Towards a global analysis of the Standard Model Effective Field Theory

Juan Rojo
VU Amsterdam & Theory group, Nikhef

Particle Physics seminar

University of Oregon at Eugene

06/08/2019

Particle Physics in the LHC precision era

Open questions in particle physics

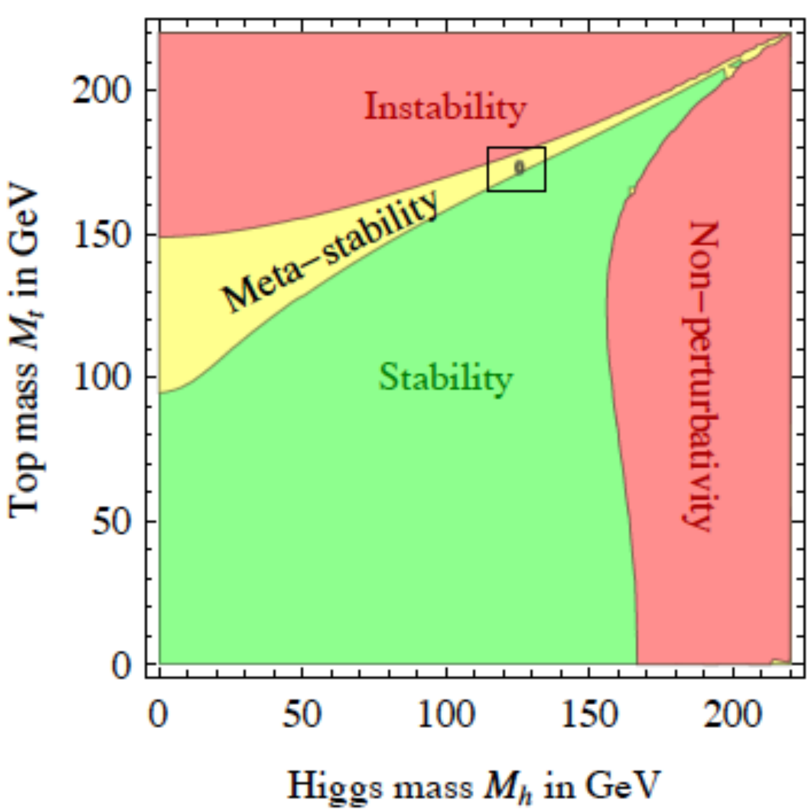
The Higgs boson

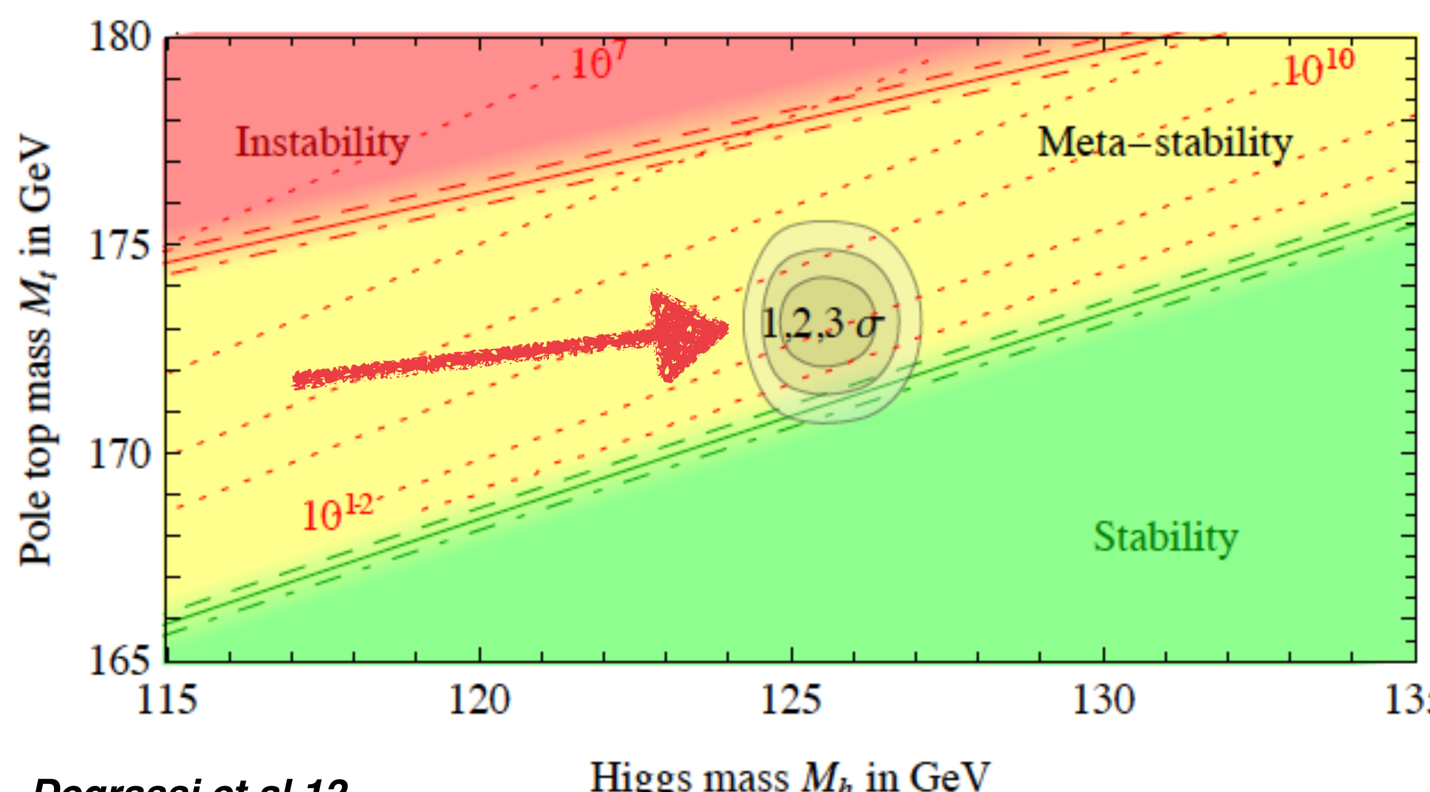
- Huge gap between **weak** and **Plank scales?**
- **Compositeness**? Non-minimal Higgs sector?
- Coupling to **Dark Matter**? Role in cosmological phase transitions?
- Is the **vacuum state of the Universe** stable?

- **10^{19} GeV** (Planck scale)

What defines the value of the weak scale? Why so different from Planck scale?

- **125 GeV** (Higgs mass)
- 1 GeV (Proton mass)





Degrassi et al 12

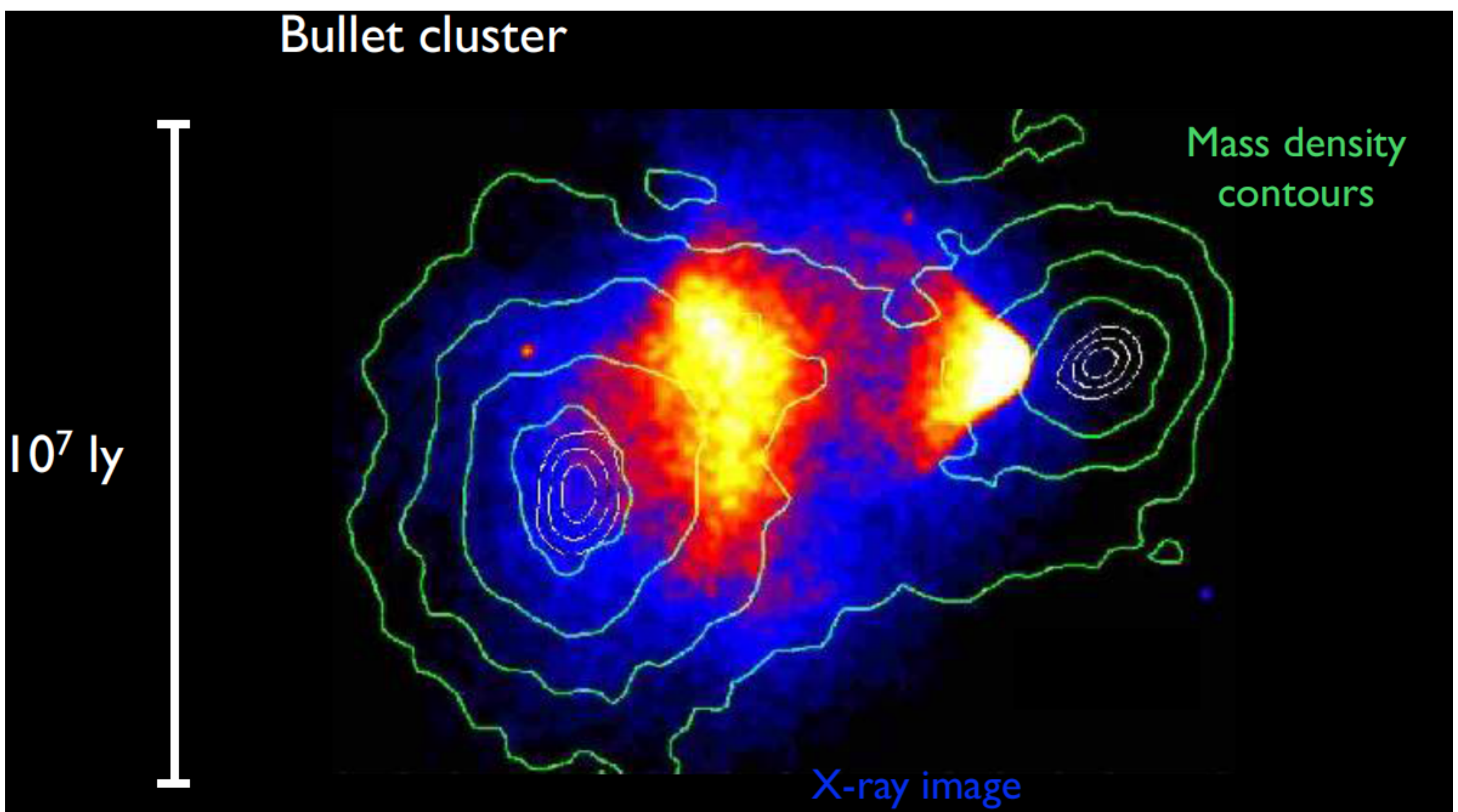
Open questions in particle physics

The Higgs boson

- Huge gap between **weak** and **Plank scales?**
- **Compositeness**? Non-minimal Higgs sector?
- Coupling to **Dark Matter**? Role in cosmological phase transitions?
- Is the **vacuum state of the Universe** stable?

Dark matter

- **Weakly interacting massive particles**? Neutrinos? Ultralight particles (axions)?
- **Interactions** with SM particles? Self-interactions?
- **Structure** of the Dark Sector?



Open questions in particle physics

The Higgs boson

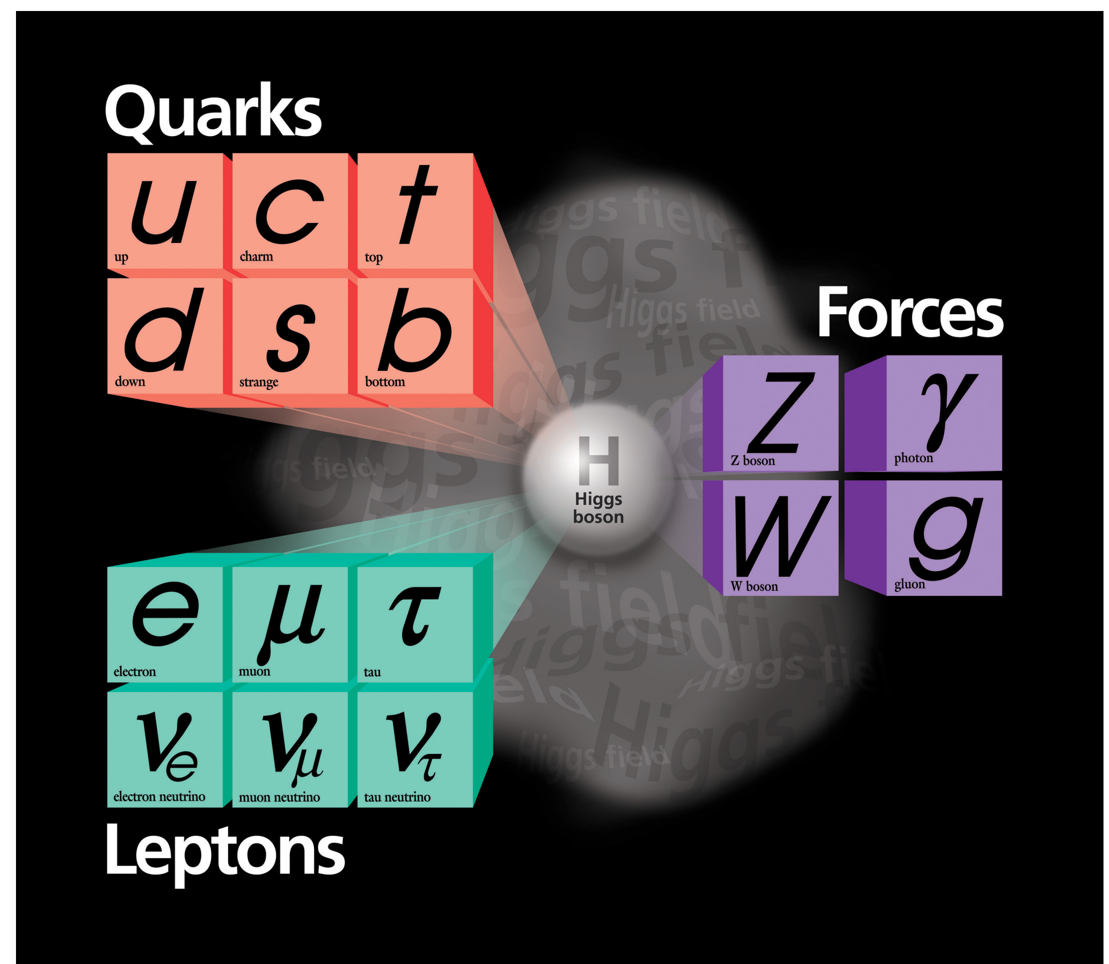
- Huge gap between **weak** and **Plank scales?**
- **Compositeness**? Non-minimal Higgs sector?
- Coupling to **Dark Matter**? Role in cosmological phase transitions?
- Is the **vacuum state of the Universe** stable?

Quarks and leptons

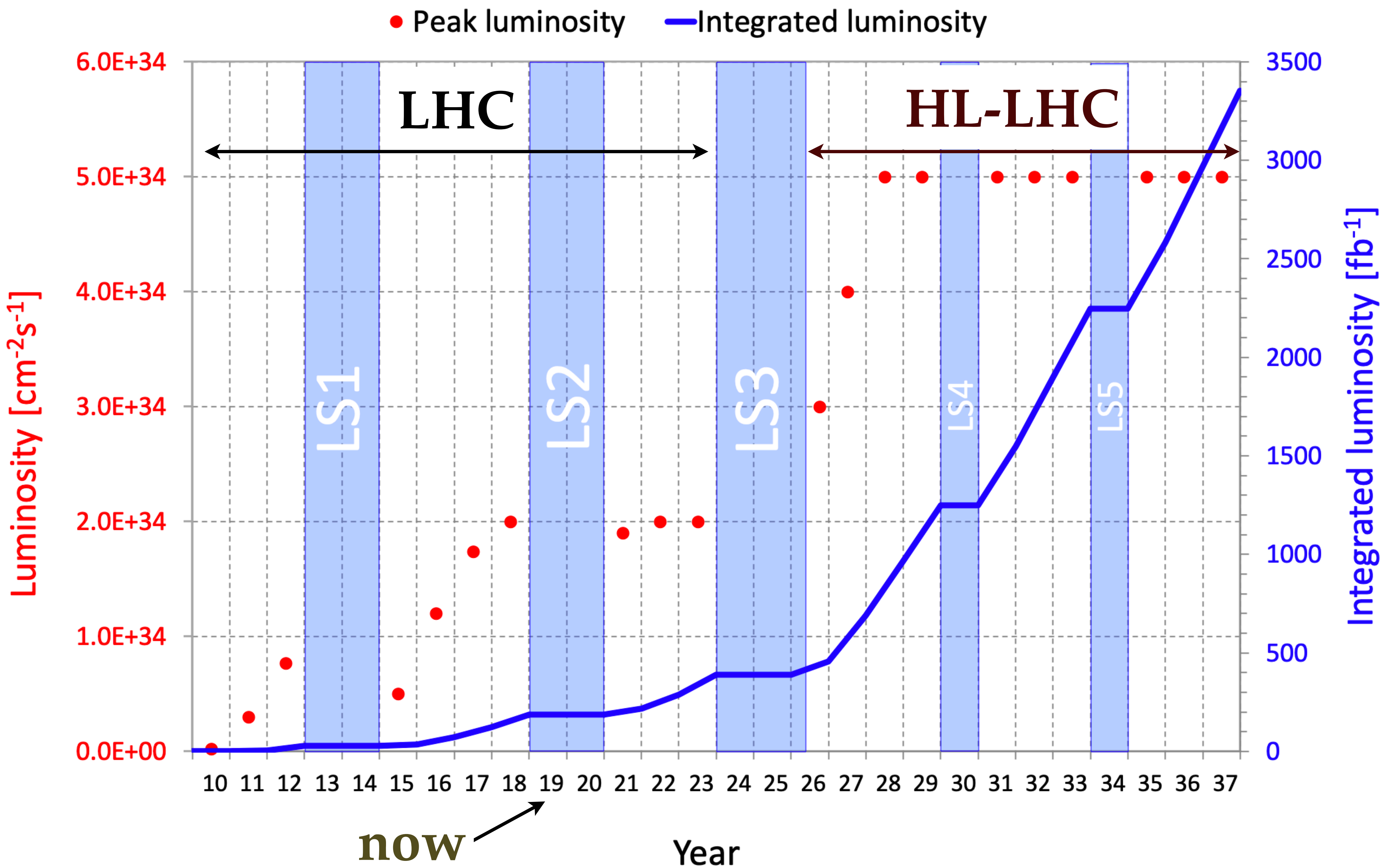
- Why **3 families?** Origin of **masses, mixings**?
- Origin of **Matter-Antimatter asymmetry?**
- Lepton Flavour **Universality?**
- Origin of **neutrino masses?** Are neutrinos Majorana or Dirac?

Dark matter

- **Weakly interacting massive particles**? Neutrinos? Ultralight particles (axions)?
- **Interactions** with SM particles? Self-interactions?
- **Structure** of the Dark Sector?



Open questions in particle physics

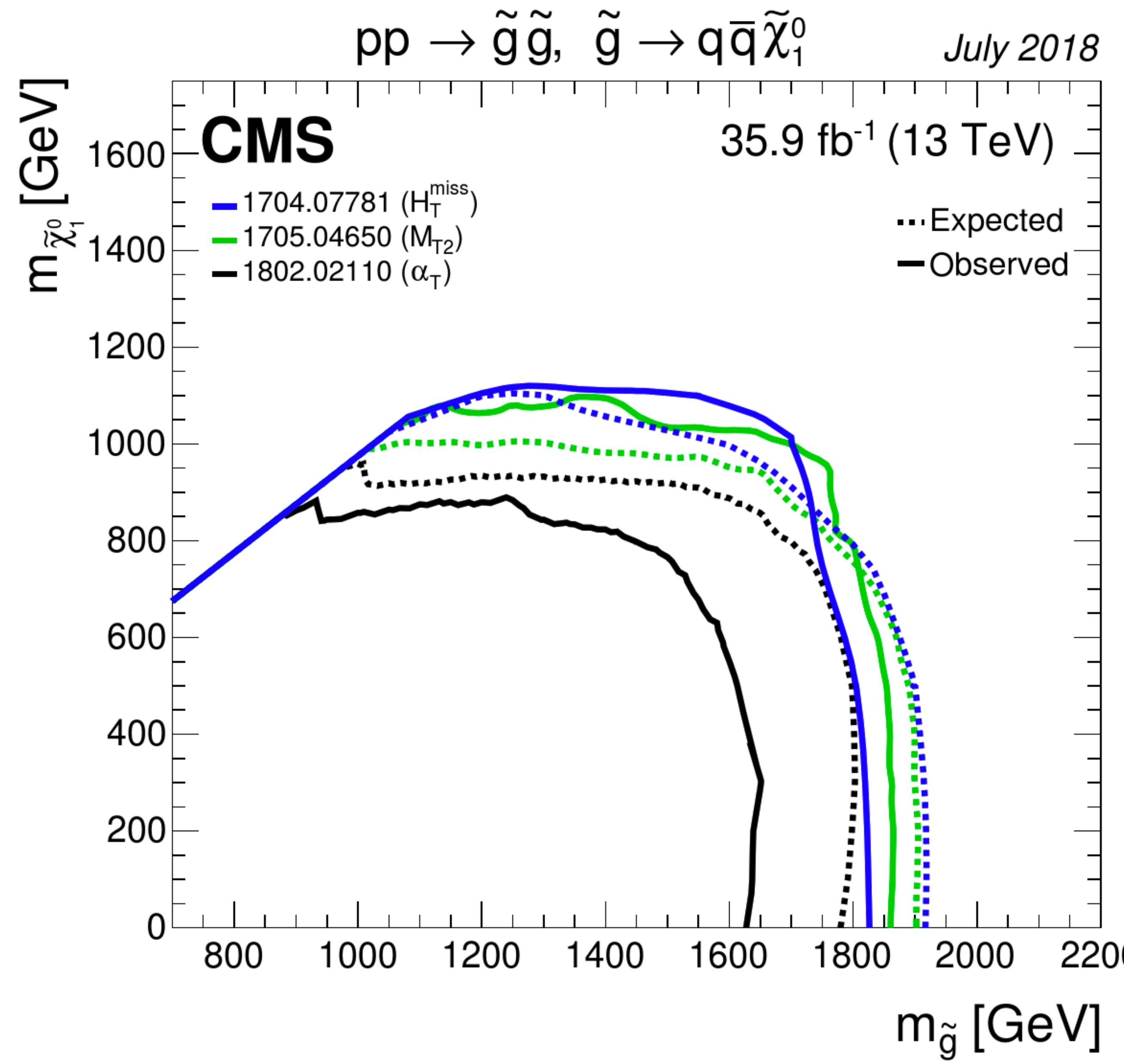


Crucial information on these fundamental questions will be provided by the LHC: the **exploration of the high-energy frontier** has just started!

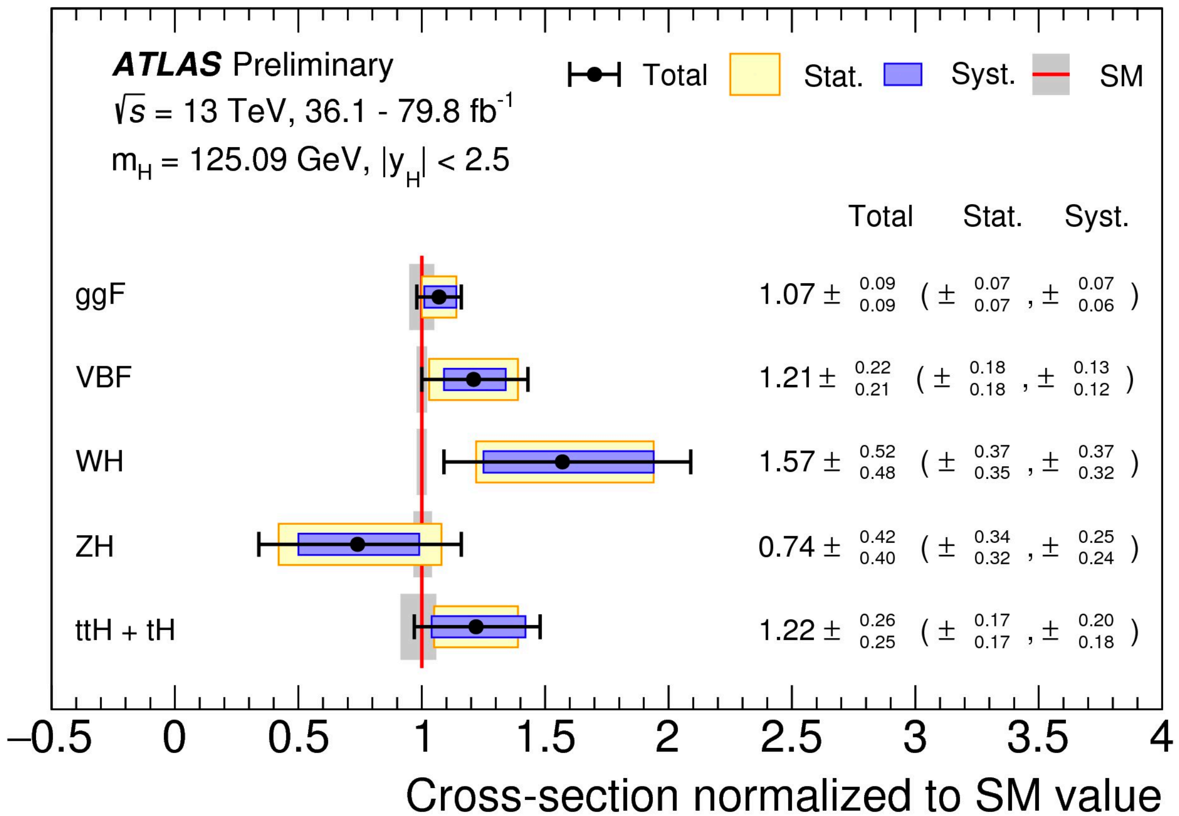
The quest for New Physics at the LHC

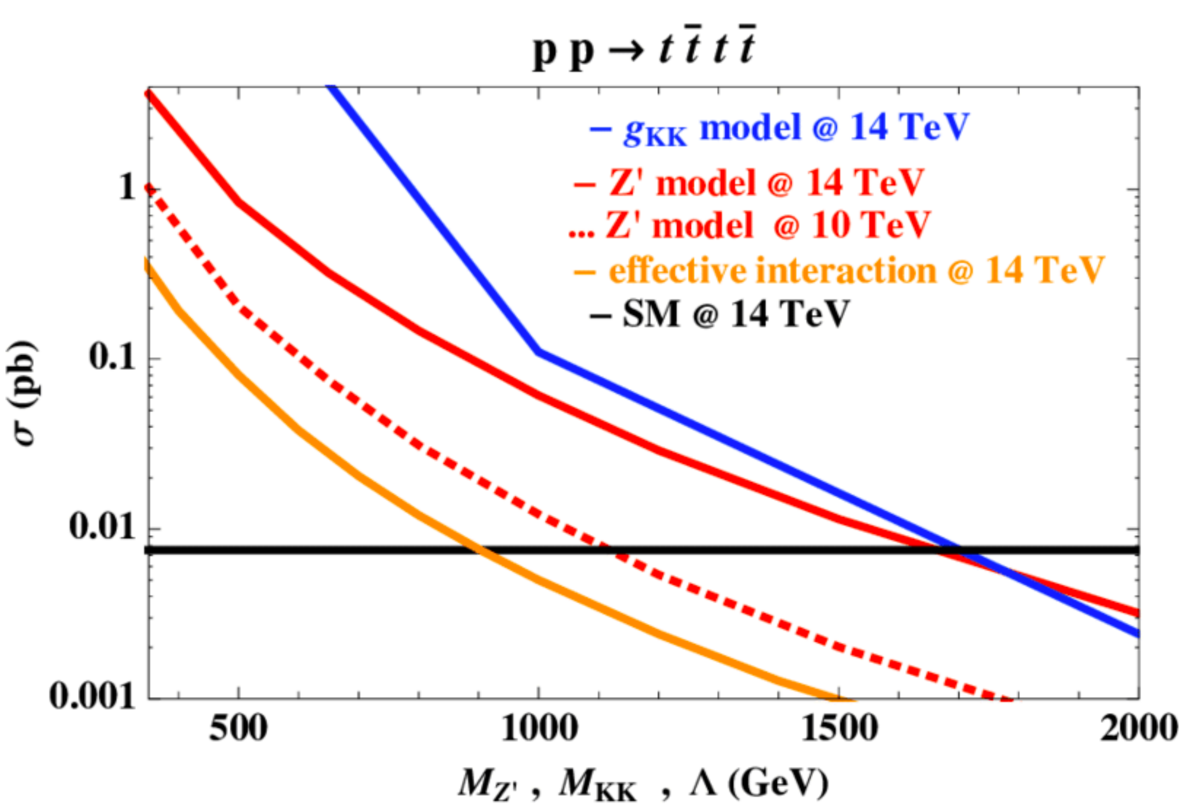
Model-Dependent Searches

- Map parameter space of **specific theories**, or specific realisations of theories (SUSY, Higgs compositeness, …)
- **Reinterpretation/recasting** challenging, since requires Monte Carlo showering, detector simulation, …
- Ad-hoc **restrictions of the BSM parameter space** to facilitate interpretation
- Sensitive to **O(1) deviations**



The quest for New Physics at the LHC





Model-Independent Searches

- ``**SM'' measurements** to constrain BSM
- Allows the use of **highest possible precision** in theory calculations
- Interpreted in **multiple BSM frameworks** (including those not thought of yet!)
- In the long-term, measurements have the **largest impact** in the HEP community
- Sensitive to **O(0.1) or O(0.01) deviations**

Towards a global SMEFT analysis

The Standard Model EFT

Systematic parametrisation of the **theory space** in vicinity of Standard Model

$$\mathcal{L}_{\mathrm{SMEFT}} = \mathcal{L}_{\mathrm{SM}} + \sum_i^{N_{d6}} \frac{c_i}{\Lambda^2} \mathcal{O}_i^{(6)} + \sum_j^{N_{d8}} \frac{b_j}{\Lambda^4} \mathcal{O}_j^{(8)} + \ldots$$

- ☑ SMEFT: **low-energy limit** of generic UV-complete theories at high energies
- ☑ Assumes **SM field content and symmetries** (except the accidental ones)
- ☑ **Complete basis** at any given mass-dimension
- ☑ **Fully renormalizable**, full-fledged QFT: can compute higher orders in QCD and EW
- ☑ Can be matched to **any BSM model** that reduces to the SM at low energies

The Standard Model EFT

Systematic parametrisation of the **theory space** in vicinity of Standard Model

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \sum_{i}^{N_{d6}} \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_{j}^{N_{d8}} \frac{b_j}{\Lambda^4}\mathcal{O}_j^{(8)} + \ldots$$

- Several operators induce **growth with the partonic centre-of-mass energy**: increased sensitivity in LHC cross-sections in the TeV region

$$\sigma(E) = \sigma_{\rm SM}(E)\left(1 + \sum_{i}^{N_{d6}} \omega_i \frac{c_i\, m_{\rm SM}^2}{\Lambda^2} + \sum_{i}^{N_{d6}} \widetilde{\omega}_i \frac{c_i\, E^2}{\Lambda^2} + \mathcal{O}\left(\Lambda^{-4}\right)\right)$$

