Machine Learning Opportunities for EFT Analyses

Maeve Madigan

Heidelberg University

UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

EFT WG 16.11.23

How to perform a global effective field theory analysis of LHC data?

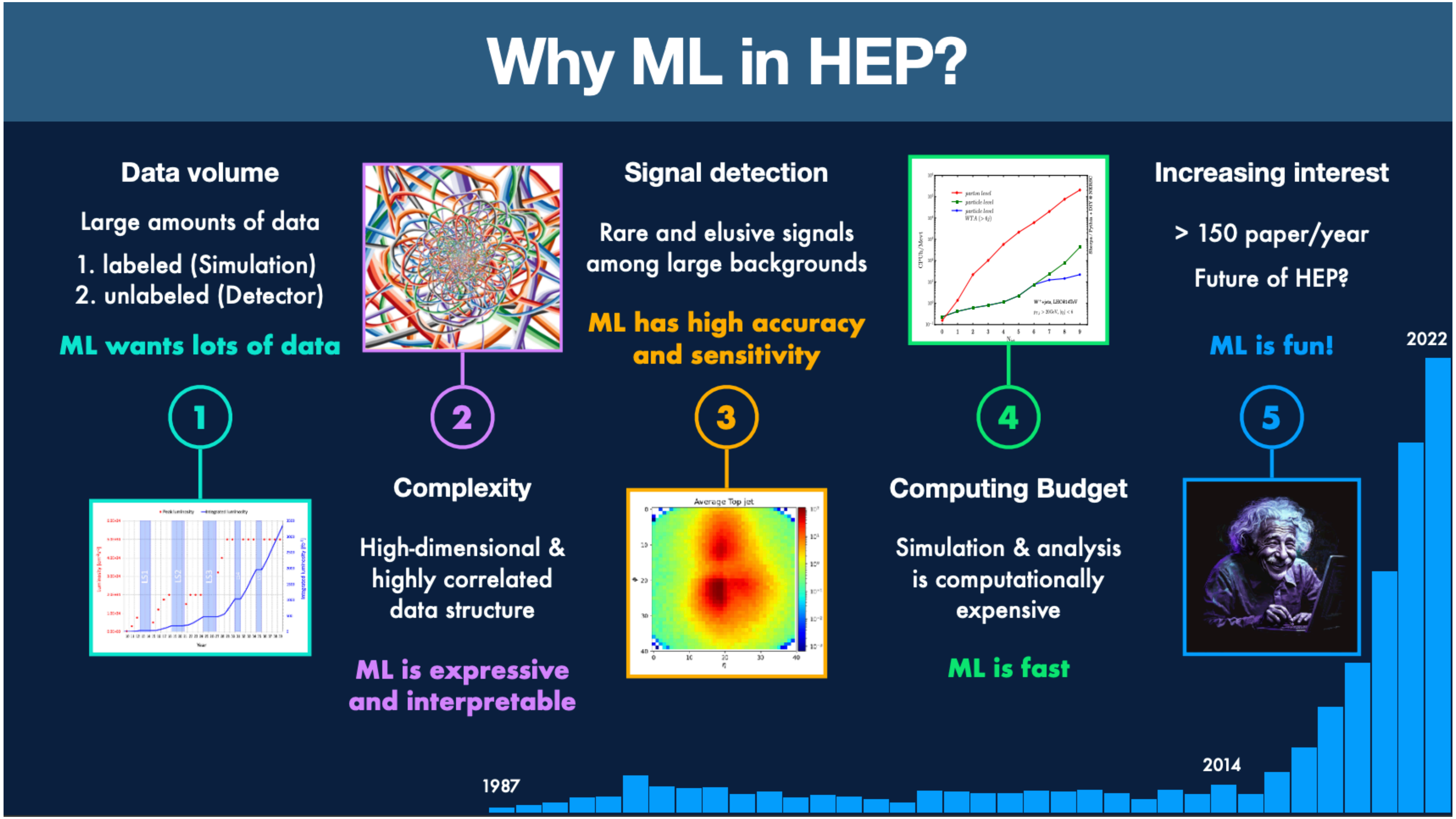
Performing a global Effective Field Theory (EFT) analysis of LHC data involves:

1. **Define EFT Framework:**
   - Choose an EFT framework relevant to the physics scenario of interest.
2. **Generate Simulated Data:**
   - Simulate expected LHC data using Monte Carlo methods based on the chosen EFT framework.
3. **Develop Analysis Pipeline:**
   - Create an analysis pipeline that includes event selection, background subtraction, and parameter estimation.
4. **Implement Machine Learning:**
   - Integrate machine learning for event classification, anomaly detection, and optimization of experimental design.
5. **Compare with Experimental Data:**
   - Compare the simulated data with actual LHC data, adjusting EFT parameters to match observed results.
6. **Statistical Analysis:**
   - Perform statistical analyses to quantify the agreement between the EFT predictions and experimental data.
7. **Iterative Refinement:**
   - Iterate through steps 2-6, refining the analysis based on feedback and incorporating additional data.
8. **Collaboration and Peer Review:**
   - Engage in collaboration with experts, undergo peer review, and ensure the transparency and reproducibility of the analysis.
9. **Publication and Interpretation:**
   - Publish the results, interpret the findings in the context of the chosen EFT, and contribute to the global understanding of particle physics.

Global SMEFT analyses of LHC data can benefit from machine learning

(according to machine learning)

From Ramon Winterhalder's talk at Pheno 2023



Machine Learning Opportunities for EFT Analyses : this talk

Overview of recent ML/EFT studies

ML4EFT: unbinned multivariate observables for global SMEFT fits

SIMUnet: global SMEFT and PDF determinations

Overview of recent ML/EFT studies

LHC EFT WG - Area 3: Observables - Opportunities and Challenges of Machine Learning for EFT analyses

Tuesday 24 Oct 2023, 15:00 → 16:40 Europe/Zurich

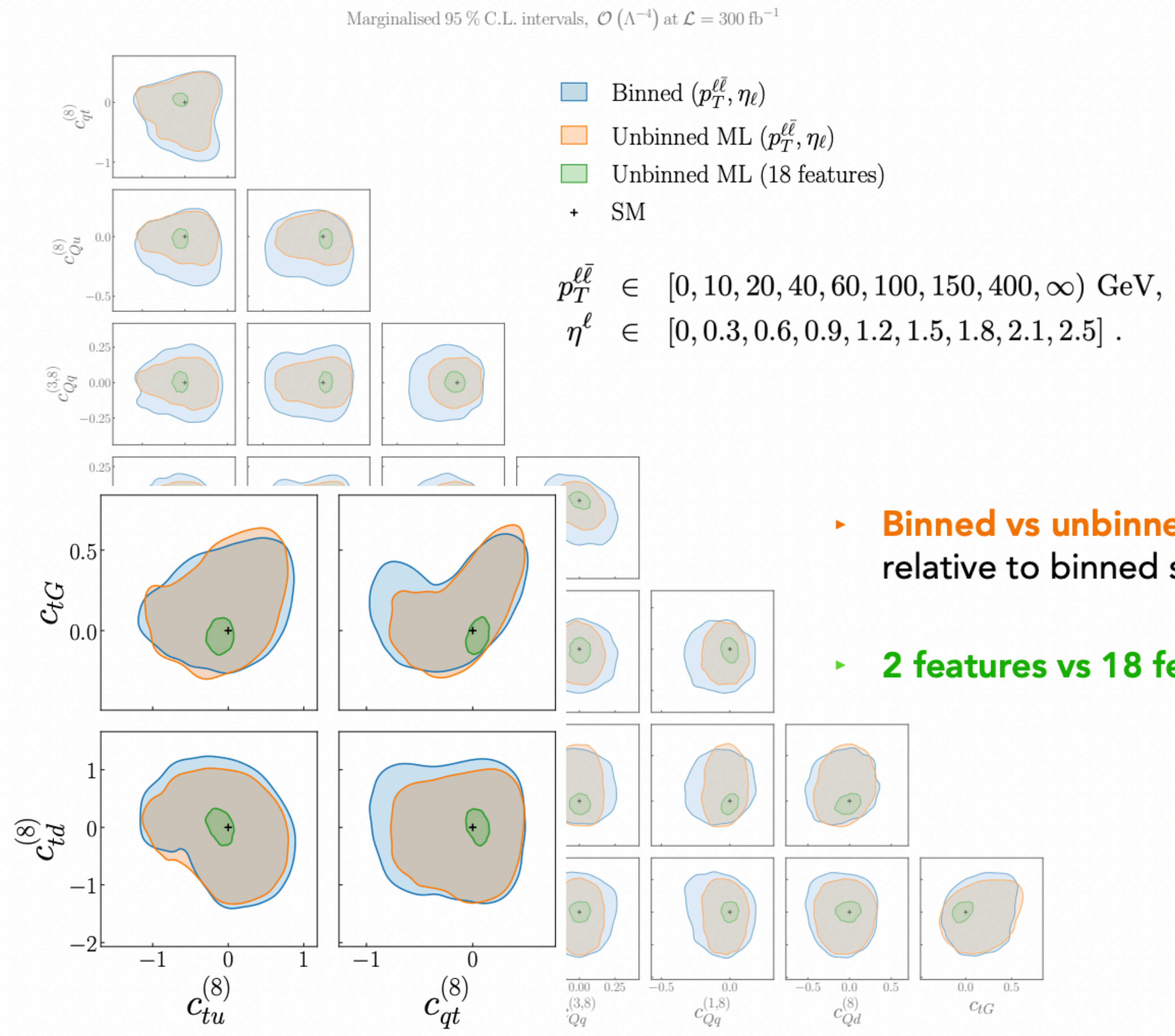
Statistically optimal observables for global SMEFT fits

Jaco ter Hoeve

2211.02058

[Link to slides]

Unbinned observables in the top sector



- **Binned vs unbinned** in $(p_T^{\ell\bar{\ell}}, \eta_\ell)$ small improvement relative to binned setup
- **2 features vs 18 features:** big increase in sensitivity

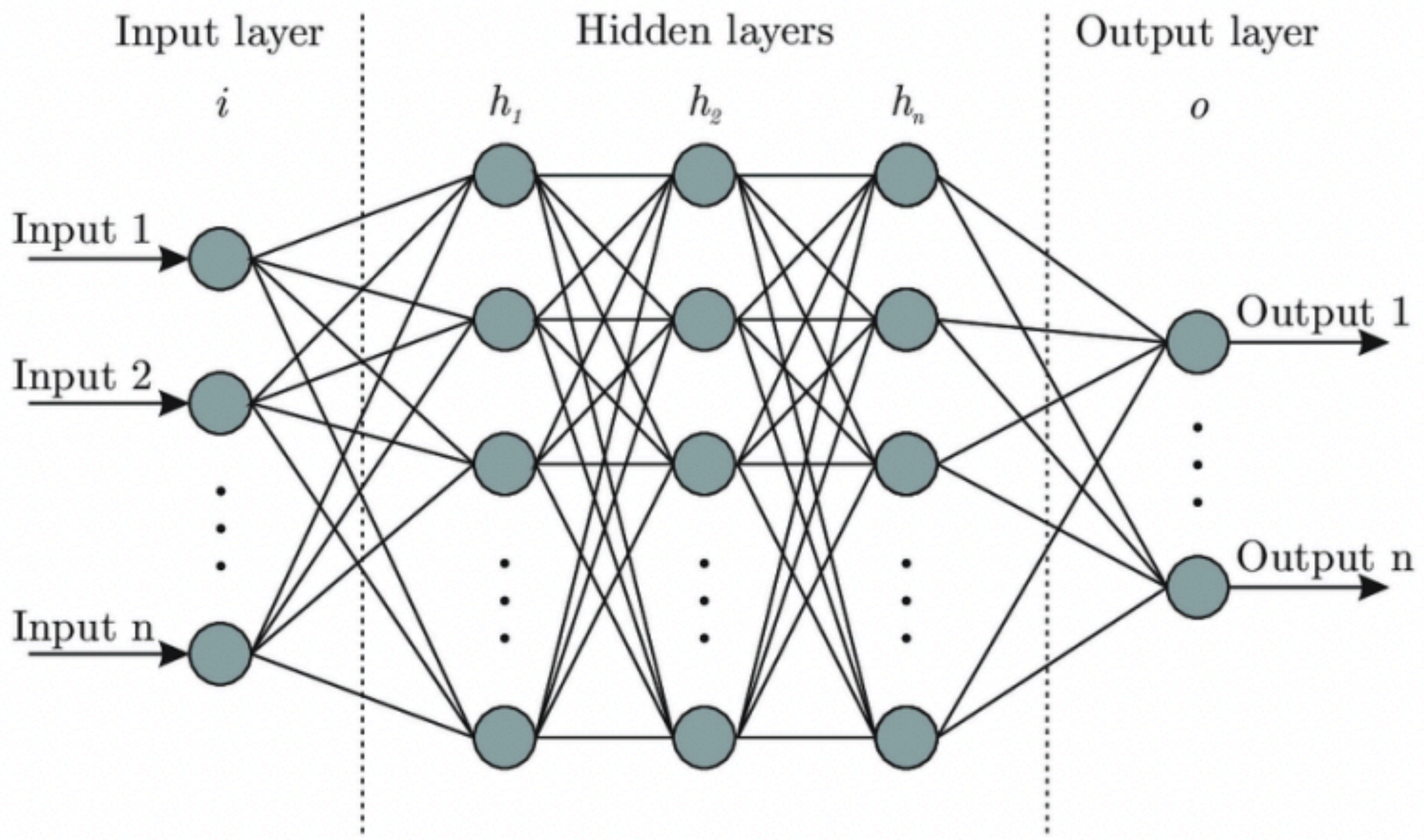
Unbinned MVA techniques for EFT analyses

Alfredo Glioti

2007.10356, 2308.05704

[Link to slides]

Basic idea: approximate $p(x|\theta)$ with **Neural Networks**: $p(x|\theta) \leftrightarrow nn(x;w)$



The result will be **fully differential** on **all observables**, quick to evaluate and it can be obtained with a relatively small amount of Monte Carlo points.

No transfer functions modeling required.

Universal and systematically improvable

$$p(x|c) = \frac{1}{\sigma}\frac{d\sigma}{dx}(c)$$

- Multivariate in all features *x*
- Extract full information on relation between data *x* and Wilson coefficients *C*
- Optimal constraints on the EFT

Unbinned MVA techniques for EFT analyses

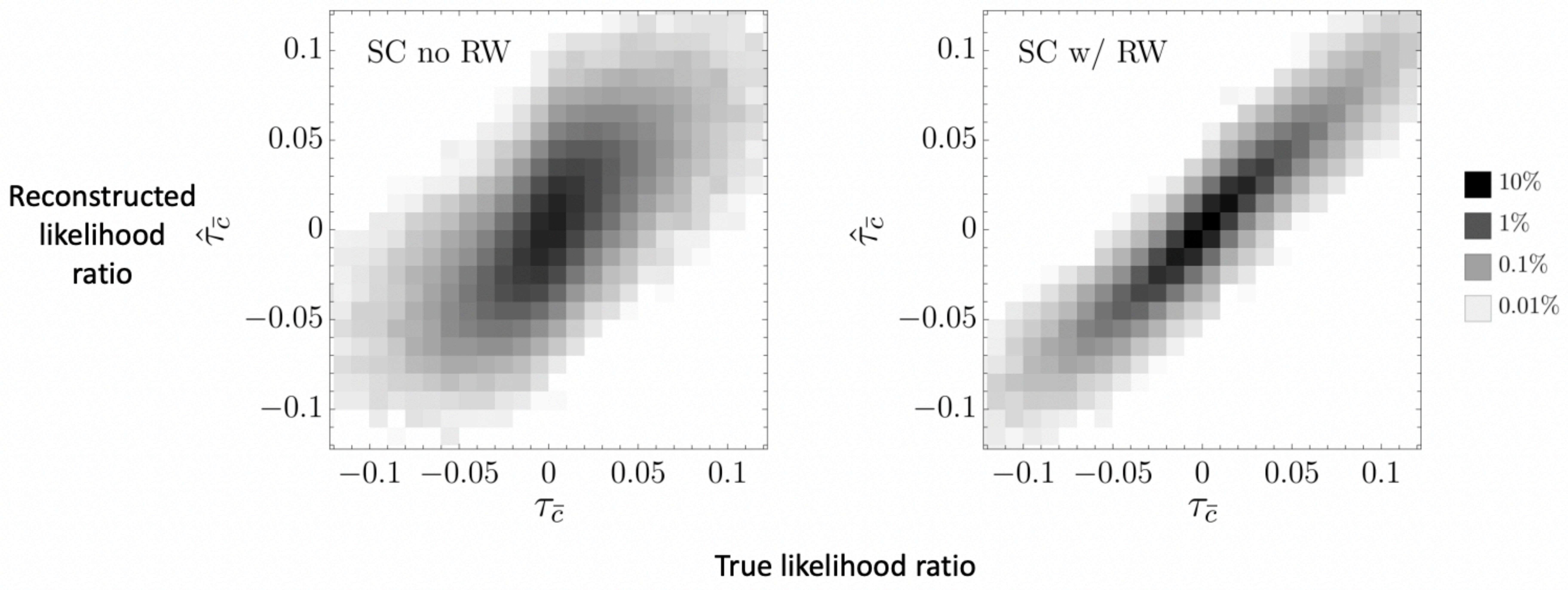
Alfredo Glioti

2007.10356, 2308.05704

Link to slides

Training on reweighted samples reduces number of training points needed and leads to a higher accuracy

Without reweighting (left) vs with reweighting (right):



TREE BOOSTING (+NEURAL NETWORKS) FOR EFT ANALYSES

R. Schöfbeck (HEPHY Vienna), Oct.. 24th, 2023, Area 3 meeting

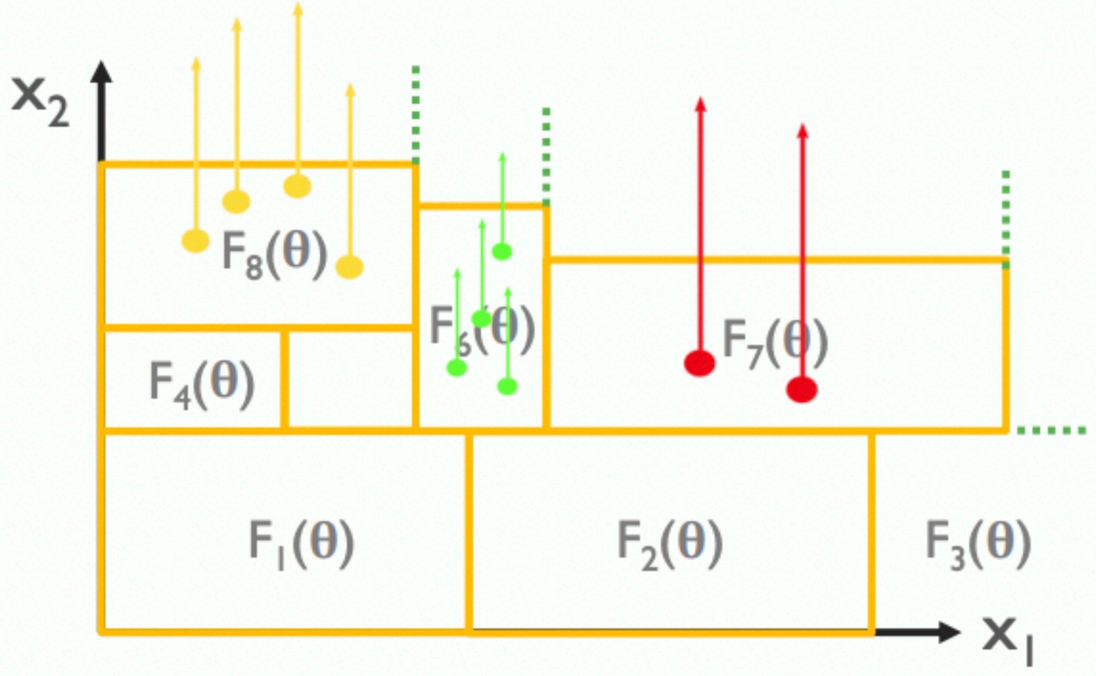
Robert Schöfbeck

[Link to slides]

A SIMPLE TREE ALGORITHM

[arXiv:2107.10859, arXiv:2205:12976]

phase-space partitioning

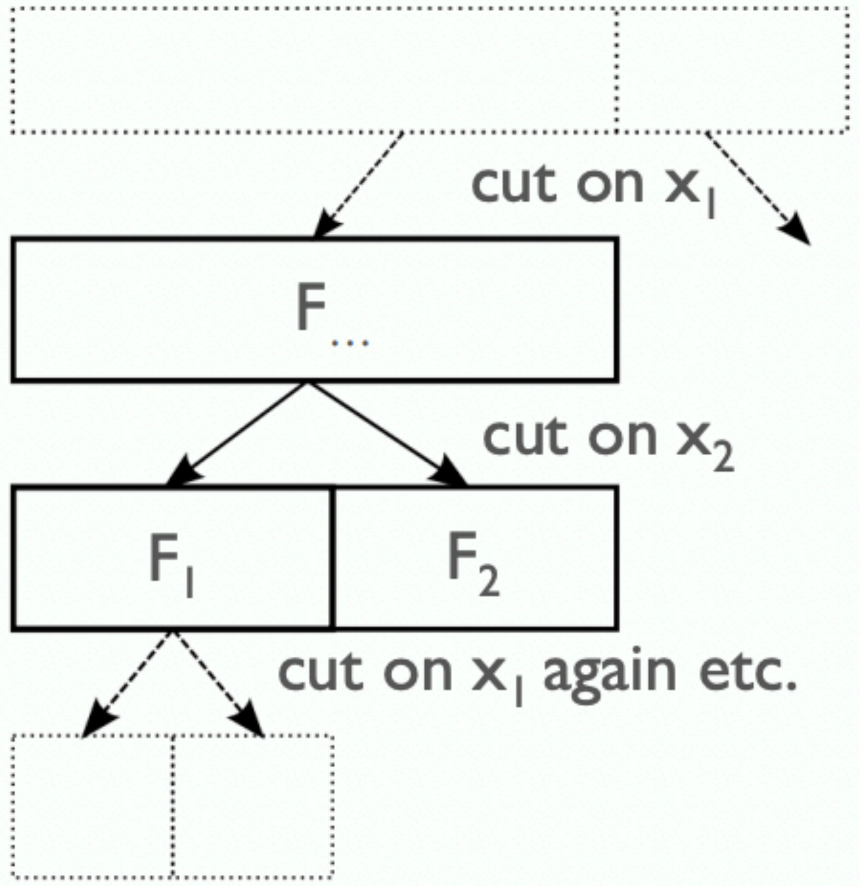


A simple tree

$$\hat{F}(\boldsymbol{x},\boldsymbol{\theta}) = \underbrace{\sum_{j\in\mathcal{J}} \overbrace{\mathbb{1}_j(\boldsymbol{x})}^{\text{index-function (non-linearity)}}}_{\text{phase space partitioning J}} \underbrace{F_j(\boldsymbol{\theta})}_{\text{prediction } F_j}$$

need to solve for partitioning J and $\{F_j\}$

training phase:
e.g. "CART" algo



- A tree is a hierarchical phase-space partitioning ($\mathcal{J}$)
  - the novelty in the Boosted Information Tree is that we associate each region j with a polynomial $F_j(\boldsymbol{\theta})$
  - Note: A tree algorithm can have an arbitrarily complicated predictive function; here it is a SMEFT polynomial
- Fitting tree: Optimize "node split positions" on some loss. Trained (e.g. greedily) on the *ensemble*.

7

TREE BOOSTING (+NEURAL NETWORKS) FOR EFT ANALYSES

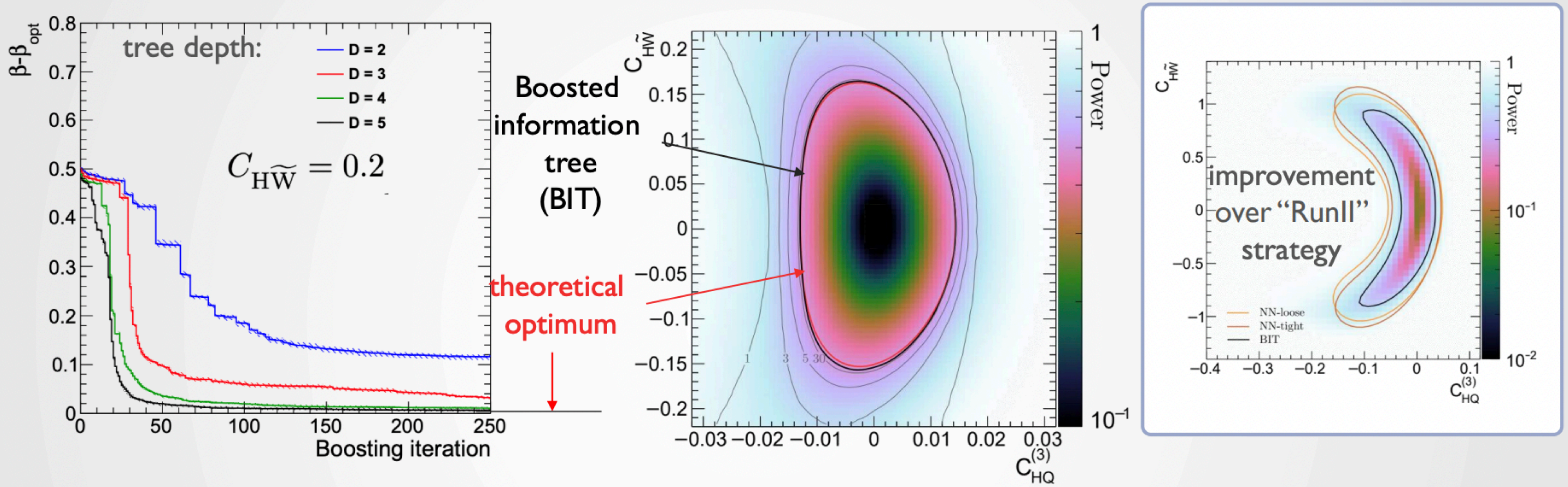
R. Schöfbeck (HEPHY Vienna), Oct.. 24th, 2023, Area 3 meeting

Robert Schöfbeck

[Link to slides]

OPTIMALITY IN TEST CASES

[arXiv:2107.10859, arXiv:2205:12976]



- Obtain parametrized classifiers with 20-40% improvements in 2D toy cases (NOT marginalized!)
- No free lunch – Analysis dependent choices are needed
  - Binned analysis: variable binning → background estimation is CPU intensive
  - Systematics treatment for unbinned analyses (beyond Higgs $M_{4\ell}$) less far developed
- Is it all worth it in higher dimensions? Yes! More examples: [ML4EFT]; full list of references in backup

12

Machine Learning for Higgs CP properties

António Jacques Costa

2112.05052

Link to slides

- CP violating operators affecting the Higgs/multiboson interactions can be probed via Effective Field Theory:

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

c_i/Λ^2 the Wilson coefficients,
Λ the scale of new physics

$$\widetilde{\mathcal{O}}_{\Phi\widetilde{B}} = \Phi^\dagger\Phi B^{\mu\nu}\widetilde{B}_{\mu\nu}\,,$$
$$\widetilde{\mathcal{O}}_{\Phi\widetilde{W}} = \Phi^\dagger\Phi W^{i\,\mu\nu}\widetilde{W}^i_{\mu\nu}\,,$$
$$\widetilde{\mathcal{O}}_{\Phi\widetilde{W}B} = \Phi^\dagger\sigma^i\widetilde{W}^{i\,\mu\nu}B_{\mu\nu}\,.$$

- Beyond-the-SM amplitude is then given by: $|\mathcal{M}_{\rm BSM}|^2 = |\mathcal{M}_{\rm SM}|^2 + \boxed{2{\rm Re}\{\mathcal{M}_{\rm SM}\mathcal{M}^*_{\rm d6}\}} + |\mathcal{M}_{\rm d6}|^2$
- Interference term leads to asymmetries in CP-odd observables

- Possible CP-odd observables
  - Statistically optimal observables
  - Angular observables – less sensitivity but easier implementation
  - Machine-learning observables – attempt to recover best sensitivity while keeping feasibility

