An Effective (Field Theory) Pathway to the New Standard Model

Juan Rojo

VU Amsterdam & Theory group, Nikhef

Nederlandse Natuurkunde Vereniging

Annual Meeting of the Subatomic Physics section

Lunteren, 4th November 2022

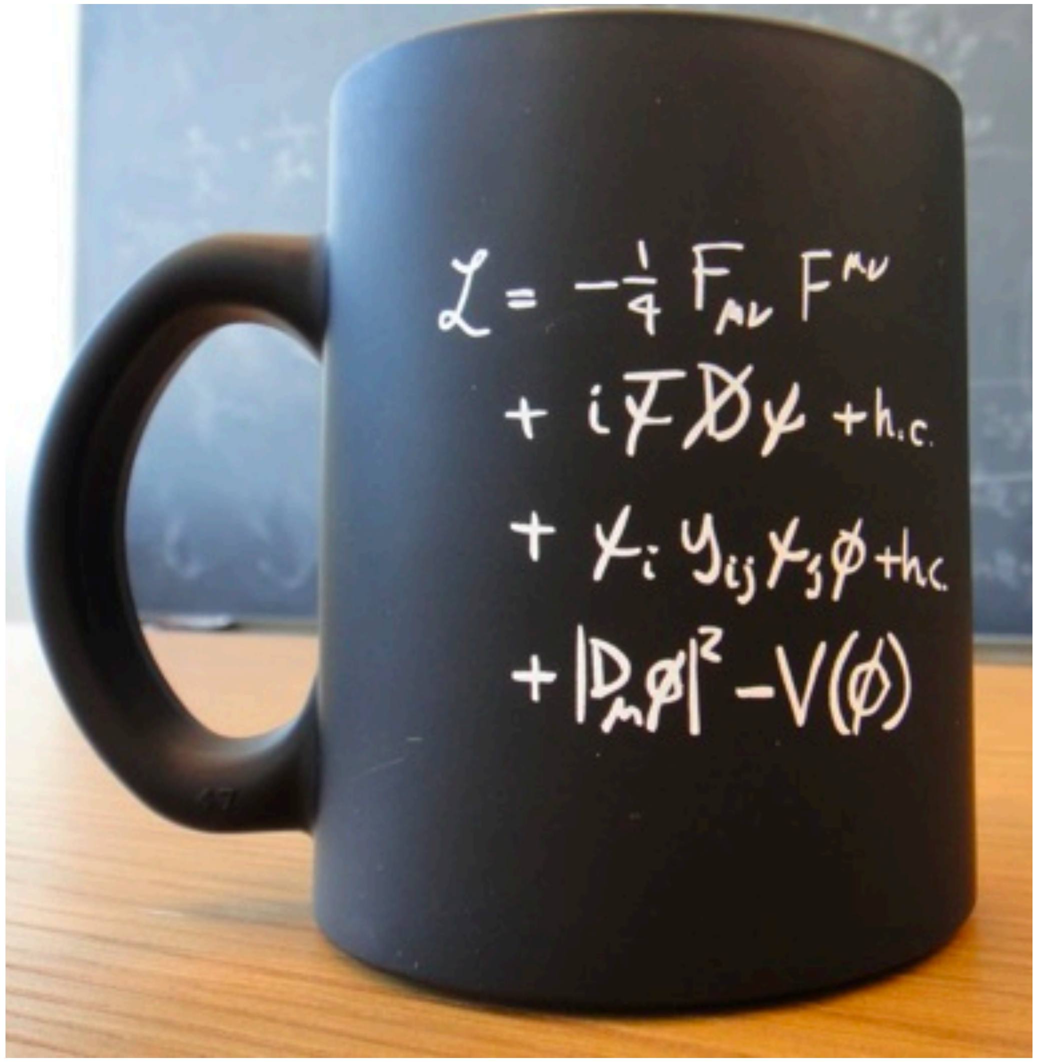
SMEFT: the new Standard Model

The Standard Model

The Standard Model is defined by:

- **Particle (matter) content**: quarks and leptons
- **Gauge** (local) symmetries and their eventual breaking mechanisms
- **Lorentz** invariance and other global symmetries
- Linearly realised $SU(2)_L$ EW symmetry breaking
- Validity up to **very high scales (renormalisability)**

e.g. Planck scale



$$\mathscr{L}_{\mathrm{SM}} = \sum_i c_i \mathcal{O}_i^{(d=4)}$$

extremely predictive framework!

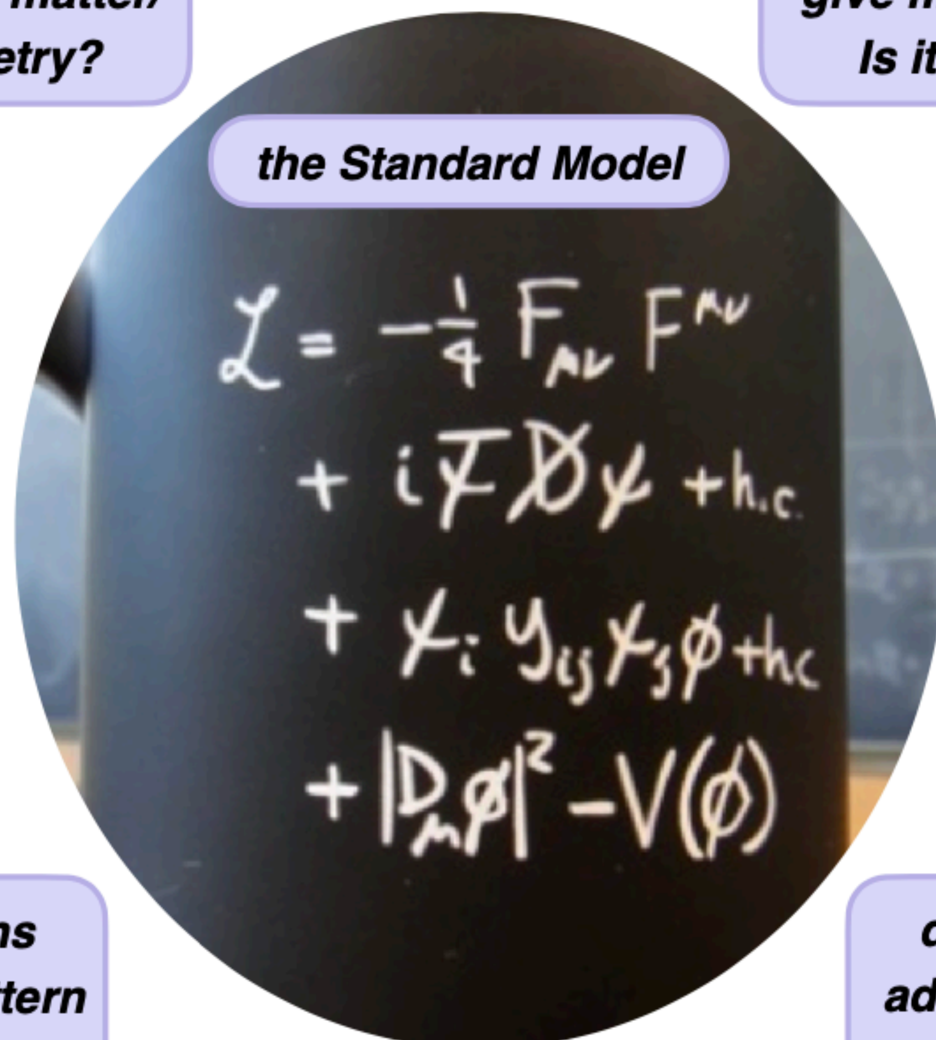
dimensionless couplings (before EW symmetry breaking)

All possible operators of **mass-dimension <=4** consistent with above requirements

The Standard Model is not the whole story

why does our Universe exhibit such a strong matter/antimatter asymmetry?

why does the Higgs mechanism give mass to elementary particles? Is it effective or fundamental?



what is the correct quantum mechanical description of gravity?

what sets the scale of neutrino masses? Do sterile neutrinos exist?

why do quarks and leptons exhibit such a disparate pattern of masses and couplings?

does Dark Matter admit an elementary particle description?

Violation of lepton flavour universality?

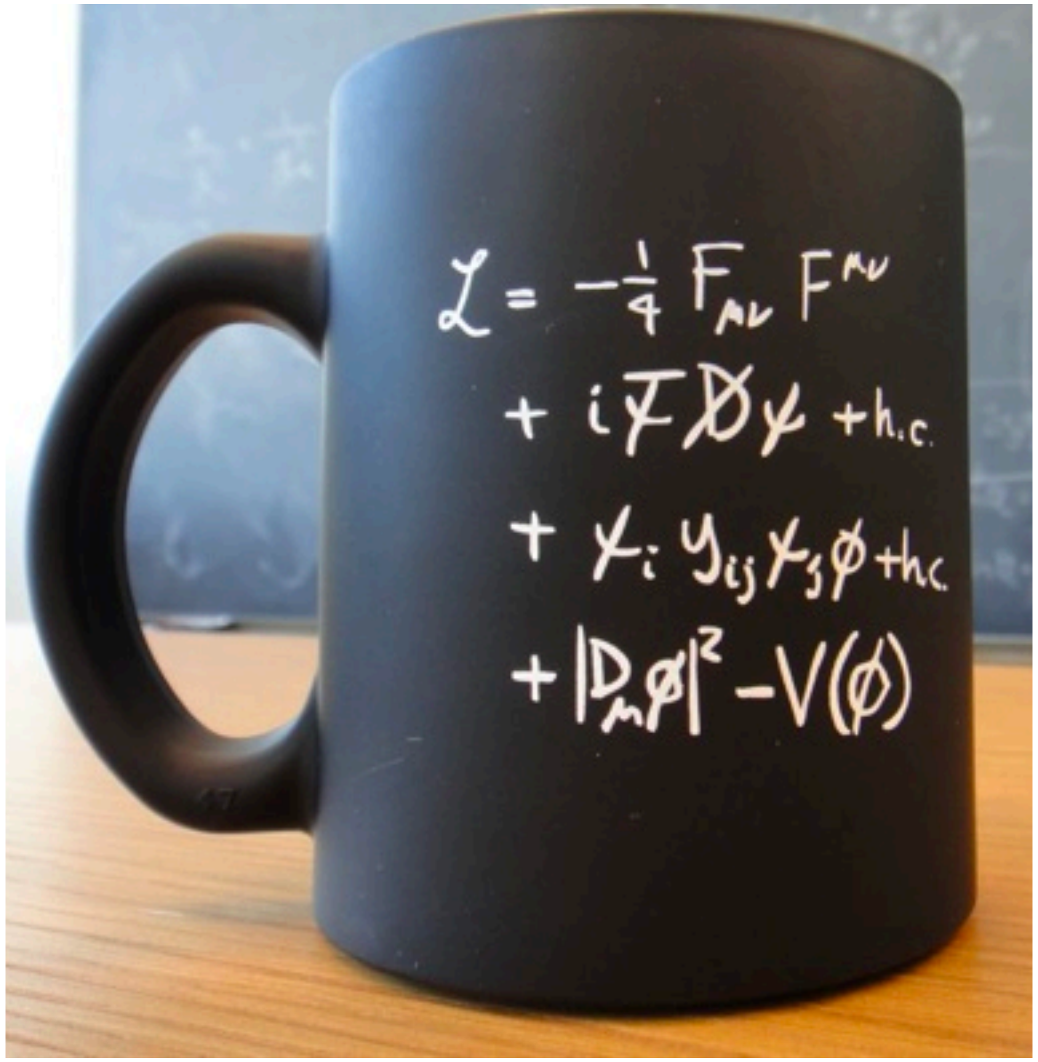
Innumerable extensions of the SM have been proposed. None of them has been validated

maybe rethink how to search for New Physics?

The Standard Model as an Effective Theory

The Standard Model EFT is defined by:

- **Particle (matter) content**: quarks and leptons
- **Gauge** (local) symmetries and their eventual breaking mechanisms
- **Lorentz** invariance and other global symmetries
- Linearly realised $SU(2)_L$ EW symmetry breaking
- ***Validity only up to certain energy scale Λ***



$$\mathcal{L}_{\text{SMEFT}}(\{c_i\},\Lambda) = \mathcal{L}_{\text{SM}} + \sum_{d=5}^{\infty}\sum_{i=1}^{N_d} c_i^{(d)} \frac{\mathcal{O}_i^{(d)}}{\Lambda^{d-4}}$$

EFT coupling constants, to be determined from **data**

All possible operators of **mass-dimension *d*** consistent with above requirements

Why the SMEFT?

$$\mathcal{L}_{\text{SMEFT}}(\{c_i\},\Lambda) = \mathcal{L}_{\text{SM}} + \sum_{d=5}^{\infty}\sum_{i=1}^{N_d} c_i^{(d)} \frac{\mathcal{O}_i^{(d)}}{\Lambda^{d-4}}$$

- ☑ **Low-energy limit** of generic UV-complete theories (with linearly realized EWSB)
- ☑ **Complete basis** at any given mass-dimension: systematic parametrisation of BSM effects
- ☑ **Fully renormalizable**, full-fledged QFT: compute higher orders in QCD and EW
- ☑ Matched to a large number of **BSM models** that reduce to the SM at low energies: exploits the full power of **SM measurements** for model-independent BSM searches
- ☑ Some operators induce **growth with the partonic centre-of-mass energy**: increased sensitivity in LHC cross-sections in the TeV region

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

The Standard Model EFT

- 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**

pure bosonic

X^3		$X^2\varphi^2$	
Q_G	$f^{ABC} G_\mu^{A\nu} G_\nu^{B\rho} G_\rho^{C\mu}$	$Q_{\varphi G}$	$\varphi^\dagger \varphi G^A_{\mu\nu} G^{A\mu\nu}$
$Q_{\widetilde{G}}$	$f^{ABC} \widetilde{G}_\mu^{A\nu} G_\nu^{B\rho} G_\rho^{C\mu}$	$Q_{\varphi B}$	$\varphi^\dagger \varphi B_{\mu\nu} B^{\mu\nu}$
Q_W	$\varepsilon^{IJK} W_\mu^{I\nu} W_\nu^{J\rho} W_\rho^{K\mu}$	$Q_{\varphi W}$	$\varphi^\dagger \varphi W^I_{\mu\nu} W^{I\mu\nu}$
$Q_{\widetilde{W}}$	$\varepsilon^{IJK} \widetilde{W}_\mu^{I\nu} W_\nu^{J\rho} W_\rho^{K\mu}$	$Q_{\varphi WB}$	$\varphi^\dagger \tau^I \varphi W^I_{\mu\nu} B^{\mu\nu}$
φ^6		$Q_{\varphi \widetilde{G}}$	$\varphi^\dagger \varphi \widetilde{G}^A_{\mu\nu} G^{A\mu\nu}$
Q_φ	$(\varphi^\dagger \varphi)^3$	$Q_{\varphi \widetilde{B}}$	$\varphi^\dagger \varphi \widetilde{B}_{\mu\nu} B^{\mu\nu}$
$\varphi^4 D^2$		$Q_{\varphi \widetilde{W}}$	$\varphi^\dagger \varphi \widetilde{W}^I_{\mu\nu} W^{I\mu\nu}$
$Q_{\varphi\Box}$	$(\varphi^\dagger \varphi)\Box(\varphi^\dagger \varphi)$	$Q_{\varphi \widetilde{W} B}$	$\varphi^\dagger \tau^I \varphi \widetilde{W}^I_{\mu\nu} B^{\mu\nu}$
$Q_{\varphi D}$	$\left(\varphi^\dagger D^\mu \varphi\right)^* \left(\varphi^\dagger D_\mu \varphi\right)$		

four-fermion operators

$(\bar{L}L)(\bar{L}L)$		$(\bar{L}L)(\bar{R}R)$	
$Q_{\ell\ell}$	$(\bar{\ell}\gamma_\mu \ell)(\bar{\ell}\gamma^\mu \ell)$	$Q_{\ell e}$	$(\bar{\ell}\gamma_\mu \ell)(\bar{e}\gamma^\mu e)$
$Q_{qq}^{(1)}$	$(\bar{q}\gamma_\mu q)(\bar{q}\gamma^\mu q)$	$Q_{\ell u}$	$(\bar{\ell}\gamma_\mu \ell)(\bar{u}\gamma^\mu u)$
$Q_{qq}^{(3)}$	$(\bar{q}\gamma_\mu \tau^I q)(\bar{q}\gamma^\mu \tau^I q)$	$Q_{\ell d}$	$(\bar{\ell}\gamma_\mu \ell)(\bar{d}\gamma^\mu d)$
$Q_{\ell q}^{(1)}$	$(\bar{\ell}\gamma_\mu \ell)(\bar{q}\gamma^\mu q)$	Q_{qe}	$(\bar{q}\gamma_\mu q)(\bar{e}\gamma^\mu e)$
$Q_{\ell q}^{(3)}$	$(\bar{\ell}\gamma_\mu \tau^I \ell)(\bar{q}\gamma^\mu \tau^I q)$	$Q_{qu}^{(1)}$	$(\bar{q}\gamma_\mu q)(\bar{u}\gamma^\mu u)$

bosonic-fermionic

$\psi^2\varphi^3$		$\psi^2\varphi^2 D$	
$Q_{u\varphi}$	$(\varphi^\dagger\varphi)(\bar{q}u\widetilde{\varphi})$	$Q_{\varphi\ell}^{(1)}$	$\left(\varphi^\dagger i\overleftrightarrow{D}_\mu \varphi\right)(\bar{\ell}\gamma^\mu \ell)$
$Q_{d\varphi}$	$(\varphi^\dagger\varphi)(\bar{q}d\varphi)$	$Q_{\varphi\ell}^{(3)}$	$\left(\varphi^\dagger i\overleftrightarrow{D}^I_\mu \varphi\right)(\bar{\ell}\tau^I\gamma^\mu \ell)$
$Q_{e\varphi}$	$(\varphi^\dagger\varphi)(\bar{\ell}e\varphi)$	$Q_{\varphi e}$	$\left(\varphi^\dagger i\overleftrightarrow{D}_\mu \varphi\right)(\bar{e}\gamma^\mu e)$
$\psi^2 X\varphi$		$Q_{\varphi q}^{(1)}$	$\left(\varphi^\dagger i\overleftrightarrow{D}_\mu \varphi\right)(\bar{q}\gamma^\mu q)$
Q_{eW}	$(\bar{\ell}\sigma^{\mu\nu}e)\tau^I\varphi W^I_{\mu\nu}$	$Q_{\varphi q}^{(3)}$	$\left(\varphi^\dagger i\overleftrightarrow{D}^I_\mu \varphi\right)(\bar{q}\tau^I\gamma^\mu q)$
Q_{eB}	$(\bar{\ell}\sigma^{\mu\nu}e)\varphi B_{\mu\nu}$	$Q_{\varphi u}$	$\left(\varphi^\dagger i\overleftrightarrow{D}_\mu \varphi\right)(\bar{u}\gamma^\mu u)$
Q_{uG}	$(\bar{q}\sigma^{\mu\nu}T^A u)\widetilde{\varphi}G^A_{\mu\nu}$	$Q_{\varphi d}$	$\left(\varphi^\dagger i\overleftrightarrow{D}_\mu \varphi\right)(\bar{d}\gamma^\mu d)$
Q_{uW}	$(\bar{q}\sigma^{\mu\nu}u)\tau^I\widetilde{\varphi}W^I_{\mu\nu}$	$Q_{\varphi ud}$	$(\widetilde{\varphi}^\dagger iD_\mu \varphi)(\bar{u}\gamma^\mu d)$