*enhanced sensitivity from **TeV-scale processes:***

unique feature of LHC

The Standard Model EFT

Systematic parametrisation of the **theory space** in vicinity of Standard Model

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \sum_i^{N_{d6}} \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_j^{N_{d8}} \frac{b_j}{\Lambda^4}\mathcal{O}_j^{(8)} + \dots$$

- Some operators induce **growth with the partonic centre-of-mass energy**: increased sensitivity in LHC cross-sections in the TeV region

$$\sigma(E) = \sigma_{\rm SM}(E)\left(1 + \sum_i^{N_{d6}} \omega_i \frac{c_i\, m_{\rm SM}^2}{\Lambda^2} + \sum_i^{N_{d6}} \widetilde{\omega}_i \frac{c_i\, E^2}{\Lambda^2} + \mathcal{O}\left(\Lambda^{-4}\right)\right)$$

- The number of SMEFT operators is large: **59 non-redundant operators at dimension 6** for one fermion generation, **2499 operators** without any flavour assumption

- A global SMEFT analysis needs to explore a **huge complicated parameter space**

Recipe for a global SMEFT analysis

Theory

(N)NLO QCD + NLO EW for SM xsecs

NLO QCD for SMEFT contributions

State-of-the-art **Parton Distributions**

Data

Higgs and **gauge boson** production

Top quark and **jet** production

Precision **LEP**, **low energy**, **flavour, ….**

Global SMEFT fit

Bounds can be compared with

specific UV completions

New data incorporated without redoing fit

Delivery

Efficient exploration of **parameter space**

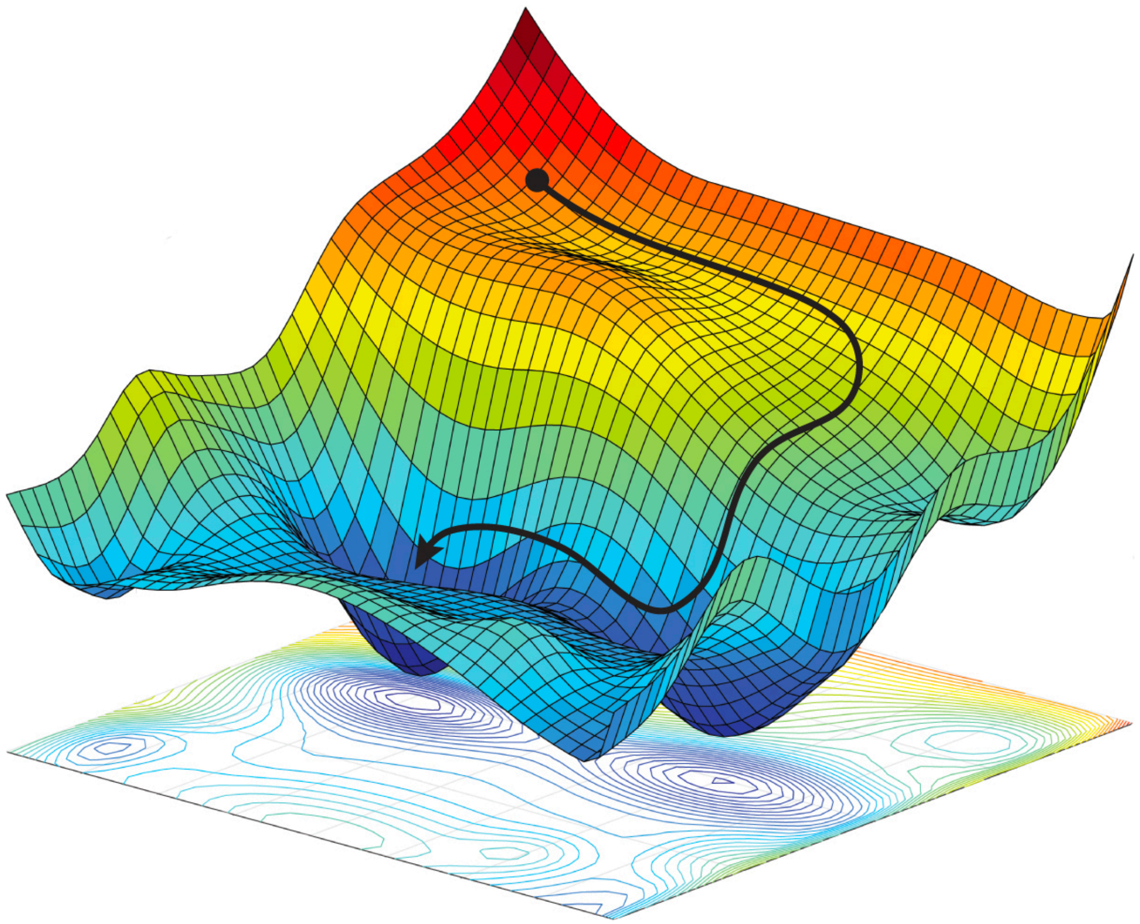
Faithful **uncertainty estimate** (exp & th)

Avoiding under- and over-fitting

Methodology

The optimisation conundrum

(Stochastic) Gradient Descent



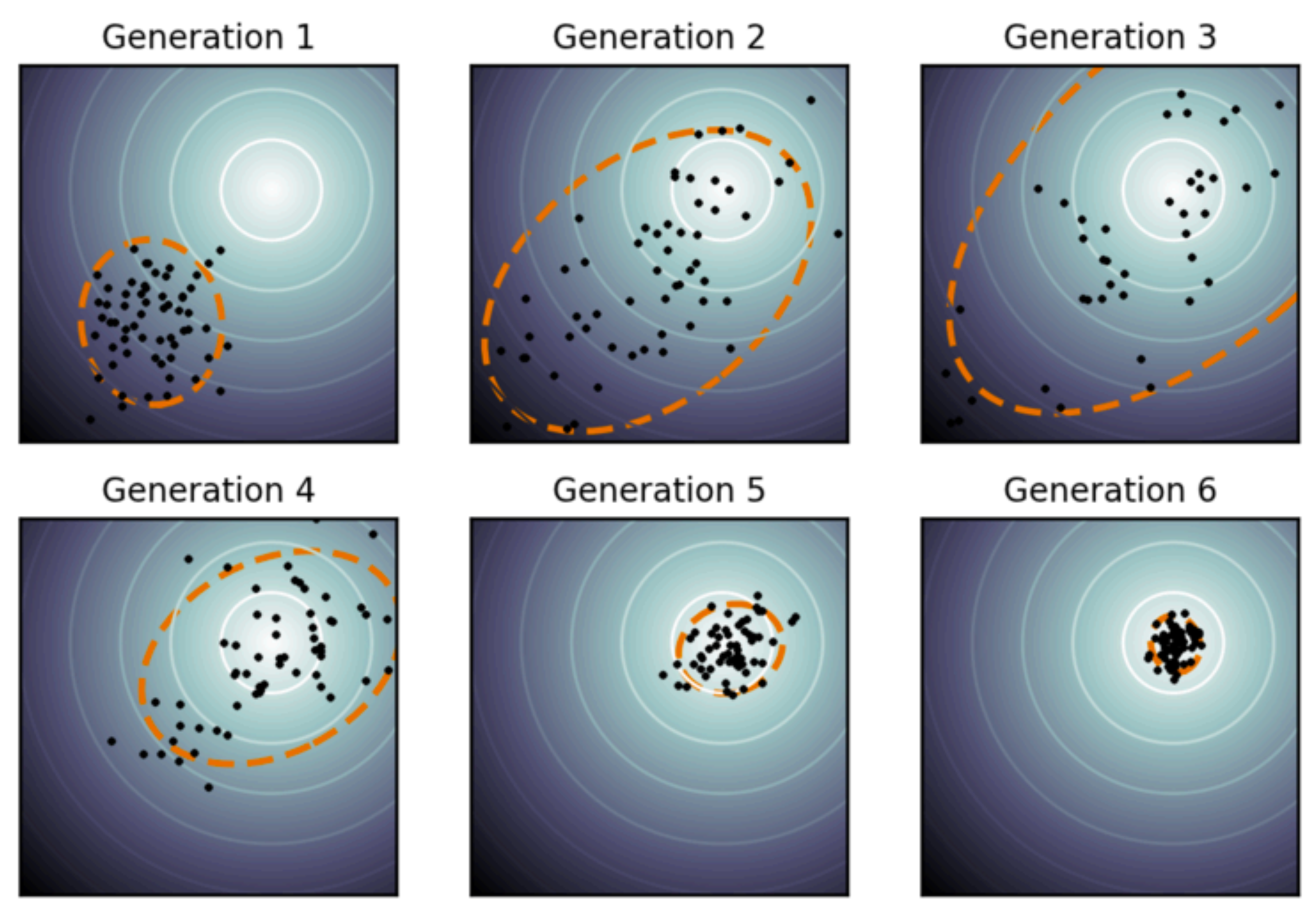
- **Deterministic algorithms**: follow the gradient of the cost function

- **Evolutionary algorithms**: act on population of solutions with random mutations and selection

- A challenge for any SMEFT global analysis is the efficient exploration of the **huge parameter space**

- Several pitfalls to be avoided: under-fitting, over-fitting, **local minima**, saddle points, ….

Genetic Algorithms



The SMEFiT framework

N. P. Hartland, F. Maltoni, E. R. Nocera, J. Rojo,
*E. Slade, E. Vryonidou, C. Zhang, **arXiv:1901.05965** (JHEP)*

The SMEFiT method

- Generate a large sample of **Monte Carlo replicas** to construct the **probability distribution** in the space of experimental data

$$\mathcal{O}_i^{(\mathrm{art})(k)} = S_{i,N}^{(k)} \mathcal{O}_i^{(\mathrm{exp})} \left(1 + \sum_{\alpha=1}^{N_{\mathrm{sys}}} r_{i,\alpha}^{(k)} \sigma_{i,c}^{(\mathrm{sys})} + r_i^{(k)} \sigma_i^{(\mathrm{stat})} \right) , \quad k = 1, \ldots, N_{\mathrm{rep}}$$

cross-section for k-th replica

central value (data)

correlated systematic uncertainties

statistical uncertainties

 # MC replicas

The SMEFiT method

- Generate a large sample of **Monte Carlo replicas** to construct the **probability distribution** in the space of experimental data

$$\mathcal{O}_i^{(\mathrm{art})(k)} = S_{i,N}^{(k)} \mathcal{O}_i^{(\mathrm{exp})} \left(1 + \sum_{\alpha=1}^{N_{\mathrm{sys}}} r_{i,\alpha}^{(k)} \sigma_{i,c}^{(\mathrm{sys})} + r_i^{(k)} \sigma_i^{(\mathrm{stat})} \right) \, , \quad k = 1, \ldots, N_{\mathrm{rep}}$$

- Construct theory calculations where the SM is **extended by SMEFT corrections**

to be determined from the data

$$\sigma_i^{\mathrm{th}}\left(\{c_n\}\right) = \sigma_{\mathrm{SM},i} + \sum_{n=1}^{N_{\mathrm{op}}} \widetilde{\sigma}_{i,n} \frac{c_n}{\Lambda^2} + \sum_{n,m=1}^{N_{\mathrm{op}}} \widetilde{\sigma}_{i,nm} \frac{c_n c_m}{\Lambda^4} \, , \quad i = 1 \, \ldots, N_{\mathrm{dat}}$$

SM: compute at (N)NLO QCD

SMEFT: compute at (N)LO QCD with aMC@NLO

The SMEFiT method

- Generate a large sample of **Monte Carlo replicas** to construct the **probability distribution** in the space of experimental data

$$\mathcal{O}_i^{(\mathrm{art})(k)} = S_{i,N}^{(k)} \mathcal{O}_i^{(\mathrm{exp})} \left(1 + \sum_{\alpha=1}^{N_{\mathrm{sys}}} r_{i,\alpha}^{(k)} \sigma_{i,c}^{(\mathrm{sys})} + r_i^{(k)} \sigma_i^{(\mathrm{stat})} \right) , \quad k = 1, \ldots, N_{\mathrm{rep}}$$

- Construct theory calculations where the SM is **extended by SMEFT corrections**

$$\mathcal{O}_i^{\mathrm{th}} \left(\{ c_n \} \right) = \sigma_{\mathrm{SM},i} + \sum_{n=1}^{N_{\mathrm{op}}} \widetilde{\sigma}_{i,n} \frac{c_n}{\Lambda^2} + \sum_{n,m=1}^{N_{\mathrm{op}}} \widetilde{\sigma}_{i,nm} \frac{c_n \, c_m}{\Lambda^4} , \quad i = 1 \ \ldots, N_{\mathrm{dat}}$$

- Determine the SMEFT coefficients **replica-by-replica** by minimising a cost function

$$E(\{ c_l^{(k)} \}) \equiv \frac{1}{N_{\mathrm{dat}}} \sum_{i,j=1}^{N_{\mathrm{dat}}} \left(\mathcal{O}_i^{(\mathrm{th})} \left(\{ c_n^{(k)} \} \right) - \mathcal{O}_i^{(\mathrm{art})(k)} \right) (\mathrm{cov}^{-1})_{ij} \left(\mathcal{O}_j^{(\mathrm{th})} \left(\{ c_n^{(k)} \} \right) - \mathcal{O}_j^{(\mathrm{art})(k)} \right)$$

The SMEFiT method

- Determine the SMEFT coefficients **replica-by-replica** by minimising a cost function

$$E(\{c_l^{(k)}\}) \equiv \frac{1}{N_{\rm dat}} \sum_{i,j=1}^{N_{\rm dat}} \left(\mathcal{O}_i^{\rm (th)} \left(\{c_n^{(k)}\} \right) - \mathcal{O}_i^{{\rm (art)}(k)} \right) ({\rm cov}^{-1})_{ij} \left(\mathcal{O}_j^{\rm (th)} \left(\{c_n^{(k)}\} \right) - \mathcal{O}_j^{{\rm (art)}(k)} \right)$$

- The covariance matrix includes **all sources of experimental errors** + some theory errors

t_0 prescription

$$({\rm cov}_{t_0})_{ij}^{\rm (exp)} \equiv \left(\sigma_i^{\rm (stat)} \right)^2 \delta_{ij} + \left(\sum_{\alpha=1}^{N_{\rm sys}} \sigma_{i,\alpha}^{\rm (sys)} \sigma_{j,\alpha}^{\rm (sys)} \mathcal{O}_i^{\rm (exp)} \mathcal{O}_j^{\rm (exp)} + \sum_{\beta=1}^{N_{\rm norm}} \sigma_{i,\beta}^{\rm (norm)} \sigma_{j,\beta}^{\rm (norm)} \mathcal{O}_i^{\rm (th,0)} \mathcal{O}_j^{\rm (th,0)} \right)$$

$${\rm cov}_{ij} = {\rm cov}_{ij}^{\rm (exp)} + {\rm cov}_{ij}^{\rm (th)}$$

th uncertainties: PDFs

can be extended to MHOUs

$${\rm cov}_{ij}^{\rm (th)} = \left\langle \mathcal{O}_i^{\rm (th)(r)} \mathcal{O}_j^{\rm (th)(r)} \right\rangle_{\rm rep} - \left\langle \mathcal{O}_i^{\rm (th)(r)} \right\rangle_{\rm rep} \left\langle \mathcal{O}_j^{\rm (th)(r)} \right\rangle_{\rm rep} ,$$

The SMEFiT method

- Determine the SMEFT coefficients **replica-by-replica** by minimising a cost function

$$E(\{c_l^{(k)}\}) \equiv \frac{1}{N_{\rm dat}} \sum_{i,j=1}^{N_{\rm dat}} \left(\mathcal{O}_i^{\rm (th)}\left(\{c_n^{(k)}\}\right) - \mathcal{O}_i^{{\rm (art)}(k)} \right) ({\rm cov}^{-1})_{ij} \left(\mathcal{O}_j^{\rm (th)}\left(\{c_n^{(k)}\}\right) - \mathcal{O}_j^{{\rm (art)}(k)} \right)$$

- The covariance matrix includes **all sources of experimental errors** + some theory errors

$${\rm cov}_{ij} = {\rm cov}_{ij}^{\rm (exp)} + {\rm cov}_{ij}^{\rm (th)}$$

- The ensemble of coefficients $\{c_l^{(k)}\}$ then provides a sampling of the **probability density** in the **SMEFT parameter space**

$$\left\langle c_l \right\rangle \equiv \frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} c_l^{(k)} \qquad \rho\left(c_i, c_j\right) = \frac{\frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} c_i^{(k)} c_j^{(k)} - \left\langle c_i \right\rangle \left\langle c_j \right\rangle}{\delta c_i \delta c_j}\,.$$

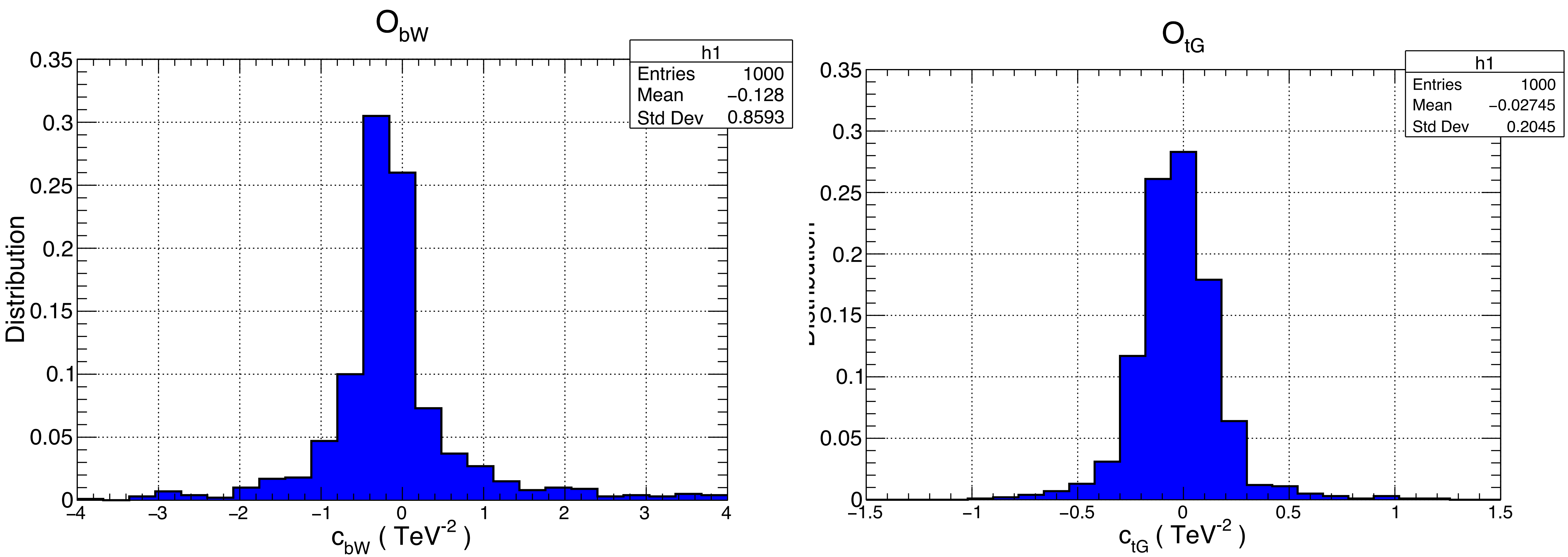
Sampling the SMEFT probability distribution

- The output of SMEFiT is a sampling of the **probability distribution** in the SMEFT space

$$\left\{ c_n^{(k)} \right\}, \quad n = 1 \ldots, N_{\rm op}, \quad k = 1 \ldots, N_{\rm rep}$$

- Used to **evaluate statistical estimators** such as variances, correlations, higher moments, …

- Distributions are **reasonably Gaussian** for well-constrained degrees of freedom ….



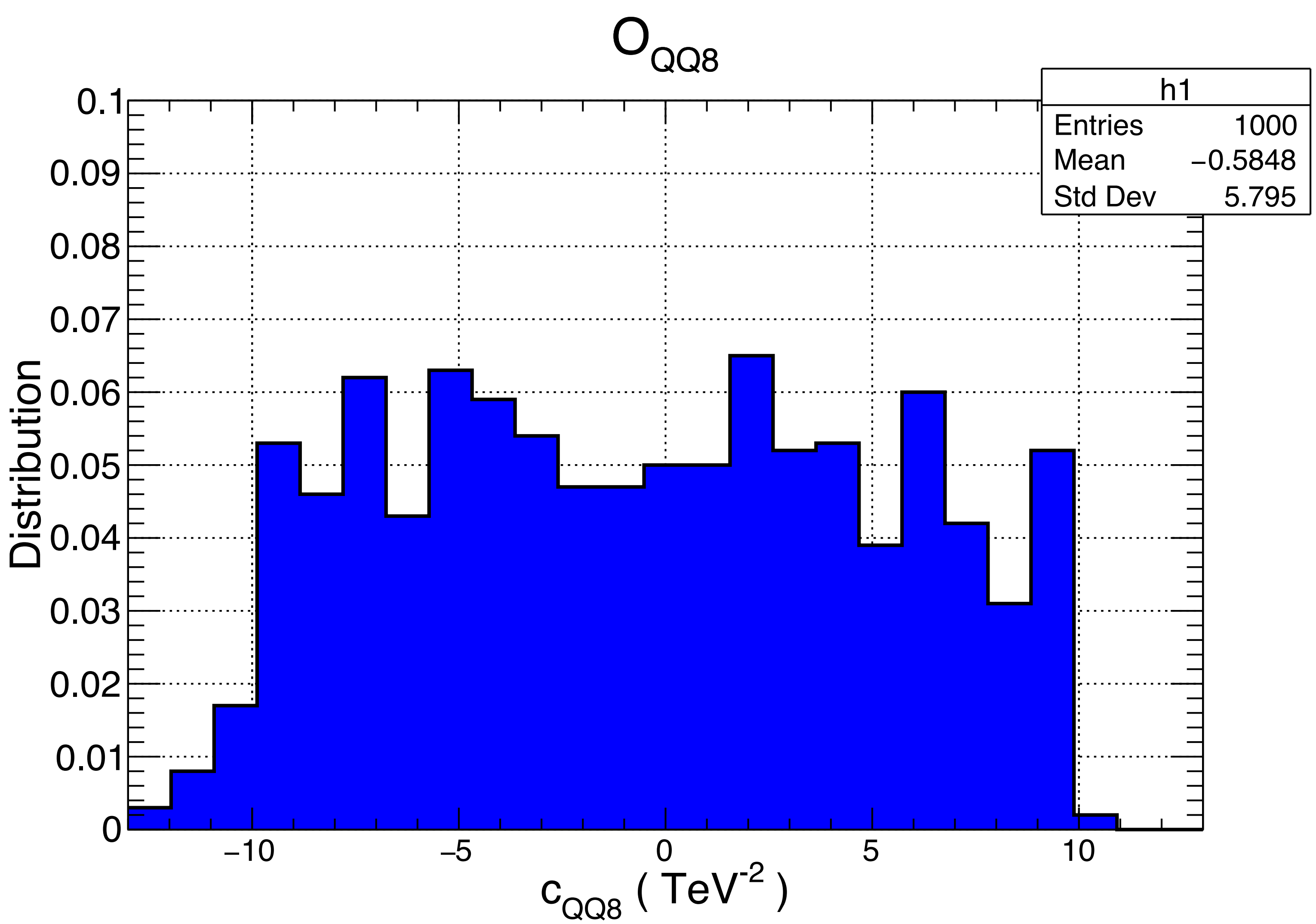
Sampling the SMEFT probability distribution

- The output of SMEFiT is a sampling of the **probability distribution** in the SMEFT space

$$\left\{ c_n^{(k)} \right\}, \quad n = 1 \ldots, N_{\rm op}\,, \quad k = 1 \ldots, N_{\rm rep}$$

- Used to **evaluate statistical estimators** such as variances, correlations, higher moments, …

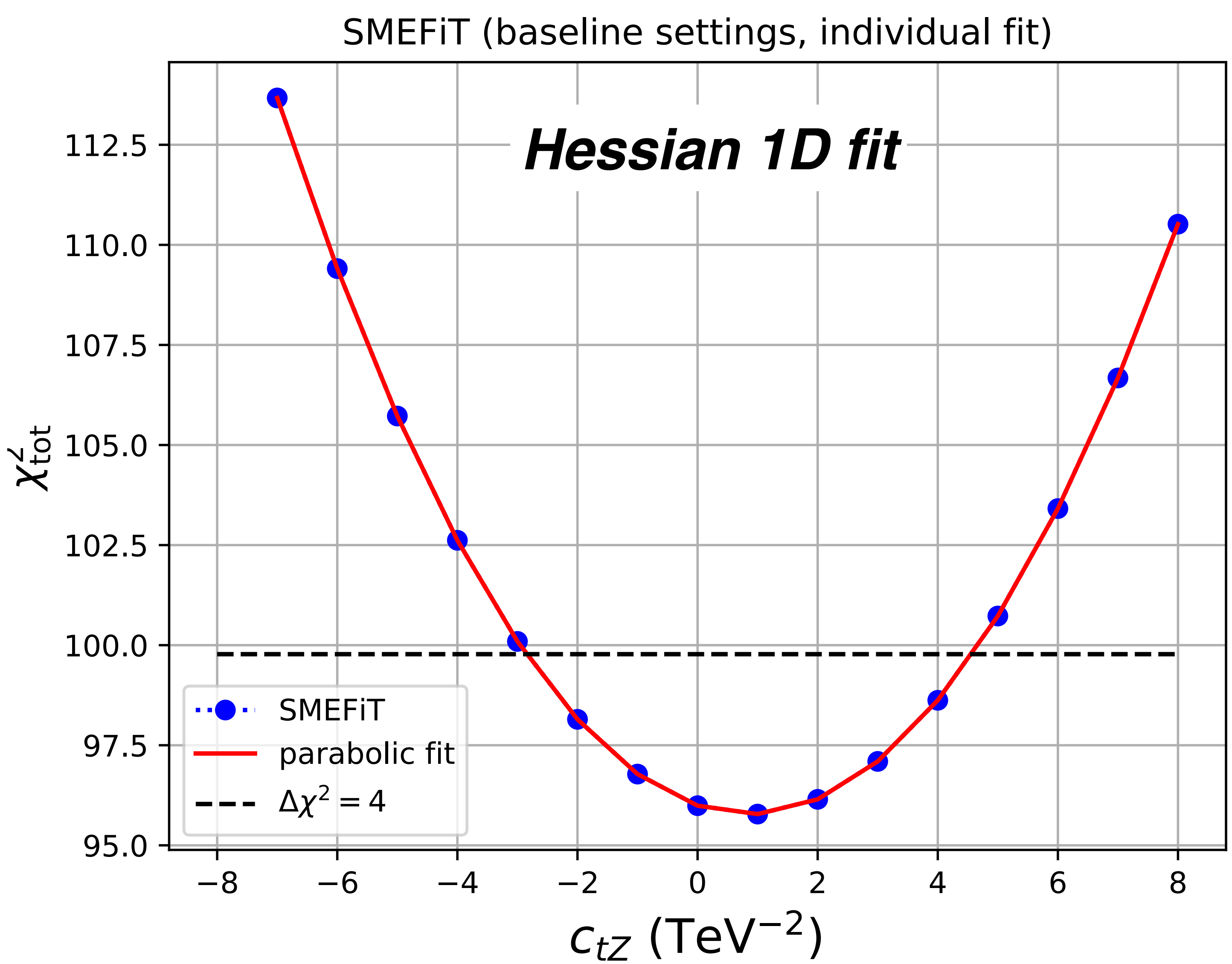
- …. but much less so for **under-constrained** or **redundant** operators



The SMEFiT method

- Uncertainties on the SMEFT degrees of freedom evaluated from **variance of MC sample**

$$\left(\delta c_n\right)^2 = \frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} \left(c_n^{(k)}\right)^2 - \left\langle c_n \right\rangle^2$$



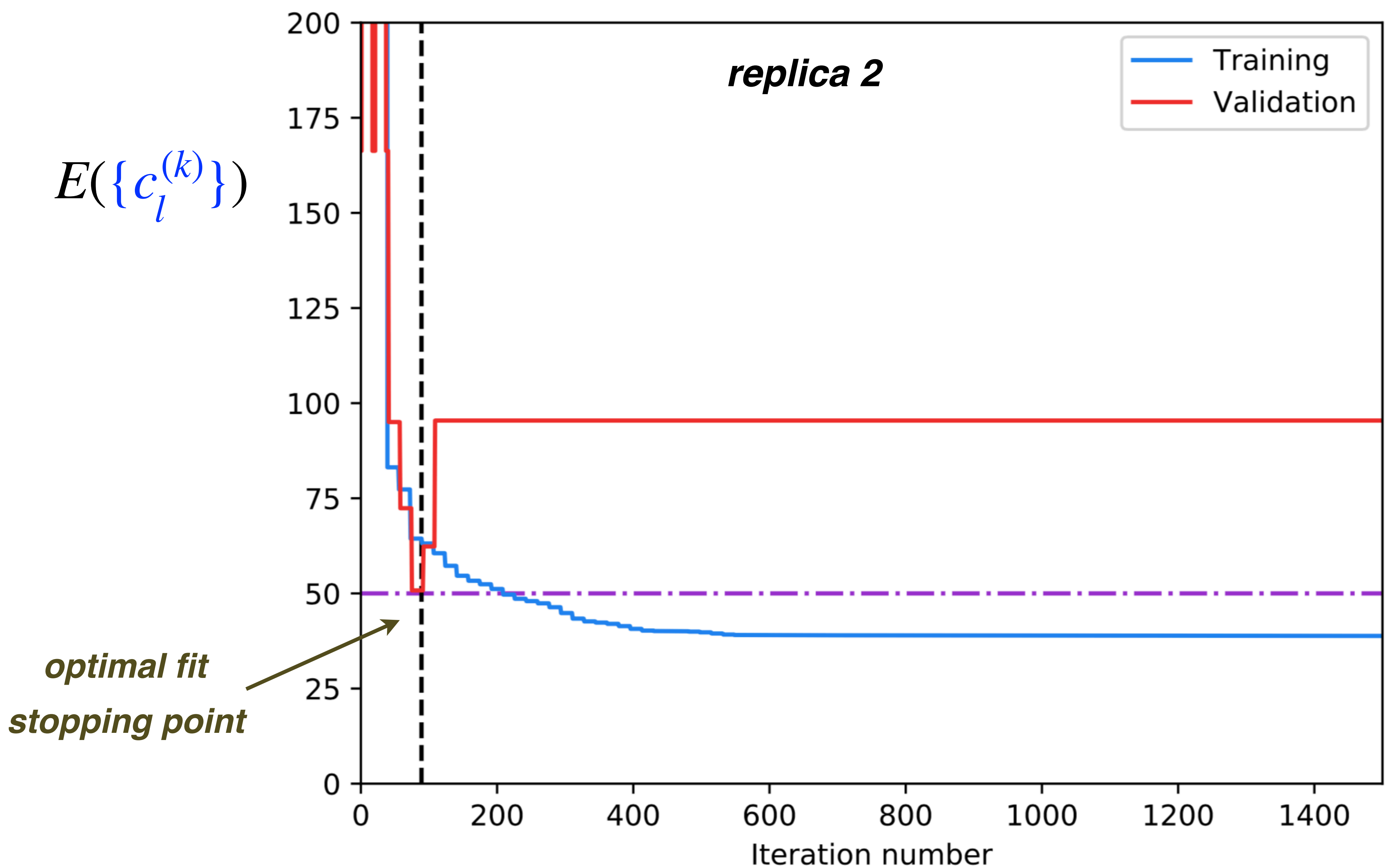
- For single-parameter fits, **Monte Carlo results** benchmarked with **Hessian method**, finding good agreement

