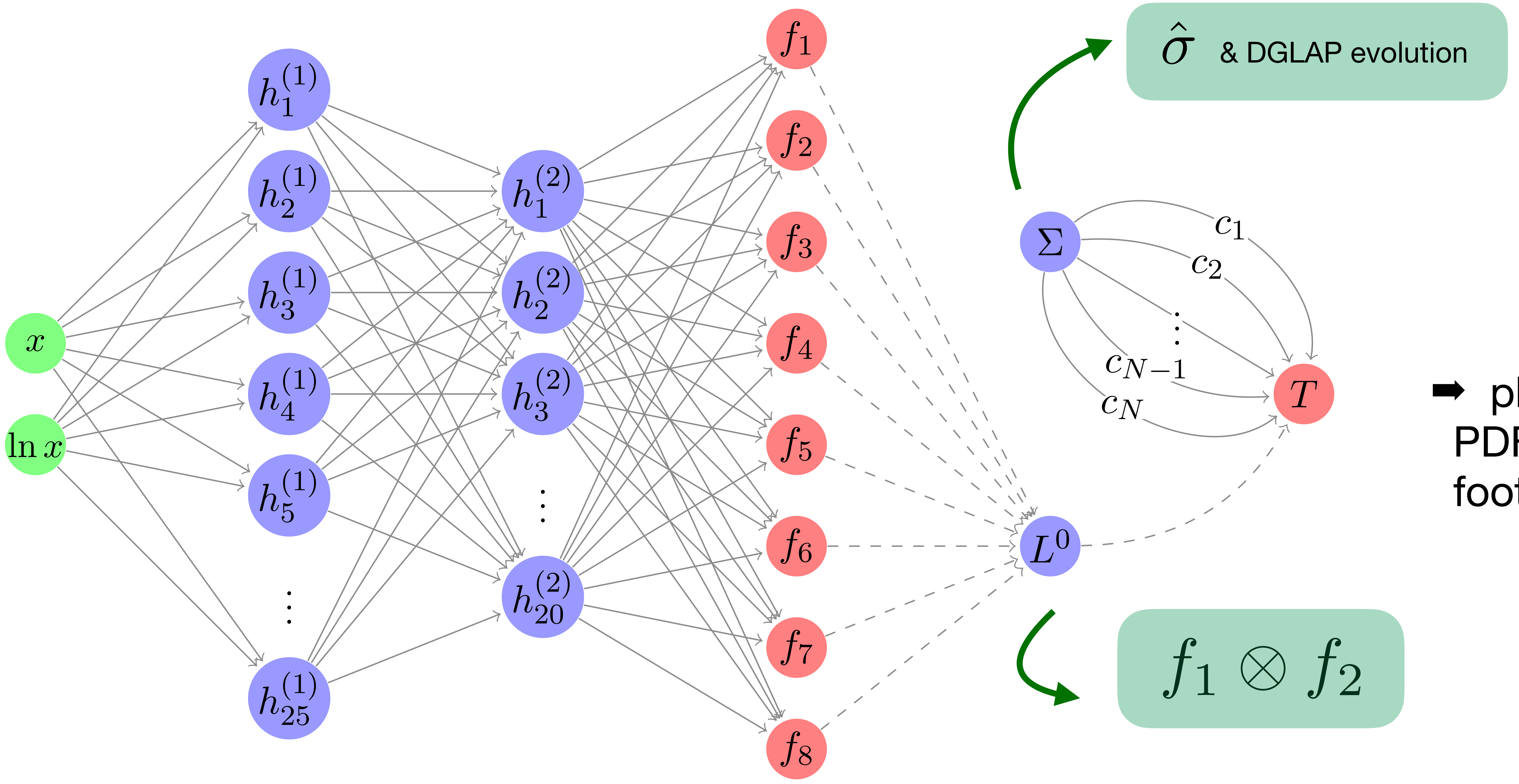
How does it work?