The Standard Model EFT

- 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**

X^3		$X^2\varphi^2$	
Q_G	$f^{ABC}G_\mu^{A\nu}G_\nu^{B\rho}G_\rho^{C\mu}$	$Q_{\varphi G}$	$\varphi^\dagger\varphi G^A_{\mu\nu}G^{A\mu\nu}$
$Q_{\widetilde{G}}$	$f^{ABC}\widetilde{G}_\mu^{A\nu}G_\nu^{B\rho}G_\rho^{C\mu}$	$Q_{\varphi B}$	$\varphi^\dagger\varphi B_{\mu\nu}B^{\mu\nu}$
Q_W	$\varepsilon^{IJK}W_\mu^{I\nu}W_\nu^{J\rho}W_\rho^{K\mu}$	$Q_{\varphi W}$	$\varphi^\dagger\varphi W^I_{\mu\nu}W^{I\mu\nu}$
$Q_{\widetilde{W}}$	$\varepsilon^{IJK}\widetilde{W}_\mu^{I\nu}W_\nu^{J\rho}W_\rho^{K\mu}$	$Q_{\varphi WB}$	$\varphi^\dagger\tau^I\varphi W^I_{\mu\nu}B^{\mu\nu}$
φ^6		$Q_{\varphi\widetilde{G}}$	$\varphi^\dagger\varphi\widetilde{G}^A_{\mu\nu}G^{A\mu\nu}$

pure bosonic

four-fermion operators

bosonic-fermionic

Fulfilling the potential of the SMEFT framework demands global analyses based on **a wide range of process** such that most (all?) **directions in the EFT parameter space** are covered

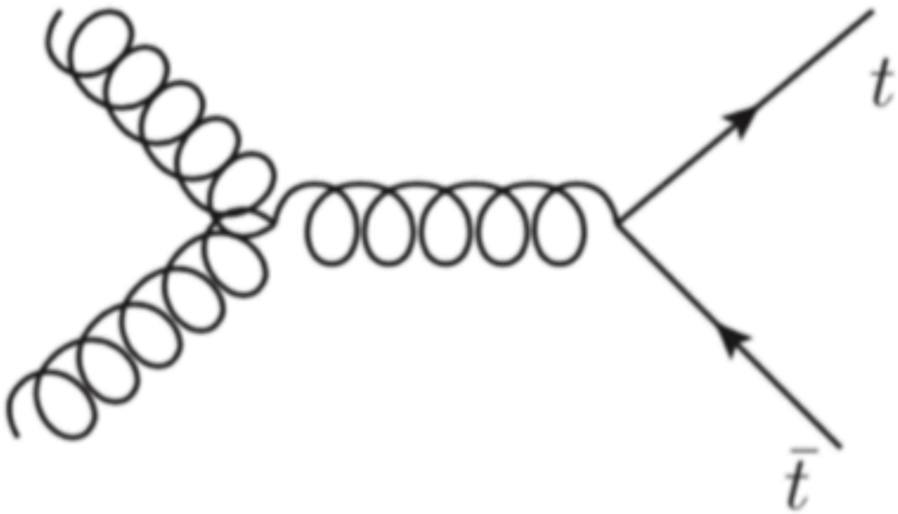
$Q_{\ell\ell}$	$(\bar\ell\gamma_\mu\ell)(\bar\ell\gamma^\mu\ell)$	$Q_{\ell e}$	$(\bar\ell\gamma_\mu\ell)(\bar e\gamma^\mu e)$
$Q_{qq}^{(1)}$	$(\bar q\gamma_\mu q)(\bar q\gamma^\mu q)$	$Q_{\ell u}$	$(\bar\ell\gamma_\mu\ell)(\bar u\gamma^\mu u)$
$Q_{qq}^{(3)}$	$(\bar q\gamma_\mu\tau^I q)(\bar q\gamma^\mu\tau^I q)$	$Q_{\ell d}$	$(\bar\ell\gamma_\mu\ell)(\bar d\gamma^\mu d)$
$Q_{\ell q}^{(1)}$	$(\bar\ell\gamma_\mu\ell)(\bar q\gamma^\mu q)$	Q_{qe}	$(\bar q\gamma_\mu q)(\bar e\gamma^\mu e)$
$Q_{\ell q}^{(3)}$	$(\bar\ell\gamma_\mu\tau^I\ell)(\bar q\gamma^\mu\tau^I q)$	$Q_{qu}^{(1)}$	$(\bar q\gamma_\mu q)(\bar u\gamma^\mu u)$

Q_{eW}	$(\bar\ell\sigma^{\mu\nu}e)\tau^I\varphi W^I_{\mu\nu}$	$Q_{\varphi q}^{(3)}$	$(\varphi^\dagger i\overleftrightarrow{D}^I_\mu\varphi)(\bar q\tau^I\gamma^\mu q)$
Q_{eB}	$(\bar\ell\sigma^{\mu\nu}e)\varphi B_{\mu\nu}$	$Q_{\varphi u}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi)(\bar u\gamma^\mu u)$
Q_{uG}	$(\bar q\sigma^{\mu\nu}T^A u)\widetilde\varphi G^A_{\mu\nu}$	$Q_{\varphi d}$	$(\varphi^\dagger i\overleftrightarrow{D}_\mu\varphi)(\bar d\gamma^\mu d)$
Q_{uW}	$(\bar q\sigma^{\mu\nu}u)\tau^I\widetilde\varphi W^I_{\mu\nu}$	$Q_{\varphi ud}$	$(\widetilde\varphi^\dagger iD_\mu\varphi)(\bar u\gamma^\mu d)$

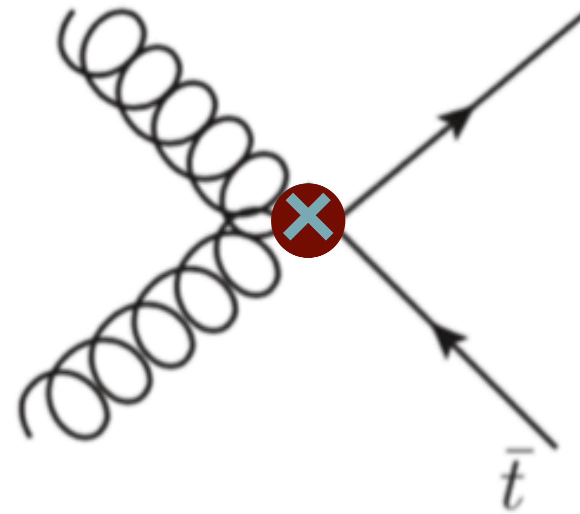
SMEFT effects in top quark pair production

Standard Model

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



+



new interactions between SM fields

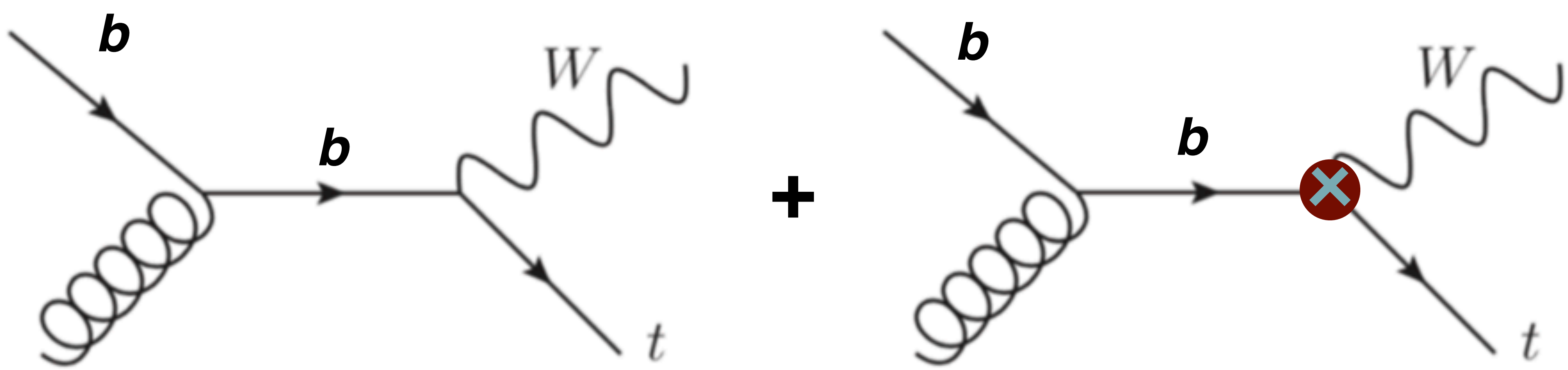
$$= \sigma_{SM} \times \left(1 + a\frac{c_{tG}}{\Lambda^2} + b\frac{c_{tG}^2}{\Lambda^4} \right)$$

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

SMEFT effects in single top production

Standard Model

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



modifications of the SM interactions

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

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

Theory calculations in the SMEFT

from Lagrangian …

$$\mathscr{L}_{\text{SMEFT}} = \mathscr{L}_{\text{SM}} + \sum_{m=1}^{N_6} \frac{c_m}{\Lambda^2} \mathcal{O}_i^{(6)} + \sum_{n=1}^{N_8} \frac{b_j}{\Lambda^4} \mathcal{O}_i^{(8)} + \ldots$$

SM — **EFT$_{d6}$** — **EFT$_{d8}$**

Theory calculations in the SMEFT

from Lagrangian …

$$\mathscr{L}_{\rm SMEFT} = \mathscr{L}_{\rm SM} + \sum_{m=1}^{N_6} \frac{c_m}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_{n=1}^{N_8} \frac{b_j}{\Lambda^4}\mathcal{O}_i^{(8)} + \ldots$$

SM **$\mathrm{EFT}_{\rm d6}$** **$\mathrm{EFT}_{\rm d8}$**

Linear EFT cross-sections:

interference SM-$\mathrm{EFT}_{\rm d6}$

Quadratic EFT cross-sections:

squares $\mathrm{EFT}_{\rm d6}$

to cross-sections ….

$$\sigma_{\rm SMEFT}\left(\boldsymbol{c},\boldsymbol{\Lambda}\right) \simeq \sigma_{\rm SM} \times \left(1 + \sum_{m=1}^{N_6} \frac{c_m}{\Lambda^2}\sigma_m^{\rm (eft)} + \sum_{m,n=1}^{N_6} \frac{c_m c_n}{\Lambda^4}\sigma_{m,n}^{\rm (eft)}\right)$$

evaluate at (N)NLO QCD + NLO EW

evaluate at NLO QCD
with **SMEFT@NLO**

Theory calculations in the SMEFT

… to constraints on the EFT parameters by comparing with data

$$\chi^2(\mathbf{c},\Lambda) = \frac{1}{n_{\rm dat}} \sum_{i,j=1}^{n_{\rm dat}} \left(\sigma_{\rm i,SMEFT}(\mathbf{c},\Lambda) - \sigma_{\rm i,exp}\right) \left(\mathrm{cov}^{-1}\right)_{ij} \left(\sigma_{\rm j,SMEFT}(\mathbf{c},\Lambda) - \sigma_{\rm j,exp}\right)$$

log-likelihood minimisation

to cross-sections ….

Linear EFT cross-sections: interference SM-EFT$_{\rm d6}$

Quadratic EFT cross-sections: squares EFT$_{\rm d6}$

$$\sigma_{\rm SMEFT}(\mathbf{c},\boldsymbol{\Lambda}) \simeq \sigma_{\rm SM} \times \left(1 + \sum_{m=1}^{N_6} \frac{c_m}{\Lambda^2}\sigma_m^{(\rm eft)} + \sum_{m,n=1}^{N_6} \frac{c_m c_n}{\Lambda^4}\sigma_{m,n}^{(\rm eft)}\right)$$

evaluate at (N)NLO QCD + NLO EW

evaluate at NLO QCD with **SMEFT@NLO**

Towards 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

Delivery

Bounds for UV completions

New data incorporated without redoing fit

Methodology

Efficient exploration of **parameter space**

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

A combined interpretation of Higgs, diboson, and top quark data in the SMEFT

N. P. Hartland et al. ``A Monte Carlo global analysis of the Standard Model Effective Field Theory: the top quark sector,'' JHEP 04 (2019), 10 [arXiv:1901.05965 [hep-ph]].

J. J. Ethier et al. [SMEFiT Collaboration], ``Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC,'' JHEP 11 (2021), 089, [arXiv:2105.00006 [hep-ph]].

The SMEFiT framework

Theory

(N)NLO QCD + NLO EW for SM xsecs

NLO QCD, both linear and quadratic terms, **with SMEFT@NLO**

State-of-the-art **parton distributions** (avoid double counting)

Data

Higgs data (signal strengths, diff, STXS), diboson LEP and LHC, all available **top quark data** from Runs I+II, VBS, more in progress

Full experimental **correlations** included

SMEFiT

Extensive **statistical toolbox** to validate results: information geometry, PCA, closure testing, …

Full **posterior probabilities** in the EFT coefficients available, likelihoods WIP

Validation

Two independent fitting methods, **MCfit** and **NestedSampling** (no reliance on linear approx) cross-check each other

Modular structure facilitates adding new datasets of better theory calculations

Methodology

Fitting methodology

MCfit generate a large sample of **Monte Carlo replicas** to construct the **probability distribution** in the space of experimental data accounting for all uncertainties

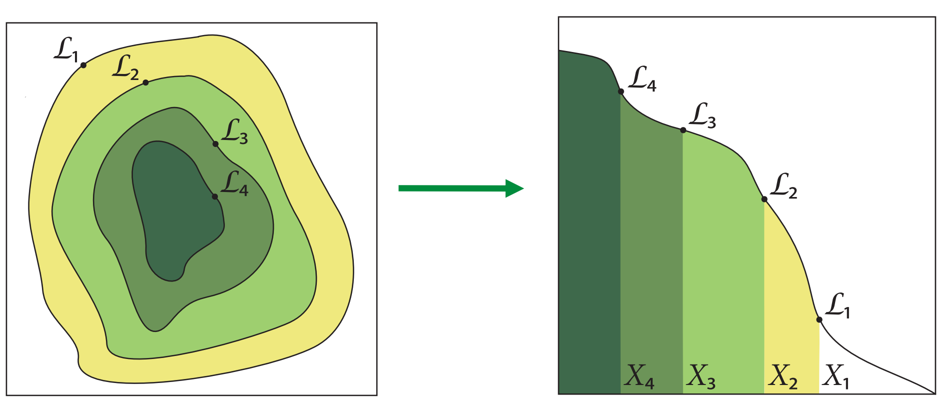
Determine the SMEFT coefficients **replica-by-replica** by minimising a cost function

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

where covariance matrix includes **all sources of experimental + theory errors**

Nested Sampling statistical mapping of the *N*-dimensional likelihood profile to 1D

$$Z = \int d^N c \mathscr{L} \left({\rm data} \,|\, \vec{c} \right) \pi(\vec{c}) = \int_0^1 dX \mathscr{L}(X)$$



- Samples directly from prior space to locate **regions of maximum likelihood**
- Main advantage: **no need for optimiser** (fitting)
- Exponential increase in runtime as prior volume increases

The SMEFiT framework

SMEFiT

Search docs

OVERVIEW:

Features

Available Datasets

THEORY:

SMEFT

References

TALKS AND LECTURES:

Talks and seminars

IMPLEMENTATION:

Fitting strategies

Nested Sampling

MCFit

RESULTS:

SMEFiT Top

SMEFiT RW

SMEFiT VBS

SMEFiT2.0

SMEFiT

Welcome to the SMEFiT website!

SMEFiT is a Python package for global analyses of particle physics data in the framework of the Standard Model Effective Field Theory (SMEFT). The SMEFT represents a powerful model-independent framework to constrain, identify, and parametrise potential deviations with respect to the predictions of the Standard Model (SM). A particularly attractive feature of the SMEFT is its capability to systematically correlate deviations from the SM between different processes. The full exploitation of the SMEFT potential for indirect New Physics searches from precision measurements requires combining the information provided by the broadest possible dataset, namely carrying out extensive global analysis which is the main purpose of SMEFiT.

Project description

The SMEFiT framework has been used in the following **scientific publications**:

- *A Monte Carlo global analysis of the Standard Model Effective Field Theory: the top quark sector*, N. P. Hartland, F. Maltoni, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou, C. Zhang [HMN+19].
- *Constraining the SMEFT with Bayesian reweighting*, S. van Beek, E. R. Nocera, J. Rojo, and E. Slade [vBNRS19].
- *SMEFT analysis of vector boson scattering and diboson data from the LHC Run II* , J. Ethier, R. Gomez-Ambrosio, G. Magni, J. Rojo [EGAMR21].
- *Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC*, J. Ethier, F. Maltoni, L. Mantani, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou, C. Zhang [EMM+21]

Results from these publications, including driver and analysis scripts, are available in the **Results** section.