- The Hessian method **numerically less stable** as dimensionality of parameter space increases

The SMEFiT method

- Since in general there will be **unconstrained/degenerate directions** in the parameter space, it is crucial to avoid overfitting (that is, fitting statistical fluctuations)

- Achieved by the **cross-validation look-back validation stopping** method



SMEFiT code structure

Stand-alone **Python code**, which exploits functionalities of the **NNPDF framework**

NNPDF code

- **Experimental data** and covariance matrices
- **NLO** APPLgrids + **NNLO** C-factors (for processes used in PDF fit)

aMC@NLO

- NLO **QCD** (benchmark)
- LO, NLO **SMEFT**
- Both **O(Λ^{-2})** and **O(Λ^{-4})** from ***d=6*** **operators**

MCFM

- NLO **QCD** (consistent choice of PDFs)
- Cross-checks of **aMC@NLO**

Python analysis code

- Assemble **theory predictions** for generic SMEFT Wilson coefficients
- **Optimisation** with Sequential Quadratic Programming (**SciPy**)
- Look-back **cross-validation stopping**
- **Monte Carlo replicas** for uncertainty propagation

The Top Quark Case

Operator basis

- We follow the same flavour assumptions as in the **LHC Top WG note**
- Minimal Flavour Violation (MFV), diagonal CKM, zero Yukawas for first two quark gens, CP conservation assumed
- Include those SMEFT dimension-6 operators of Warsaw basis with **at least one top quark**
- The fit includes a total of **34 independent degrees of freedom**
- Include both **interference** and **quadratic contributions** from these operators

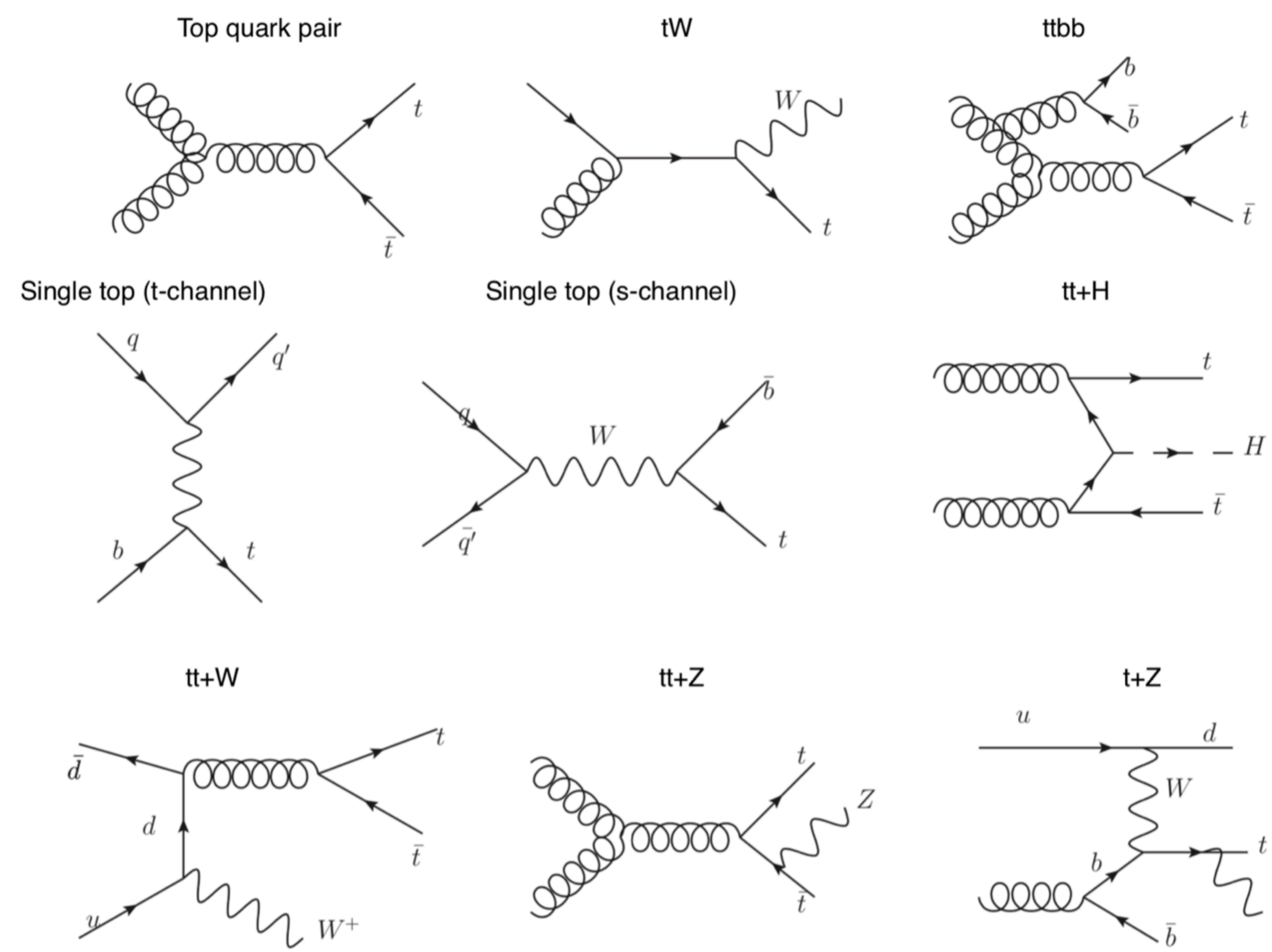
Class	Notation	Degree of Freedom	Operator Definition
$QQQQ$ **4-heavy**	OQQ1	c^1_{QQ}	$2C_{qq}^{1(3333)} - \frac{2}{3}C_{qq}^{3(3333)}$
	OQQ8	c^8_{QQ}	$8C_{qq}^{3(3333)}$
	OQt1	c^1_{Qt}	$C_{qu}^{1(3333)}$
	OQt8	c^8_{Qt}	$C_{qu}^{8(3333)}$
	OQb1	c^1_{Qb}	$C_{qd}^{1(3333)}$
	OQb8	c^8_{Qb}	$C_{qd}^{8(3333)}$
	Ott1	c^1_{tt}	$C_{uu}^{(3333)}$
	Otb1	c^1_{tb}	$C_{ud}^{1(3333)}$
	Otb8	c^8_{tb}	$C_{ud}^{8(3333)}$
	OQtQb1	c^1_{QtQb}	$C_{quqd}^{1(3333)}$
	OQtQb8	c^8_{QtQb}	$C_{quqd}^{8(3333)}$
$QQqq$ **2-heavy-2-light**	O81qq	$c^{1,8}_{Qq}$	$C_{qq}^{1(i33i)} + 3C_{qq}^{3(i33i)}$
	O11qq	$c^{1,1}_{Qq}$	$C_{qq}^{1(ii33)} + \frac{1}{6}C_{qq}^{1(i33i)} + \frac{1}{2}C_{qq}^{3(i33i)}$
	O83qq	$c^{3,8}_{Qq}$	$C_{qq}^{1(i33i)} - C_{qq}^{3(i33i)}$
	O13qq	$c^{3,1}_{Qq}$	$C_{qq}^{3(ii33)} + \frac{1}{6}(C_{qq}^{1(i33i)} - C_{qq}^{3(i33i)})$
	O8qt	c^8_{tq}	$C_{qu}^{8(ii33)}$
	O1qt	c^1_{tq}	$C_{qu}^{1(ii33)}$
	O8ut	c^8_{tu}	$2C_{uu}^{(i33i)}$
	O1ut	c^1_{tu}	$C_{uu}^{(ii33)} + \frac{1}{3}C_{uu}^{(i33i)}$
	O8qu	c^8_{Qu}	$C_{qu}^{8(33ii)}$
	O1qu	c^1_{Qu}	$C_{qu}^{1(33ii)}$
	O8dt	c^8_{td}	$C_{ud}^{8(33ii)}$
	O1dt	c^1_{td}	$C_{ud}^{1(33ii)}$
	O8qd	c^8_{Qd}	$C_{qd}^{8(33ii)}$
	O1qd	c^1_{Qd}	$C_{qd}^{1(33ii)}$
$QQ+V,G,\varphi$ **2-heavy + V/h**	OtG	c_{tG}	$\mathrm{Re}\{C_{uG}^{(33)}\}$
	OtW	c_{tW}	$\mathrm{Re}\{C_{uW}^{(33)}\}$
	ObW	c_{bW}	$\mathrm{Re}\{C_{dW}^{(33)}\}$
	OtZ	c_{tZ}	$\mathrm{Re}\{-s_W C_{uB}^{(33)} + c_W C_{uW}^{(33)}\}$
	Off	$c_{\varphi tb}$	$\mathrm{Re}\{C_{\varphi ud}^{(33)}\}$
	Ofq3	$c^3_{\varphi Q}$	$C_{\varphi q}^{3(33)}$
	OpQM	$c^-_{\varphi Q}$	$C_{\varphi q}^{1(33)} - C_{\varphi q}^{3(33)}$
	Opt	$c_{\varphi t}$	$C_{\varphi u}^{(33)}$
	Otp	$c_{t\varphi}$	$\mathrm{Re}\{C_{u\varphi}^{(33)}\}$

The top quark sector of the SMEFT

Notation	Sensitivity at $\mathcal{O}(\Lambda^{-2})$ ($\mathcal{O}(\Lambda^{-4})$)								
	$t\bar{t}$	single-top	tW	tZ	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}t\bar{t}$	$t\bar{t}b\bar{b}$
OQQ1								✓	✓
OQQ8								✓	✓
OQt1								✓	✓
OQt8								✓	✓
OQb1								(✓)	✓
OQb8								(✓)	✓
Ott1								✓	✓
Otb1								(✓)	✓
Otb8								✓	✓
OQtQb1									
OQtQb8									
O81qq	✓				✓	✓	✓	✓	✓
O11qq	✓				(✓)	(✓)	(✓)	✓	✓
O83qq	✓	✓		(✓)	✓	✓	✓	✓	✓
O13qq	✓	✓		✓	(✓)	(✓)	(✓)	✓	✓
O8qt	✓				✓	✓	✓	✓	✓
O1qt	✓				(✓)	(✓)	(✓)	✓	✓
O8ut	✓					✓	✓	✓	✓
O1ut	✓					(✓)	(✓)	✓	✓
O8qu	✓					✓	✓	✓	✓
O1qu	✓					(✓)	(✓)	✓	✓
O8dt	✓					✓	✓	✓	✓
O1dt	✓					(✓)	(✓)	✓	✓
O8qd	✓					✓	✓	✓	✓
O1qd	✓					(✓)	(✓)	✓	✓
OtG	✓				✓	✓	✓	✓	✓
OtW		✓	✓	✓					
ObW		(✓)	(✓)						
OtZ				✓		✓			
Off		(✓)	(✓)	(✓)					
Ofq3		✓	✓	✓		✓			
OpQM				✓		✓			
Opt				✓		✓	✓		
Otp							✓		

A large number of different dimension-6 SMEFT operators modify **top production at LHC**

$$\sigma_i^{\rm th}\left(\{c_n\}\right) = \sigma_{{\rm SM},i} + \sum_{n=1}^{N_{\rm op}} \widetilde{\sigma}_{i,n}\frac{c_n}{\Lambda^2} + \sum_{n,m=1}^{N_{\rm op}} \widetilde{\sigma}_{i,nm}\frac{c_n c_m}{\Lambda^4}$$

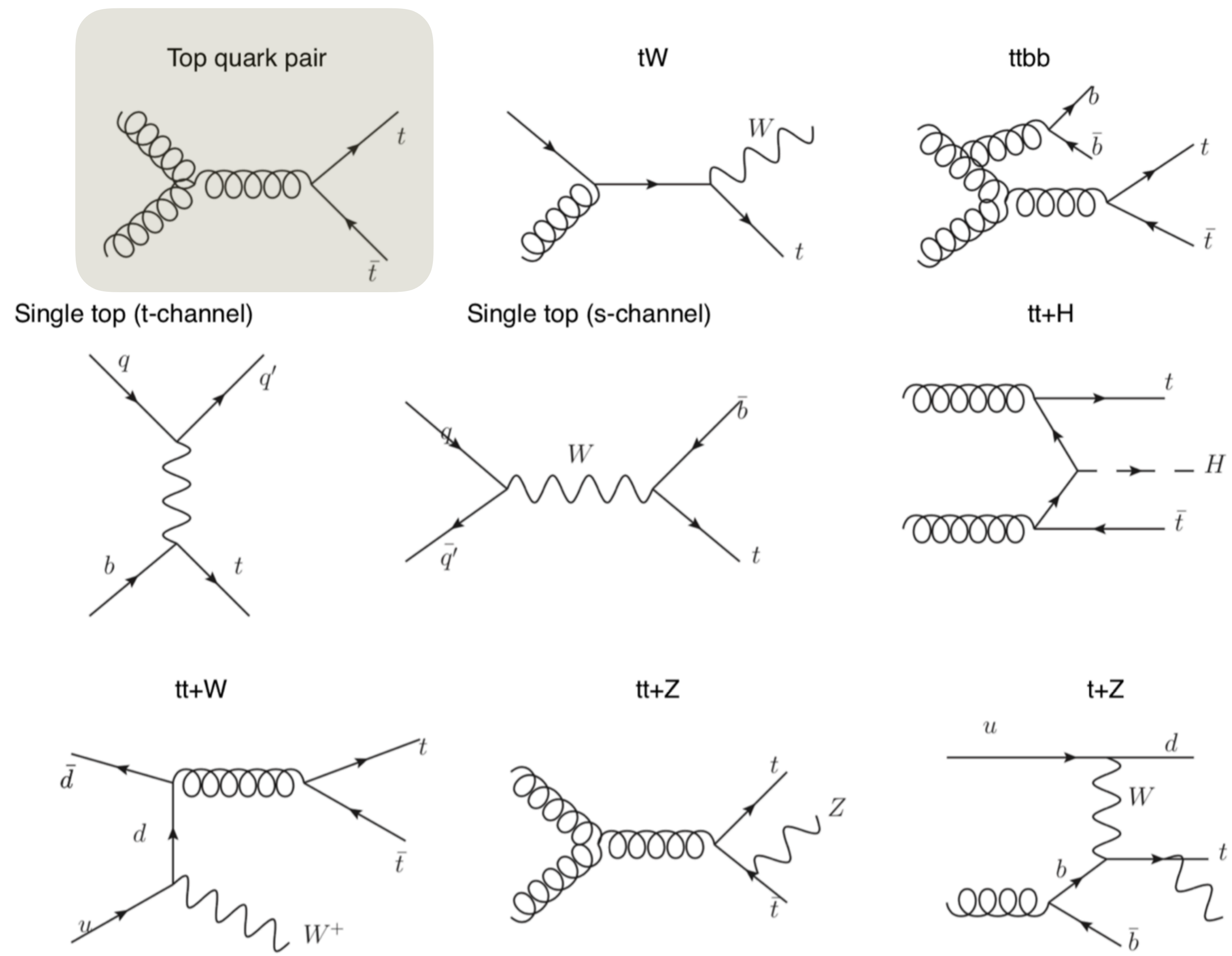


The top quark sector of the SMEFT

Notation	Sensitivity at $\mathcal{O}(\Lambda^{-2})$ ($\mathcal{O}(\Lambda^{-4})$)								
	$t\bar{t}$	single-top	tW	tZ	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}t\bar{t}$	$t\bar{t}b\bar{b}$
OQQ1								✓	✓
OQQ8								✓	✓
OQt1								✓	✓
OQt8								✓	✓
OQb1								(✓)	✓
OQb8								(✓)	✓
Ott1								✓	✓
Otb1								(✓)	✓
Otb8								✓	✓
OQtQb1									
OQtQb8									
O81qq	✓				✓	✓	✓	✓	✓
O11qq	✓				(✓)	(✓)	(✓)	✓	✓
O83qq	✓	✓		(✓)	✓	✓	✓	✓	✓
O13qq	✓	✓		✓	(✓)	(✓)	(✓)	✓	✓
O8qt	✓				✓	✓	✓	✓	✓
O1qt	✓				(✓)	(✓)	(✓)	✓	✓
O8ut	✓					✓	✓	✓	✓
O1ut	✓					(✓)	(✓)	✓	✓
O8qu	✓					✓	✓	✓	✓
O1qu	✓					(✓)	(✓)	✓	✓
O8dt	✓					✓	✓	✓	✓
O1dt	✓					(✓)	(✓)	✓	✓
O8qd	✓					✓	✓	✓	✓
O1qd	✓					(✓)	(✓)	✓	✓
OtG	✓				✓	✓	✓	✓	✓
OtW		✓	✓	✓					
ObW		(✓)	(✓)						
OtZ				✓		✓			
Off		(✓)	(✓)	(✓)					
Ofq3		✓	✓	✓		✓			
OpQM				✓		✓			
Opt				✓		✓	✓		
Otp							✓		

A large number of different dimension-6 SMEFT operators modify **top production at LHC**

$$\sigma_i^{\rm th}\left(\{c_n\}\right) = \sigma_{{\rm SM},i} + \sum_{n=1}^{N_{\rm op}} \widetilde{\sigma}_{i,n}\frac{c_n}{\Lambda^2} + \sum_{n,m=1}^{N_{\rm op}} \widetilde{\sigma}_{i,nm}\frac{c_n c_m}{\Lambda^4}$$

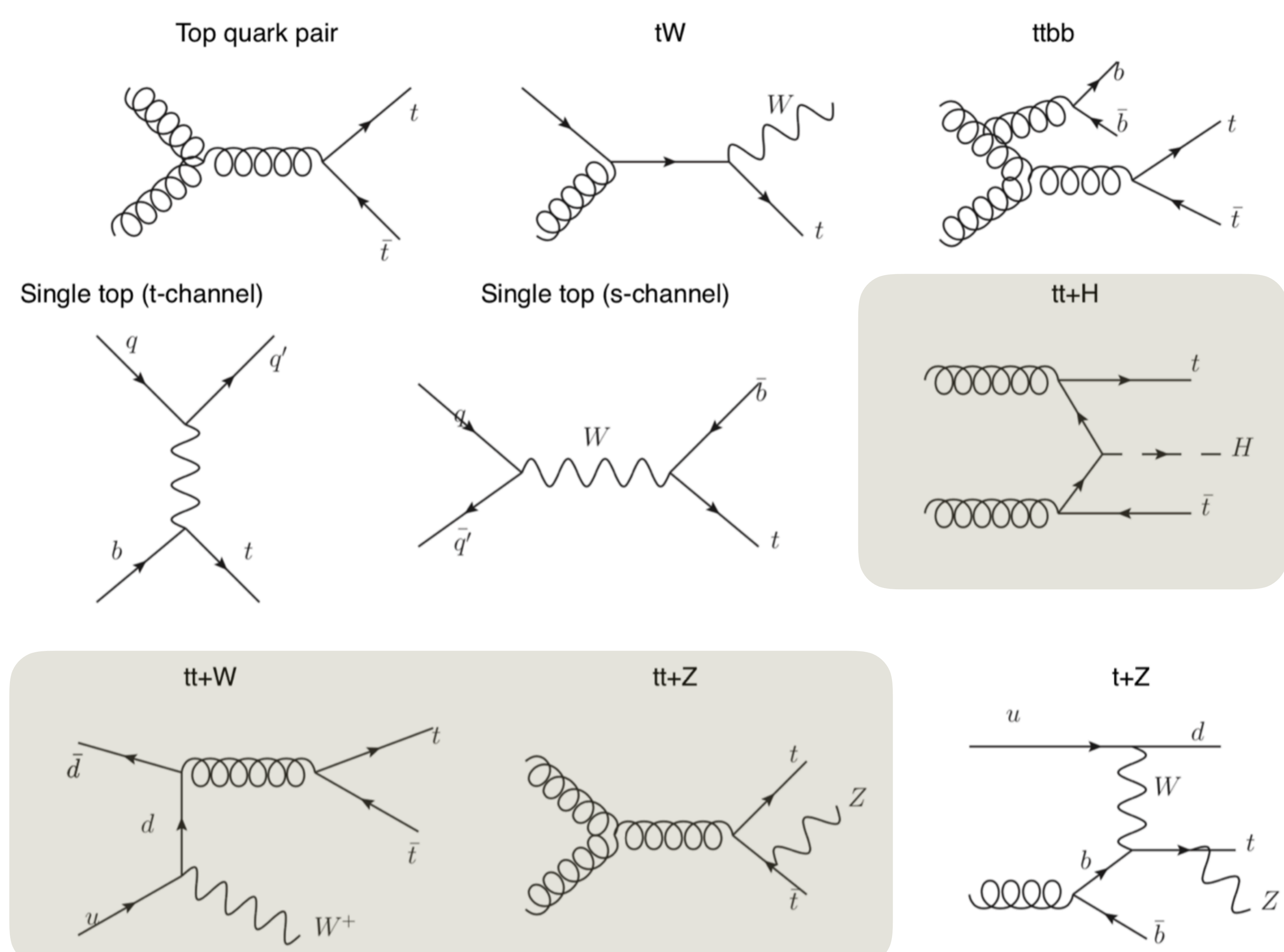


The top quark sector of the SMEFT

Notation	Sensitivity at $\mathcal{O}(\Lambda^{-2})$ ($\mathcal{O}(\Lambda^{-4})$)								
	$t\bar{t}$	single-top	tW	tZ	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}t\bar{t}$	$t\bar{t}b\bar{b}$
OQQ1								✓	✓
OQQ8								✓	✓
OQt1								✓	✓
OQt8								✓	✓
OQb1								(✓)	✓
OQb8								(✓)	✓
Ott1								✓	✓
Otb1								(✓)	✓
Otb8								✓	✓
OQtQb1									
OQtQb8									
O81qq	✓				✓	✓	✓	✓	✓
O11qq	✓				(✓)	(✓)	(✓)	✓	✓
O83qq	✓	✓		(✓)	✓	✓	✓	✓	✓
O13qq	✓	✓		✓	(✓)	(✓)	(✓)	✓	✓
O8qt	✓				✓	✓	✓	✓	✓
O1qt	✓				(✓)	(✓)	(✓)	✓	✓
O8ut	✓					✓	✓	✓	✓
O1ut	✓					(✓)	(✓)	✓	✓
O8qu	✓					✓	✓	✓	✓
O1qu	✓					(✓)	(✓)	✓	✓
O8dt	✓					✓	✓	✓	✓
O1dt	✓					(✓)	(✓)	✓	✓
O8qd	✓					✓	✓	✓	✓
O1qd	✓					(✓)	(✓)	✓	✓
OtG	✓				✓	✓	✓	✓	✓
OtW		✓	✓	✓					
ObW		(✓)	(✓)						
OtZ				✓		✓			
Off		(✓)	(✓)	(✓)					
Ofq3		✓	✓	✓		✓			
OpQM				✓		✓			
Opt				✓		✓	✓		
Otp							✓		

A large number of different dimension-6 SMEFT operators modify **top production at LHC**

$$\sigma_i^{\rm th}\left(\{c_n\}\right) = \sigma_{{\rm SM},i} + \sum_{n=1}^{N_{\rm op}} \widetilde{\sigma}_{i,n}\frac{c_n}{\Lambda^2} + \sum_{n,m=1}^{N_{\rm op}} \widetilde{\sigma}_{i,nm}\frac{c_n c_m}{\Lambda^4}$$

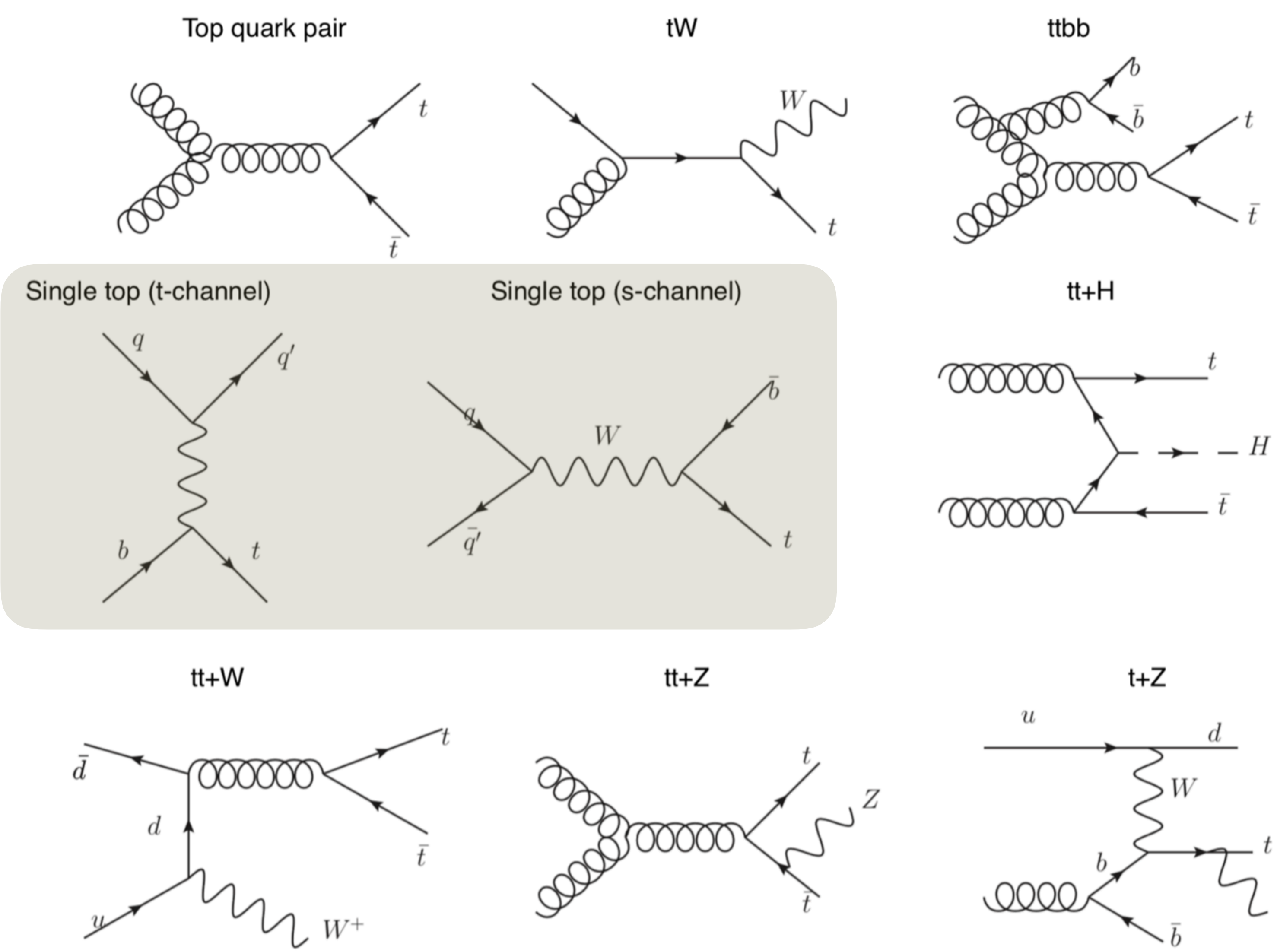


The top quark sector of the SMEFT

Notation	Sensitivity at $\mathcal{O}(\Lambda^{-2})$ ($\mathcal{O}(\Lambda^{-4})$)								
	$t\bar{t}$	single-top	tW	tZ	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}t\bar{t}$	$t\bar{t}b\bar{b}$
OQQ1								✓	✓
OQQ8								✓	✓
OQt1								✓	✓
OQt8								✓	✓
OQb1								(✓)	✓
OQb8								(✓)	✓
Ott1								✓	✓
Otb1								(✓)	✓
Otb8								✓	✓
OQtQb1									
OQtQb8									
O81qq	✓				✓	✓	✓	✓	✓
O11qq	✓				(✓)	(✓)	(✓)	✓	✓
O83qq	✓	✓		(✓)	✓	✓	✓	✓	✓
O13qq	✓	✓		✓	(✓)	(✓)	(✓)	✓	✓
O8qt	✓				✓	✓	✓	✓	✓
O1qt	✓				(✓)	(✓)	(✓)	✓	✓
O8ut	✓					✓	✓	✓	✓
O1ut	✓					(✓)	(✓)	✓	✓
O8qu	✓					✓	✓	✓	✓
O1qu	✓					(✓)	(✓)	✓	✓
O8dt	✓					✓	✓	✓	✓
O1dt	✓					(✓)	(✓)	✓	✓
O8qd	✓					✓	✓	✓	✓
O1qd	✓					(✓)	(✓)	✓	✓
OtG	✓				✓	✓	✓	✓	✓
OtW		✓	✓	✓					
ObW		(✓)	(✓)						
OtZ				✓		✓			
Off		(✓)	(✓)	(✓)					
Ofq3		✓	✓	✓		✓			
OpQM				✓		✓			
Opt				✓		✓	✓		
Otp							✓		

A large number of different dimension-6 SMEFT operators modify **top production at LHC**

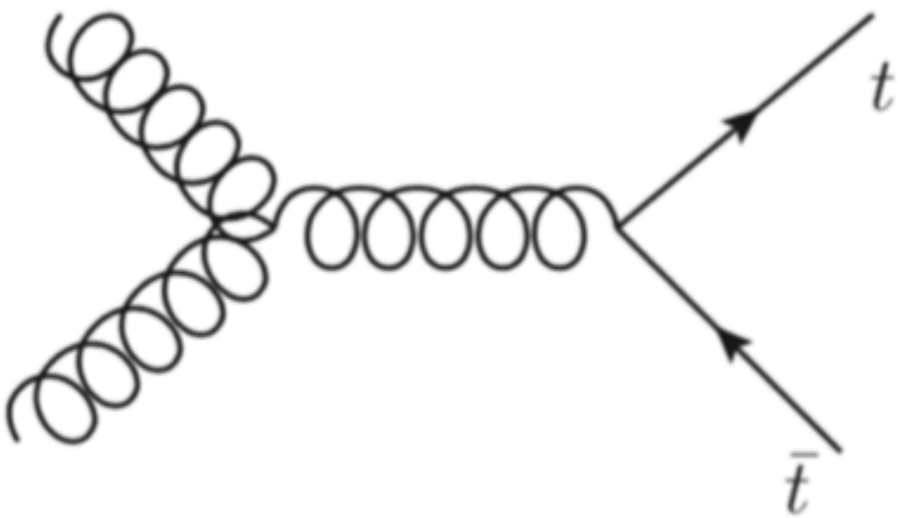
$$\sigma_i^{\rm th}\left(\{c_n\}\right) = \sigma_{{\rm SM},i} + \sum_{n=1}^{N_{\rm op}} \widetilde{\sigma}_{i,n}\frac{c_n}{\Lambda^2} + \sum_{n,m=1}^{N_{\rm op}} \widetilde{\sigma}_{i,nm}\frac{c_n c_m}{\Lambda^4}$$



