Staying on Top of SMEFT-Likelihood Analyses

Improving global SMEFT analyses using public likelihoods and more

Nikita Schmal

03/12/2024, 8th General Meeting of the LHC EFT Working Group

Based on
[2208.08454] Ilaria Brivio, Sebastian Brugisser, Nina Elmer, Emma Geoffray, Michel Luchmann
[2312.12502] Nina Elmer, Maeve Madigan, Tilman Plehn, Nikita Schmal
[2411.00942] Theo Heimel, Tilman Plehn, Nikita Schmal

UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

A brief history of SFITTER

- Used for various global SMEFT analyses
  - Higgs, EWPOs, Di-Boson data [1505.05516, 1812.07587, 2208.08454]
  - Top data [1910.03606, 2312.12502]
- **Fully correlated** systematic uncertainties within experiments
- Allows for both **profiling and marginalization** methods
- Mapping of likelihood using MCMC (not anymore)

SMEFT in SFitter

- **SMEFT**: model agnostic approach for BSM

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

- operators of **dimension 6**, contributions up to quadratic order
- SMEFT predictions expressed via **simple bilinear operation**

$$p_i^{(b)} = W_{ijk} C_j^{(b)} \tilde{C}_k^{(b)} + B_i$$

- Further assumptions depending on specific analysis

SFITTER dataset

[2312.12502]

Consider **Top** sector with **22 Wilson coefficients**

- 122 datapoints
- many distributions (including boosted top)
- includes $t\bar{t}$, $t\bar{t}Z$, $t\bar{t}W$ and SingleTop
- also top decays, charge asymmetries

[2208.08454]

and **Higgs** sector with **20 Wilson coefficients**

- 311 Higgs datapoints
- 43 Di-Boson datapoints
- 14 EWPOs (linear SMEFT contr.)
- 4 high kinematic measurements

SFITTER dataset

[2312.12502]

Consider Top sector with 22 Wilson coefficients

- 122 datapoints
- many distributions (including boosted top)
- includes $t\bar{t}$, $t\bar{t}Z$, $t\bar{t}W$ and SingleTop
- also top decays, charge asymmetries

Includes measurements with **public likelihoods**

[2208.08454]

and Higgs sector with 20 Wilson coefficients

- 311 Higgs datapoints
- 43 Di-Boson datapoints
- 14 EWPOs (linear SMEFT contr.)
- 4 high kinematic measurements

SFITTER likelihood

Single measurement likelihood

$$L_{\text{excl}} = \text{Pois}(d\,|\,p(C,\theta,b))\ \text{Pois}(b_{CR}\,|\,bk)\ \prod_i \mathscr{C}_i(\theta_i,\sigma_i)$$

$\mathscr{C}_i(\theta_i,\sigma_i)$ →

- **Systematic:** **fully correlated** Gaussian
- **Statistical:** uncorrelated Gaussian
- **Theory:** $F(x\,|\,\sigma,\mu) = \frac{1}{2\sigma}\Theta\left[x-(\mu-\sigma)\right]\Theta\left[(\mu+\sigma)-x\right]$

- Remove nuisance parameters via **profiling** or **marginalization**

$$L_{\text{prof}} = \max_{\theta,b} \mathscr{L}_{\text{excl}} \qquad L_{\text{marg}} = \int d\theta db\ \mathscr{L}_{\text{excl}}$$

SFitter likelihood

Concerning correlations

- Global analyses combine various different analyses

$$\mathcal{L}_{\text{excl,full}} = \prod_c \text{Pois}(d_c|p_c)\text{Pois}(b_{CR_c}|b_c\, k_c) \prod_i \mathcal{C}(\theta_{i,c}, \sigma_{i,c})$$

Systematic uncertainties
Beam
Background (Separate for each channel)
ETmis
Jets
Leptons
LightTagging
Luminosity
Pileup
Trigger
Tune
bTagging
partonShower
tTagging
tauTagging

- Take into account correlations

$$C_{ij} = \frac{\sum_{\text{syst}} \rho_{ij}\ \sigma_{i,\text{syst}}\sigma_{j,\text{syst}}}{\sigma_{i,\text{exp}}\sigma_{j,\text{exp}}} \qquad \text{with} \qquad \sigma^2_{i,\text{exp}} = \sum_{\text{syst}} \sigma^2_{i,\text{syst}} + \sum_{\text{pois}} \sigma^2_{i,\text{pois}}$$

