MARIA UBIALI
UNIVERSITY OF CAMBRIDGE

PDF AND (SM)EFT INTERPLAY

OUTLINE

- Introduction:
  - ➡ PDF and SMEFT fits
  - ➡ Time to study interplay between SMEFT and PDF fits

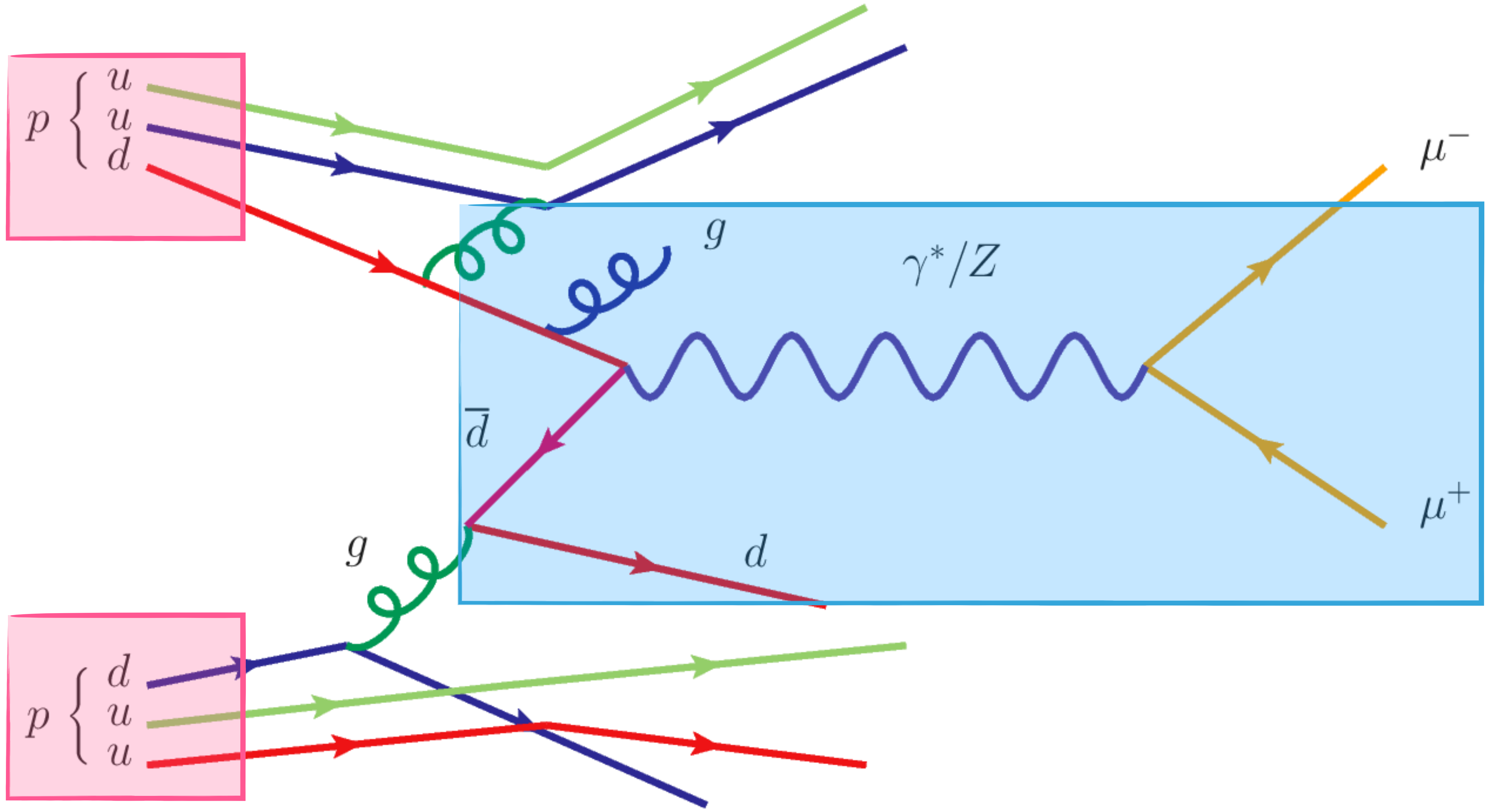
- PDFs in the SMEFT from high-mass Drell-Yan tails
  - ➡ Analysis settings
  - ➡ Oblique operators: constraints from Run I and Run II data
  - ➡ Oblique operators: projections at the HL-LHC
  - ➡ Flavour specific four-fermion operator

- Conclusions and outlook

INTRODUCTION

THEORETICAL PREDICTIONS AT THE LHC

Parton Distribution Functions



Hard Scattering: Perturbative QCD + EW

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

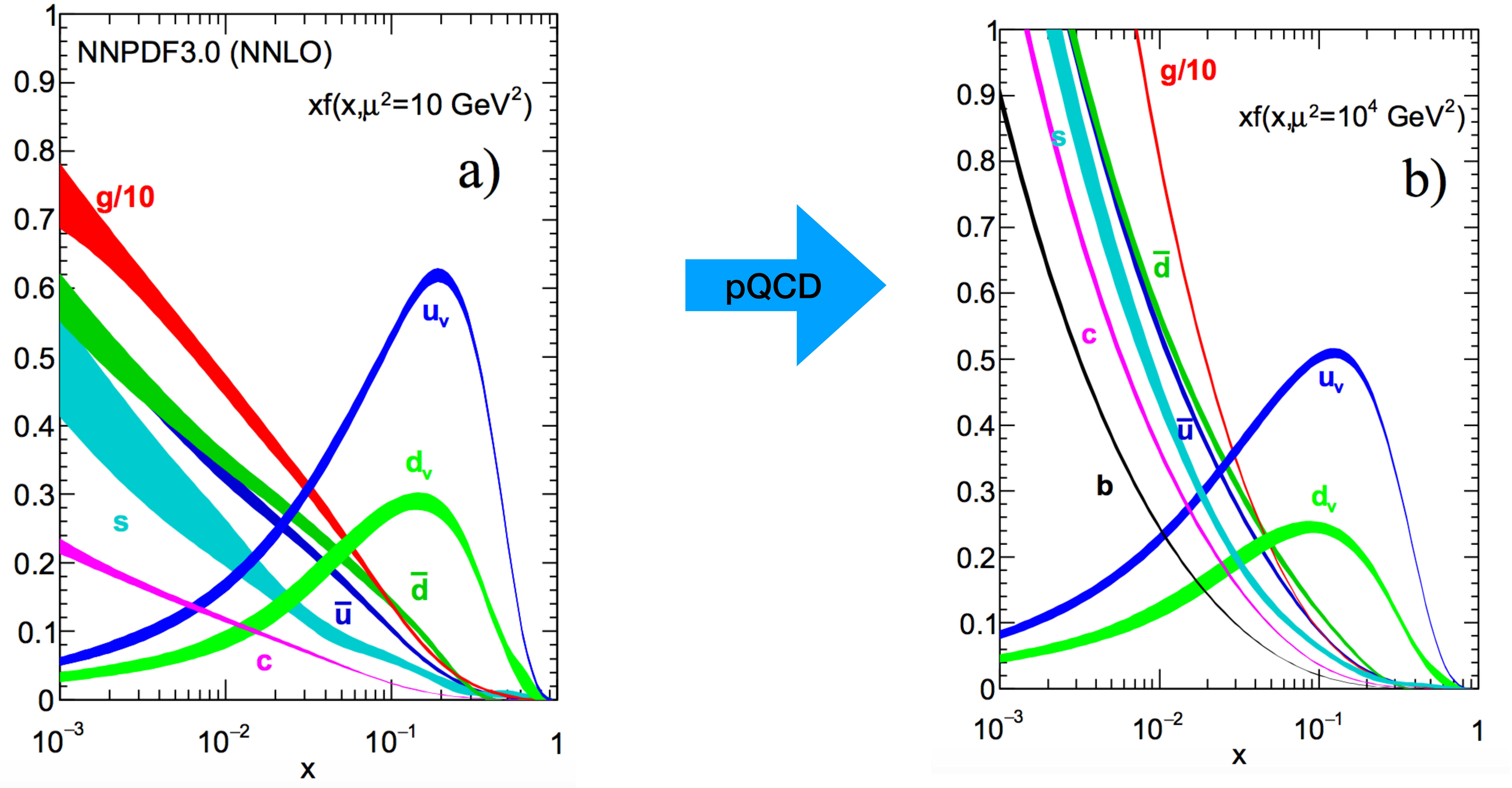
PARTON DISTRIBUTION FUNCTIONS

$$f_i(x, \mu)$$

Data

Perturbative QCD

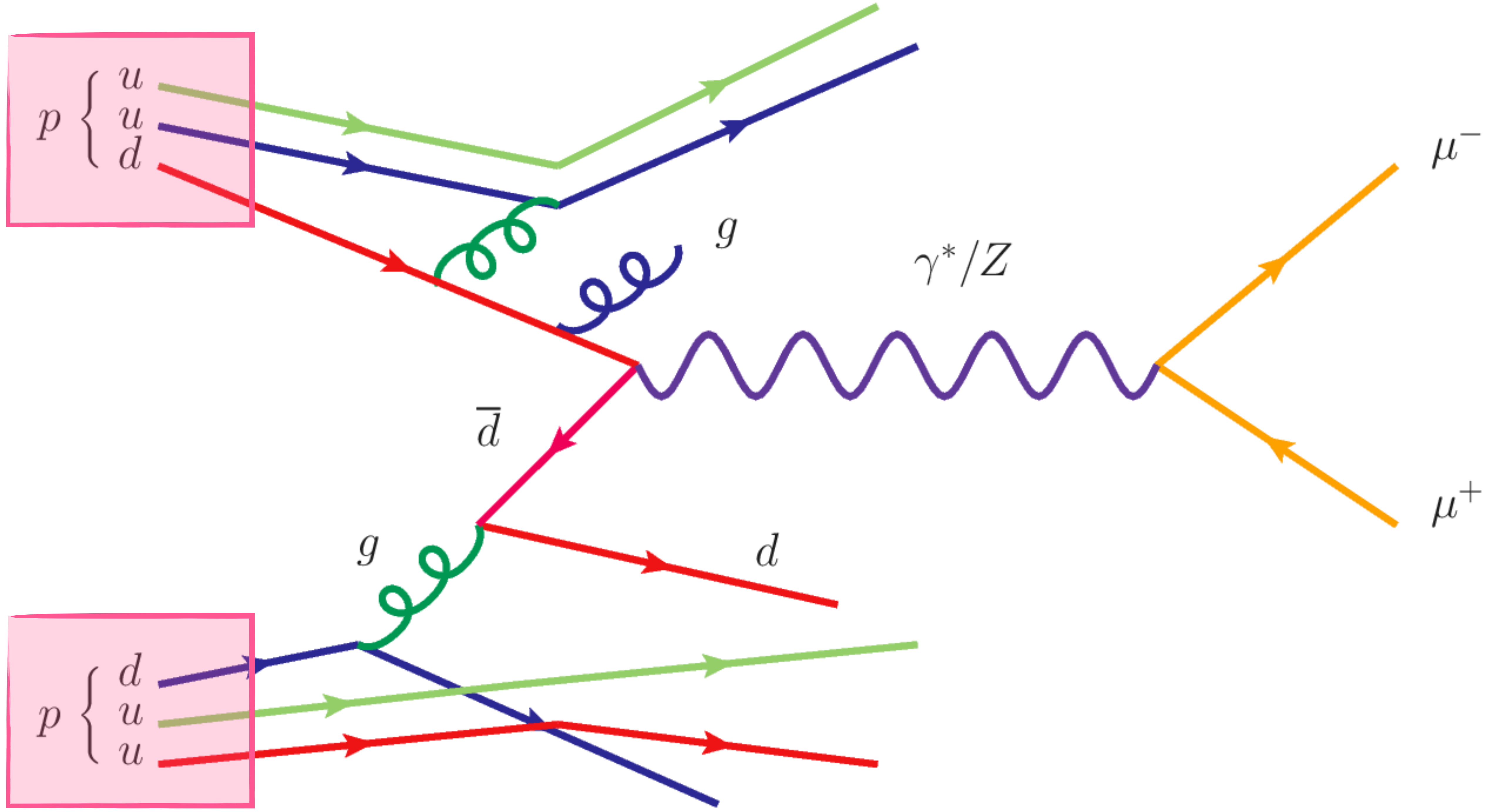
Hadronic scale: global fit of PDFs



High scale: input to the LHC

PDG 2016

PDF FITS



PDF fits:
Extract universal parton distribution functions from data and propagate data uncertainty in PDF uncertainty

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

$T_i(\{\theta_k\})$

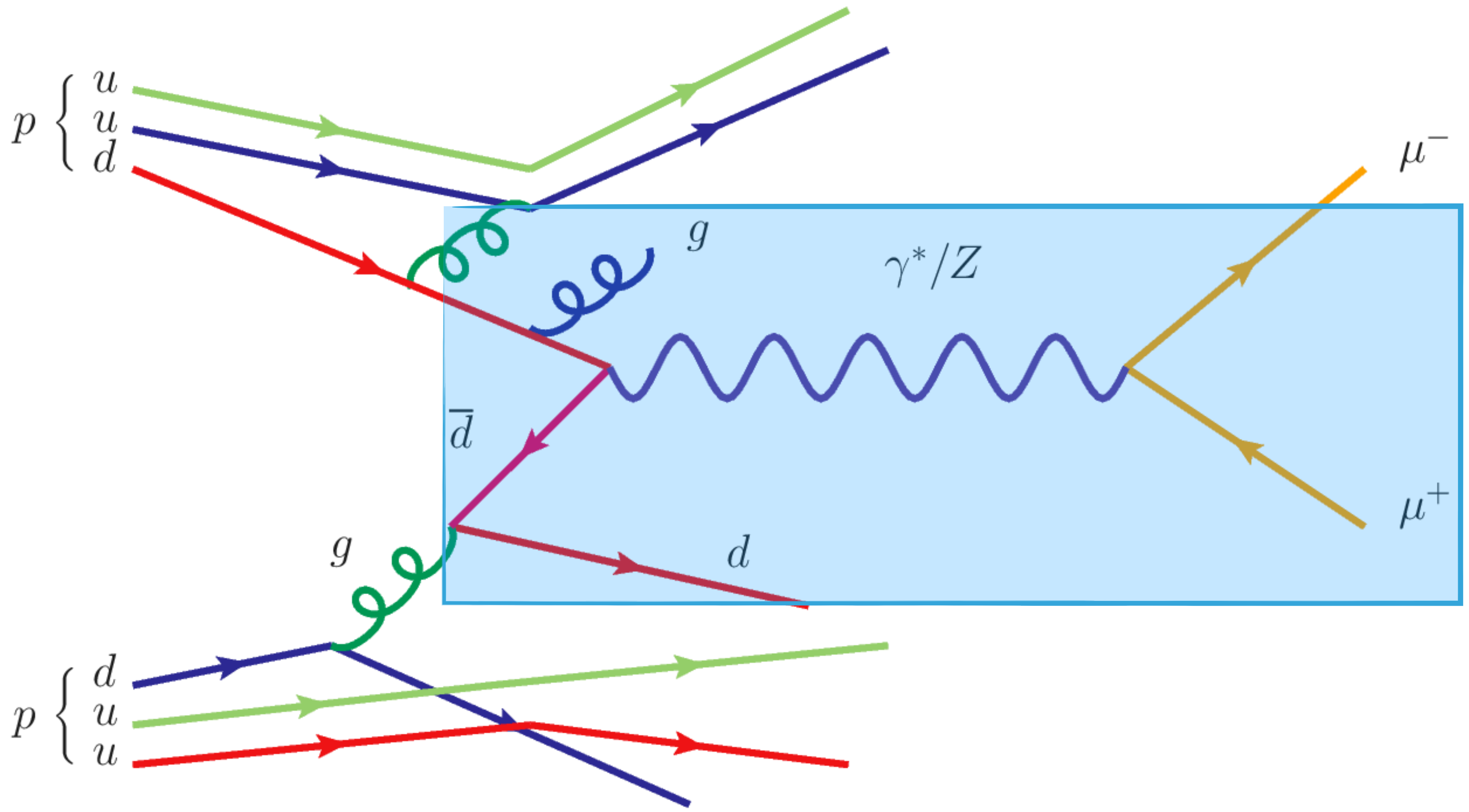
T $$d\sigma^{pp\to ab} = \sum_{i,j} f_i \otimes f_j \otimes d\hat{\sigma}^{ij\to ab} + \ldots$$

$f(\{\theta_k\})$

Assume SM theory predictions

Fit PDF parameters $\{\theta_k\}$

SMEFT FITS



SMEFT fits:
Treat the Standard Model as the low energy, IR limit of some UV complete theory and extract bounds on Wilson coefficients c_i from data