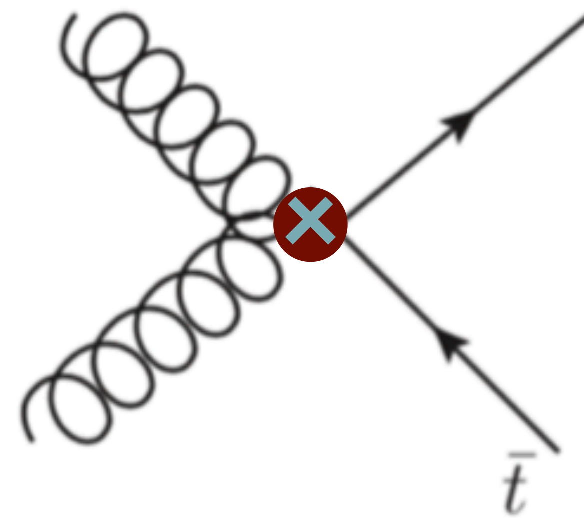
SMEFT effects

Standard Model

$${}^{\ddagger}O_{uG}^{(ij)} = (\bar{q}_i \sigma^{\mu\nu} T^A u_j)\, \tilde{\varphi} G^A_{\mu\nu} \,,$$



+



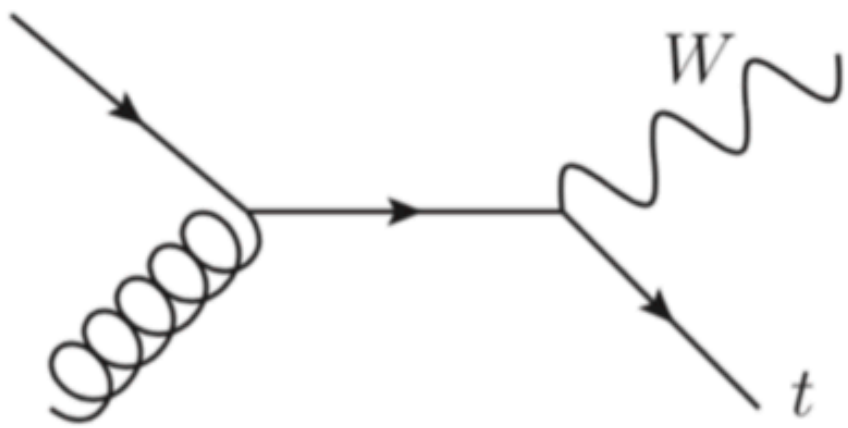
$$= \sigma_{SM} \times \left(1 + a\frac{c_{tG}}{\Lambda^2} + b\frac{c_{tG}^2}{\Lambda^4} \right)$$

SM: N(NLO) QCD **interference** **squared**

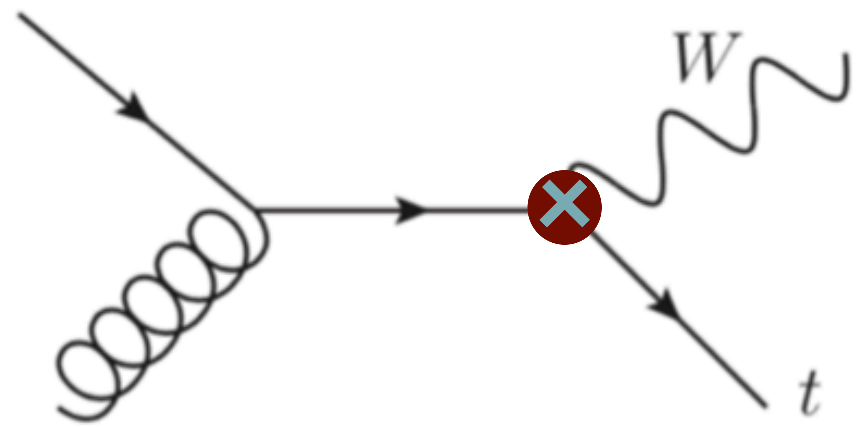
SMEFT effects

Standard Model

$$O_{uW}^{(ij)} = (\bar{q}_i \sigma^{\mu\nu} \tau^I u_j)\, \tilde{\varphi} W^I_{\mu\nu}$$



+



$$= \sigma_{SM} \times \left(1 + a \frac{c_{tW}}{\Lambda^2} + b \frac{c_{tW}^2}{\Lambda^4} \right)$$

SM: N(NLO) QCD **interference** **squared**

Input dataset (I)

Process	Dataset	$\sqrt{s}$	Info	Observables	$N_{\rm dat}$	Ref
$t\bar{t}$	`ATLAS_tt_8TeV_ljets`	**8 TeV**	lepton+jets	$d\sigma/d\lvert y_t\rvert$, $d\sigma/dp_t^T$, $d\sigma/dm_{t\bar{t}}$, $d\sigma/d\lvert y_{t\bar{t}}\rvert$	5, 8, 7, 5	[77]
$t\bar{t}$	`CMS_tt_8TeV_ljets`	**8 TeV**	lepton+jets	$d\sigma/dy_t$, $d\sigma/dp_t^T$, $d\sigma/dm_{t\bar{t}}$, $d\sigma/dy_{t\bar{t}}$	10, 8, 7, 10	[78]
$t\bar{t}$	`CMS_tt2D_8TeV_dilep`	**8 TeV**	dileptons	$d^2\sigma/dy_t dp_t^T$, $d^2\sigma/dy_t dm_{t\bar{t}}$, $d^2\sigma/dp_{t\bar{t}}^T dm_{t\bar{t}}$, $d^2\sigma/dy_{t\bar{t}} dm_{t\bar{t}}$	16, 16, 16, 16	[79]
$t\bar{t}$	`CMS_tt_13TeV_ljets`	**13 TeV**	lepton+jets	$d\sigma/d\lvert y_t\rvert$, $d\sigma/dp_t^T$, $d\sigma/dm_{t\bar{t}}$, $d\sigma/d\lvert y_{t\bar{t}}\rvert$	7, 9, 8, 6	[83]
$t\bar{t}$	`CMS_tt_13TeV_ljets2`	**13 TeV**	lepton+jets	$d\sigma/d\lvert y_t\rvert$, $d\sigma/dp_t^T$, $d\sigma/dm_{t\bar{t}}$, $d\sigma/d\lvert y_{t\bar{t}}\rvert$	11, 12, 10, 10	[85]
$t\bar{t}$	`CMS_tt_13TeV_dilep`	**13 TeV**	dileptons	$d\sigma/dy_t$, $d\sigma/dp_t^T$, $d\sigma/dm_{t\bar{t}}$, $d\sigma/dy_{t\bar{t}}$	8, 6, 6, 8	[86]
$t\bar{t}$	`ATLASCMS_AcMtt_8TeV`	**8 TeV**	Asymm comb	$A_C(m_{t\bar{t}})$, Eq. (3.1)	6	[80]
$t\bar{t}$	`ATLAS_WhelF_8TeV`	**8 TeV**	W helicity fract	F_0, F_L, F_R	3	[81]
$t\bar{t}$	`CMS_WhelF_8TeV`	**8 TeV**	W helicity fract	F_0, F_L, F_R	3	[82]

Input dataset (II)

Process	Dataset	$\sqrt{s}$	Info	Observables	$N_{\rm dat}$	Ref
Single t	`CMS_t_tch_8TeV_inc`	**8 TeV**	t-channel	$\sigma_{\rm tot}(t), \sigma_{\rm tot}(\bar{t})$ (R_t)	2 (1)	[95]
Single t	`CMS_t_sch_8TeV`	**8 TeV**	s-channel	$\sigma_{\rm tot}(t+\bar{t})$	1	[96]
Single t	`ATLAS_t_sch_8TeV`	**8 TeV**	s-channel	$\sigma_{\rm tot}(t+\bar{t})$	1	[97]
Single t	`ATLAS_t_tch_8TeV`	**8 TeV**	t-channel	$d\sigma(tq)/dp_T^t, d\sigma(\bar{t}q)/dp_T^{\bar{t}}$ $d\sigma(tq)/dy_t, d\sigma(\bar{t}q)/dy_t$	5, 4 4, 4	[98]
Single t	`ATLAS_t_tch_13TeV`	**13 TeV**	t-channel	$\sigma_{\rm tot}(t), \sigma_{\rm tot}(\bar{t})$ (R_t)	2 (1)	[99]
Single t	`CMS_t_tch_13TeV_inc`	**13 TeV**	t-channel	$\sigma_{\rm tot}(t+\bar{t})$ (R_t)	1 (1)	[100]
Single t	`CMS_t_tch_8TeV_dif`	**8 TeV**	t-channel	$d\sigma/dp_T^{(t+\bar{t})}$, $d\sigma/d\lvert y^{(t+\bar{t})}\rvert$	6 6	[101]
Single t	`CMS_t_tch_13TeV_dif`	**13 TeV**	t-channel	$d\sigma/dp_T^{(t+\bar{t})}$, $d\sigma/d\lvert y^{(t+\bar{t})}\rvert$	4 4	[102]
tW	`ATLAS_tW_inc_8TeV`	**8 TeV**	inclusive	$\sigma_{\rm tot}(tW)$	1	[103]
tW	`CMS_tW_inc_8TeV`	**8 TeV**	inclusive	$\sigma_{\rm tot}(tW)$	1	[104]
tW	`ATLAS_tW_inc_13TeV`	**13 TeV**	inclusive	$\sigma_{\rm tot}(tW)$	1	[105]
tW	`CMS_tW_inc_13TeV`	**13 TeV**	inclusive	$\sigma_{\rm tot}(tW)$	1	[106]
tZ	`CMS_tZ_inc_13TeV`	**13 TeV**	inclusive	$\sigma_{\rm fid}(Wbl^+l^-q)$	1	[107]
tZ	`ATLAS_tZ_inc_13TeV`	**13 TeV**	inclusive	$\sigma_{\rm tot}(tZq)$	1	[108]

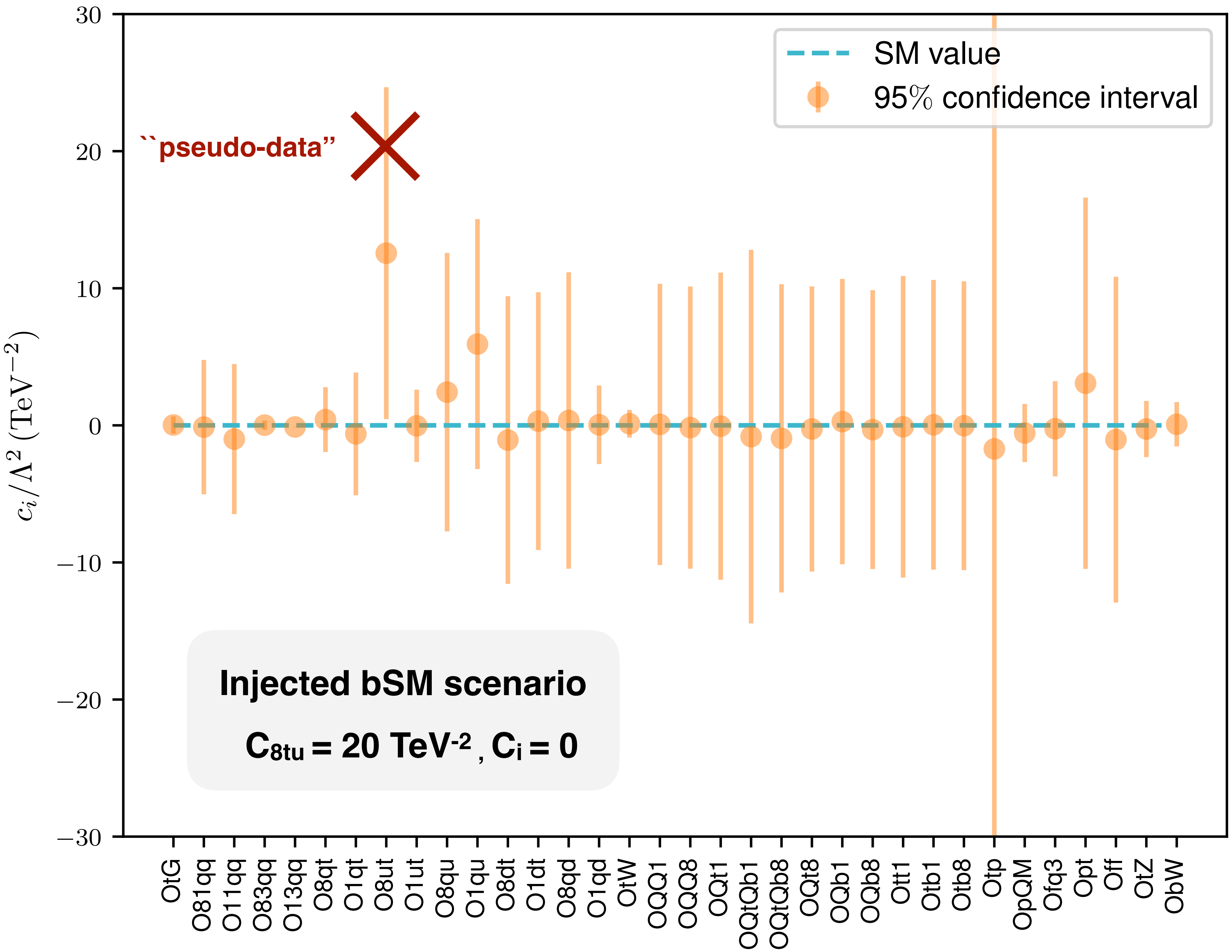
Input dataset (III)

Process	Dataset	$\sqrt{s}$	Info	Observables	$N_{\rm dat}$	Ref
$t\bar{t}b\bar{b}$	`CMS_ttbb_13TeV`	**13 TeV**	total xsec	$\sigma_{\rm tot}(t\bar{t}b\bar{b})$	1	[87]
$t\bar{t}t\bar{t}$	`CMS_tttt_13TeV`	**13 TeV**	total xsec	$\sigma_{\rm tot}(t\bar{t}t\bar{t})$	1	[88]
$t\bar{t}Z$	`CMS_ttZ_8_13TeV`	**8+13 TeV**	total xsec	$\sigma_{\rm tot}(t\bar{t}Z)$	2	[89, 90]
$t\bar{t}Z$	`ATLAS_ttZ_8_13TeV`	**8+13 TeV**	total xsec	$\sigma_{\rm tot}(t\bar{t}Z)$	2	[91, 92]
$t\bar{t}W$	`CMS_ttW_8_13TeV`	**8+13 TeV**	total xsec	$\sigma_{\rm tot}(t\bar{t}W)$	2	[89, 90]
$t\bar{t}W$	`ATLAS_ttW_8_13TeV`	**8+13 TeV**	total xsec	$\sigma_{\rm tot}(t\bar{t}W)$	2	[91, 92]
$t\bar{t}H$	`CMS_tth_13TeV`	**13 TeV**	signal strength	$\mu_{t\bar{t}H}$	1	[93]
$t\bar{t}H$	`ATLAS_tth_13TeV`	**13 TeV**	total xsec	$\sigma_{\rm tot}(t\bar{t}H)$	1	[94]

The fit includes more than **100 cross-section measurements** at 8 and 13 TeV from **10 different top-quark production processes**

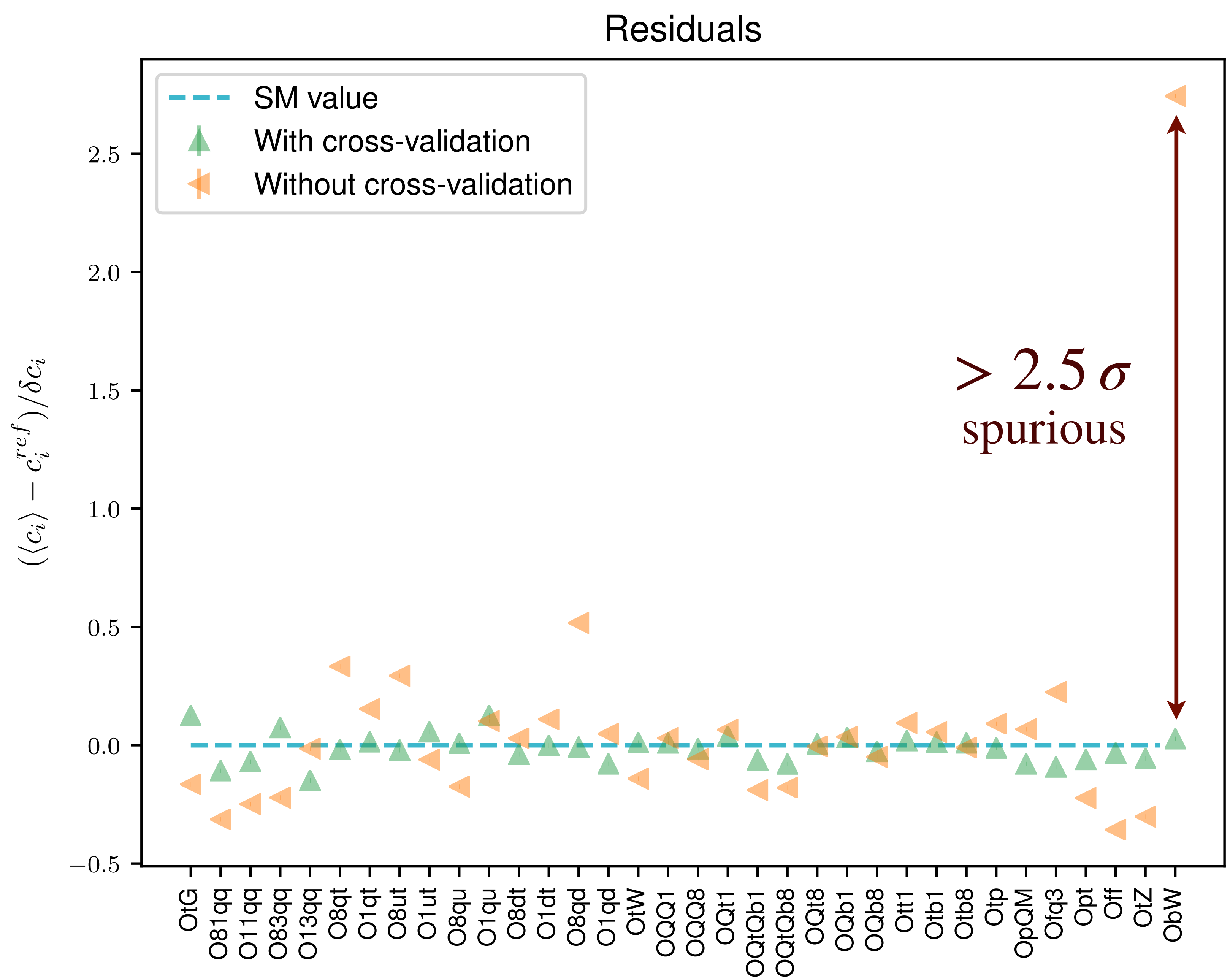
Closure Tests

- Generate **pseudo-data based** on a given scenario (SM or BSM) and check that the correct (known) results are reproduced after the fit
- Allows quantifying the **expected statistical significance** for BSM deviations



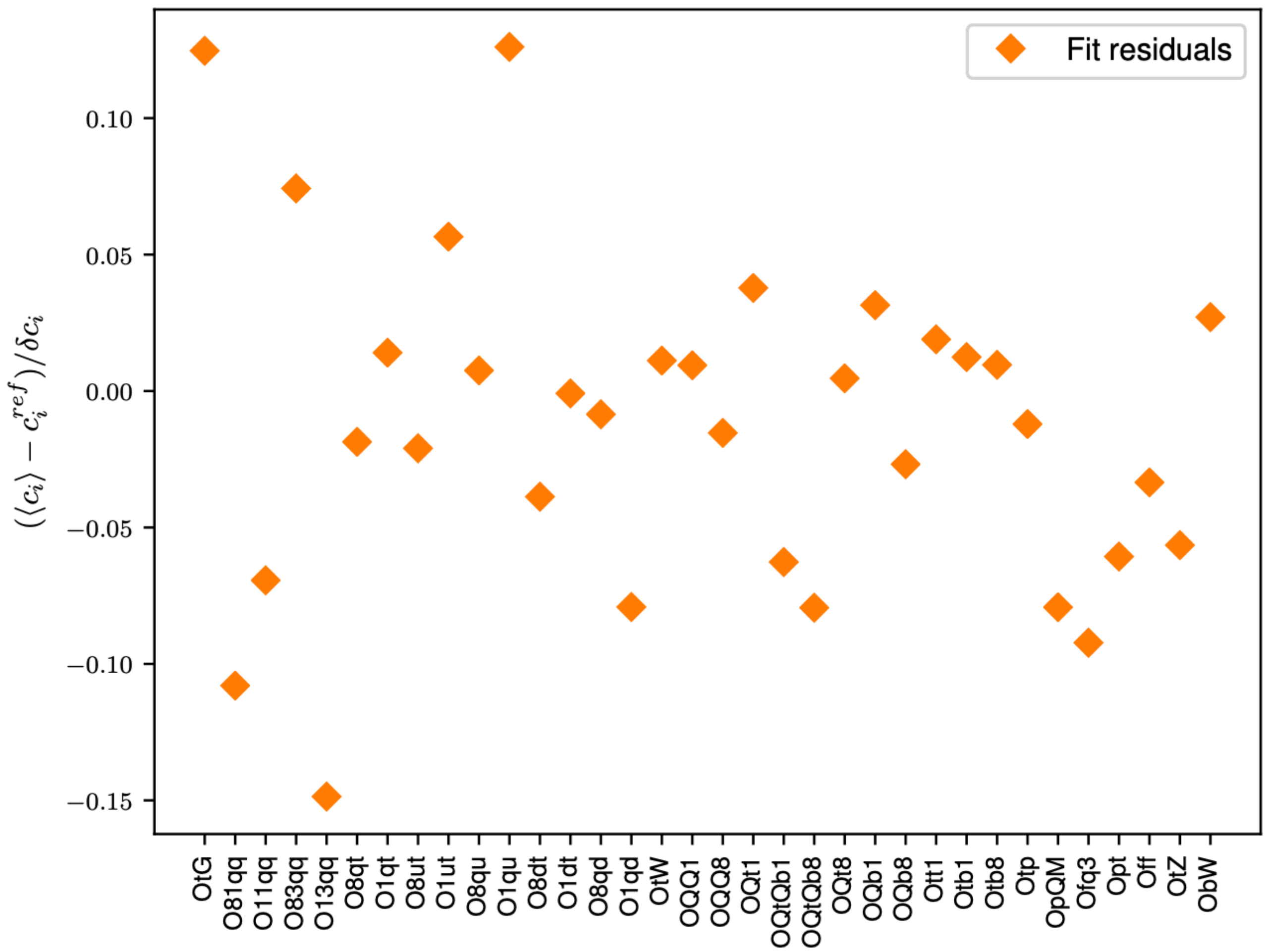
Cross-validation

- Since N_{par} is not too different from N_{dat}, overfitting will take place for an efficient optimiser
- **Artificial tensions with the SM** are likely to be generated by overfitting!
- Fit residuals **consistent with true result (SM)** only with cross-validation

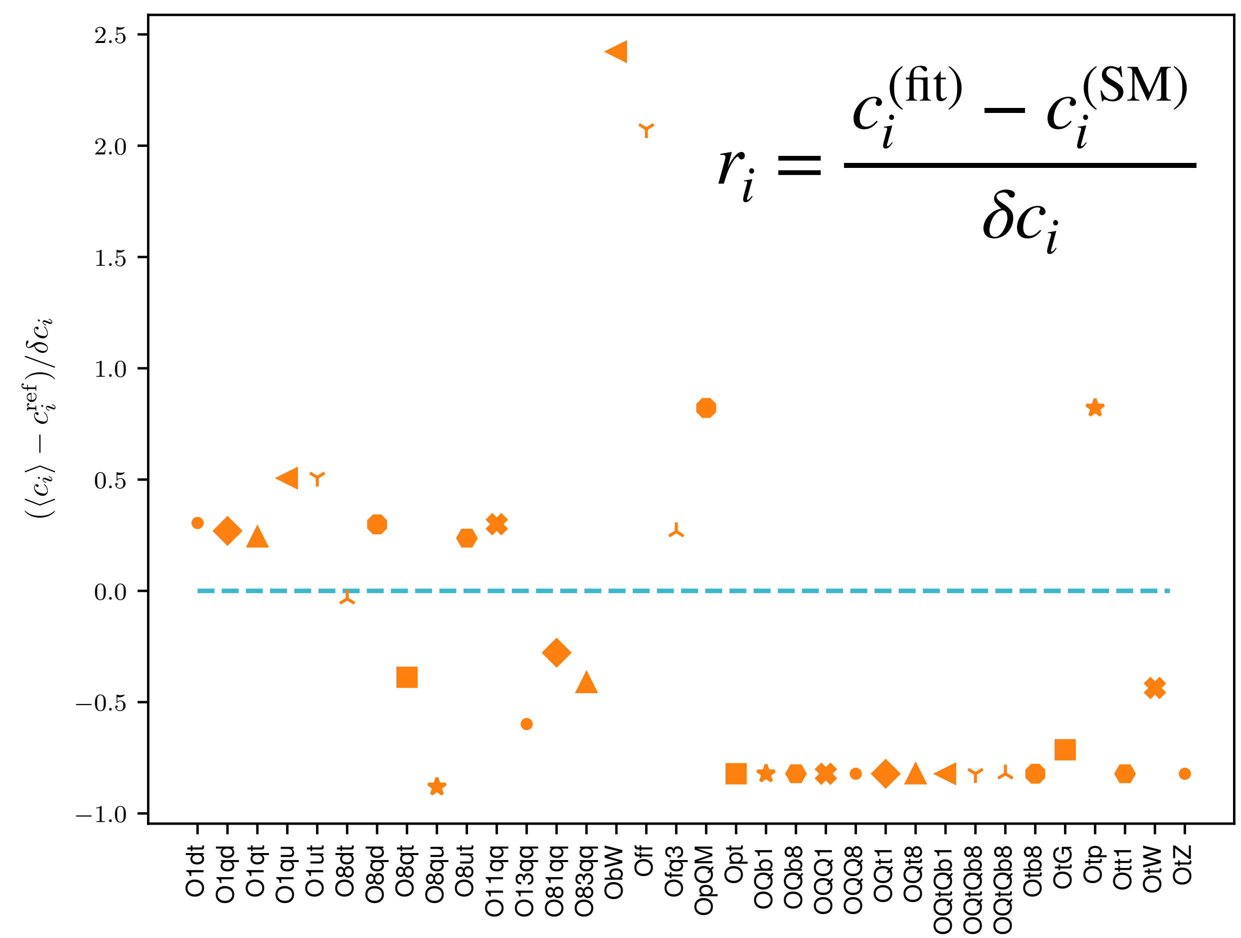


Global vs individual fits

Global fits (all params fitted)



Individual fits (one param at a time)



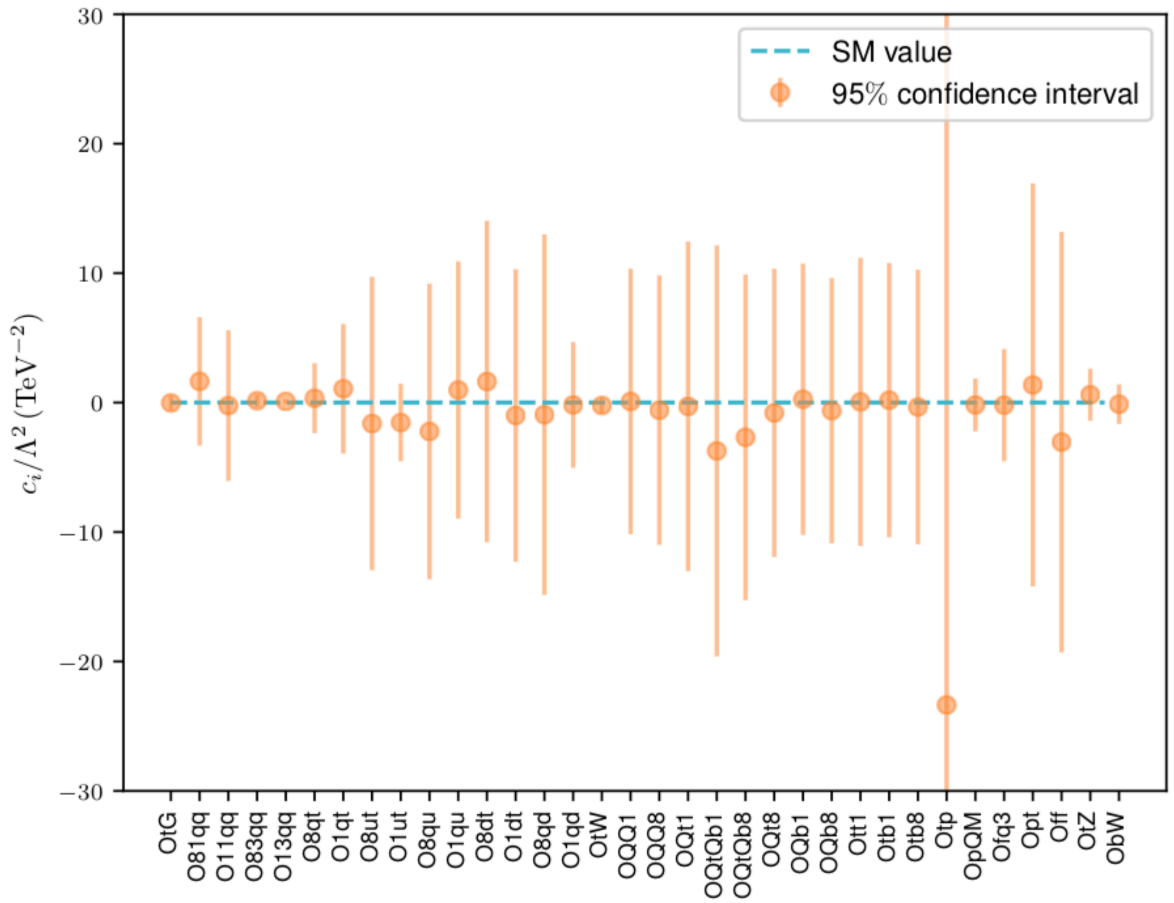
- If each operator was a truly **independent random variable**, we would expect that at least **2 operators** have residuals **|r| > 1** (bounds are 95% CL)
- This is far from being the case when all operators are fitted simultaneously
- Explained by **correlations between operators + degeneracies in parameter space**: much larger fluctuations if we fit one operator at a time