Machine Learning for Higgs CP properties

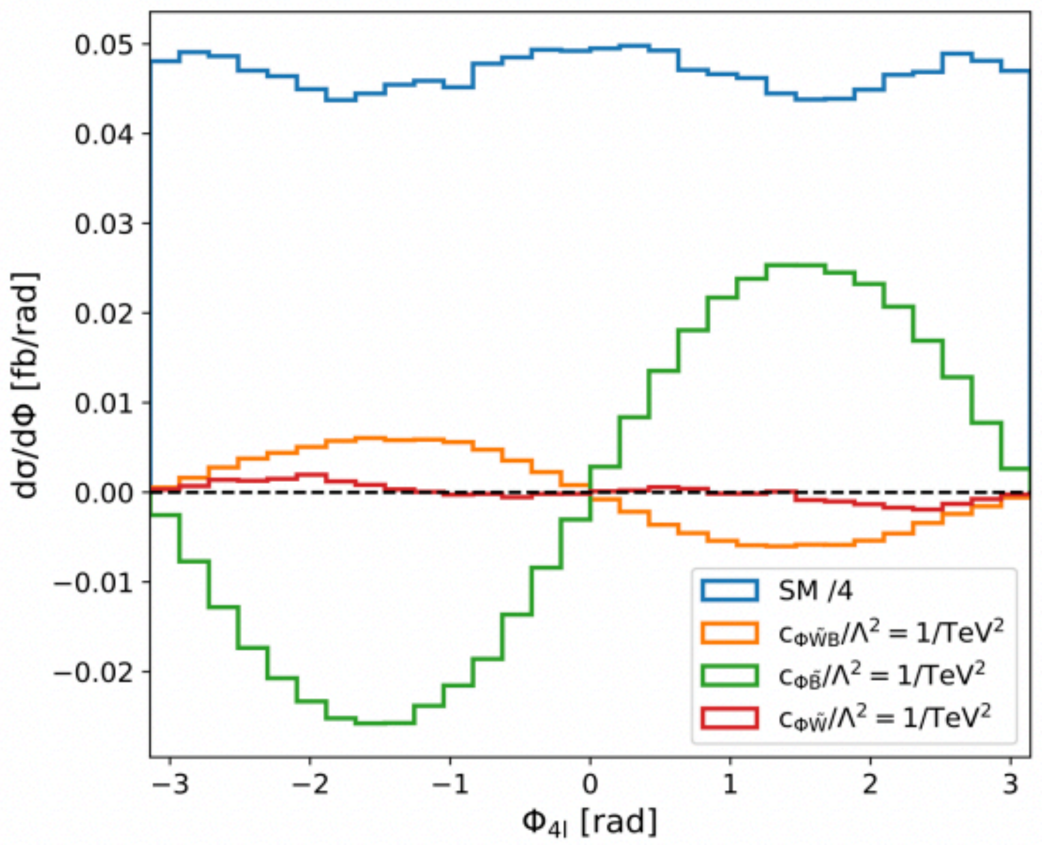
António Jacques Costa

2112.05052

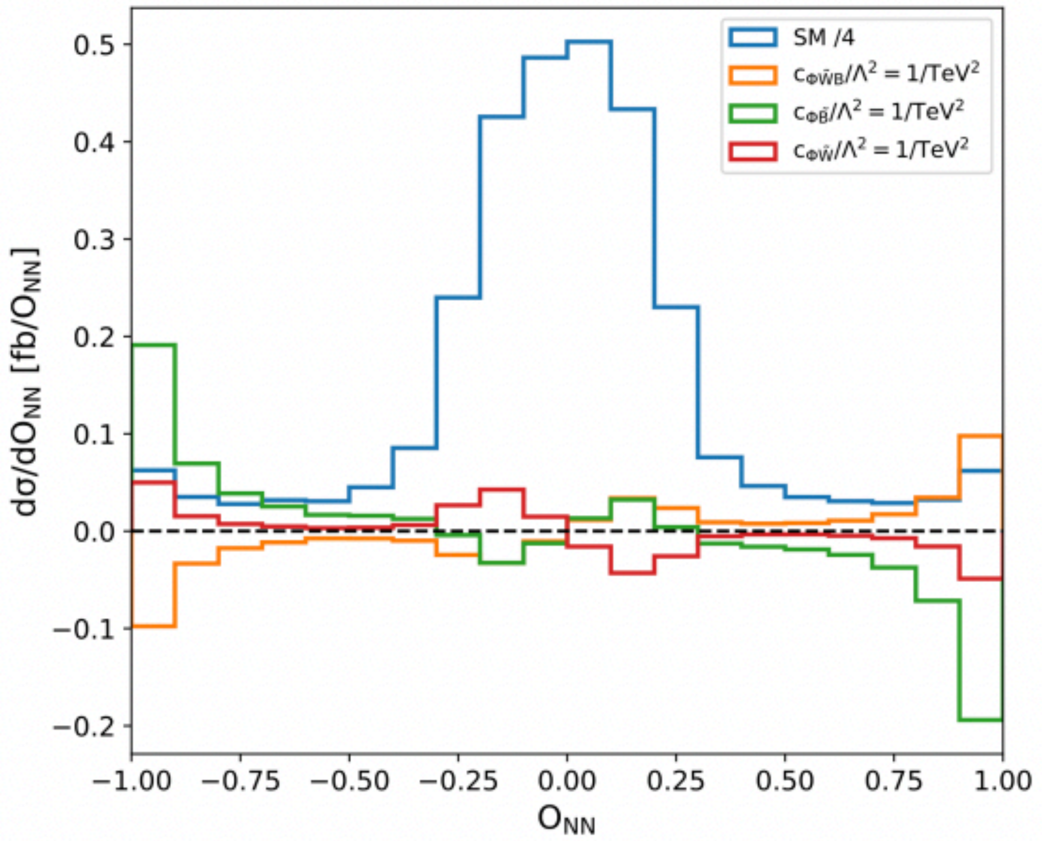
Link to slides

Neural network-based observable: $h \to 4\ell$ results

$\phi_{4\ell}$



O_{NN}



- Expected 95% confidence intervals for the three Wilson coefficients given an integrated luminosity of 139 fb^{-1}

CP-odd observable	$c_{\Phi\tilde{W}B}/\Lambda^2$ [TeV^{-2}]	$c_{\Phi\tilde{B}}/\Lambda^2$ [TeV^{-2}]	$c_{\Phi\tilde{W}}/\Lambda^2$ [TeV^{-2}]
$\Phi_{4\ell}$	$[-6.2, 6.2]$	$[-1.4, 1.4]$	$[-30, 30]$
$\Phi_{4\ell}, m_{12}$	$[-1.9, 1.9]$	$[-0.85, 0.85]$	$[-3.7, 3.7]$
O_{NN} (binary)	$[-1.5, 1.5]$	$[-0.75, 0.75]$	$[-3.0, 3.0]$
O_{NN} (multi-class)	$[-1.4, 1.4]$	$[-0.71, 0.71]$	$[-2.7, 2.7]$

- Factor 2 to 10 improvement using O_{NN} in sensitivity to Wilson coefficients
- Considerable gain can be recovered by two-dimensional fit to $\phi_{4\ell}$ and m_{12}

Reusing Neural Networks: Experiences and suggestions for EFT cases

Tomasz Procter

Link to slides

Initial experiences with reinterpretation

★

- So far: two publicly available LHC analysis NN networks - both from ATLAS SUSY:
  - ANA-SUSY-2019-04 (RPV SUSY search in lep+jets final state)
  - ANA-SUSY-2018-30 (gluino pair production in multi-b final states)
- This is a new type of experimental output - the experiments are still feeling out how to publicise this.
- ANA-SUSY-2018-30 has worked well in multiple frameworks (rivet, gambit, checkmate, ...).
- Several key features made this work:
  - Lots of extra info (ordering, units, usage example) - would have been impossible without SimpleAnalysis[1].
  - All inputs are easily accessible to reinterpretation tools
    - Lepton/jet kinematics, MET, btag yes/no
    - No detector-level variables (including continuous btag score)

Cut	Paper	Rivet
0-lep	80.0	83.7
$\Delta\phi^{4j}_{min}\geq0.6$	52.5	54.6
2800-1400 NN Cut	21.7	23.9
$\Delta\phi^{4j}_{min}\geq0.6$	52.5	54.6
2300-1000 NN Cut	21.3	23.3
$\Delta\phi^{4j}_{min}\geq0.4$	61.1	63.8
2100-1600 NN Cut	6.20	6.50
$\Delta\phi^{4j}_{min}\geq0.4$	61.1	63.8
2000-1800 NN Cut	0.192	0.204

Tomasz Procter, LHC EFT WG Area 3, October 2023

[1] Simplified ATLAS SUSY analysis framework

5

Reusing Neural Networks: Experiences and suggestions for EFT cases

Tomasz Procter

Link to slides

Initial experiences with reinterpretation

- So far: two publicly available LHC analysis NN networks - both from ATLAS SUSY:
  - ANA-SUSY-2019-04 (RPV SUSY search in lep+jets final state)
  - ANA-SUSY-2018-30 (gluino pair production in multi-b final states)
- This is a new type of experimental output - the experiments are still feeling out how to publicise this.

Cut	Paper	Rivet
0-lep	80.0	83.7
$\Delta\phi^{4j}_{min} \geq 0.6$	52.5	54.6
2800-1400 NN Cut	21.7	23.9
$\Delta\phi^{4j}_{min} \geq 0.6$	52.5	54.6

- Key rule:

  Reinterpretation is easiest when the analysis team think about it from the start
  - Make sure models can be saved in a preservable format.
  - Example code snippets, metadata is very important.
  - Think about choice of inputs:
    - Do we need to use efficiencies/surrogates instead?

ML4EFT: unbinned multivariate observables for global SMEFT fits

2211.02058 Raquel Gomez Ambrosio, Jaco ter Hoeve, MM, Juan Rojo, Veronica Sanz

Why Unbinned Measurements?

'Presenting Unbinned Differential Cross Section Results', Arratia et al, 2109.13243

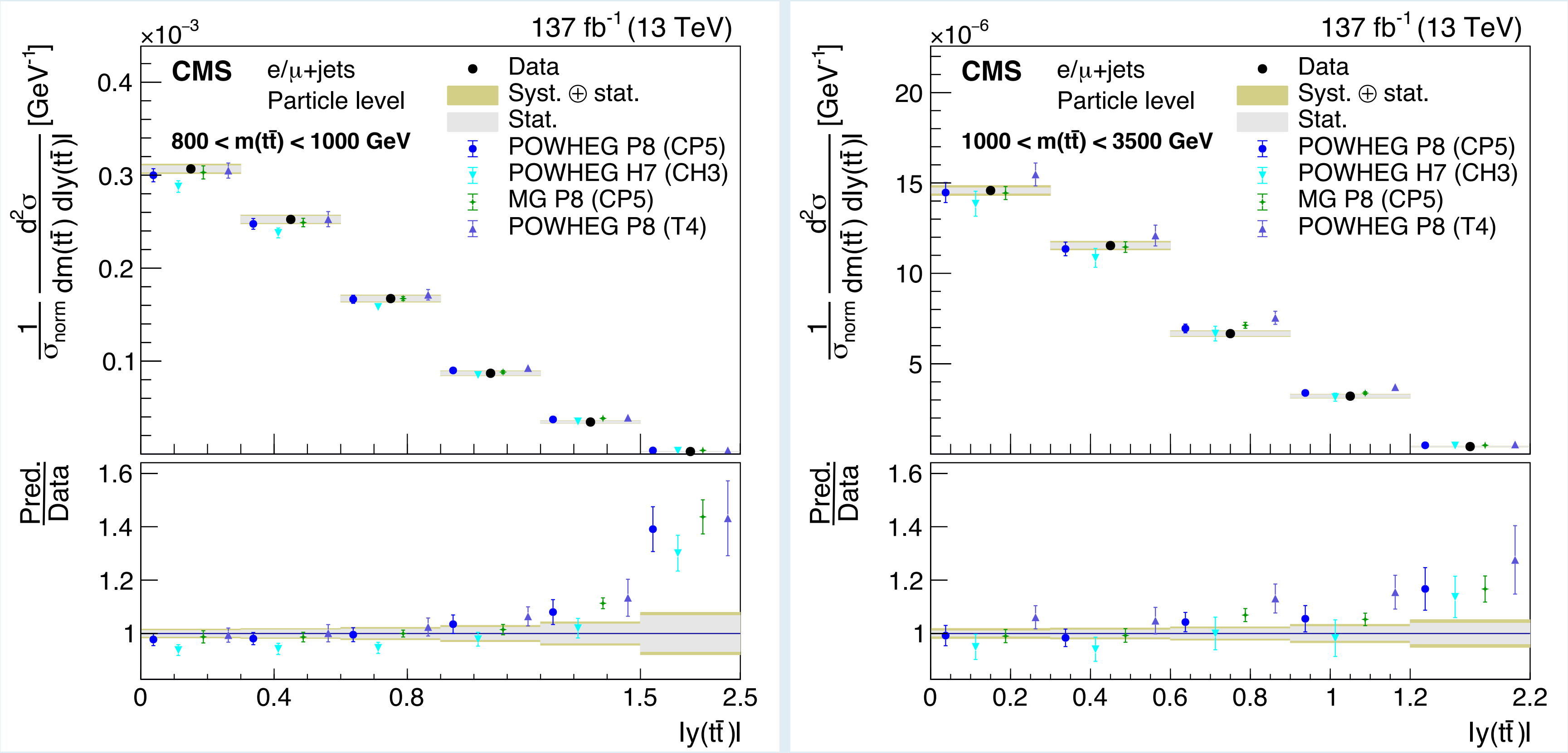
1. **Inference-aware binning:**
   - optimal choice of binning can be made at the time of *each* statistical analysis or global fit

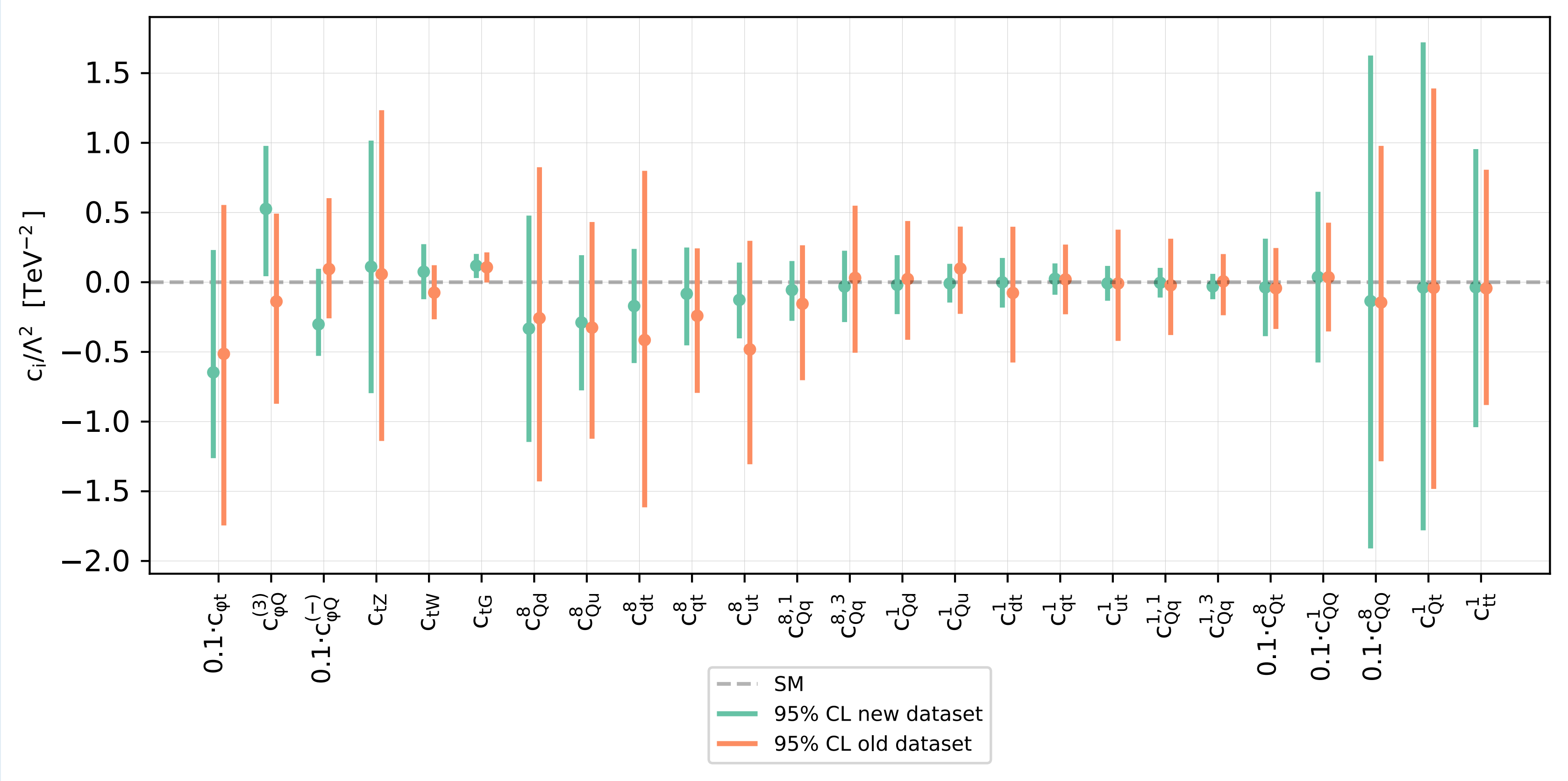
Why Unbinned Measurements?

'Presenting Unbinned Differential Cross Section Results', Arratia et al, 2109.13243

1. **Inference-aware binning:**
   - optimal choice of binning can be made at the time of *each* statistical analysis or global fit

Typically we **reinterpret** measurements optimised for SM measurements or NP resonance searches





e.g. CMS measurement of top pair production in the l+jets channel 2108.02803

Why Unbinned Measurements?

'Presenting Unbinned Differential Cross Section Results', Arratia et al, 2109.13243

1. **Inference-aware binning:**
   - optimal choice of binning can be made at the time of *each* statistical analysis or global fit

2. **Derivative measurements:**
   - given measurements of features $x_1, ..., x_n$, 'post-hoc measurement' of $f(x_1, ..., x_n)$ possible

Why Unbinned Measurements?

'Presenting Unbinned Differential Cross Section Results', Arratia et al, 2109.13243

1. **Inference-aware binning:**
   - optimal choice of binning can be made at the time of *each* statistical analysis or global fit

2. **Derivative measurements:**
   - given measurements of features $x_1, ..., x_n$, 'post-hoc measurement' of $f(x_1, ..., x_n)$ possible

3. **Extension to higher dimensions:**
   - ML-based unbinned unfolding techniques well-suited to multiple features

ML4EFT

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

2211.02058 Raquel Gomez Ambrosio, Jaco ter Hoeve, MM, Juan Rojo, Veronica Sanz

Open-source NN-based python framework for the integration of **unbinned multivariate observables** into global SMEFT interpretations.

Goal: to provide optimal constraints on the SMEFT

ML4EFT

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

2211.02058 Raquel Gomez Ambrosio, Jaco ter Hoeve, MM, Juan Rojo, Veronica Sanz

Open-source NN-based python framework for the integration of **unbinned multivariate observables** into global SMEFT interpretations.

Goal: to provide optimal constraints on the SMEFT

Diagnostic tool:

What is the information loss given a particular choice of bins?

Projections:

If unbinned data are made available, how will SMEFT constraints improve?

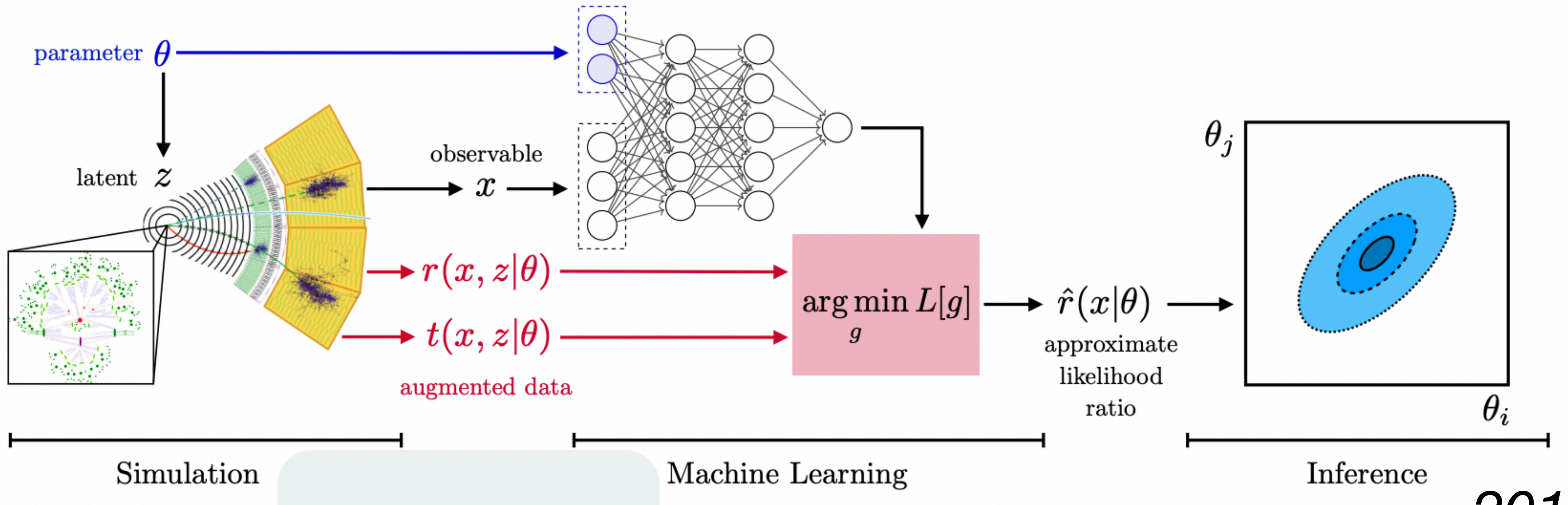
ML4EFT

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

2211.02058 Raquel Gomez Ambrosio, Jaco ter Hoeve, MM, Juan Rojo, Veronica Sanz

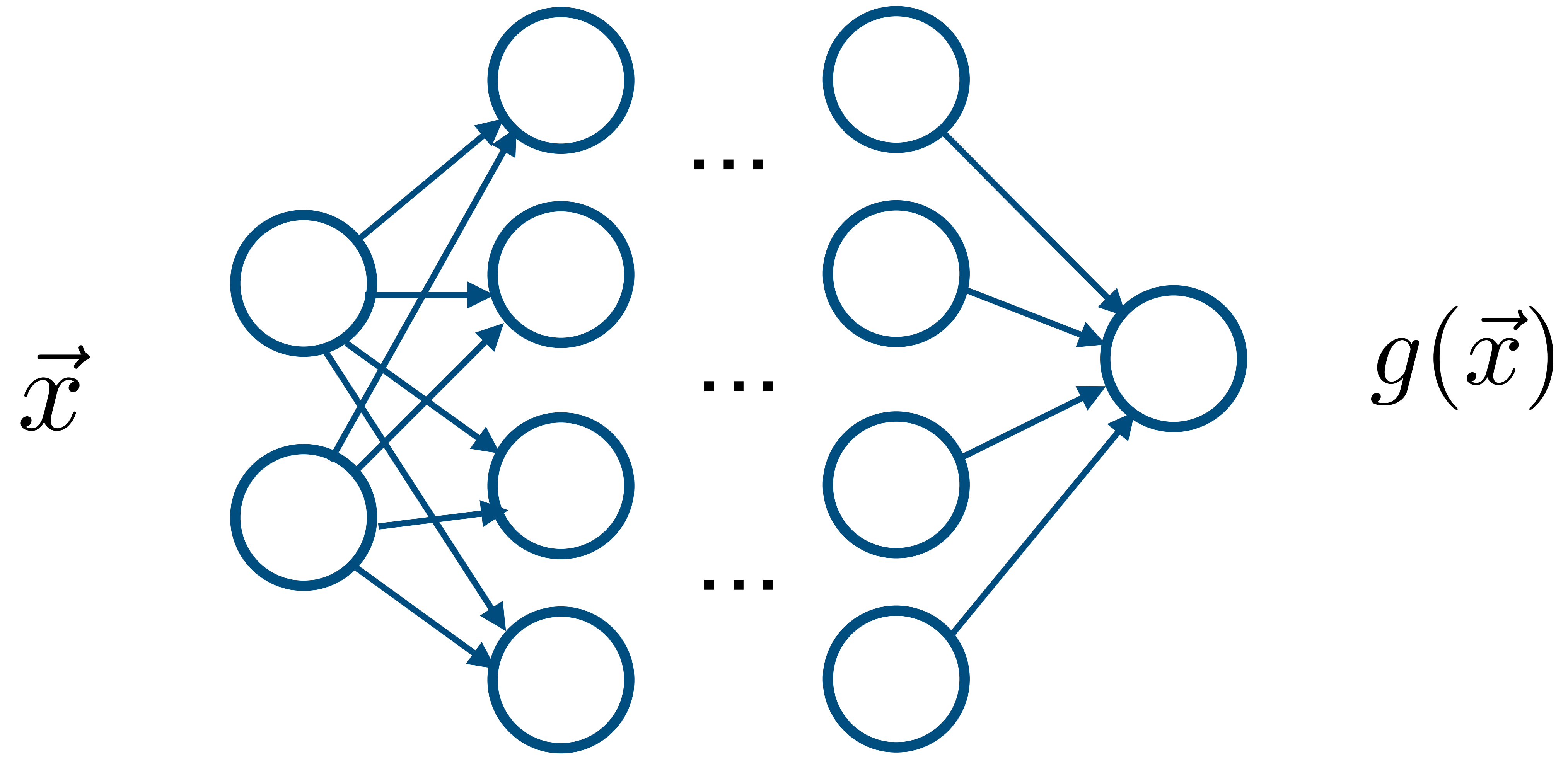
Open-source NN-based python framework for the integration of **unbinned multivariate observables** into global SMEFT interpretations.