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

$T_i(\{c_k\})$

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

Assume SM PDFs

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \cdots$$

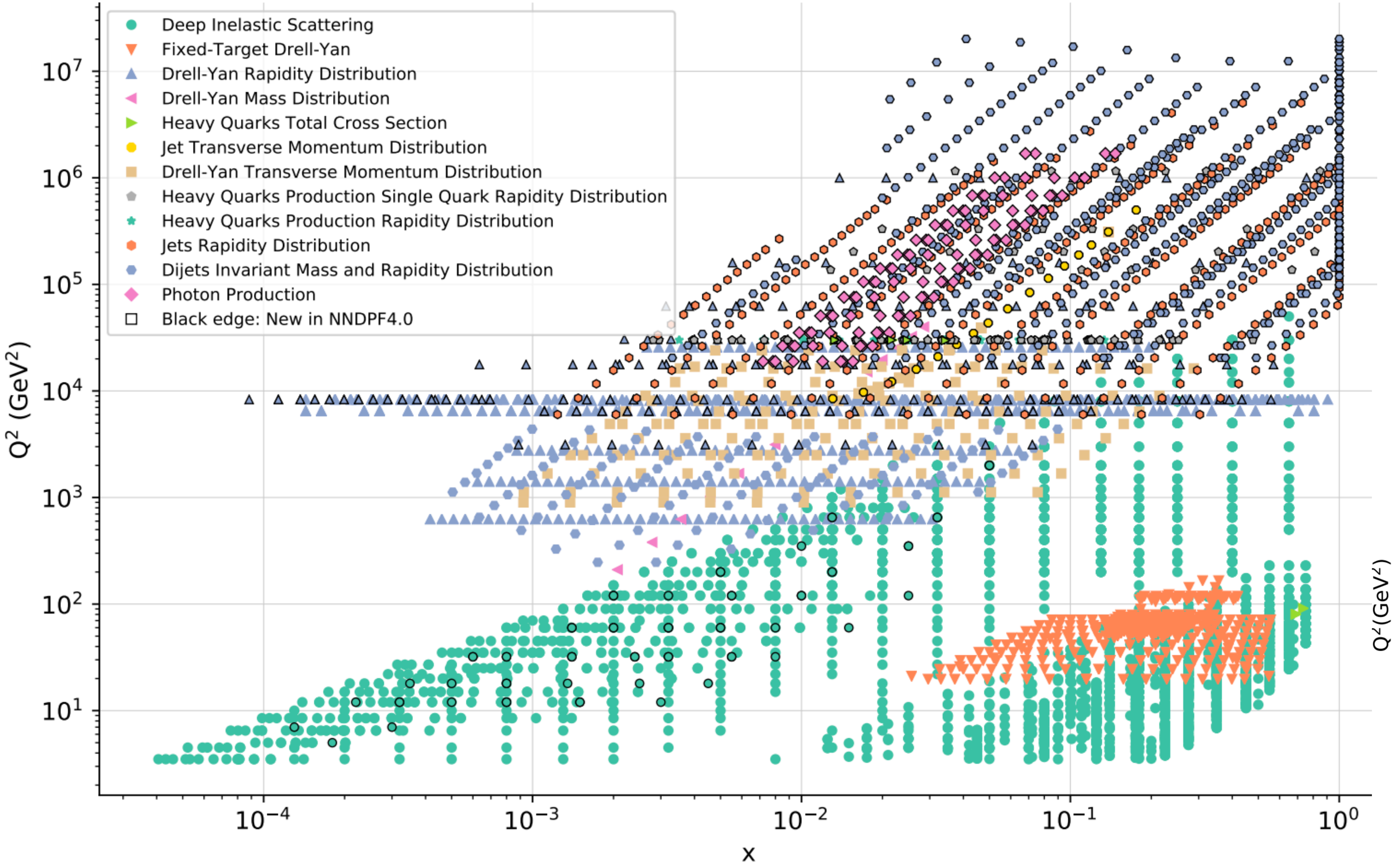
Fit of Wilson Coefficients $\{c_k\}$

PDF AND SMEFT INTERPLAY

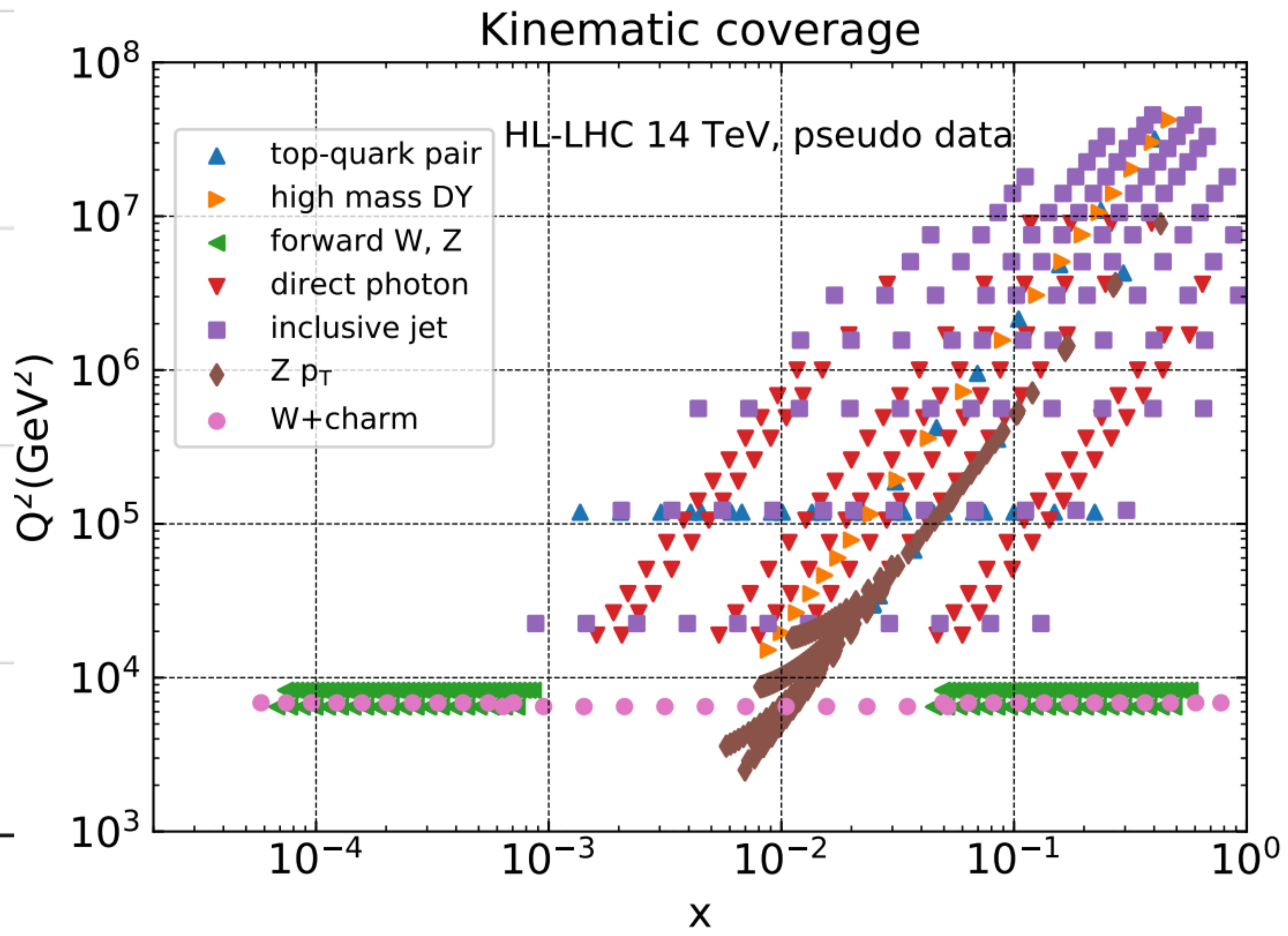
- PDFs are low-scale quantities extracted from experimental data, without considering any potential high-scale contamination due to new physics.
- Model-independent parametrisation of new physics are performed by assuming a priori that PDFs are SM-like.
- In principle low-scale physics is separable from high-scale physics, BUT the complexity of the LHC environment might well intertwine them.

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

PDF AND SMEFT INTERPLAY



- ➡ Top pair production and single top data included in SMEFT analysis [Hartland et al 1901.05965] [Ellis et al 2012.02779]
- ➡ Dijets data in [Bordone et al 2103.10332] [Alioli et al 1706.03068]
- ➡ Drell-Yan data in [Farina et al 1609.08157, Torre et al 2008.12978]
- ➡ Inclusive jets in [Alte et al 1711.07484]
- ➡ Overlap enhanced in HL-LHC projections [Abdul Khalek et al1810.03639]



A FEW COMPELLING QUESTIONS

- 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 the same operators that I am fitting?

T

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

$f(\{\theta_k\})$

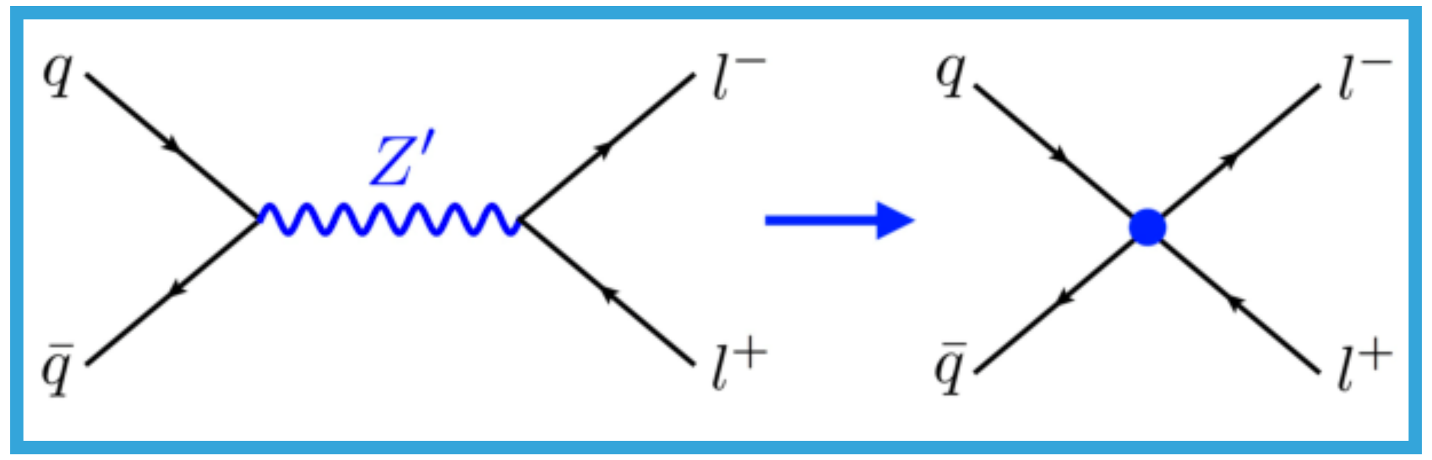
$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \sum_i \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \cdots$$

Simultaneous fits can shed light on their interplay

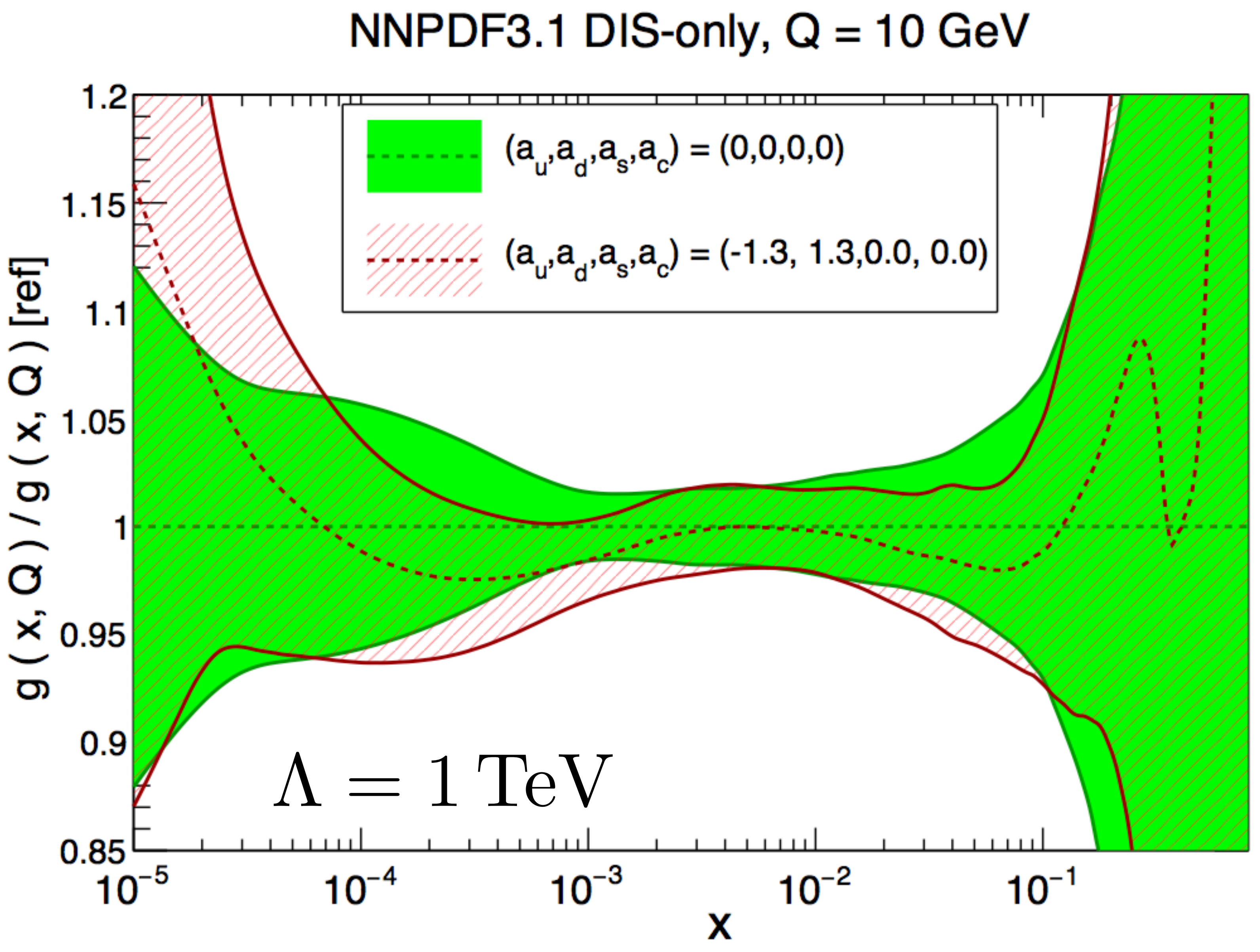
$T(\{\theta_k\},\{c_i\})$

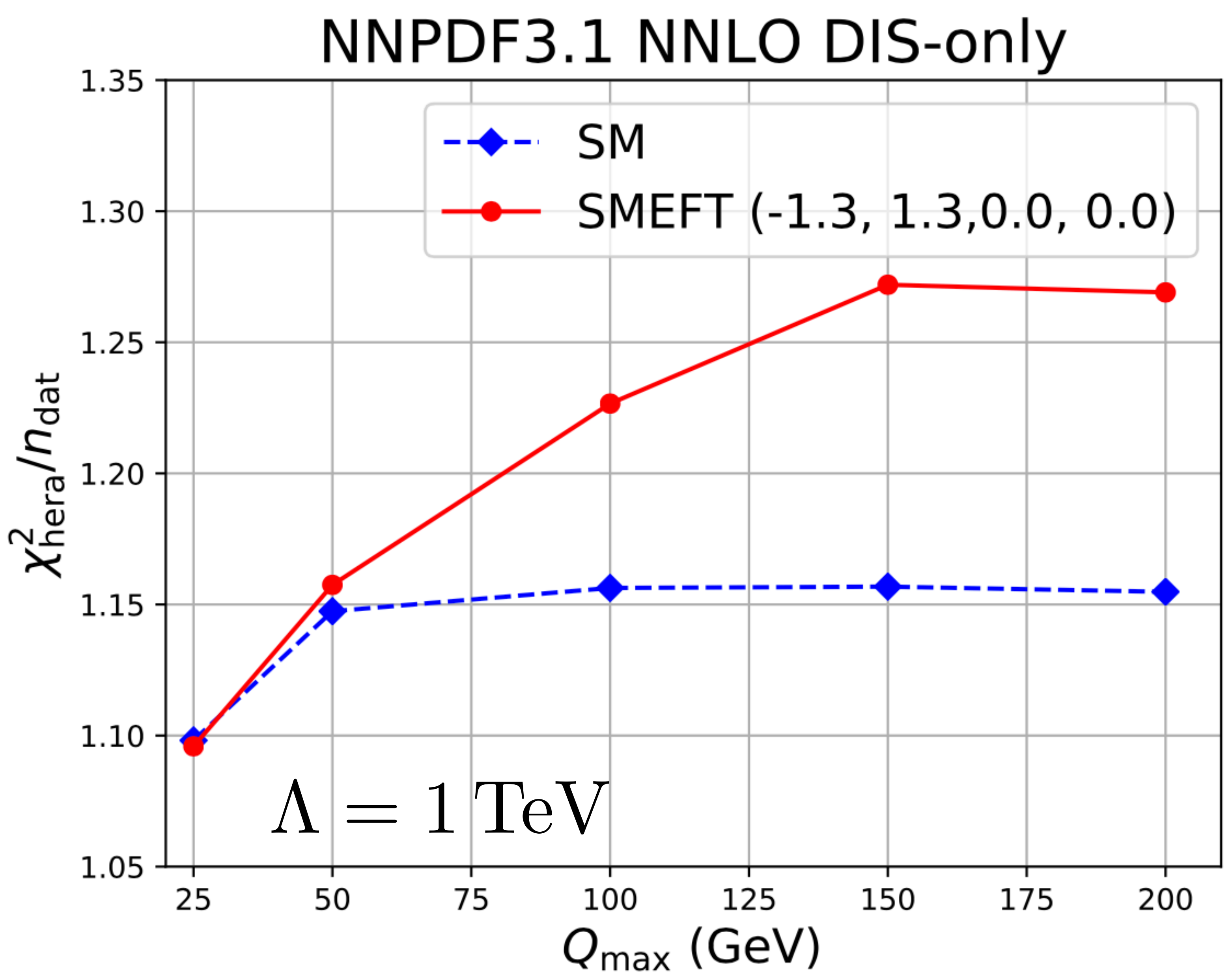
CASE-STUDY: DEEP INELASTIC SCATTERING

- 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 strong LEP constraints
- PDF fits based on DIS only data (Q ≲ 200 GeV for HERA data)