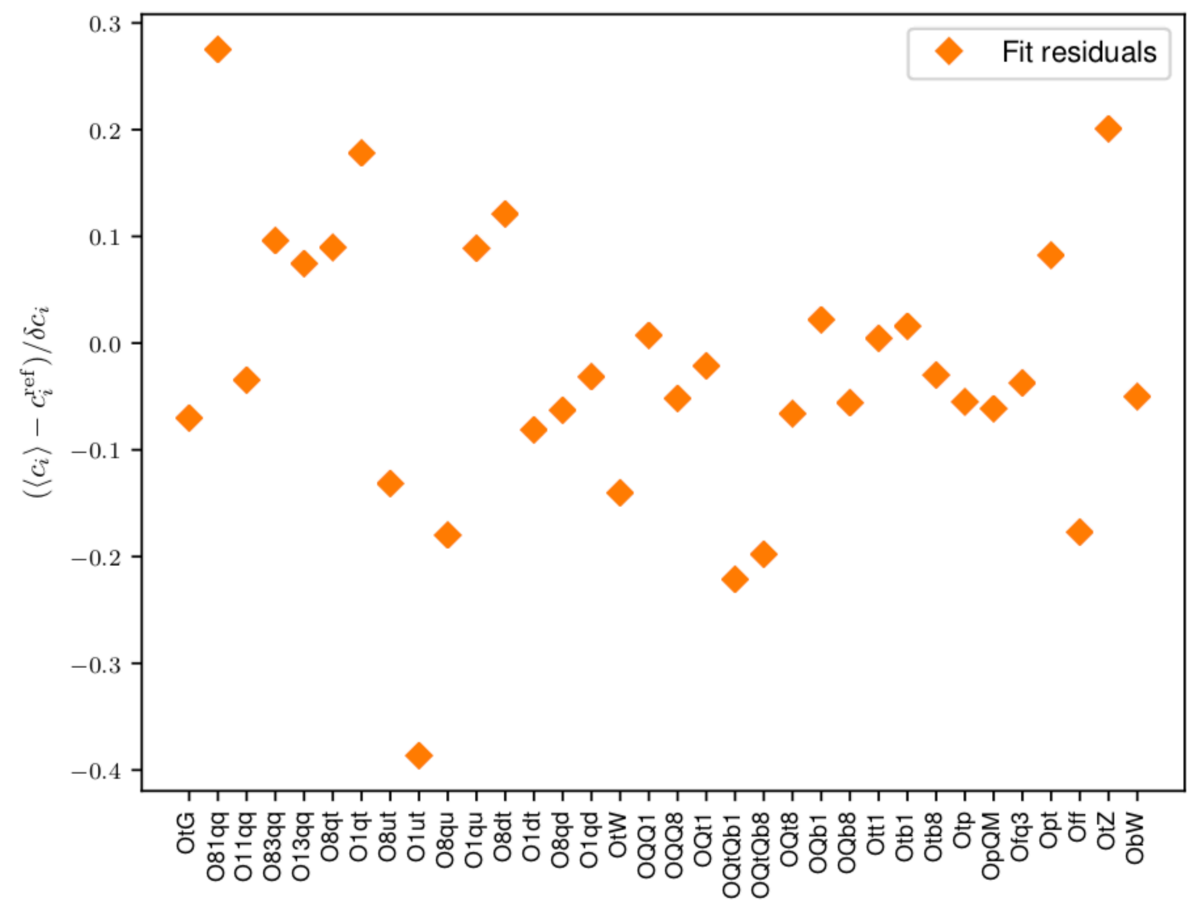
Fit quality

- **Good agreement** between theory (SM and SMEFT) and data for most datasets
- For the **103 fitted cross-sections**, we find χ^2/n_{dat} *of* **1.11 (1.06)** before (after) fit
- Including SMEFT effects improves agreement with data: need to quantify how **significant** this improvement is

Dataset	$\chi^2/n_{\rm dat}$ (prior)	$\chi^2/n_{\rm dat}$ (fit)	$n_{\rm dat}$
ATLAS_tt_8TeV_ljets [$m_{t\bar{t}}$]	1.51	1.25	7
CMS_tt_8TeV_ljets [$y_{t\bar{t}}$]	1.17	1.17	10
CMS_tt2D_8TeV_dilep [$(m_{t\bar{t}}, y_t)$]	1.38	1.38	16
CMS_tt_13TeV_ljets2 [$m_{t\bar{t}}$]	1.09	1.28	8
CMS_tt_13TeV_dilep [$m_{t\bar{t}}$]	1.34	1.42	6
CMS_tt_13TeV_ljets_2016 [$m_{t\bar{t}}$]	1.87	1.87	10
ATLAS_WhelF_8TeV	1.98	0.27	3
CMS_WhelF_8TeV	0.31	1.18	3
CMS_ttbb_13TeV	5.00	1.29	1
CMS_tttt_13TeV	0.05	0.02	1
ATLAS_tth_13TeV	1.61	0.55	1
CMS_tth_13TeV	0.34	0.01	1
ATLAS_ttZ_8TeV	1.32	5.29	1
ATLAS_ttZ_13TeV	0.01	1.06	1
CMS_ttZ_8TeV	0.04	0.06	1
CMS_ttZ_13TeV	0.90	0.67	1
ATLAS_ttW_8TeV	1.34	0.27	1
ATLAS_ttW_13TeV	0.82	0.65	1
CMS_ttW_8TeV	1.54	0.54	1
CMS_ttW_13TeV	0.03	0.09	1
CMS_t_tch_8TeV_dif	0.11	0.32	6
ATLAS_t_tch_8TeV [y_t]	0.91	0.43	4
ATLAS_t_tch_8TeV [$y_{\bar{t}}$]	0.39	0.45	4
ATLAS_t_sch_8TeV	0.08	1.92	1
ATLAS_t_tch_13TeV	0.02	0.09	2
CMS_t_tch_13TeV_dif [y_t]	0.46	0.49	4
CMS_t_sch_8TeV	1.26	0.76	1
ATLAS_tW_inc_8TeV	0.02	0.06	1
CMS_tW_inc_8TeV	0.00	0.07	1
ATLAS_tW_inc_13TeV	0.52	0.82	1
CMS_tW_inc_13TeV	4.29	1.68	1
ATLAS_tZ_inc_13TeV	0.00	0.00	1
CMS_tZ_inc_13TeV	0.66	0.34	1
Total	**1.11**	**1.06**	**103**

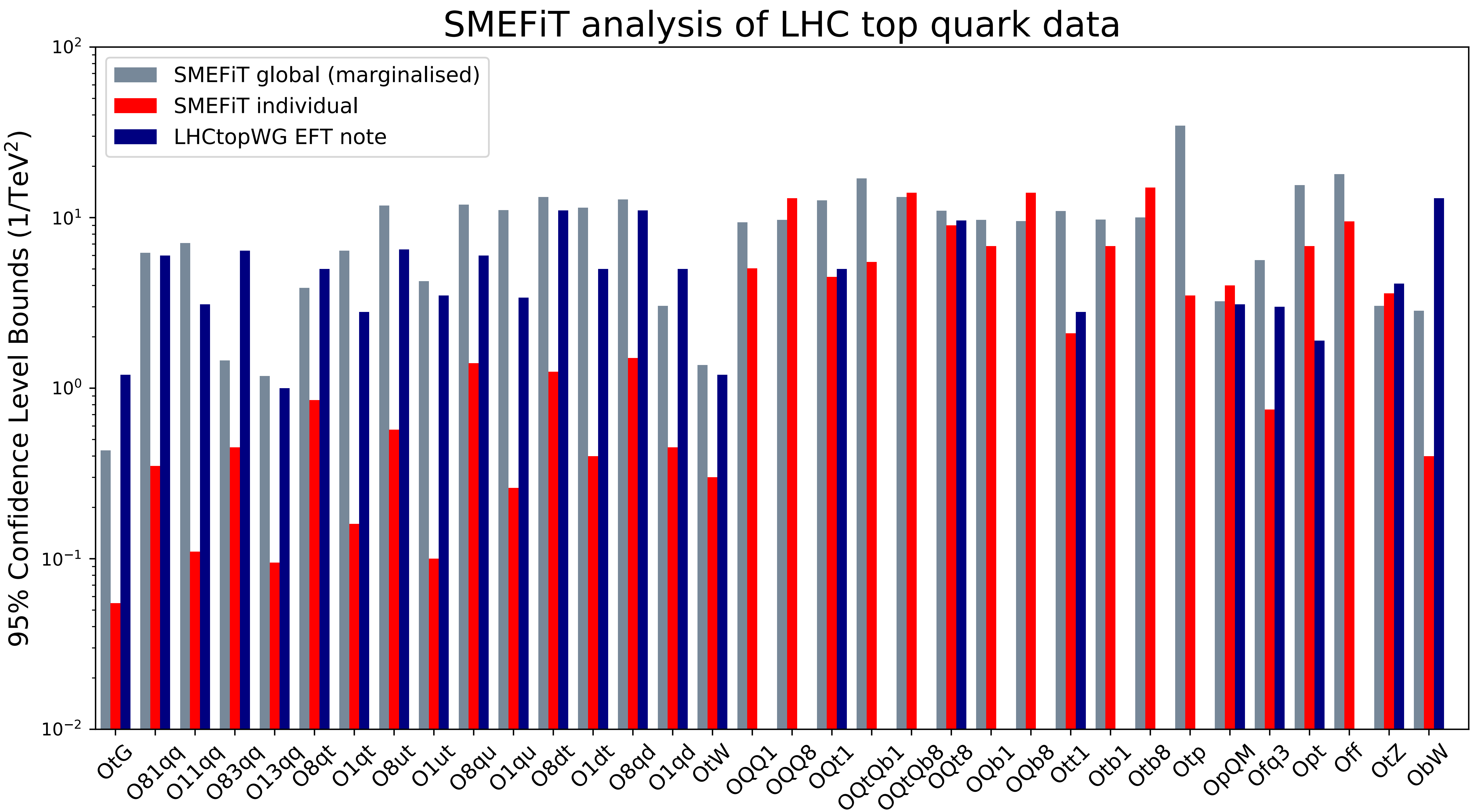
SMEFiT results





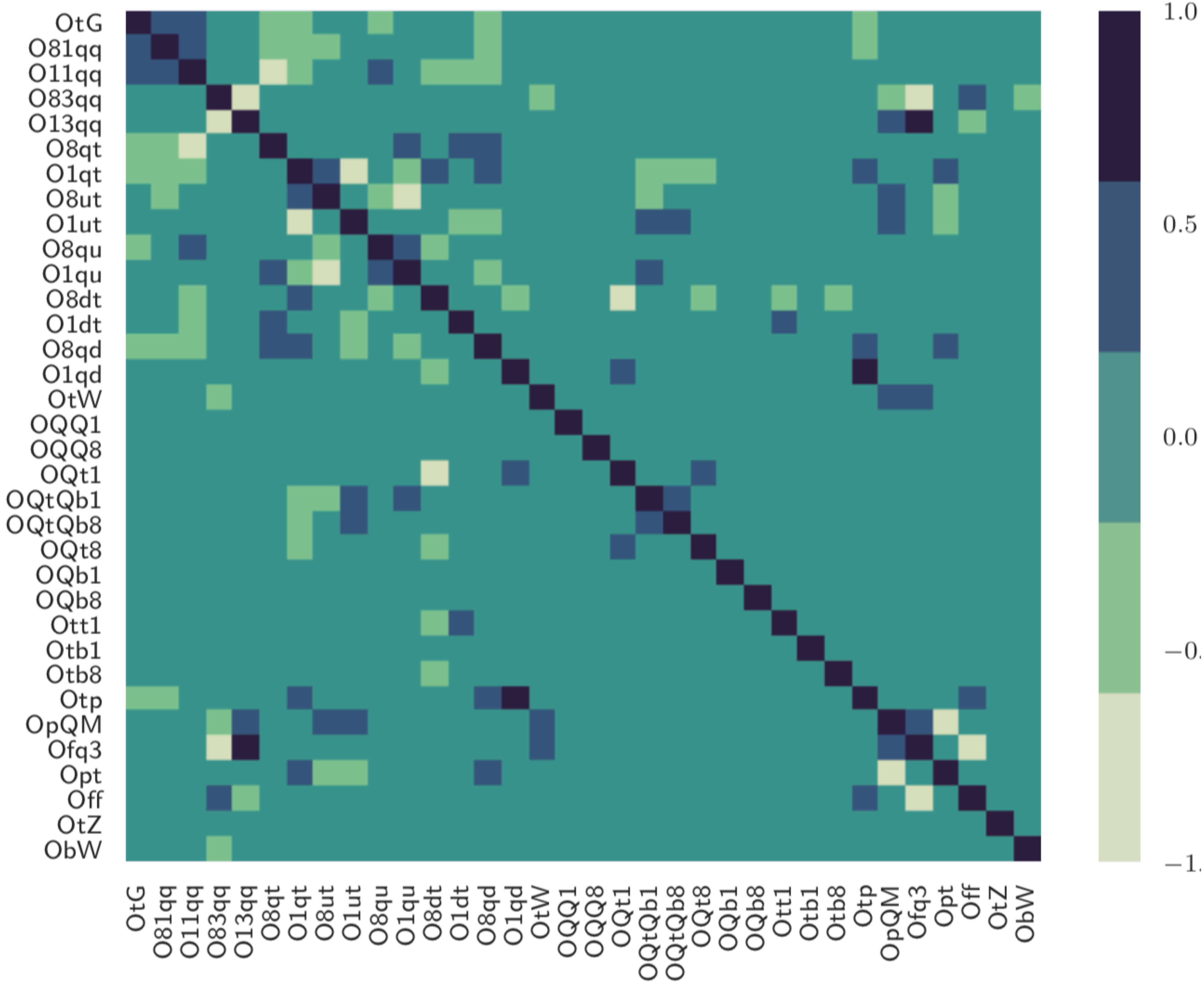
- Agreement with the SM expectation **within uncertainties**
- Bounds on individual operators are in general largely **correlated among them**
- Large differences between the bounds obtained from each operator

Comparison with 1D fits and previous bounds

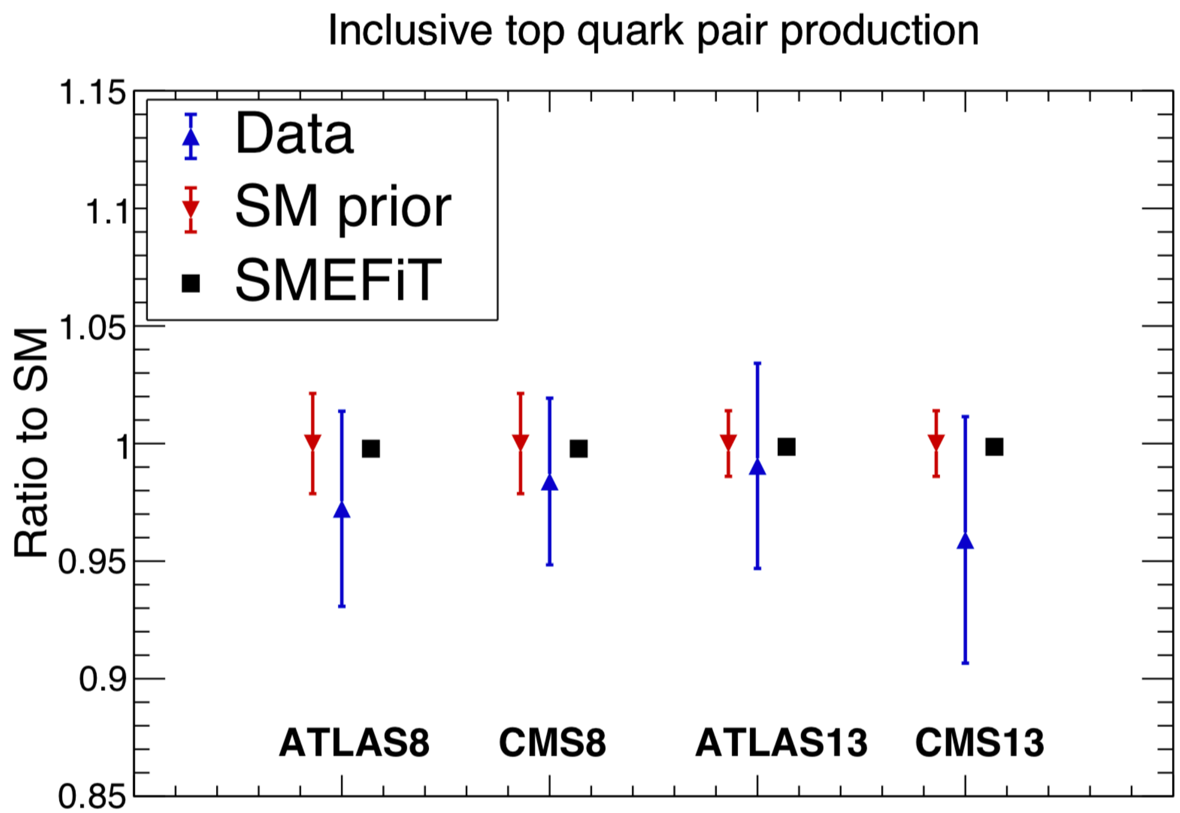


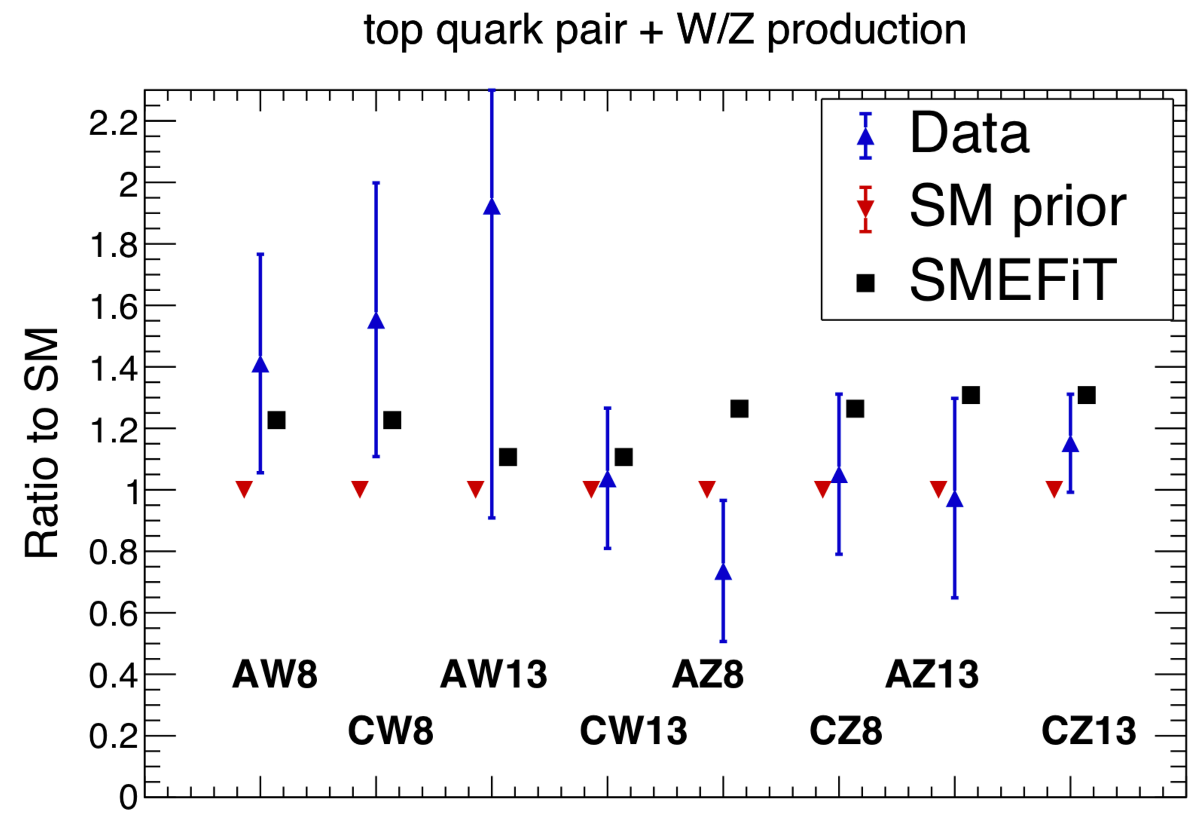
- **Improvement** found (more stringent bounds) in most fitted degrees of freedom
- For some specific operators **our bounds are the first ones** to be reported
- Individual bounds can dramatically **overestimate** the actual (marginalised) bounds

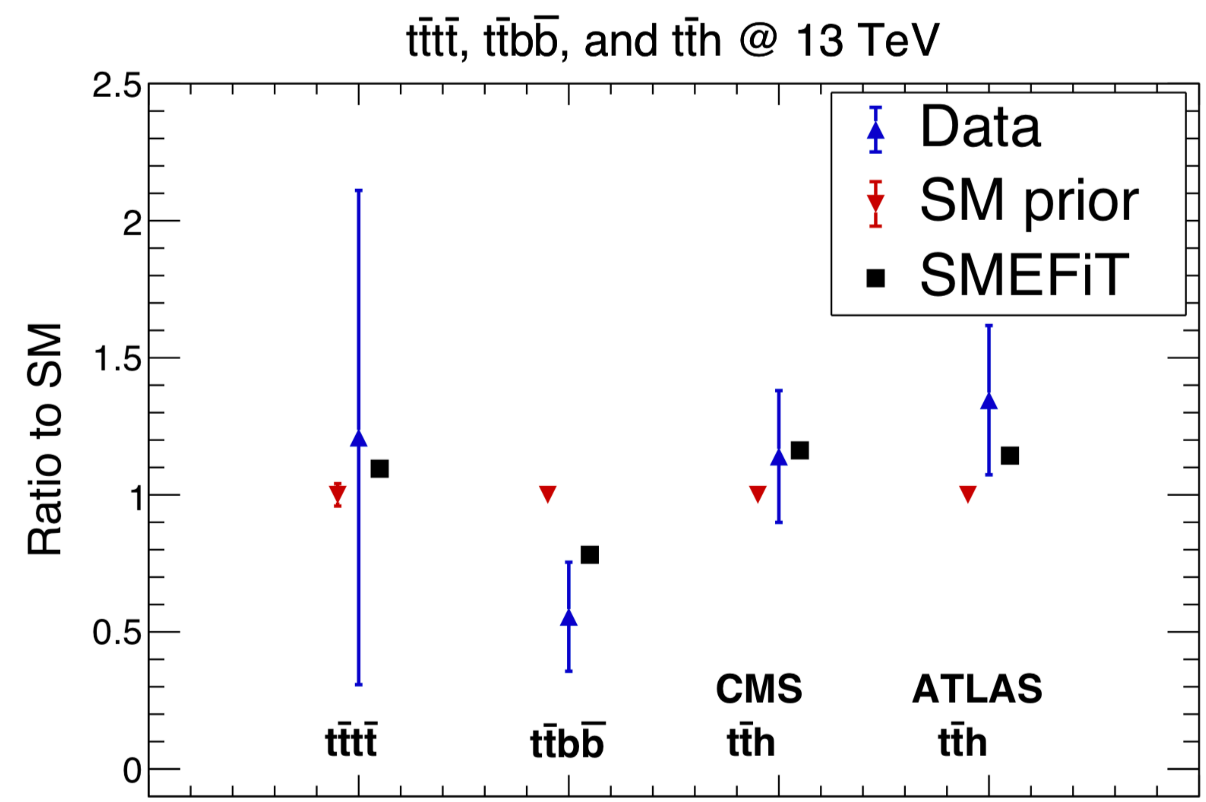
Correlation map



Comparison with data





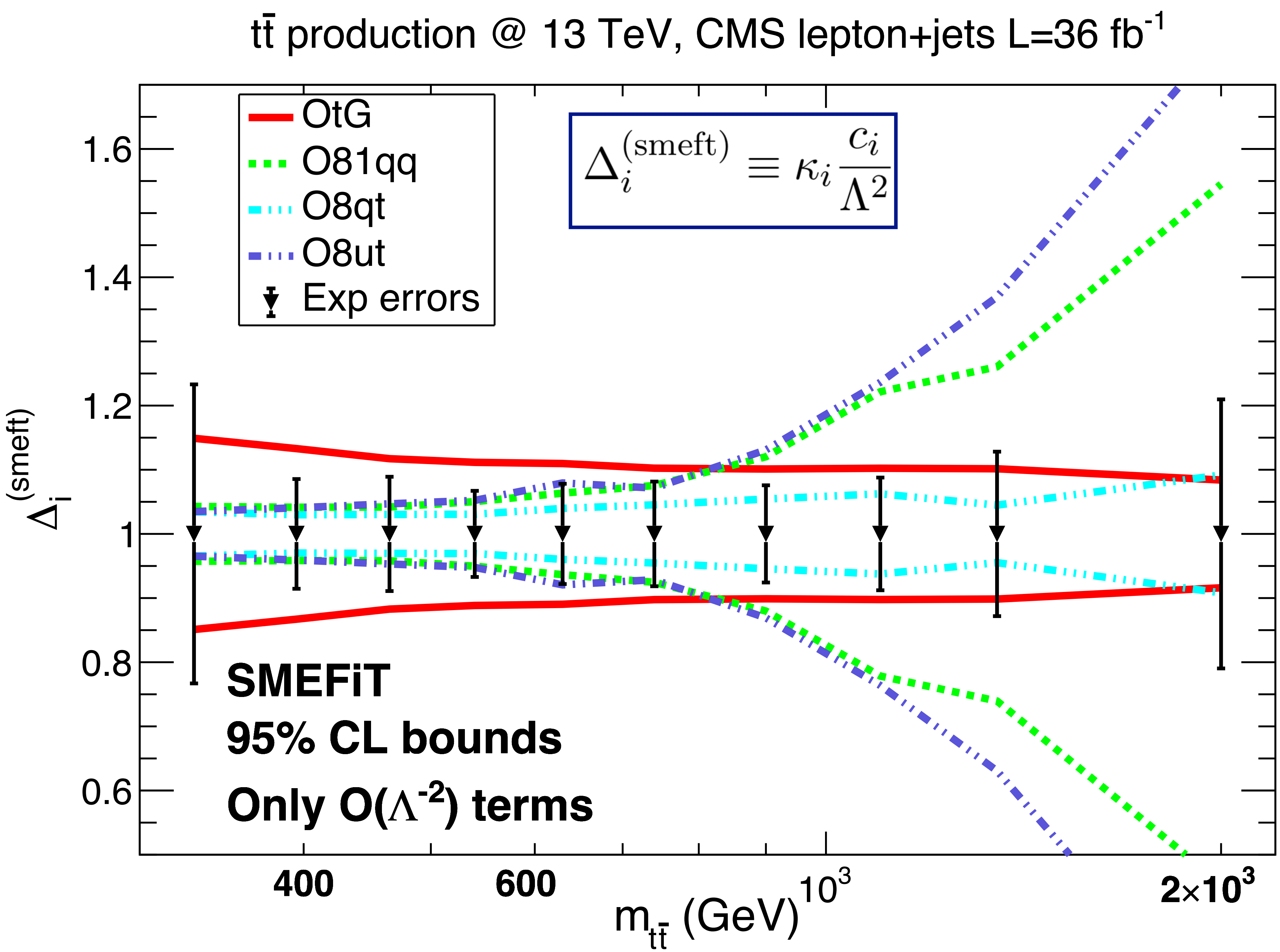


- The **best-fit SMEFT-induced shift** wrt the SM calculation depends on the process

- For inclusive top quark pair and single top, the SMEFT shifts are < 2%

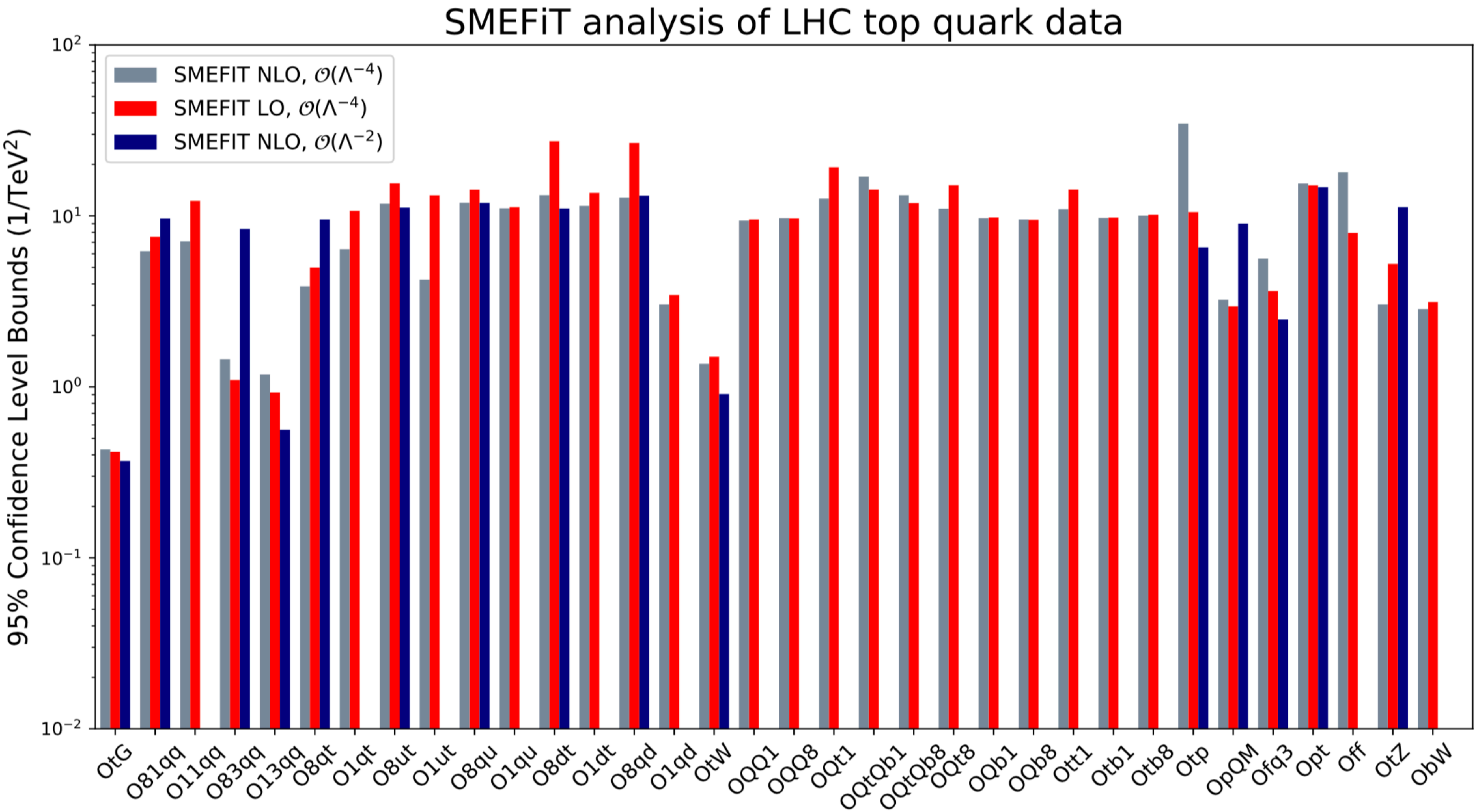
- For ***tttt, ttbbb***, and ***tth*** the SMEFT shifts can be as large as 20% (reflecting the larger experimental errors)

High-energy behaviour



Energy-growing effects enhance sensitivity to SMEFT effects with **TeV-scale cross-sections** but need to be careful to ensure **validity of EFT description**