Related work:
- 2007.10356 *Parameterized classifiers for SMEFT* A. Glioti et al.
- 2308.05704 *Boosted likelihood learning with event reweighting* A. Glioti et al
- 2205.12976 *Learning the EFT likelihood with tree boosting* R. Schöfbeck et al
- *MadMiner* (J.Brehmer, K.Cranmer, G.Louppe et al.) [1907.10621, 1805.00020, …]
- + many others



2010.06439

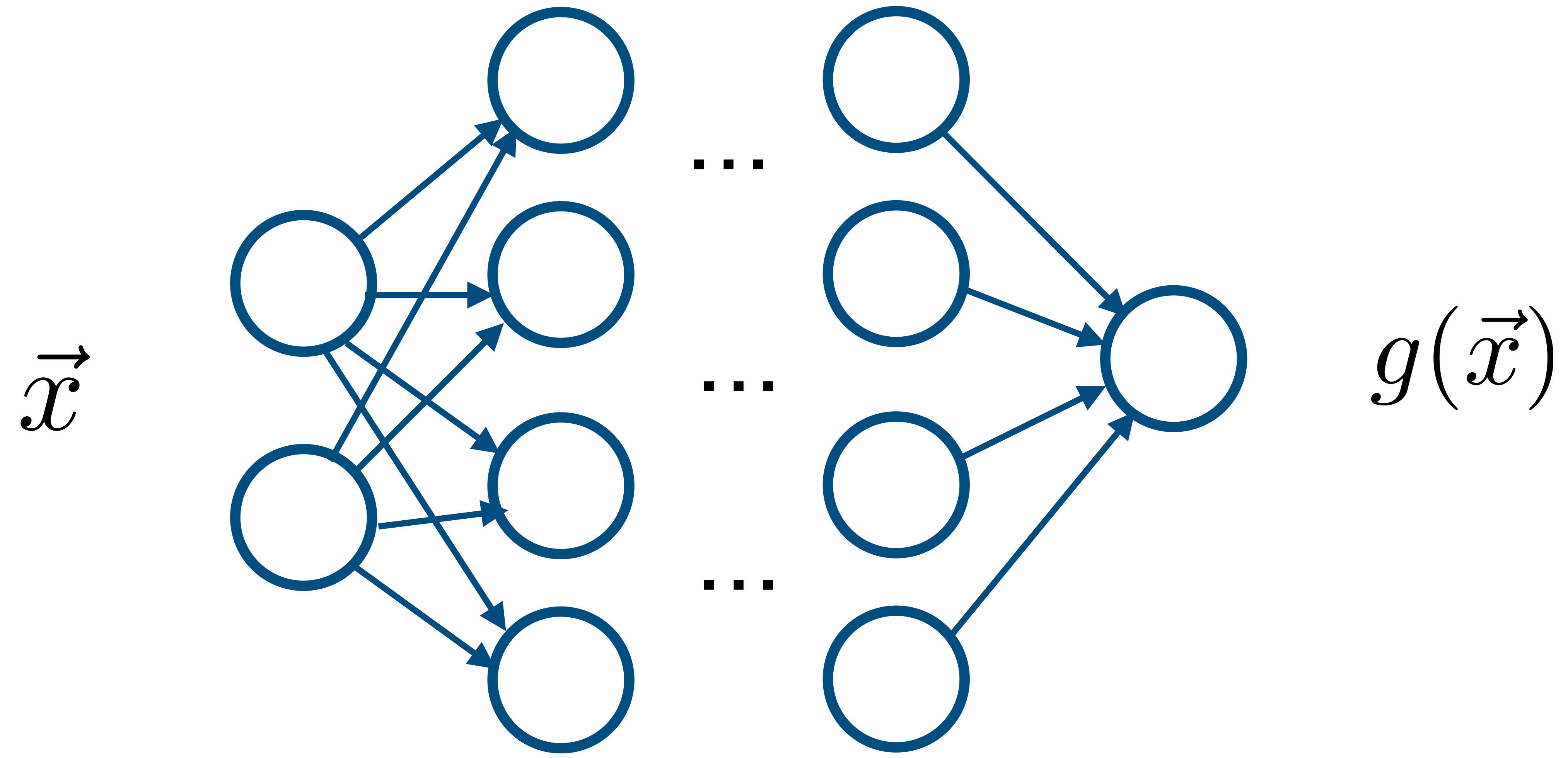
The ‘classifier trick’



$g(\vec{x}) = 0$

$g(\vec{x}) = 1$

The ‘classifier trick’

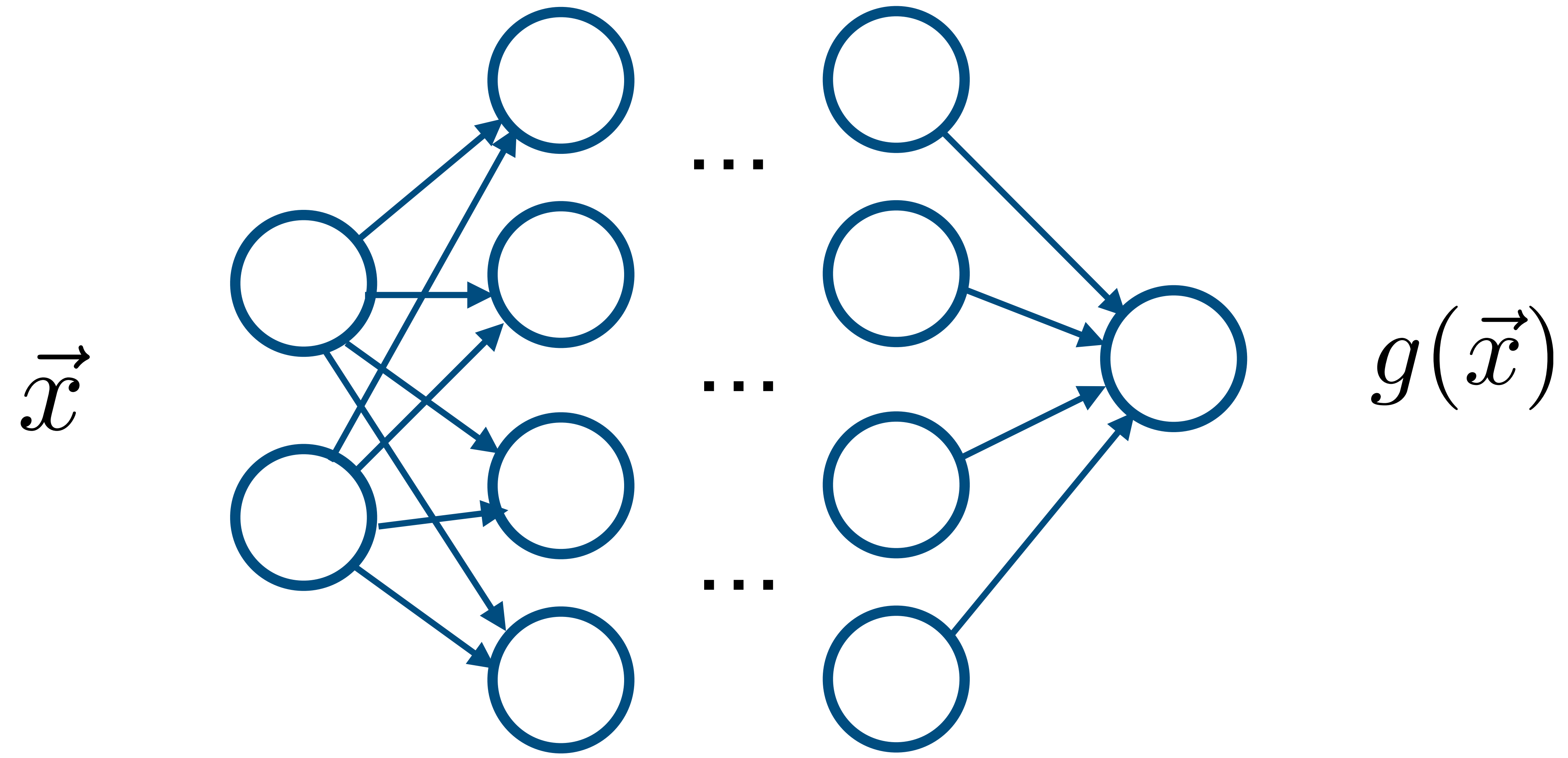


$g(\vec{x}) = 0$

$g(\vec{x}) = 1$

$$\vec{x} = \{m_{t\bar{t}}, p_T^{\ell_1}, p_T^{\ell_2}, \Delta\eta_{\ell_1,\ell_2}, \Delta\phi_{\ell_1,\ell_2}, ...\}$$

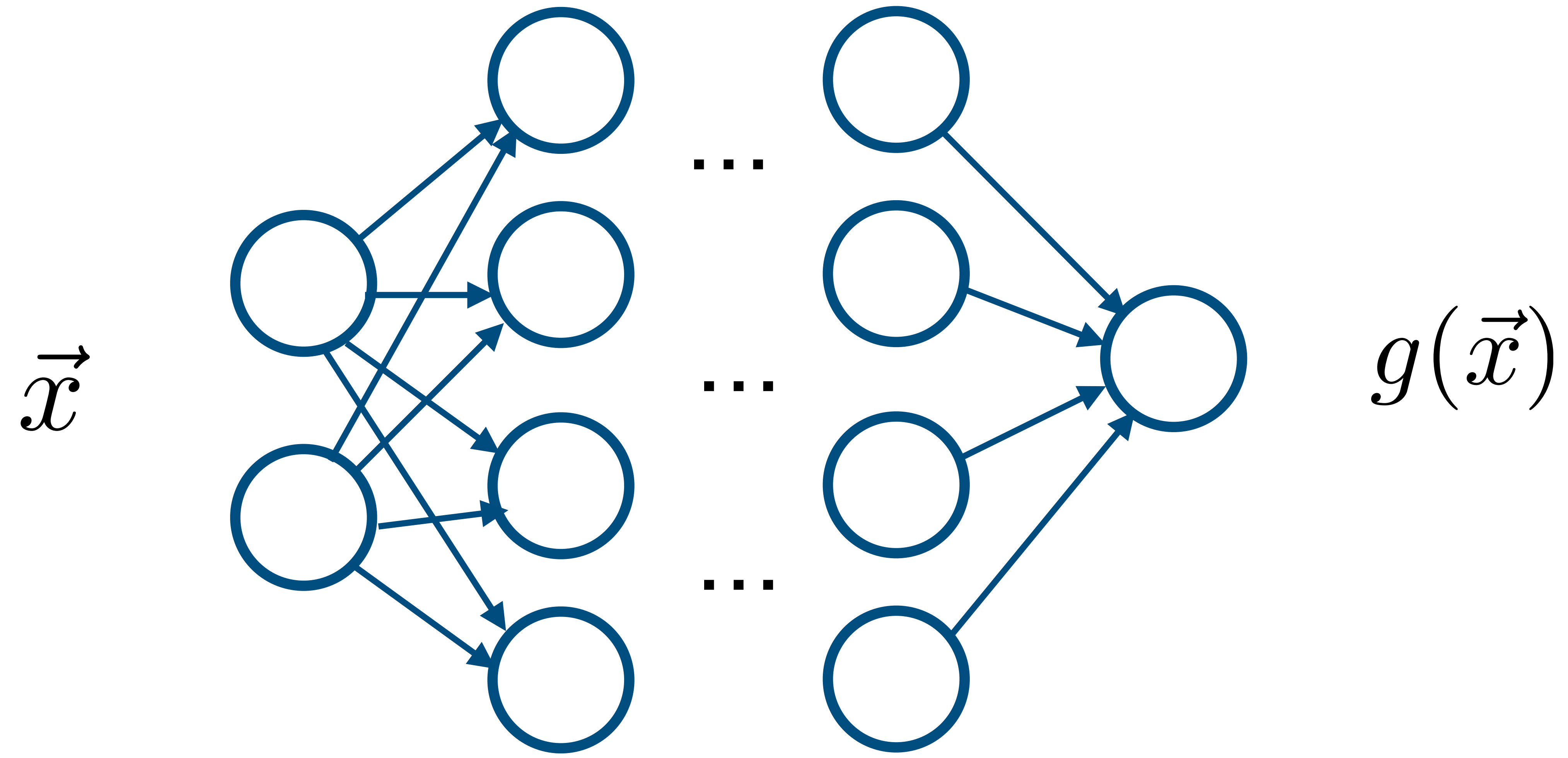
The ‘classifier trick’



$g(\vec{x}) = 0$ **SMEFT**

$g(\vec{x}) = 1$ **SM**

The ‘classifier trick’



$g(\vec{x}) = 0$ **SMEFT**

$g(\vec{x}) = 1$ **SM**

Loss function:

$$L[g(\boldsymbol{x},\boldsymbol{c})] = -\int d\boldsymbol{x} \frac{d\sigma(\boldsymbol{x},\boldsymbol{c})}{d\boldsymbol{x}} \log(1 - g(\boldsymbol{x},\boldsymbol{c})) - \int d\boldsymbol{x} \frac{d\sigma(\boldsymbol{x},\boldsymbol{0})}{d\boldsymbol{x}} \log g(\boldsymbol{x},\boldsymbol{c})$$

SMEFT **SM**

The ‘classifier trick’

$$L[g(\boldsymbol{x},\boldsymbol{c})] = -\int d\boldsymbol{x}\,\frac{d\sigma(\boldsymbol{x},\boldsymbol{c})}{d\boldsymbol{x}}\log(1-g(\boldsymbol{x},\boldsymbol{c})) - \int d\boldsymbol{x}\,\frac{d\sigma(\boldsymbol{x},\boldsymbol{0})}{d\boldsymbol{x}}\log g(\boldsymbol{x},\boldsymbol{c})$$

The 'classifier trick'

$$L[g(\boldsymbol{x},\boldsymbol{c})] = -\int d\boldsymbol{x} \frac{d\sigma(\boldsymbol{x},\boldsymbol{c})}{d\boldsymbol{x}} \log(1 - g(\boldsymbol{x},\boldsymbol{c})) - \int d\boldsymbol{x} \frac{d\sigma(\boldsymbol{x},\boldsymbol{0})}{d\boldsymbol{x}} \log g(\boldsymbol{x},\boldsymbol{c})$$

$$\frac{\delta L}{\delta g} = 0 \Rightarrow \boxed{g(\mathbf{x},\mathbf{c})} = \left(1 + \frac{d\sigma(\mathbf{x},\mathbf{c})}{dx} \Big/ \frac{d\sigma(\mathbf{x},\mathbf{0})}{dx}\right)^{-1} \equiv \frac{1}{1 + r_\sigma(\mathbf{x},\mathbf{c})}$$

The ‘classifier trick’

$$L[g(\boldsymbol{x},\boldsymbol{c})] = -\int d\boldsymbol{x} \frac{d\sigma(\boldsymbol{x},\boldsymbol{c})}{d\boldsymbol{x}} \log(1-g(\boldsymbol{x},\boldsymbol{c})) - \int d\boldsymbol{x} \frac{d\sigma(\boldsymbol{x},\boldsymbol{0})}{d\boldsymbol{x}} \log g(\boldsymbol{x},\boldsymbol{c})$$

$$\frac{\delta L}{\delta g} = 0 \Rightarrow \boxed{g(\mathbf{x},\mathbf{c})} = \left(1 + \frac{d\sigma(\mathbf{x},\mathbf{c})}{dx} \Big/ \frac{d\sigma(\mathbf{x},\mathbf{0})}{dx}\right)^{-1} \equiv \frac{1}{1+r_\sigma(\mathbf{x},\mathbf{c})}$$

or: $$\frac{d\sigma(x,c)/dx}{d\sigma(x,0)/dx} = \frac{1-g(x,c)}{g(x,c)}$$

Parametrised classifier

$$r_\sigma(x,c) = \frac{d\sigma(x,c)/dx}{d\sigma(x,0)/dx} = \frac{1-g(x,c)}{g(x,c)}$$

Exploit the polynomial structure of the SMEFT when defining the classifier *g*:

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

c.f. $k_{\rm i}(C) = 1 + r_{\rm lin,i}C + r_{\rm quad,i}C^2 \quad i = 1, ..., n_{bins}$

Parametrised classifier

$$r_\sigma(x,c) = \frac{d\sigma(x,c)/dx}{d\sigma(x,0)/dx} = \frac{1-g(x,c)}{g(x,c)}$$

Exploit the polynomial structure of the SMEFT when defining the classifier *g*:

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

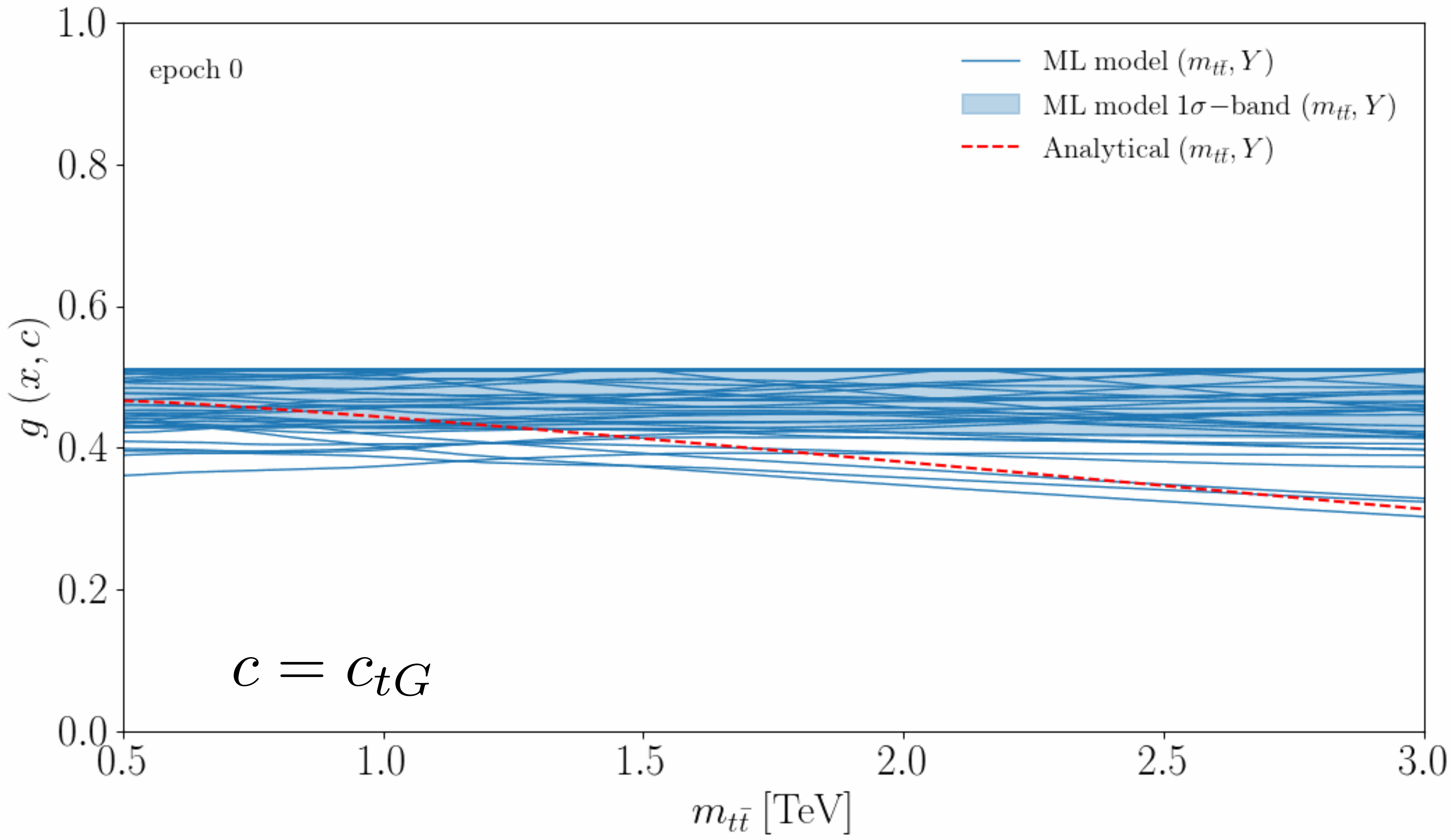
Parallelisable: generate a training sample with only c_i and learn only $\mathrm{NN}^i(\mathbf{x})$

well-suited to global fits of many SMEFT coefficients

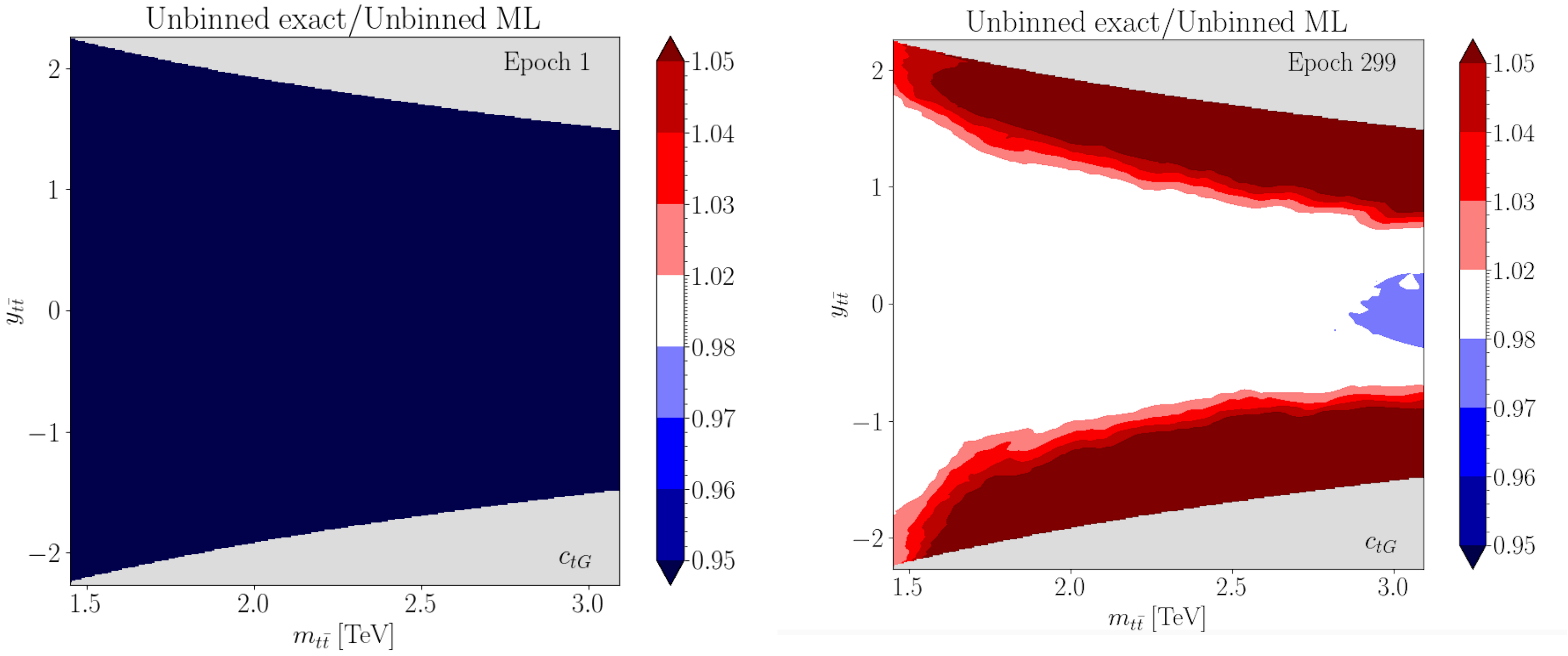
Proof of concept: top quark pair production

Validation against the analytical calculation of $\frac{d^2\sigma}{dm_{t\bar{t}}dy_{t\bar{t}}}$ for **parton-level** $t\bar{t}$ production

Train multiple instances of **g** to quantify the impact of finite training data samples



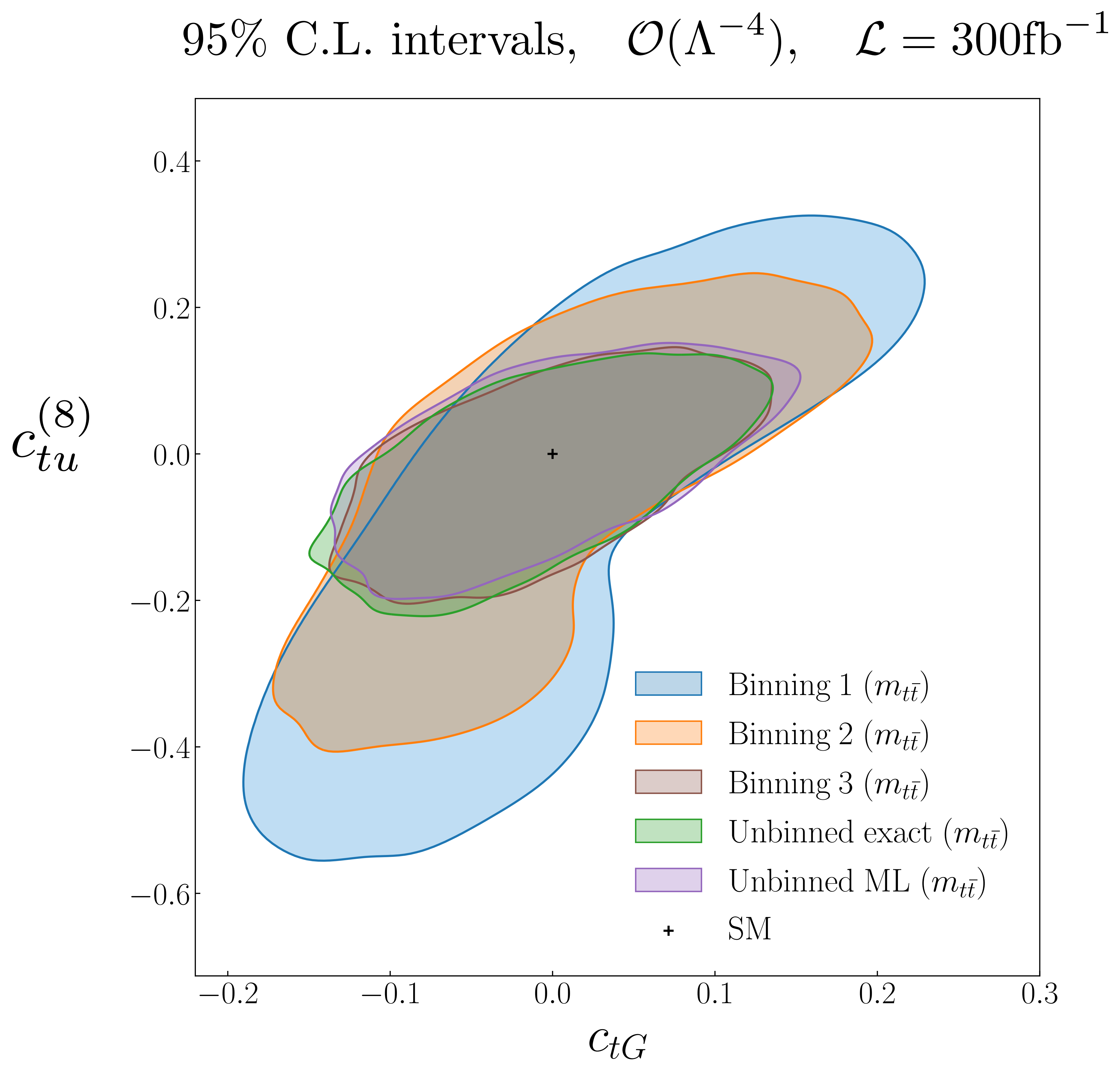
Proof of concept: top quark pair production



Proof of concept: top quark pair production

Validation against the analytical calculation of $\frac{d^2\sigma}{dm_{t\bar{t}}dy_{t\bar{t}}}$ for **parton-level** $t\bar{t}$ production

- Unbinned exact and unbinned ML agree:
  - **validation of methodology**



Proof of concept: top quark pair production

Validation against the analytical calculation of $\frac{d^2\sigma}{dm_{t\bar{t}}dy_{t\bar{t}}}$ for **parton-level** $t\bar{t}$ production

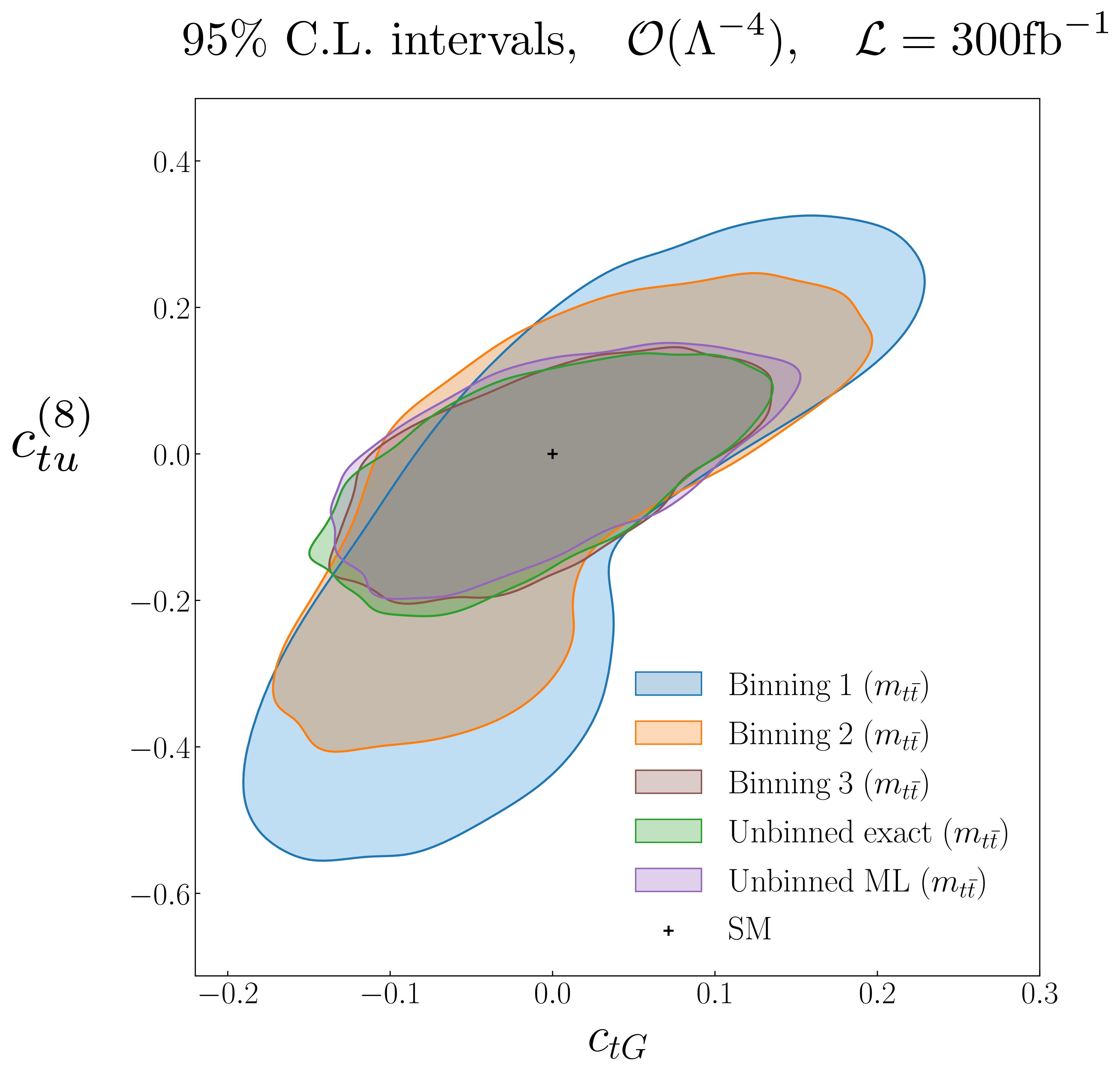
- Unbinned exact and unbinned ML agree:
  - **validation of methodology**

Binning 1 -> binning 2 -> binning 3

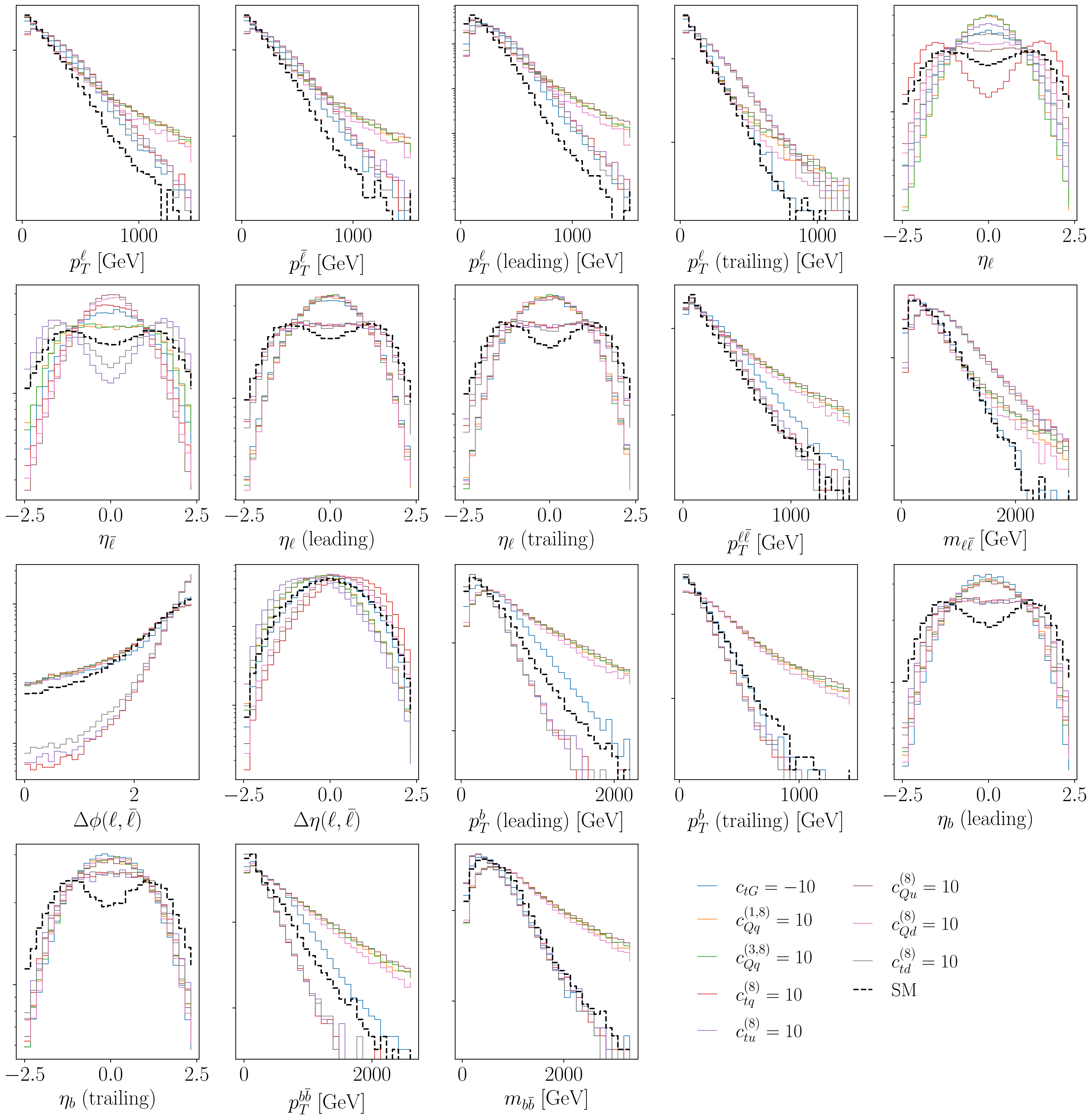
finer binning in $m_{t\bar{t}}$ →

- Binning converges to unbinned constraints with finer binning:

binning 3: $m_{t\bar{t}} \in [1.45, 1.5, 1.55, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3.0, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4.0, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 5.0, \infty)$ TeV.