Team description

The **SMEFiT collaboration** is currently composed by the following members:

- Jaco ter Hoeve, *VU Amsterdam and Nikhef Theory Group*
- Giacomo Magni, *VU Amsterdam and Nikhef Theory Group*
- Fabio Maltoni, *Centre for Cosmology, Particle Physics and Phenomenology Louvain and University of Bologna*
- Luca Mantani, *Centre for Cosmology, Particle Physics and Phenomenology Louvain*
- Emanuele Roberto Nocera, *Higgs Center for Theoretical Physics, University of Edinburgh*
- Juan Rojo, *VU Amsterdam and Nikhef Theory Group*
- Eleni Vryonidou, *University of Manchester*

https://lhcfitnikhef.github.io/SMEFT/

Operator basis and flavour assumptions

Class	N_{dof}	Independent DOFs	DoF in EWPOs
four-quark (two-light-two-heavy)	14	$c_{Qq}^{1,8}$, $c_{Qq}^{1,1}$, $c_{Qq}^{3,8}$, $c_{Qq}^{3,1}$, c_{tq}^{8}, c_{tq}^{1}, c_{tu}^{8}, c_{tu}^{1}, c_{Qu}^{8}, c_{Qu}^{1}, c_{td}^{8}, c_{td}^{1}, c_{Qd}^{8}, c_{Qd}^{1}	
four-quark (four-heavy)	5	c_{QQ}^{1}, c_{QQ}^{8}, c_{Qt}^{1}, c_{Qt}^{8}, c_{tt}^{1}	
four-lepton	1		$c_{\ell\ell}$
two-fermion (+ bosonic fields)	23	$c_{t\varphi}$, c_{tG}, $c_{b\varphi}$, $c_{c\varphi}$, $c_{\tau\varphi}$, c_{tW}, c_{tZ}, $c_{\varphi Q}^{(3)}$, $c_{\varphi Q}^{(-)}$, $c_{\varphi t}$	$c_{\varphi\ell_1}^{(1)}$, $c_{\varphi\ell_1}^{(3)}$, $c_{\varphi\ell_2}^{(1)}$, $c_{\varphi\ell_2}^{(3)}$, $c_{\varphi\ell_3}^{(1)}$, $c_{\varphi\ell_3}^{(3)}$, $c_{\varphi e}$, $c_{\varphi\mu}$, $c_{\varphi\tau}$, $c_{\varphi q}^{(3)}$, $c_{\varphi q}^{(-)}$, $c_{\varphi u}$, $c_{\varphi d}$
Purely bosonic	7	$c_{\varphi G}$, $c_{\varphi B}$, $c_{\varphi W}$, $c_{\varphi d}$, c_{WWW}	$c_{\varphi WB}$, $c_{\varphi D}$
Total	50 (36 independent)	34	16 (2 independent)

- **Dim-6 SMEFT operators** modifying Higgs, dibosons, and top quark properties: **36** (14) **independent** (dependent) **DoFs**

- Flavour assumption is **MFV**, with $\mathrm{U}(2)_q \times \mathrm{U}(2)_u \times \mathrm{U}(3)_d$ in quark sector (special role for top quark) and $\left(\mathrm{U}(1)_\ell \times \mathrm{U}(1)_e\right)^3$ in lepton sector

- Constraints from **LEP EWPOs** imposed via restrictions in parameter space

$$
\begin{pmatrix} c_{\varphi\ell_i}^{(3)} \\ c_{\varphi\ell_i}^{(1)} \\ c_{\varphi e/\mu/\tau} \\ c_{\varphi q}^{(-)} \\ c_{\varphi q}^{(3)} \\ c_{\varphi u} \\ c_{\varphi d} \\ c_{\ell\ell} \end{pmatrix} = \begin{pmatrix} -\frac{1}{t_W} & -\frac{1}{4t_W^2} \\ 0 & -\frac{1}{4} \\ 0 & -\frac{1}{2} \\ \frac{1}{t_W} & \frac{1}{4s_W^2} - \frac{1}{6} \\ -\frac{1}{t_W} & -\frac{1}{4t_W^2} \\ 0 & \frac{1}{3} \\ 0 & -\frac{1}{6} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} c_{\varphi WB} \\ c_{\varphi D} \end{pmatrix}
$$

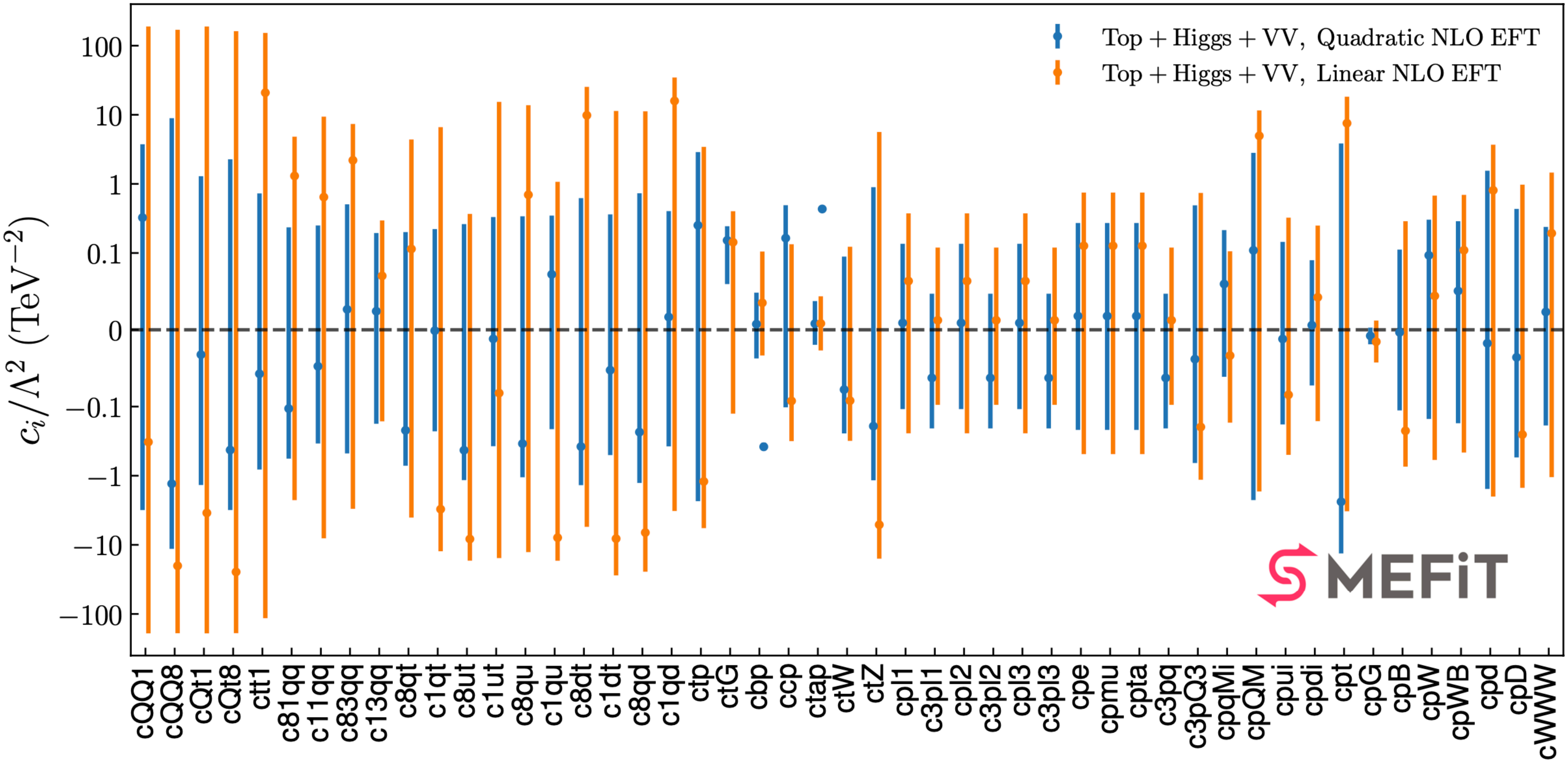
Experimental data

Category	Processes	n_{dat}
Top quark production	$t\bar{t}$ (inclusive) **(incl LHC charge asy)**	94
	$t\bar{t}Z$, $t\bar{t}W$ **(incl ptZ in ttZ)**	14
	single top (inclusive)	27
	tZ, tW	9
	$t\bar{t}t\bar{t}$, $t\bar{t}b\bar{b}$	6
	Total	**150**
Higgs production and decay	Run I signal strengths	22
	Run II signal strengths	40
	Run II, differential distributions & STXS	35
	Total	**97**
Diboson production	LEP-2 **(WW)**	40
	LHC **(WW & WZ)**	30
	Total	**70**
Baseline dataset	**Total**	**317**

+ systematic assessment of fit results **wrt dataset variations**:

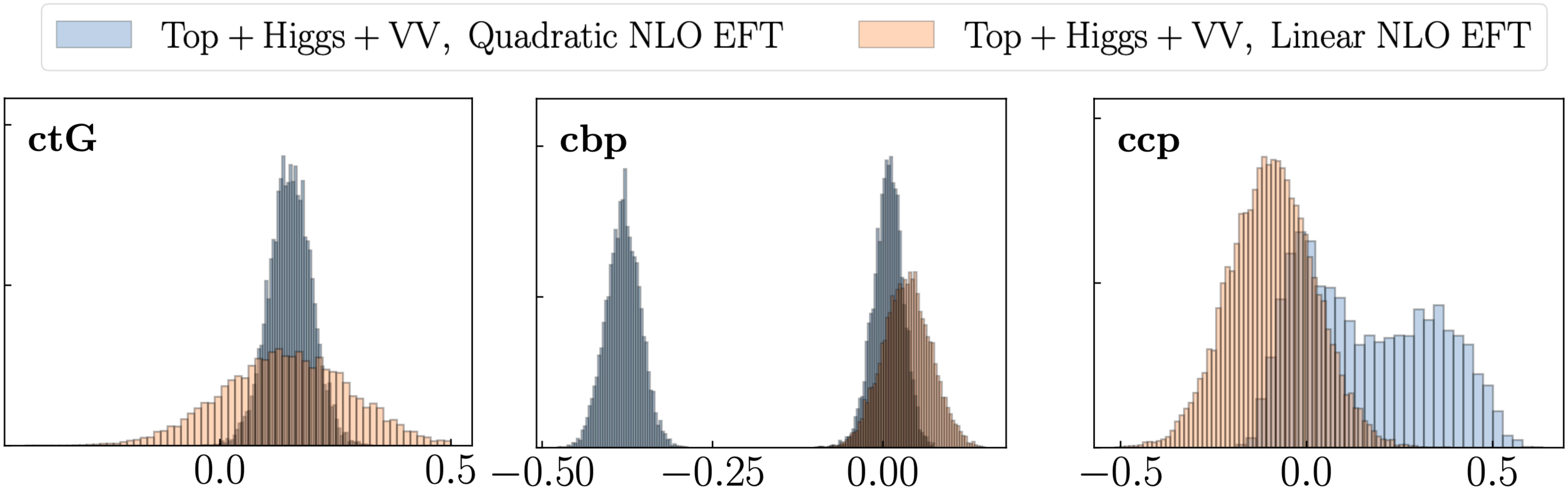
Higgs-only fit, top-only fit, no high-E data, no diboson data …

Results: global fit



- Agreement with SM at 95% CL for all EFT coefficients except for **ctG** in quadratic fit
- Quadratic corrections bring in sensitivity (more stringent bounds) *e.g.* for four-fermion operators
- Some DoFs exhibit a second ``BSM-like" solution in the quartic fit

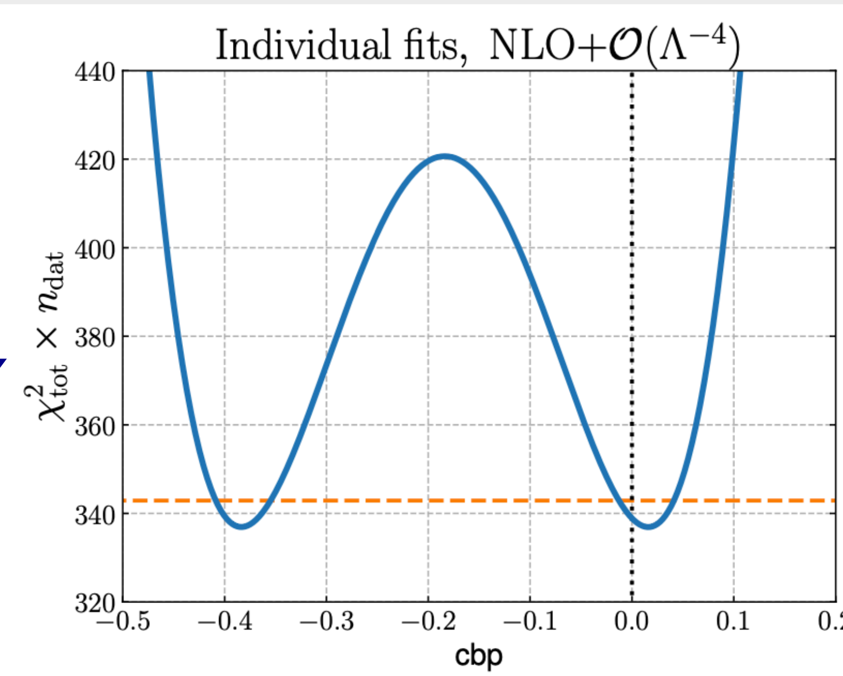
Results: global fit

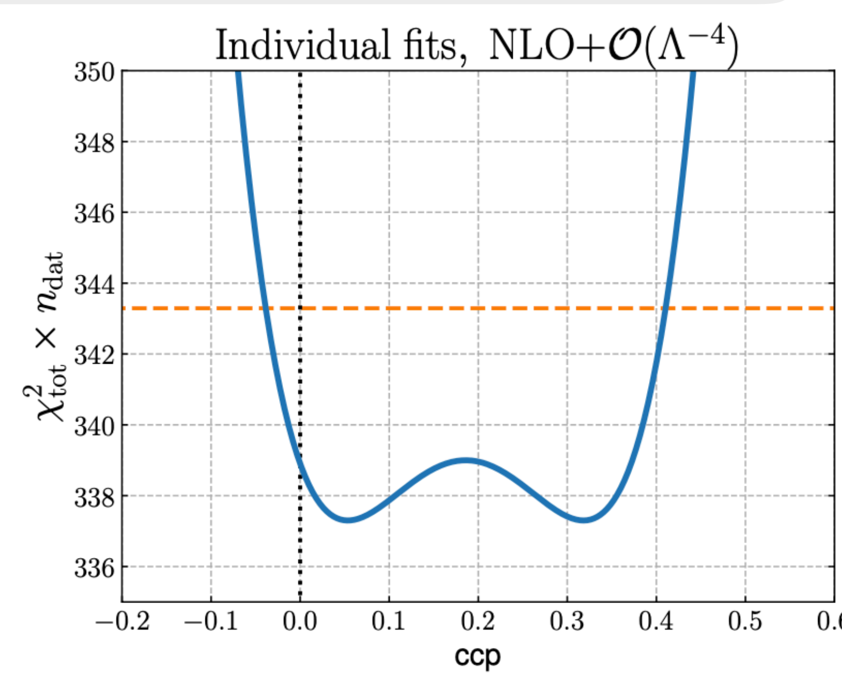


chromo-magnetic operator ***bottom Yukawa*** ***charm Yukawa***

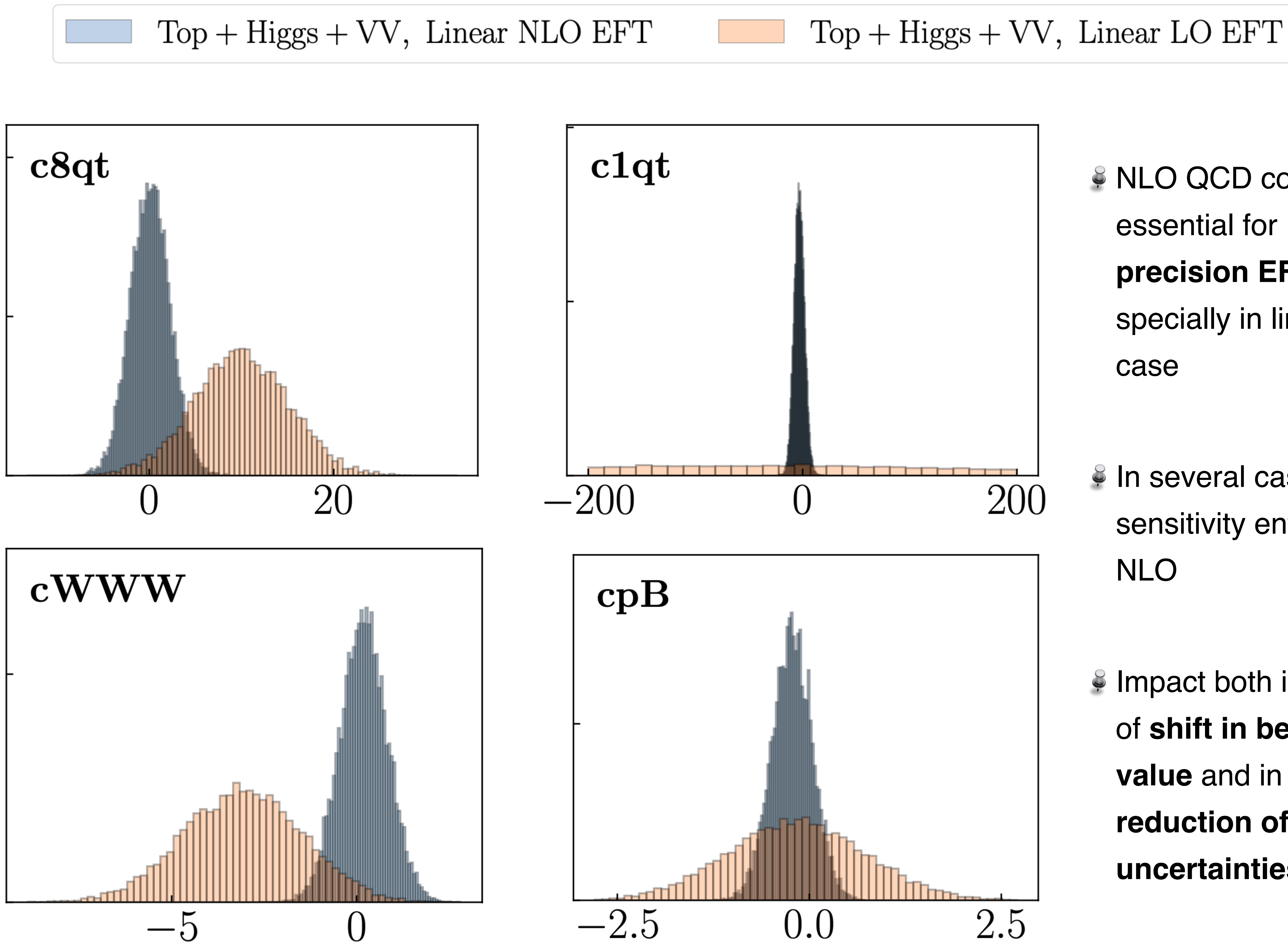
in general, sensitivity of fit results to inclusion of **quadratic EFT corrections**