SIMUnet

S. Iranipour, M. Ubiali, 2201.07240

$$\sigma = f_1 \otimes f_2 \otimes \hat{\sigma}_{SM}$$

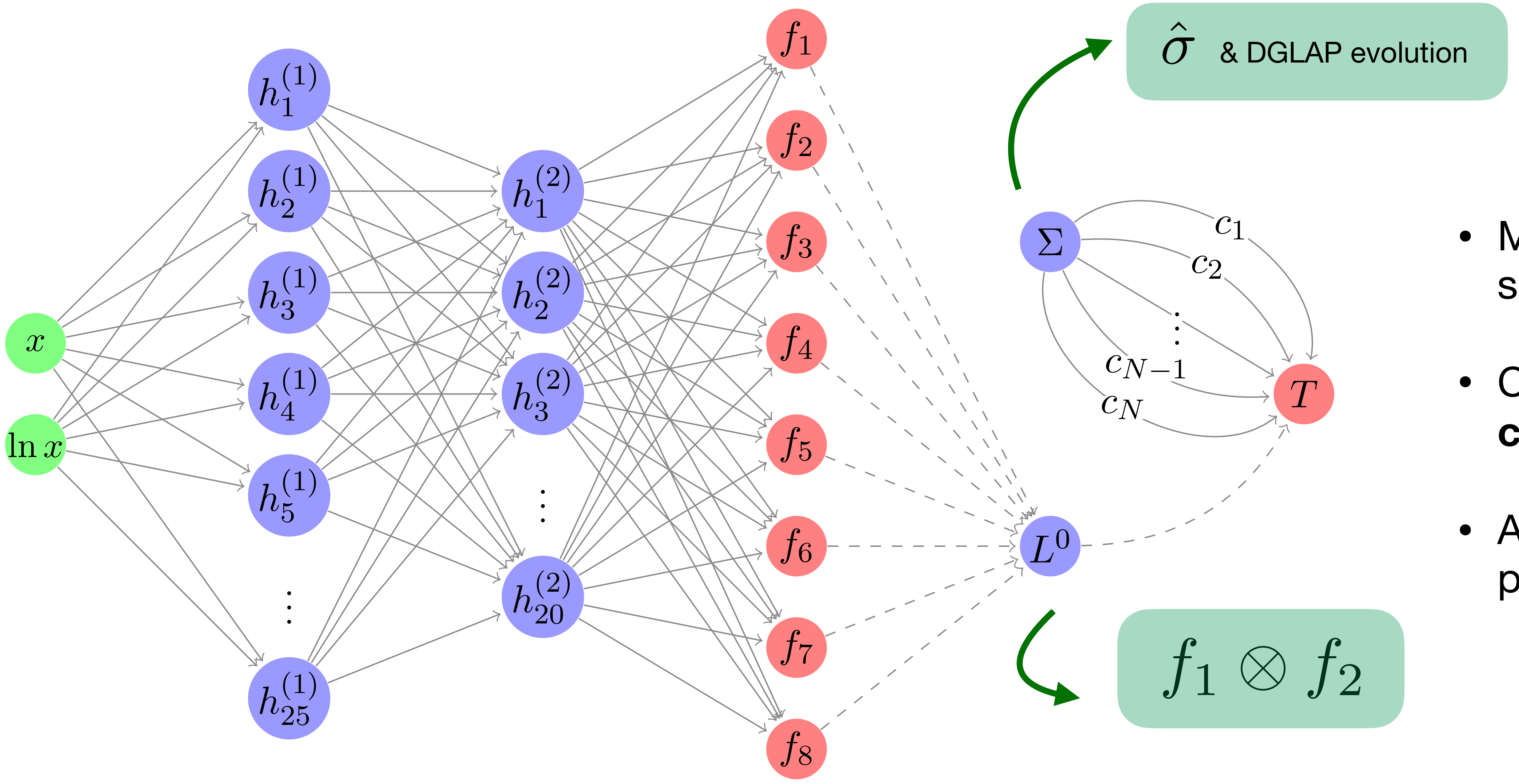


➡ places SMEFT parameters and PDF parameters on the same footing

SIMUnet

S. Iranipour, M. Ubiali, 2201.07240

$$\sigma = f_1 \otimes f_2 \otimes \hat{\sigma}_{SM}$$



- Much more efficient than a grid scan
- Capable of handling **more SMEFT coefficients**
- Already benchmarked for the W,Y parameters in high mass DY