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





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

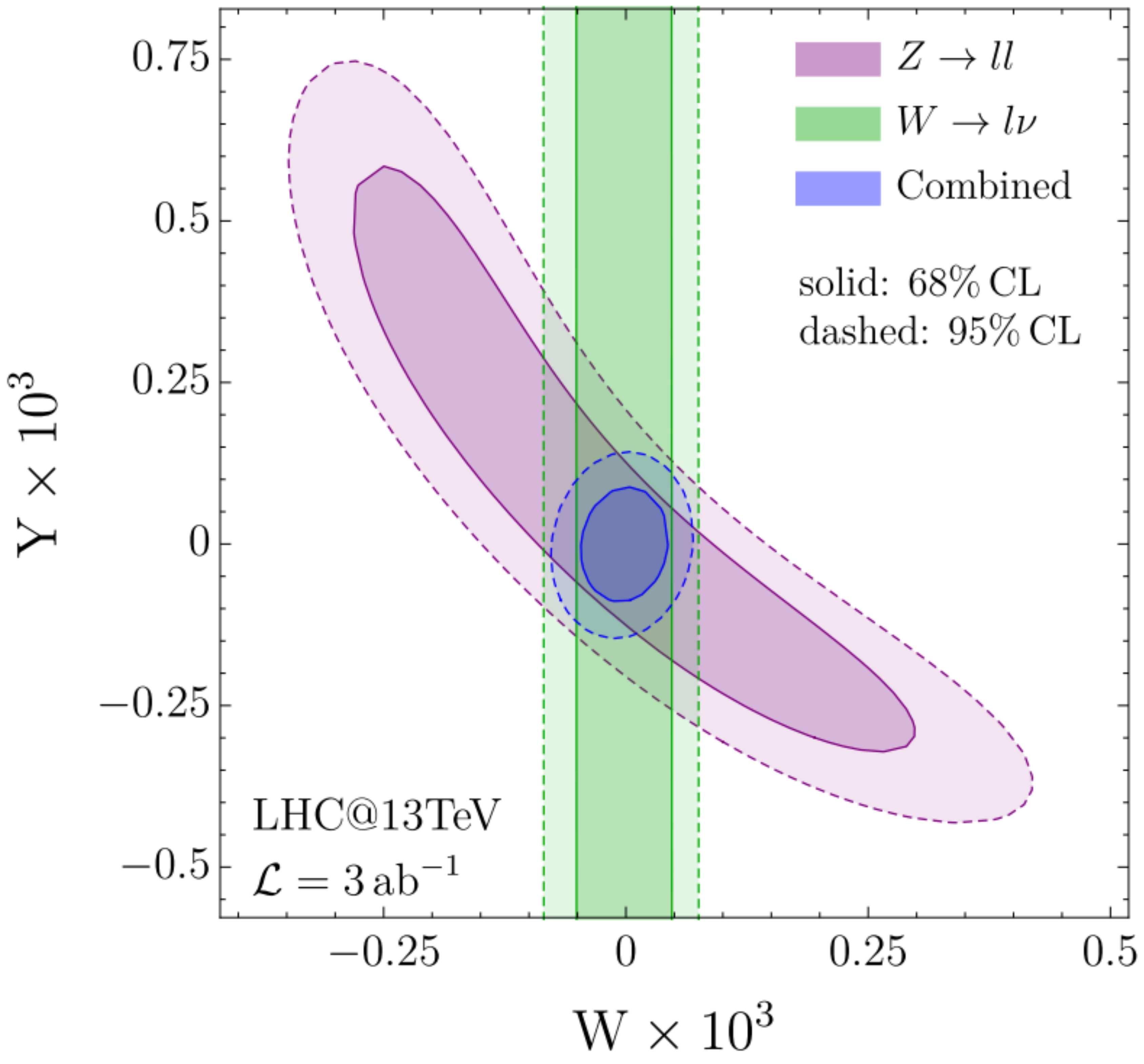
HIGH-MASS DRELL-YAN TAILS

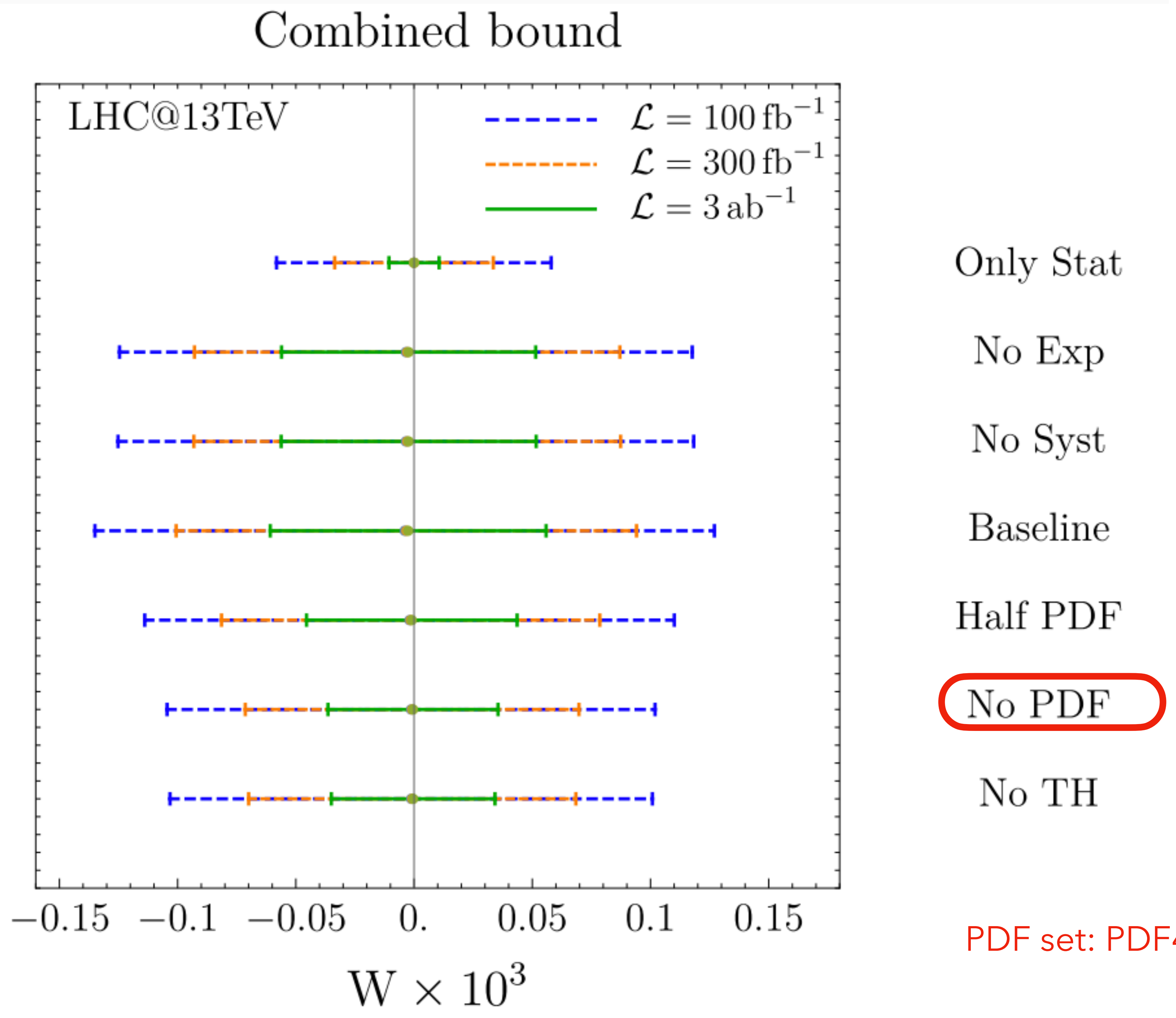
W&Y INTERPRETATION OF HIGH-ENERGY DRELL-YAN MEASUREMENTS

- Case study at higher energy: EW oblique corrections in high-mass NC and CC Drell-Yan tails.
- They 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]

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

Torre et al, 2008.12978





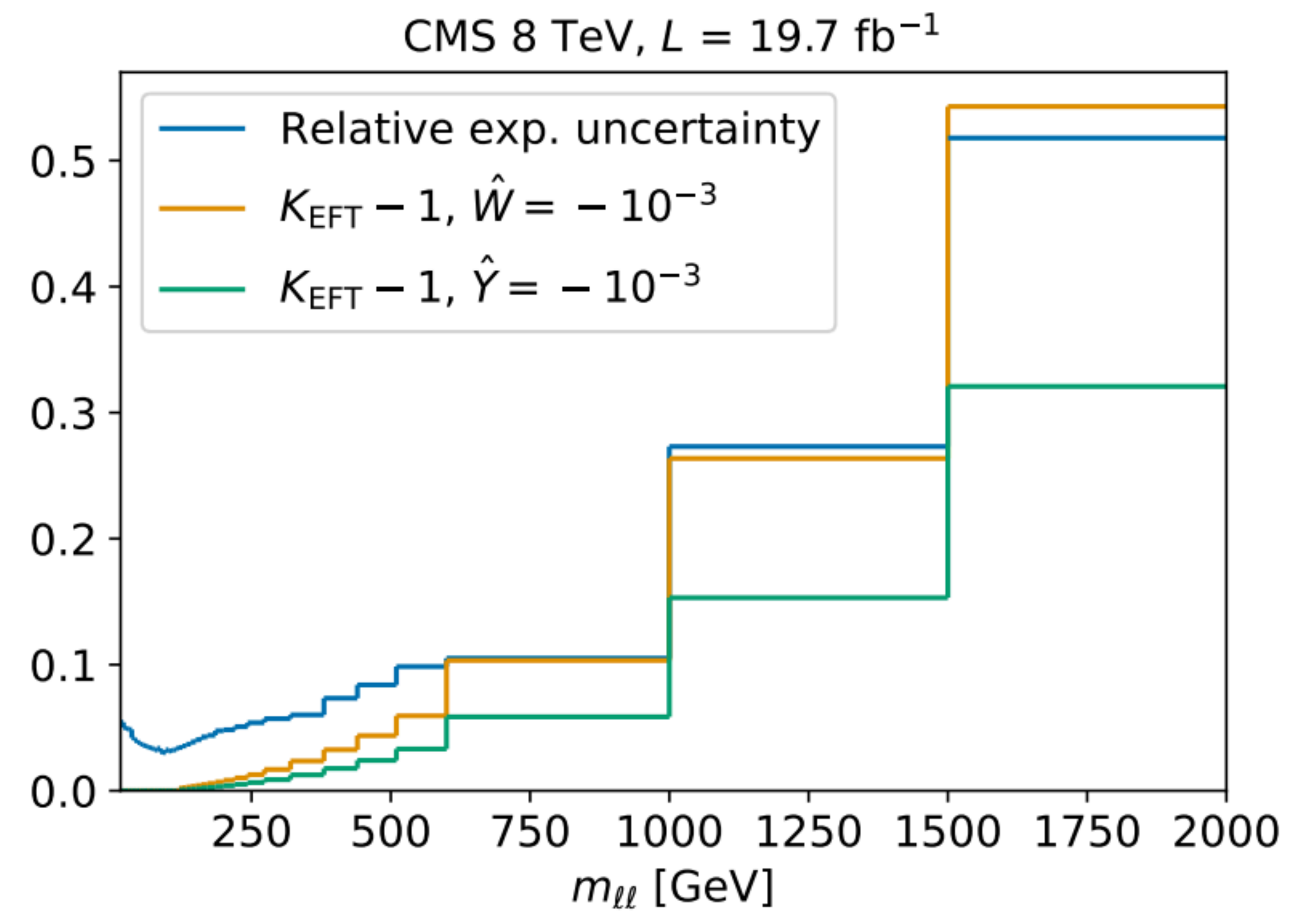
PDF set: PDF4LHC15_nlo_30_pdfas

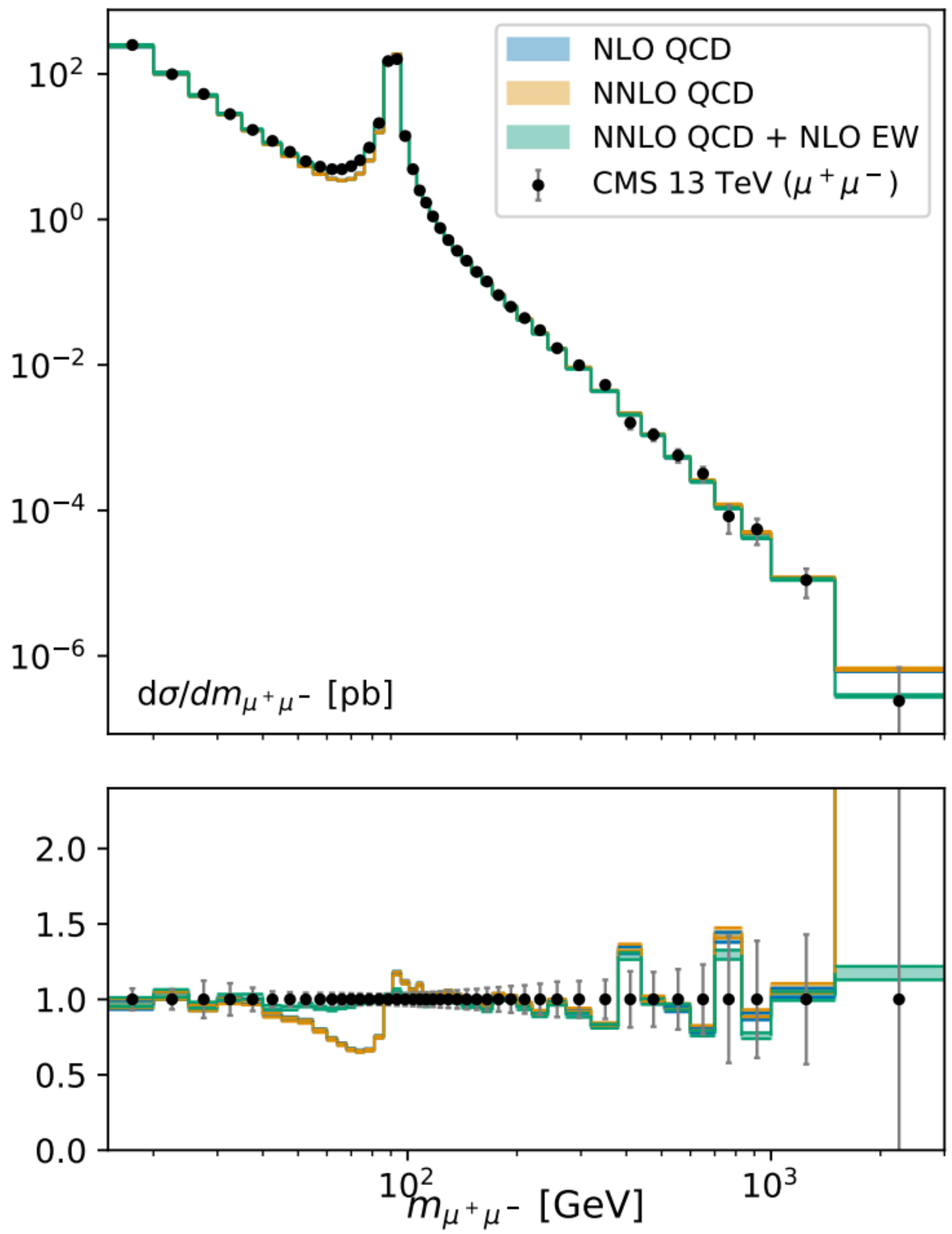
ANALYSIS SETTINGS

- 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]

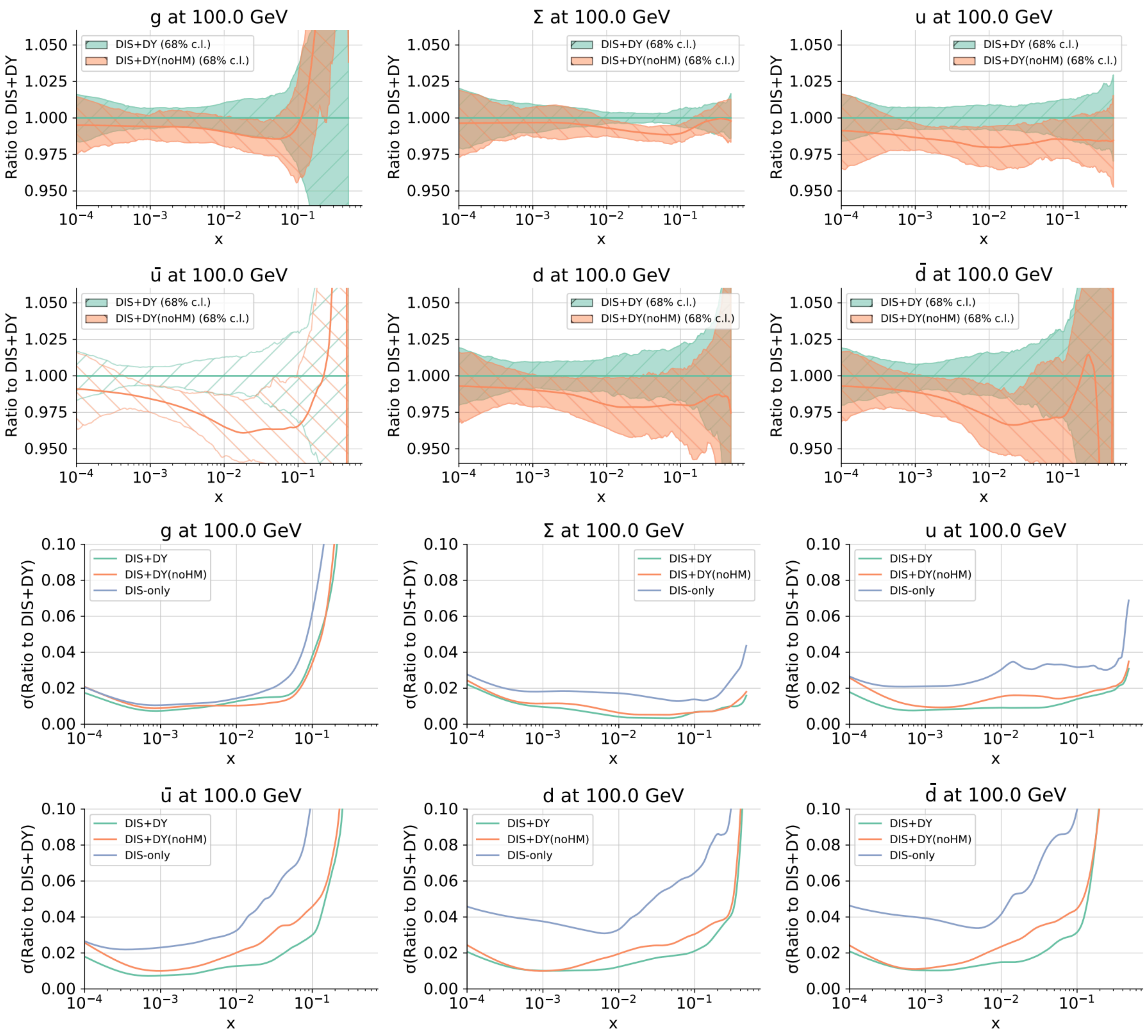
- Settings:
  ➡ PDF fit based on DIS (~3000 data points), Drell-Yan on-shell and low-mass data from ATLAS, CMS and LHCb (~600 data points)
  ➡ + Run I and II ATLAS and CMS high mass NC Drell-Yan data (~300 data points)
  ➡ SM predictions at NNLO QCD + NLO EW and SMEFT corrections added via local K-factors