Dependence on theory settings



Accounting for the **quadratic O(Λ^{-4}) terms** strengthens bounds for several operators

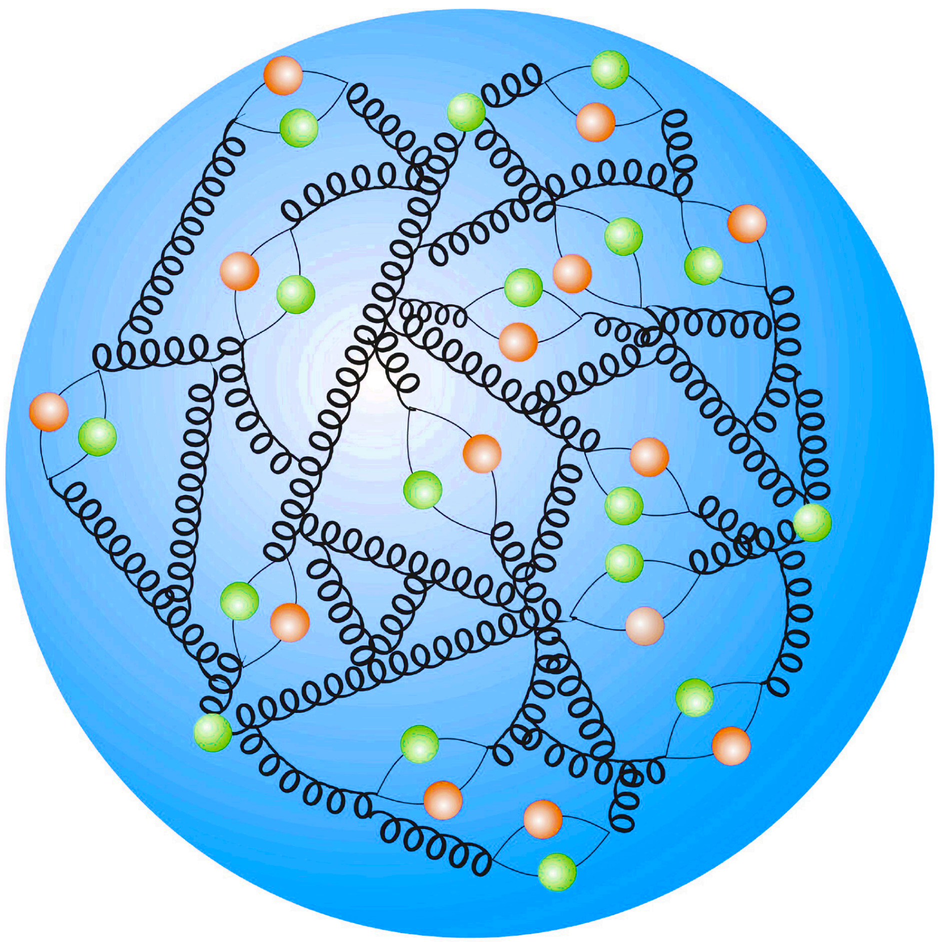
Simultaneous fits of the SMEFT and the proton structure

Based on S. Carrazza, C. Degrande, S. Iranipour, J. Rojo, M. Ubiali

arXiv:1905.05215, *submitted to PRL*

Proton structure: parton distributions

Proton energy divided among constituents: **quarks** and **gluons**



Parton Distribution Functions (PDFs)

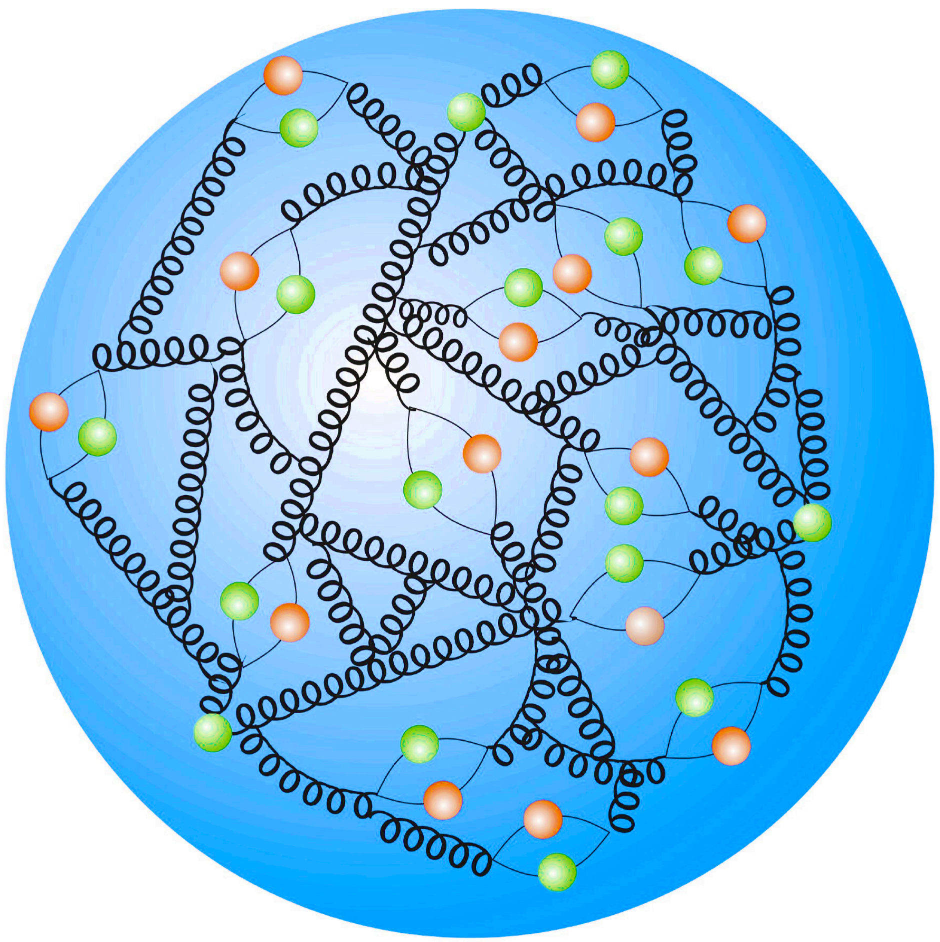
↓

Determine from **data:**

Global QCD analysis

Proton structure: parton distributions

Proton energy divided among constituents: **quarks** and **gluons**



Parton Distribution Functions (PDFs)

Determine from **data:**

Global QCD analysis

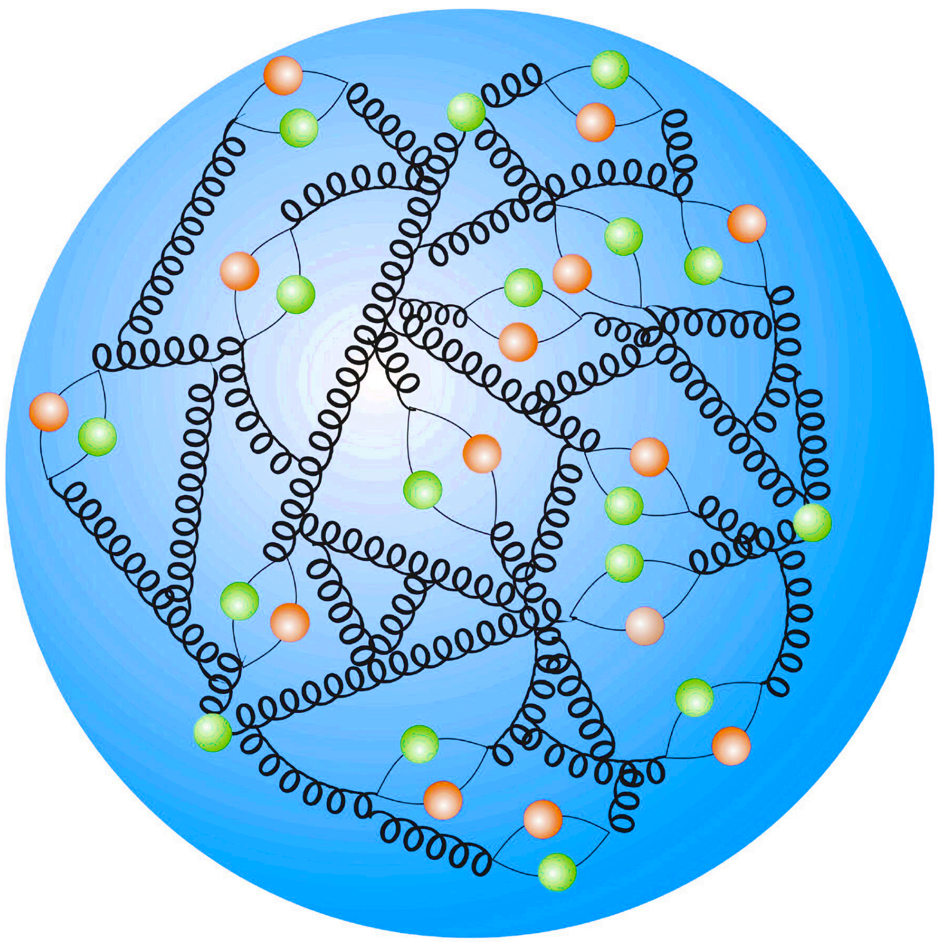
Mass? Spin?

Heavy quark content?

Novel QCD dynamics?

Proton structure: parton distributions

Proton energy divided among constituents: **quarks** and **gluons**



Parton Distribution Functions (PDFs)

Determine from **data:**

Global QCD analysis

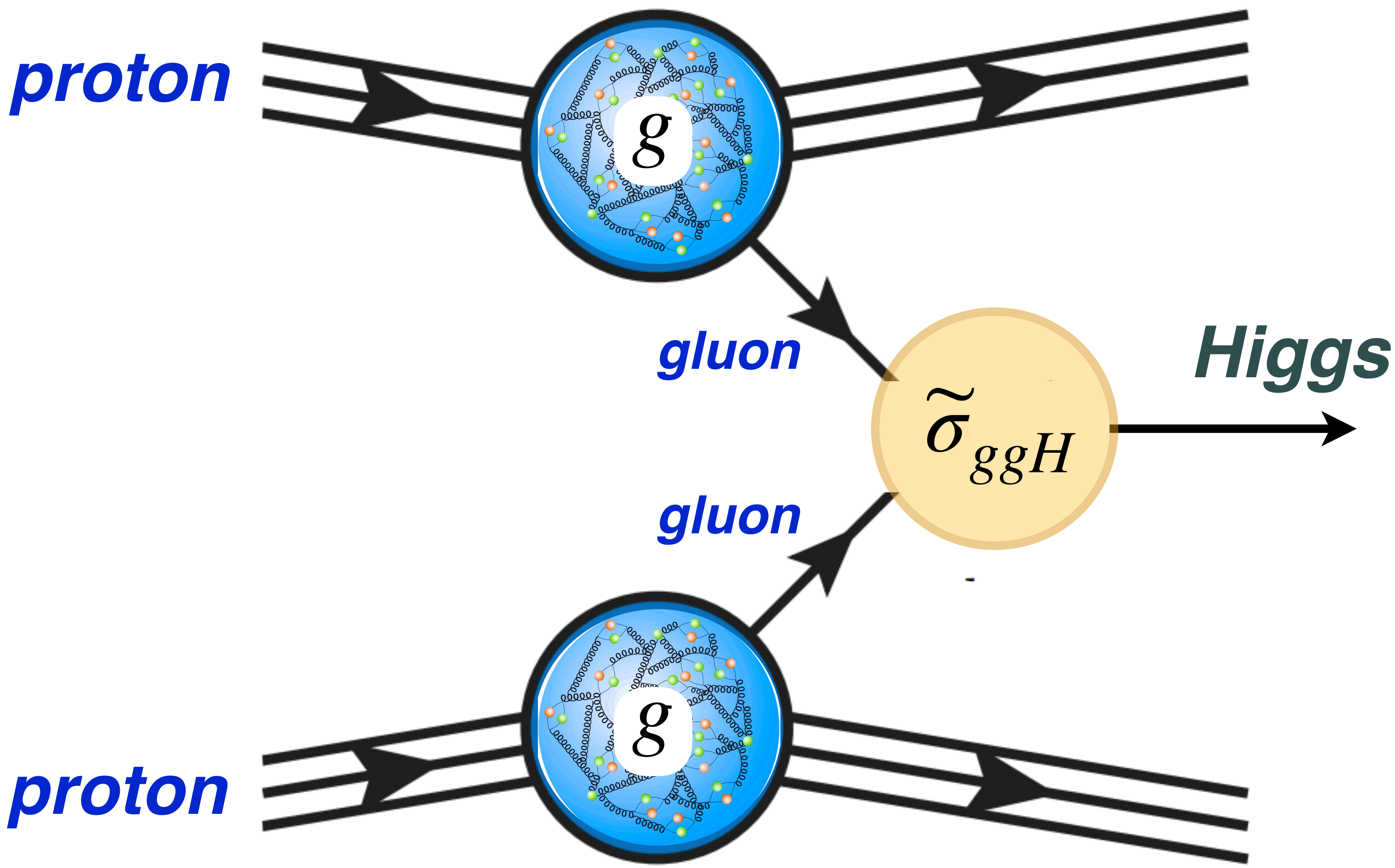
Mass? Spin?
Heavy quark content?
Novel QCD dynamics?

Theoretical predictions for LHC, RHIC, IceCube?

Proton structure: parton distributions

$$N_{\mathrm{LHC}}(H) \sim g \otimes g \otimes \widetilde{\sigma}_{ggH}$$

Parton Distributions



All-order structure: **QCD factorisation theorems**

Parton Distributions

$$g(x, Q)$$

Probability of **finding a gluon inside a proton**, carrying a fraction ***x*** of the proton momentum, when probed with energy ***Q***

x: fraction of proton momentum carried by gluon

Energy of hard-scattering reaction**:** inverse of resolution length

Dependence on *x* fixed by **non-perturbative QCD dynamics**: extract from experimental data

- **Energy conservation**: momentum sum rule

$$\int_0^1 dx\, x \left(\sum_{i=1}^{n_f} \left[q_i((x,Q^2) + \bar{q}_i(x,Q^2) \right] + g(x,Q^2) \right) = 1$$

- **Quark number conservation**: valence sum rules

$$\int_0^1 dx\, \left(u(x,Q^2) + \bar{u}(x,Q^2) \right) = 2$$

Parton Distributions

$$g(x, Q)$$

Probability of **finding a gluon inside a proton**, carrying a fraction ***x*** of the proton momentum, when probed with energy ***Q***

x: fraction of proton momentum carried by gluon

Energy of hard-scattering reaction**:** inverse of resolution length

Dependence on ***Q*** fixed by **perturbative QCD dynamics**: computed up to $\mathcal{O}\left(\alpha_s^4\right)$

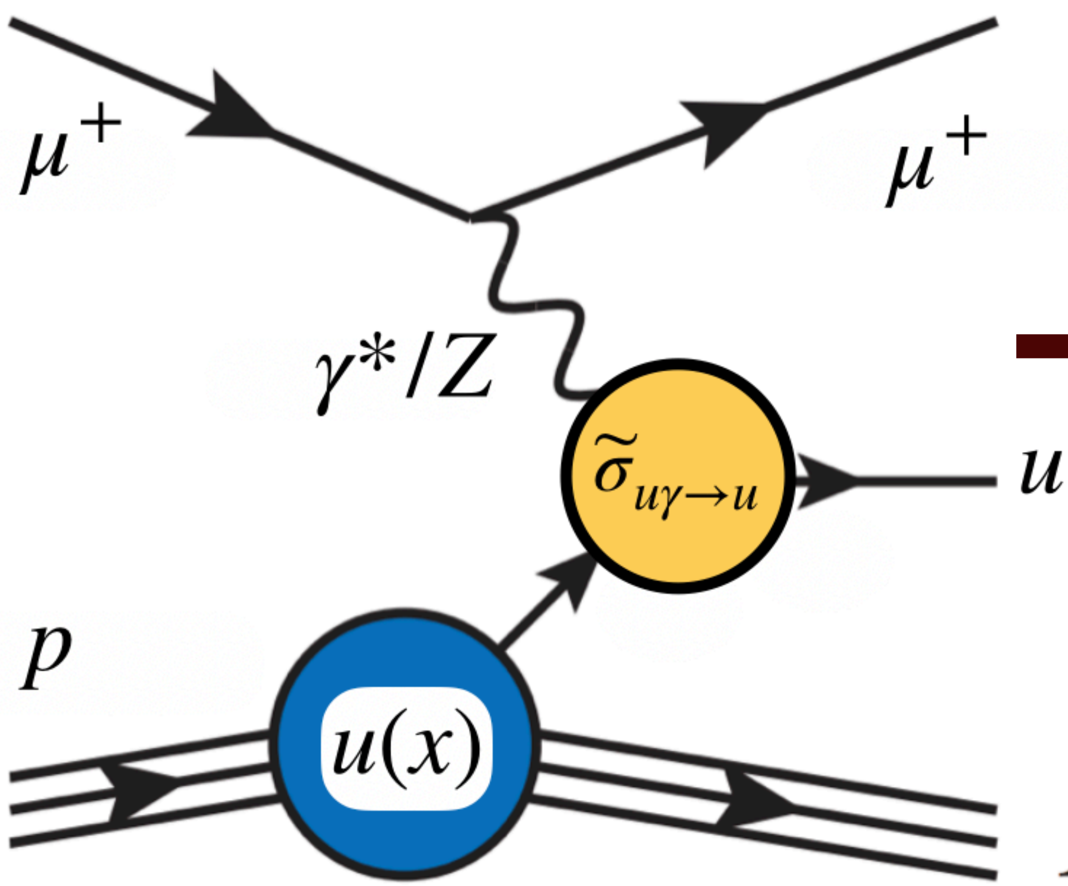
$$\frac{\partial}{\partial \ln Q^2} q_i(x, Q^2) = \int_x^1 \frac{dz}{z} P_{ij}\left(\frac{x}{z}, \alpha_s(Q^2)\right) q_j(z, Q^2)$$

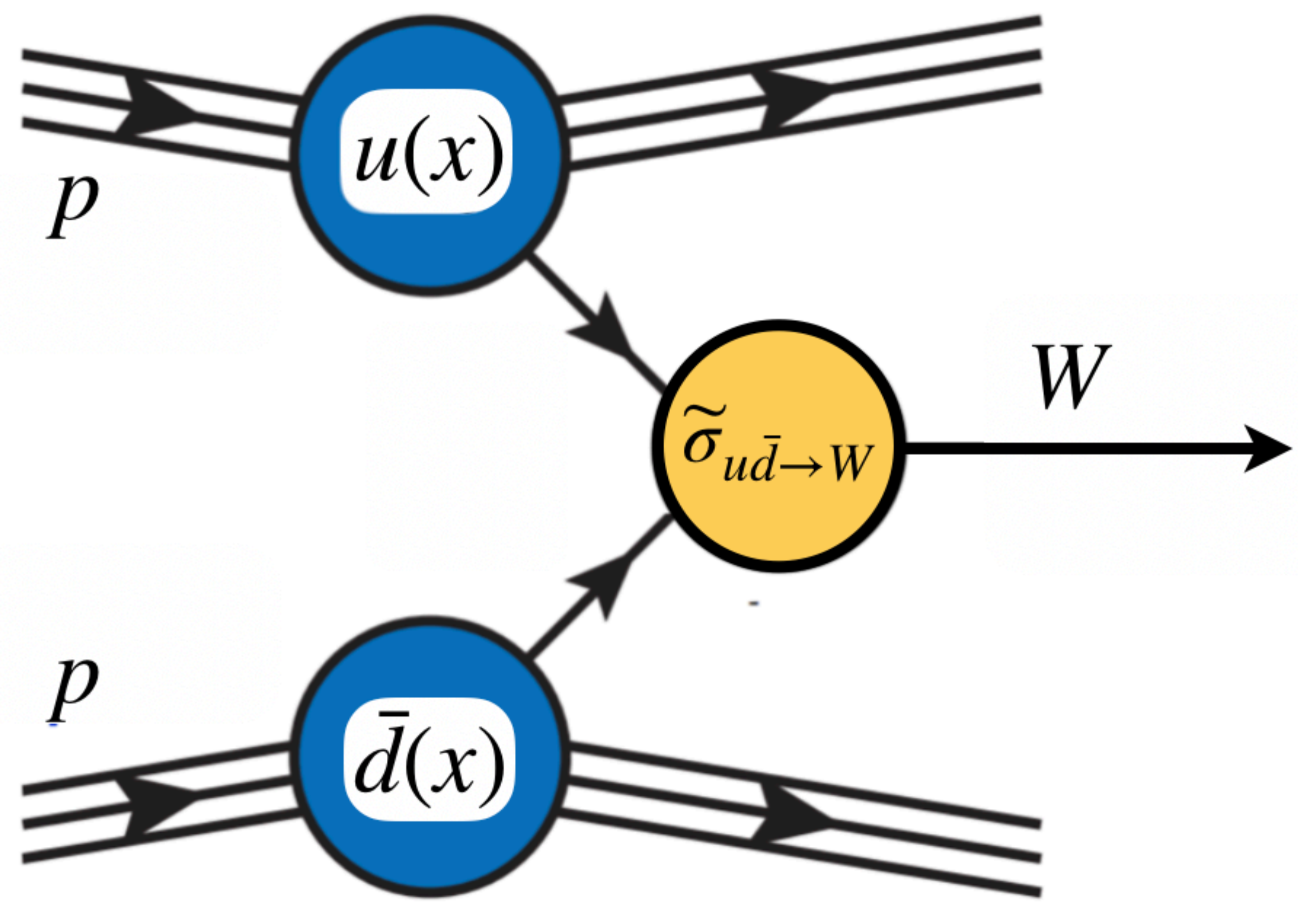
***DGLAP** parton evolution equations*

The Global QCD analysis paradigm

QCD factorisation theorems: **PDF universality**

$$\sigma_{l\,p\to\mu\,X} = \widetilde{\sigma}_{u\gamma\to u} \otimes u(x) \;\;\rightarrow\;\; \sigma_{p\,p\to W} = \widetilde{\sigma}_{u\bar{d}\to W} \otimes u(x) \otimes \bar{d}(x)$$

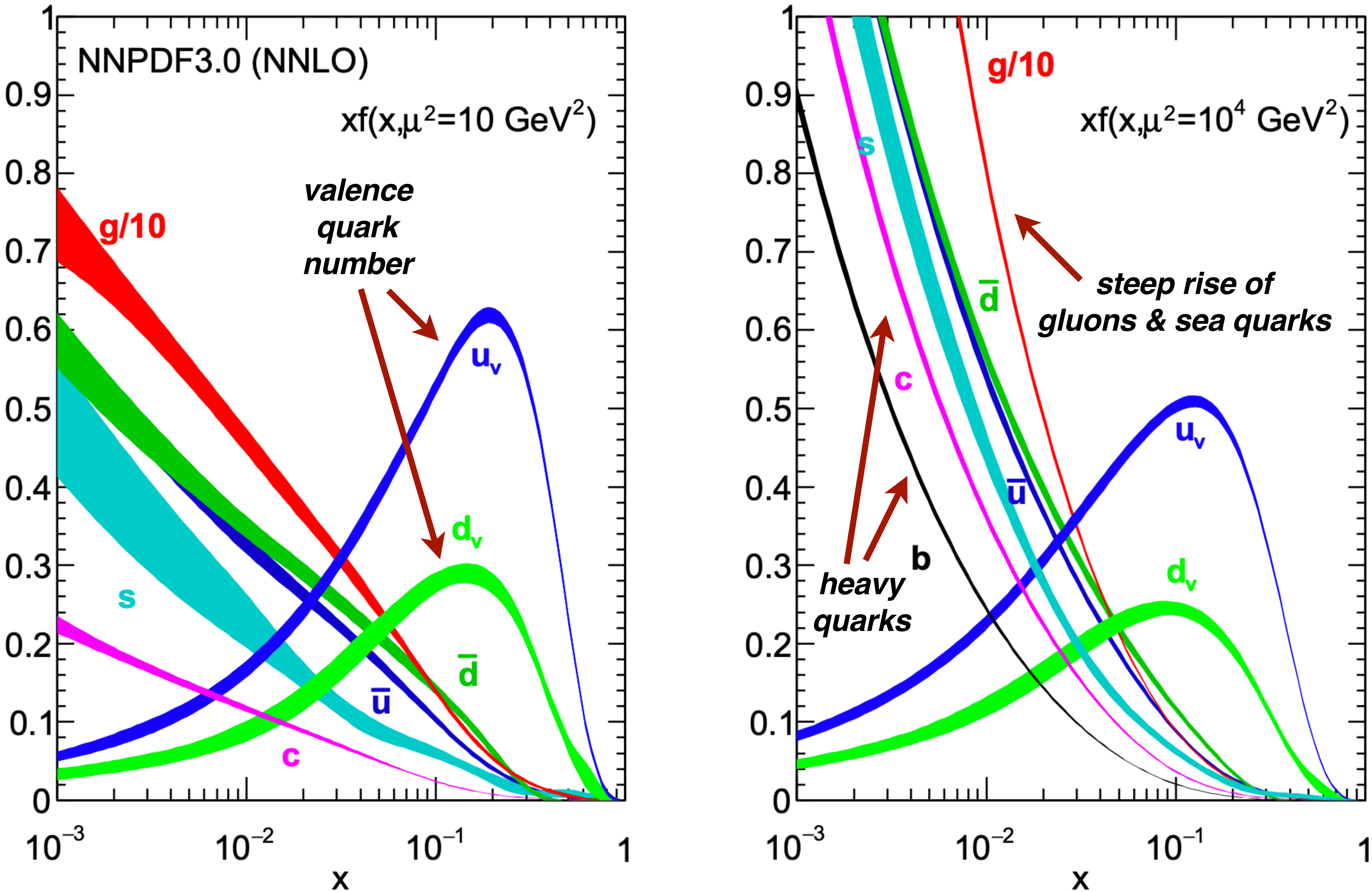




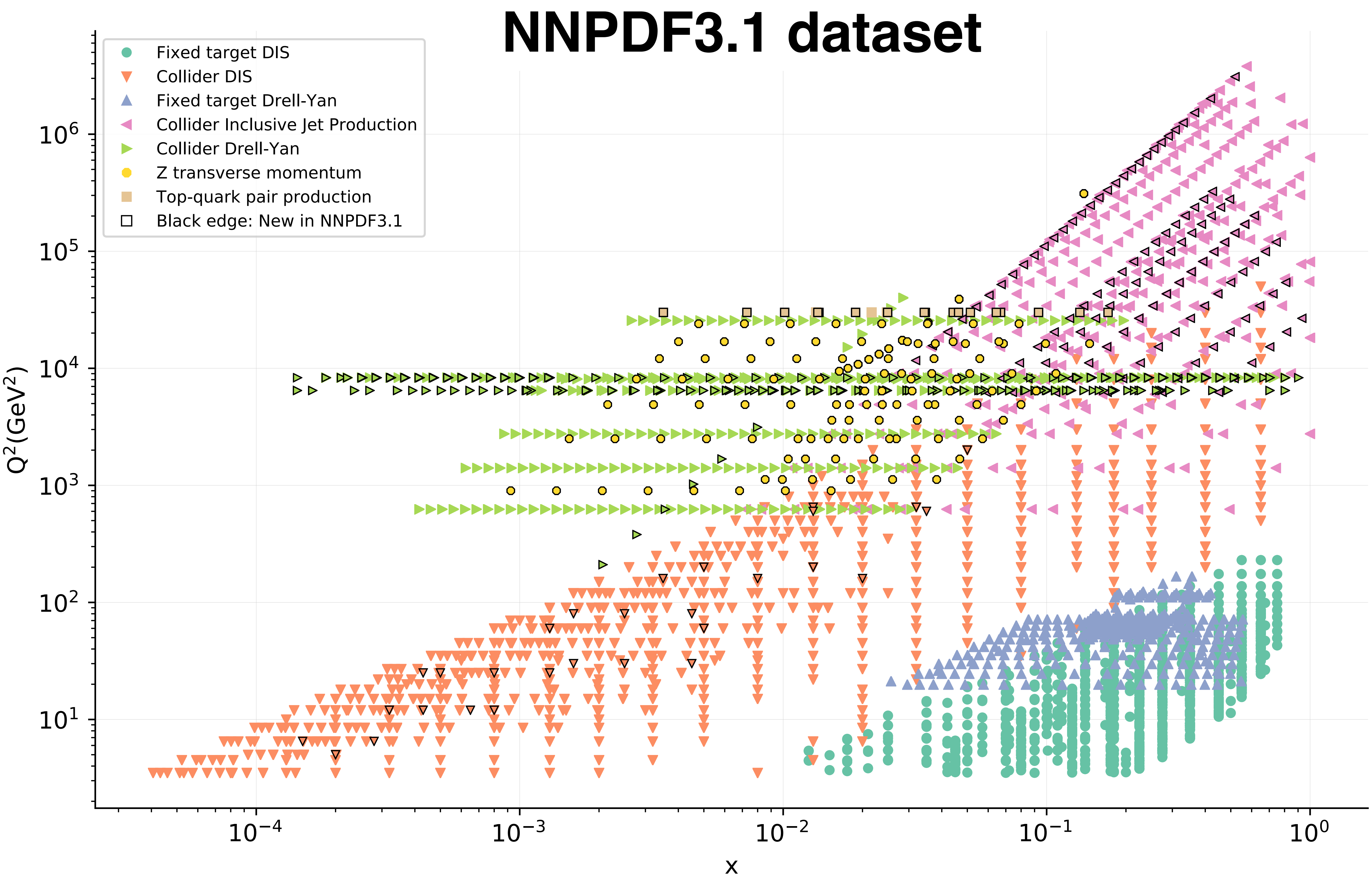
*Determine PDFs from **deep-inelastic scattering**…*