1-parameter fits



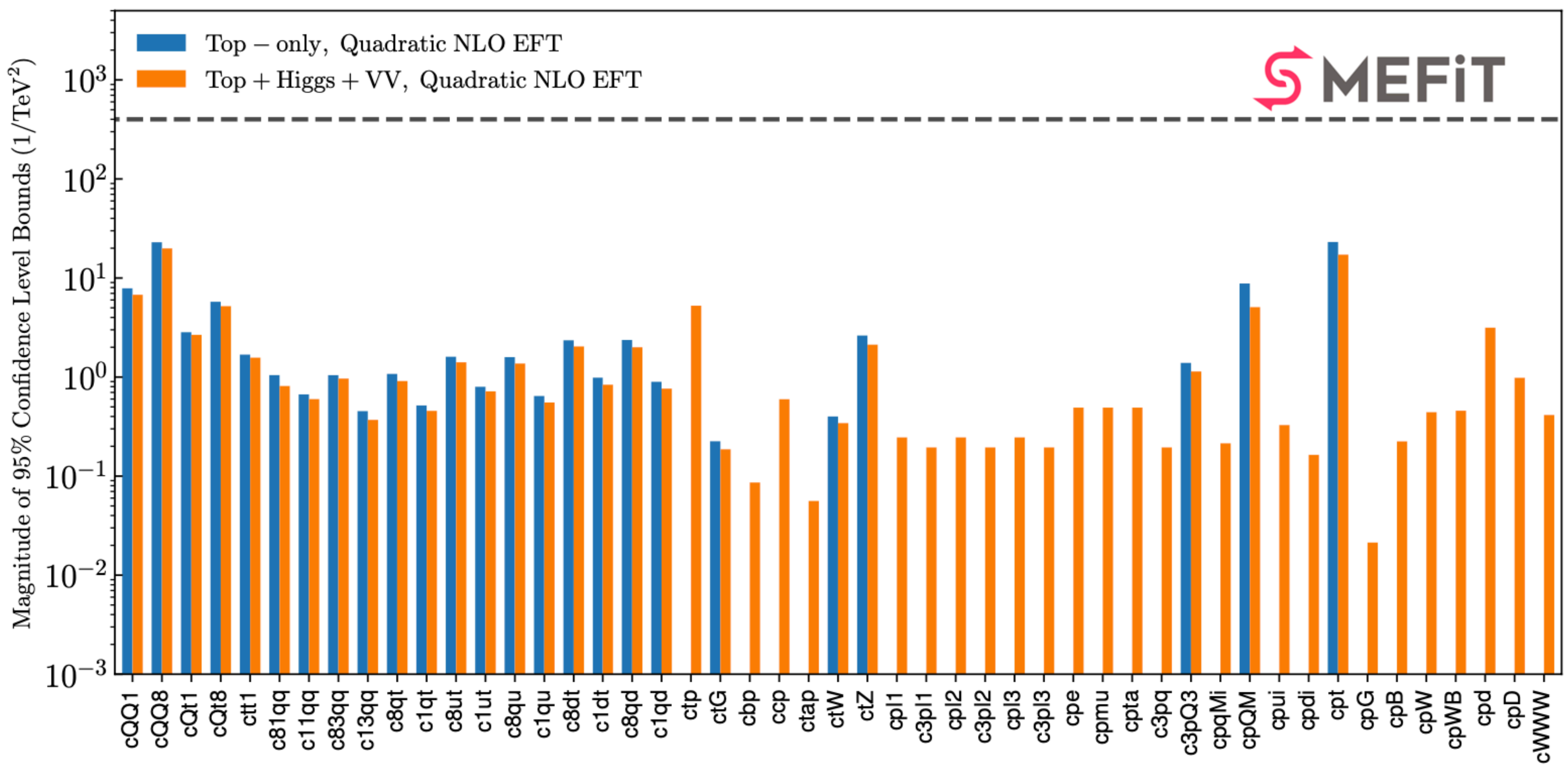


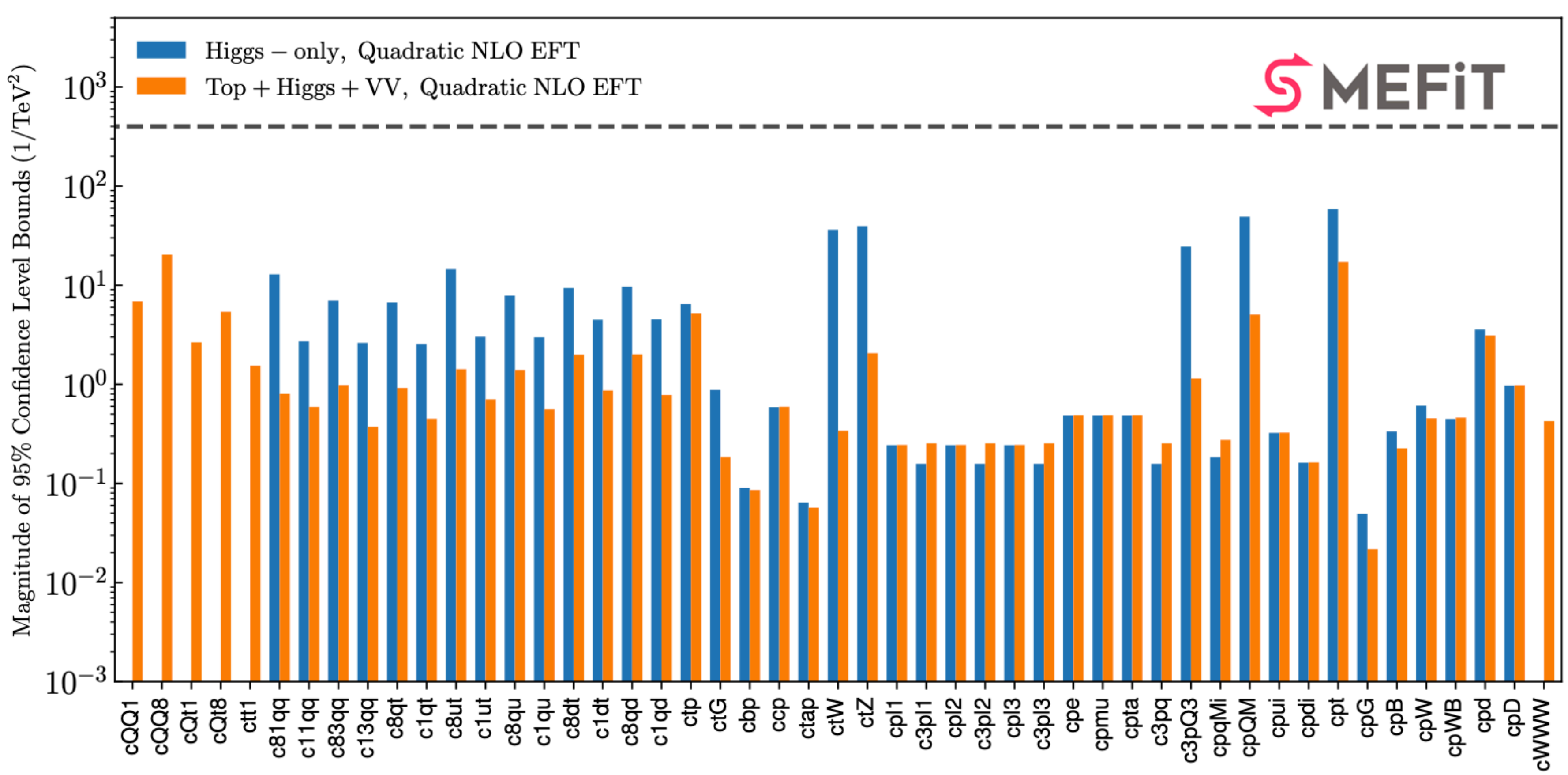
Results: impact of NLO corrections



- NLO QCD corrections essential for **precision EFT fits**, specially in linear case
- In several cases new sensitivity enters at NLO
- Impact both in terms of **shift in best-fit value** and in **reduction of fit uncertainties**

Results: dataset dependence





- Global fits consistent, but **more accurate**, with top-only or Higgs-only fit
- Top data **boosts the Higgs EFT fit** all across the board
- Diboson data only constraints **cWWW**
- Fit results stable upon **removal of high energy bins** (E > 1 TeV)

The SMEFiT framework reloaded

- The SMEFiT framework has been completely rewritten and **released as an open source general EFT fitting toolbox**
- It allows reproducing results of the global SMEFiT analyses, adding new datasets or improved theory calculations, quantifying the impact of future measurements …
- As an application example, we **reproduce the results of the ATLAS EFT interpretation of Higgs measurements** based om the full Run II dataset

ATL-PHYS-PUB-2022-037

SMEFiT

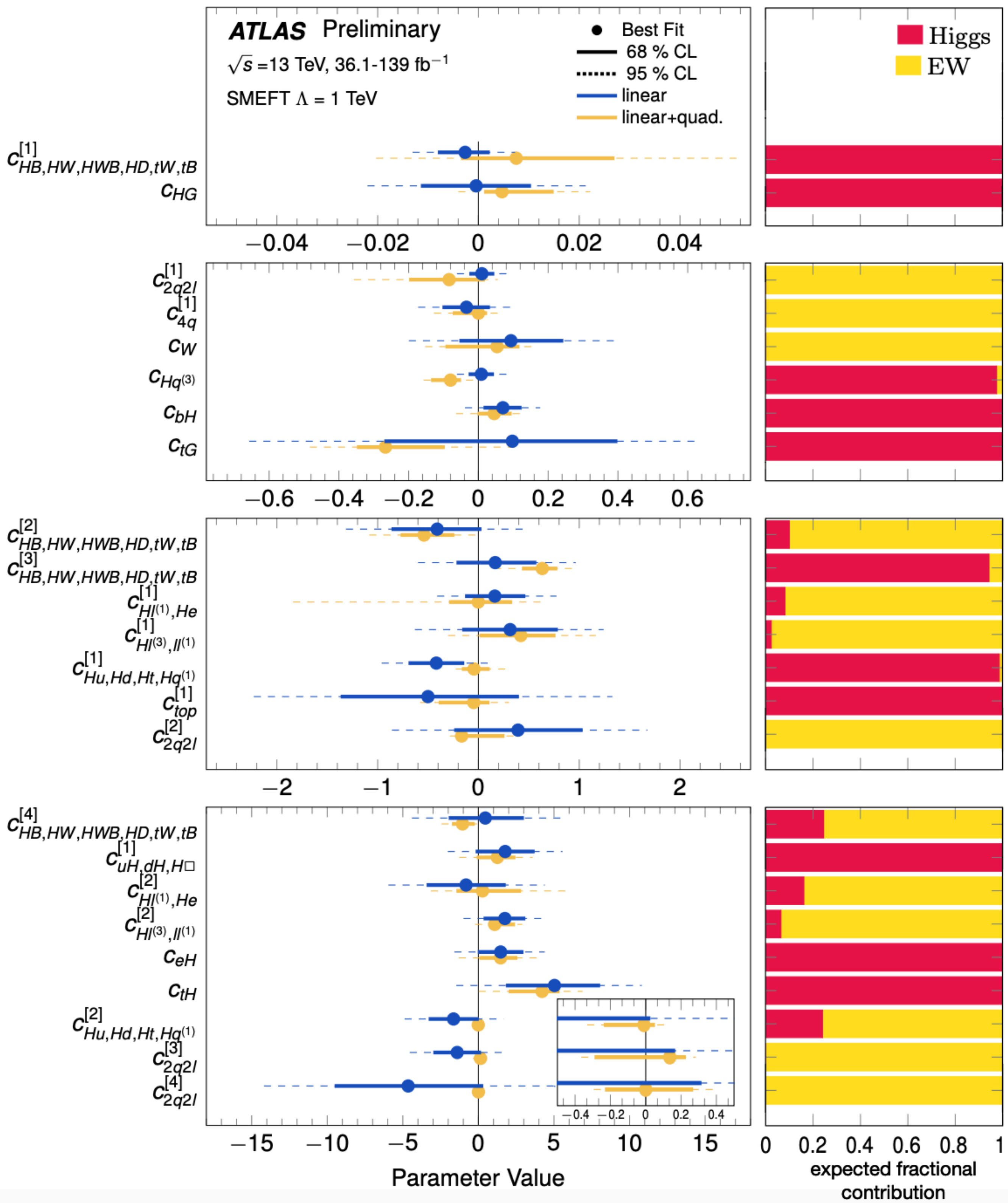
Nikhef-2022-aaa

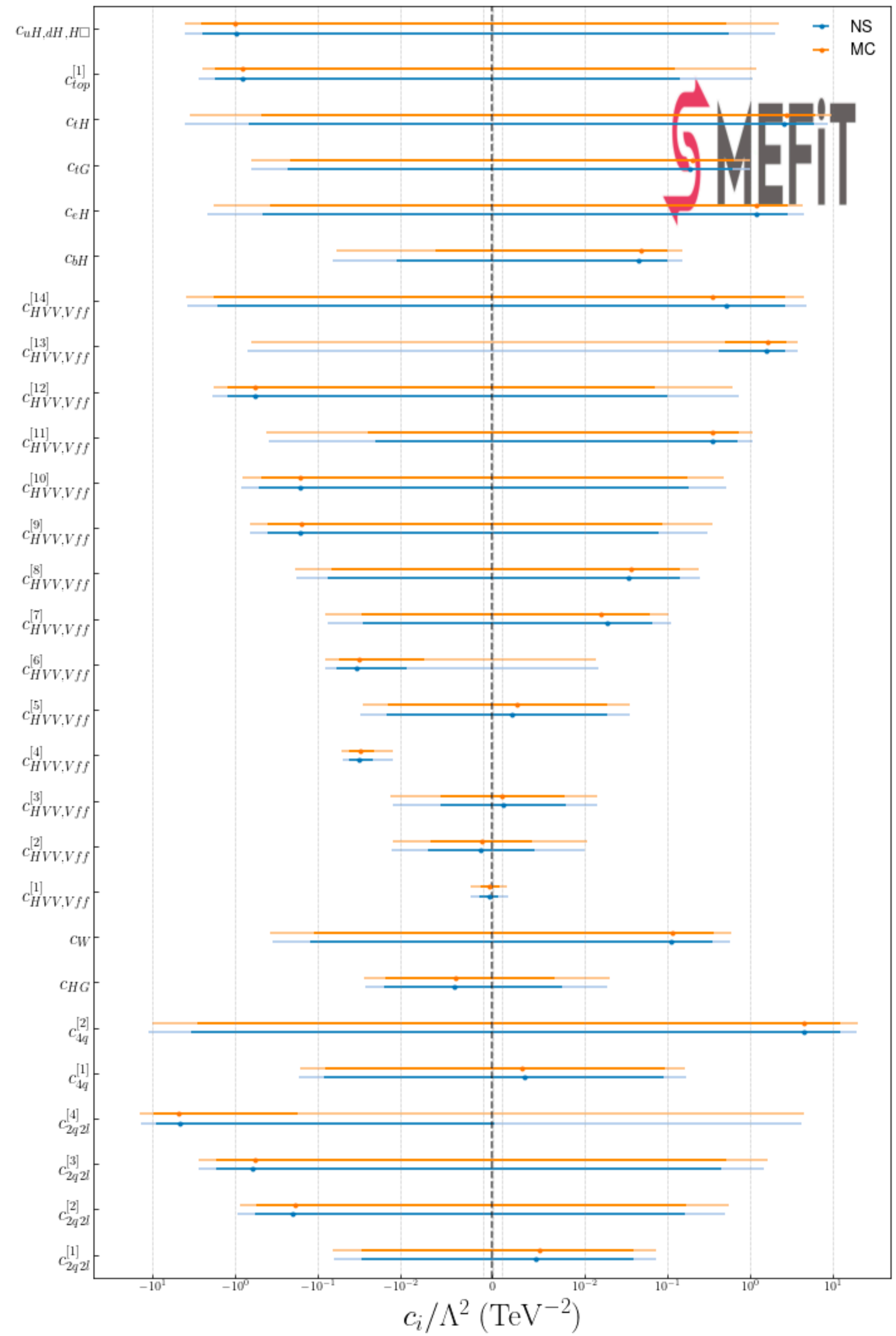
SMEFiT: a flexible toolbox for global interpretations of particle physics data with effective field theories

https://lhcfitnikhef.github.io/smefit_release/

Tommaso Giani, Giacomo Magni, and Juan Rojo

The SMEFiT framework reloaded





For the same inputs, we reproduce the **ATLAS linear EFT** results in **two different fitting bases**

Statistically optimal observables for global SMEFT fits

R. Gomez-Ambrosio, J. ter Hoeve, M. Madigan, J. Rojo, V. Sanz, ``***Unbinned multivariate observables for global SMEFT analyses from machine learning***'', submitted yesterday to the arXiv

https://lhcfitnikhef.github.io/ML4EFT/

Statistically optimal observables for EFTs

Which kind of measurement is **most sensitive to SMEFT operators?**

Difficult question to answer in general since SMEFT-sensitive measurements can be:

- Inclusive or (1,2,3, …)-differential (in which specific variables?)
- Binned (choice of binning?) or unbinned
- Unfolded at parton level, at particle level, or at detector level

relevant to many other extractions of SM & BSM parameters from data

Our approach:

deploy **unbinned multivariate measurements** to determine the best sensitivity that a given process can have on SMEFT operators by means of **machine learning techniques**