Exp.	$\sqrt{s}$ (TeV)	Ref.	$\mathcal{L}$ (fb^{-1})	Channel	1D/2D	$n_{\rm dat}$	$m_{\ell\ell}^{\rm max}$ (TeV)
ATLAS	7	[117]	4.9	e^-e^+	1D	13	[1.0, 1.5]
ATLAS (*)	8	[83]	20.3	$\ell^-\ell^+$	2D	46	[0.5, 1.5]
CMS	7	[118]	9.3	$\mu^-\mu^+$	2D	127	[0.2, 1.5]
CMS (*)	8	[84]	19.7	$\ell^-\ell^+$	1D	41	[1.5, 2.0]
CMS (*)	13	[119]	5.1	e^-e^+, $\mu^-\mu^+$ $\ell^-\ell^+$	1D	43, 43 43	[1.5, 3.0]
Total						**270 (313)**	





INPUT PDFS @ RUN I AND RUN II



- High mass Drell-Yan data have a pull on light quark and anti-quark PDFs compared to a fit that does not include them
- PDF uncertainties much reduced as compared to DIS-only fit
- Light quark and antiquark uncertainties further reduced by inclusion of high-mass Drell-Yan data

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]

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

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

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

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

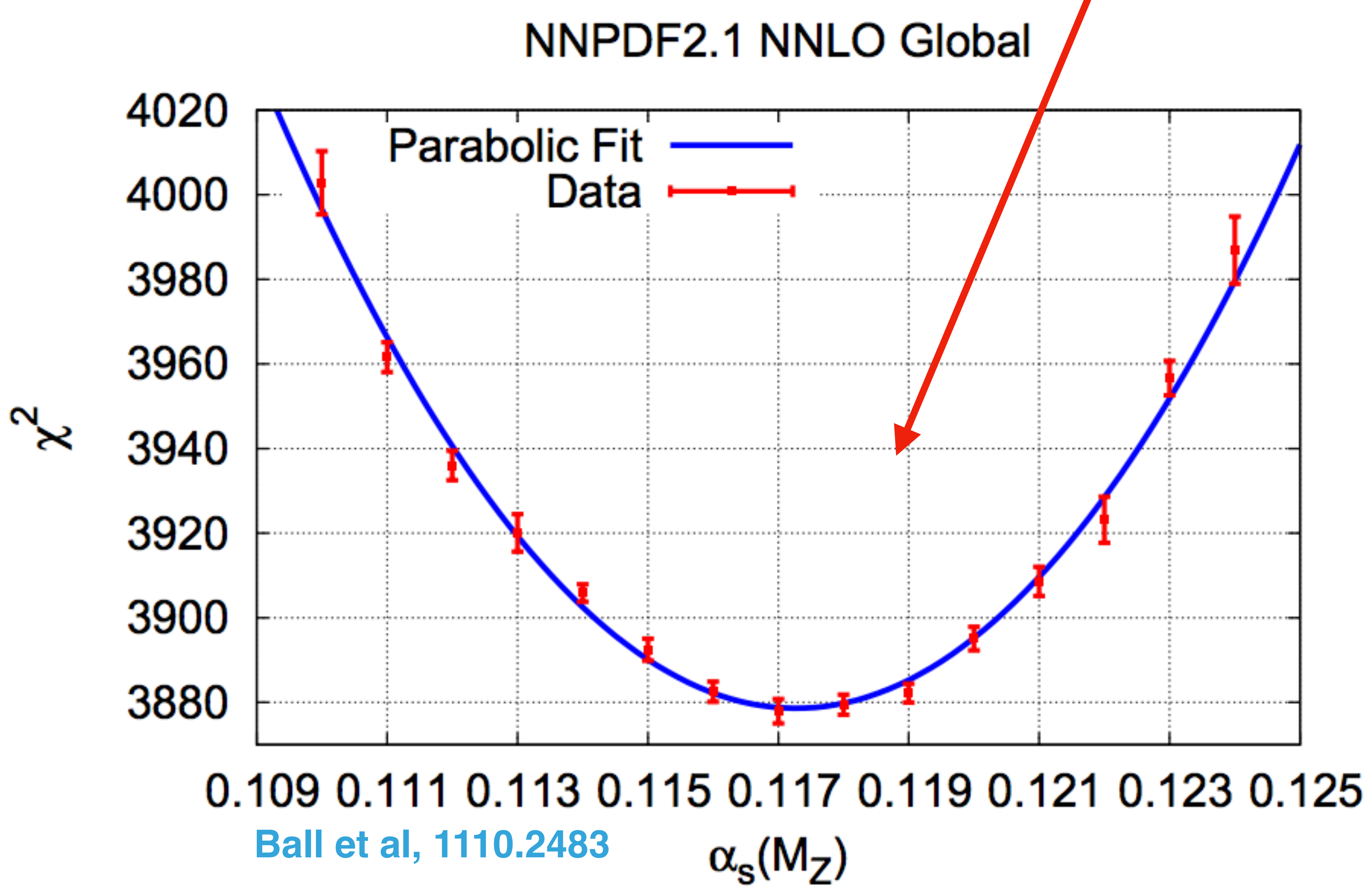
$$T = f_{1,\text{BSM}} \otimes f_{2,\text{BSM}} \otimes \hat{\sigma}_{\text{BSM}}$$

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

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



Ball et al, 1110.2483

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

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

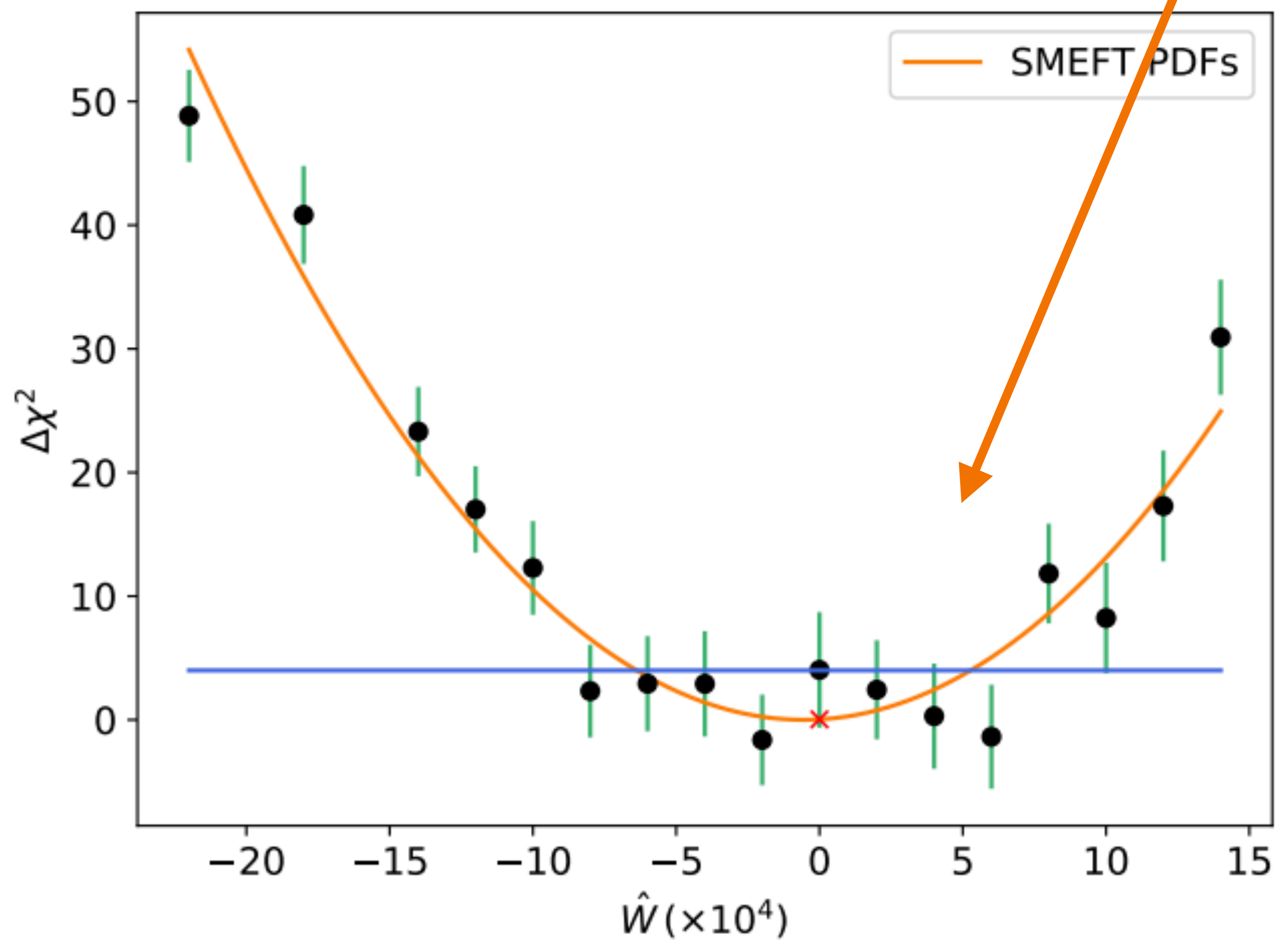
$$T = f_{1,\rm BSM} \otimes f_{2,\rm BSM} \otimes \hat{\sigma}_{\rm BSM}$$

SMEFT PDFs / Simultaneous fit

ANALYSIS METHODOLOGY

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

- 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



Greljo et al, 2104.02723

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

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

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

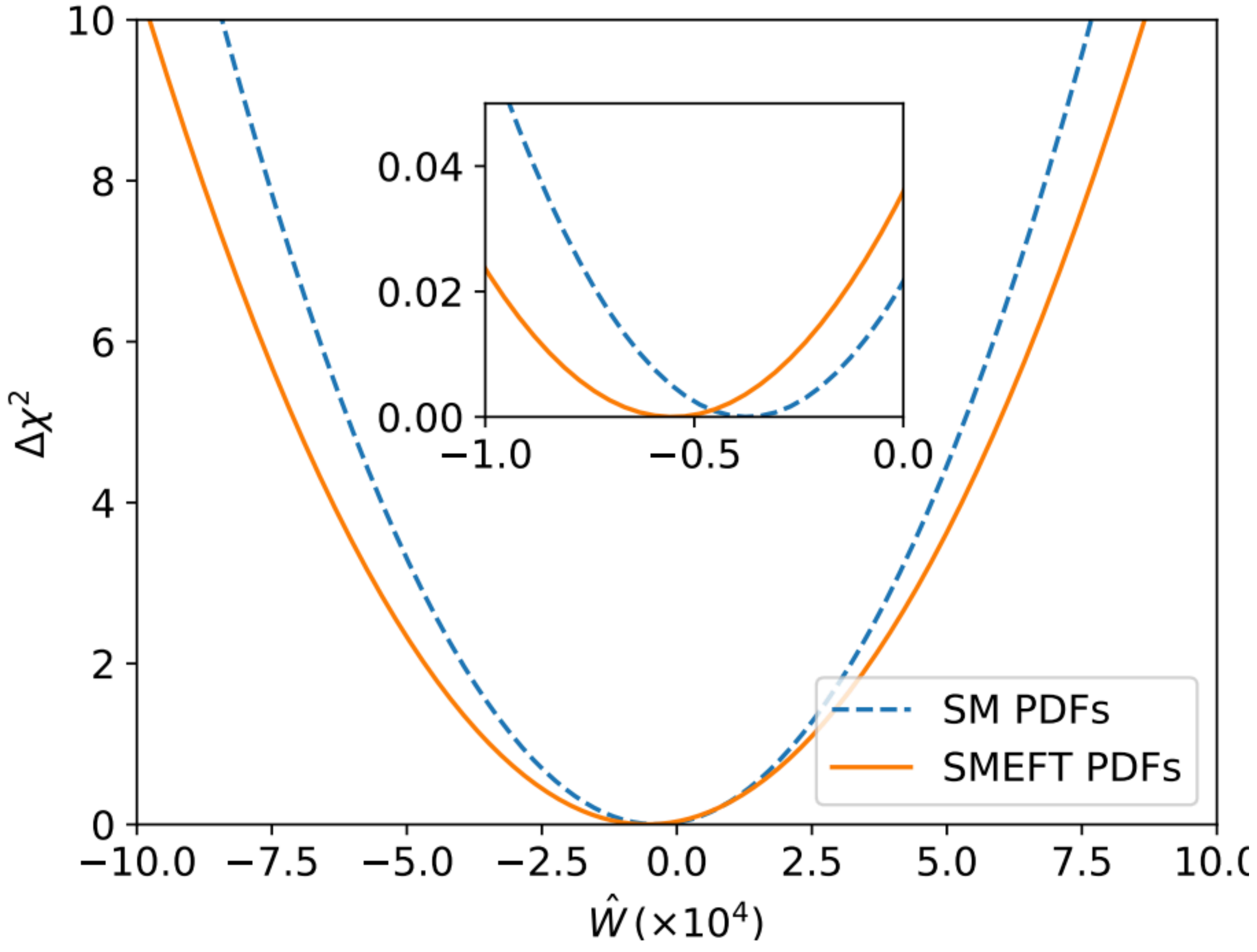
SM PDFs

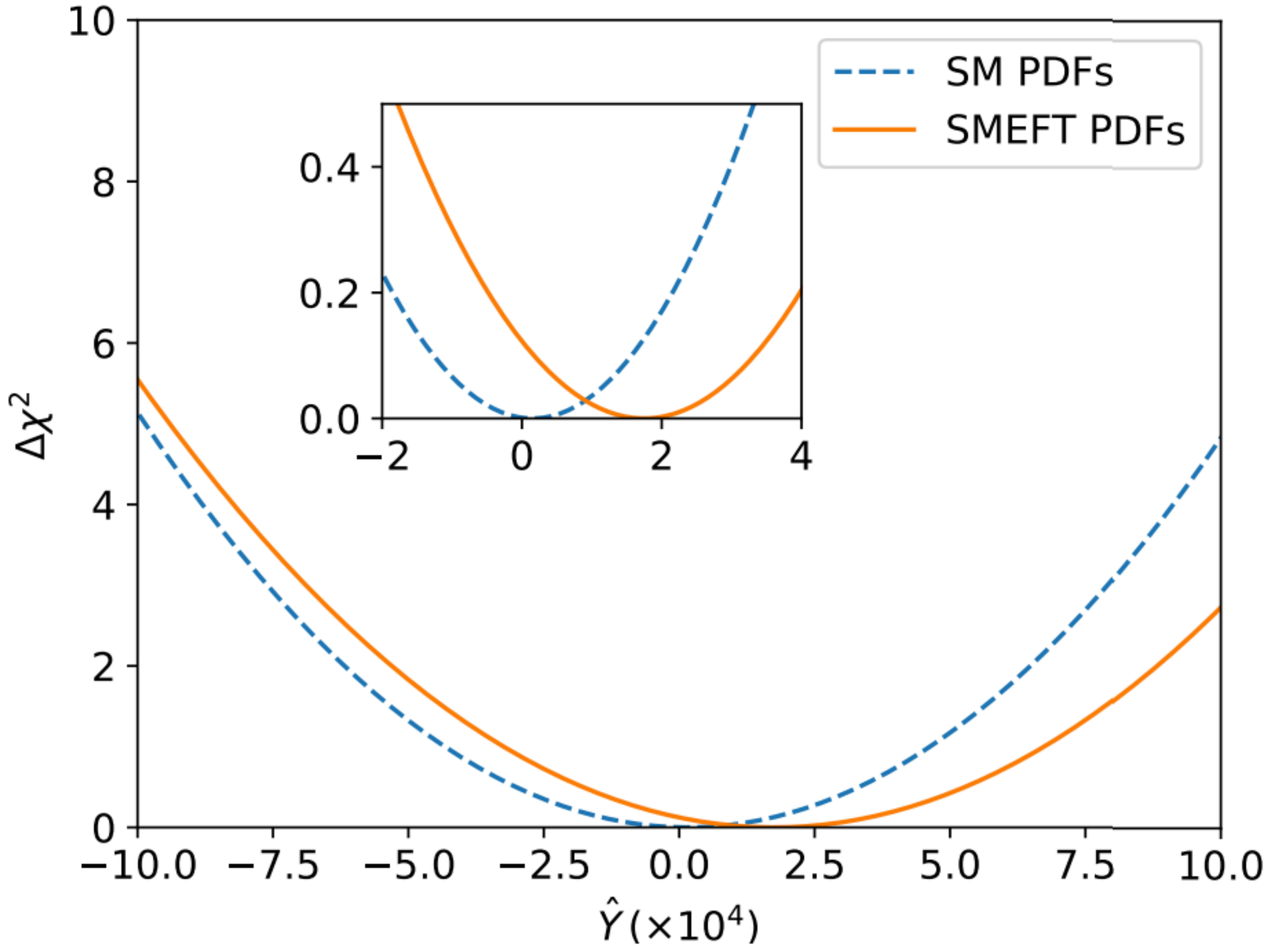
1. Take data, make theoretical predictions accounting for operator in partonic cross section and PDFs.
2. Compute chi2 as a function of WCs (Wilson Coefficients)
3. Minimise chi2 and find best-fit and C.L.s of WCs
4. Extract bounds