- **Assumption:** Fully correlated systematics between measurements

SFITTER likelihood

Concerning correlations

- Global analyses combine various different analyses

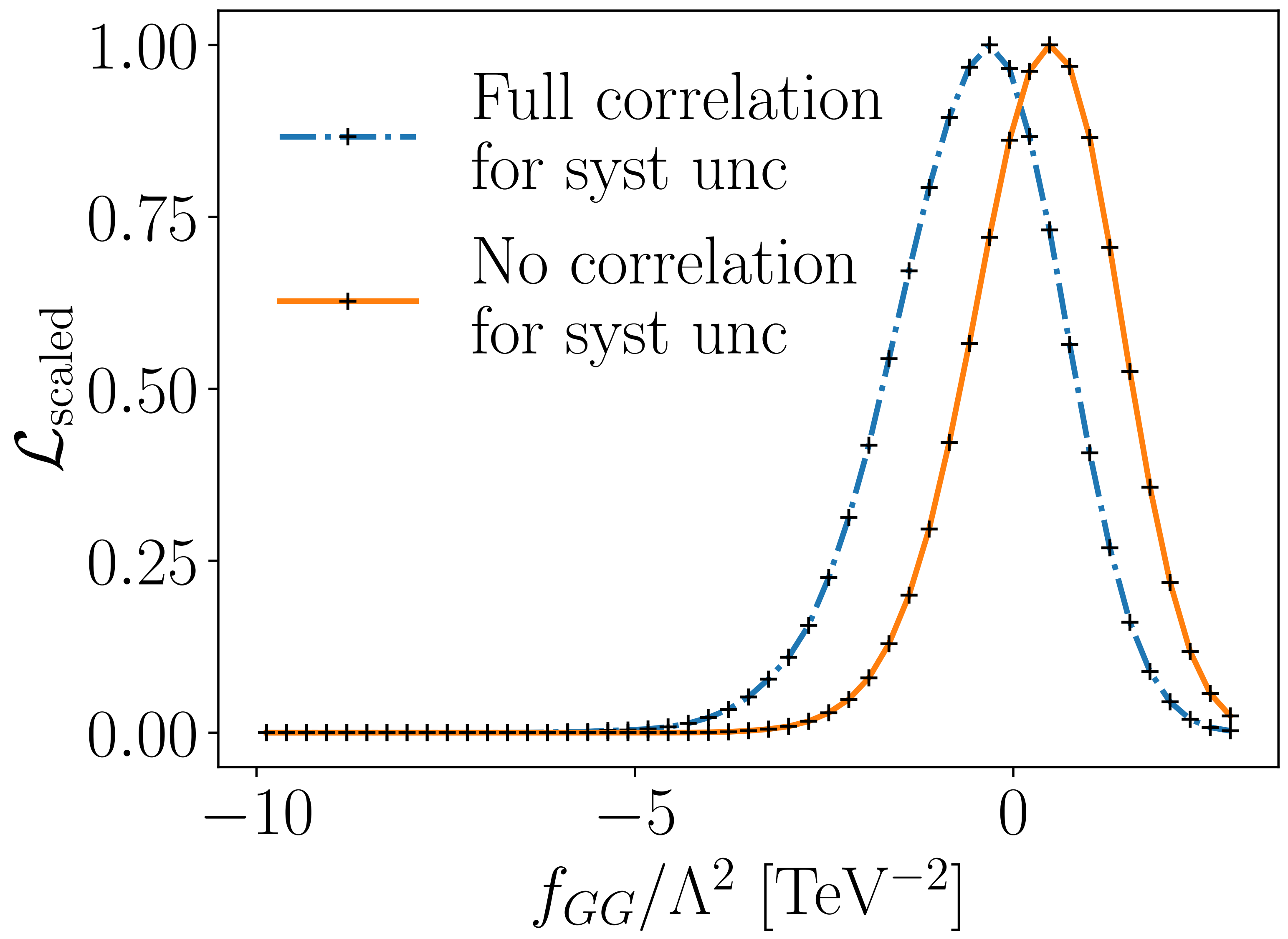
$$\mathcal{L}_{\text{excl,full}} = \prod_c \text{Pois}(d_c|p_c)\text{Pois}(b_{CR_c}|b_c\, k_c) \prod_i \mathcal{C}(\theta_{i,c}, \sigma_{i,c})$$

- Take into account correlations

$$C_{ij} = \frac{\sum_{\text{syst}} \rho_{ij}\ \sigma_{i,\text{syst}}\sigma_{j,\text{syst}}}{\sigma_{i,\text{exp}}\sigma_{j,\text{exp}}} \qquad \text{with} \qquad \sigma^2_{i,\text{exp}} = \sum_{\text{syst}} \sigma^2_{i,\text{syst}} + \sum_{\text{pois}} \sigma^2_{i,\text{pois}}$$

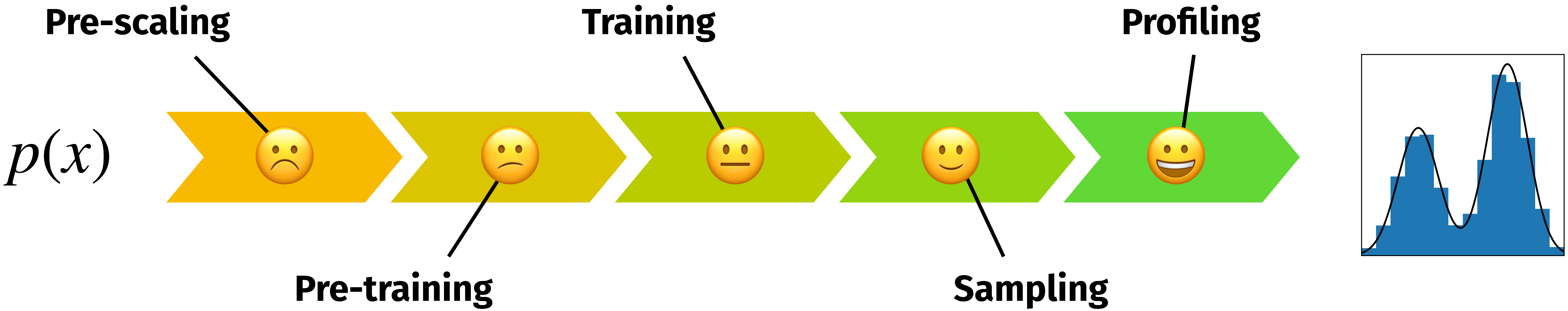
- **Assumption:** Fully correlated systematics between measurements

[arXiv:2208.08454]