Unbinned observables in the top sector



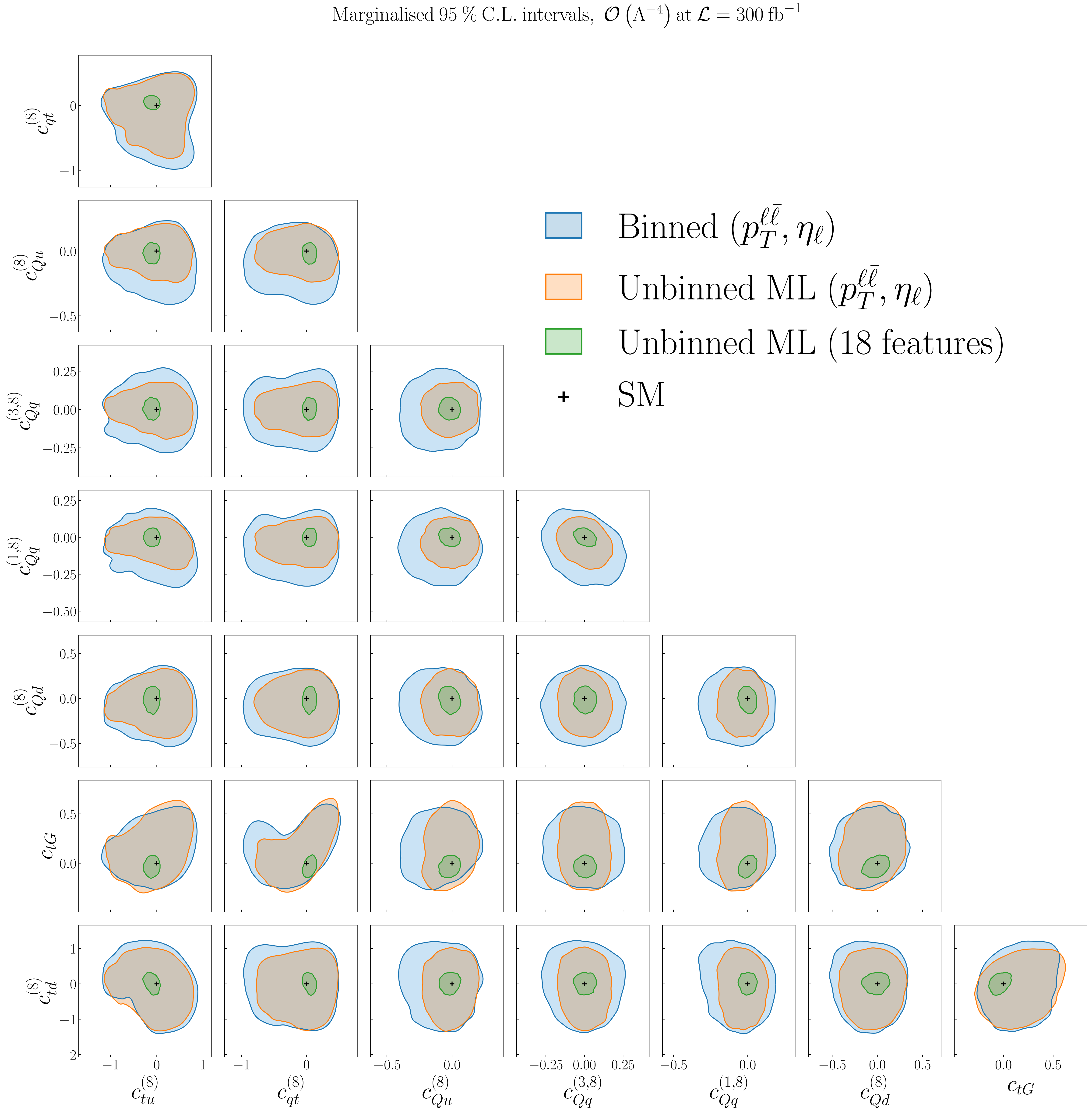
Particle-level top quark pair production in the dileptonic channel:

$$pp \to t\bar{t} \to \ell^+\ell^- b\bar{b}\nu_\ell\bar{\nu_\ell}$$

Constraints on 8 SMEFT operators:

$$O_{tG} + \text{4-fermion operators}$$

Unbinned observables in the top sector



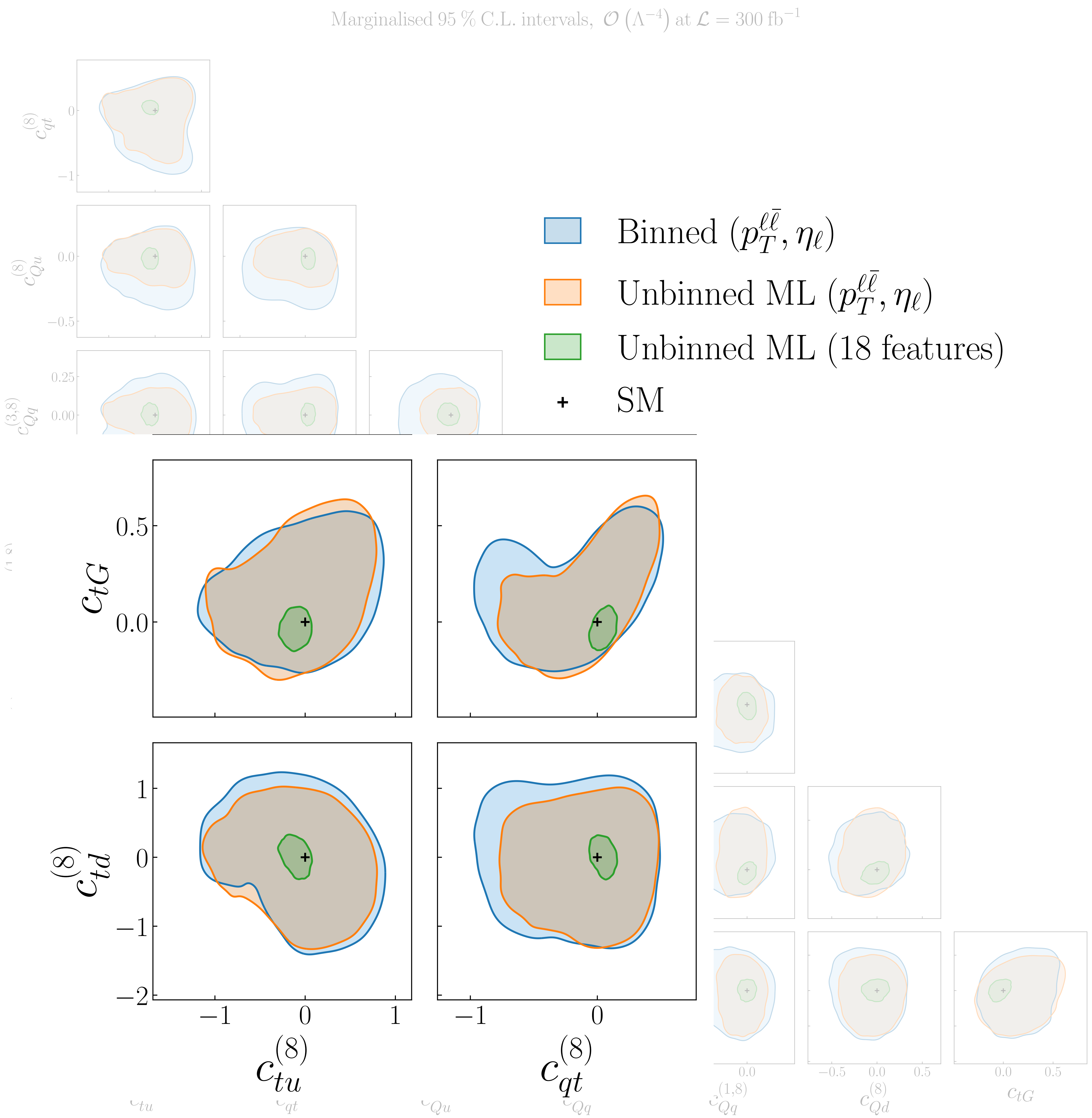
Particle-level top quark pair production in the dileptonic channel:

$$pp \to t\bar{t} \to \ell^+\ell^- b\bar{b}\nu_\ell\bar{\nu_\ell}$$

Constraints on 8 SMEFT operators:

$$O_{tG} + \text{4-fermion operators}$$

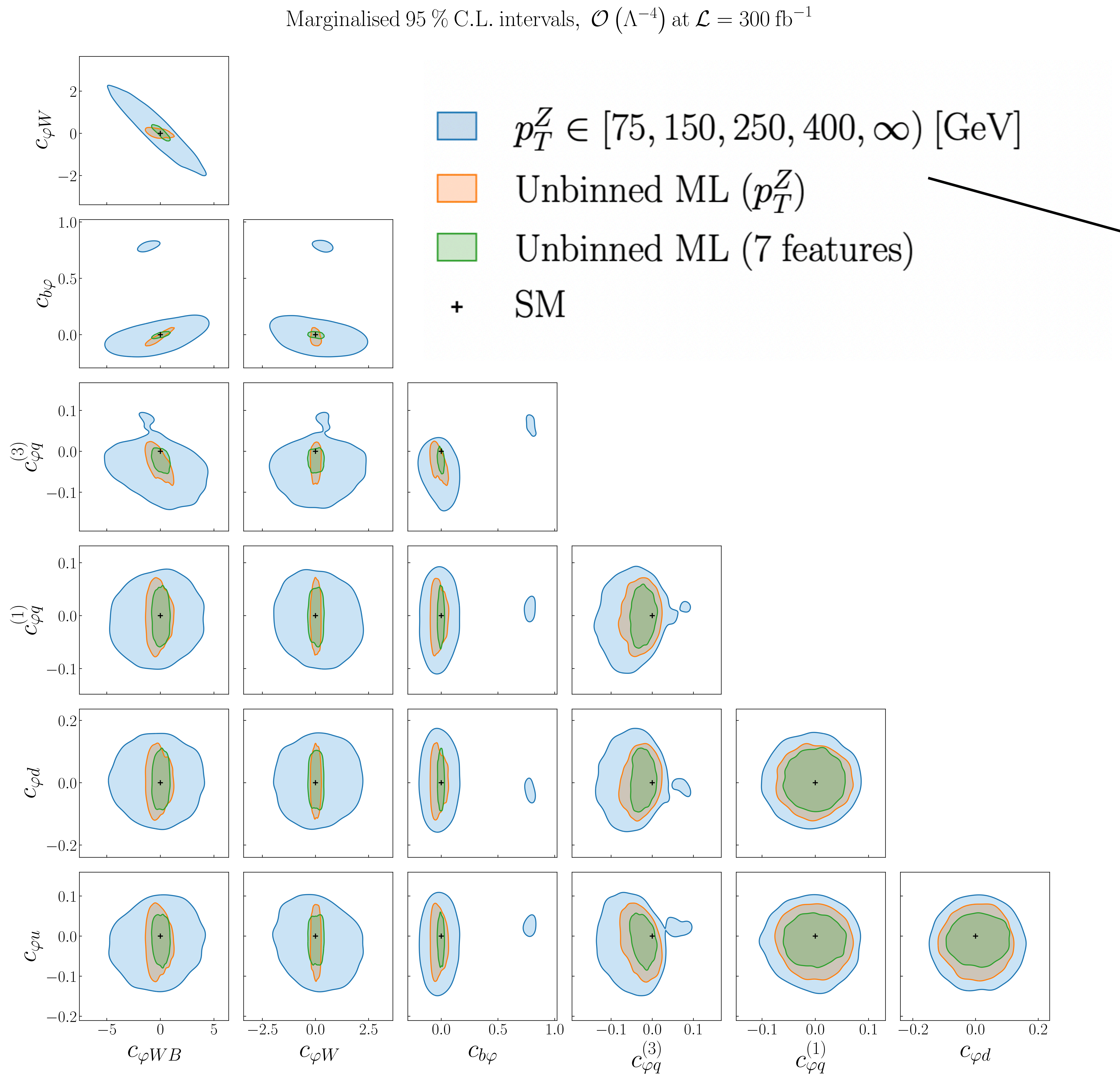
Unbinned observables in the top sector



Binned vs unbinned in $(p_T^{\ell\bar{\ell}}, \eta_\ell)$: small improvement from unbinned measurements, relative to nominal choice of bins

2 features vs 18 features: vast improvement in constraining power

Unbinned observables in the Higgs sector



$$pp \to hZ \to b\bar{b}\ell^+\ell^-$$

STXS binning - see also 1908.06980, Brehmer et. al

Constraints on 7 SMEFT coefficients:

$$c_{\varphi u}, c_{\varphi d}, c_{\varphi q}^{(1)}, c_{\varphi q}^{(3)}, c_{\varphi W}, c_{\varphi WB}, c_{b\varphi}$$

Future directions

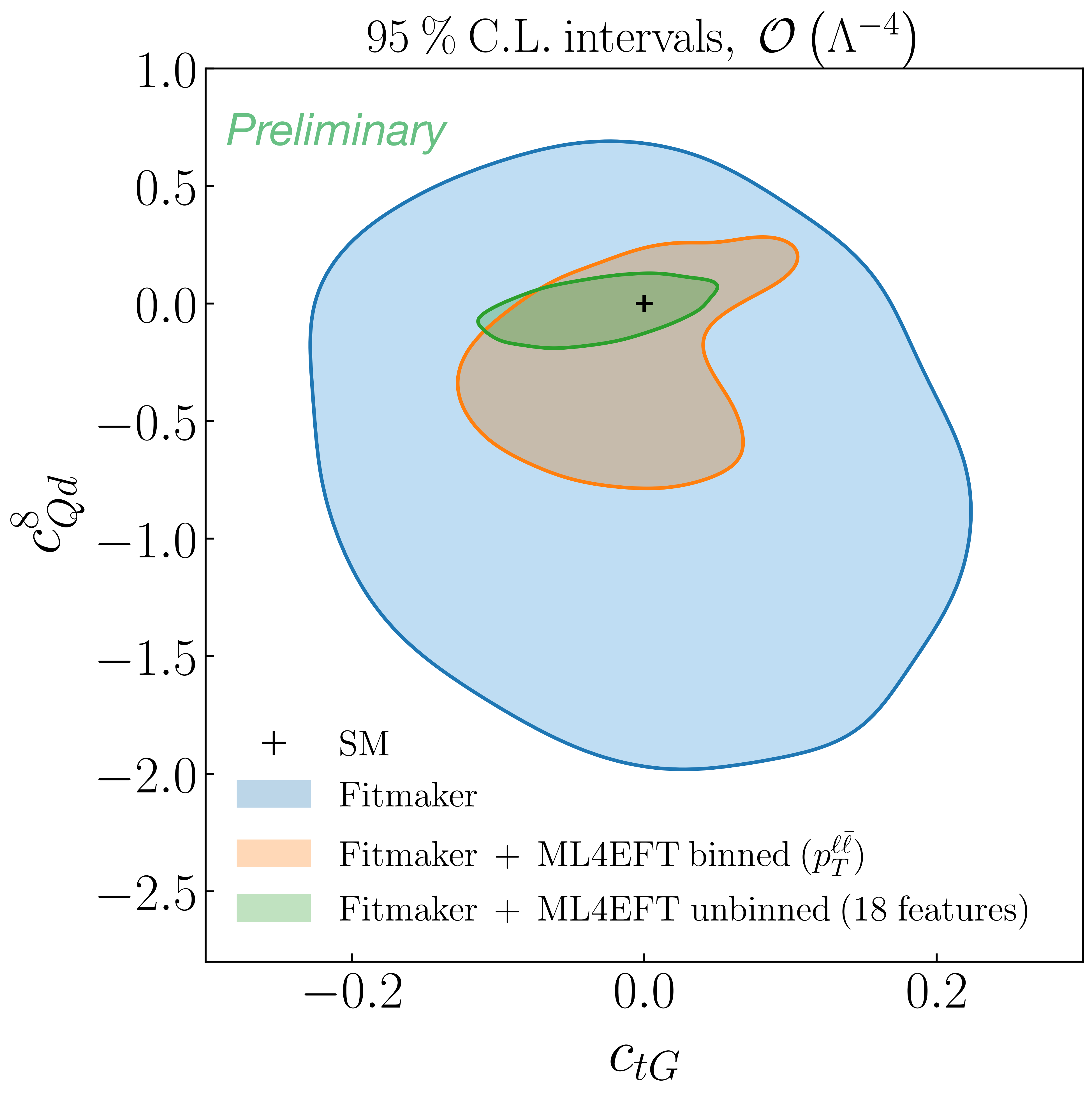
- treatment of systematic uncertainties

- New unbinned measurements can be combined alongside existing binned measurements:

$$\log\mathcal{L}(c) = \sum_{k=1}^{N_D^{(\mathrm{unbinned})}} \log\mathcal{L}_k^{\mathrm{unbinned}}(c) + \sum_{k=1}^{N_D^{(\mathrm{binned})}} \log\mathcal{L}_k^{\mathrm{binned}}(c)$$

- incorporate parton-showered observables

Work in progress, Pim Herbschleb, Jaco ter Hoeve



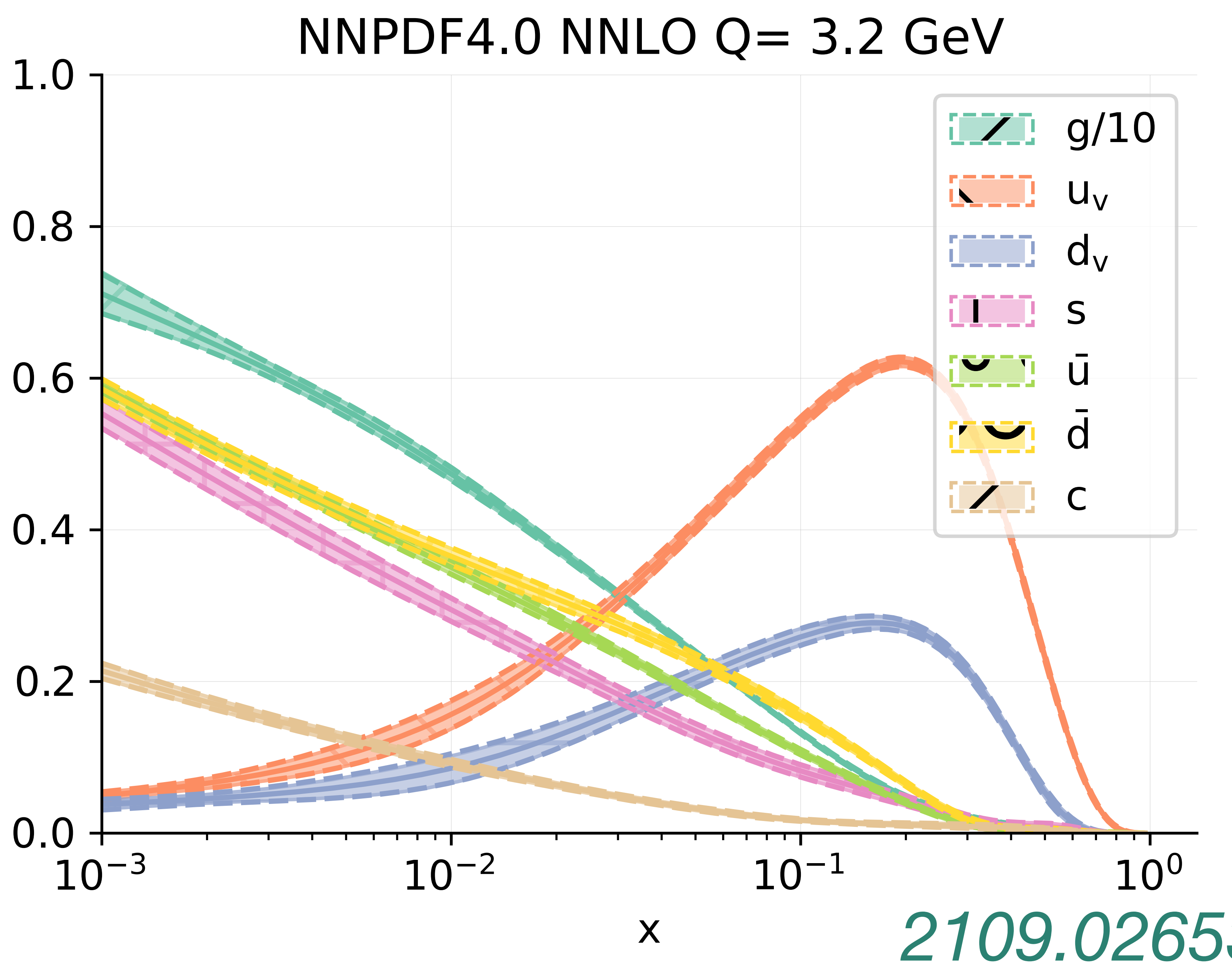
Work in progress, Jaco ter Hoeve, MM

SIMUnet: an open-source tool for the simultaneous fit of PDFs and SMEFT coefficients

Work in progress by Elie Hammou, Maeve Madigan, Luca Mantani, James Moore, Manuel Morales Alvarado, Mark Nestor Costantini, Maria Ubiali

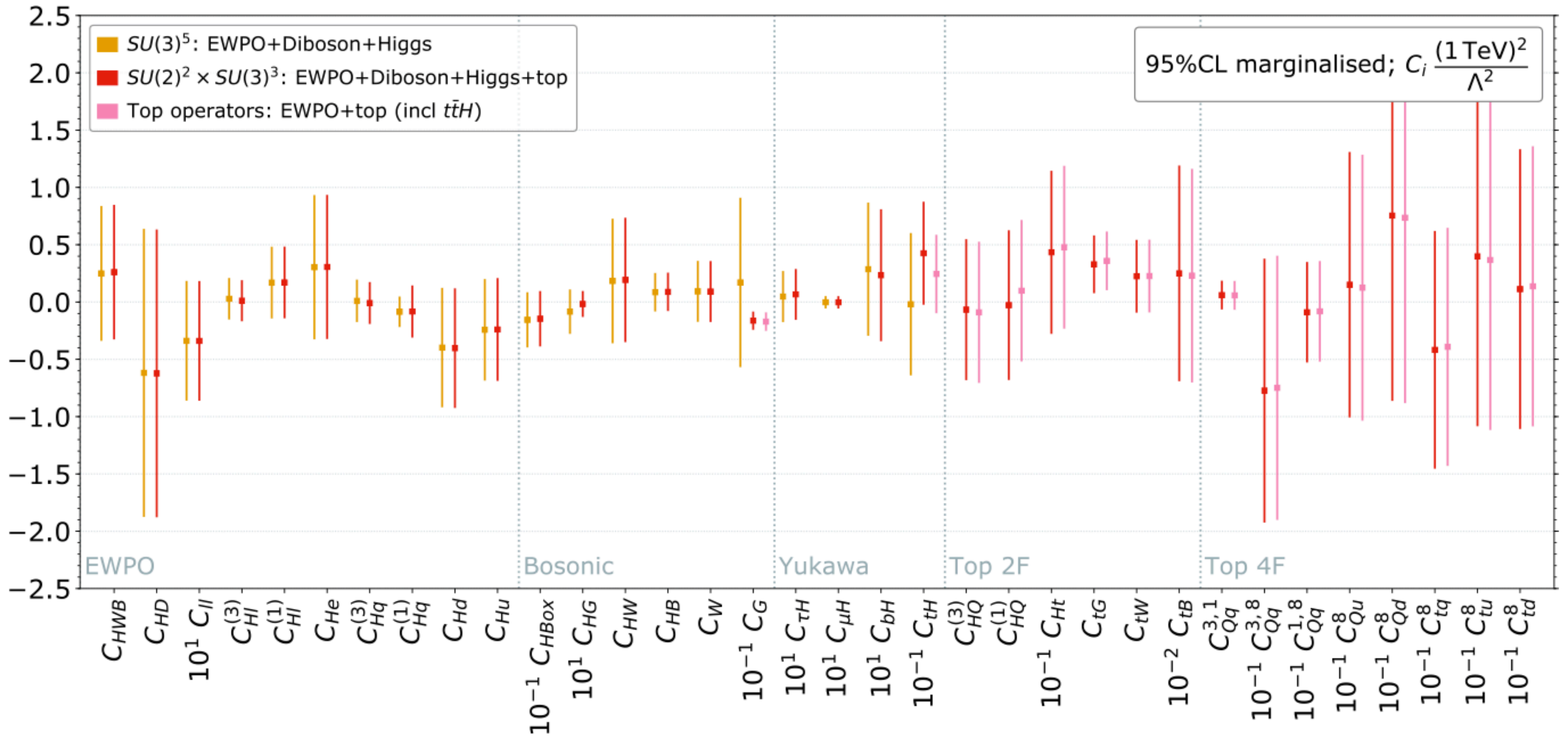
PDF-EFT Interplay

PDF fits



2109.02653

SMEFT Fits and BSM searches



2012.02779

PDF-EFT Interplay

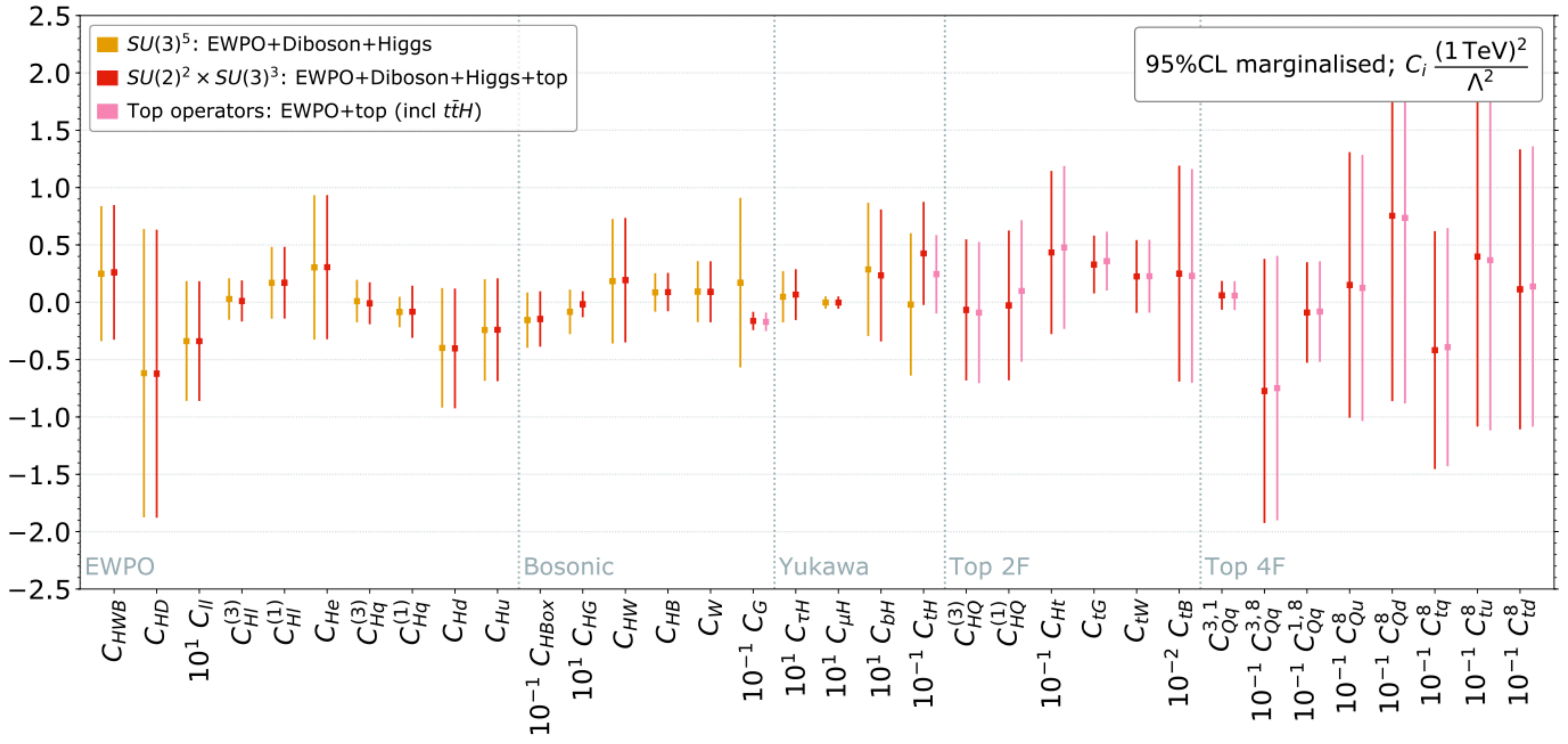
PDF fits

BSM parameters are kept fixed:

$$\sigma(\bar{c},\theta) = f_1(\theta) \otimes f_2(\theta) \otimes \hat{\sigma}(\bar{c})$$