Gaussian likelihood

observed event counts — *predicted event counts*

$$\mathcal{L}(\boldsymbol{n};\boldsymbol{\nu}(\boldsymbol{c})) = \prod_{i=1}^{N_b} \exp\left[-\frac{1}{2}\frac{(n_i - \nu_i(\boldsymbol{c}))^2}{\nu_i(\boldsymbol{c})}\right]$$

retains full information on event-by-event kinematics

Unbinned multivariate likelihood

sum over events — *event probability* — *event kinematics*

$$\mathcal{L}(\boldsymbol{c}) = \frac{\nu_{\rm tot}(\boldsymbol{c})^{N_{\rm ev}}}{N_{\rm ev}!}e^{-\nu_{\rm tot}(\boldsymbol{c})}\prod_{i=1}^{N_{\rm ev}} f_\sigma\left(\boldsymbol{x}_i, \boldsymbol{c}\right) \qquad f_\sigma\left(\boldsymbol{x}, \boldsymbol{c}\right) = \frac{1}{\sigma_{\rm fid}(\boldsymbol{c})}\frac{d\sigma(\boldsymbol{x},\boldsymbol{c})}{d\boldsymbol{x}}\,,$$

Statistically optimal observables for EFTs

Which kind of measurement is **most sensitive to SMEFT operators?**

Difficult question to answer in general since SMEFT-sensitive measurements can be:

- Inclusive or (1,2,3, …)-differential (in which specific variables?)
- Binned (choice of binning?) or unbinned
- Unfolded at parton level, at particle level, or at detector level

relevant to many other extractions of SM & BSM parameters from data

Our approach:

deploy **unbinned multivariate measurements** to determine the best sensitivity that a given process can have on SMEFT operators by means of **machine learning techniques**

Challenges:

- Parameter inference requires knowledge of the likelihood for **any value of the EFT coefficients**
- Evaluation of likelihood functions **computationally costly due to high dimensionality** both of the space of kinematic features $\boldsymbol{x}$ and of EFT parameters $\boldsymbol{c}$

Solution:

- Neural networks as **universal unbiased interpolants** to parametrise high-dimensional likelihoods

Statistically optimal observables from ML

the dependence of the cross-section on **kinematic variables and all EFT coefficients**

$$r_\sigma(\boldsymbol{x}, \boldsymbol{c}) \equiv \frac{f_\sigma(\boldsymbol{x}, \boldsymbol{c})}{f_\sigma(\boldsymbol{x}, \boldsymbol{0})} = 1 + \sum_{j=1}^{n_{\rm eft}} r_\sigma^{(j)}(\boldsymbol{x}) c_j + \sum_{j=1}^{n_{\rm eft}} \sum_{k \geq j}^{n_{\rm eft}} r_\sigma^{(j,k)}(\boldsymbol{x}) c_j c_k$$

parametrised with **neural networks** trained to Monte Carlo simulations & benchmarked with exact calculations

$$\hat{r}_\sigma(\boldsymbol{x}, \boldsymbol{c}) = 1 + \sum_{j=1}^{n_{\rm eft}} \mathrm{NN}^{(j)}(\boldsymbol{x}) c_j + \sum_{j=1}^{n_{\rm eft}} \sum_{k \geq j}^{n_{\rm eft}} \mathrm{NN}^{(j,k)}(\boldsymbol{x}) c_j c_k$$

extendable to **arbitrary number** of kinematic variables and EFT coefficients: training can be parallelised

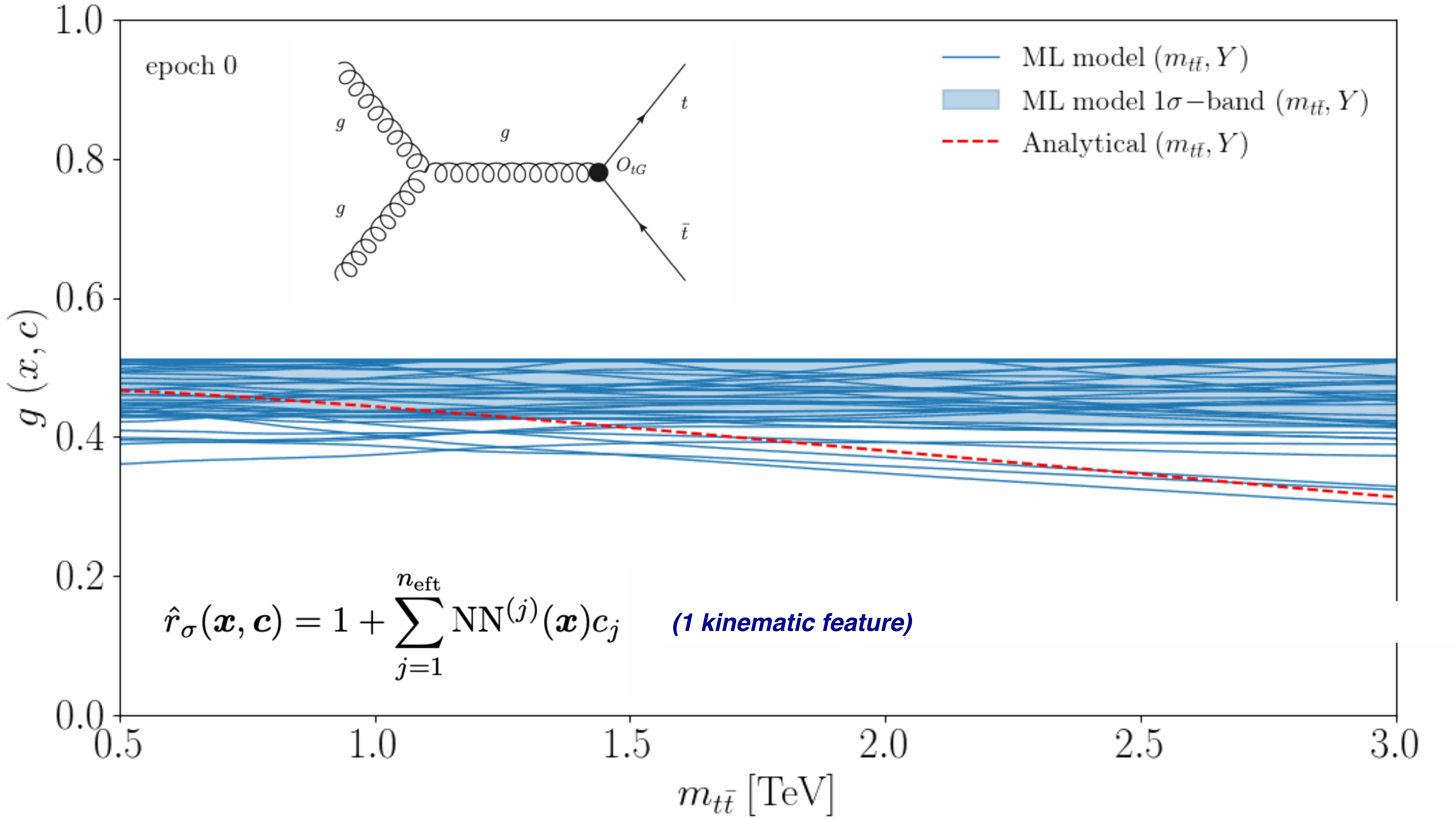
methodological uncertainties (e.g. finite training samples) assess with the **replica method**

$$\hat{r}_\sigma^{(i)}(\boldsymbol{x}, \boldsymbol{c}) \equiv 1 + \sum_{j=1}^{n_{\rm eft}} \mathrm{NN}_i^{(j)}(\boldsymbol{x}) c_j + \sum_{j=1}^{n_{\rm eft}} \sum_{k \geq j}^{n_{\rm eft}} \mathrm{NN}_i^{(j,k)}(\boldsymbol{x}) c_j c_k \, , \qquad i = 1, \ldots, N_{\rm rep}$$

each replica trained to an independent set of MC events

representation of the probability distribution in the space of ML models

Neural network training



$$\hat{r}_\sigma(\boldsymbol{x}, \boldsymbol{c}) = 1 + \sum_{j=1}^{n_{\rm eft}} \mathrm{NN}^{(j)}(\boldsymbol{x}) c_j$$

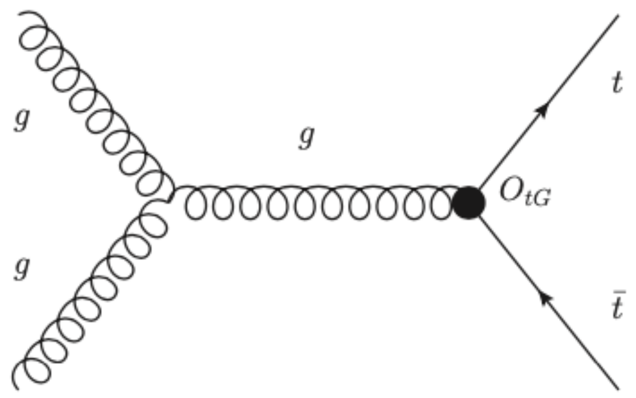
(1 kinematic feature)

NN training by minimising cross-entropy loss function

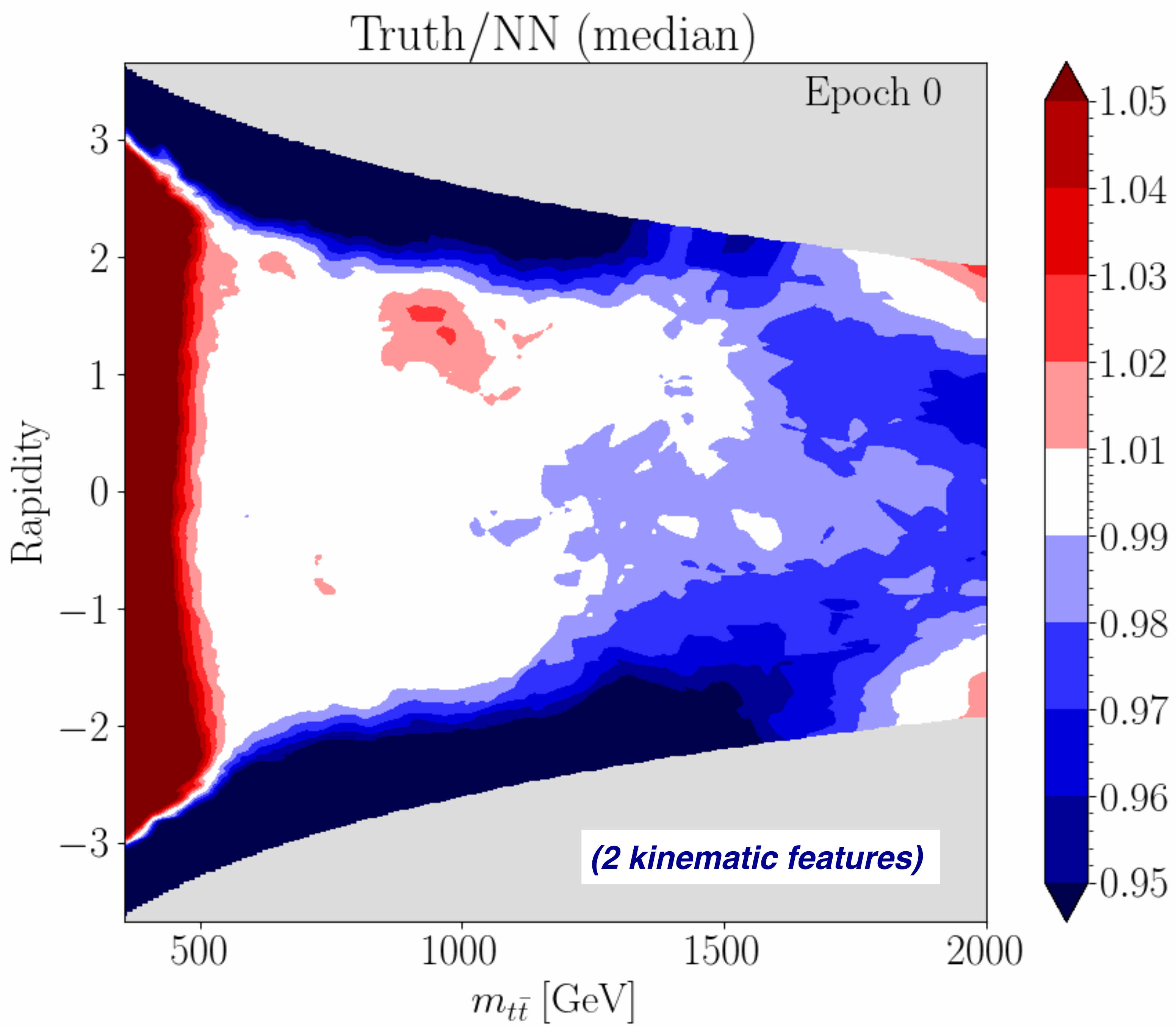
$$L[g(\boldsymbol{x}, \boldsymbol{c})] = -\sigma_{\rm fid}(\boldsymbol{c}) \sum_{i=1}^{N_{\rm ev}} \log(1 - g(\boldsymbol{x}_i, \boldsymbol{c})) - \sigma_{\rm fid}(\boldsymbol{0}) \sum_{j=1}^{N_{\rm ev}} \log g(\boldsymbol{x}_j, \boldsymbol{c})$$

$$g = (1 + r_\sigma)^{-1}$$

Neural network training



$$\hat{r}_\sigma(\boldsymbol{x},\boldsymbol{c}) = 1 + \sum_{j=1}^{n_{\rm eft}} {\rm NN}^{(j)}(\boldsymbol{x}) c_j$$



NN training by minimising cross-entropy loss function

$$L[g(\boldsymbol{x},\boldsymbol{c})] = -\sigma_{\rm fid}(\boldsymbol{c}) \sum_{i=1}^{N_{\rm ev}} \log(1 - g(\boldsymbol{x}_i,\boldsymbol{c})) - \sigma_{\rm fid}(\boldsymbol{0}) \sum_{j=1}^{N_{\rm ev}} \log g(\boldsymbol{x}_j,\boldsymbol{c})$$

$$g = (1 + r_\sigma)^{-1}$$

Neural network training

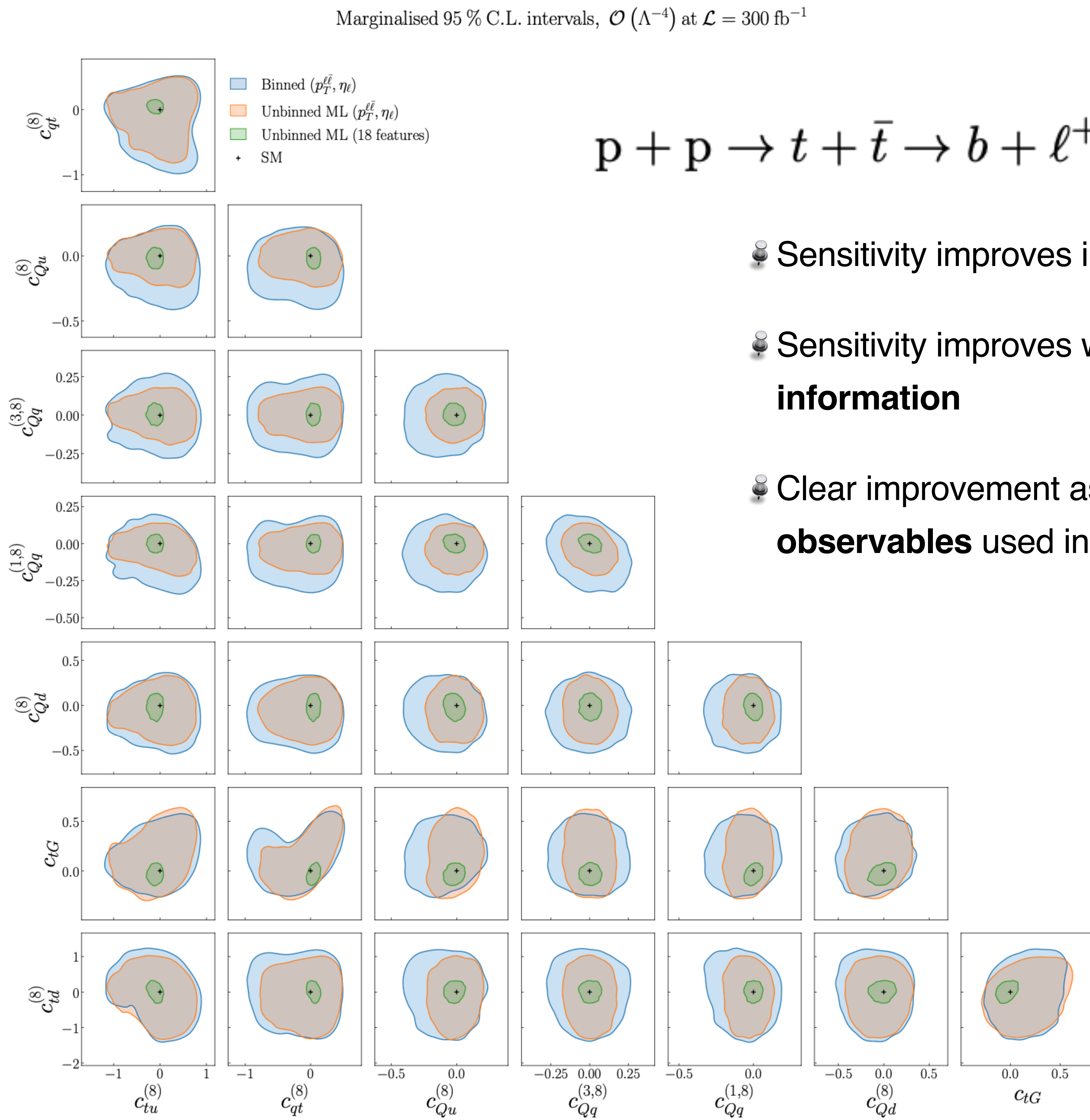
Process	N_{rep}	$c_j^{(\text{tr})}$	$\tilde{N}_{\text{ev}}$ (per replica)	N_{nn}	#trainings
$pp \to t\bar{t}$	50	$c_{tG} = -10$ $c_{tu}^{(8)} = 10$	10^5	4	200
$pp \to t\bar{t} \to b\bar{b}\ell^+\ell^-\nu_\ell\bar{\nu}_\ell$	25	$c_{td}^{(8)} = 10$ $c_{Qd}^{(8)} = 10$ $c_{Qq}^{(1,8)} = 10$ $c_{Qq}^{(3,8)} = 10$ $c_{Qu}^{(8)} = 10$ $c_{tG} = -10$ $c_{qt}^{(8)} = 10$ $c_{tu}^{(8)} = 10$	10^5	40	1000
$pp \to hZ \to b\bar{b}\ell^+\ell^-$	50	$c_{\varphi u} = 10$ $c_{\varphi d} = -10$ $c_{\varphi q}^{(1)} = -10$ $c_{\varphi q}^{(3)} = 10$ $c_{b\varphi} = -10$ $c_{\varphi W} = 10$ $c_{\varphi WB} = 10$	10^5	30	1500