PDF-EFT in the top sector

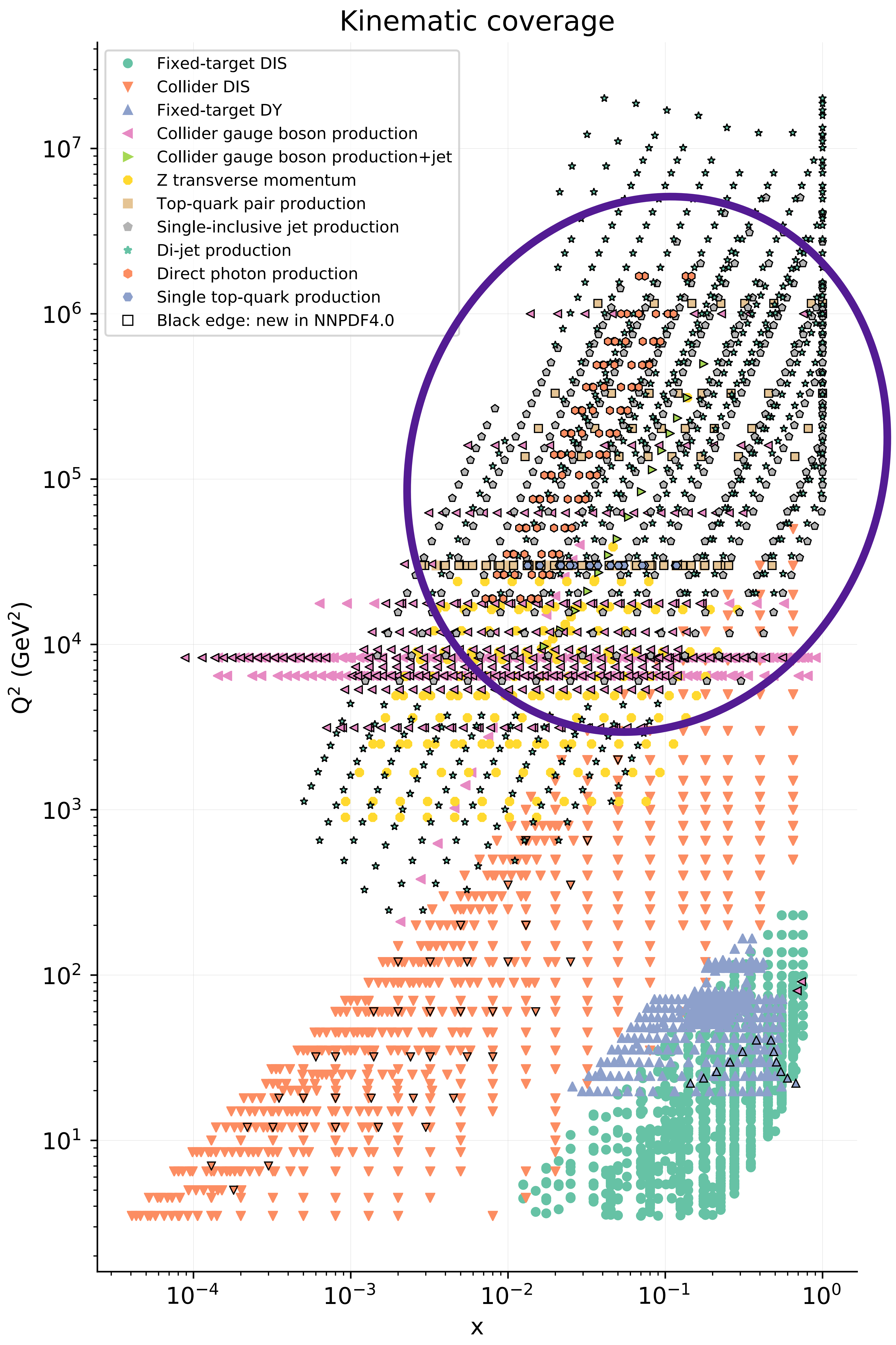
work in progress

LHC Run II data from the top sector has already been used to constrain both the PDFs and SMEFT.