The five steps to happiness

- We finally extract constraints on the WCs from this likelihood
- SFitter makes use of MCMC, but that is **slow**
- **Alternative:** Train a **normalizing flow** to speed this up
  - You should read [2411.00942]



Public likelihoods

[2006.13076]

EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

ATLAS EXPERIMENT

CERN

Phys. Lett. B 810 (2020) 135797
DOI: 10.1016/j.physletb.2020.135797
CERN-EP-2020-096
10th November 2020

Measurement of the $t\bar{t}$ production cross-section in the lepton+jets channel at $\sqrt{s}$ = 13 TeV with the ATLAS experiment

$t\bar{t}$

$t\bar{t}Z$

[2209.08990]

Eur. Phys. J. C (2021) 81:737
https://doi.org/10.1140/epjc/s10052-021-09439-4
THE EUROPEAN PHYSICAL JOURNAL C
Check for updates
Regular Article - Experimental Physics

Measurements of the inclusive and differential production cross sections of a top-quark–antiquark pair in association with a Z boson at $\sqrt{s} = 13$ TeV with the ATLAS detector

ATLAS Collaboration*
CERN, 1211 Geneva 23, Switzerland

[2103.12603]

EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

ATLAS EXPERIMENT

CERN

JHEP 06 (2023) 191
DOI: 10.1007/JHEP06(2023)191
CERN-EP-2022-116
6th July 2023

Measurement of single top-quark production in the s-channel in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector

s-channel single-top

Public likelihoods

Quick overview

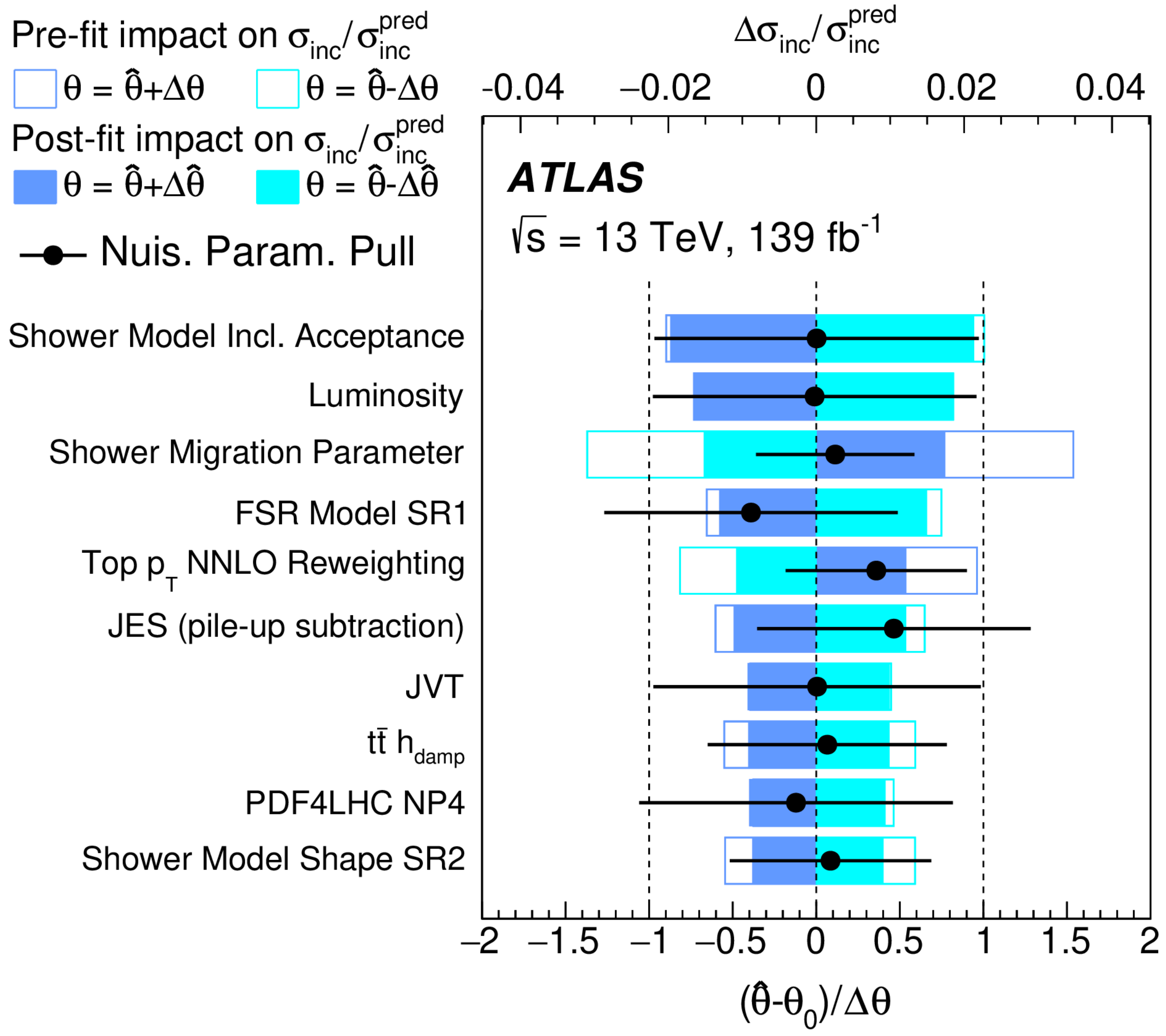
- Likelihood published in the **HistFactory** format

$$\mathcal{L}(n_{cb}, a_\chi | \eta, \chi) = \prod_{c \in \text{channels}} \prod_{b \in \text{bins}} \text{Pois}(n_{cb} | \nu_{cb}(\eta, \chi)) \prod_{\chi \in \vec{\chi}} \mathcal{C}_\chi(a_\chi | \chi)$$