- The number of NN to be trained **grows quadratically with number of EFT parameters**, yet is fully paralellizable

- A realistic scenario requires training **several thousands of NNs**, each with between 10 and 20 input kinematic features

process	features	hidden layers	learning rate	n_{batch}	time (min)
$pp \to t\bar{t}$	$m_{t\bar{t}}$	$25 \times 25 \times 25$	10^{-3}	5	17.3 ± 13.9
	$m_{t\bar{t}}, y_{t\bar{t}}$	$25 \times 25 \times 25$	10^{-3}	5	16.4 ± 12.7
$pp \to t\bar{t} \to b\bar{b}\ell^+\ell^-\nu_\ell\bar{\nu}_\ell$	$p_T^{\ell\bar{\ell}}$	$25 \times 25 \times 25$	10^{-3}	1	46.8 ± 35.0
	$p_T^{\ell\bar{\ell}}, \eta_\ell$	$25 \times 25 \times 25$	10^{-3}	1	53.7 ± 29.9
	18	$100 \times 100 \times 100$	10^{-4}	50	5.4 ± 2.7
$pp \to hZ \to b\bar{b}\ell^+\ell^-$	p_T^Z	$100 \times 100 \times 100$	10^{-3}	100	9.4 ± 9.0
	7	$100 \times 100 \times 100$	10^{-4}	50	14.1 ± 8.7

Results: top quark pair production

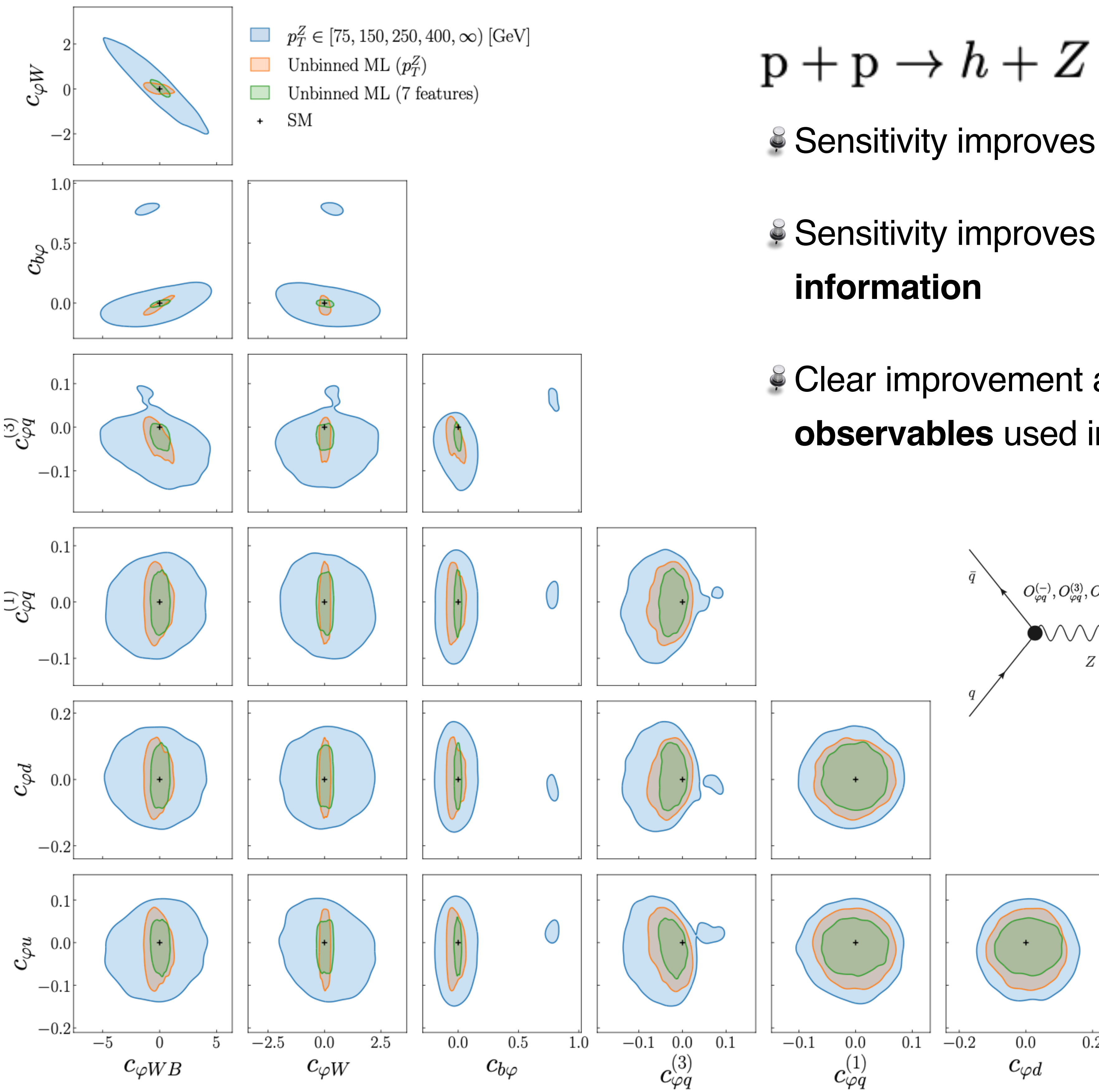


$$\text{p} + \text{p} \to t + \bar{t} \to b + \ell^+ + \nu_\ell + \bar{b} + \ell^- + \bar{\nu}_\ell$$

- Sensitivity improves in **unbinned analysis**
- Sensitivity improves when **using all kinematic information**
- Clear improvement as compared to **traditional observables** used in EFT fits

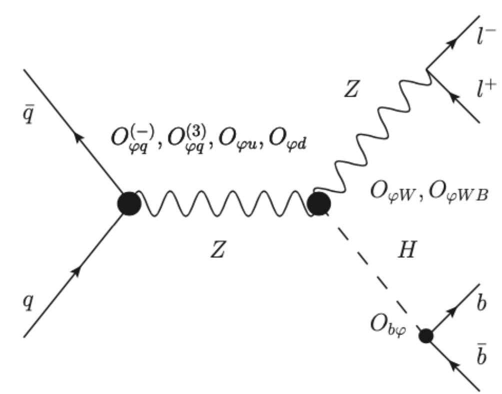
Results: Higgs+Z production

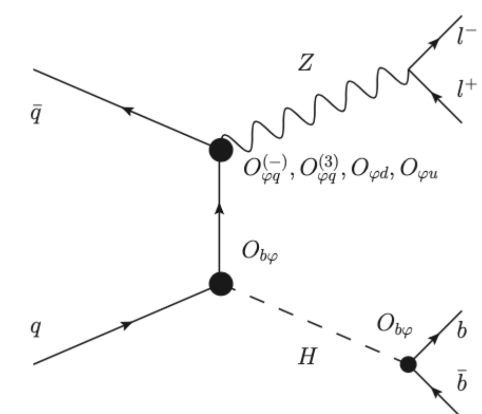
Marginalised 95 % C.L. intervals, $\mathcal{O}\left(\Lambda^{-4}\right)$ at $\mathcal{L} = 300\ \text{fb}^{-1}$



$$\mathrm{p} + \mathrm{p} \rightarrow h + Z \rightarrow b + \bar{b} + \ell^{+} + \ell^{-}$$

- Sensitivity improves in **unbinned analysis**
- Sensitivity improves when **using all kinematic information**
- Clear improvement as compared to **traditional observables** used in EFT fits





Can New Physics Hide Inside the Proton?

S. Carrazza et al,``Can New Physics hide inside the proton?,'' Phys. Rev. Lett. 23 (2019) no.13, 132001, [arXiv:1905.05215 [hep-ph]].

A. Greljo et al, ``Parton distributions in the SMEFT from high-energy Drell-Yan tails,'' JHEP 07 (2021), 122 [arXiv:2104.02723 [hep-ph]].

Can New Physics hide inside the proton?

``How can you be sure you are not reabsorbing BSM physics into PDF fits?"

Assuming the **SM**, the theory calculations that enter a global PDF fit are:

$$\sigma_{\rm LHC}(\boldsymbol{\theta}) \propto \sum_{ij=u,d,g,\ldots} \int_{M^2}^{s} d\hat{s}\; \mathscr{L}_{ij}(\hat{s}, s, \boldsymbol{\theta})\, \widetilde{\sigma}_{{\rm SM},ij}(\hat{s}, \alpha_s(M))$$

SM PDFs

However in the case of BSM physics, here parametrised by the **SMEFT,** the correct expression is:

SMEFT PDFs

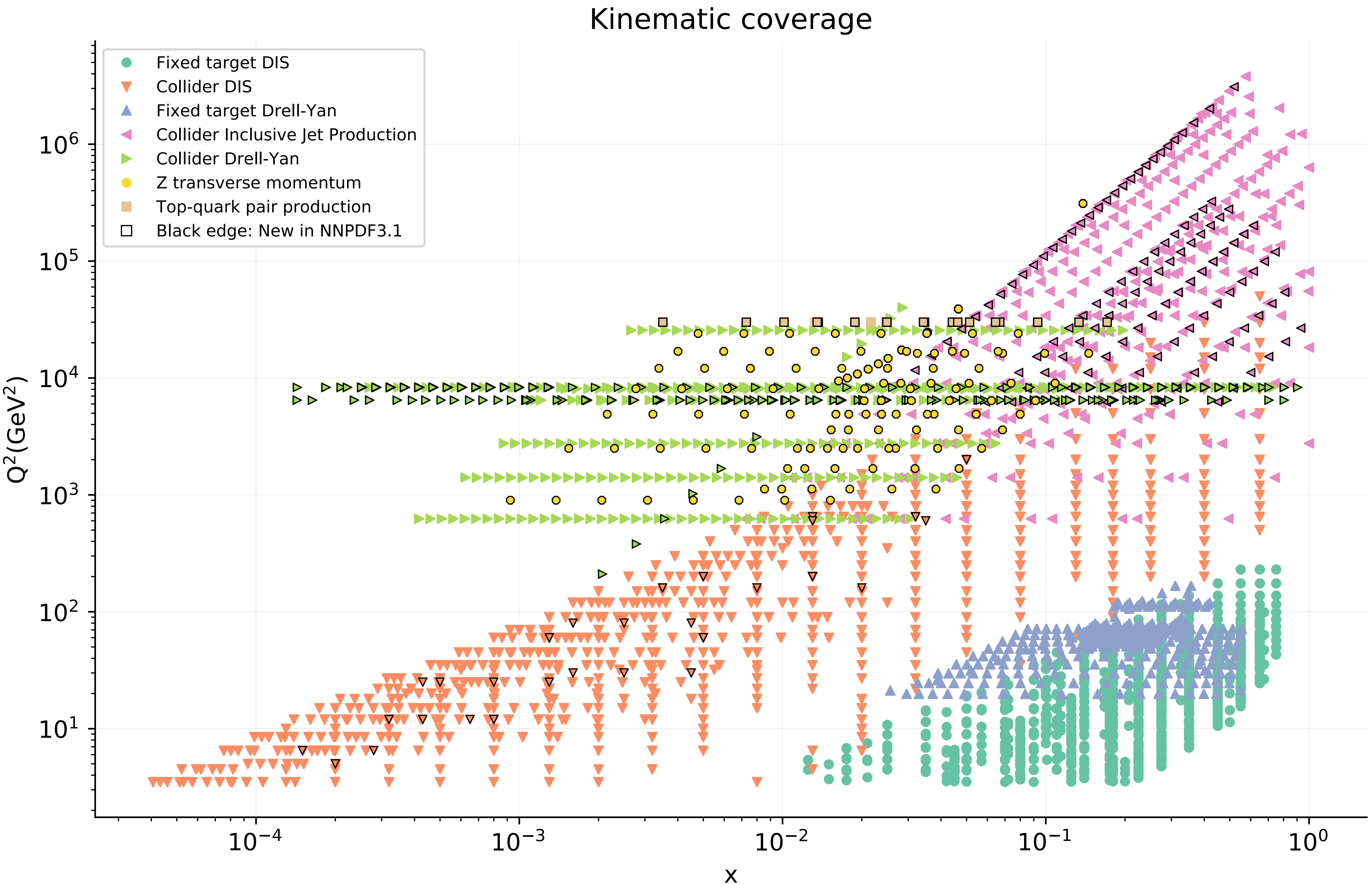
$$\sigma_{\rm LHC}\left(\boldsymbol{c}, \Lambda, \boldsymbol{\theta}\right) \simeq \left(\int_{M^2}^{s} d\hat{s}\; \mathscr{L}_{ij}(\hat{s}, s, \boldsymbol{\theta})\, \widetilde{\sigma}_{{\rm SM},ij}(\hat{s}, \alpha_s(M)) \right) \times \left(1 + \sum_{m=1}^{N_6} c_m \frac{\kappa_m}{\Lambda^2} + \sum_{m,n=1}^{N_6} c_m c_n \frac{\kappa_{mn}}{\Lambda^4} \right),$$

PDF parameters

SMEFT coefficients

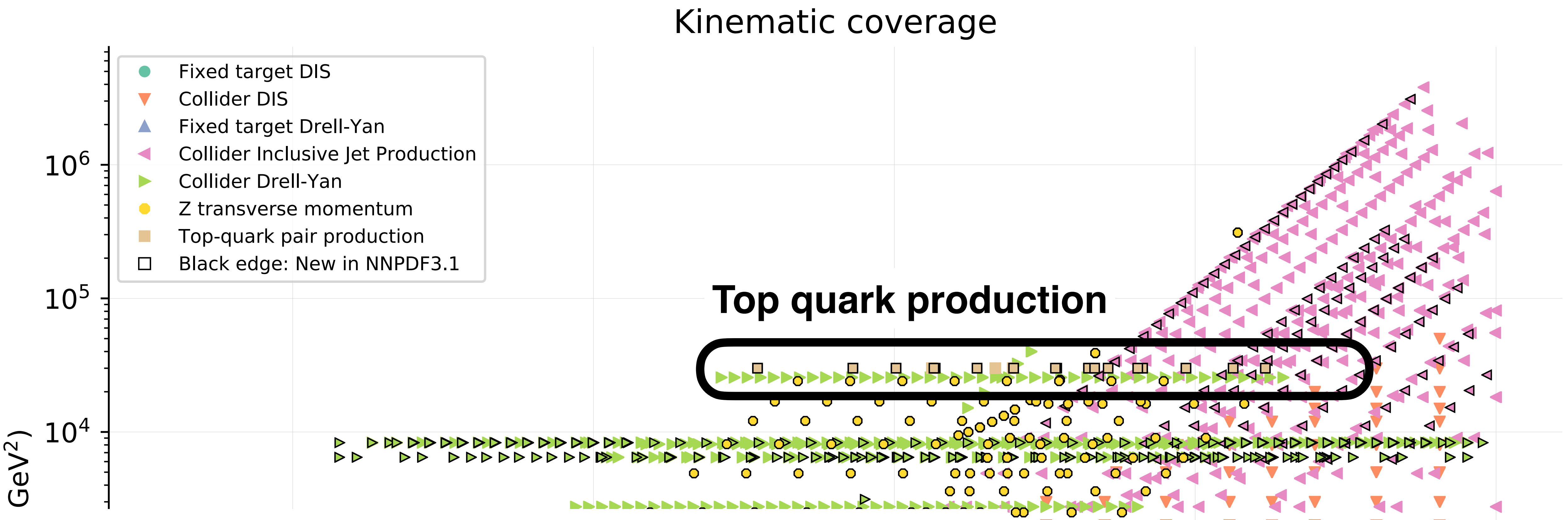
How different are ``SM PDFs" & ``SMEFT PDFs"? Can we quantify the risk of **fitting away BSM** in PDFs?