*… and use them to compute predictions for **proton-proton collisions***

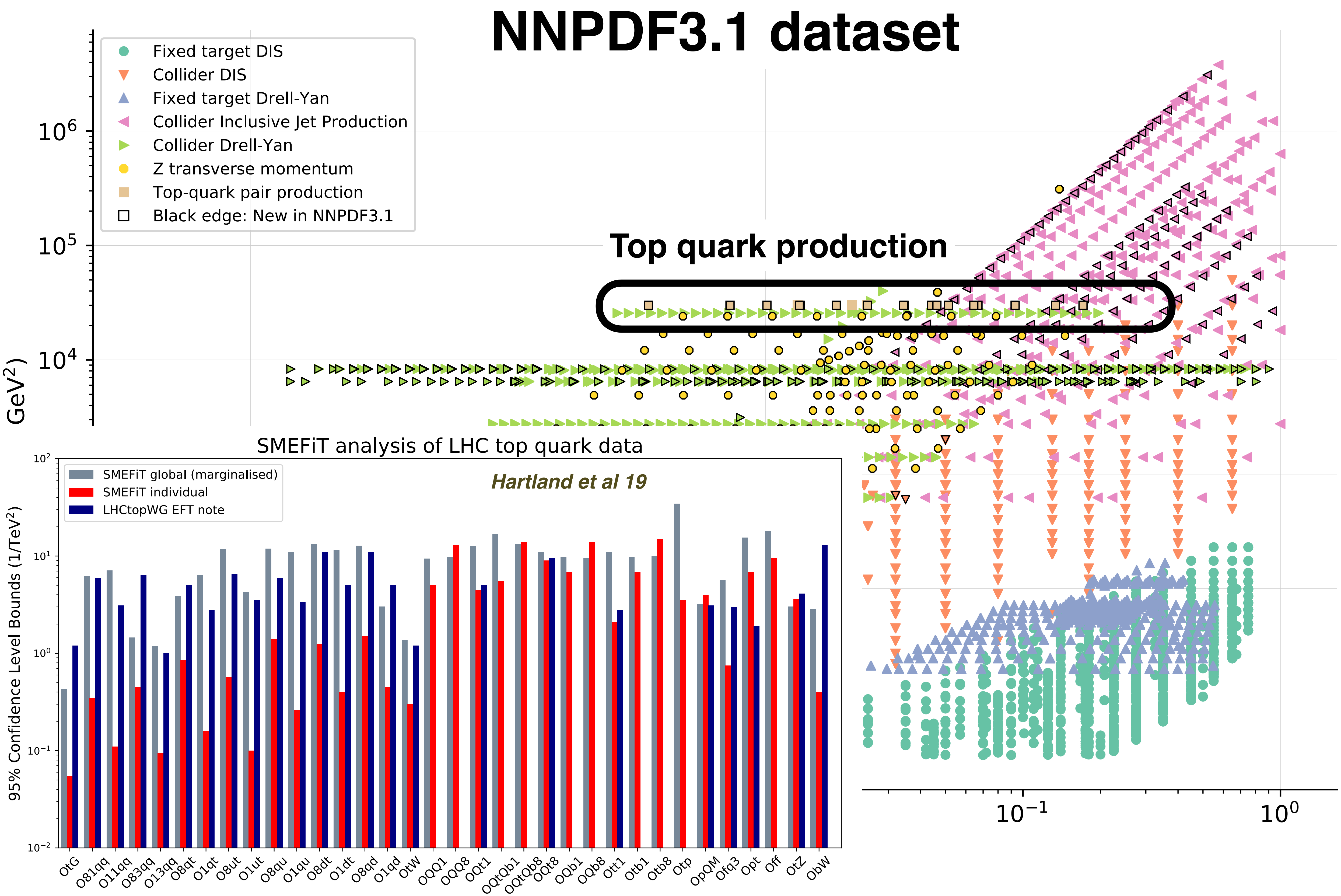
A proton structure snapshop



SMEFT & PDFs

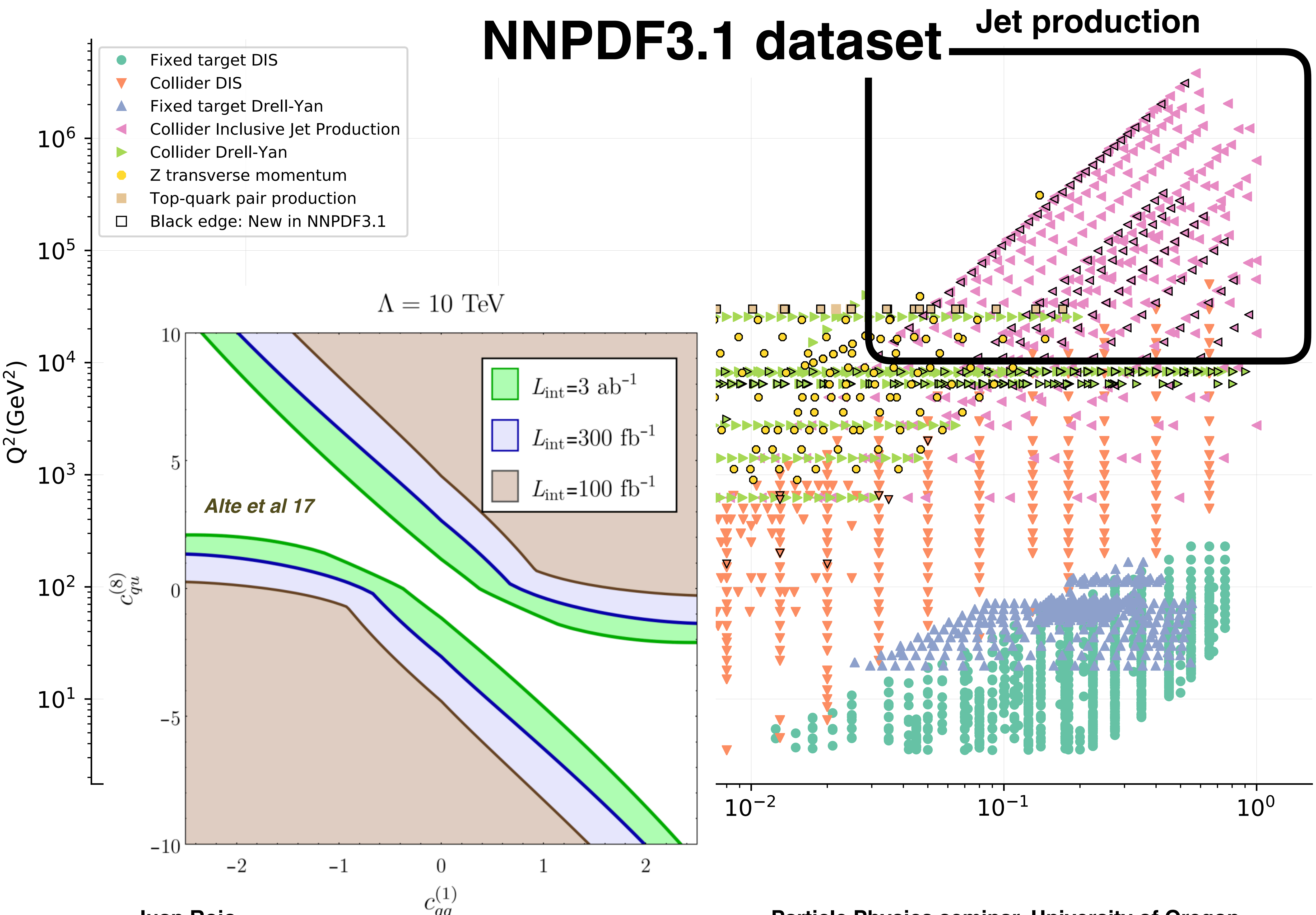

NNPDF3.1 dataset
Fixed target DIS
Collider DIS
Fixed target Drell-Yan
Collider Inclusive Jet Production
Collider Drell-Yan
Z transverse momentum
Top-quark pair production
Black edge: New in NNPDF3.1
$Q^2(\mathrm{GeV}^2)$
10^1
10^2
10^3
10^4
10^5
10^6
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^0
x

SMEFT & PDFs



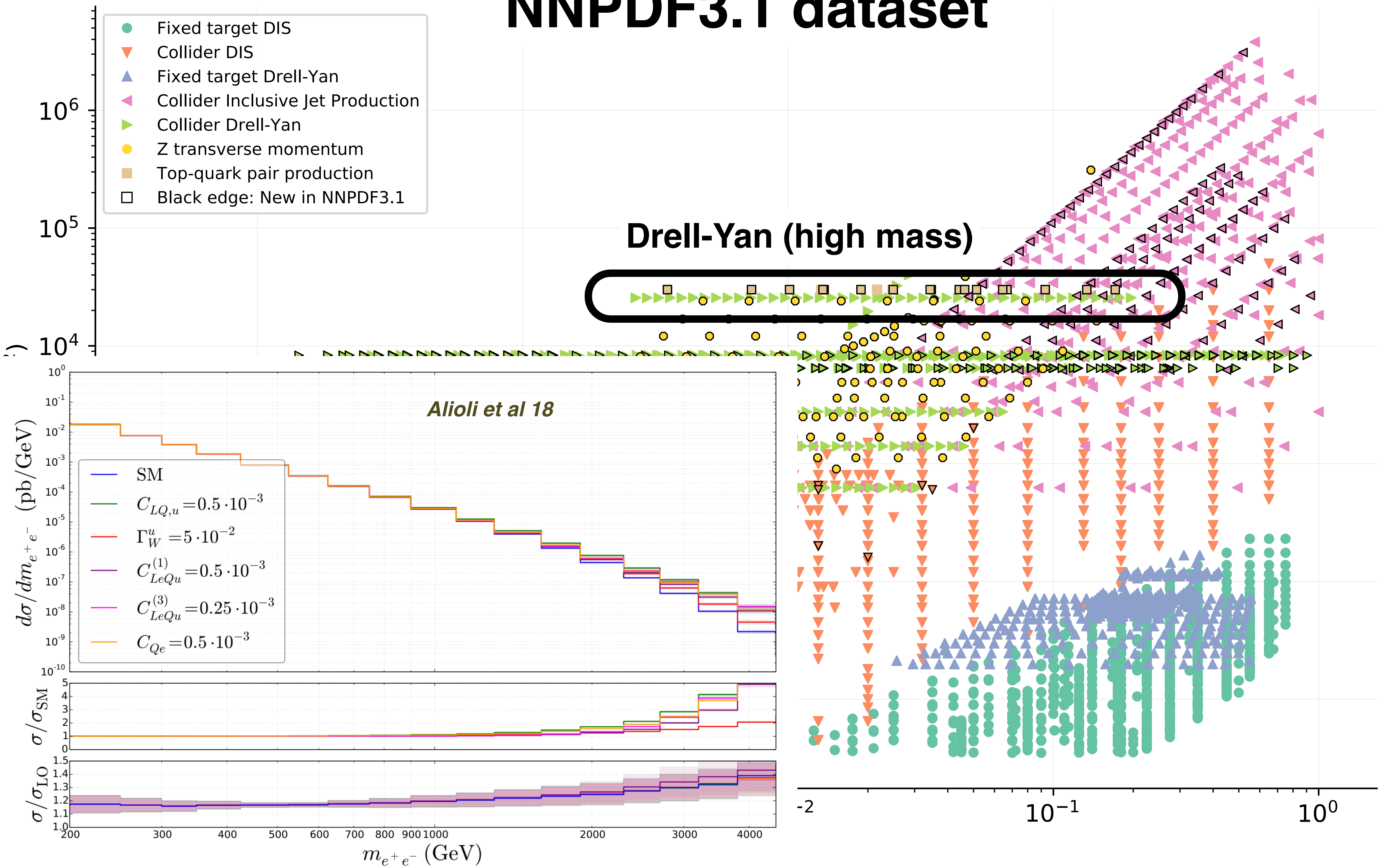
NNPDF3.1 dataset
Fixed target DIS
Collider DIS
Fixed target Drell-Yan
Collider Inclusive Jet Production
Collider Drell-Yan
Z transverse momentum
Top-quark pair production
Black edge: New in NNPDF3.1
Top quark production
10^6
10^5
10^4
GeV^2)
10^{-1}
10^0
SMEFiT analysis of LHC top quark data
SMEFiT global (marginalised)
SMEFiT individual
LHCtopWG EFT note
Hartland et al 19
95% Confidence Level Bounds (1/TeV^2)
10^2
10^1
10^0
10^{-1}
10^{-2}
OtG O81qq O11qq O83qq O13qq O8qt O1qt O8ut O1ut O8qu O1qu O8dt O1dt O8qd O1qd OtW OQQ1 OQQ8 OQt1 OQtQb1 OQtQb8 OQt8 OQb1 OQb8 Ott1 Otb1 Otb8 Otp OpQM Ofq3 Opt Off OtZ ObW

SMEFT & PDFs



NNPDF3.1 dataset
Jet production
Fixed target DIS
Collider DIS
Fixed target Drell-Yan
Collider Inclusive Jet Production
Collider Drell-Yan
Z transverse momentum
Top-quark pair production
Black edge: New in NNPDF3.1
$Q^2(\mathrm{GeV}^2)$
10^1
10^2
10^3
10^4
10^5
10^6
10^{-2}
10^{-1}
10^0
$\Lambda = 10$ TeV
$L_{\rm int}$=3 ab^{-1}
$L_{\rm int}$=300 fb^{-1}
$L_{\rm int}$=100 fb^{-1}
Alte et al 17
$c_{qu}^{(8)}$
$c_{qq}^{(1)}$
10
5
0
−5
−10
−2
−1
0
1
2

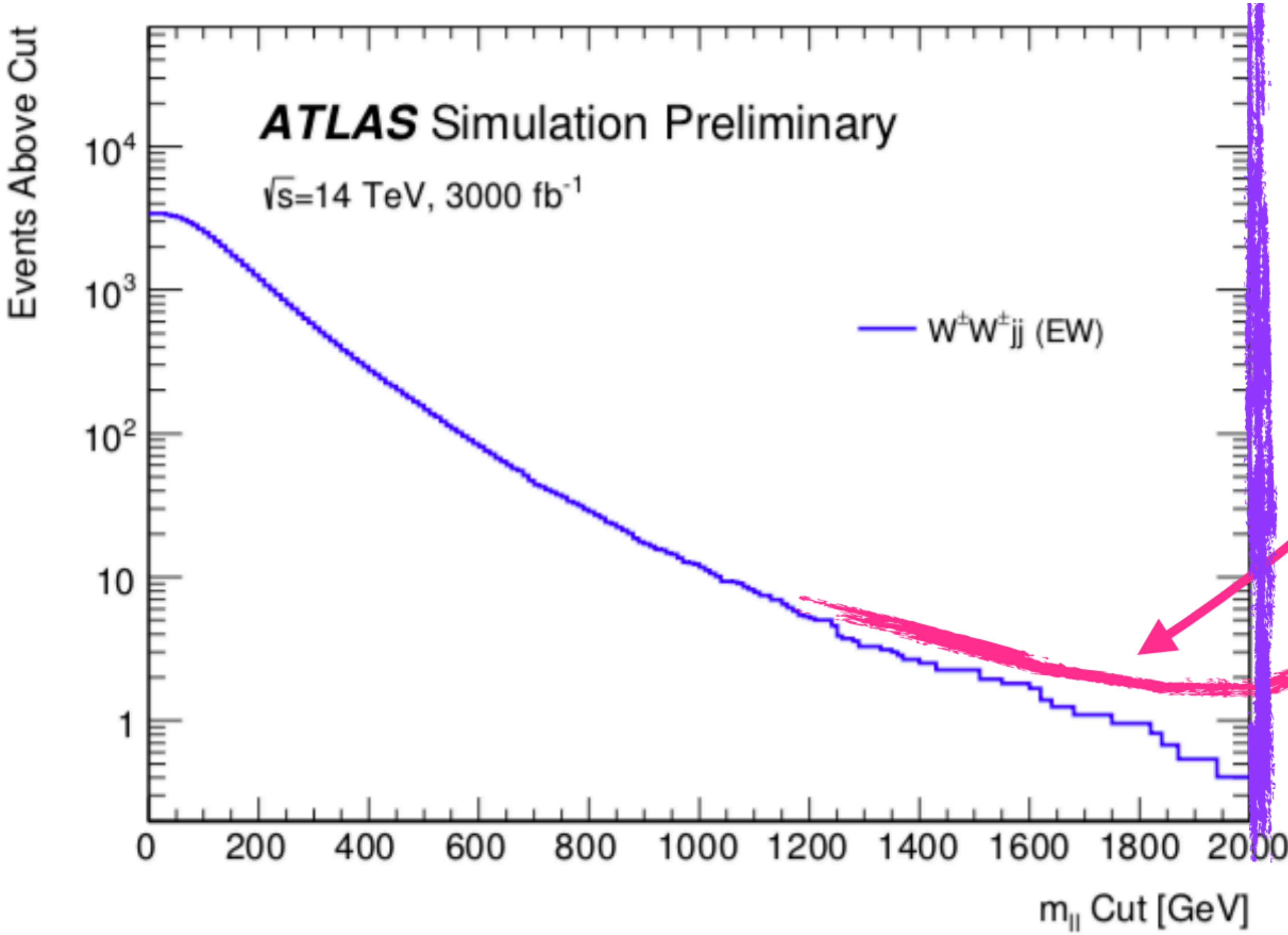
SMEFT & PDFs



NNPDF3.1 dataset
Fixed target DIS
Collider DIS
Fixed target Drell-Yan
Collider Inclusive Jet Production
Collider Drell-Yan
Z transverse momentum
Top-quark pair production
Black edge: New in NNPDF3.1
Drell-Yan (high mass)
10^6
10^5
10^4
10^{-1}
10^0
Alioli et al 18
SM
$C_{LQ,u} = 0.5 \cdot 10^{-3}$
$\Gamma_W^u = 5 \cdot 10^{-2}$
$C_{LeQu}^{(1)} = 0.5 \cdot 10^{-3}$
$C_{LeQu}^{(3)} = 0.25 \cdot 10^{-3}$
$C_{Qe} = 0.5 \cdot 10^{-3}$
$d\sigma/dm_{e^+e^-}$ (pb/GeV)
$\sigma/\sigma_{\mathrm{SM}}$
$\sigma/\sigma_{\mathrm{LO}}$
$m_{e^+e^-}$ (GeV)

Why do we need better PDFs?

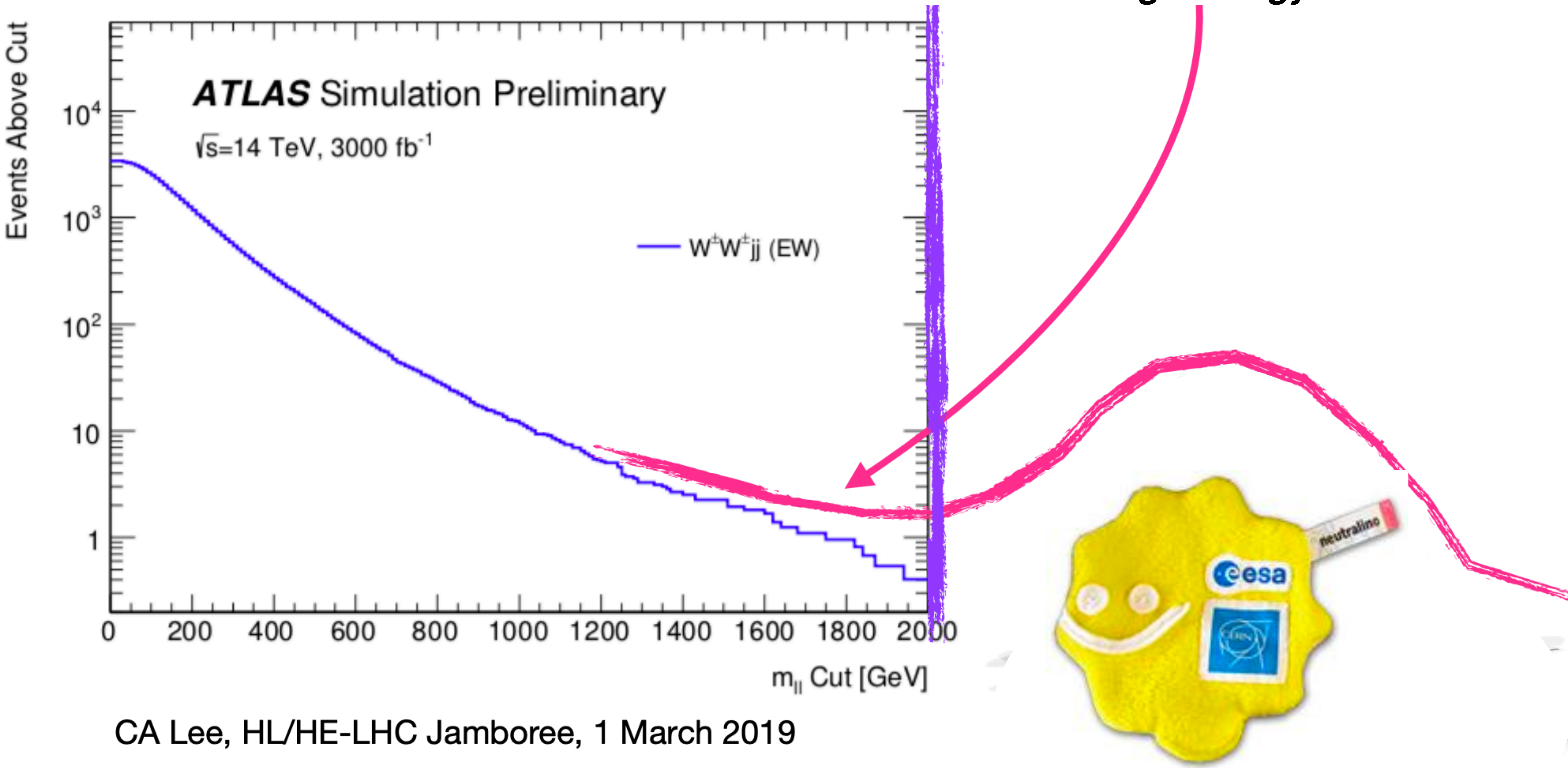
Deviation from SM prediction in high energy tails?



CA Lee, HL/HE-LHC Jamboree, 1 March 2019

Why do we need better PDFs?

Deviation from SM prediction in high energy tails?

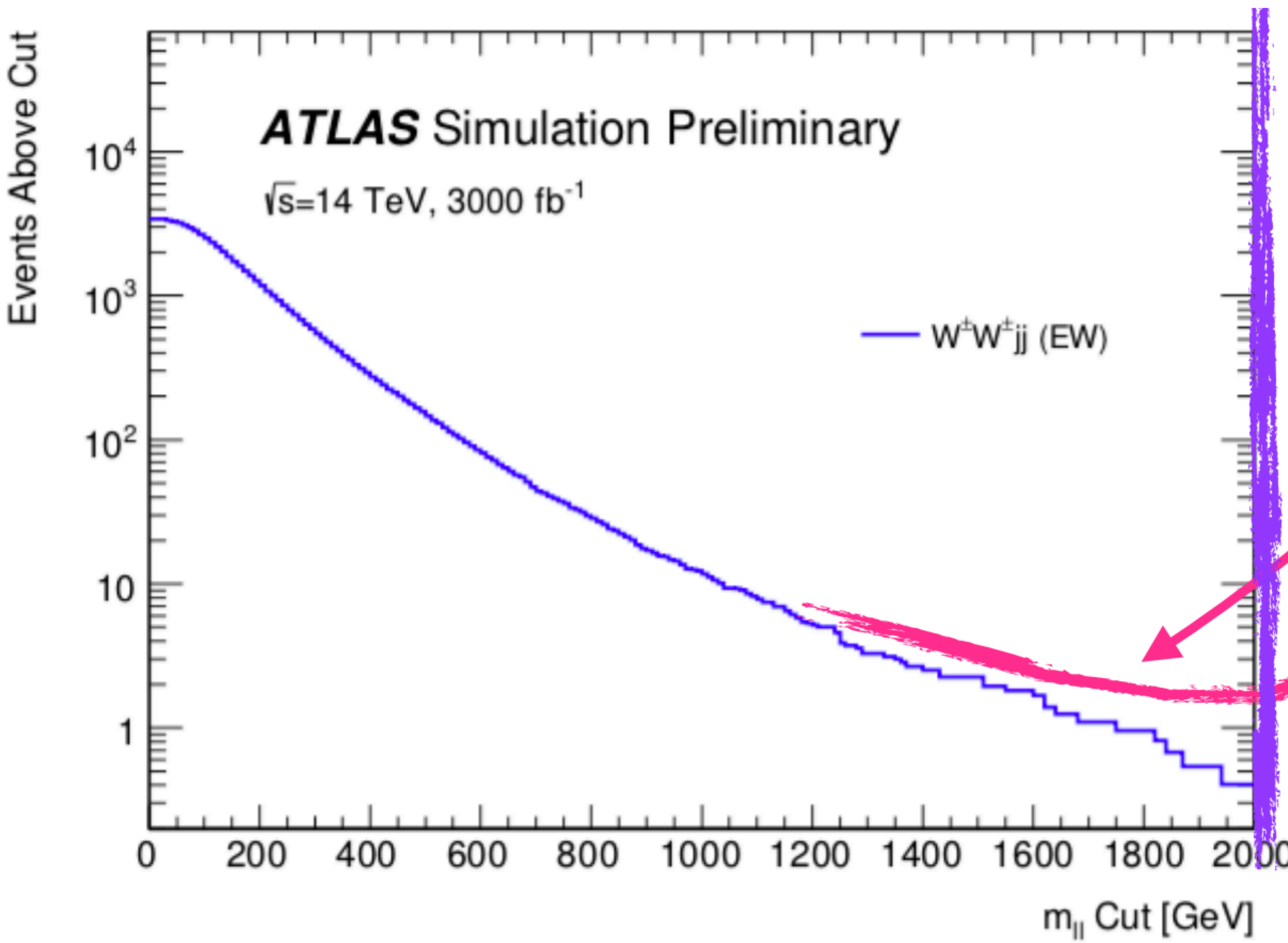


CA Lee, HL/HE-LHC Jamboree, 1 March 2019

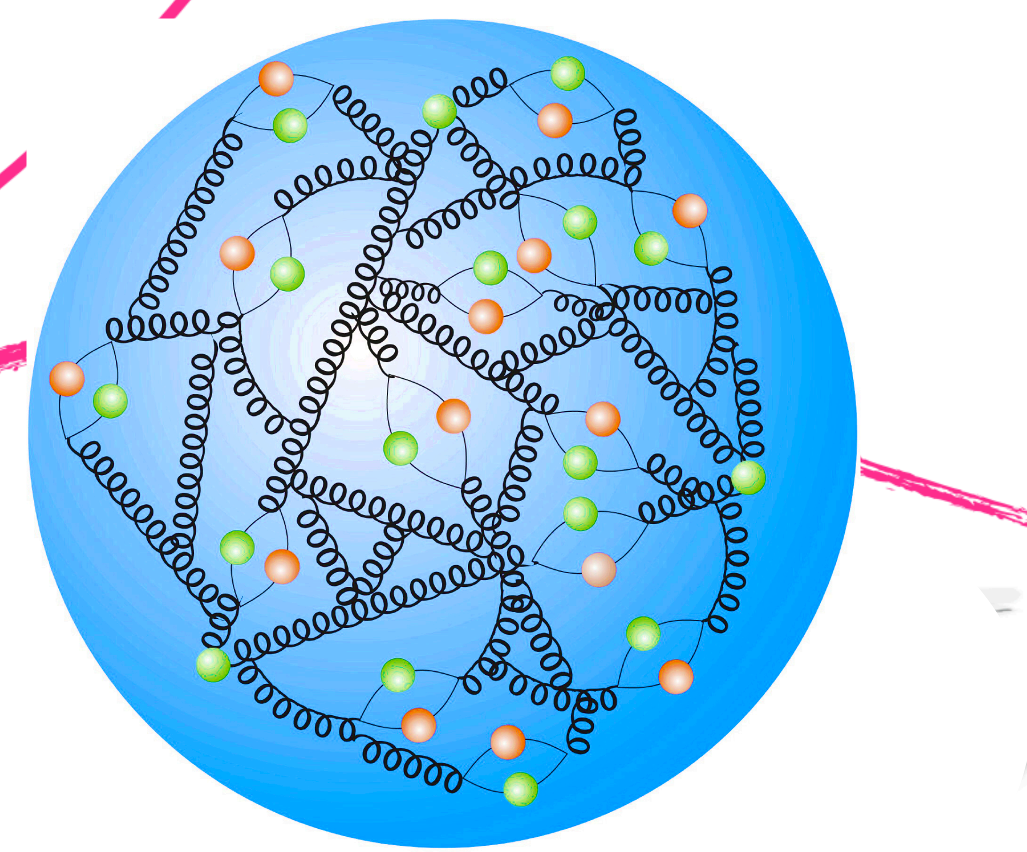
SMEFT interpretation: from a massive particle at high energies …

Why do we need better PDFs?

Deviation from SM prediction in high energy tails?



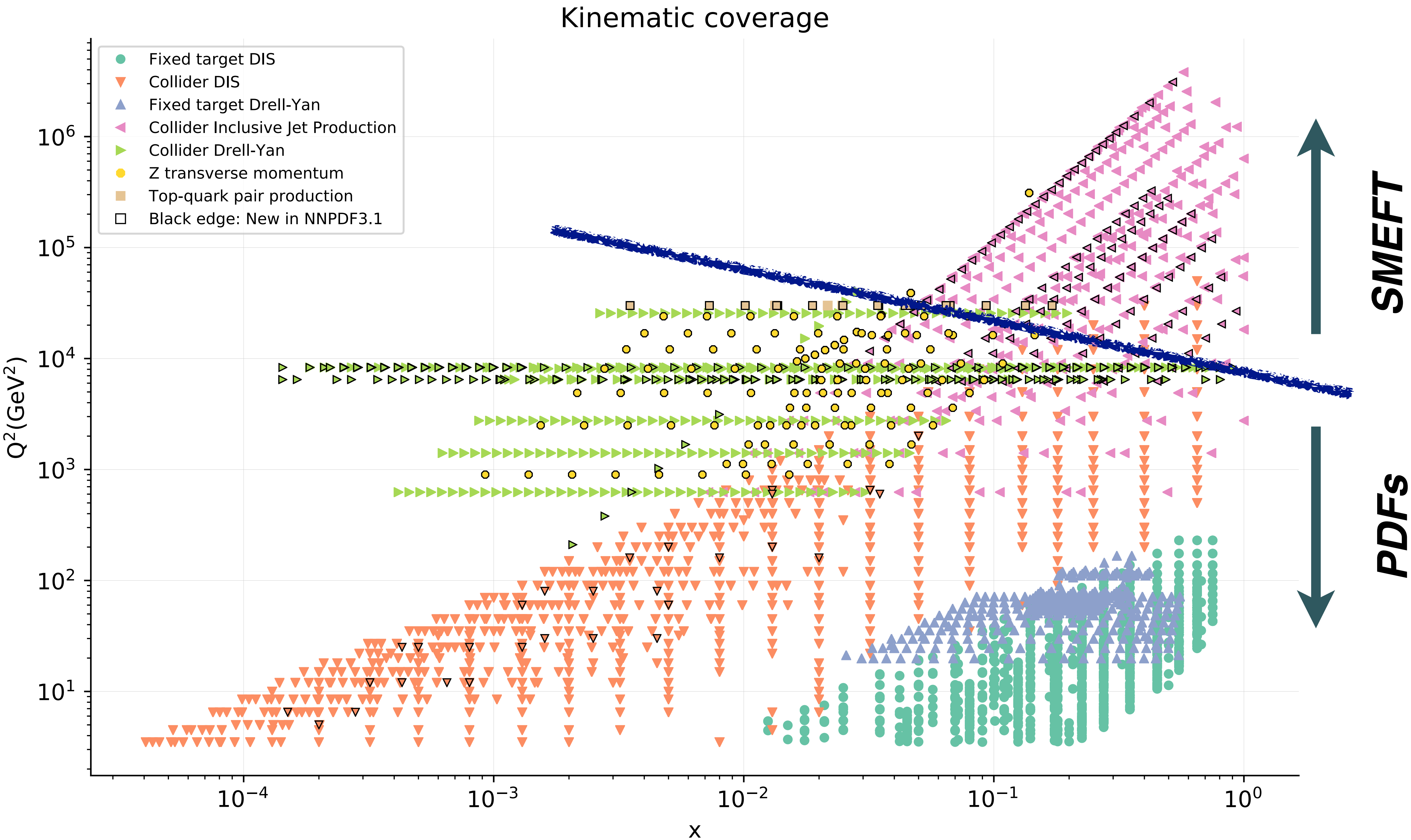
CA Lee, HL/HE-LHC Jamboree, 1 March 2019



…or reflecting our limited understating of proton structure?