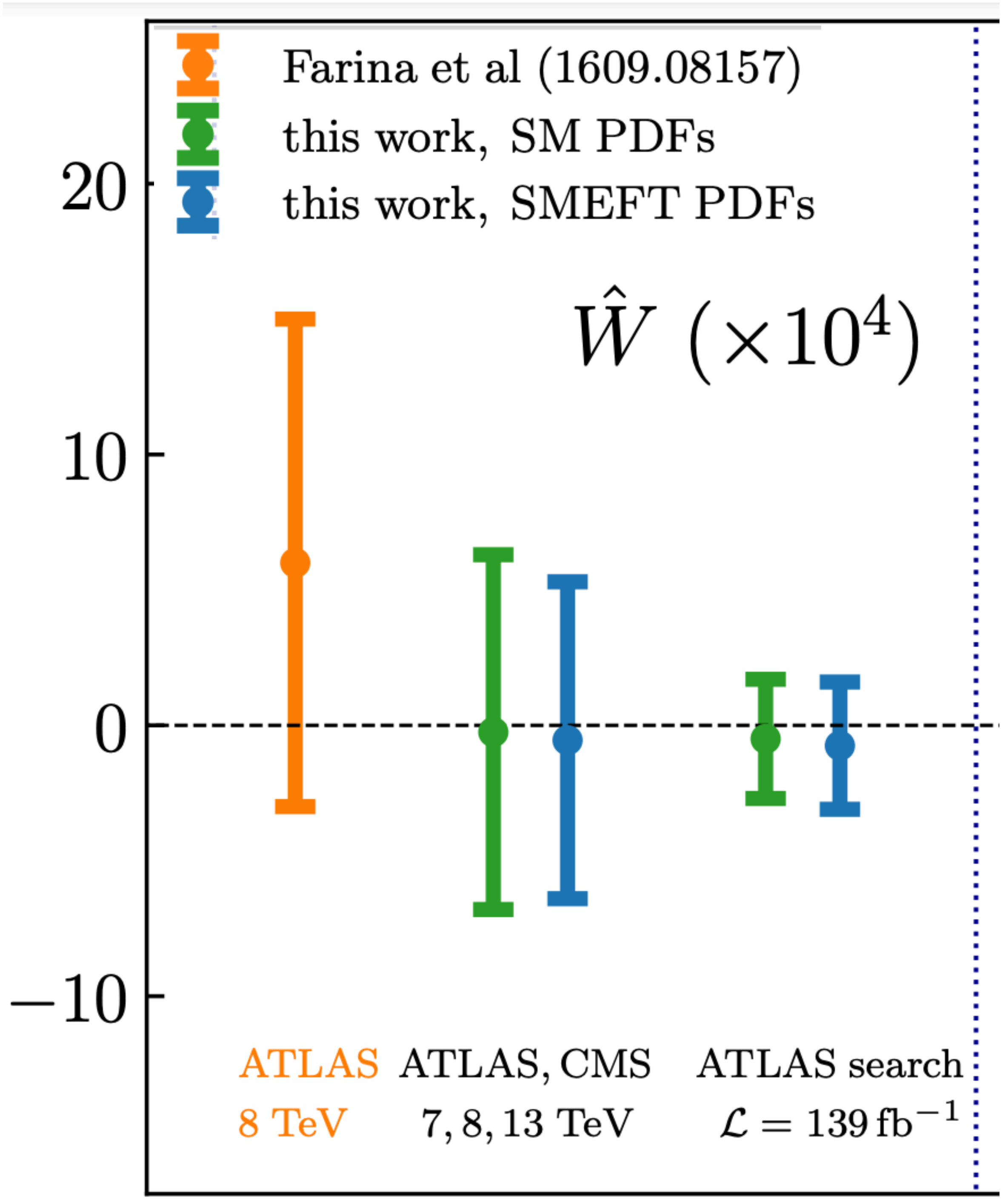
$$T = f_{1,\rm BSM} \otimes f_{2,\rm BSM} \otimes \hat{\sigma}_{\rm BSM}$$

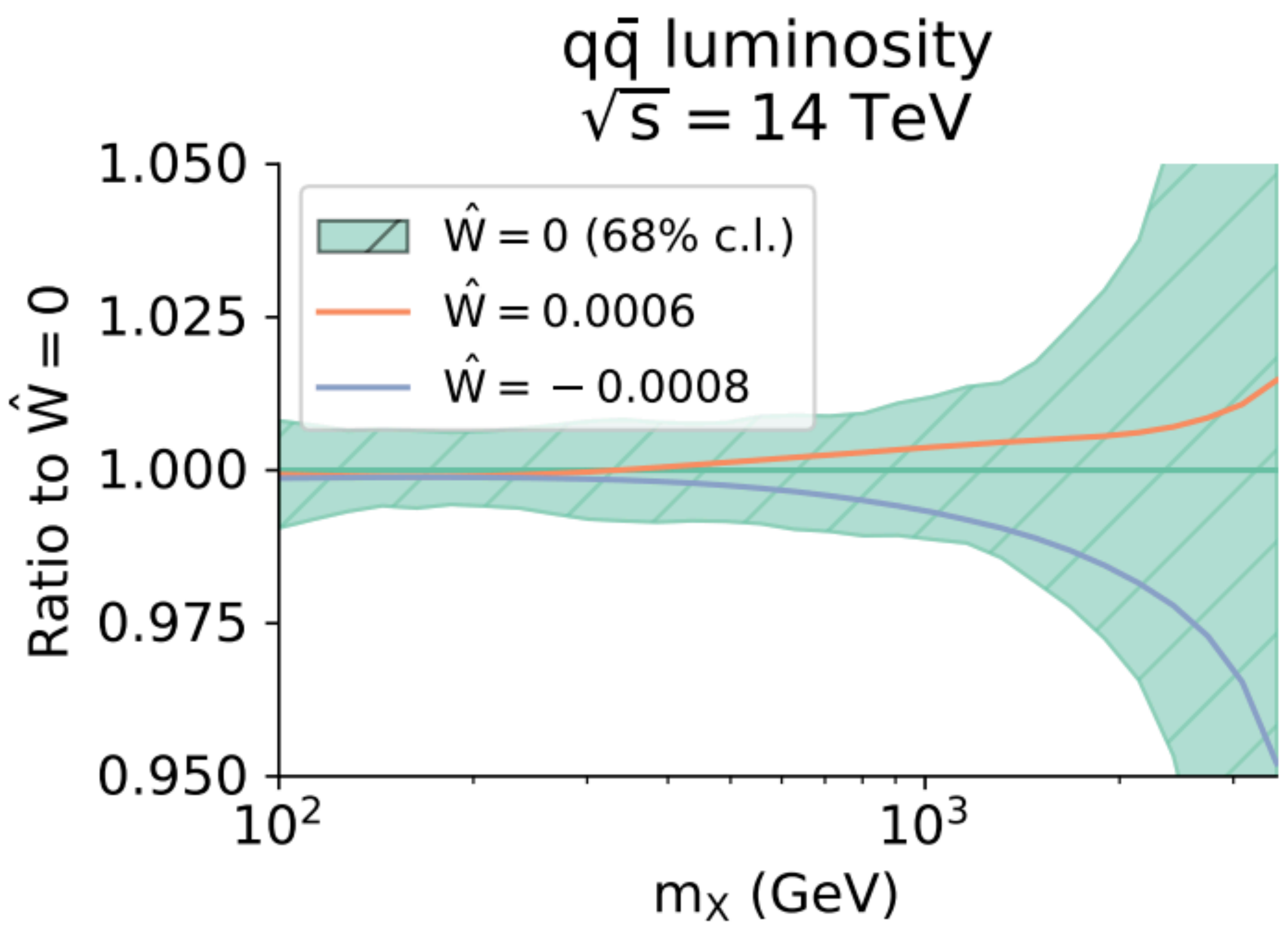
SMEFT PDFs / Simultaneous fit

INTERPLAY @ RUN I AND RUN II









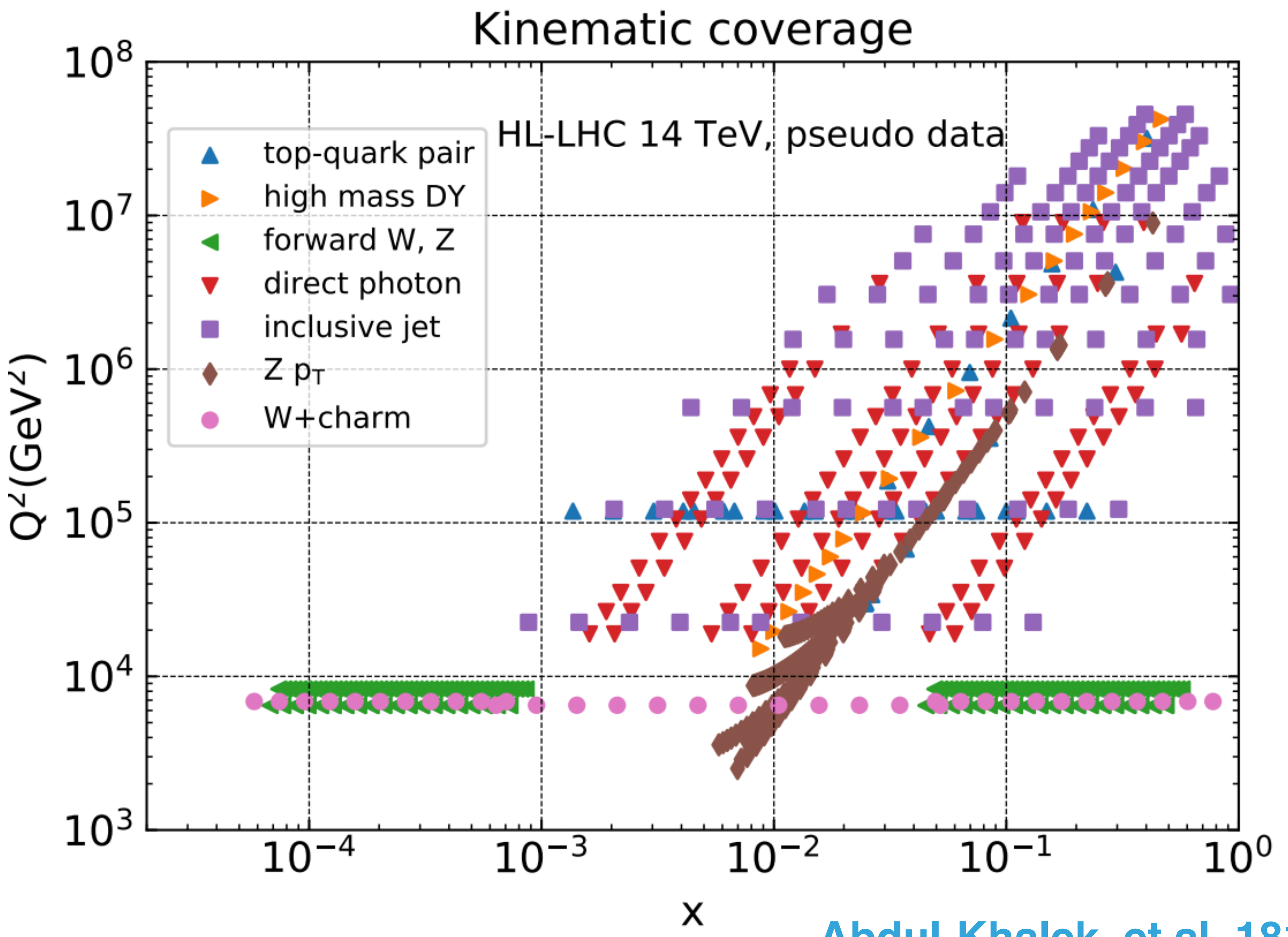
- With current data, PDFs are moderately affected by inclusion of non-zero W and Y coefficients in the fit, mostly quark-antiquark luminosity within uncertainties
- Broadening of individual bounds on W and Y once SMEFT PDFs are used (i.e. PDFs that have been fitted with consistent values of W and Y) 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)

INPUT PDF @ HL-LHC

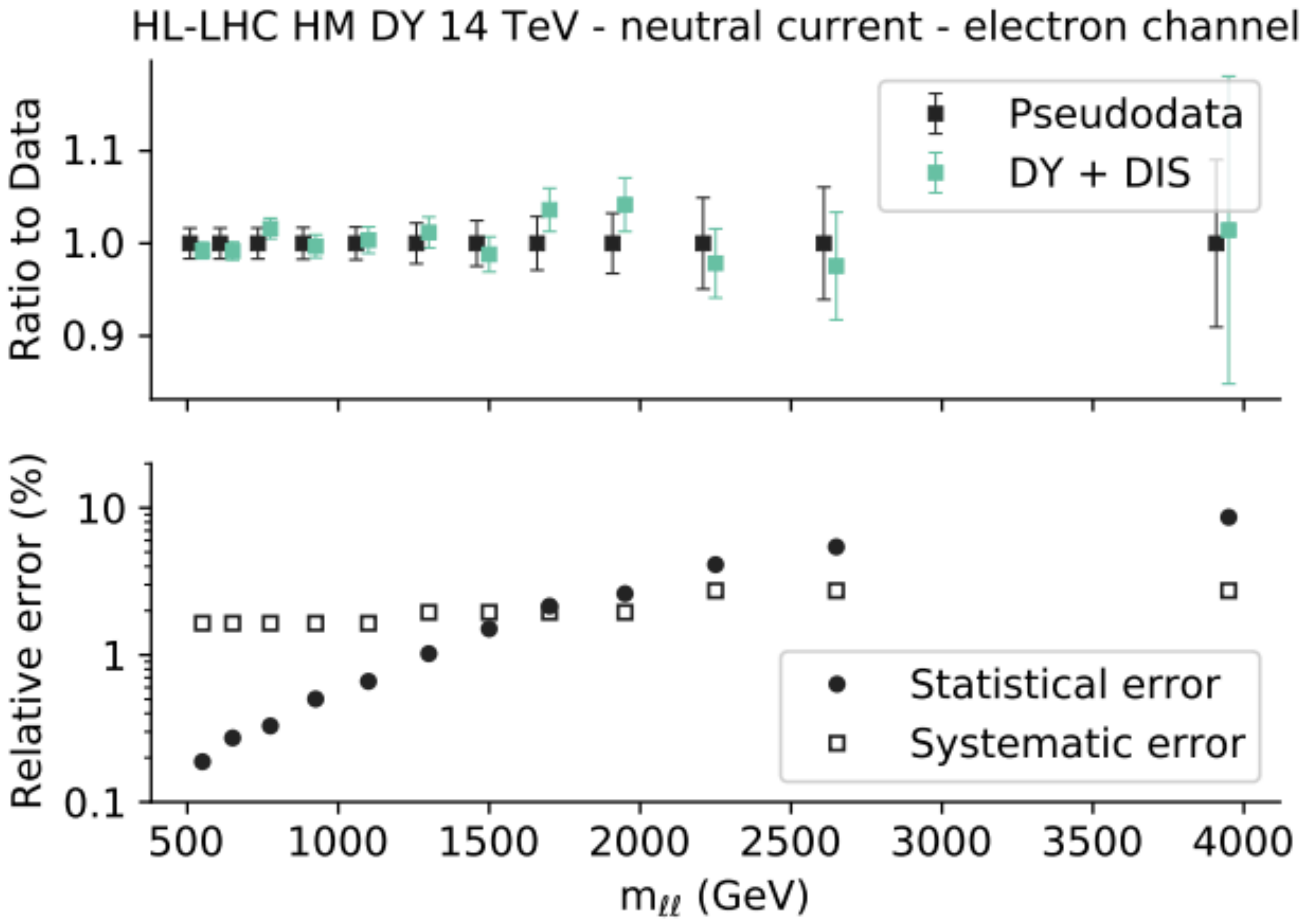
- Add HL-LHC projections for both NC and CC in PDF fit