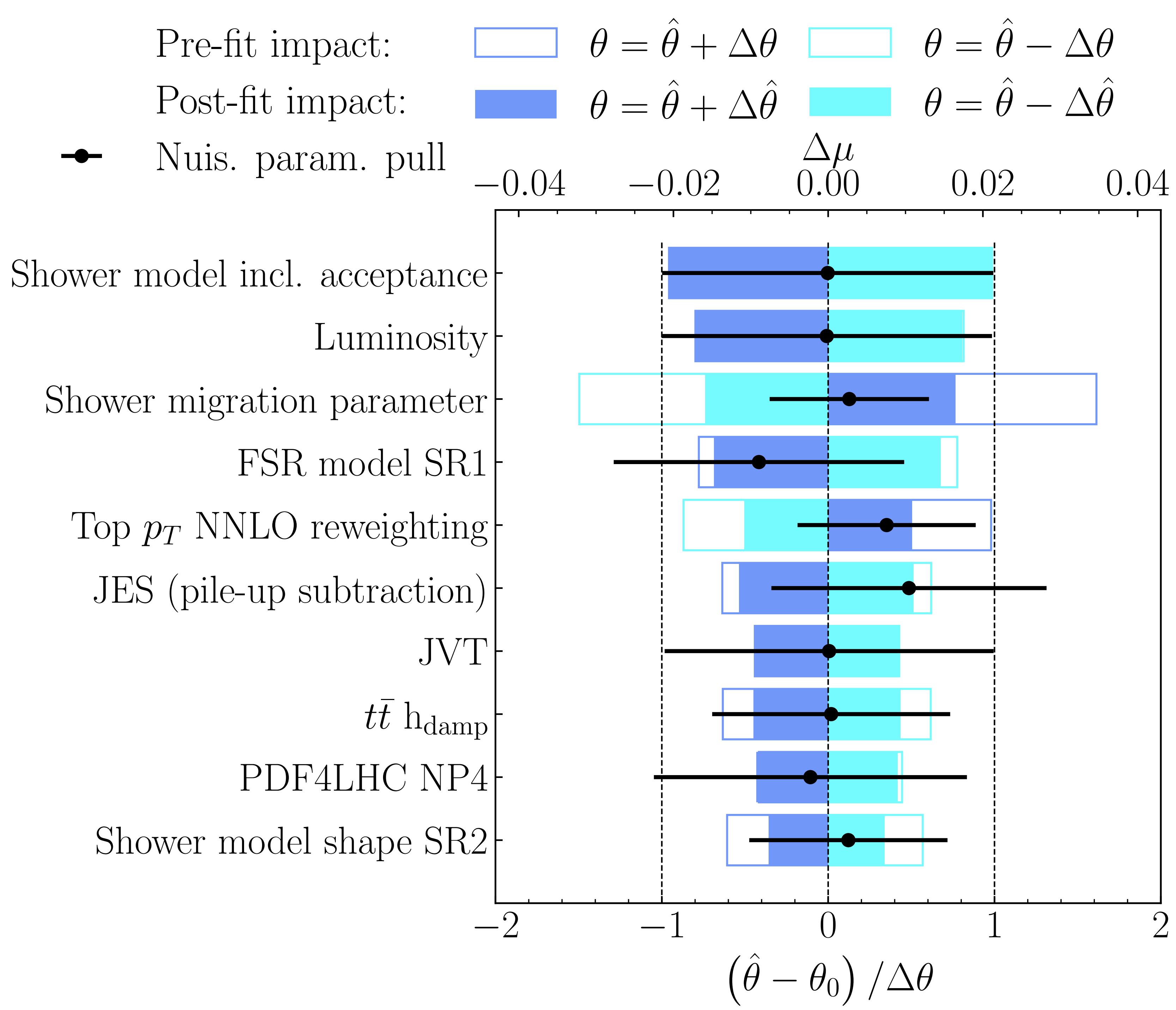
- There is a large number of different nuisance parameters
- Analysed using dedicated python libraries such as **pyhf** and **cabinetry**
  - **Question:** How to make use of this within SFITTER?

Public likelihoods

Reproduction



Original: arXiv:2006.13076



Reproduction: arXiv:2312.12502

cabinetry

pyhf ∂ifferentiable ℒikelihoods

Public likelihoods

Uncertainties

- **Previously:** Uncertainties taken as given in the paper
- **Now:** Uncertainties extracted from profiling in pyhf
  - Implemented into SFITTER using the constraint terms
- **Problem:** Difficult to automate due to inconsistent naming conventions

Uncertainty	Reproduced $\frac{\Delta\sigma_{t\bar{t}Z}}{\sigma_{t\bar{t}Z}}$ [%]	Paper $\frac{\Delta\sigma_{t\bar{t}Z}}{\sigma_{t\bar{t}Z}}$ [%]
ttZ parton shower	3.1	3.1
tWZ modeling	2.9	2.9
b-tagging	2.9	2.9
WZ/ZZ + jets modeling	2.7	2.8
tZq modeling	2.6	2.6
Lepton	2.3	2.3
Luminosity	2.2	2.2
Jets + E_T^{miss}	2.1	2.1
Fake leptons	2.1	2.1
$t\bar{t}Z$ ISR	1.7	1.6
$t\bar{t}Z\mu_F$ and μ_r scales	0.9	0.9
Other backgrounds	0.8	0.7
Pile-up	0.7	0.7
$t\bar{t}Z$ PDF	0.2	0.2
Stat	5.2	5.2