(see for example 2109.02653, *2012.02779, 2105.00006)*

Working towards a simultaneous fit of the PDFs and 16 dimension-6 SMEFT operators, using data from

$t\bar{t}$ (incl. A_C), $t\bar{t} + X$, single t, $tZ, tW, ...$



Conclusions

We have studied the interplay of PDF and EFT effects in DY and DIS data.

Using data from LHC Run I and II, the effect of the interplay is visible but still within PDF uncertainties.

At the HL-LHC:

- Not accounting for the interplay may lead to artificially precise constraints on the EFT.
- Conservative PDFs still lead to stronger bounds than SMEFT PDFs.

Next steps:

- Use the new SIMUnet methodology to investigate this interplay at the level of a global fit in the top sector
- Further investigation into the definition of conservative PDF sets i.e. cutting data out of the PDF fits

Conclusions

We have studied the interplay of PDF and EFT effects in DY and DIS data.

Using data from LHC Run I and II, the effect of the interplay is visible but still within PDF uncertainties.

At the HL-LHC:

- Not accounting for the interplay may lead to artificially precise constraints on the EFT.
- Conservative PDFs still lead to stronger bounds than SMEFT PDFs.

Next steps:

- Use the new SIMUnet methodology to investigate this interplay at the level of a global fit in the top sector
- Further investigation into the definition of conservative PDF sets i.e. cutting data out of the PDF fits

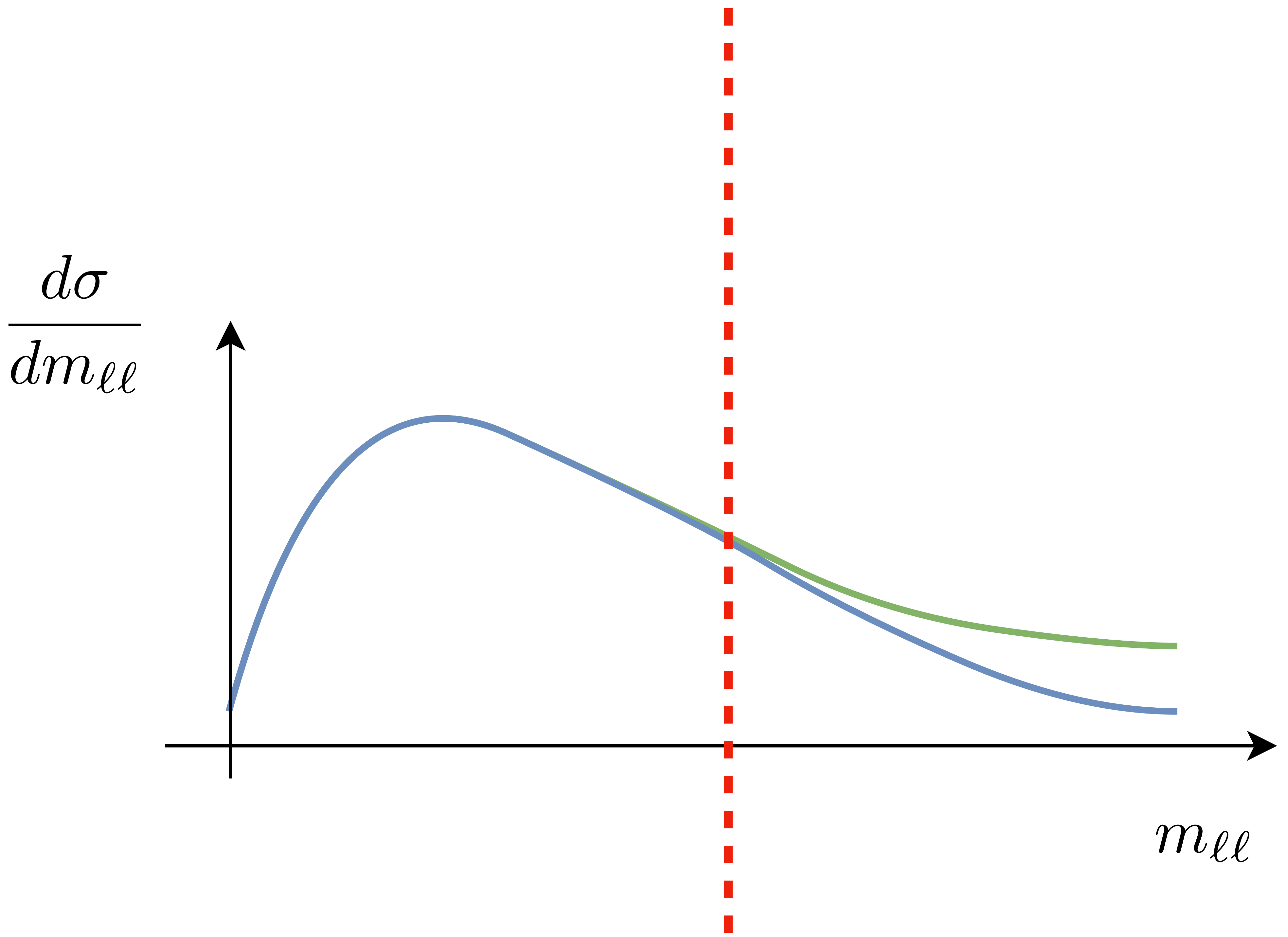
Thank you for listening!