SMEFT & PDFs

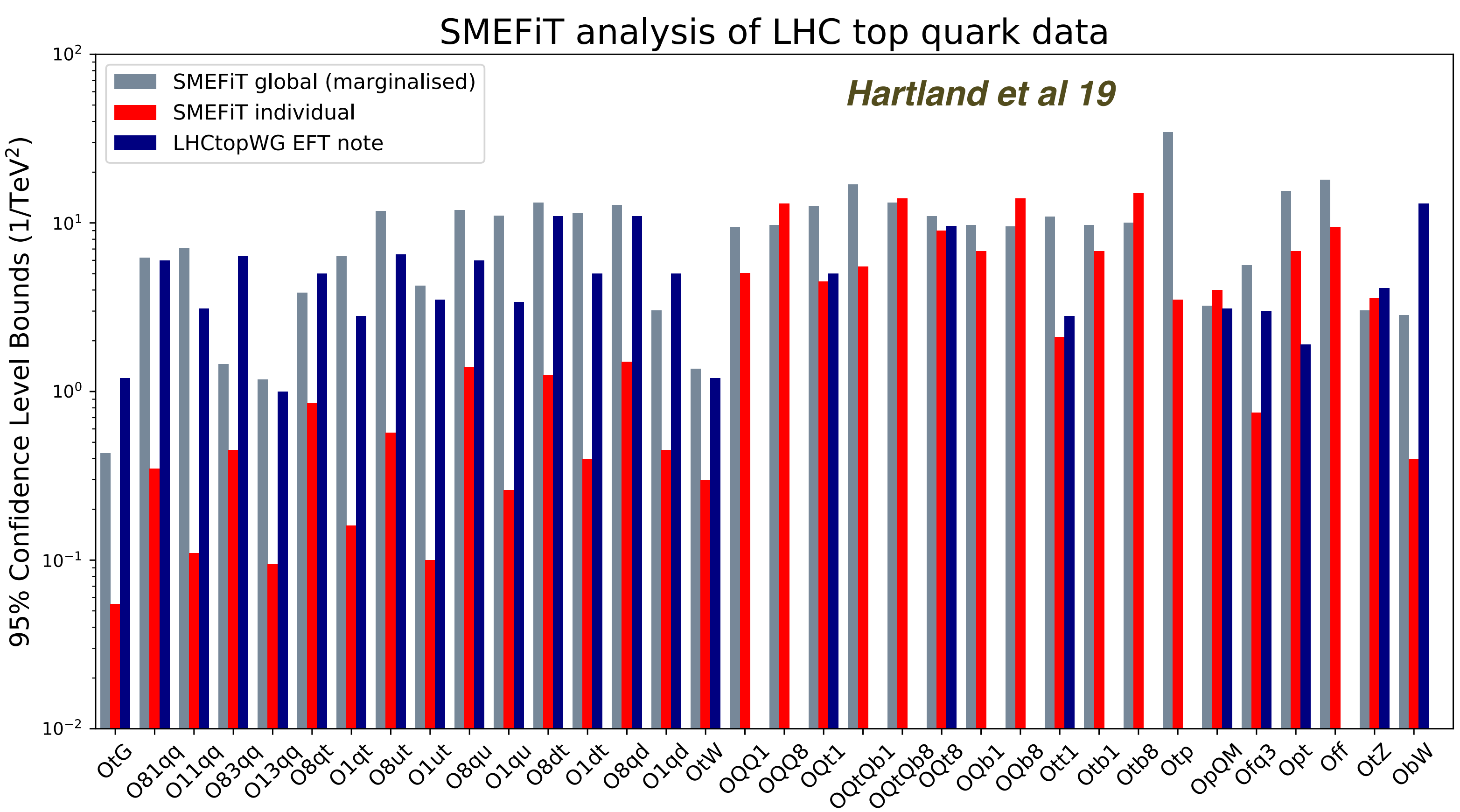


Kinematic coverage
Fixed target DIS
Collider DIS
Fixed target Drell-Yan
Collider Inclusive Jet Production
Collider Drell-Yan
Z transverse momentum
Top-quark pair production
Black edge: New in NNPDF3.1
$Q^2(\mathrm{GeV}^2)$
10^1
10^2
10^3
10^4
10^5
10^6
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^0
x

SMEFT & PDFs

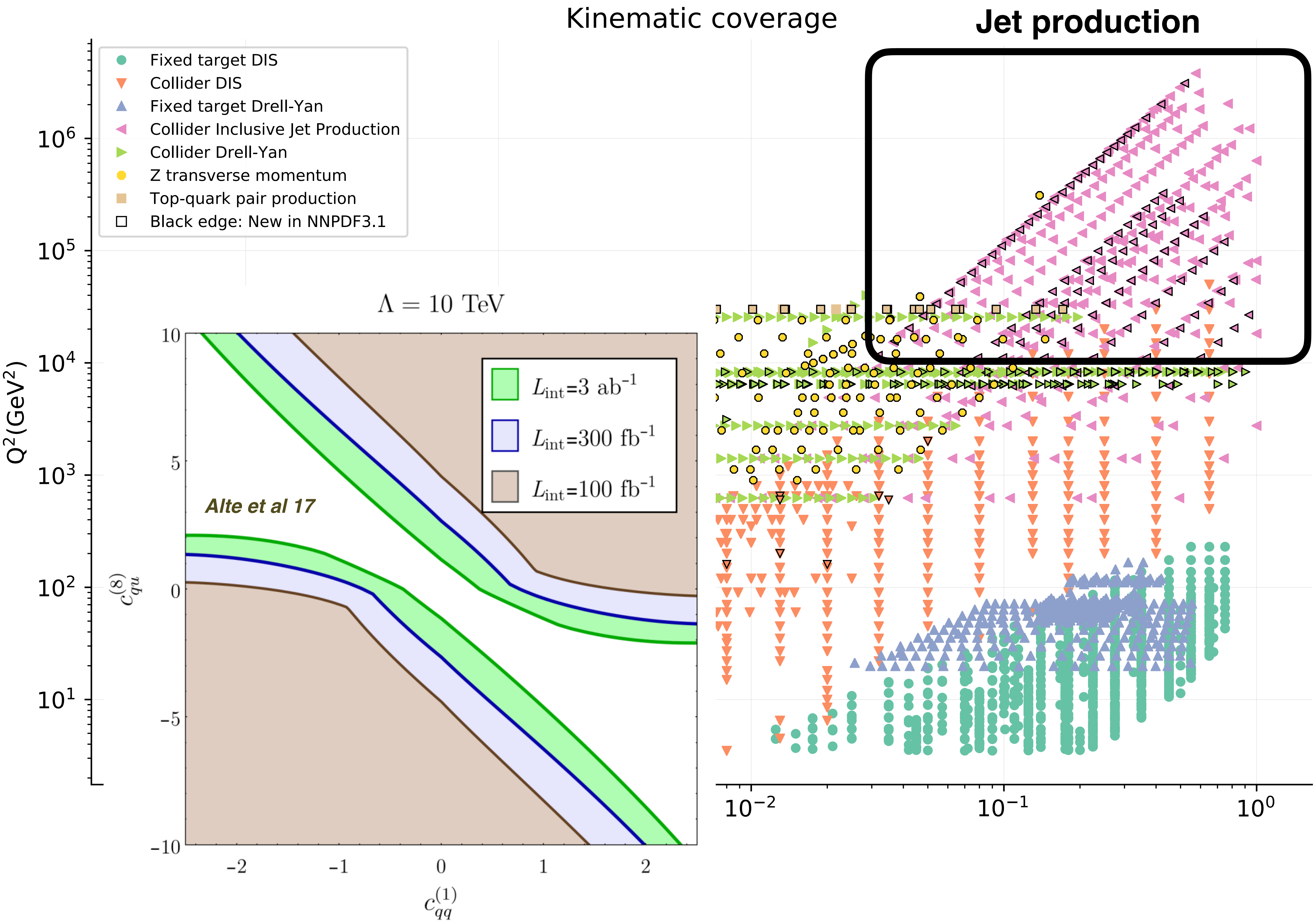


Kinematic coverage
Fixed target DIS
Collider DIS
Fixed target Drell-Yan
Collider Inclusive Jet Production
Collider Drell-Yan
Z transverse momentum
Top-quark pair production
Black edge: New in NNPDF3.1
Top quark production
GeV^2)
10^4
10^5
10^6
10^{-1}
10^0



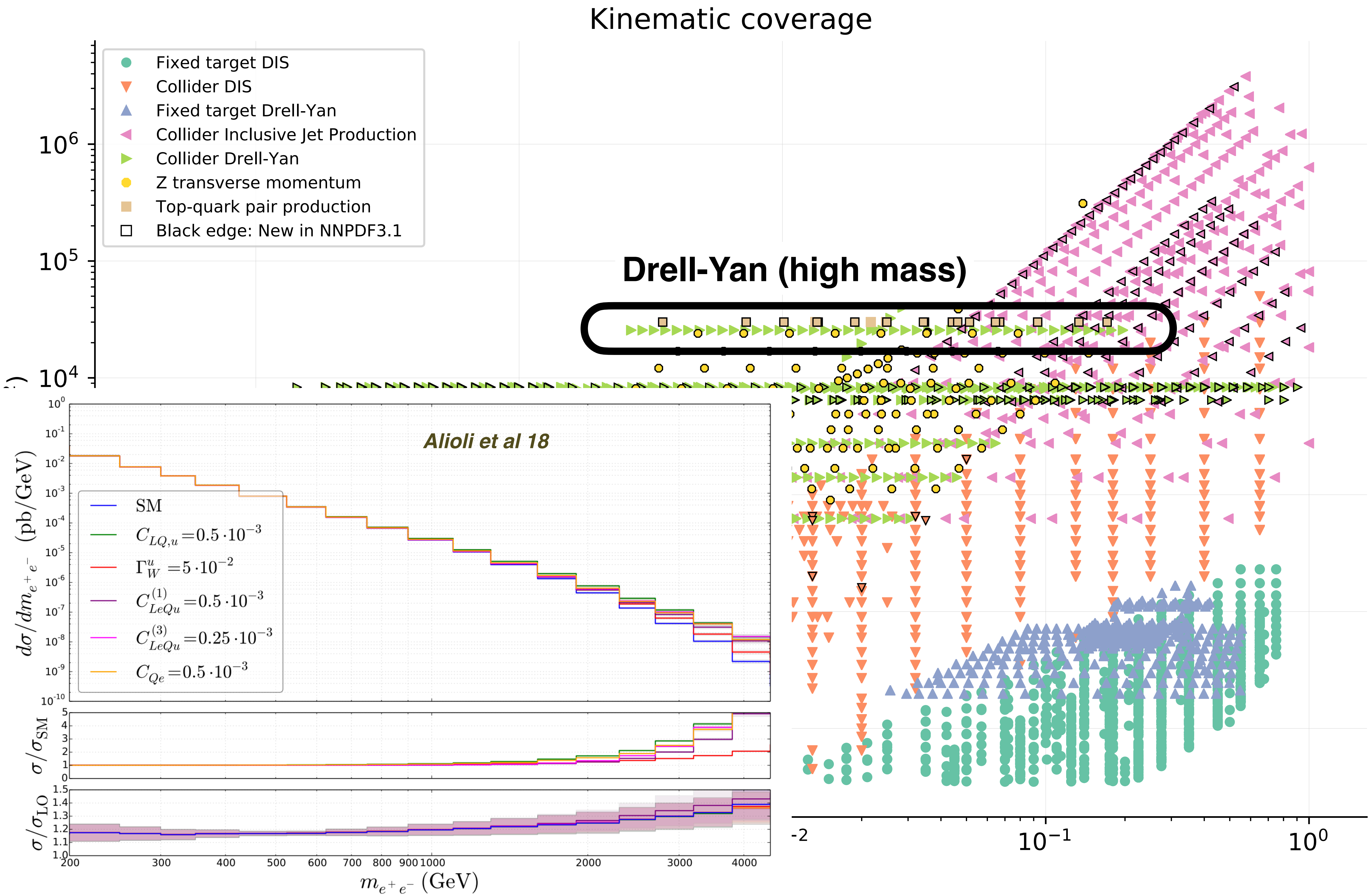
SMEFiT analysis of LHC top quark data
SMEFiT global (marginalised)
SMEFiT individual
LHCtopWG EFT note
Hartland et al 19
95% Confidence Level Bounds (1/TeV^2)
OtG O81qq O11qq O83qq O13qq O8qt O1qt O8ut O1ut O8qu O1qu O8dt O1dt O8qd O1qd OtW OQQ1 OQQ8 OQt1 OQtQb1 OQtQb8 OQt8 OQb1 OQb8 Ott1 Otb1 Otb8 Otp OpQM Ofq3 Opt Off OtZ ObW

SMEFT & PDFs



Kinematic coverage
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
Q²(GeV²)
10⁶
10⁵
10⁴
10³
10²
10¹
10⁻²
10⁻¹
10⁰
Λ = 10 TeV
L_int=3 ab⁻¹
L_int=300 fb⁻¹
L_int=100 fb⁻¹
Alte et al 17
c_qu^(8)
c_qq^(1)
10
5
0
−5
−10
−2
−1
0
1
2

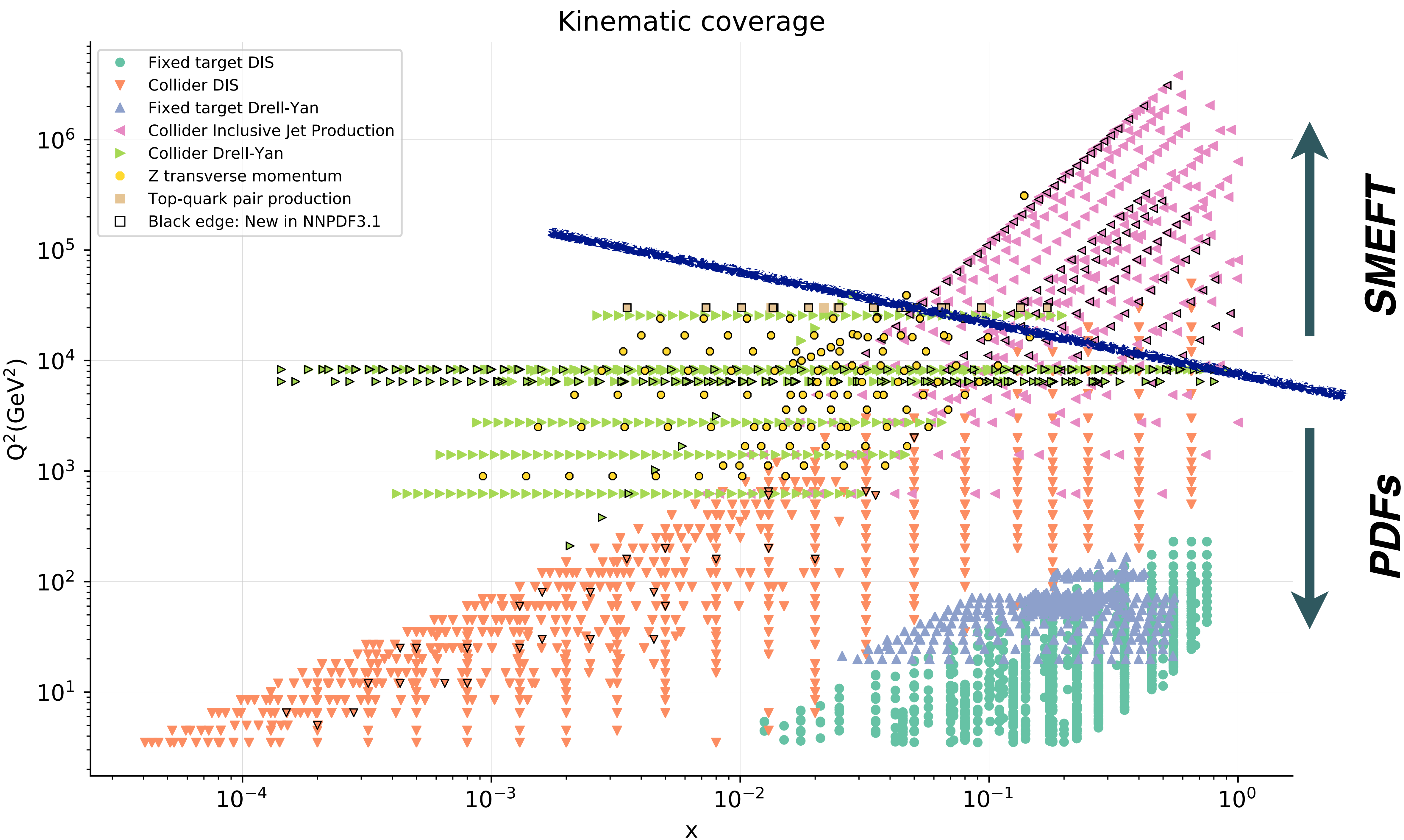
SMEFT & PDFs



Kinematic coverage
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)
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)
σ/σ_{SM}
σ/σ_{LO}
$m_{e^+e^-}$ (GeV)

SMEFT & PDFs

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



significant **information loss on PDFs**, specially in crucial large-x region

SMEFT PDFs from high-E Drell-Yan

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)**	

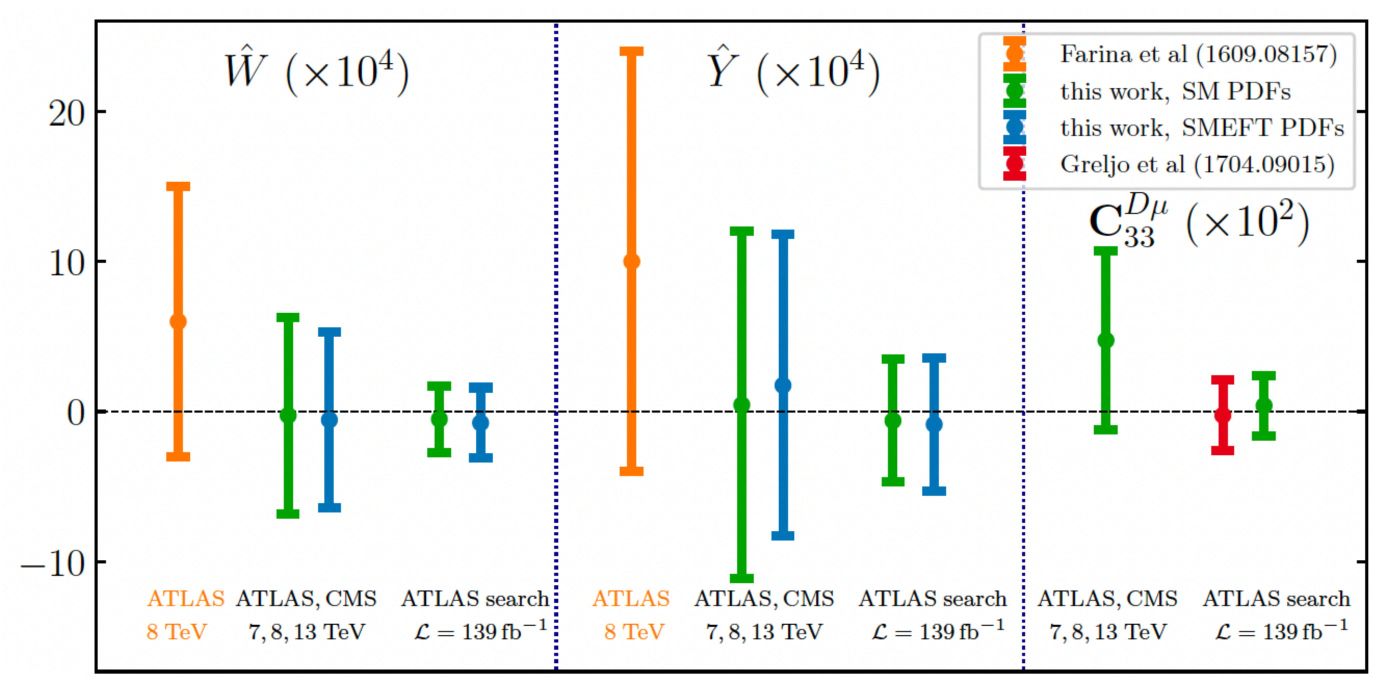
Extract PDFs from global fit where **high-mass DY cross-sections** account for EFT effects in two benchmark scenarios