Typically PDF fits assume the SM:
$\bar{c} = 0$

SMEFT Fits and BSM searches



2012.02779

PDF-EFT Interplay

PDF fits

BSM parameters are kept fixed:

$$\sigma(\bar{c}, \theta) = f_1(\theta) \otimes f_2(\theta) \otimes \hat{\sigma}(\bar{c})$$

Typically PDF fits assume the SM:

$$\bar{c} = 0$$

SMEFT Fits and BSM searches

PDF parameters are fixed:

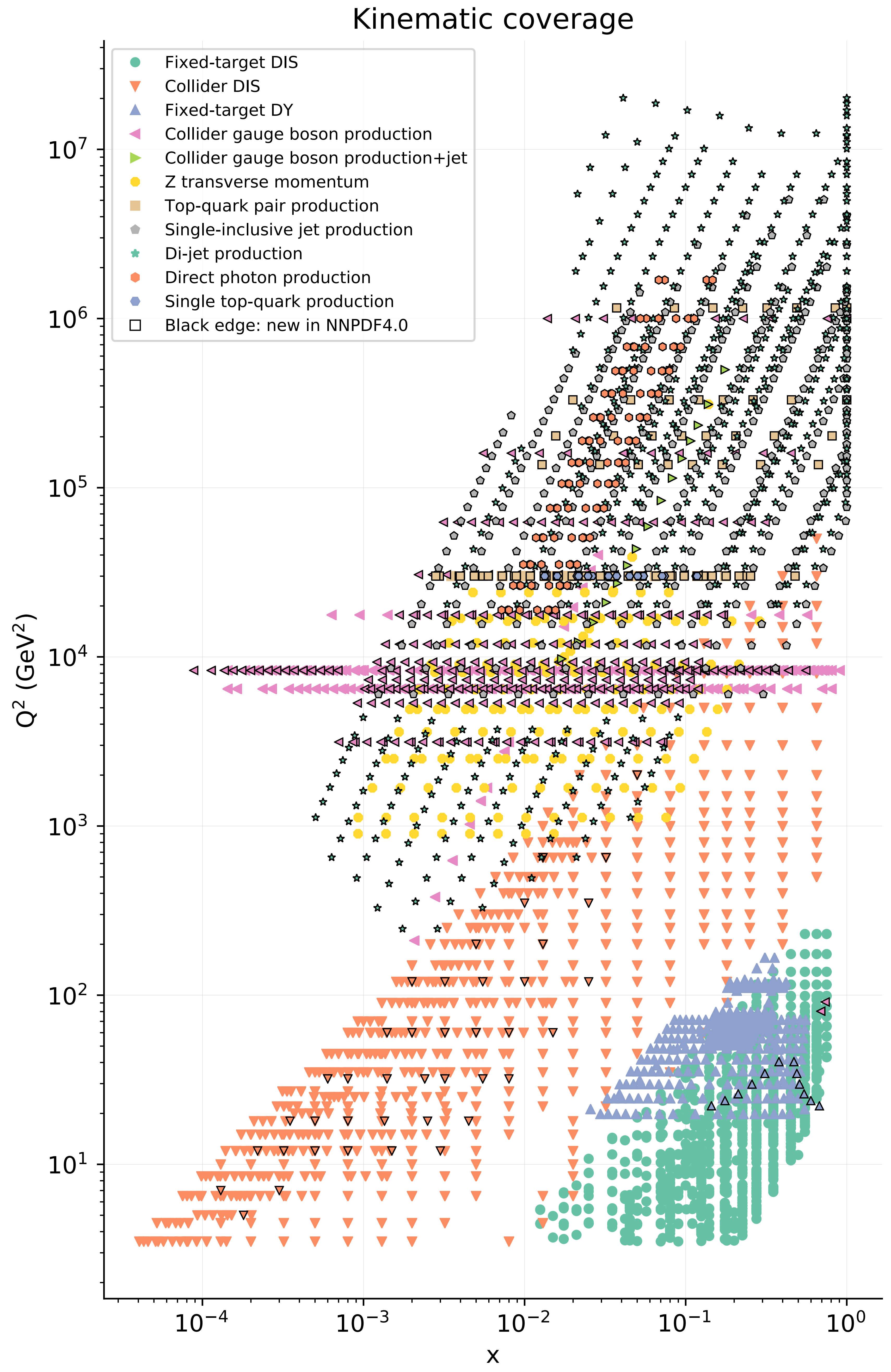
$$\sigma(c, \bar{\theta}) = f_1(\bar{\theta}) \otimes f_2(\bar{\theta}) \otimes \hat{\sigma}(c)$$

PDFs used in BSM searches rely on SM assumptions

Data overlap

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

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

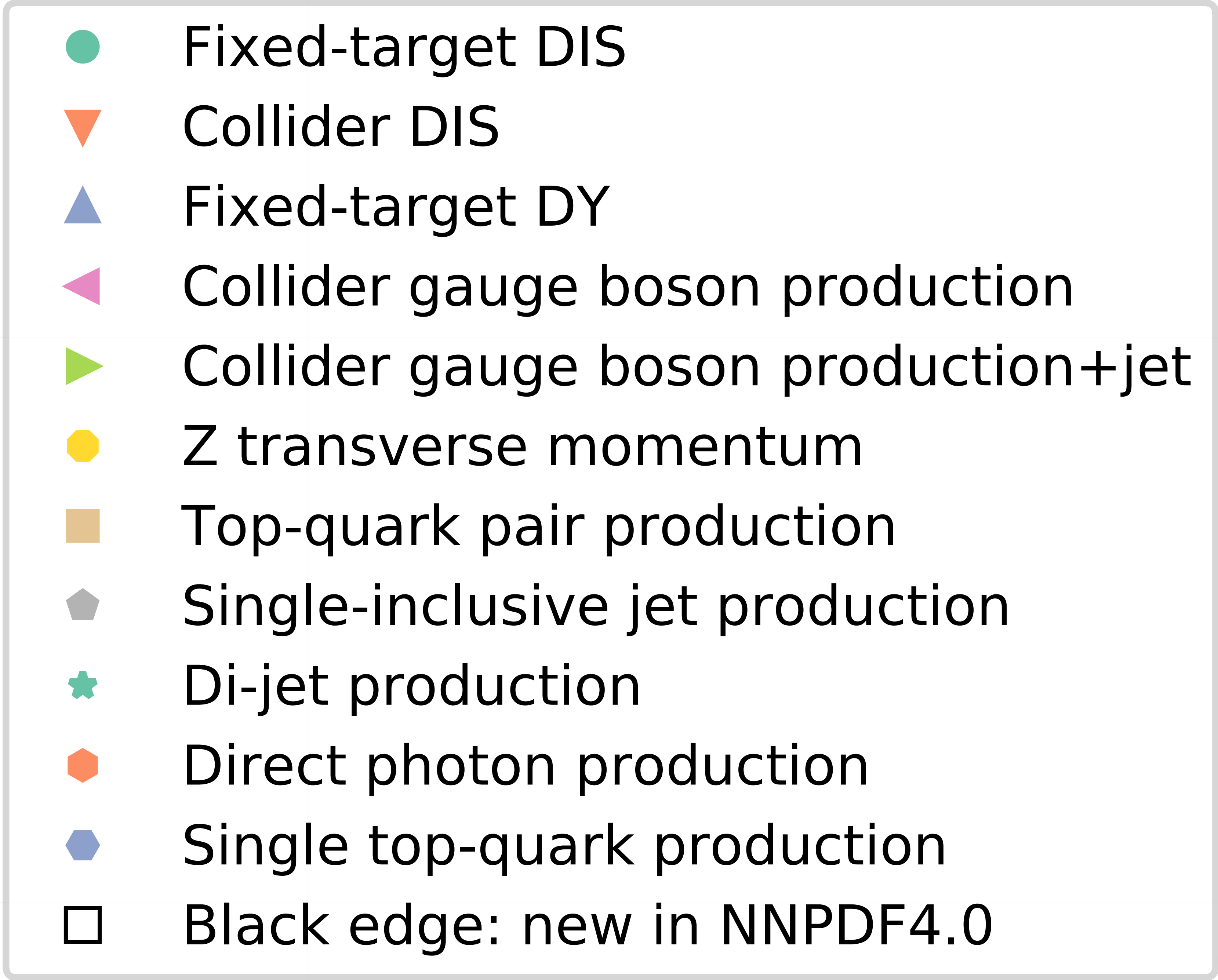


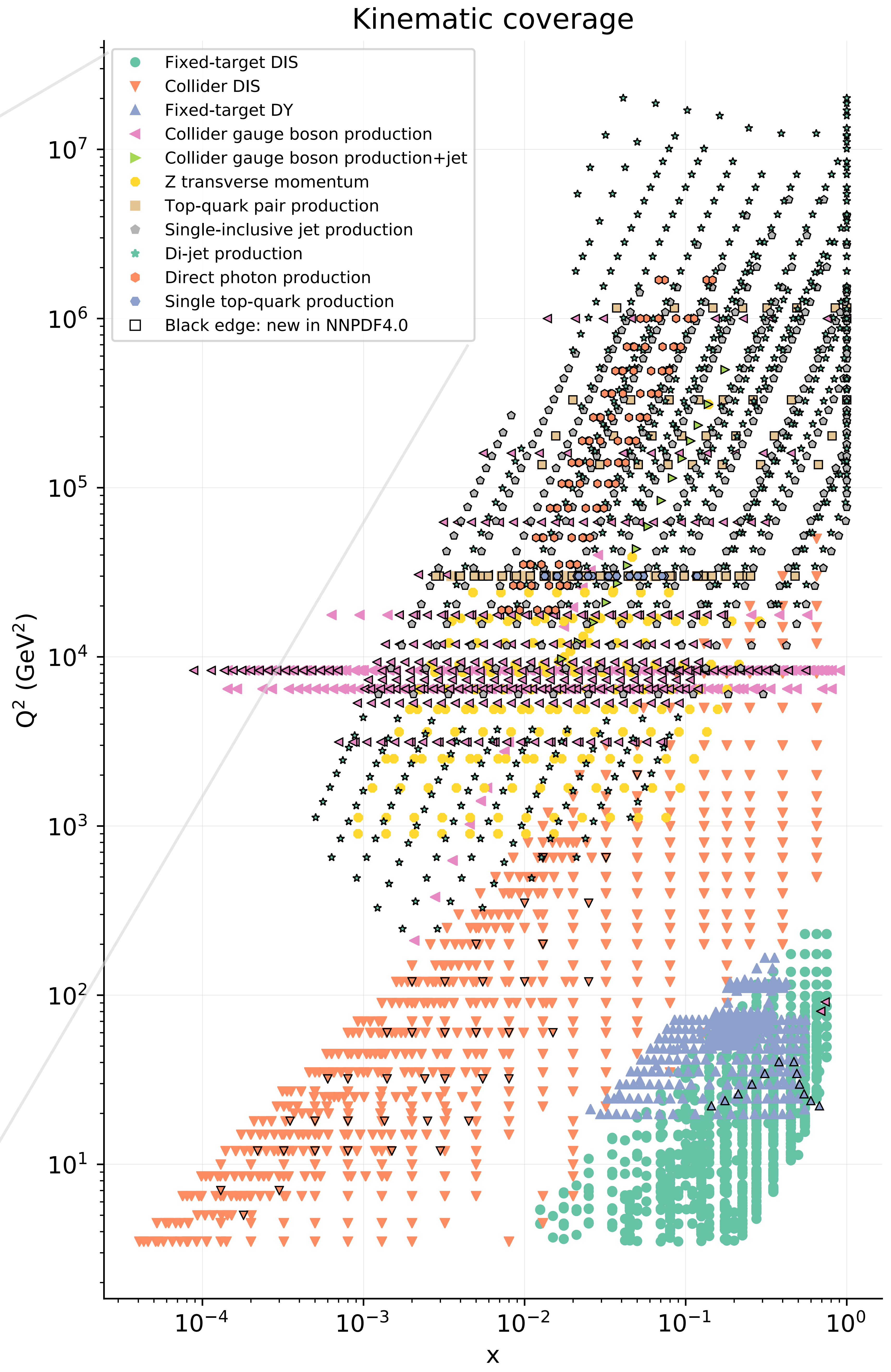
Data overlap

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

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

Data included in NNPDF4.0, [2109.02653]:



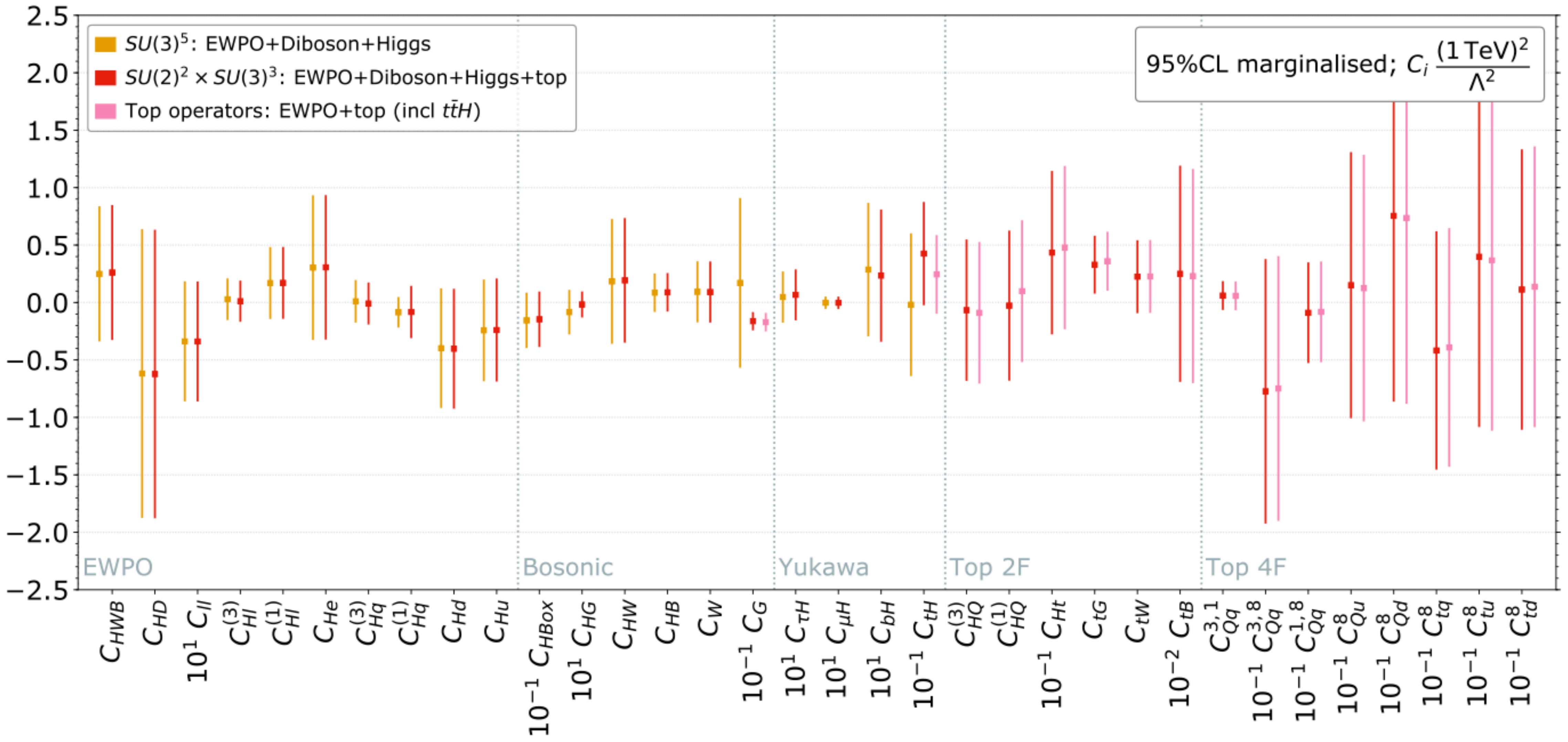


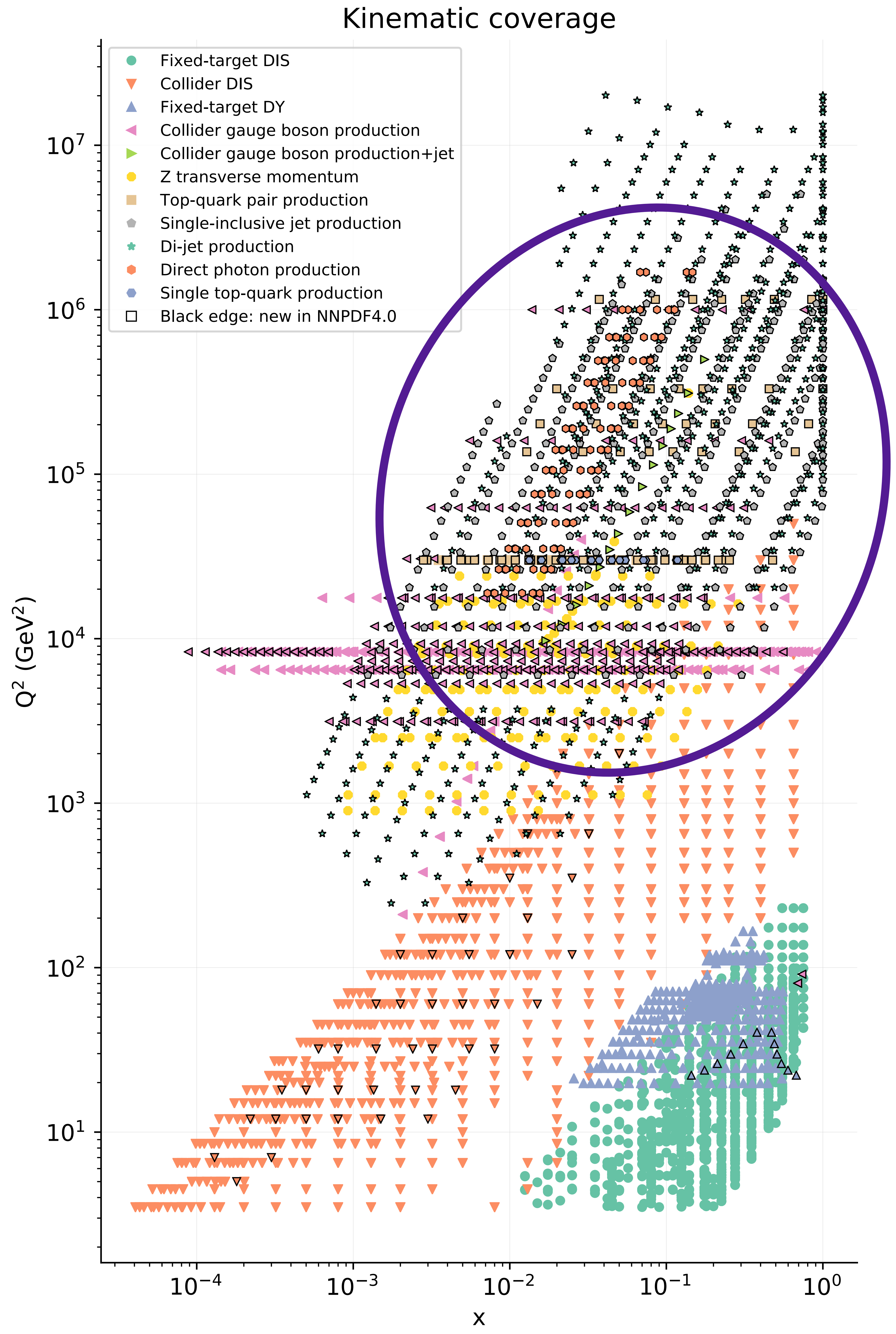
Data overlap

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

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

- e.g. Top quark data used to fit the SMEFT in the global fit of *2012.02779, J. Ellis, MM, K. Mimasu, V. Sanz, T. You*



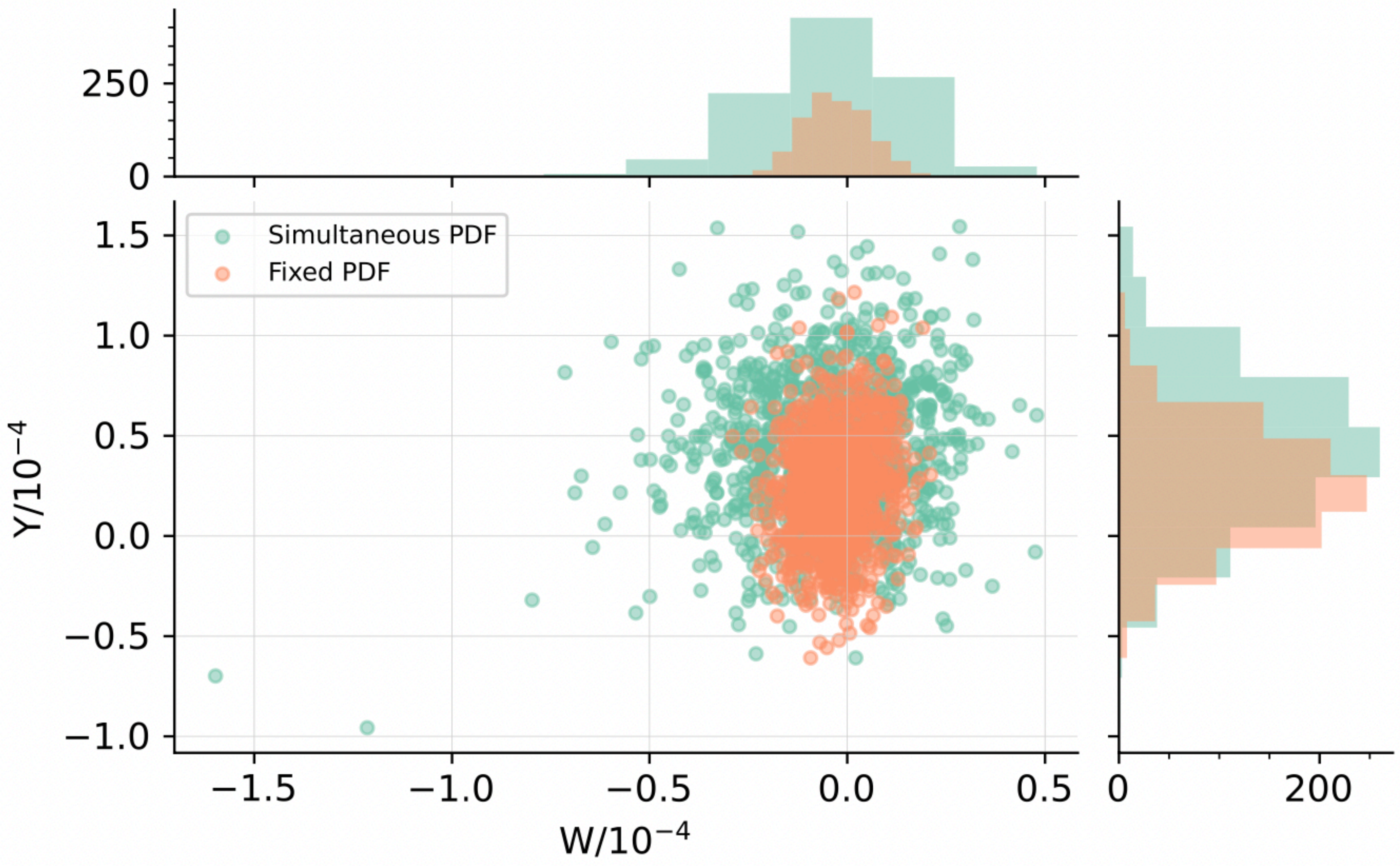


Simultaneous PDF and SMEFT determinations

S. Iranipour, M. Ubiali, 2201.07240

High-mass Drell-Yan

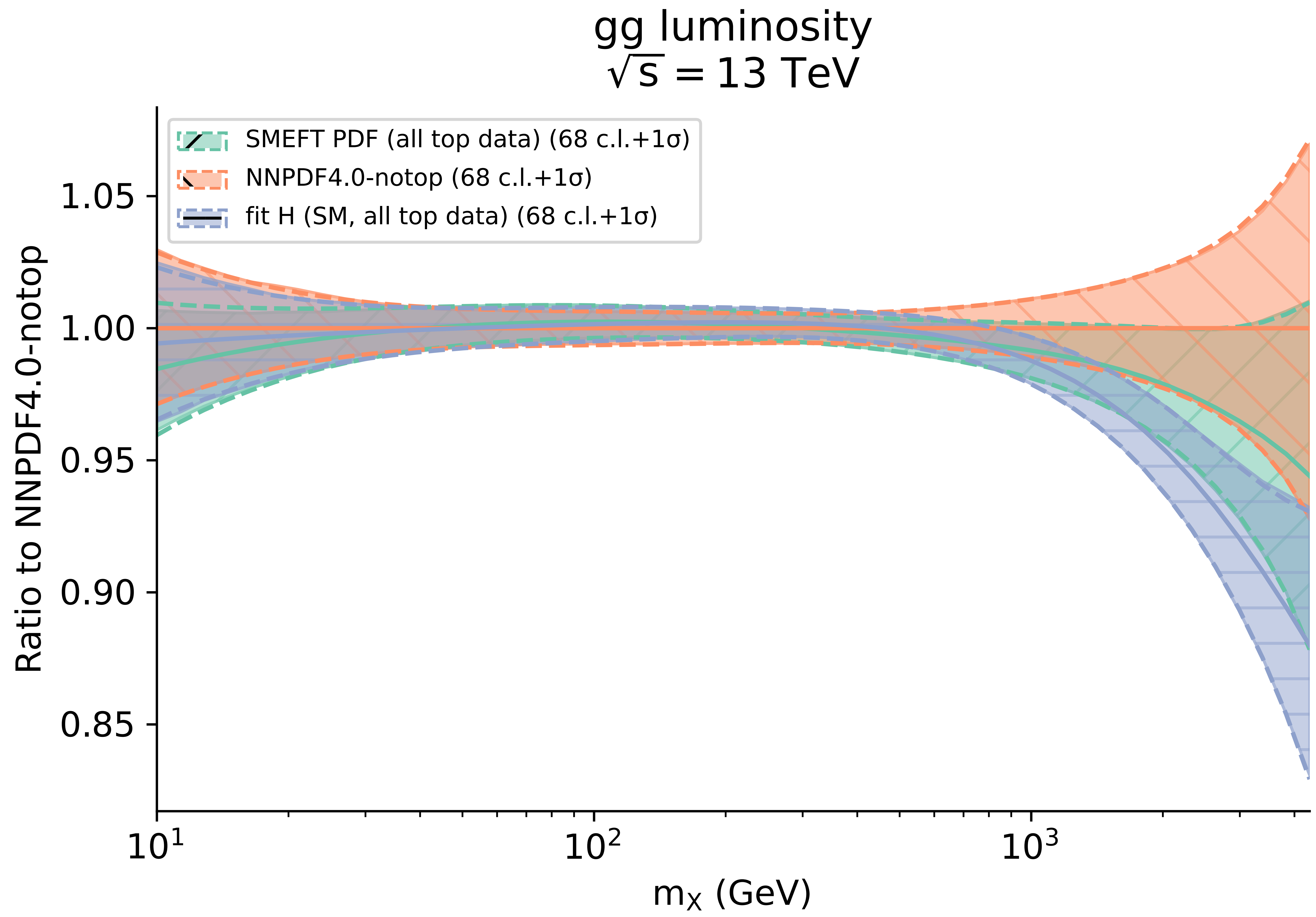
Neglecting PDF-EFT interplay at the **HL-LHC** leads to a significant overestimate of EFT constraints