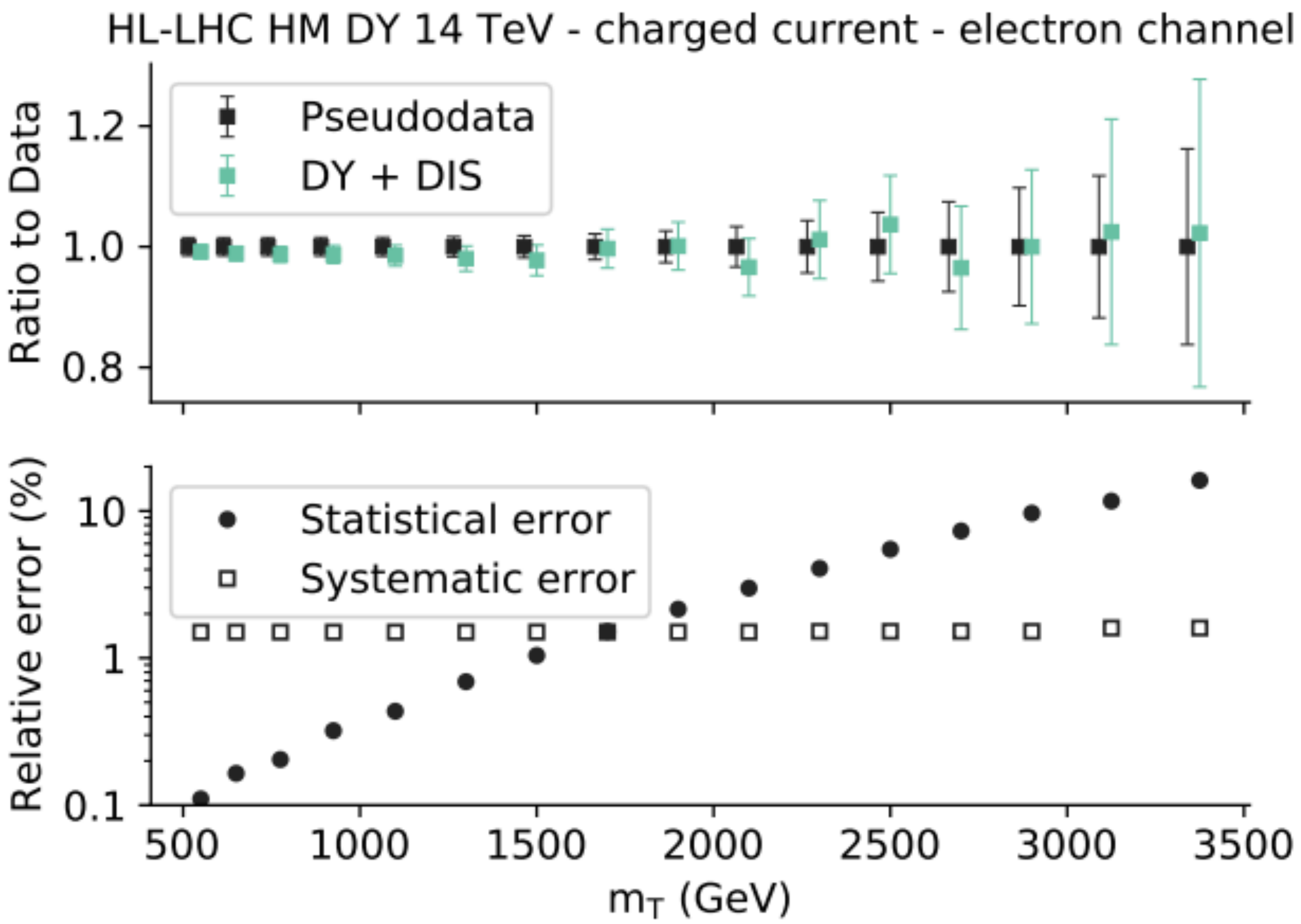
$$\sigma_i^{\rm hllhc} \equiv \sigma_i^{\rm th}\left(1 + \lambda\delta_{\mathcal{L}}^{\rm exp} + r_i\delta_{{\rm tot},i}^{\rm exp}\right), \qquad i = 1, \ldots, n_{\rm bin}$$

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



Abdul-Khalek et al, 1810.03639





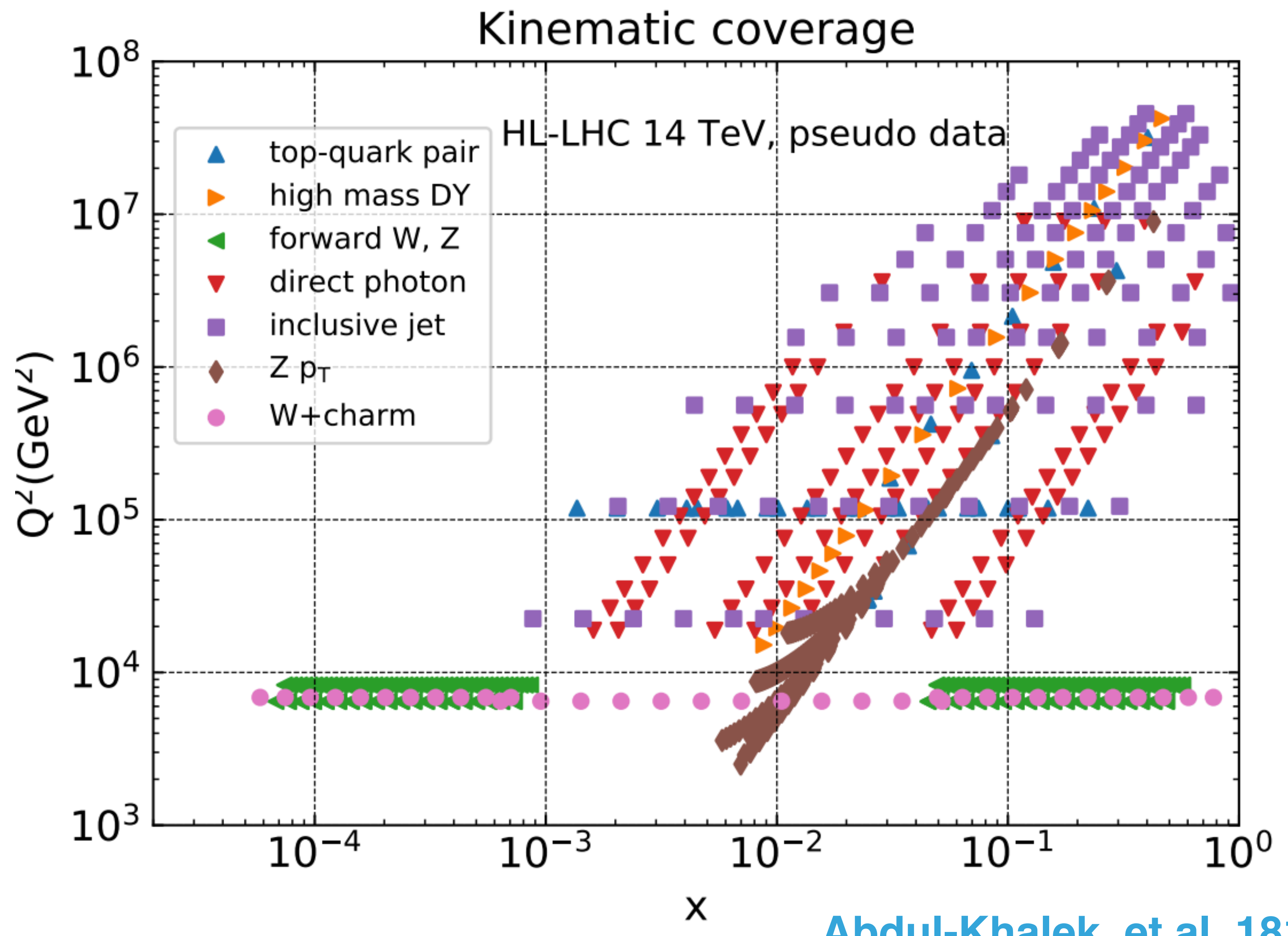
+ same for muon channel

INPUT PDF @ HL-LHC

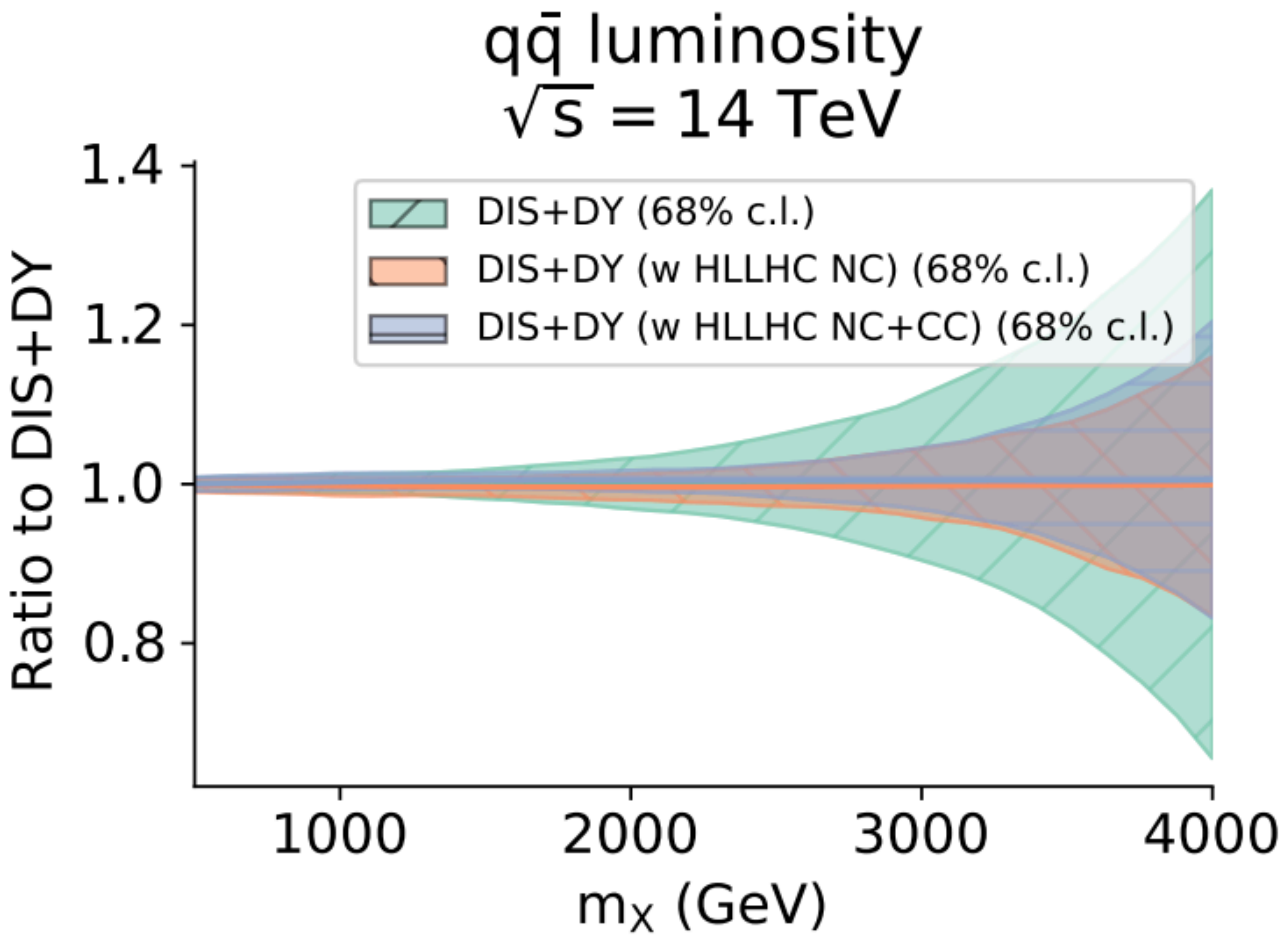
- Add HL-LHC projections for both NC and CC in PDF fit

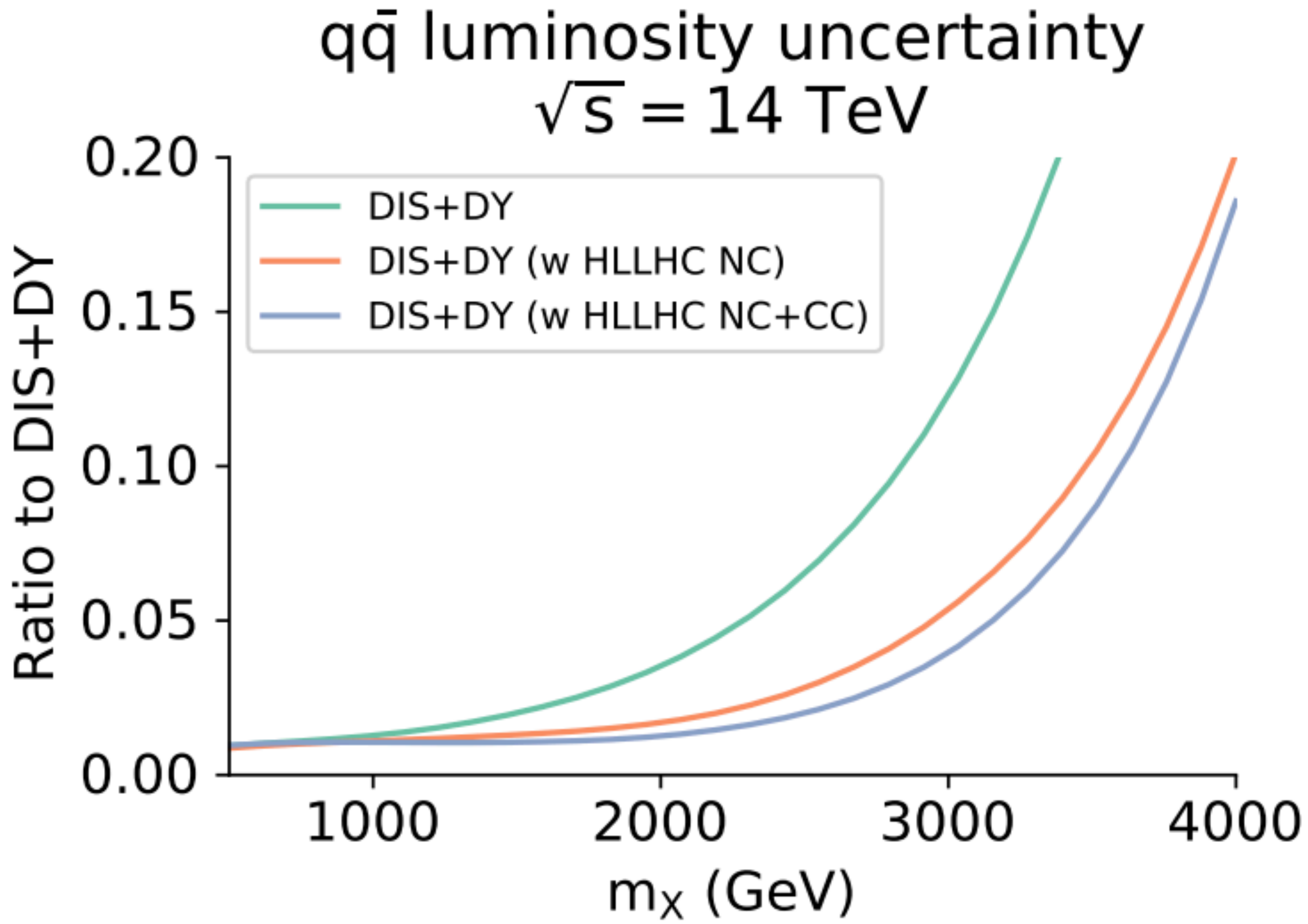
$$\sigma_i^{\rm hllhc} \equiv \sigma_i^{\rm th}\left(1 + \lambda\delta_{\mathcal{L}}^{\rm exp} + r_i\delta_{{\rm tot},i}^{\rm exp}\right)\,, \qquad i = 1,\ldots,n_{\rm bin}$$