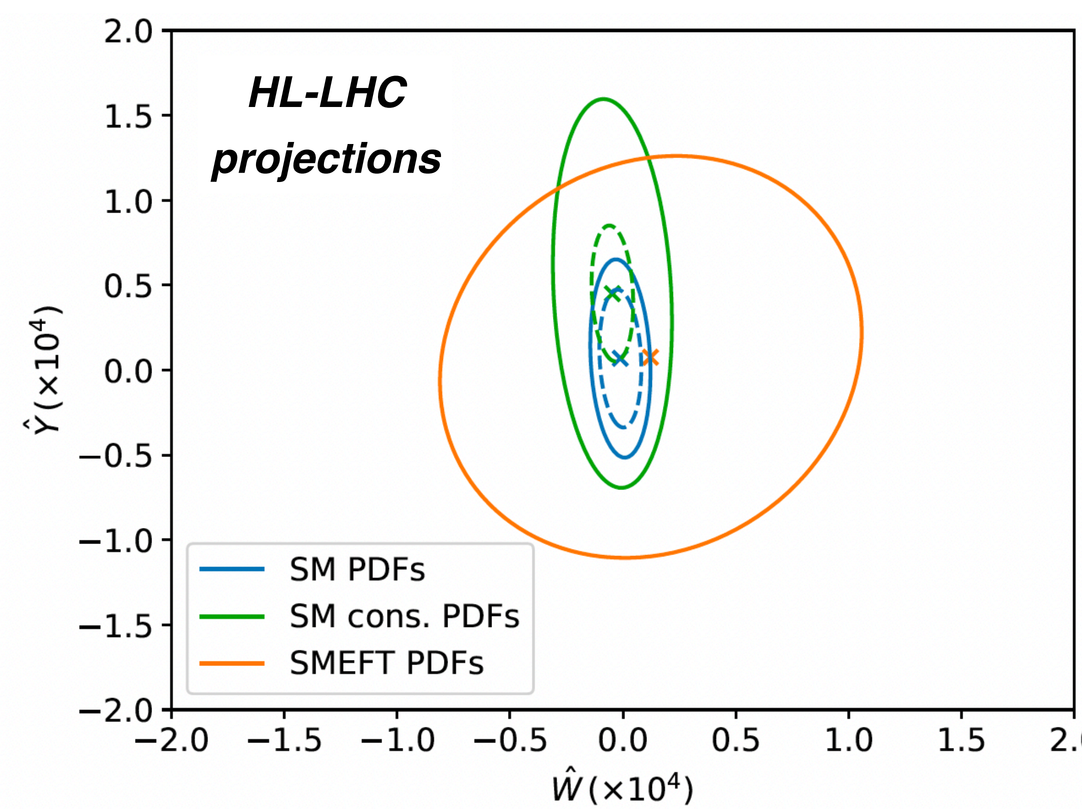
$$d\sigma_{\rm SMEFT} = d\sigma_{\rm SM} \times K_{\rm EFT}$$

$$K_{\rm EFT} = 1 + \sum_{n=1}^{n_{\rm op}} c_n R^{(n)}_{\rm SMEFT} + \sum_{n,m=1}^{n_{\rm op}} c_n c_m R^{(n,m)}_{\rm SMEFT}$$

Available data: limited interplay between PDF and EFT fits, best constraints from **searches**

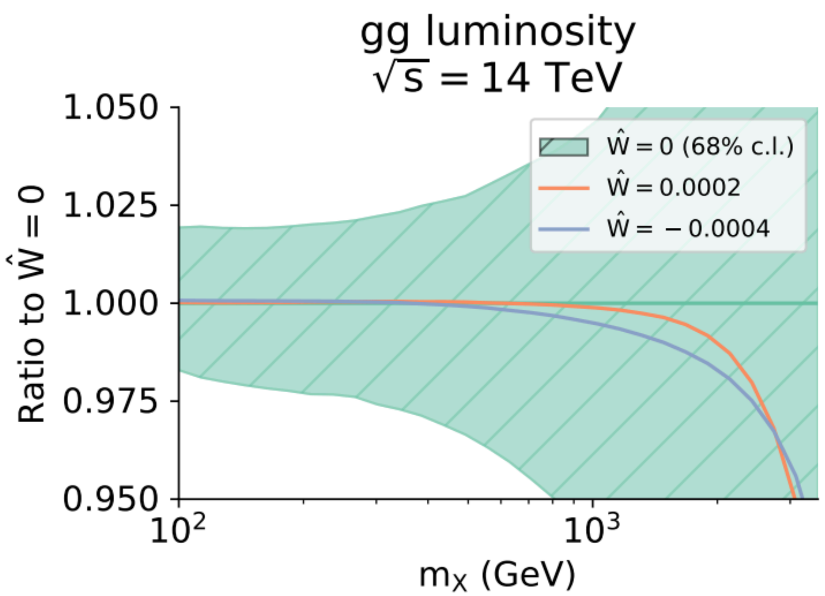
HL-LHC: EFT effects, if present, would be **reabsorbed into PDFs**

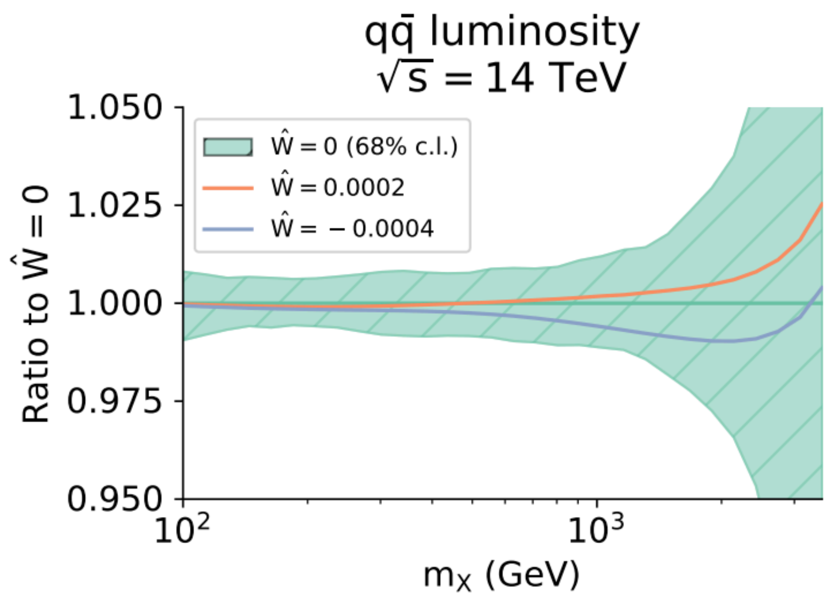


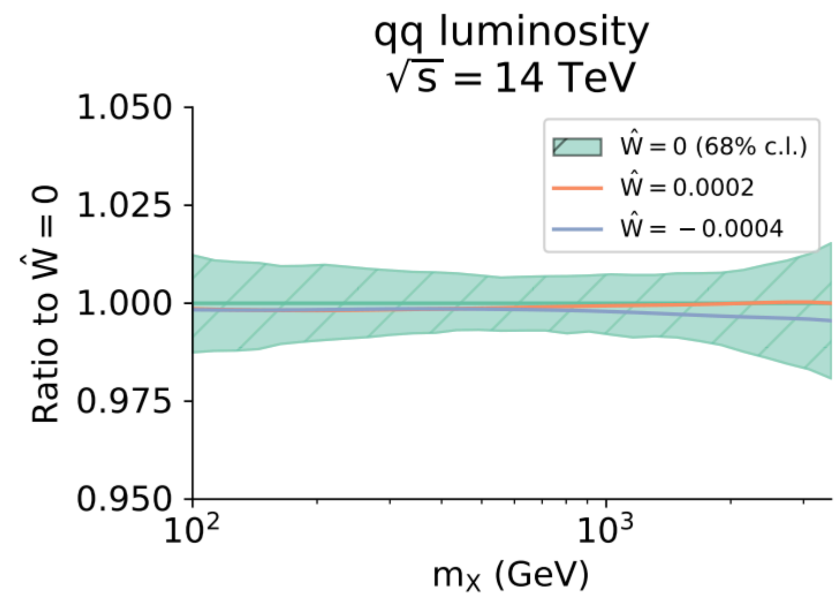


SMEFT PDFs from high-E Drell-Yan

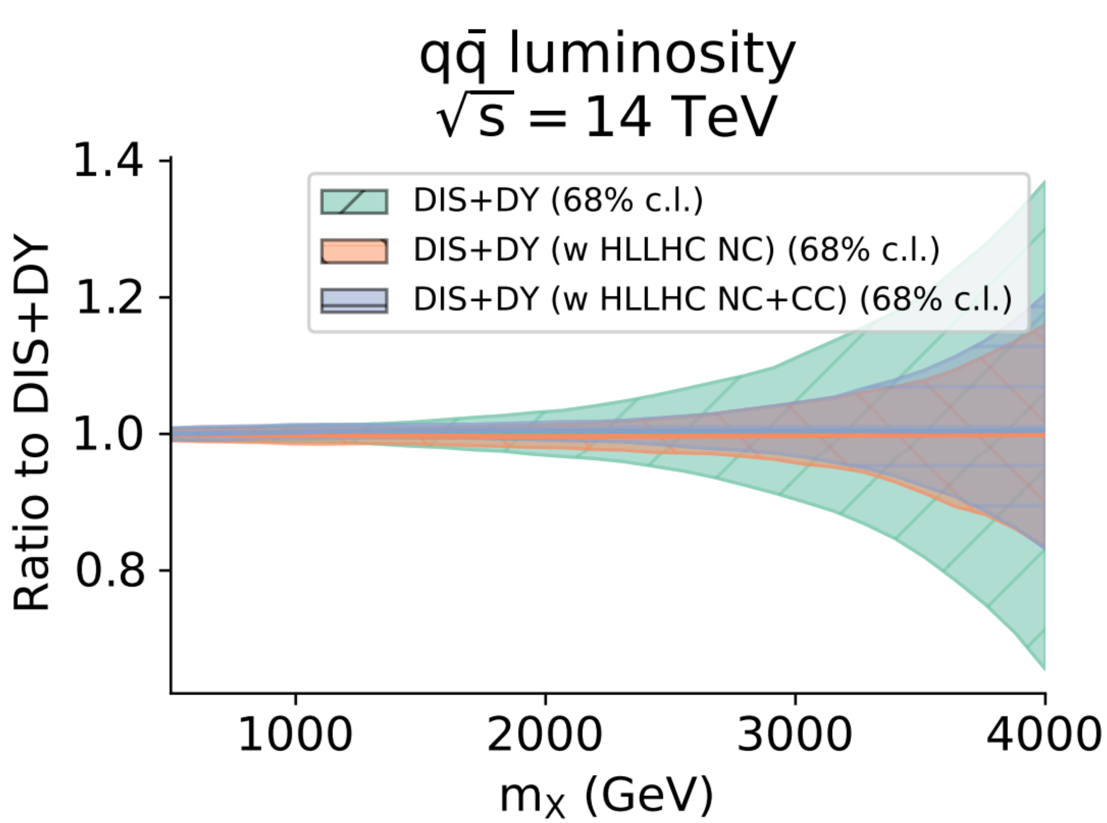
with current (published) DY data, interplay between PDF and EFT effects **moderate ….**

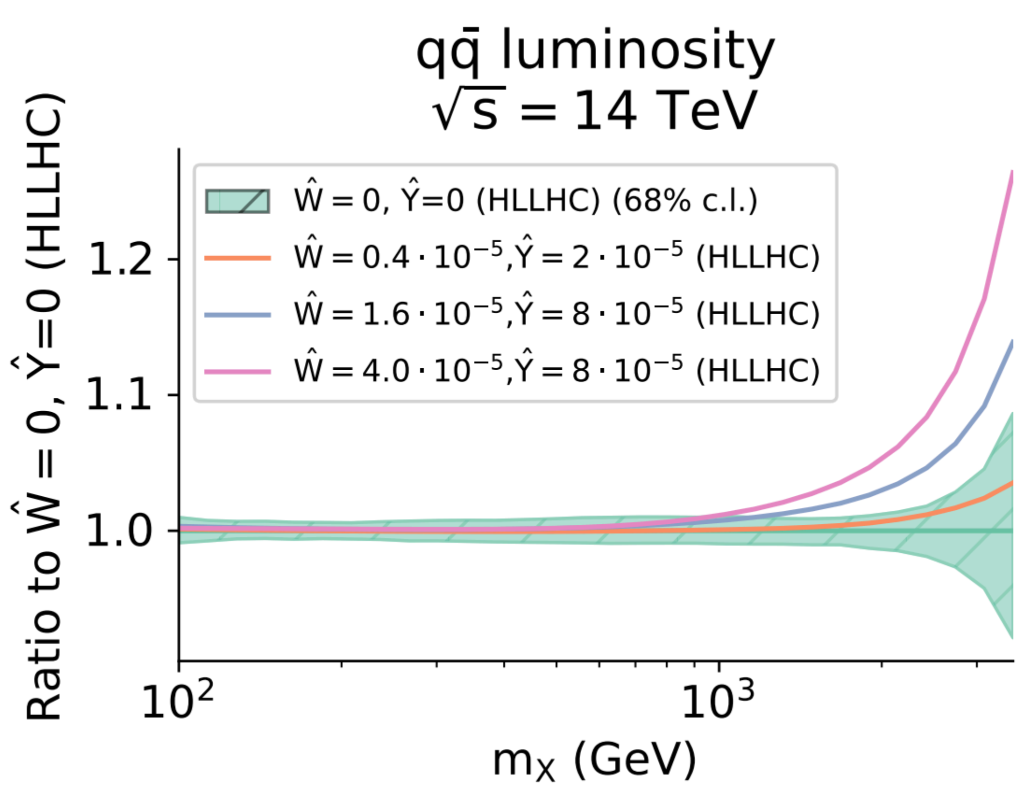






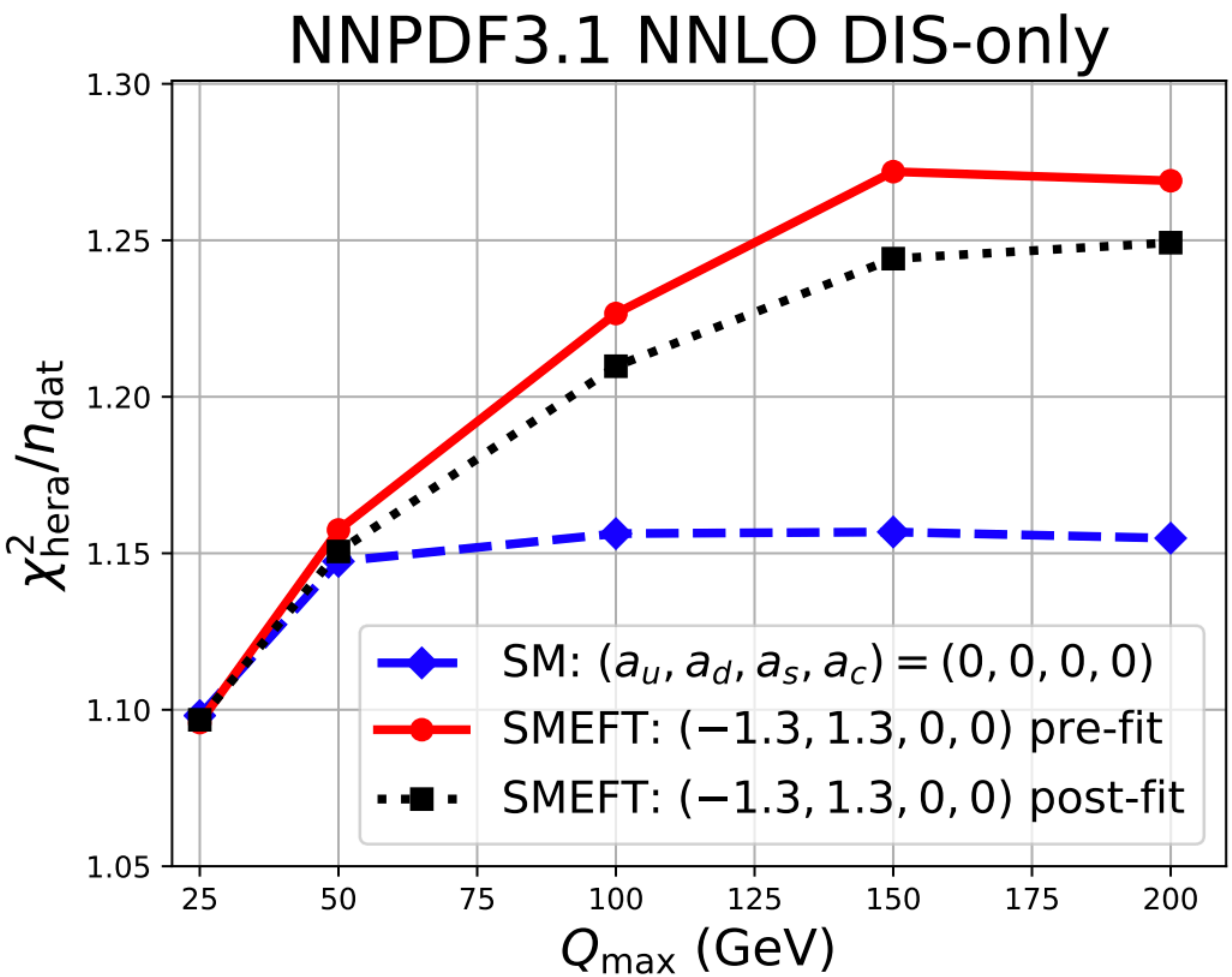
… while at the HL-LHC EFT effects may be **reabsorbed into the PDFs**: careful separation required





Fingerprinting EFT effects

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



Summary and outlook

- The EFT framework provides a robust strategy to interpret particle physics data in terms of new BSM phenomena while **minimising model assumptions**

- Only within a **global SMEFT interpretation** it is possible to compare with largest possible class of UV-complete theories and to reduce assumptions i.e. concerning flavour structure

- The SMEFiT framework has been successfully deployed for the **most extensive SMEFT analysis of LHC data** to date based on state-of-the-art EFT calculations

- Ongoing work includes adding more processes, constructing optimally-sensitive observables with ML, matching to UV complete models, accounting for flavour and low-energy constraints …

Summary and outlook

- The EFT framework provides a robust strategy to interpret particle physics data in terms of new BSM phenomena while **minimising model assumptions**

- Only within a **global SMEFT interpretation** it i… …st possible class of UV-complete theories and t… …structure

- The SM… …e **most extensive SMEFT analy**… …e-art EFT calculations

- Ongoin… …es adding more processes, constructing optimally-sensitive observables with ML, matching to UV complete models, accounting for flavour and low-energy constraints …

Thanks for your attention!