Backup

Conservative PDFs

Could we improve the SM PDF fits by removing the high-mass data from PDF fits?

- not in the spirit of global fits
- still have a theoretical inconsistency due to SM assumptions
- **but** much easier than doing a simultaneous PDF-SMEFT fit

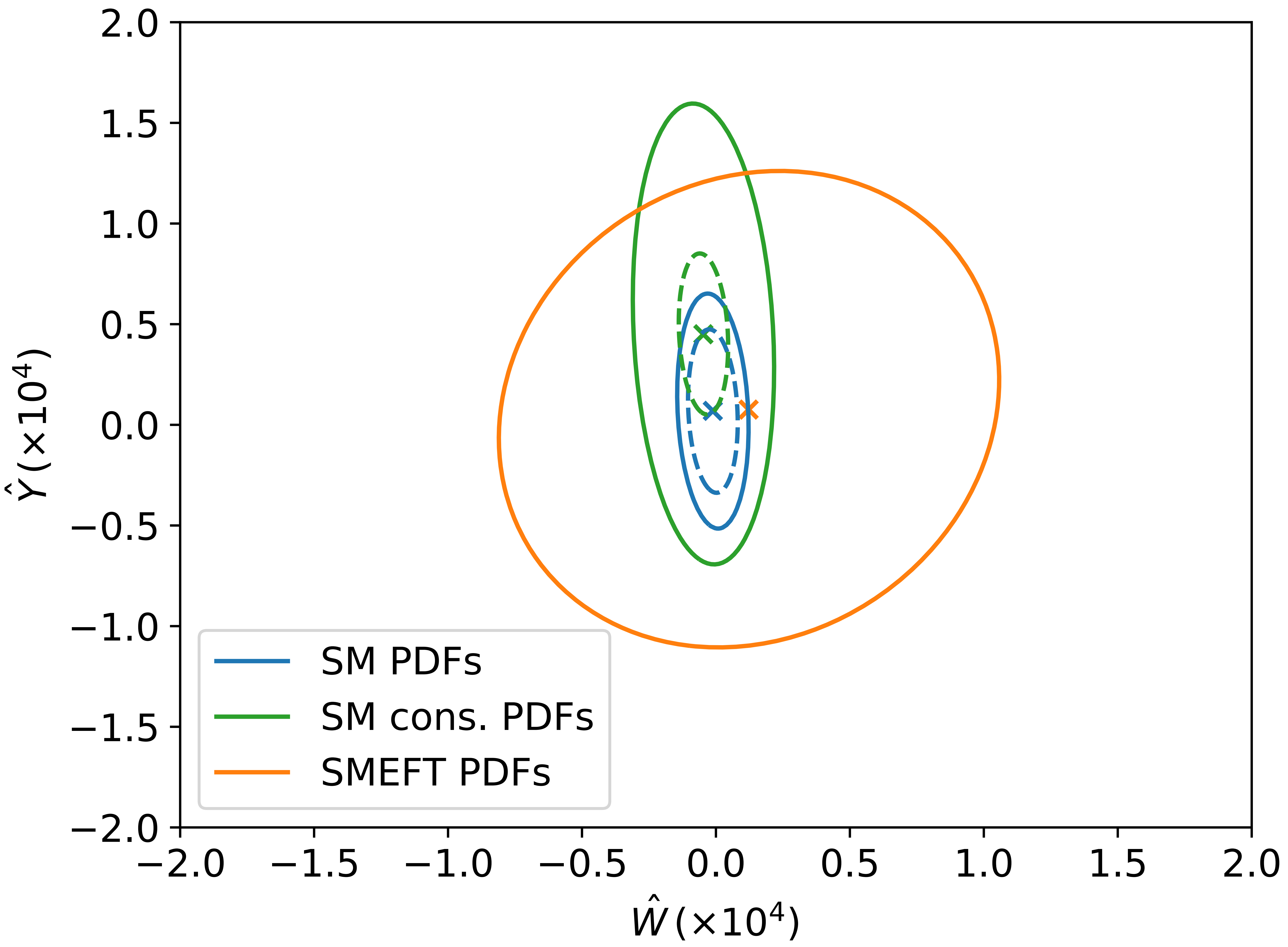


Conservative PDFs

Conservative PDFs:

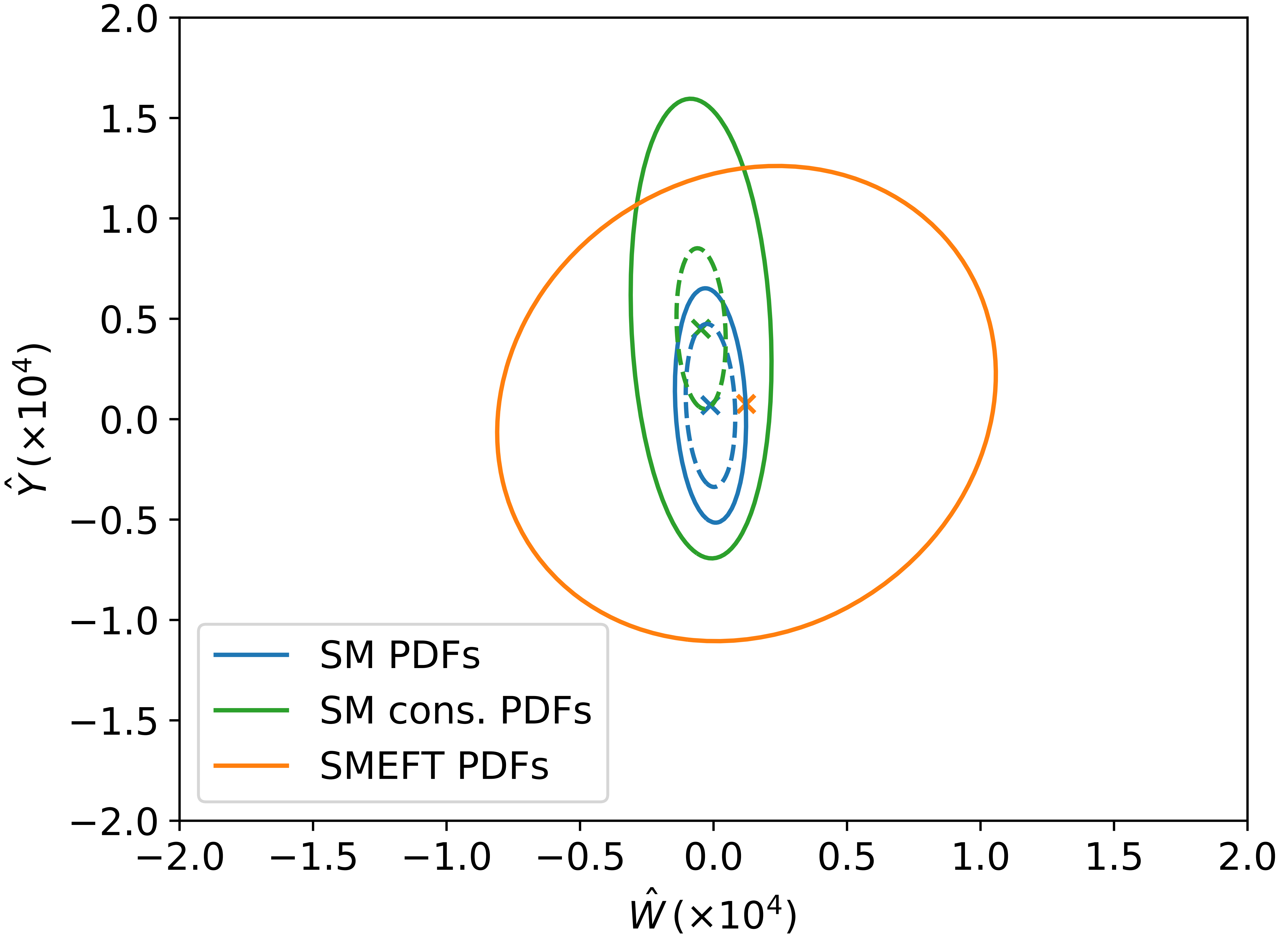
- assume the SM
- are fit to data which does not receive large SMEFT corrections
  (i.e. no HL-LHC data, no high-mass DY data)

dashed: not including PDF uncertainties



Conservative PDFs

dashed: not including PDF uncertainties



Conservative PDFs:

- assume the SM
- are fit to data which does not receive large SMEFT corrections (i.e. no HL-LHC data, no high-mass DY data)

Comparing green to orange:

- the constraints using SM conservative PDFs are closer to those using SMEFT PDFs
- still overestimating the constraints, especially in the $\hat{W}$ direction

HL-LHC projections

We generate theory predictions as before, assuming

$$\sqrt{s} = 14 \text{ TeV} \qquad \mathcal{L} = 6 \text{ ab}^{-1} \longrightarrow$$ ATLAS & CMS

Acceptance cuts and systematics uncertainties are modelled on reference measurements:

- NC DY: CMS 13 TeV measurement of the DY differential cross section, *1310.7291*
- CC DY: ATLAS 13 TeV search for *W'* bosons in the dilepton channel, *1906.05609*

HL-LHC projections

We produce pseudodata **under the assumption of the SM:** $\sigma_i^{exp} = \sigma_i^{th}(1 + \lambda\delta_{\mathcal{L}} + r_i\delta_{exp,i})$

where

- $\lambda, r_i \sim \mathcal{N}(0,1)$
- Luminosity uncertainty, correlated across all bins: $\delta_{\mathcal{L}} = 1.5\%$
- Statistical & systematic uncertainties: $\delta_{exp,i} = \sqrt{\delta^2_{stat,i} + (f_{red}\delta_{syst,i})^2}$

$f_{red} = 0.2$ → *'optimism factor'*