Public likelihoods

Uncertainties

Uncertainty	Reproduced $\frac{\Delta\sigma_{t\bar{t}Z}}{\sigma_{t\bar{t}Z}}$ [%]	Paper $\frac{\Delta\sigma_{t\bar{t}Z}}{\sigma_{t\bar{t}Z}}$ [%]
ttZ parton shower	3.1	3.1
tWZ modeling	2.9	2.9
b-tagging	2.9	2.9
WZ/ZZ + jets modeling	2.7	2.8
tZq modeling	2.6	2.6
Lepton	2.3	2.3
Luminosity	2.2	2.2
Jets + E_T^{miss}	2.1	2.1
Fake leptons	2.1	2.1
$t\bar{t}Z$ ISR	1.7	1.6
$t\bar{t}Z\mu_F$ and μ_r scales	0.9	0.9
Other backgrounds	0.8	0.7
Pile-up	0.7	0.7
$t\bar{t}Z$ PDF	0.2	0.2
Stat	5.2	5.2

Assign →

Systematic uncertainties
Beam
Background (Separate for each channel)
ETmis
Jets
Leptons
LightTagging
Luminosity
Pileup
Trigger
Tune
bTagging
partonShower
tTagging
tauTagging

- **Previously:** Possibly incompatible groups, how to correlate?

Public likelihoods

Uncertainties

Uncertainty	Reproduced $\frac{\Delta\sigma_{t\bar{t}Z}}{\sigma_{t\bar{t}Z}}$ [%]	Paper $\frac{\Delta\sigma_{t\bar{t}Z}}{\sigma_{t\bar{t}Z}}$ [%]
ttZ parton shower	3.1	3.1
tWZ modeling	2.9	2.9
b-tagging	2.9	2.9
WZ/ZZ + jets modeling	2.7	2.8
tZq modeling	2.6	2.6
Lepton	2.3	2.3
Luminosity	2.2	2.2
Jets + E_T^{miss}	2.1	2.1
Fake leptons	2.1	2.1
$t\bar{t}Z$ ISR	1.7	1.6
$t\bar{t}Z\mu_F$ and μ_r scales	0.9	0.9
Other backgrounds	0.8	0.7
Pile-up	0.7	0.7
$t\bar{t}Z$ PDF	0.2	0.2
Stat	5.2	5.2

Assign →

Systematic uncertainties
Beam
Background (Separate for each channel)
ETmis
Jets
Leptons
LightTagging
Luminosity
Pileup
Trigger
Tune
bTagging
partonShower
tTagging
tauTagging

- **Previously:** Possibly incompatible groups, how to correlate?

Public likelihoods

Uncertainties

Uncertainty	Reproduced $\frac{\Delta\sigma_{t\bar{t}Z}}{\sigma_{t\bar{t}Z}}$ [%]	Paper $\frac{\Delta\sigma_{t\bar{t}Z}}{\sigma_{t\bar{t}Z}}$ [%]
ttZ parton shower	3.1	3.1
tWZ modeling	2.9	2.9
b-tagging	2.9	2.9
WZ/ZZ + jets modeling	2.7	2.8
tZq modeling	2.6	2.6
Lepton	2.3	2.3
Luminosity	2.2	2.2
Jets + E_T^{miss}	2.1	2.1
Fake leptons	2.1	2.1
$t\bar{t}Z$ ISR	1.7	1.6
$t\bar{t}Z \mu_F$ and μ_r scales	0.9	0.9
Other backgrounds	0.8	0.7
Pile-up	0.7	0.7
$t\bar{t}Z$ PDF	0.2	0.2
Stat	5.2	5.2

Assign →

Systematic uncertainties
Beam
Background (Separate for each channel)
ETmis
Jets
Leptons
LightTagging
Luminosity
Pileup
Trigger
Tune
bTagging
partonShower
tTagging
tauTagging

- **Previously:** Possibly incompatible groups, how to correlate?

Public likelihoods

Uncertainties

Uncertainty	Reproduced $\frac{\Delta\sigma_{t\bar{t}Z}}{\sigma_{t\bar{t}Z}}$ [%]	Paper $\frac{\Delta\sigma_{t\bar{t}Z}}{\sigma_{t\bar{t}Z}}$ [%]
ttZ parton shower	3.1	3.1
tWZ modeling	2.9	2.9
b-tagging	2.9	2.9
WZ/ZZ + jets modeling	2.7	2.8
tZq modeling	2.6	2.6
Lepton	2.3	2.3
Luminosity	2.2	2.2
Jets + E_T^{miss}	2.1	2.1
Fake leptons	2.1	2.1
$t\bar{t}Z$ ISR	1.7	1.6
$t\bar{t}Z \mu_F$ and μ_r scales	0.9	0.9
Other backgrounds	0.8	0.7
Pile-up	0.7	0.7
$t\bar{t}Z$ PDF	0.2	0.2
Stat	5.2	5.2

Assign →

Systematic uncertainties
Beam
Background (Separate for each channel)
? ETmis
Jets
Leptons
LightTagging
Luminosity
Pileup
Trigger
Tune
bTagging
partonShower
tTagging
tauTagging

- **Previously:** Possibly incompatible groups, how to correlate?