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



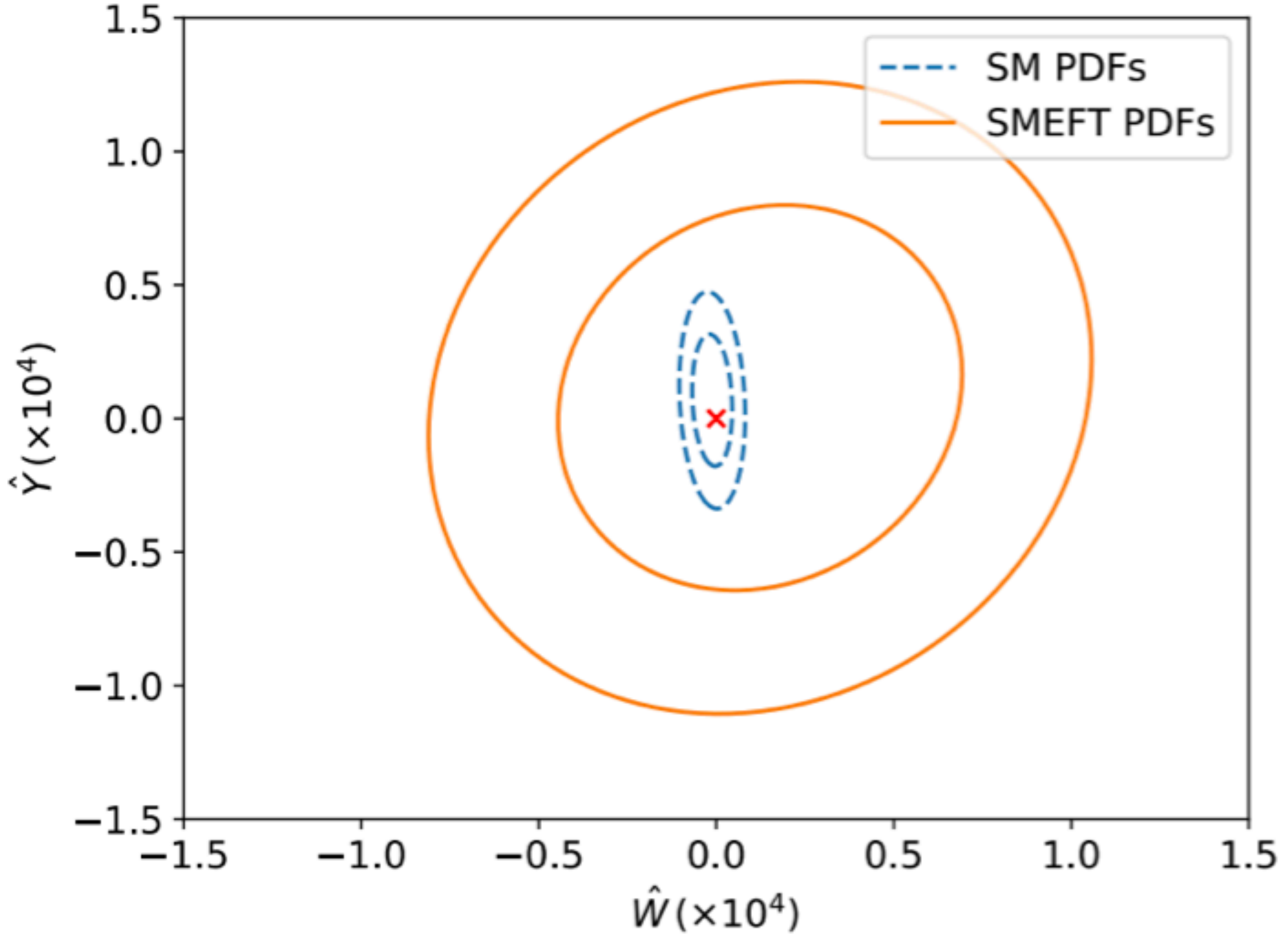
Abdul-Khalek et al, 1810.03639





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



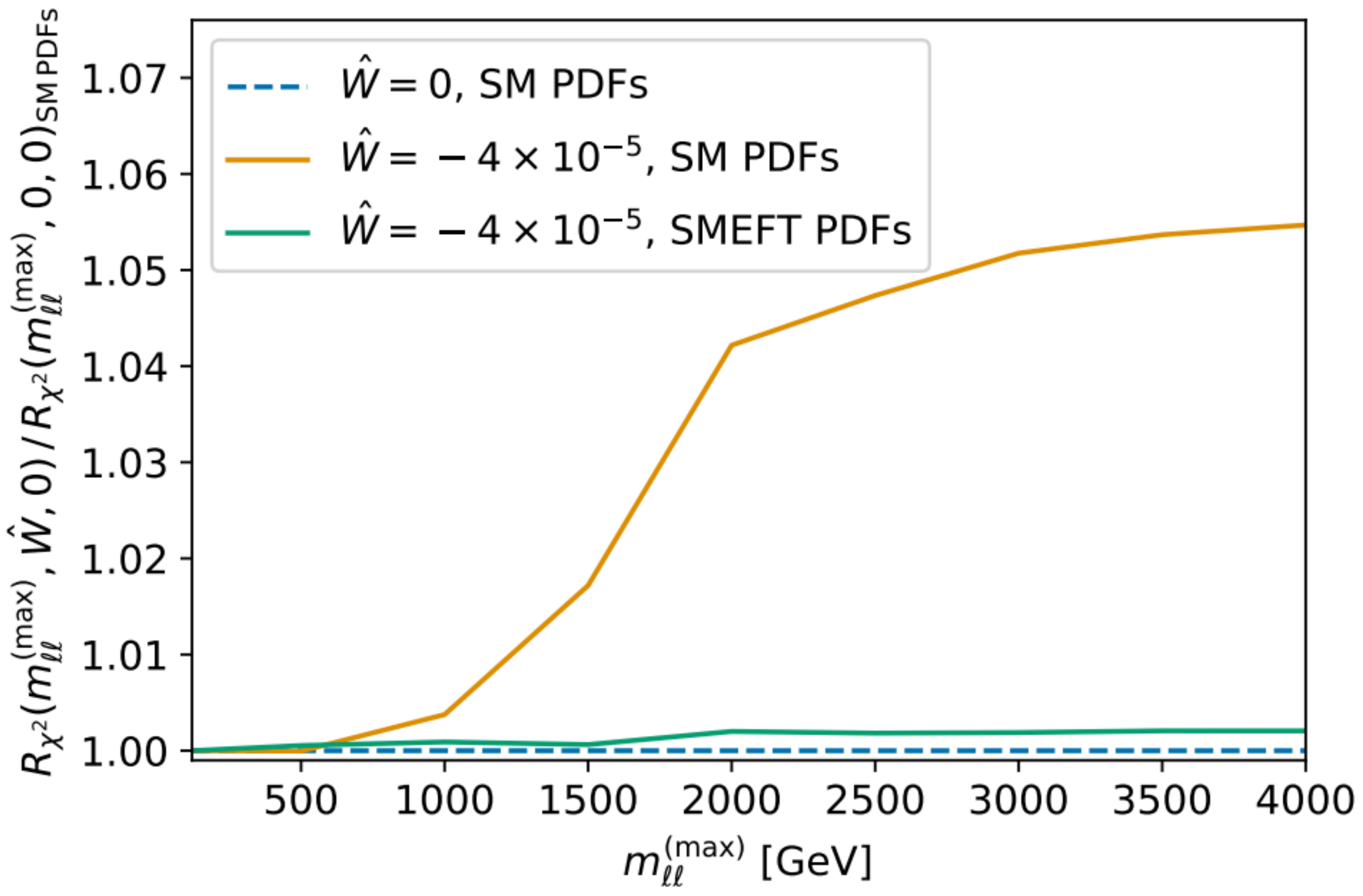
	SM PDFs	SMEFT PDFs	best-fit shift	broadening
$\hat{W} \times 10^5$ (68% CL)	$[-0.7, 0.5]$	$[-4.5, 6.9]$	1.3	850%
	$[-1.0, 0.9]$		1.3	500%
$\hat{W} \times 10^5$ (95% CL)	$[-1.0, 0.8]$	$[-8.1, 10.6]$	1.4	940%
	$[-1.4, 1.2]$		1.4	620%
$\hat{Y} \times 10^5$ (68% CL)	$[-1.8, 3.2]$	$[-6.4, 8.0]$	0.1	190%
	$[-3.7, 4.7]$		0.3	70%
$\hat{Y} \times 10^5$ (95% CL)	$[-3.4, 4.7]$	$[-11.1, 12.6]$	0.1	190%
	$[-5.3, 6.3]$		0.3	110%

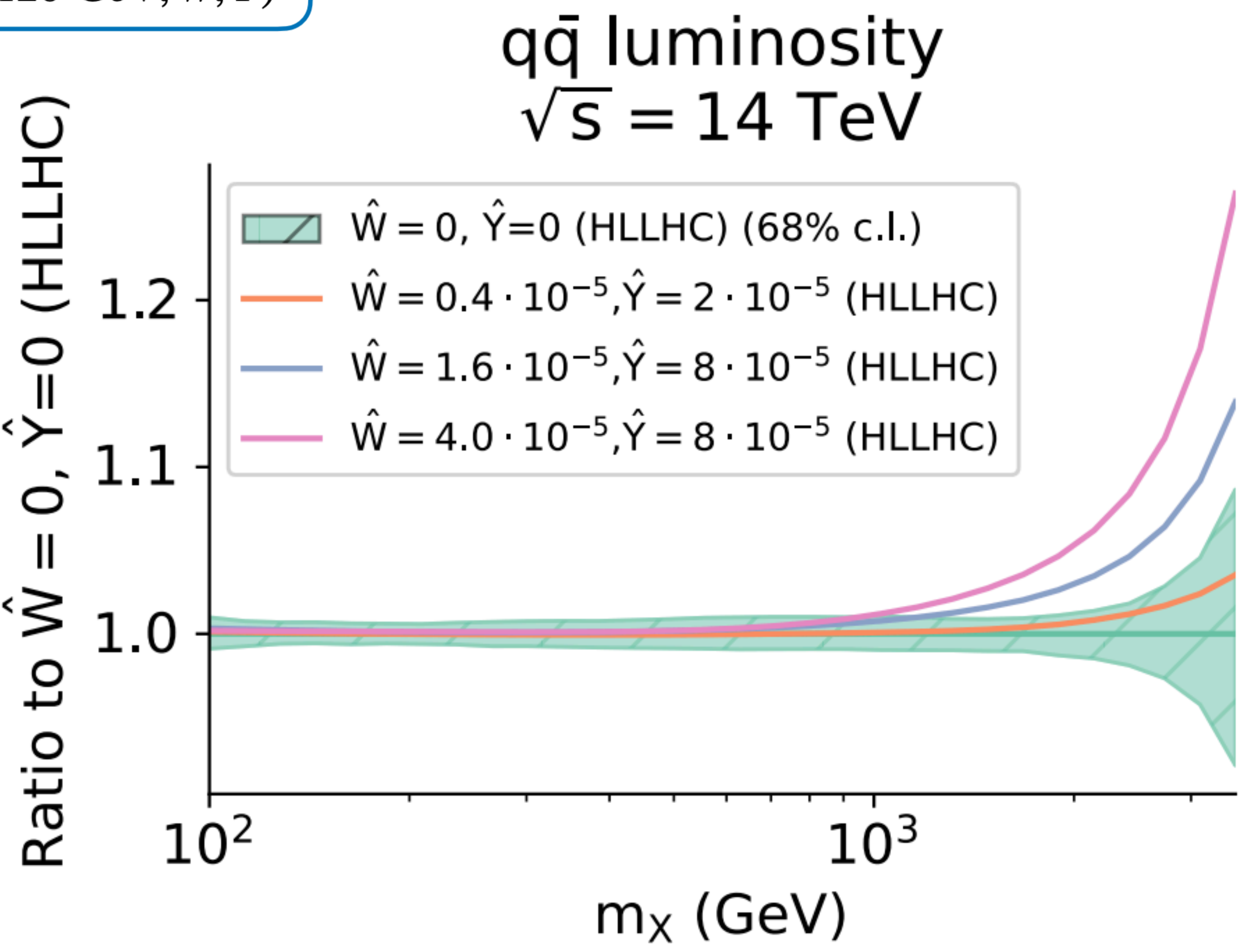
Greljo et al, 2104.02723

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!

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

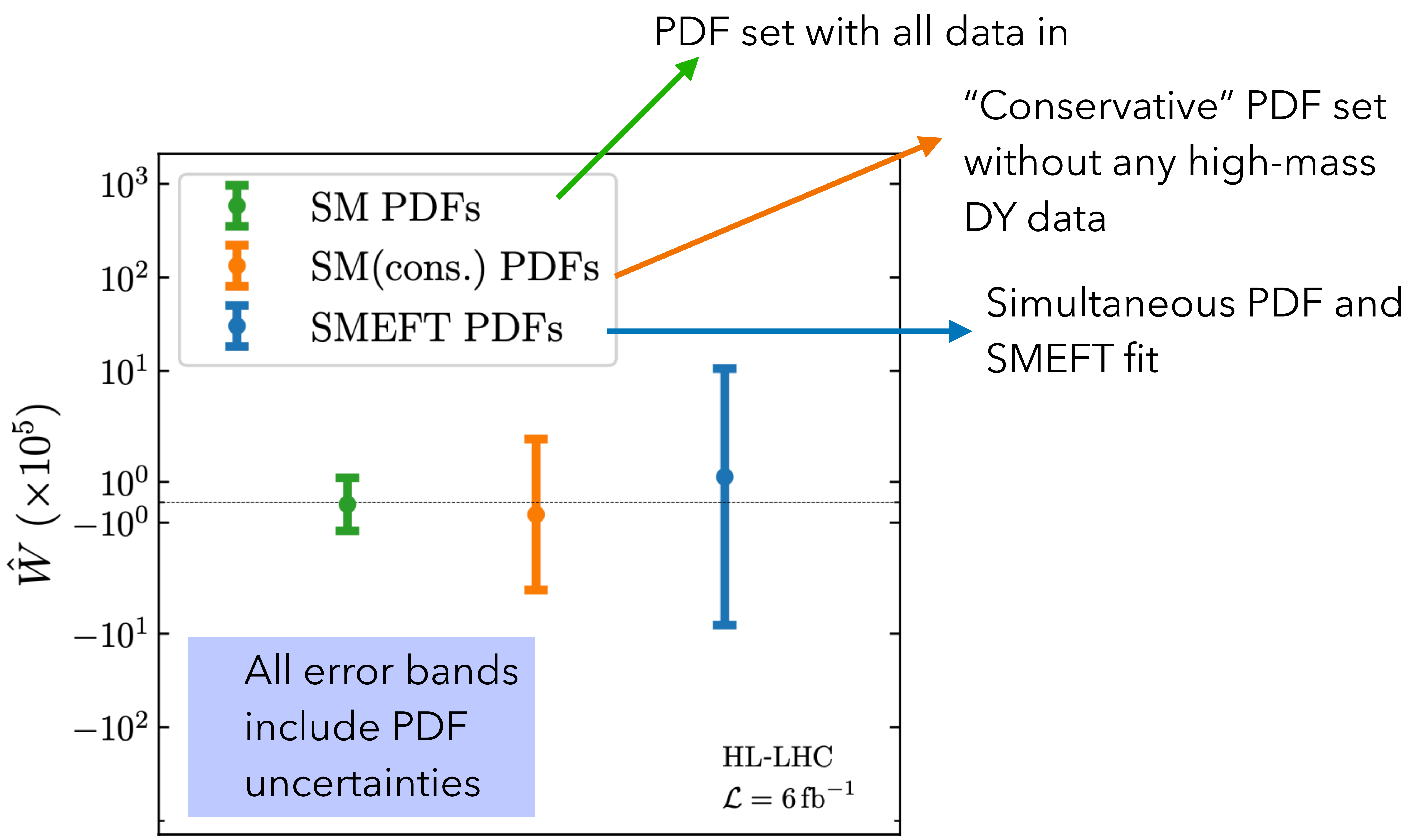




Greljo et al, 2104.02723

INTERPLAY @ HL-LHC

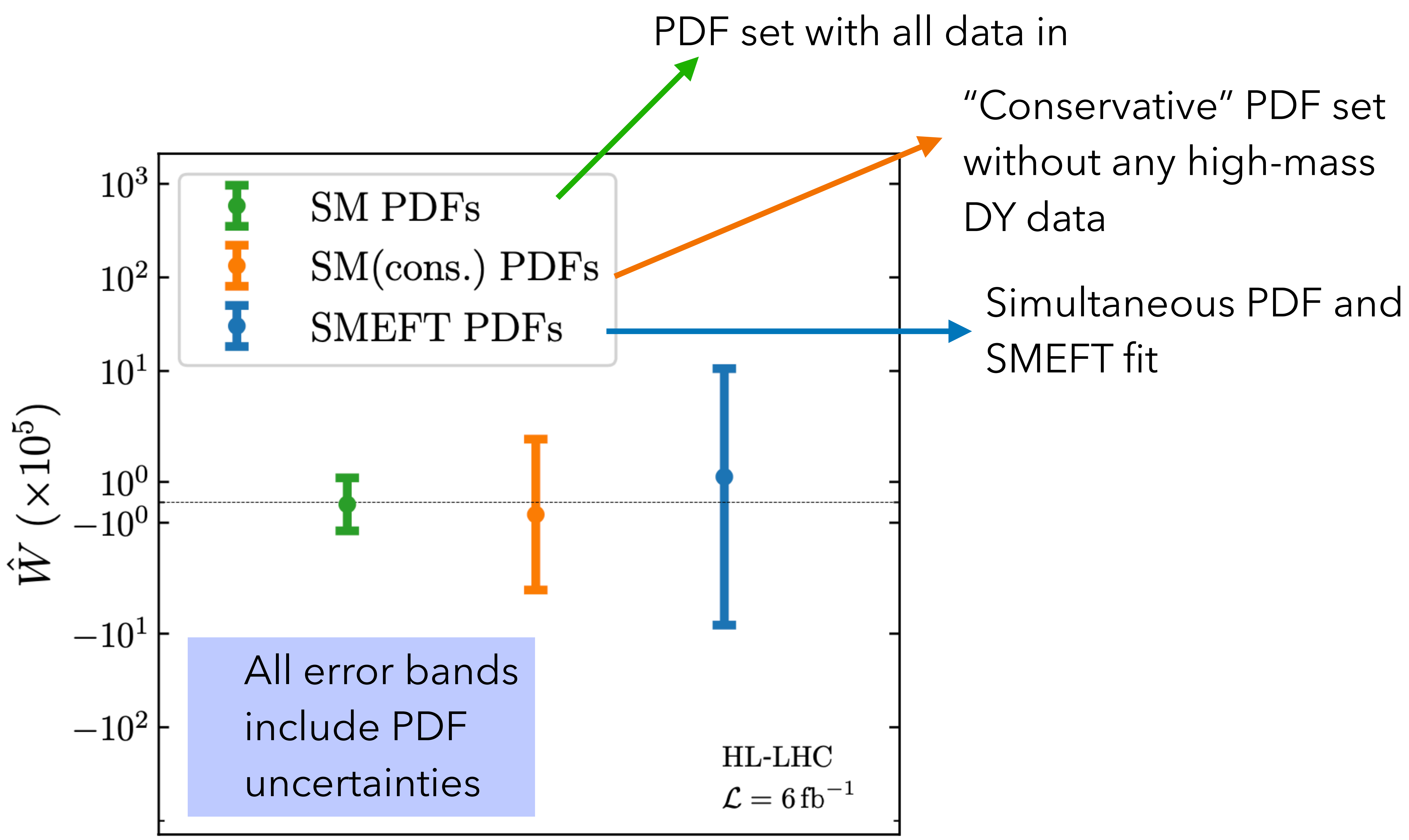
- Compare Wilson coefficients bounds from HL-LHC projections assuming SM PDFs (that include NC+CC data) to the bounds on the same Wilson coefficients obtained from a simultaneous fit of PDFs and Wilson coefficients
- Not accounting for interplay (using PDFs as a black box) leads to over-constrained bounds
- PDFs do absorb effect of new physics in this case!
- What if we use a clean or "conservative" set of PDFs that does not include any high-mass Drell-Yan data?