Simultaneous PDF and SMEFT determinations

Kassabov et. al: 2303.06159

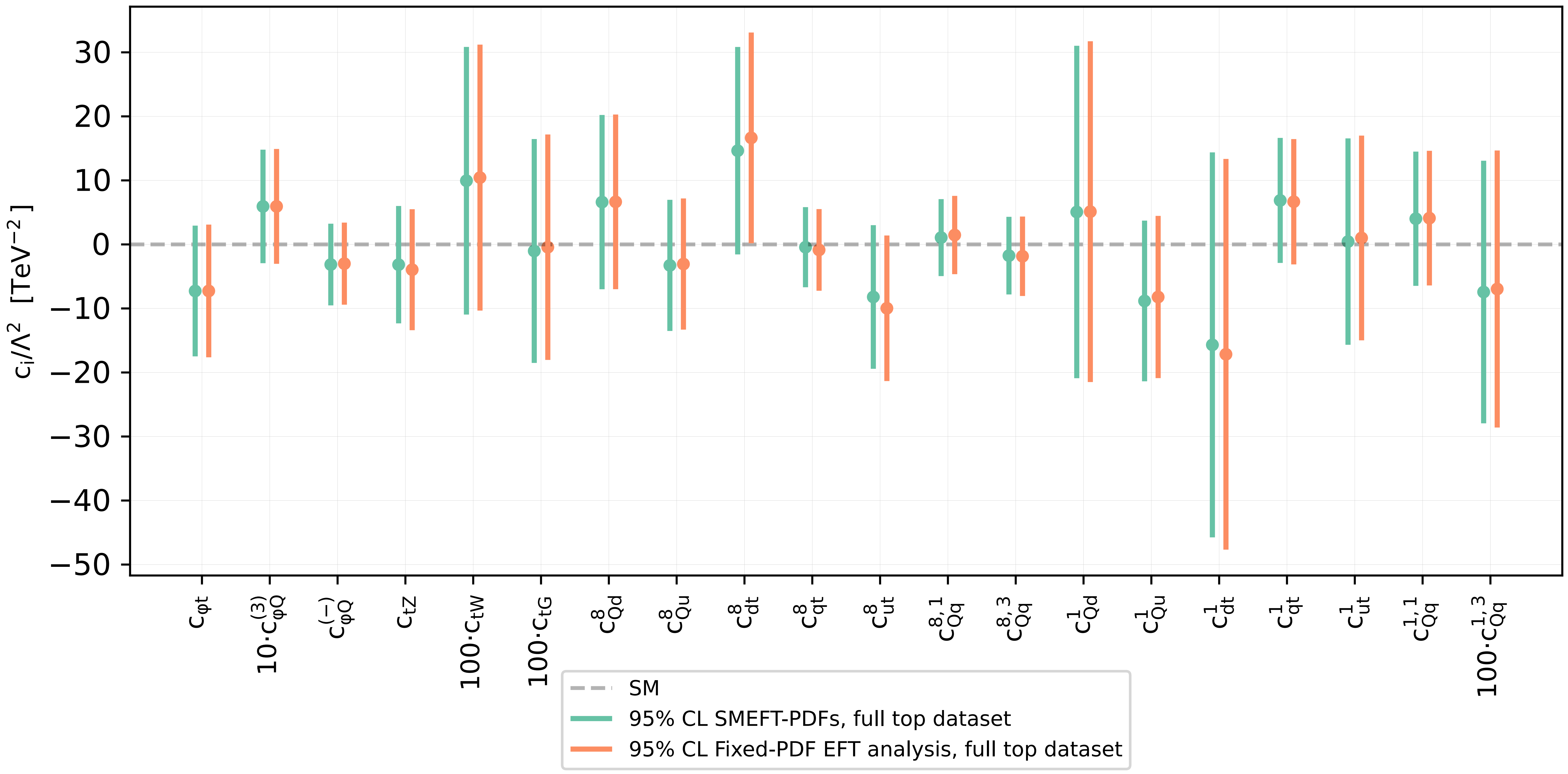
Top quark data



Simultaneous PDF and SMEFT determinations

Kassabov et. al: 2303.06159

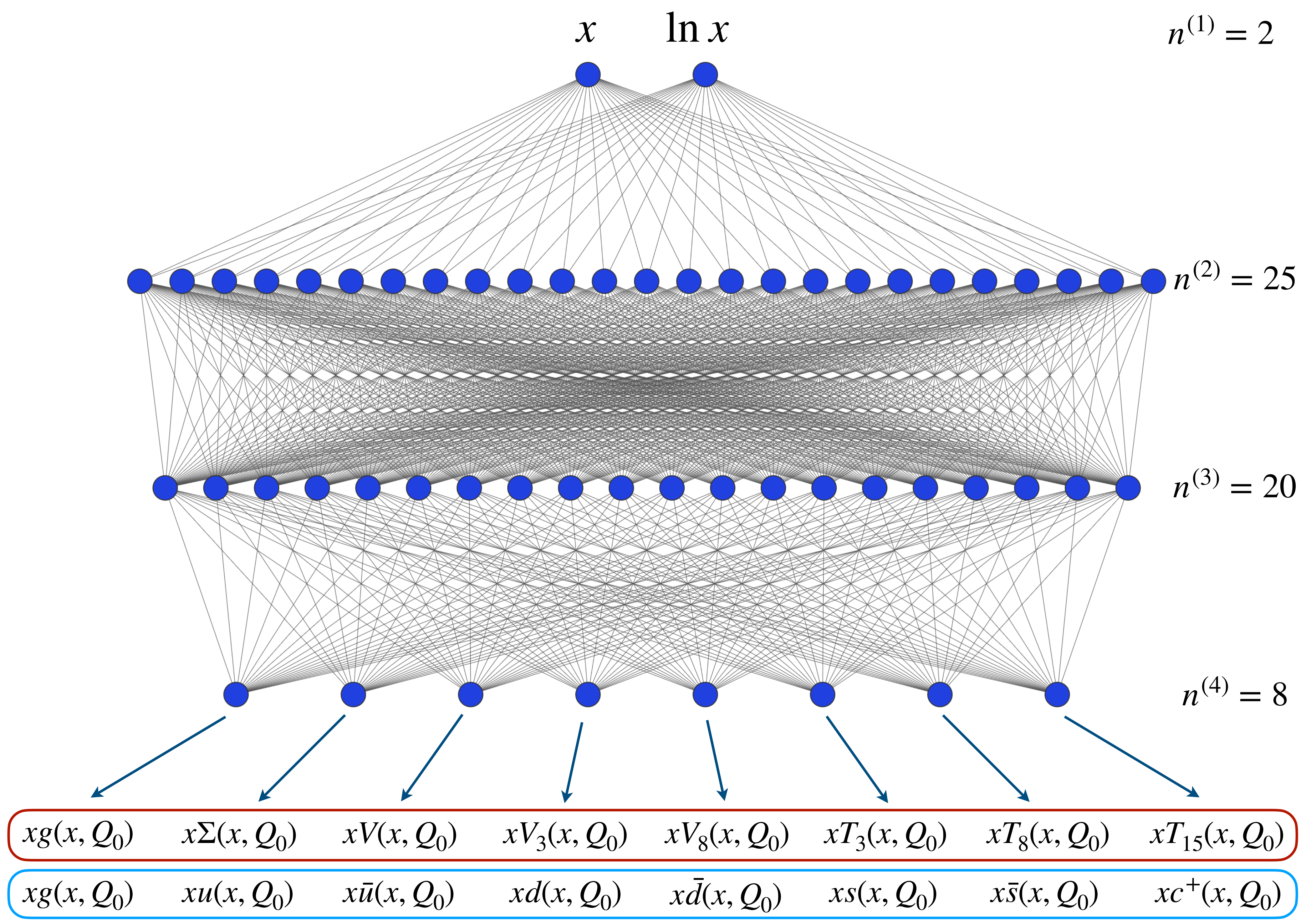
Top quark data



SIMUnet methodology

An extension of the NNPDF framework

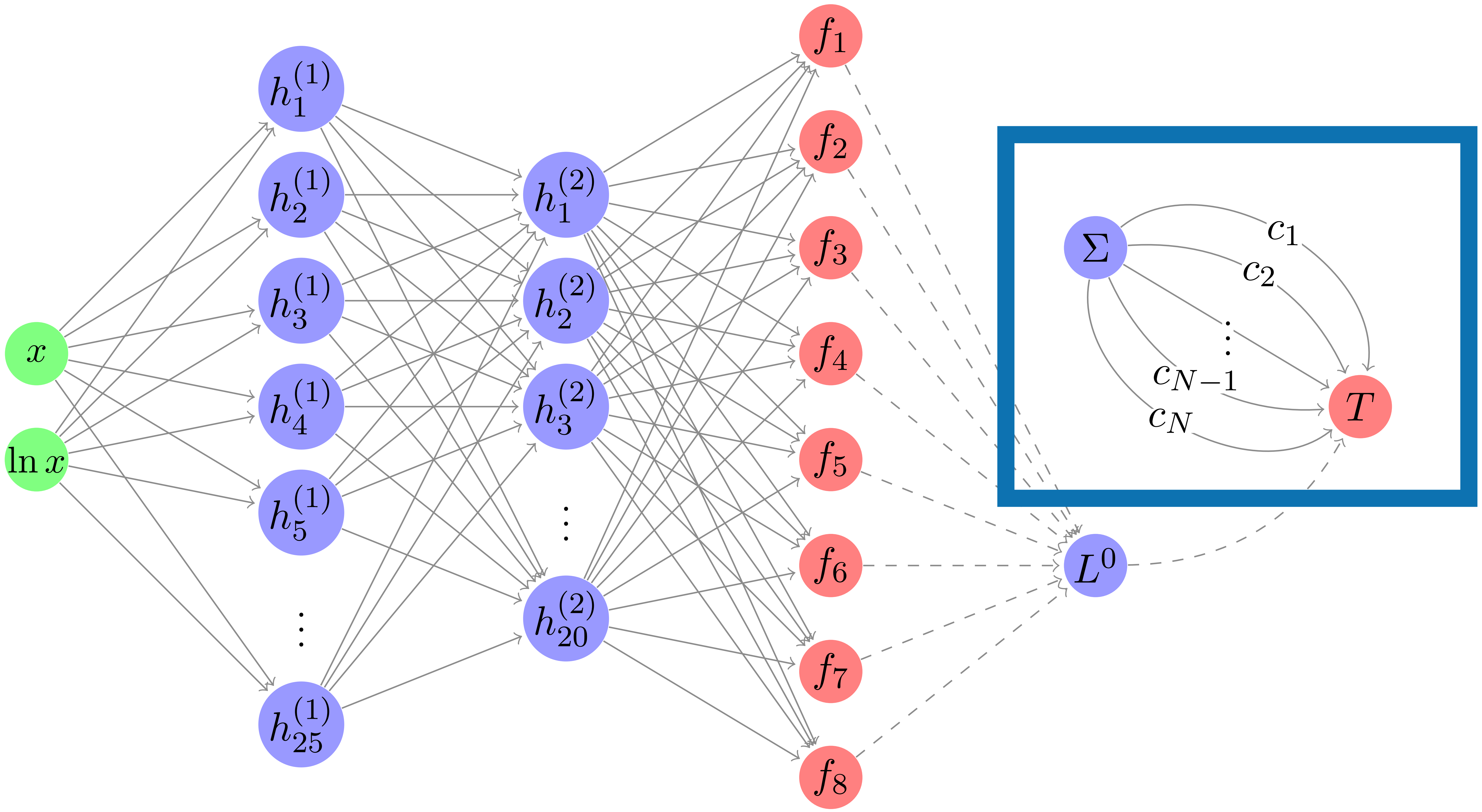
- PDFs parameterised by a neural network



Ball et. al, NNPDF4.0, 2109.02653

SIMUnet methodology

Additional layer accounts for dependence of partonic cross section on Wilson coefficients via k-factor approximation

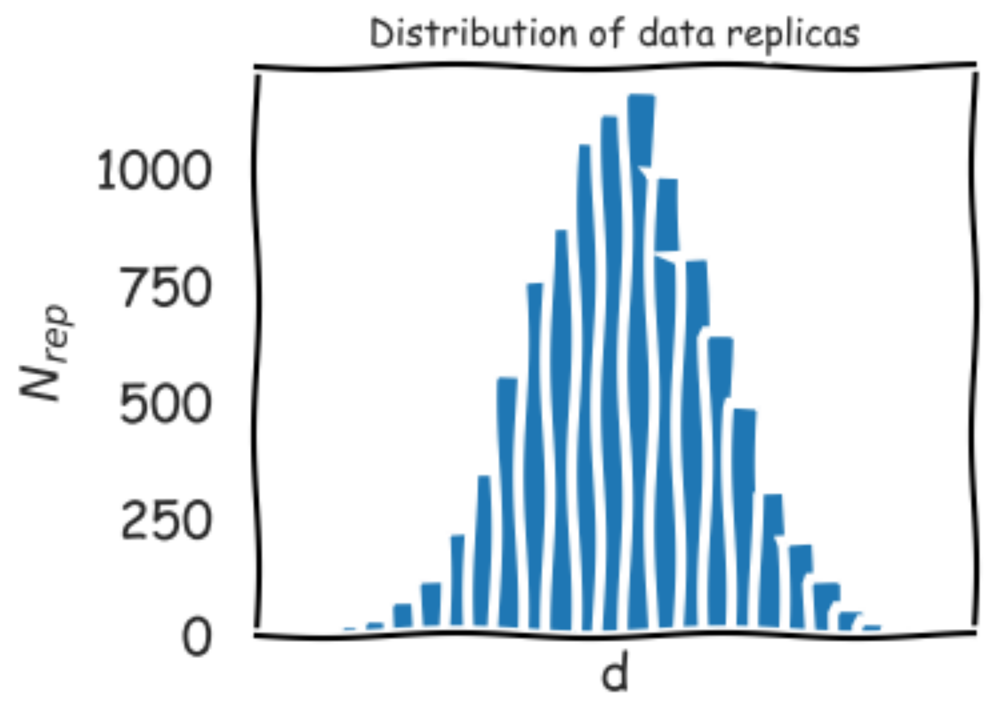


SIMUnet methodology

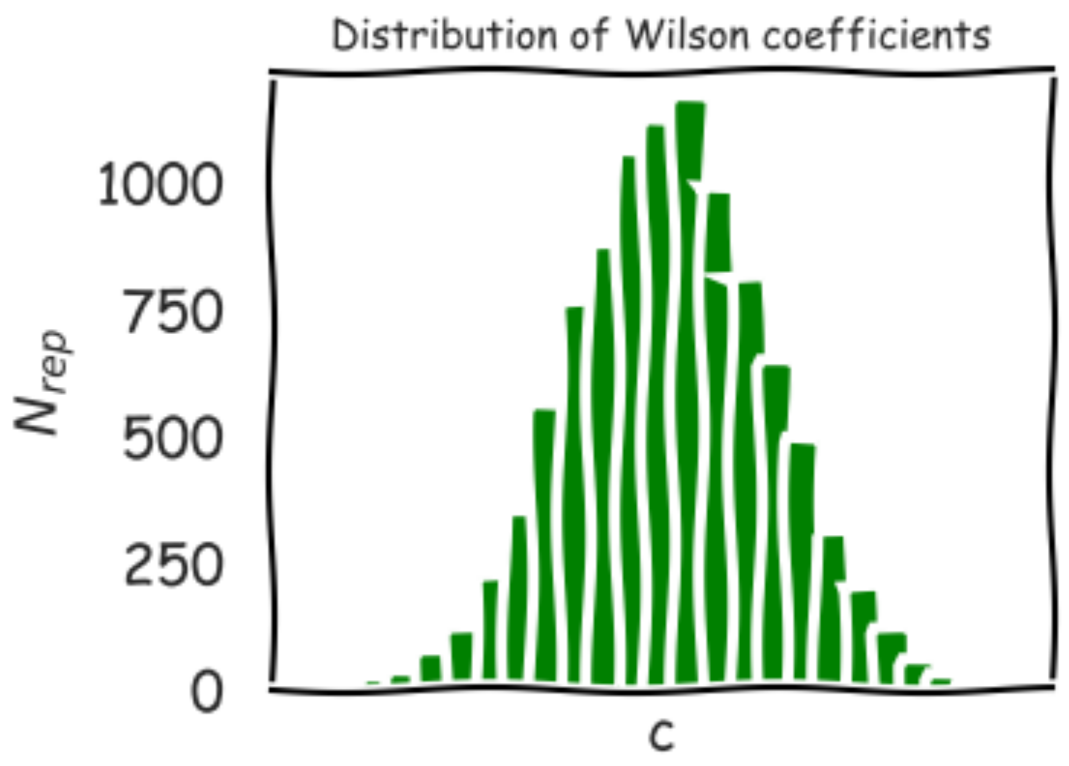
An extension of the NNPDF framework

- PDFs parameterised by a neural network
- Propagates uncertainties from data to NN parameters using the Monte Carlo replica method