Public likelihoods

Uncertainties

Uncertainty	Reproduced $\frac{\Delta\sigma_{t\bar{t}Z}}{\sigma_{t\bar{t}Z}}$ [%]	Paper $\frac{\Delta\sigma_{t\bar{t}Z}}{\sigma_{t\bar{t}Z}}$ [%]
ttZ parton shower	3.1	3.1
tWZ modeling	2.9	2.9
b-tagging	2.9	2.9
WZ/ZZ + jets modeling	2.7	2.8
tZq modeling	2.6	2.6
Lepton	2.3	2.3
Luminosity	2.2	2.2
Jets + E_T^{miss}	2.1	2.1
Fake leptons	2.1	2.1
$t\bar{t}Z$ ISR	1.7	1.6
$t\bar{t}Z \mu_F$ and μ_r scales	0.9	0.9
Other backgrounds	0.8	0.7
Pile-up	0.7	0.7
$t\bar{t}Z$ PDF	0.2	0.2
Stat	5.2	5.2

Assign →

Systematic uncertainties
Beam
Background (Separate for each channel)
ETmis
Jets
Leptons
LightTagging
Luminosity
Pileup
Trigger
Tune
bTagging
partonShower
tTagging
tauTagging

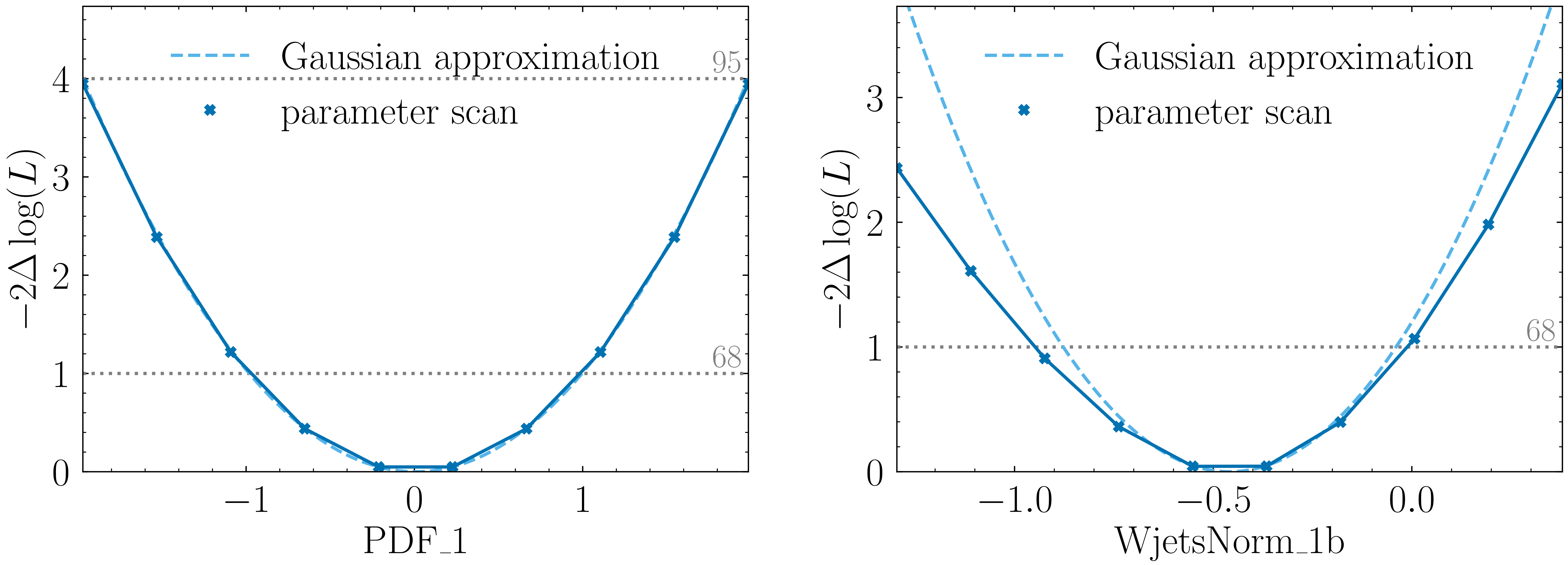
?