Naive approach

Separate LHC data into **input for PDF fits** and **input for SMEFT studies?**



Can we do better?

Simultaneous PDF & SMEFT fits

Our goal: constrain **simultaneously** both the PDFs and SMEFT degrees of freedom

Proof of concept: DIS-only fits where SM theory is **augmented** by *d=6* SMEFT operators

$$\mathscr{L}_{\rm SMEFT} = \mathscr{L}_{\rm SM} + \sum_{q=u,d,s,c} \frac{a_q}{\Lambda^2} \left(\bar{l}_R \gamma^\mu l_R\right) \left(\bar{q}_R \gamma_\mu q_R\right)$$

which can arise *e.g.* from a ***Z' boson*** with non-universal couplings to quarks

These SMEFT operators modify the DIS structure functions and thus **affect the PDF fit**

$$\Delta F_2^{\rm SMEFT} \supset \frac{x}{12e^4}\left(4a_u e^2 \frac{Q^2}{\Lambda^2}(1+4K_Z \sin^4\theta_W) + 3a_u^2 \frac{Q^4}{\Lambda^4}\right)\left(u+\bar{u}\right)$$

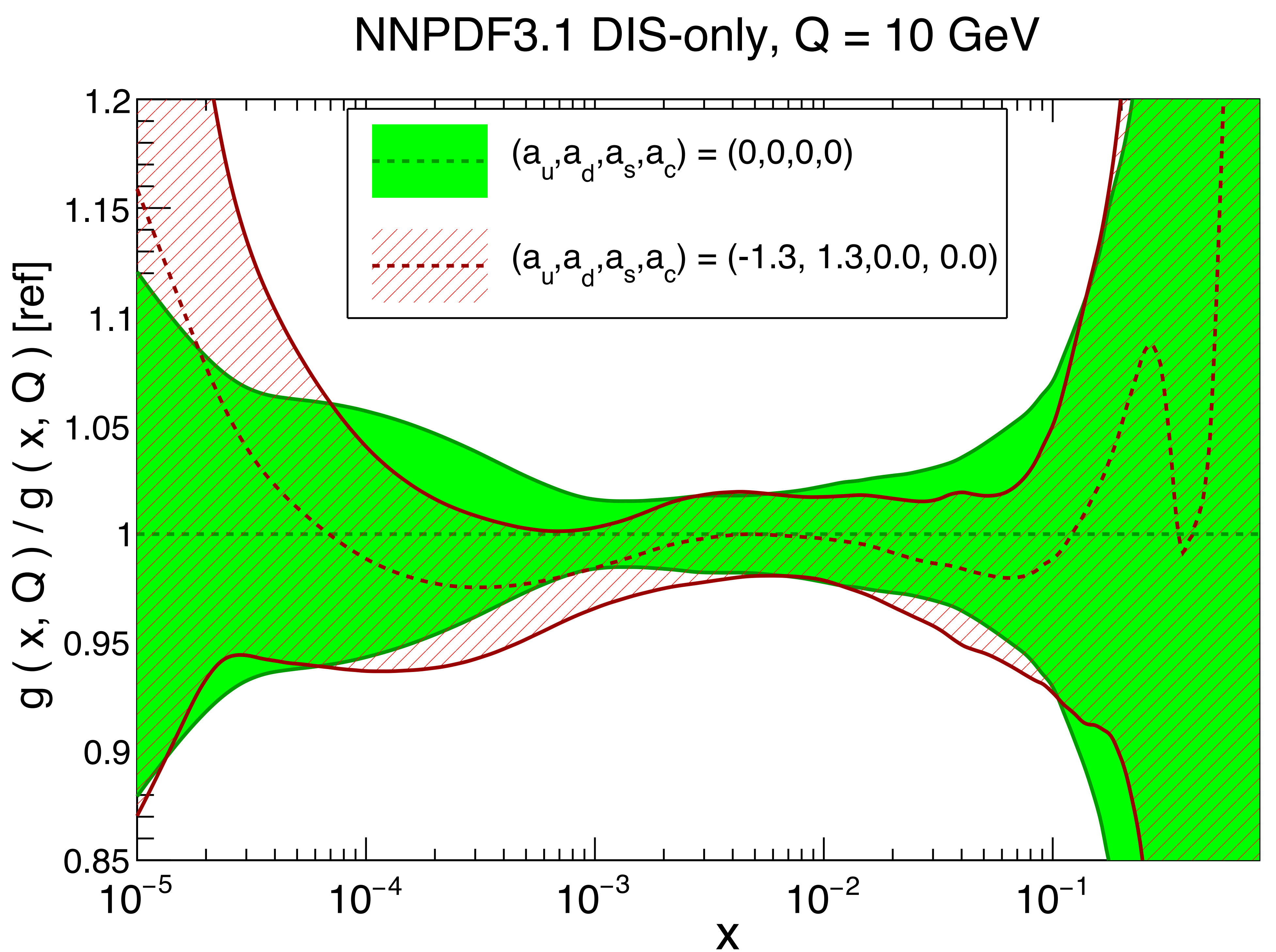
SMEFT effects enhanced by Q^2 :
constrain from HERA data

from interference with SM

from squared amplitude

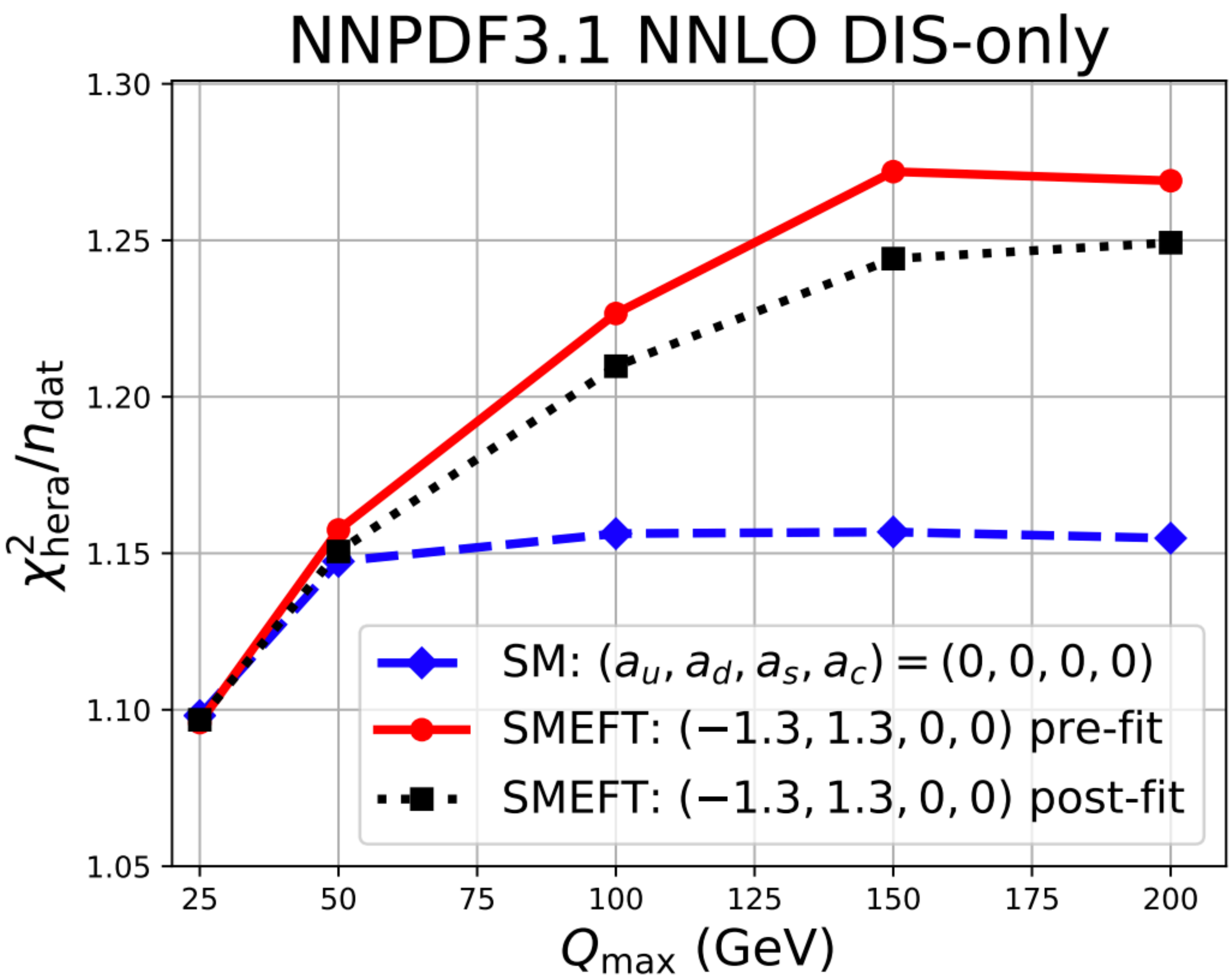
Impact on the PDFs

For a large region of the allowed parameter space,
SMEFT effects can be partially (but not completely) **reabsorbed into the PDFs**



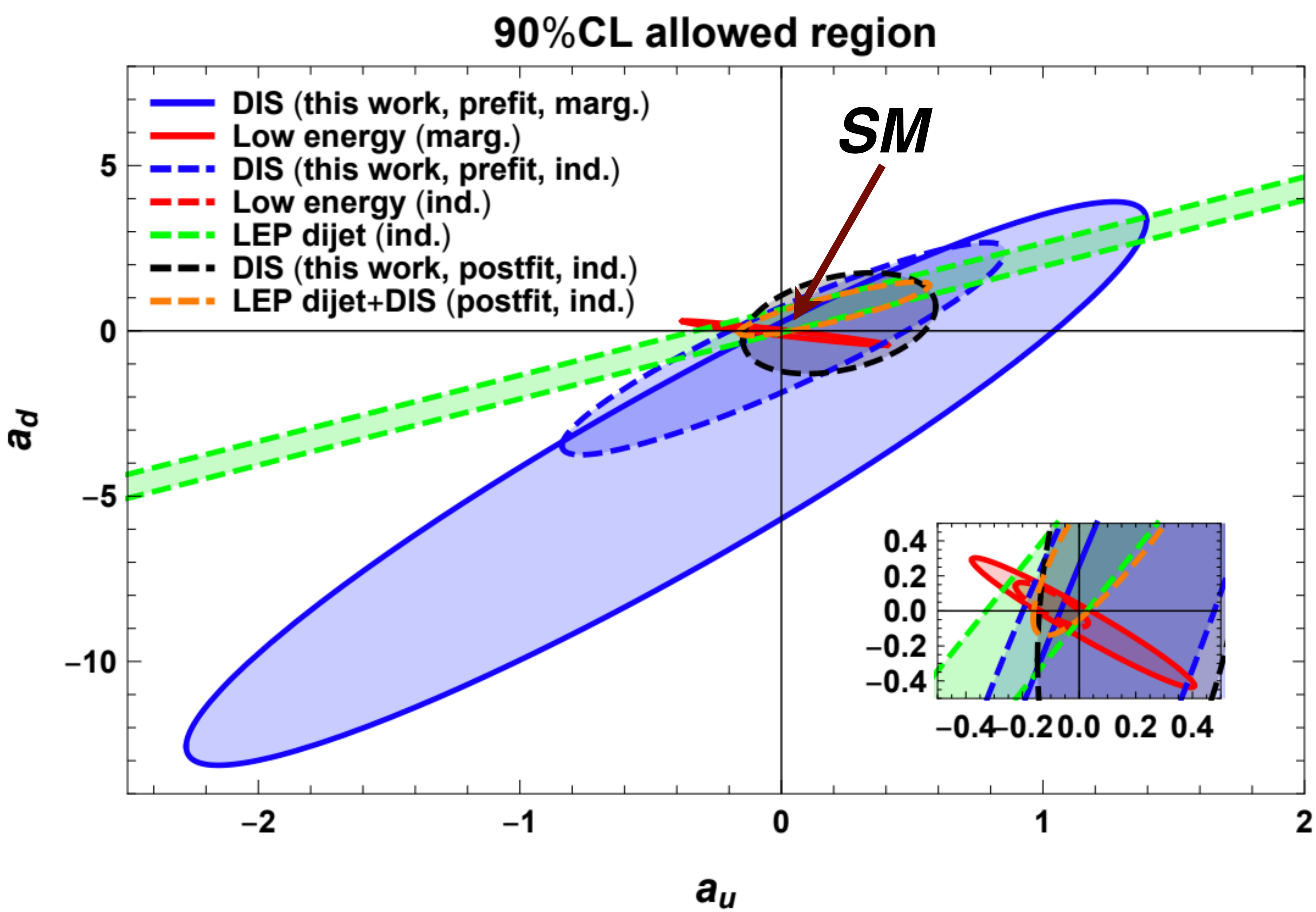
Fingerprinting BSM effects

Tell-tale sign of SMEFT effects: **rapid variation with *Q*** (DGLAP evolution slower)



Fingerprinting BSM effects

One can compare **bounds on SMEFT degrees of freedom** in the joint fit as compared to the usual approach where PDFs are kept fixed



Ultimate goal (HL-LHC timescale!): **simultaneous PDF & SMEFT global analysis**

Constraining the SMEFT with Bayesian inference

S. van Beek, E. R. Nocera, J. Rojo, E. Slade,

arXiv:1906.05296 *(submitted to SciPost)*

Bayesian reweighting

- Under many circumstances, one would like to **quantify the impact of a new measurement** in the SMEFT parameter space **without having to redo the full analysis**

- One would also like to quantify (and compare) the **amount of information** contained in current and (possible) future measurements

Bayesian Inference tells us how to update (``reweight”) the SMEFiT probability distribution with the information provided by **new measurements**

$$\omega_k \propto \left(\chi^2_k\right)^{(n_{\rm dat}-1)/2} \exp\left(-\chi^2_k/2\right) \ , \quad k=1,\ldots,N_{\rm rep}$$

weight of k-th replica

number of data points in new data

total χ^2 of new data for k-th replica

MC replicas of a prior fit

- Extensive **validation** of reweighing by comparison with **direct fits** carried out in the PDF case. What about the SMEFT parameter space?

Bayesian reweighting

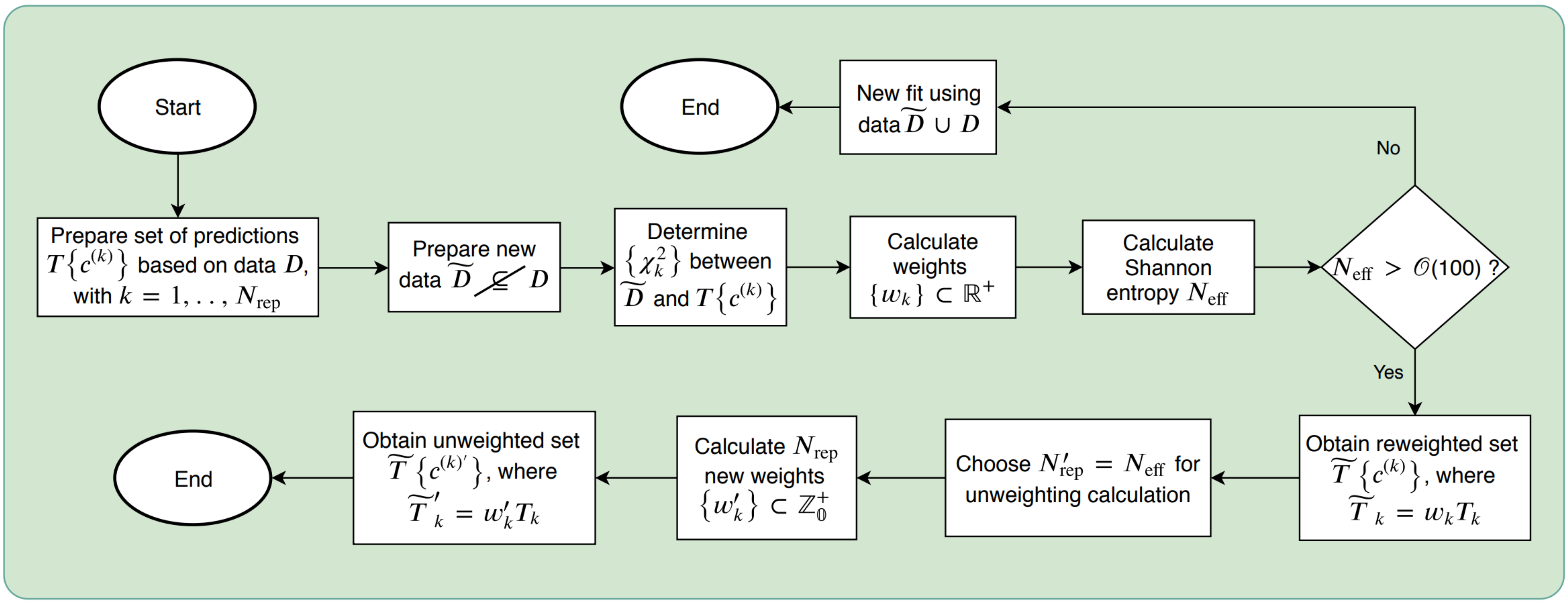
- Start from a variant of SMEFiT which excludes **LHC single top production data**
- To ensure sufficient statistics, this prior is constructed with $\boldsymbol{N_{rep}}$ **= 10000 MC replicas**
- Then add different combinations of single top data either by **reweighting** or by a **direct fit** and compare the results
- The amount of new information in each case is quantified by Shannon's entropy: the **effective number of replicas**

$$N_{\rm eff} = \exp\left(\frac{1}{N_{\rm rep}}\sum_{k=1}^{N_{\rm rep}} \omega_k \ln \frac{N_{\rm rep}}{\omega_k}\right)$$

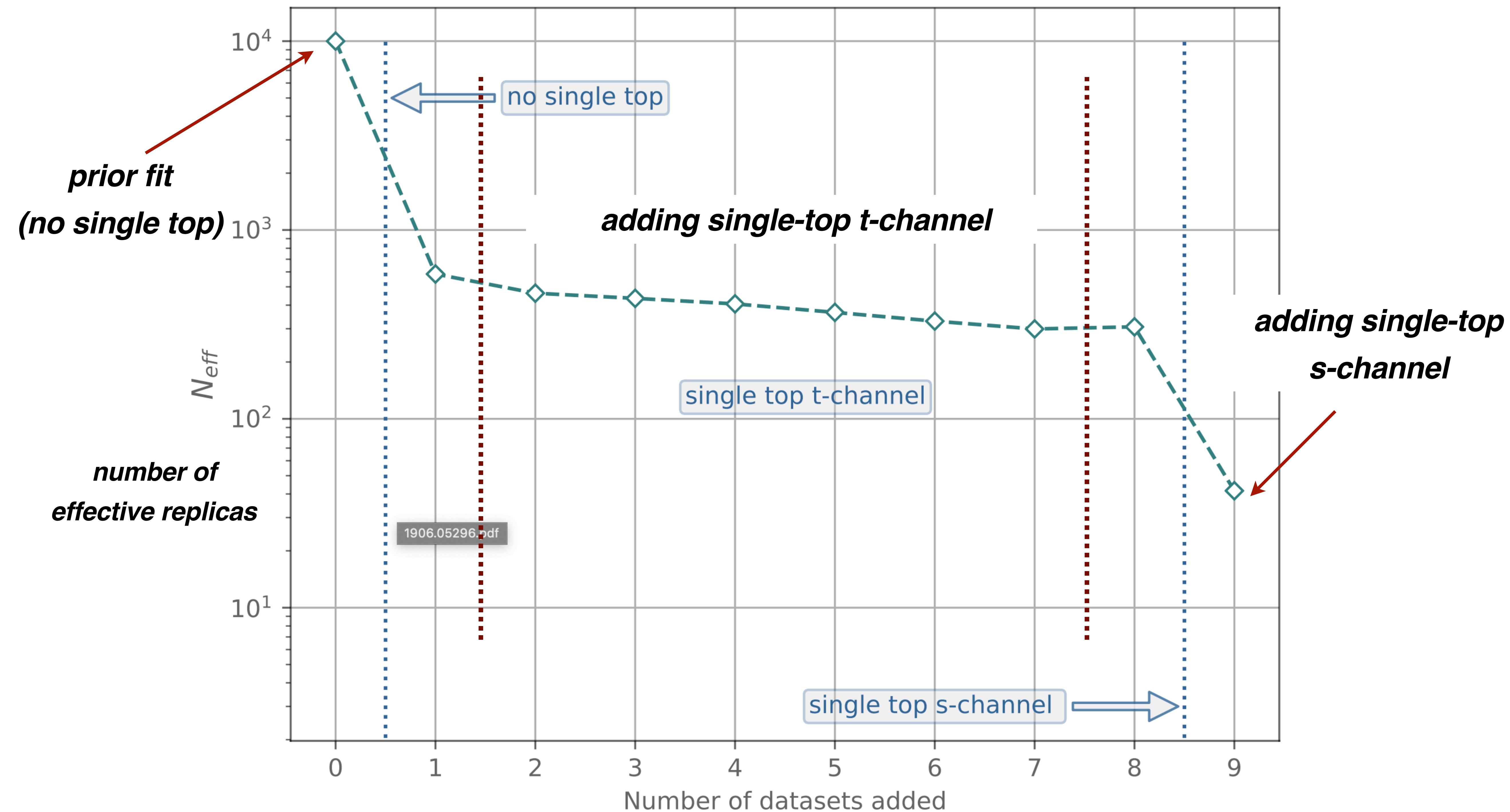
- For Bayesian reweighting to be used reliably, one requires that $\boldsymbol{N_{eff} > 50}$**,** else we run **out of statistics** and a direct refit is required

Bayesian reweighting

- To identify which SMEFT directions are more constrained by the new data, evaluate the Kolmogorov-Smirnov statistic between the **prior** and **reweighted probability distributions:** the larger the KS-statistic, the larger the effect of the new data
- Note that information can be added *(i)* due to **new direct constraints** and/or *(ii)* by **breaking degeneracies** in the parameter space



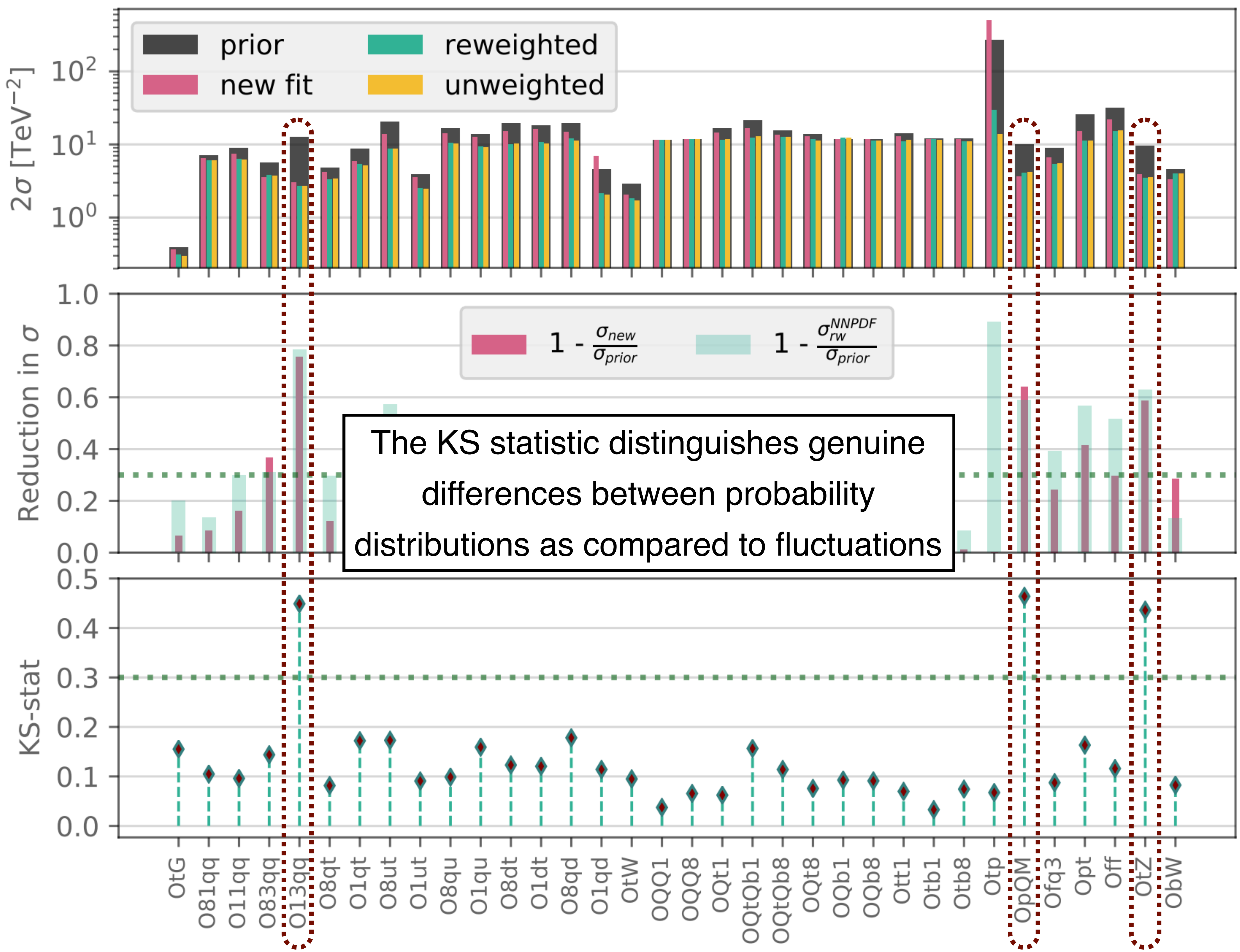
Reweighting efficiency



Significant amount of **new information** each time new process added via reweighting: marked decrease in effective number of replicas

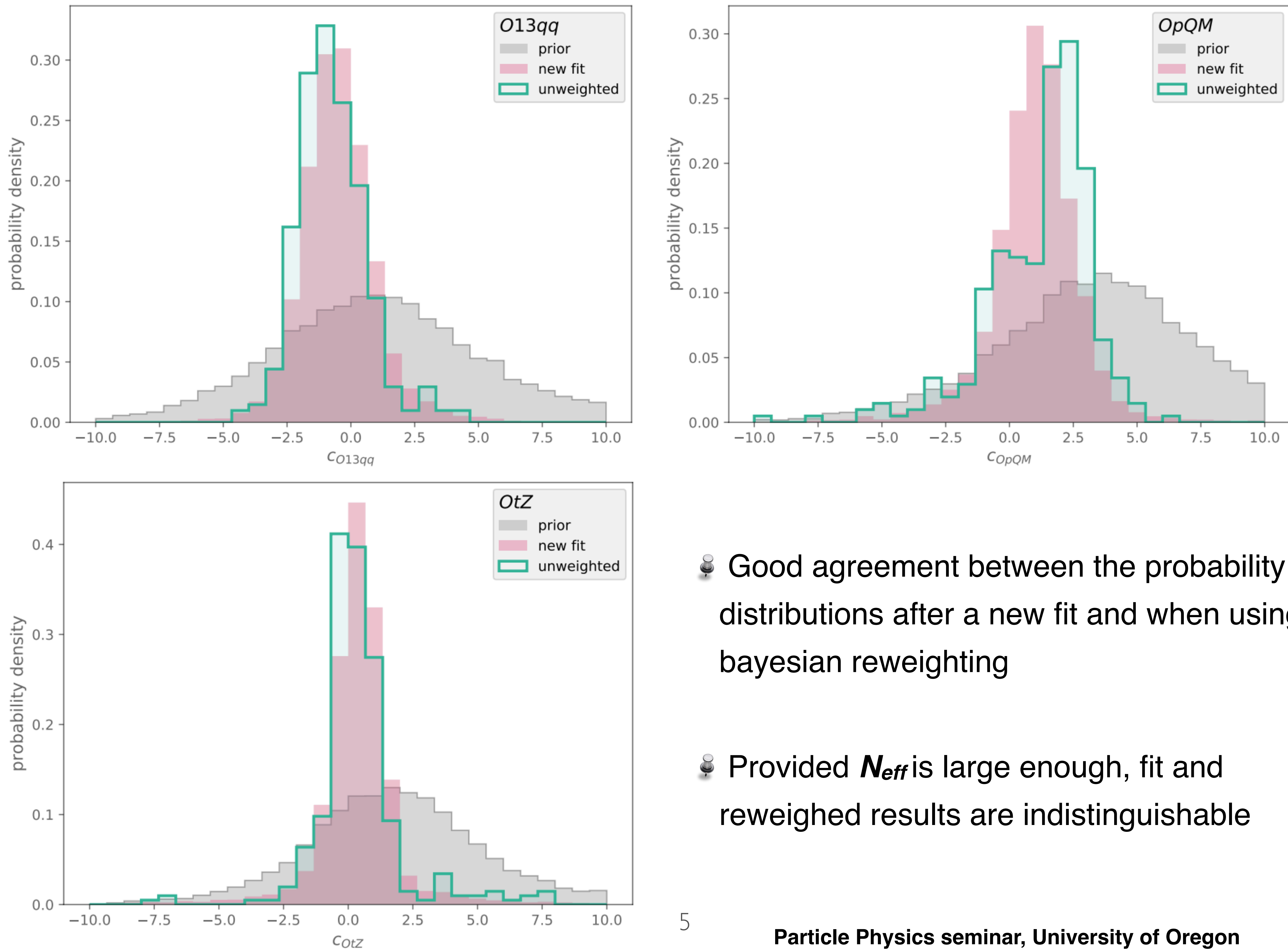
Results: adding single top t-channel

Results: adding single top t-channel



The KS statistic distinguishes genuine differences between probability distributions as compared to fluctuations

Results: adding single top t-channel



- Good agreement between the probability distributions after a new fit and when using bayesian reweighting

- Provided $\boldsymbol{N_{eff}}$ is large enough, fit and reweighed results are indistinguishable

Results: adding single top s-channel

Summary and outlook

- **SMEFiT** is a novel framework, suitable for global analyses of the SMEFT, which exploits expertise inherited from PDF fits

- As a proof-of-concept, applied this framework to the determination of the constraints in the SMEFT parameter space provided by **LHC top quark data**

- **Improved constraints** compared to previous studies (first-ever bounds in some cases)

- The **simultaneous determination of PDFs and SMEFT** degrees of freedom will be required to fully exploit the LHC potential

- Demonstrated the applicability of **Bayesian reweighting** for the *a posteriori* inclusion of the constraints from new measurements on SMEFiT without need of redoing fit

- Next steps (in progress): enlarge the operator fitting basis and include **additional LHC cross-sections** (Higgs, electroweak, jets) as well as flavour and low-energy observables, and explore implications for specific **UV-complete models**