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!
- What if we use a clean or "conservative" set of PDFs that does not include any high-mass Drell-Yan data?

Note that HL-LHC projections based on pseudo-data. If new physics was there, simultaneous fits of SMEFT and PDFs could point to a different minimum, and not only to larger uncertainties!

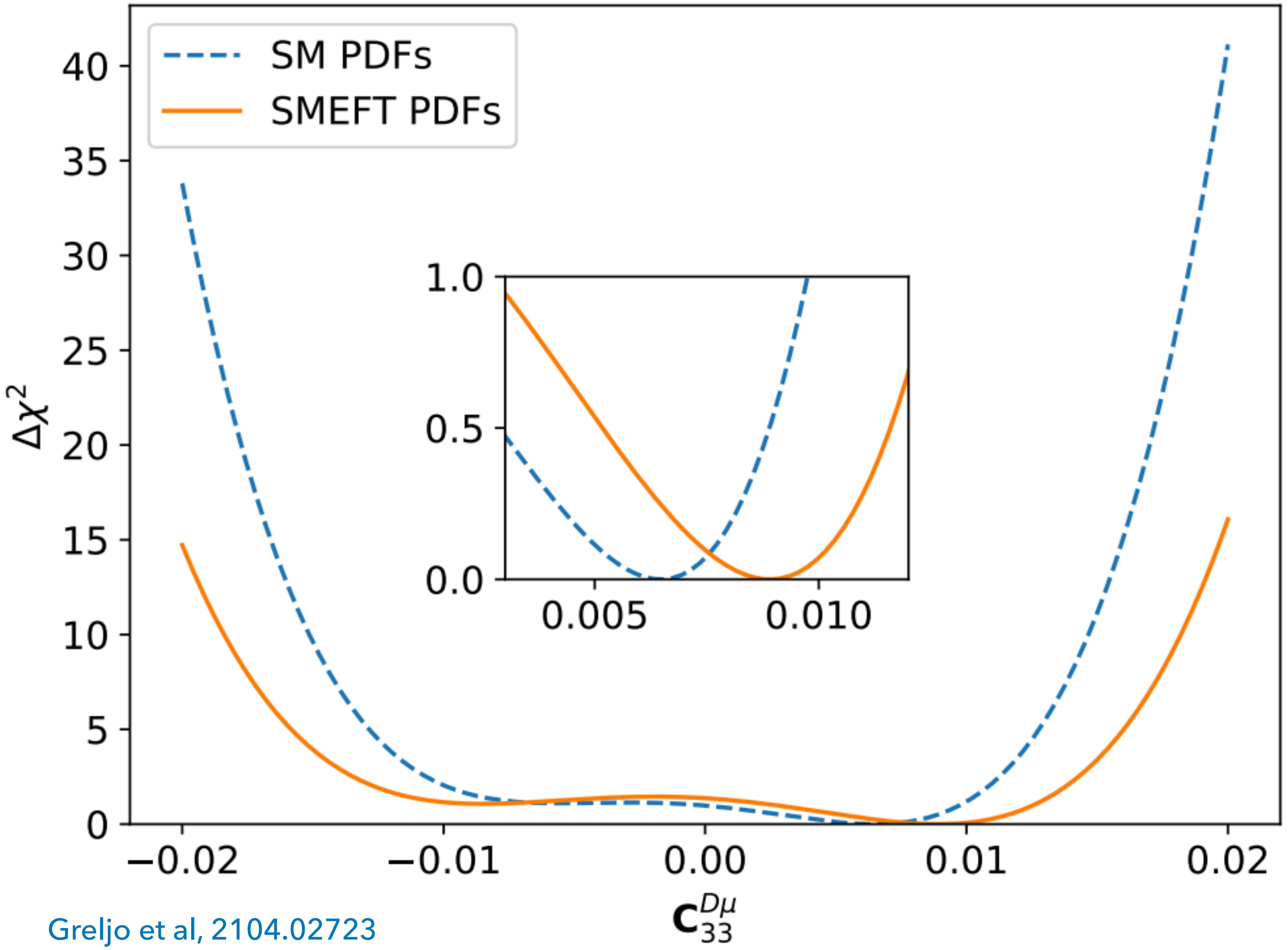


FLAVOUR-SPECIFIC 4-FERMIONS OPERATORS COUPLING MUONS AND B-QUARKS

- Consider a scenario with a single non-zero WC among gauge invariant four-fermion operators built from the SM quark and lepton SU(2)$_L$ doublets
- If the observed deviations in R(K$^{(*)}$) due to new physics, generically expect $|C^{D\mu}_{33}| \gtrsim 0.001$

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \frac{\mathbf{C}^{D\mu}_{33}}{v^2}(\bar{d}^3_L\gamma_\mu d^3_L)(\bar{\mu}_L\gamma^\mu\mu_L)$$

	SM PDFs	SMEFT PDFs	best-fit shift	broadening
$\mathbf{C}^{D\mu}_{33} \times 10^2$ (68% CL)	$[-0.1, 1.1]$	$[-0.3, 1.2]$	0.06	25%
$\mathbf{C}^{D\mu}_{33} \times 10^2$ (95% CL)	$[-1.0, 1.2]$	$[-1.2, 1.4]$	0.06	18%



Greljo et al, 2104.02723

- From PDF point of view, new physics only in Drell-Yan muon data and PDF constrained by Drell-Yan electron data

- Measurements in separate leptonic final states is of utmost importance to test BSM scenarios that account for violations of Lepton Flavour Universality.

CONCLUSIONS AND OUTLOOK

- Time to study the interplay between indirect new physics searches via EFT fits and PDFs

- Current status:
  - ➡ Comparing results of simultaneous fit and of fits assuming SM PDFs starts shedding some lights on the issue
  - ➡ Run I and Run II high-energy Drell-Yan data: the effect of the interplay is visible but is still within PDF uncertainties
  - ➡ Search data: bounds broaden compared to SM PDFs (beyond PDF uncertainties)
  - ➡ HL-LHC: Not accounting for interplay (using PDFs as a black box) leads to over-constrained bounds
  - ➡ HL-LHC: Conservative PDFs still yield stronger bounds than simultaneous fit

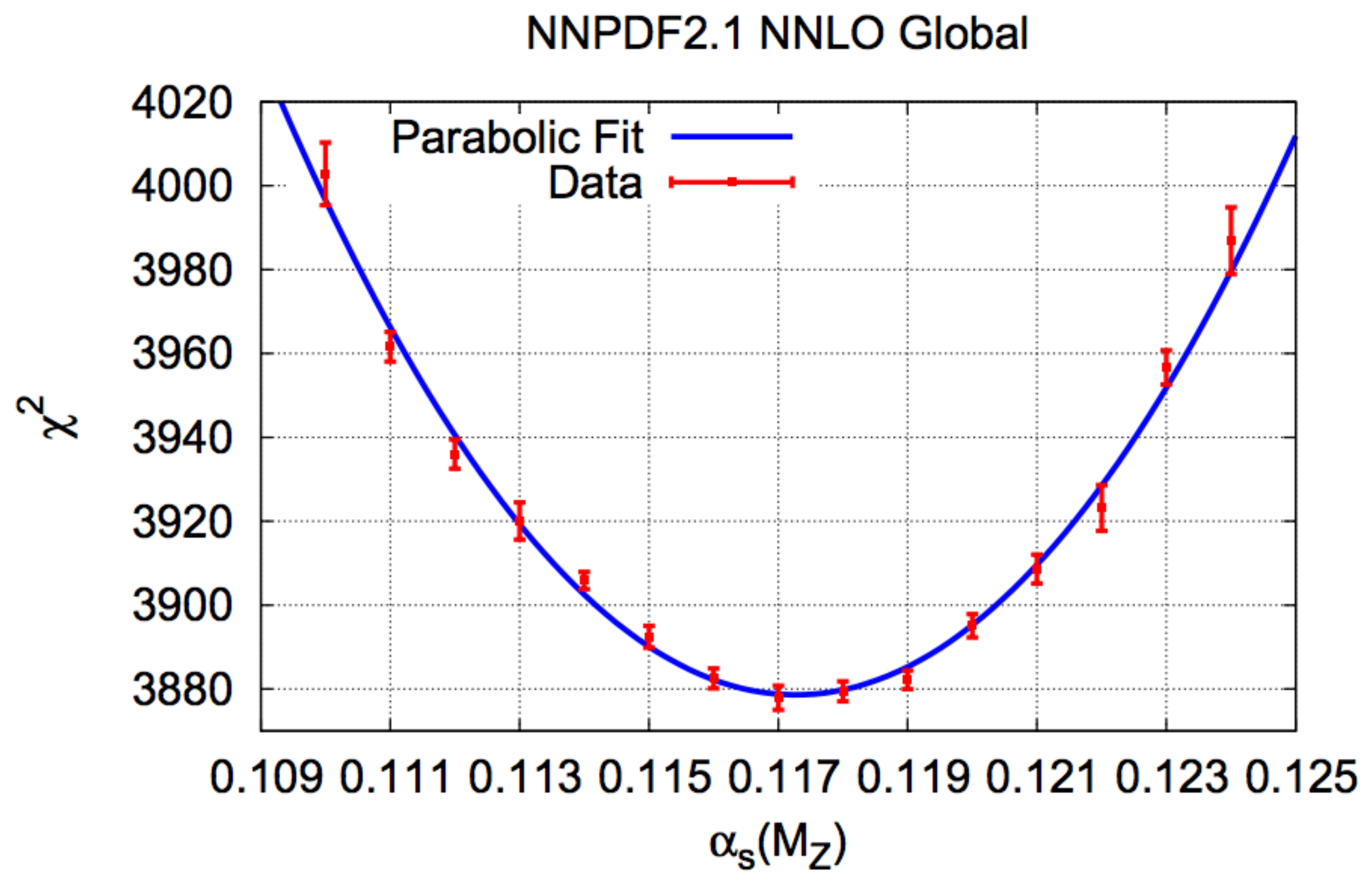
- The way ahead:
  - ➡ The preferred avenue ahead is to be able to perform simultaneous fits (like for PDFs and α_s)
  - ➡ Current methodology not devised to deal with many operators
  - ➡ More powerful methodology is work in progress
  - ➡ In parallel a more careful investigation of definition of conservative PDF sets & account for PDF uncertainties
  - ➡ Also, would be important to disentangle large-x from high-energy / low-energy (LHCb) as well as scaling behaviour (ratios at different centre of mass energies?)

THANK YOU FOR YOUR ATTENTION!

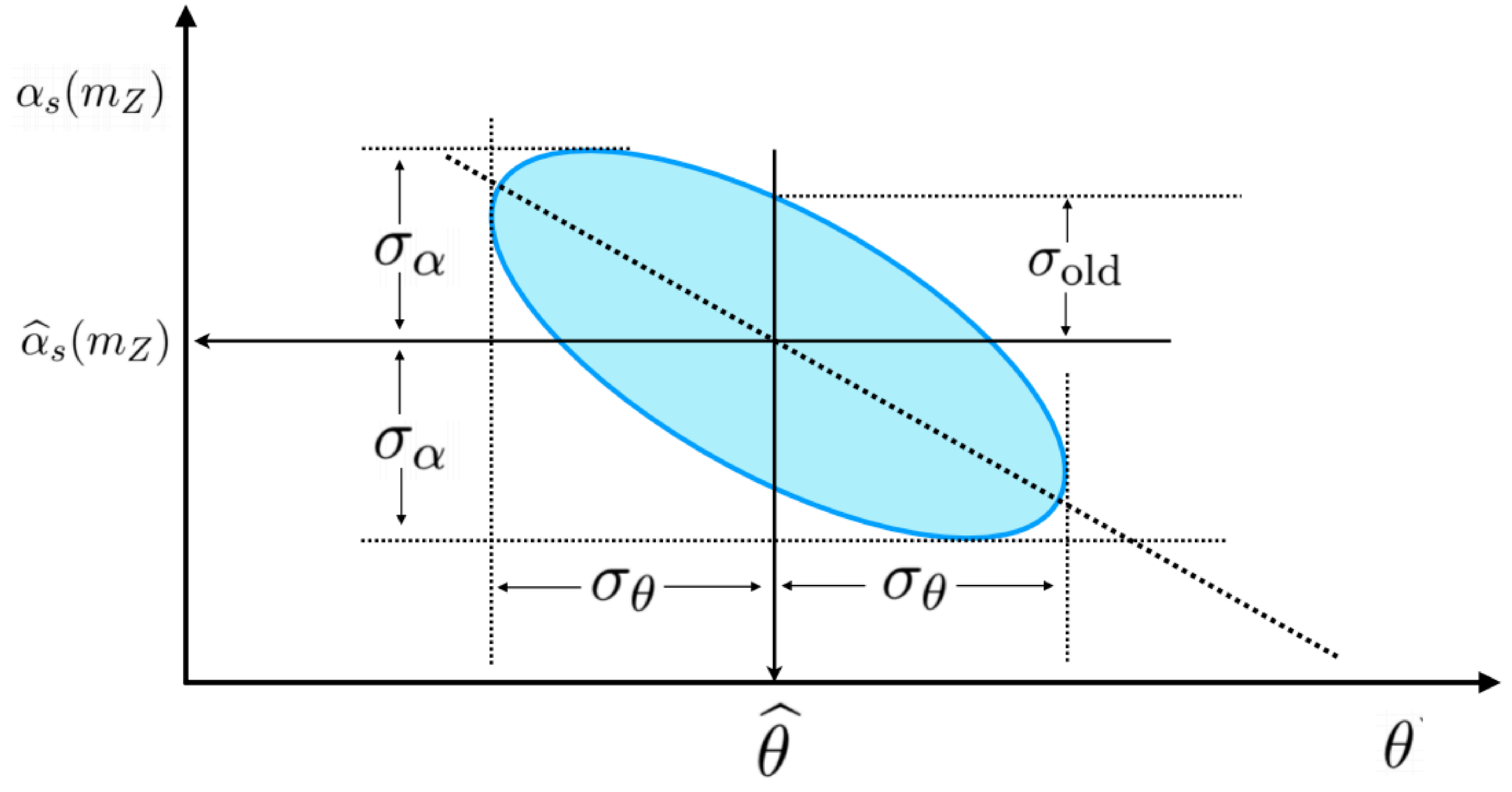
EXTRA MATERIAL

SIMULTANEOUS FITS

- 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 et al, 1110.2483

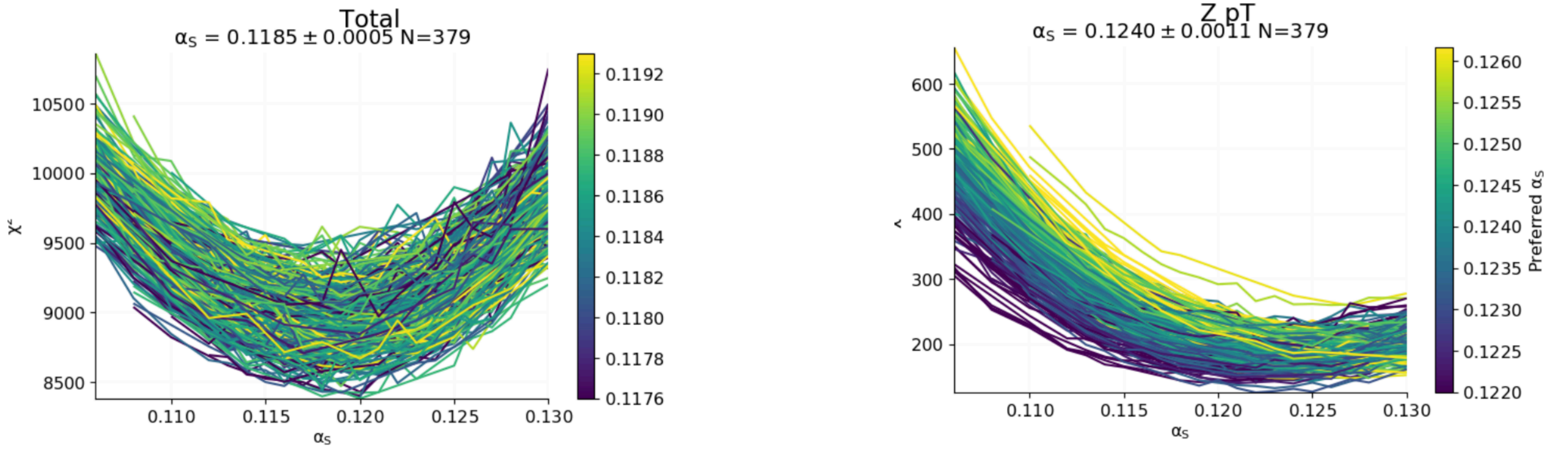


Ball, Carrazza, Del Debbio, Forte, Kassabov, Rojo, Slade, MU 1802.03398

- 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

- However note that at the current level of precision, the determination of the strong coupling constant from the precise measurement of a process at the LHC might be problematic
- Given the strong correlation between PDFs of the proton and α_s, only simultaneous determination of α_s along with the PDFs gives reliable result [Forte, Kassabov 2001.04986]



These results point towards the need of new generation of global fits, in which all ingredients that enter theoretical predictions are treated consistently.