- **Now:** Separate the nuisance parameters in profile likelihood fit

Public likelihoods

Parameter scans with cabinetry

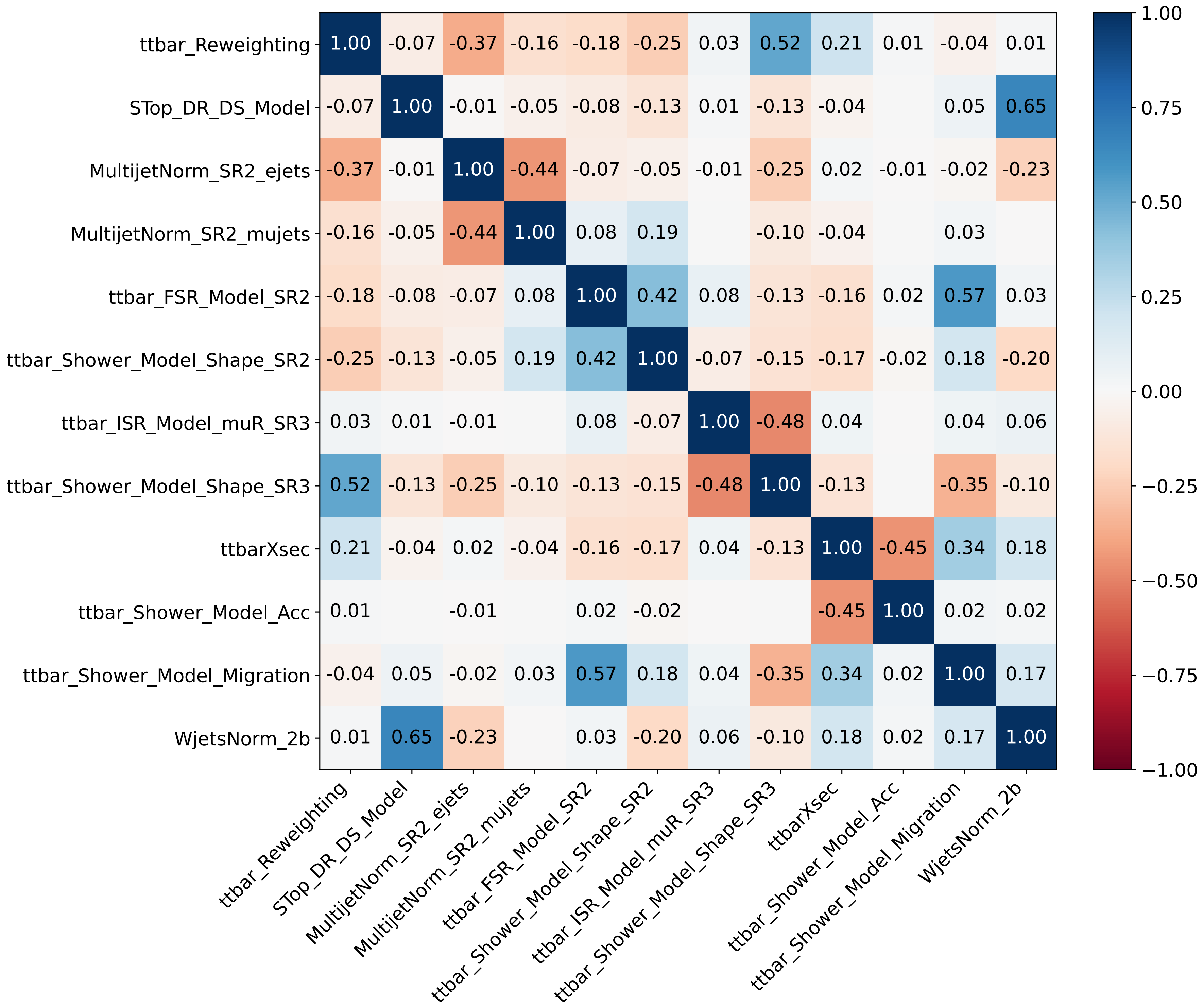
[arXiv:2312.12502]



- NPs all behave very Gaussian, only a small number of exceptions
  - **Validates Gaussian** constraint terms for systematics