$$d_k \sim \mathcal{N}(d, \sigma)$$

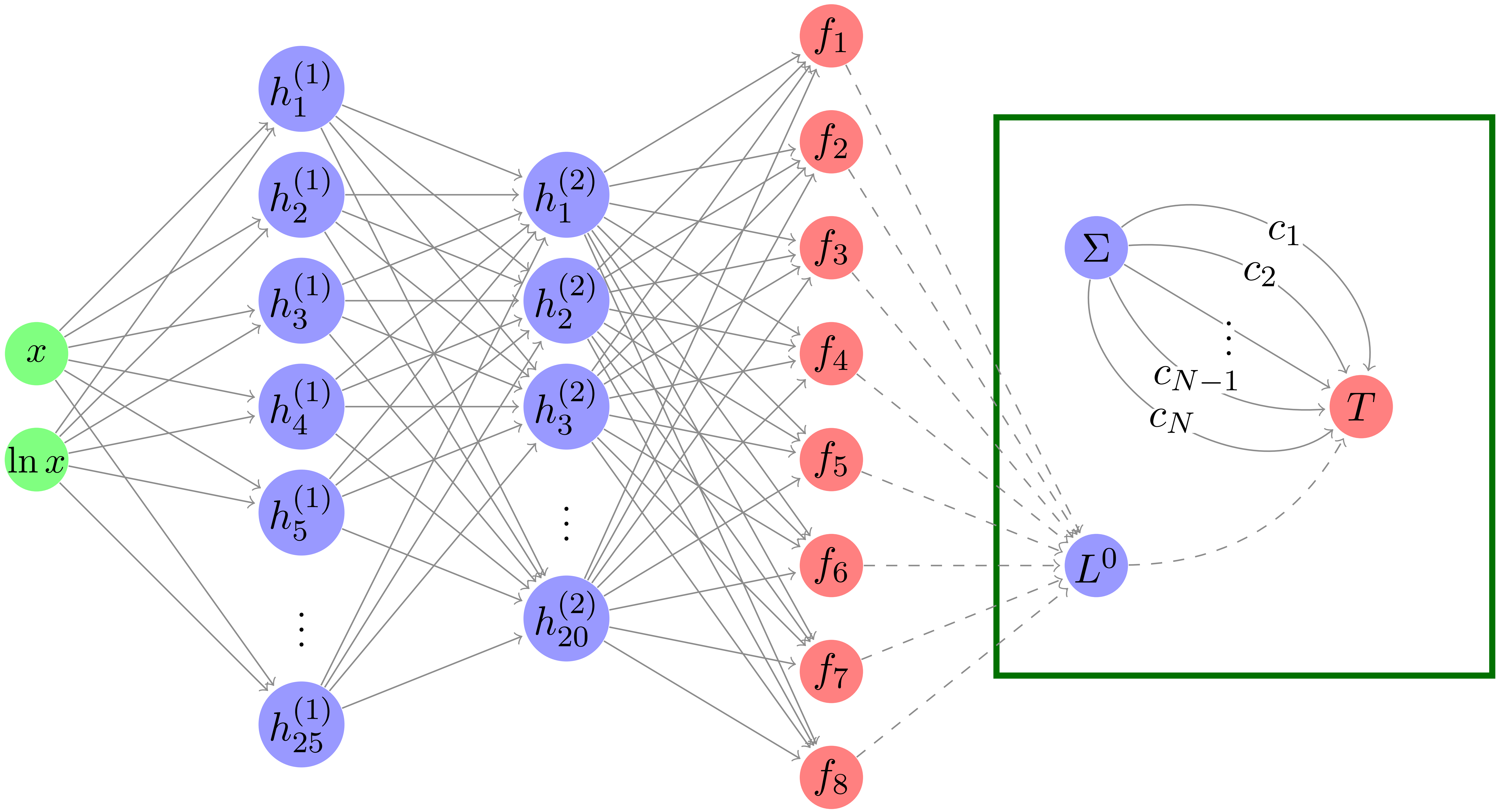


$$c_k = \arg \min_{c} \chi^2(c, d_k)$$



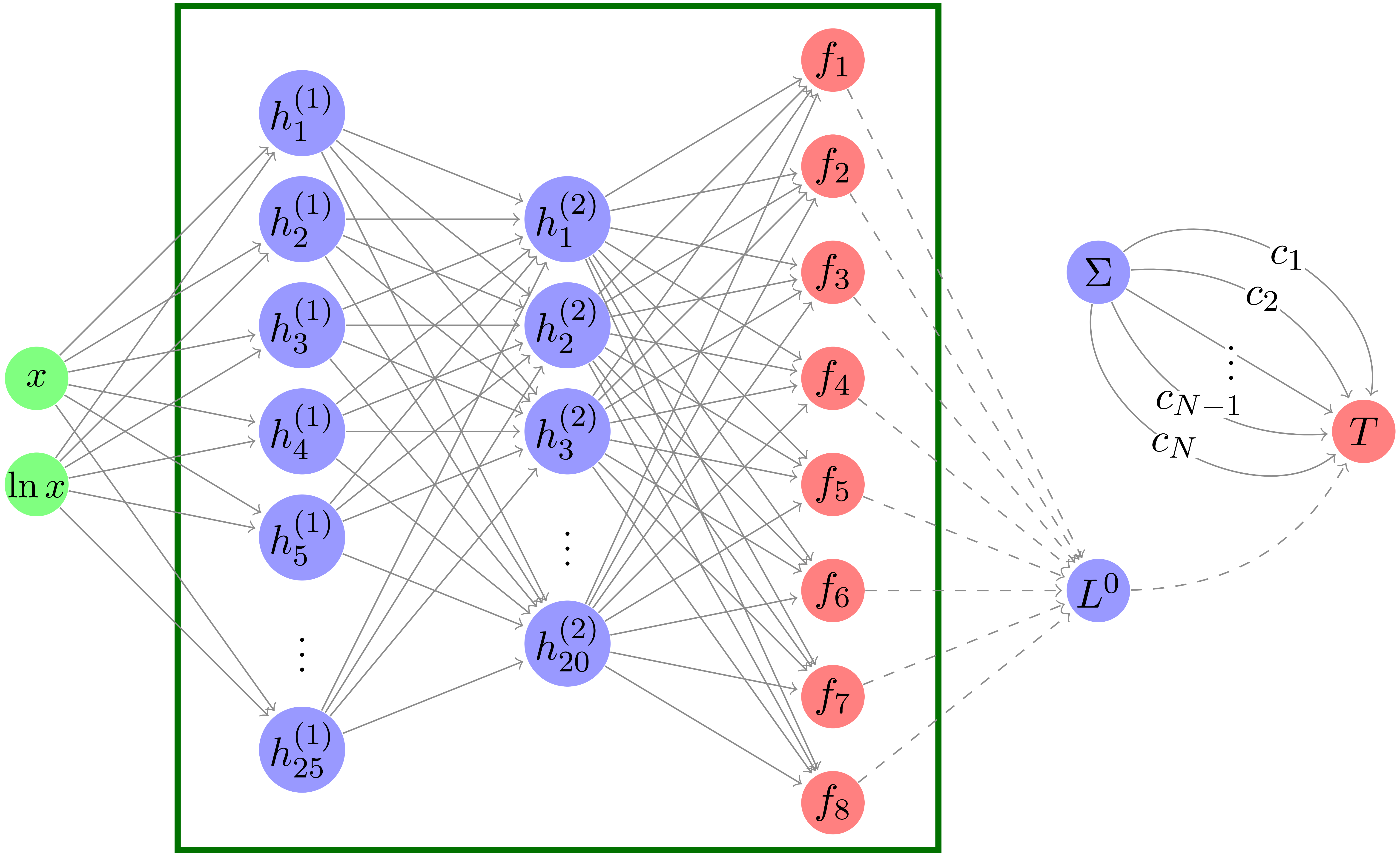
SIMUnet fits: SMEFT

Train only the final layer: reproduce **SMEFT-only fits**



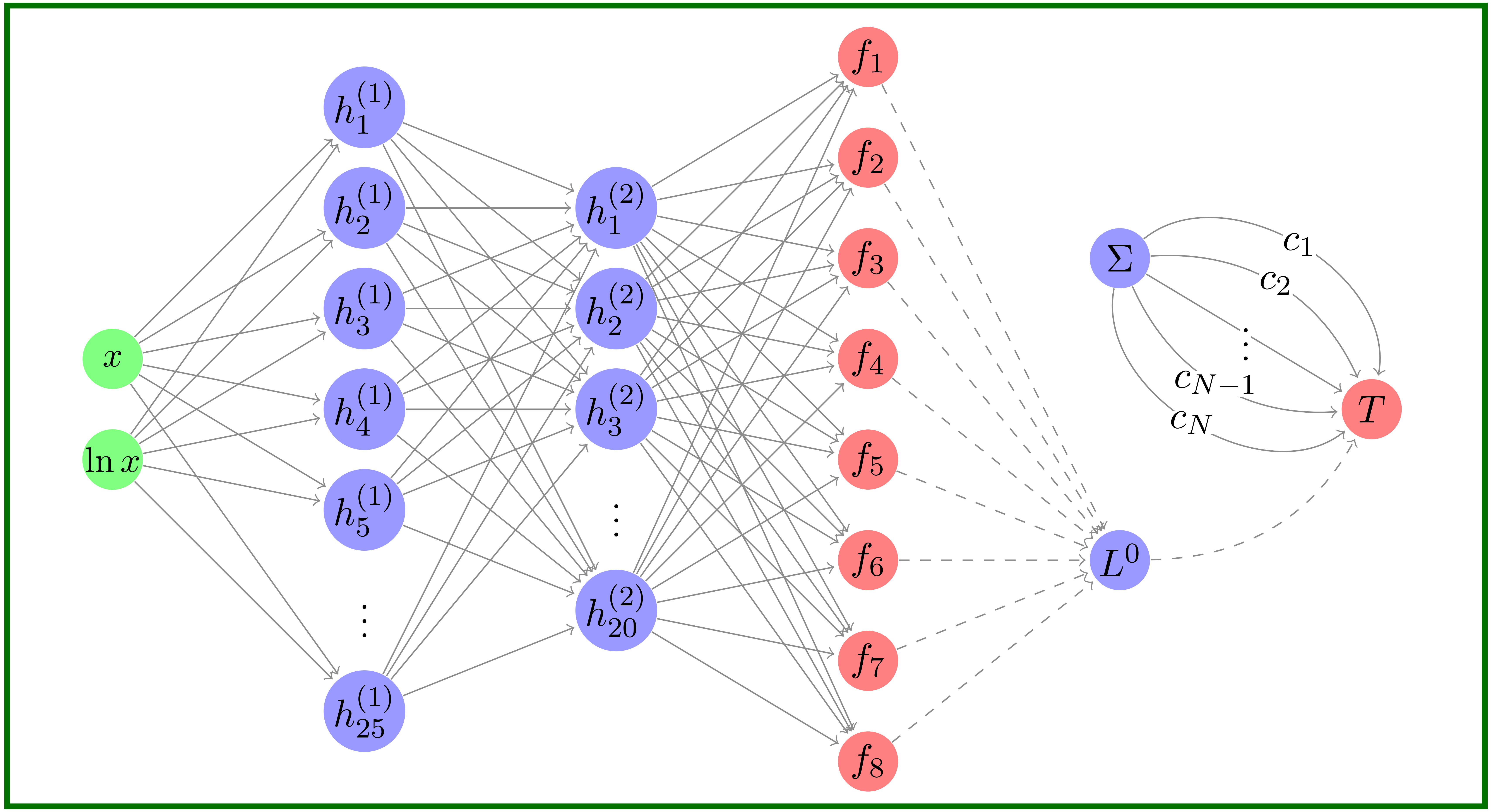
SIMUnet fits: PDFs

Train only the PDF NN weights on all data: reproduce **PDF-only fits**



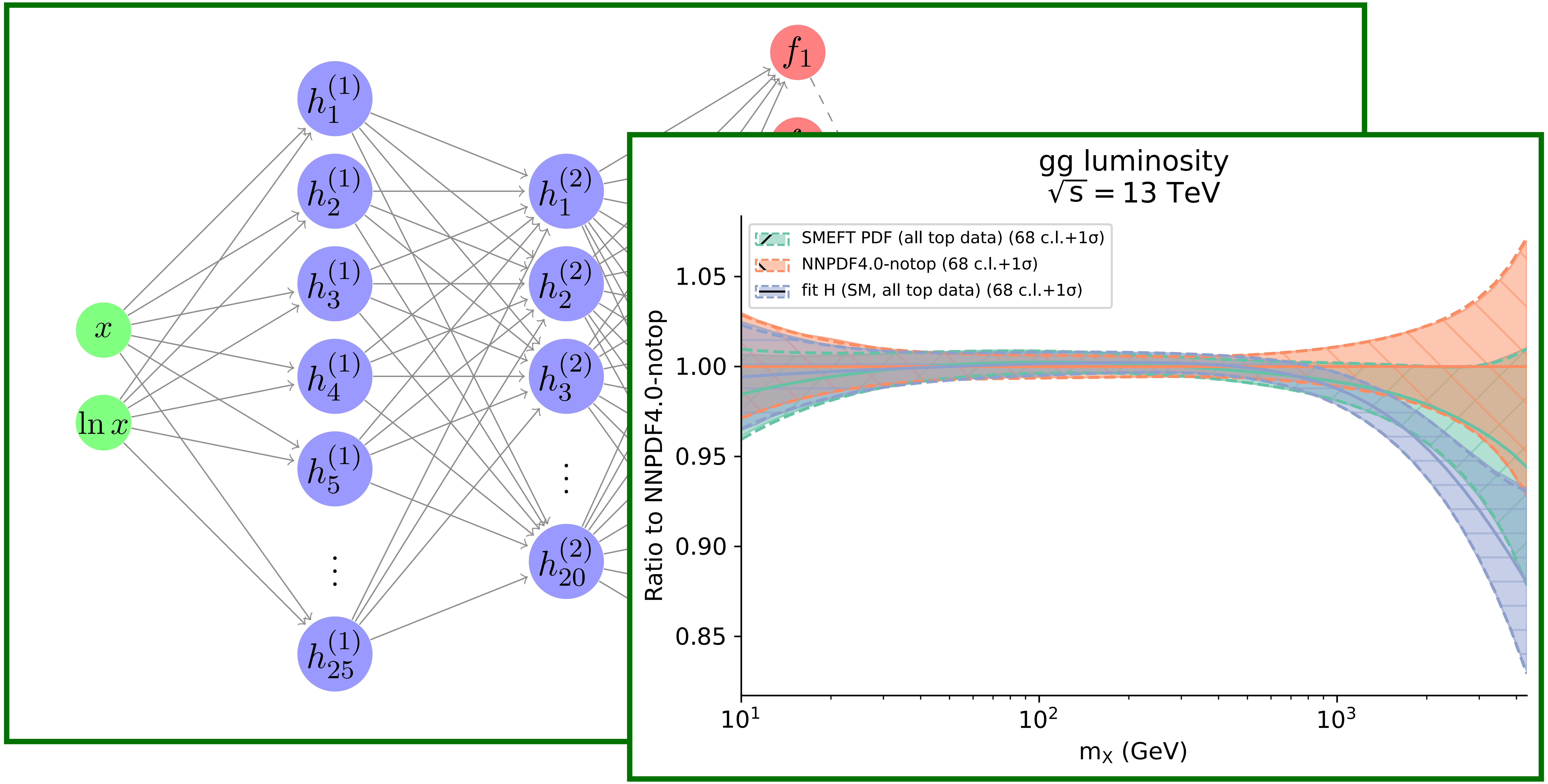
SIMUnet fits: SMEFT and PDFs

Train everything: **simultaneous fit**



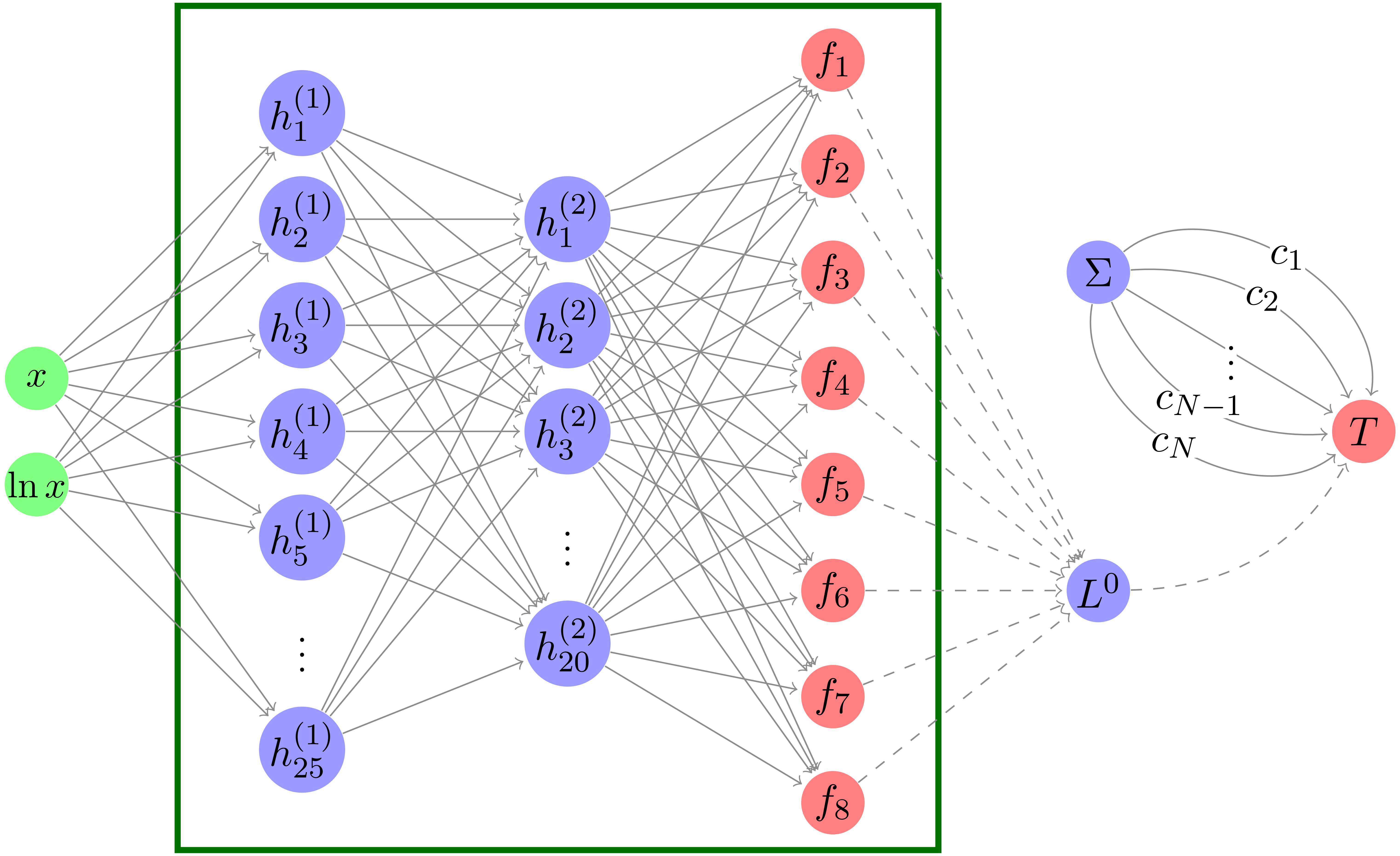
SIMUnet fits: SMEFT and PDFs

Train everything: **simultaneous fit**



SIMUnet fits: new physics contamination

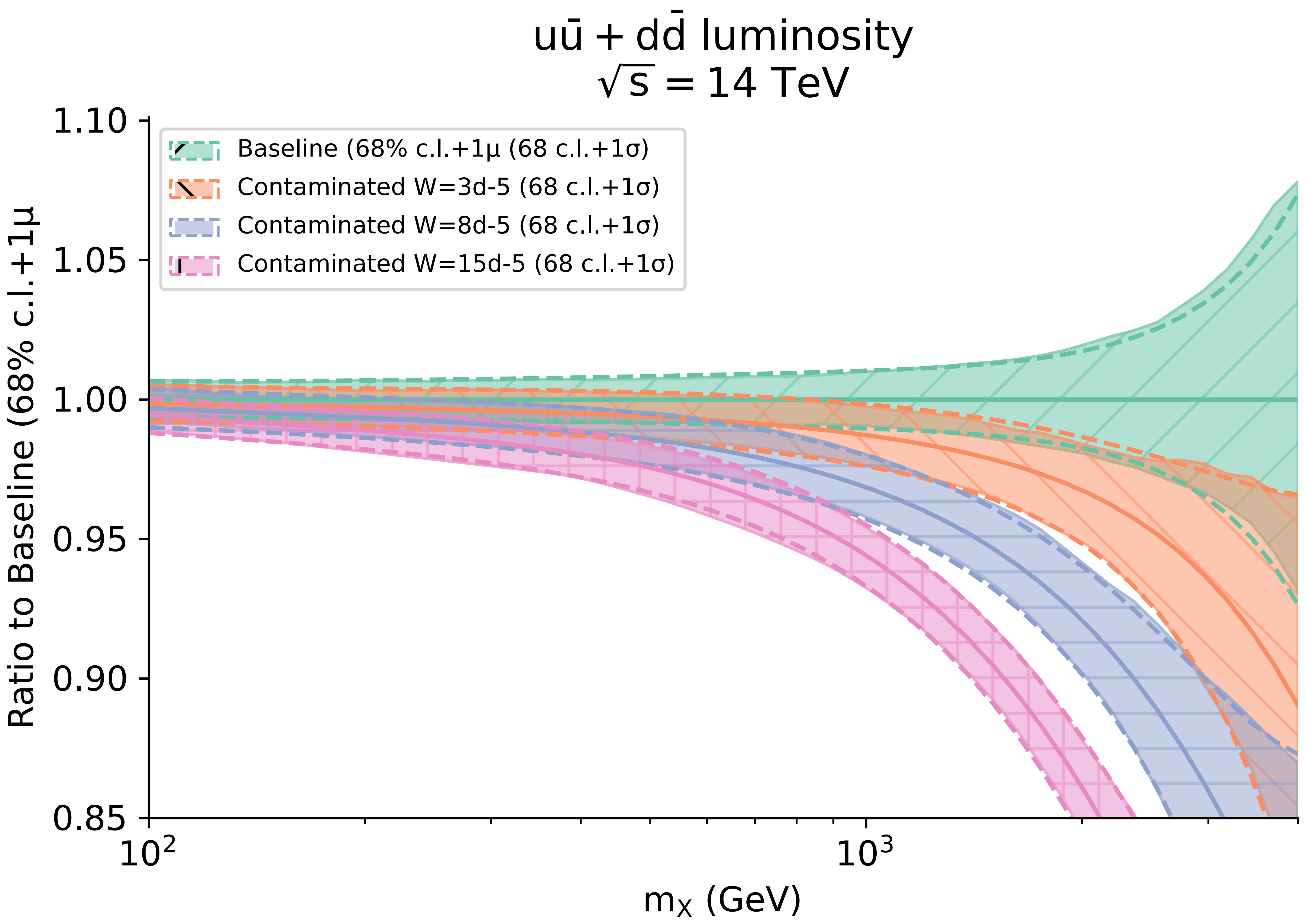
Fit only the PDF to **pseudodata modified by new physics effects** and assess the fit quality: **is new physics absorbed?**

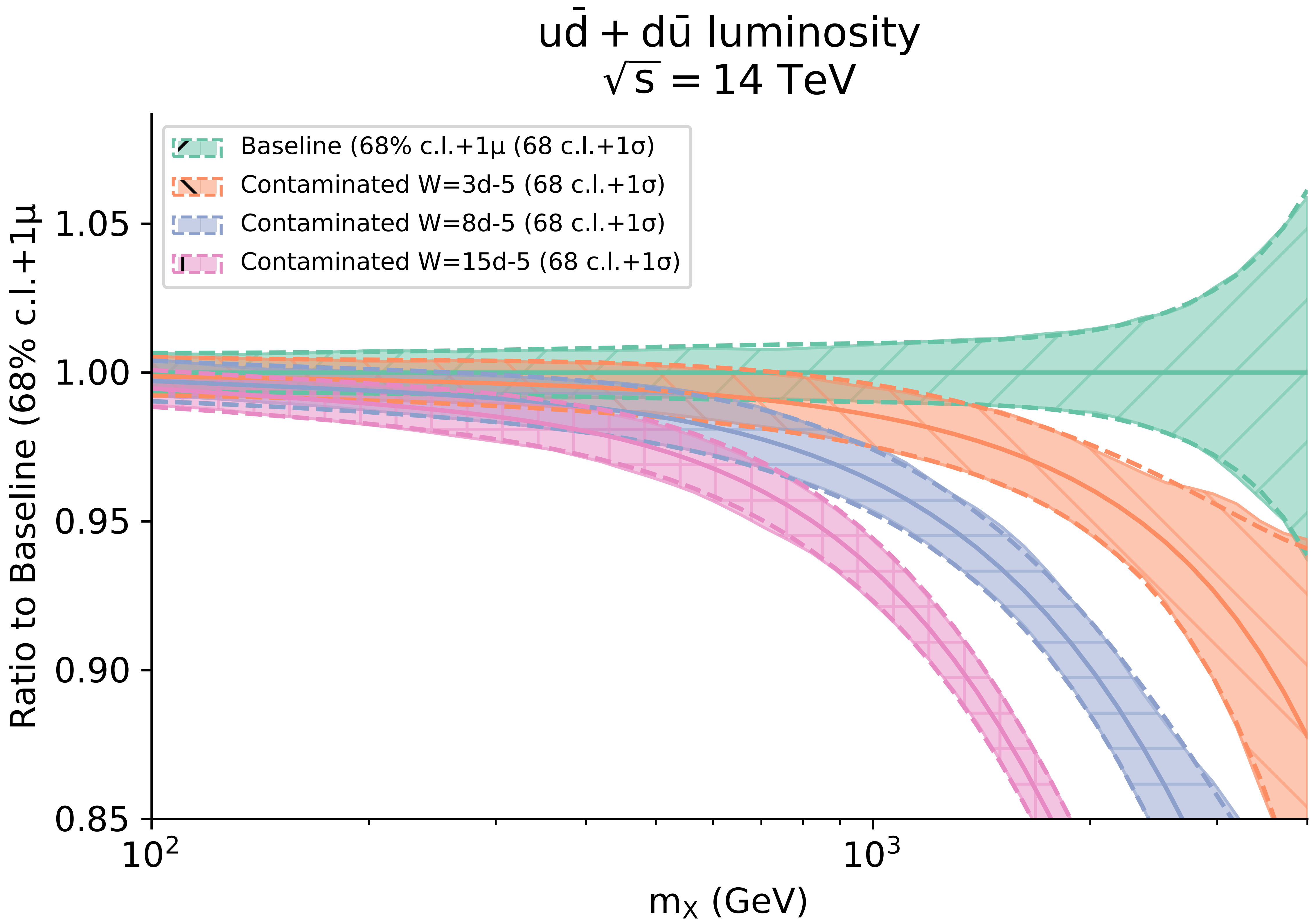


SIMUnet fits: new physics contamination

Fit only the PDF to **pseudodata modified by new physics effects** and assess the fit quality: **is new physics absorbed?**

e.g. HL-LHC high mass DY, *E. Hammou et. al 2307.10370*

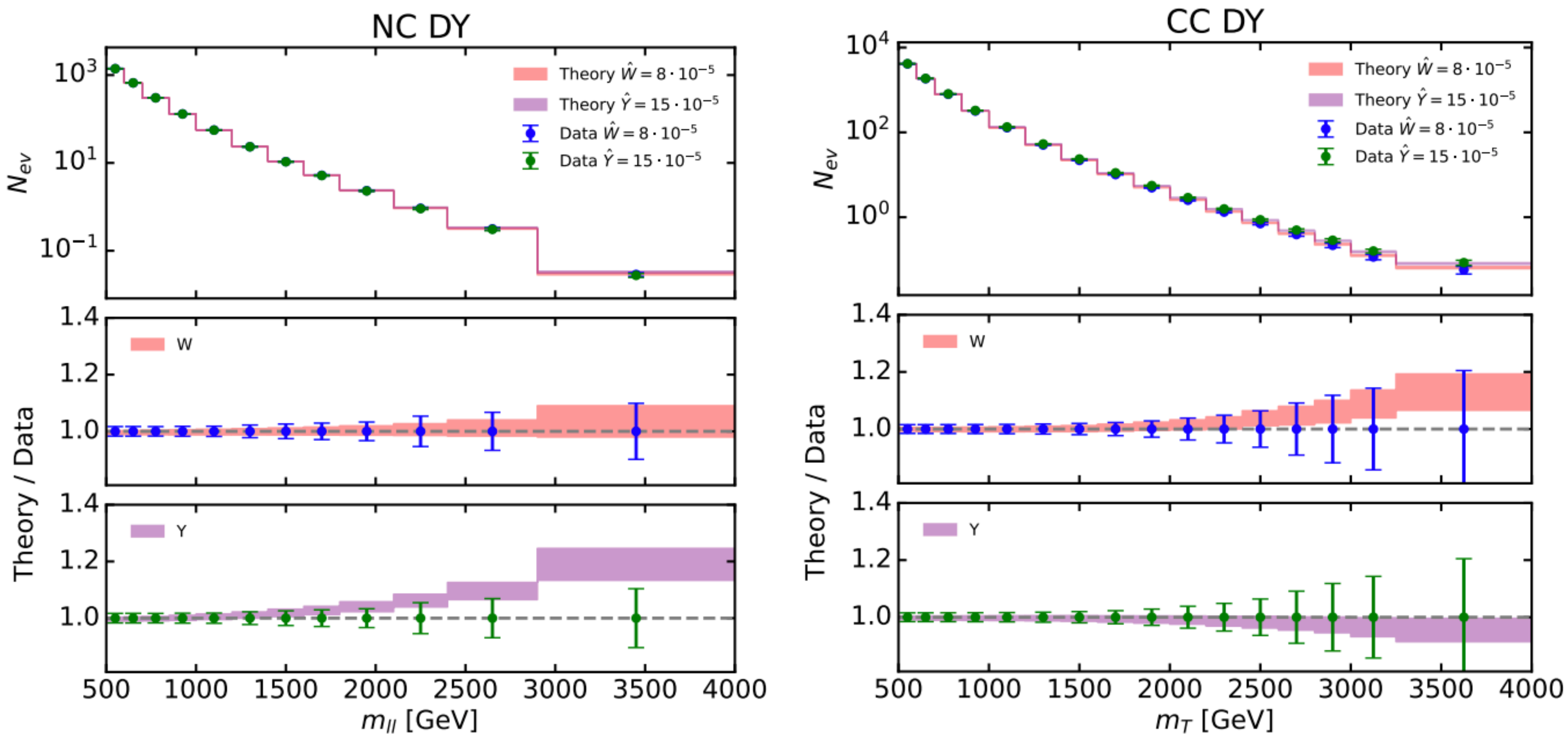




SIMUnet fits: new physics contamination

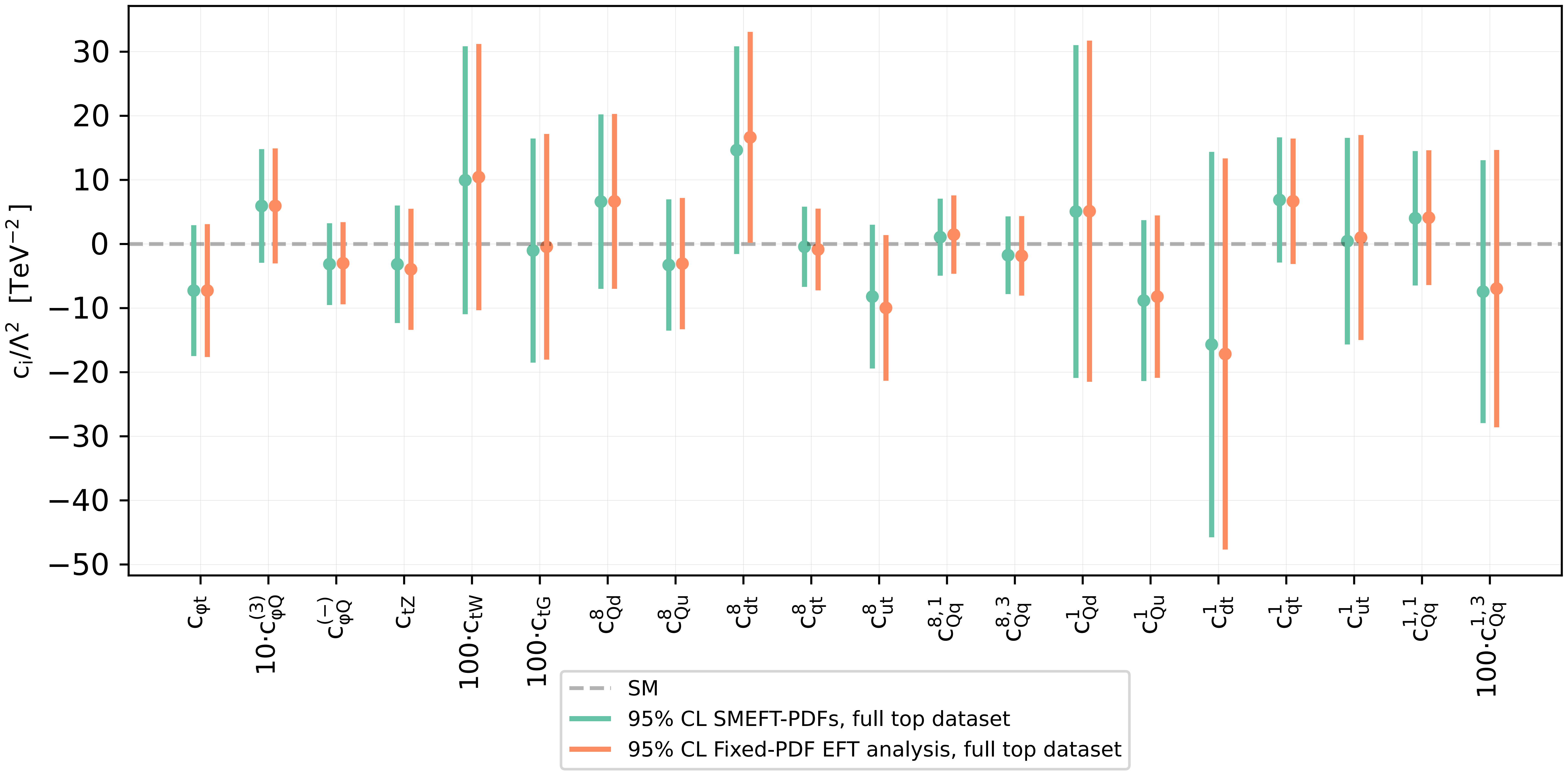
Fit only the PDF to **pseudodata modified by new physics effects** and assess the fit quality: **is new physics absorbed?**

e.g. HL-LHC high mass DY, *E. Hammou et. al 2307.10370*



The SIMUnet release

- **Simultaneous fits of PDFs and linear SMEFT effects**

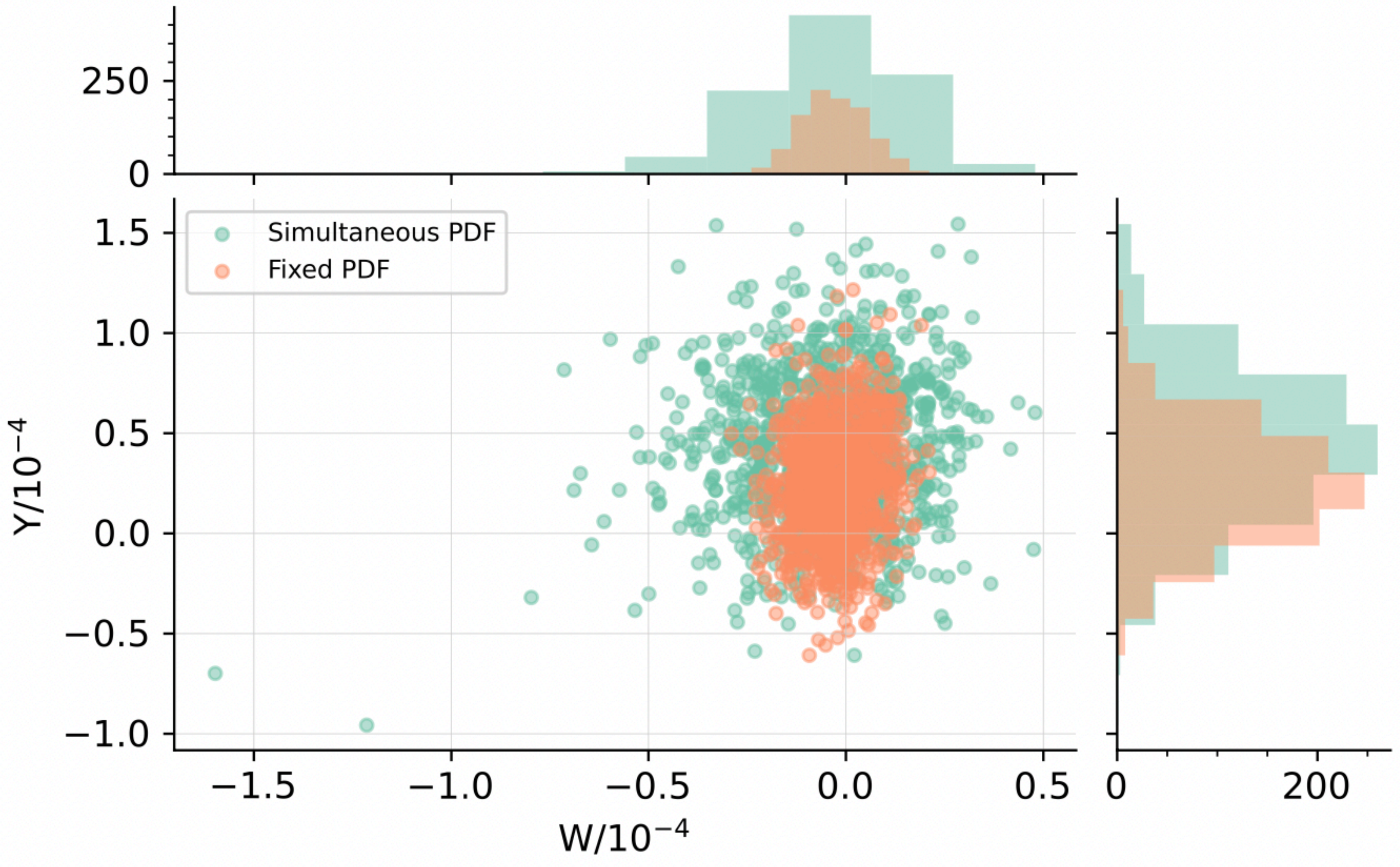


The SIMUnet release

- Simultaneous fits of PDFs and linear SMEFT effects

- **+ Fits of any linear combinations of Wilson coefficients**

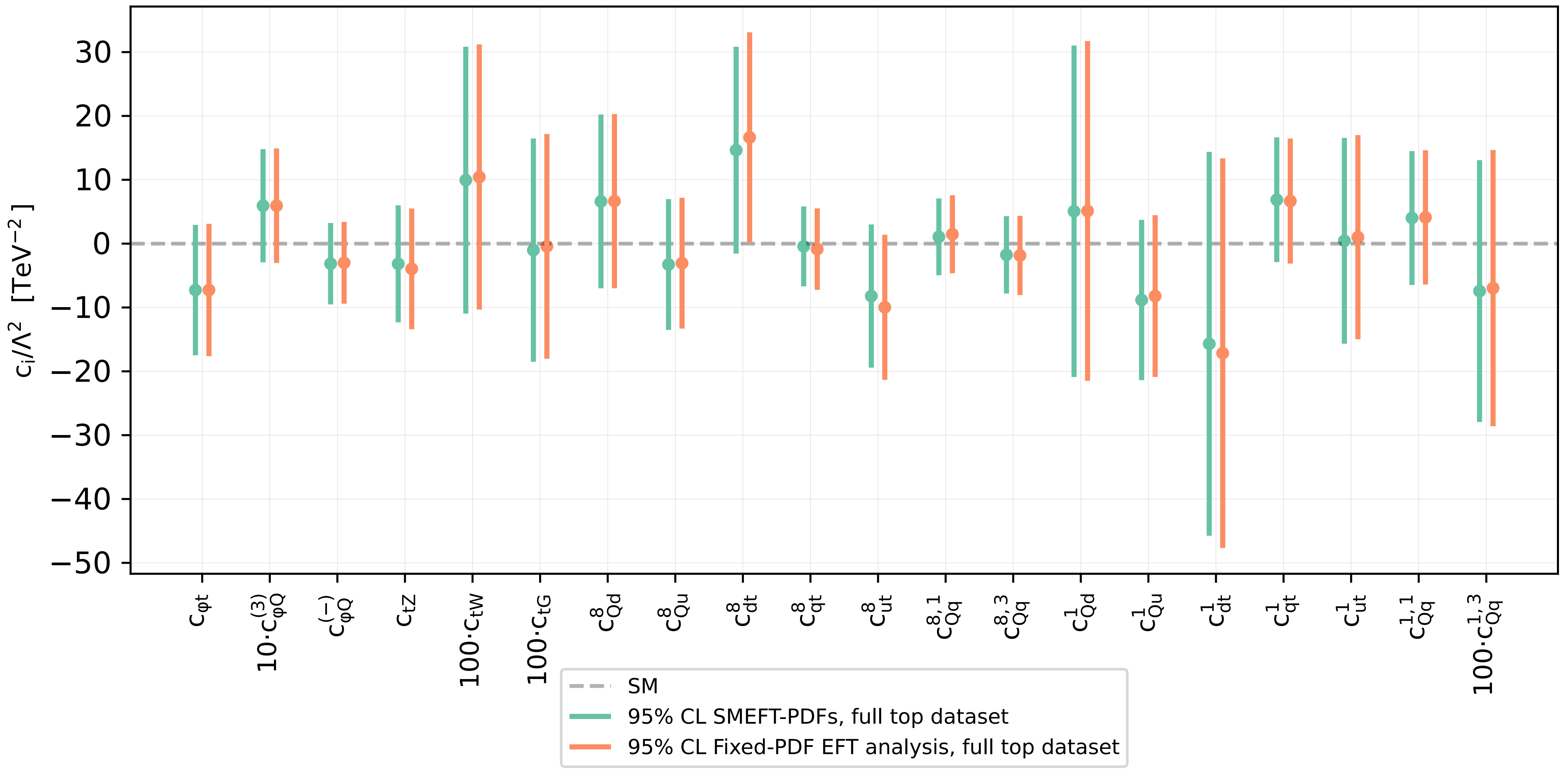
e.g. electroweak oblique parameters *W,Y*



The SIMUnet release

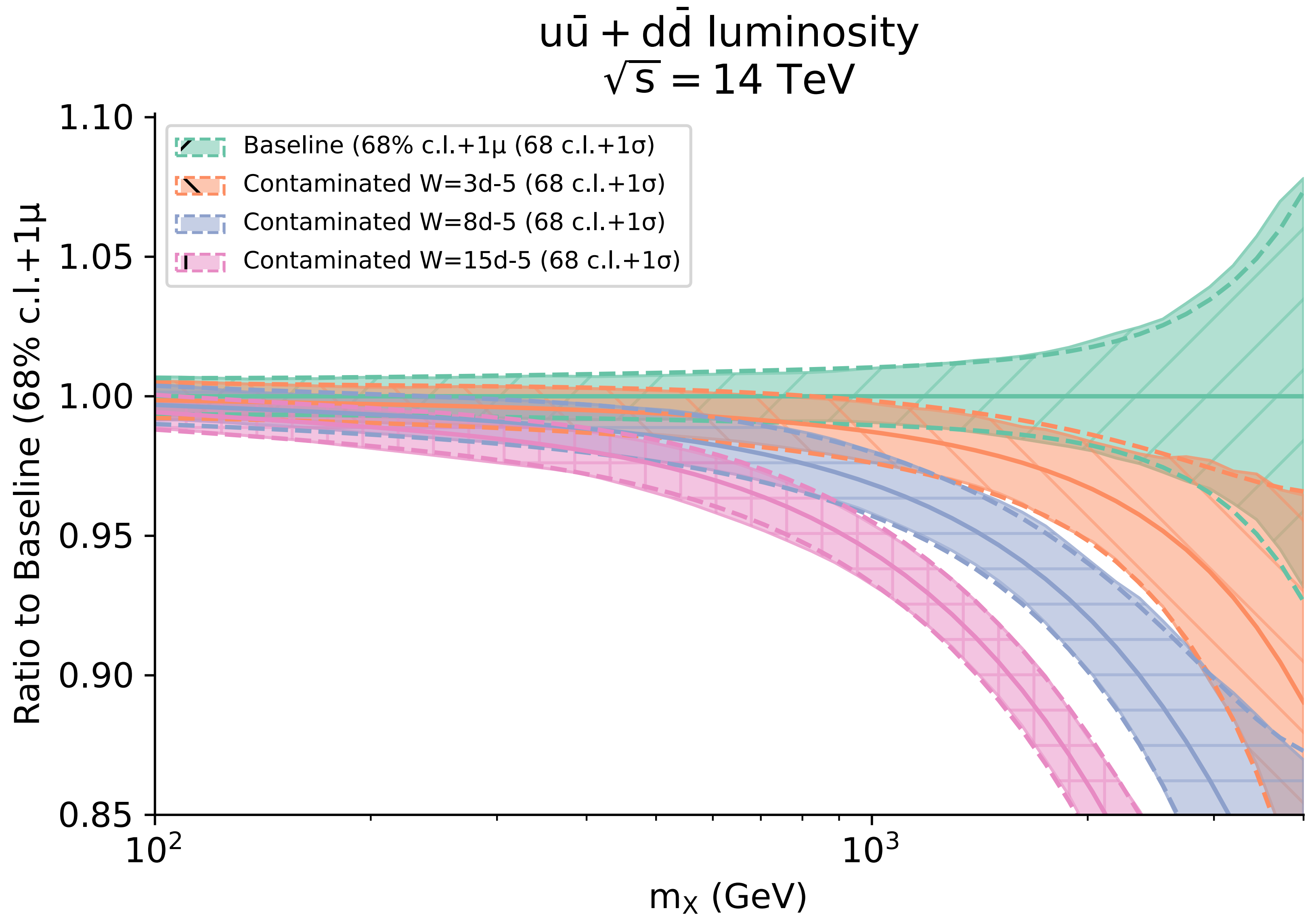
- Simultaneous fits of PDFs and linear SMEFT effects
- + Fits of linear combinations of Wilson coefficients
- **PDF-independent observables**

  E.g. measurements of *W* polarisations in top decay, electroweak precision observables



The SIMUnet release

- Simultaneous fits of PDFs and linear SMEFT effects
- + Fits of linear combinations of Wilson coefficients
- PDF-independent observables
- **Tests for new physics absorption**



The SIMUnet release

- Simultaneous fits of PDFs and linear SMEFT effects
- + Fits of linear combinations of Wilson coefficients
- PDF-independent observables
- Tests for new physics absorption

+ new data from the Higgs, diboson, electroweak, Drell-Yan and top sectors

+ Tutorials, website and documentation

Conclusions

Many examples of the use of ML in EFT analyses

multivariate and unbinned analyses using parametrised classifiers and tree-boosting algorithms; NN-based observables; importance of re-interpretability

https://indico.cern.ch/event/1331690/

Unbinned multivariate observables for global SMEFT analyses from parametrised classifiers - optimal SMEFT constraints: **ML4EFT**

Simultaneous determinations of PDFs and the SMEFT made possible by **SIMUnet**

See also the HEP ML Living Review: https://iml-wg.github.io/HEPML-LivingReview/

Conclusions

Thank you for listening!

Many examples of the use of ML in EFT analyses

multivariate and unbinned analyses using parametrised classifiers and tree-boosting algorithms; NN-based observables; importance of re-interpretability

https://indico.cern.ch/event/1331690/

Unbinned multivariate observables for global SMEFT analyses from parametrised classifiers - optimal SMEFT constraints: **ML4EFT**

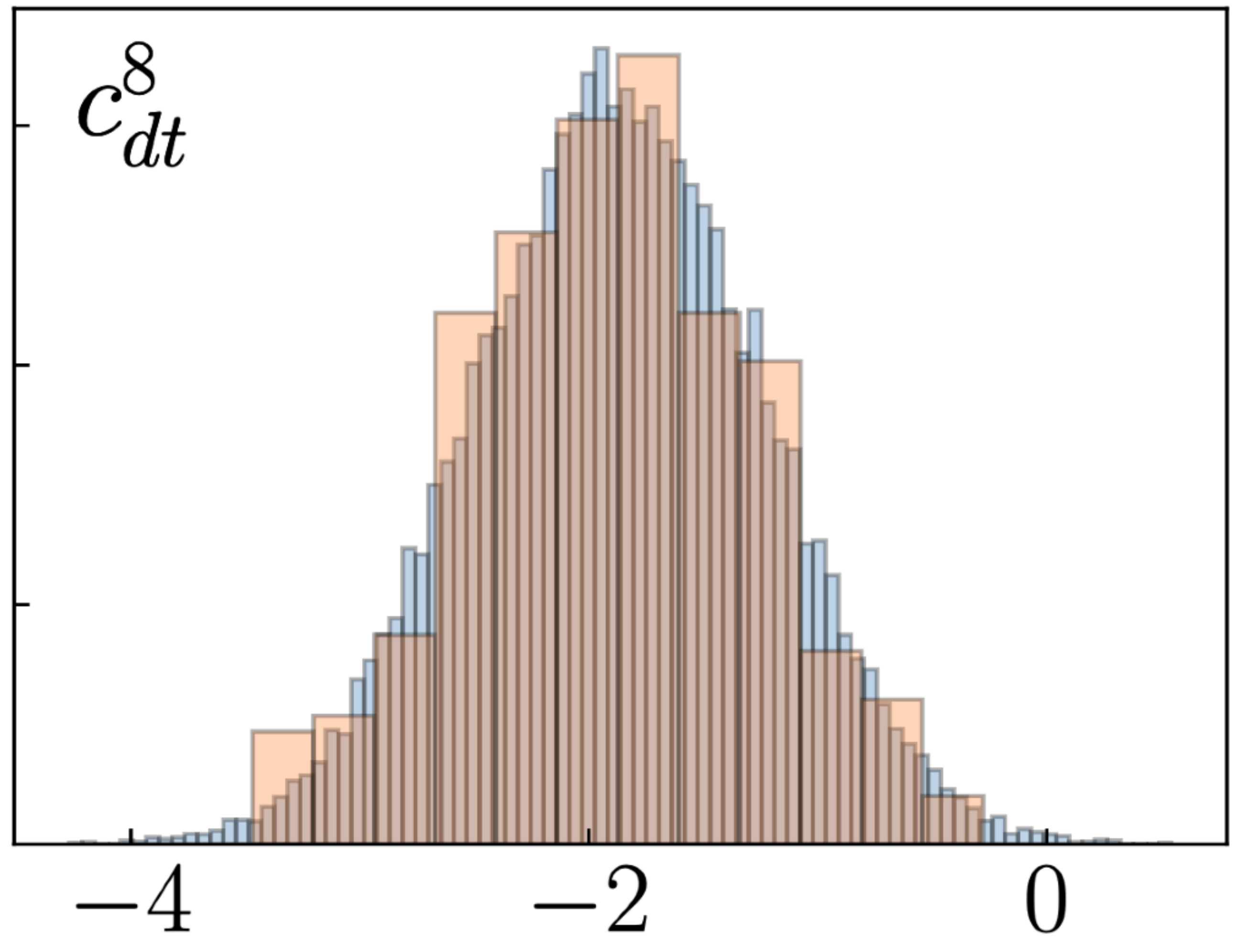
Simultaneous determinations of PDFs and the SMEFT made possible by **SIMUnet**

See also the HEP ML Living Review: https://iml-wg.github.io/HEPML-LivingReview/

Backup

The Monte Carlo Replica Method

Consider fitting 1 Wilson coefficient *c* to 1 datapoint σ_{exp}: define $\chi^2 = \dfrac{(\sigma(c) - \sigma_{\text{exp}})^2}{\delta\sigma^2}$



1. Resample: $\tilde{\sigma}_{\text{exp}} \sim \mathcal{N}(\sigma_{\text{exp}}, \Sigma)$

2. Minimise: $\bar{c} = \arg\min_c \dfrac{(\sigma(c) - \tilde{\sigma}_{\text{exp}})^2}{\delta\sigma^2}$

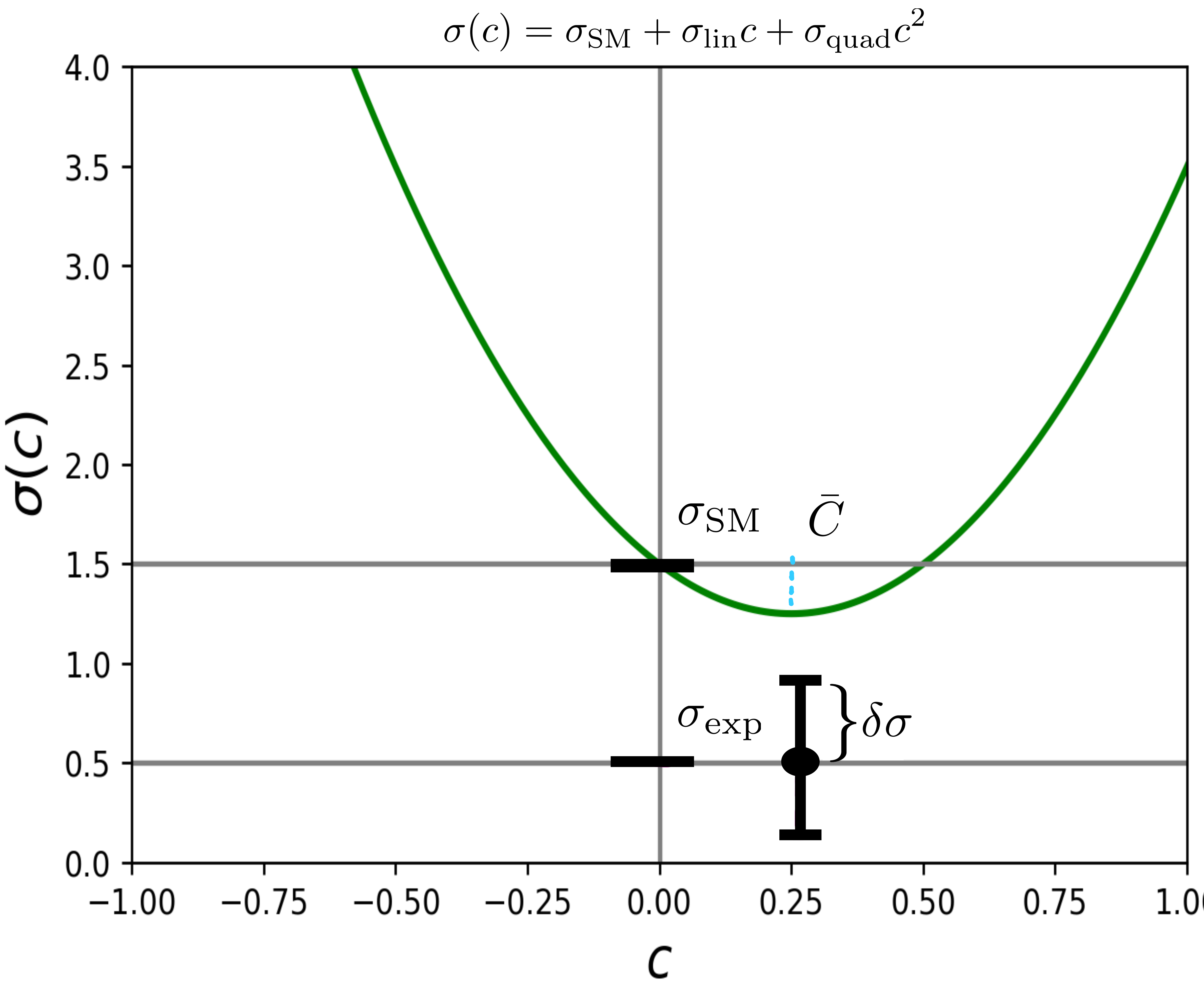
3. Repeat, and treat the sample $\{\bar{c}\}$ as a sample from the Bayesian posterior $p(c|D)$

- Often used in the context of PDF fitting and SMEFT fitting, e.g. 2109.02653, 1901.05965

The Monte Carlo Replica Method

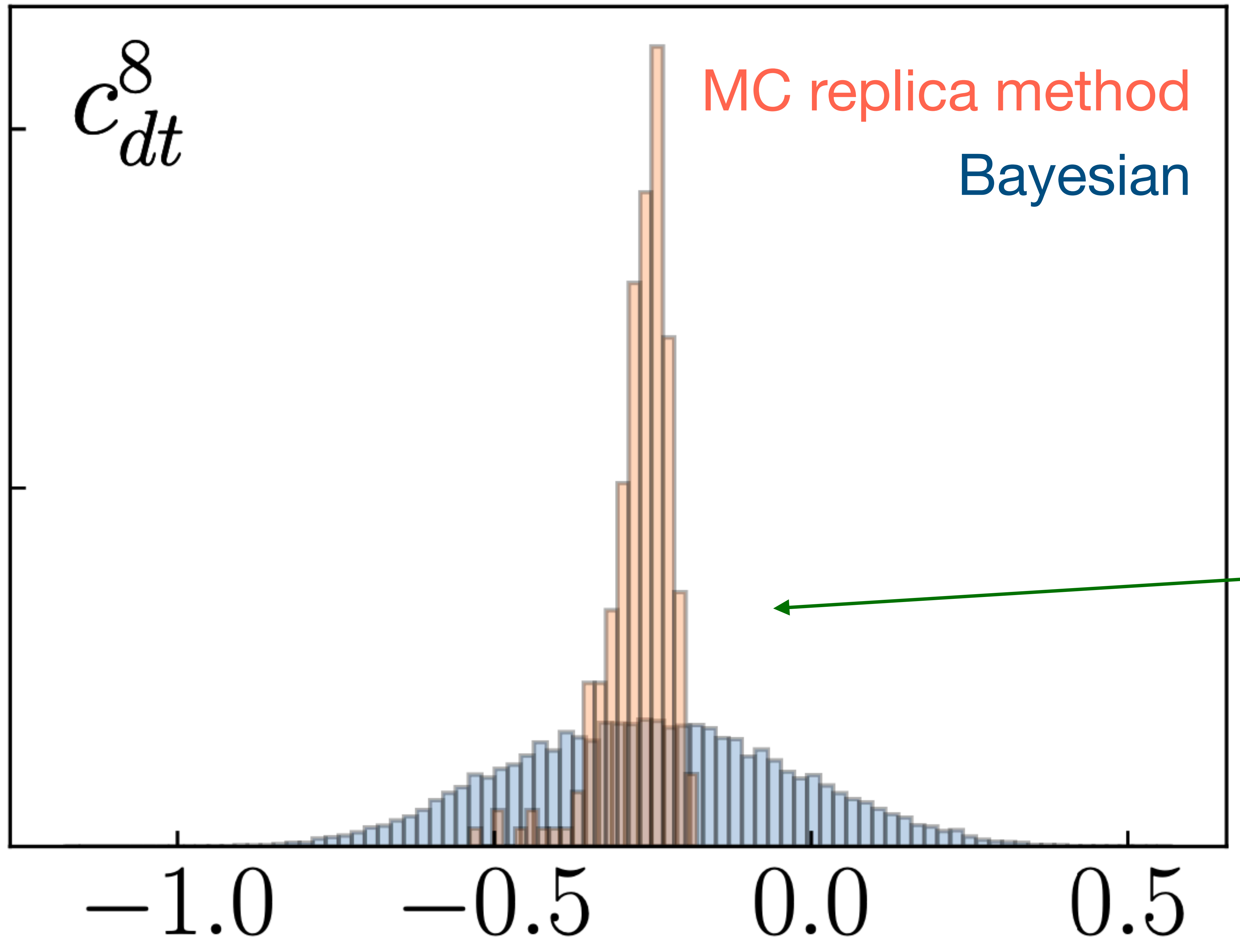
Problem: in the presence of a quadratic theory, often the minimum χ^2 will be given by the same $\bar{c}$.

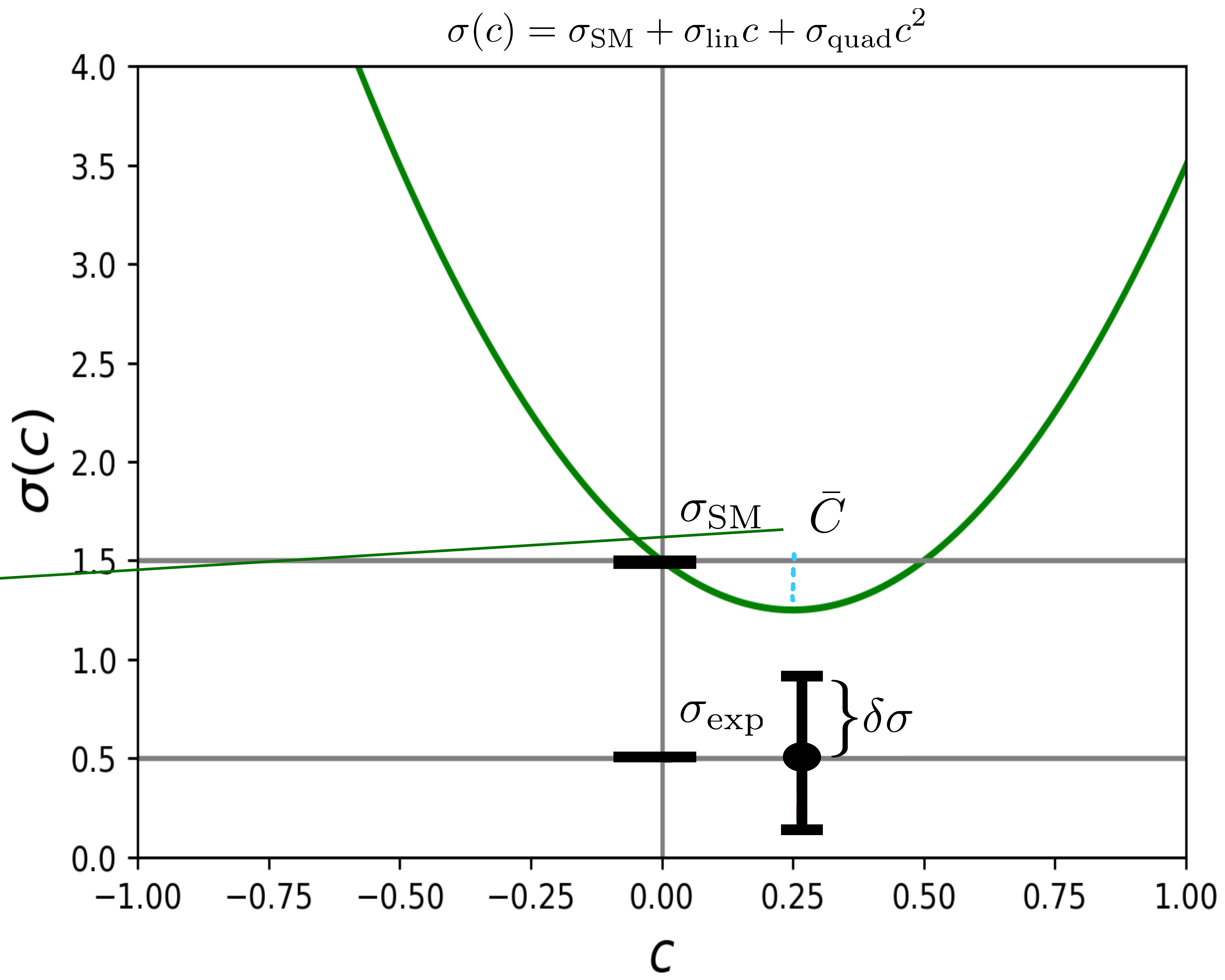
2. Minimise: $\bar{c} = \arg\,\min_c \dfrac{(\sigma(c) - \tilde{\sigma}_{\rm exp})^2}{\delta\sigma^2}$



The Monte Carlo Replica Method

Problem: in the presence of a quadratic theory, often the minimum χ^2 will be given by the same $\bar{c}$.





ML4EFT

For parameter estimation, we would like to be able to calculate the **likelihood:** $$\mathcal{L}(D|\mathbf{c}) \propto \prod_{i=1}^{N_{ev}} f_\sigma(\mathbf{x}_i, \mathbf{c})$$

where $$f_\sigma(\mathbf{x}, \mathbf{c}) = \frac{1}{\sigma(\mathbf{x}, \mathbf{c})} \frac{d\sigma(\mathbf{x}, \mathbf{c})}{d\mathbf{x}}$$

$$D = \{\mathbf{x}_i\} \qquad \mathbf{x}_i = \{m_{t\bar{t}}, p_T^{\ell_1}, p_T^{\ell_2}, \Delta\eta_{\ell_1,\ell_2}, \Delta\phi_{\ell_1,\ell_2}, ...\}$$

multi-differential cross section in **all features**

However: analytical calculation of $\mathcal{L}$ is intractable in most realistic cases.

Instead: approximate $\mathcal{L}$ using neural networks

ML4EFT

Train a classifier ***g*** to distinguish the SM from the SMEFT:

$$L[g(\mathbf{x},\mathbf{c})] = -\sum_{i=1}^{N_{\rm ev}^{\rm SMEFT}} \frac{d\sigma(\mathbf{x}_i,\mathbf{c})}{dx}\log(1-g(\mathbf{x}_i,\mathbf{c})) - \sum_{i=1}^{N_{\rm ev}^{\rm SM}} \frac{d\sigma(\mathbf{x}_i,\mathbf{0})}{dx}\log(g(\mathbf{x}_i,\mathbf{0}))$$

SMEFT training pseudodata sample

SM training pseudo data sample

ML4EFT

Train a classifier ***g*** to distinguish the SM from the SMEFT:

$$L[g(\mathbf{x},\mathbf{c})] = -\sum_{i=1}^{N_{\rm ev}^{\rm SMEFT}} \frac{d\sigma(\mathbf{x}_i,\mathbf{c})}{dx}\log(1-g(\mathbf{x}_i,\mathbf{c})) - \sum_{i=1}^{N_{\rm ev}^{\rm SM}} \frac{d\sigma(\mathbf{x}_i,\mathbf{0})}{dx}\log(g(\mathbf{x}_i,\mathbf{0}))$$

SMEFT training pseudodata sample

SM training pseudo data sample

$$\frac{\delta L}{\delta g} = 0 \Rightarrow g(\mathbf{x},\mathbf{c}) = \left(1 + \frac{d\sigma(\mathbf{x},\mathbf{c})}{dx}\Big/\frac{d\sigma(\mathbf{x},\mathbf{0})}{dx}\right)^{-1} \equiv \frac{1}{1+r_\sigma(\mathbf{x},\mathbf{c})}$$

ML4EFT

Train a classifier ***g*** to distinguish the SM from the SMEFT:

$$L[g(\mathbf{x},\mathbf{c})] = -\sum_{i=1}^{N_{\rm ev}^{\rm SMEFT}} \frac{d\sigma(\mathbf{x}_i,\mathbf{c})}{dx}\log(1-g(\mathbf{x}_i,\mathbf{c})) - \sum_{i=1}^{N_{\rm ev}^{\rm SM}} \frac{d\sigma(\mathbf{x}_i,\mathbf{0})}{dx}\log(g(\mathbf{x}_i,\mathbf{0}))$$

SMEFT training pseudodata sample

SM training pseudo data sample

$$\frac{\delta L}{\delta g} = 0 \Rightarrow g(\mathbf{x},\mathbf{c}) = \left(1 + \frac{d\sigma(\mathbf{x},\mathbf{c})}{dx}\Big/\frac{d\sigma(\mathbf{x},\mathbf{0})}{dx}\right)^{-1} \equiv \frac{1}{1+r_\sigma(\mathbf{x},\mathbf{c})}$$

In the limit of infinite training samples, the decision boundary is 1:1 with the likelihood