Public likelihoods

Concerning correlations

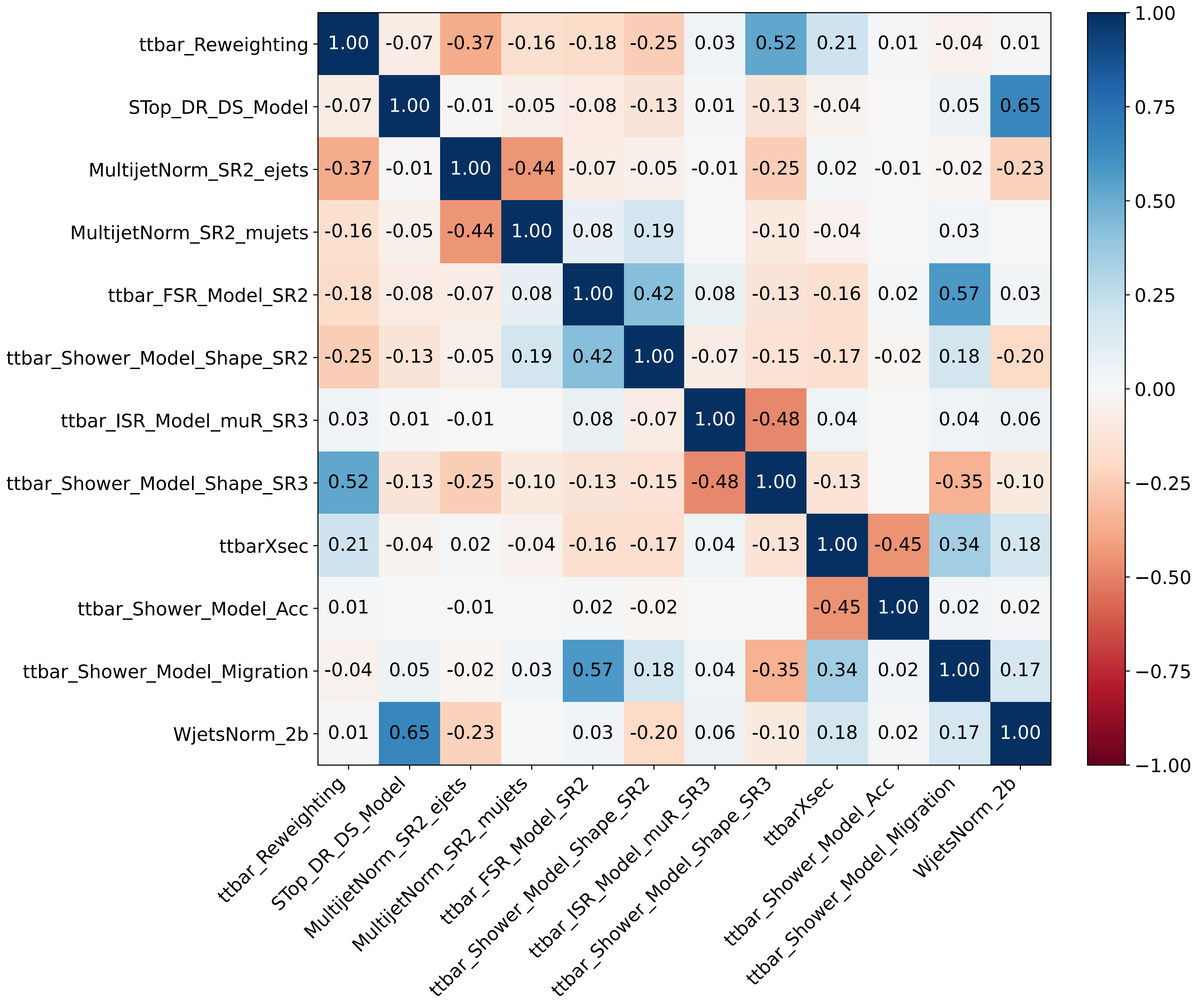


- **Currently:** No correlations between uncertainties within a measurement
  - Correlations of systematics included in SFITTER are negligibly small

Public likelihoods

Concerning correlations

[arXiv:2312.12502]

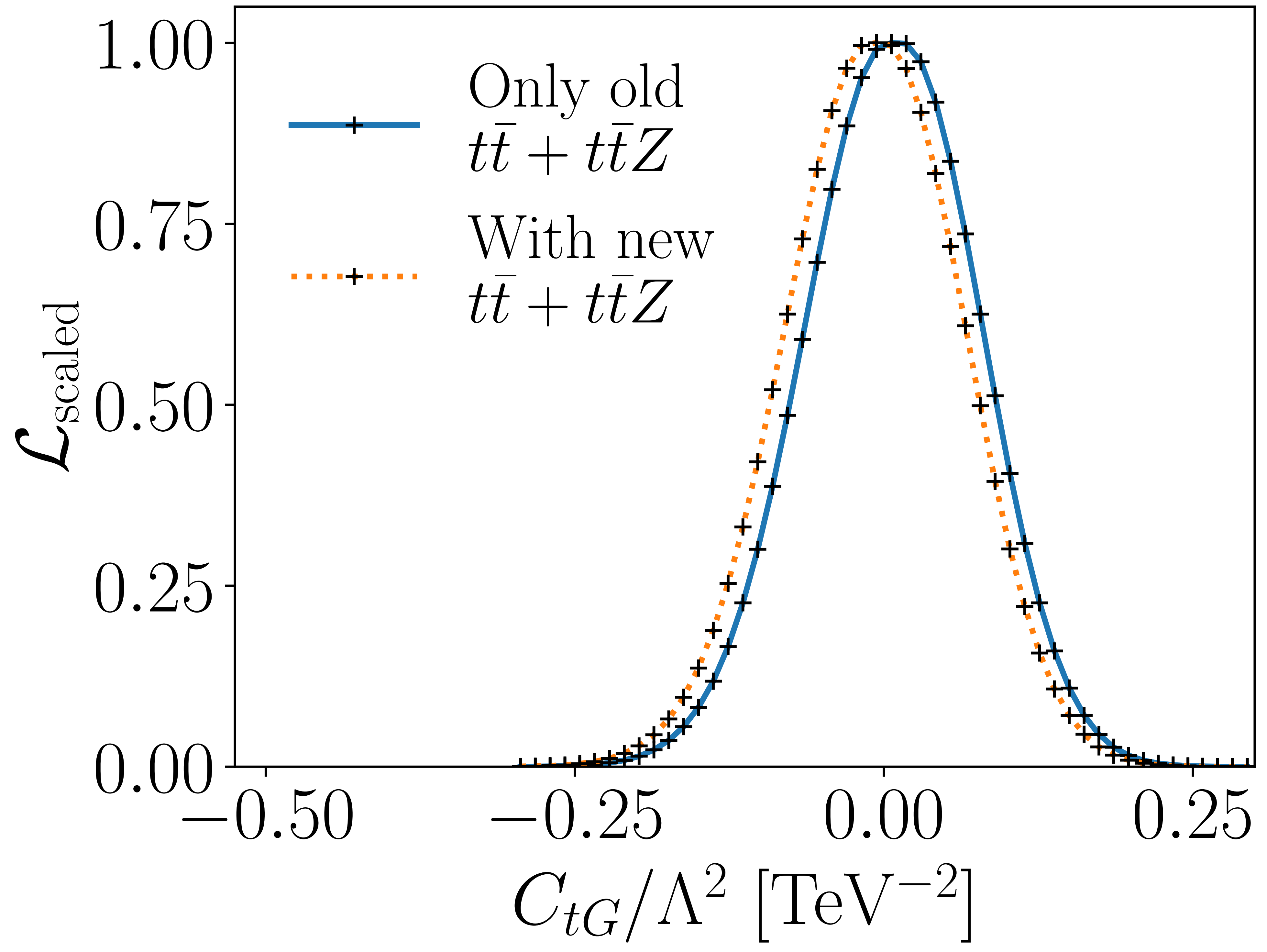


- **Currently:** No correlations between uncertainties within a measurement
  - Correlations of systematics included in SFITTER are negligibly small

Public likelihoods

Effect of new public likelihoods (1D fit)

- Constraints **shift slightly** after including new measurements
- Measurements **barely affect** strength of constraint
  - Only included **total cross sections,** distributions more promising

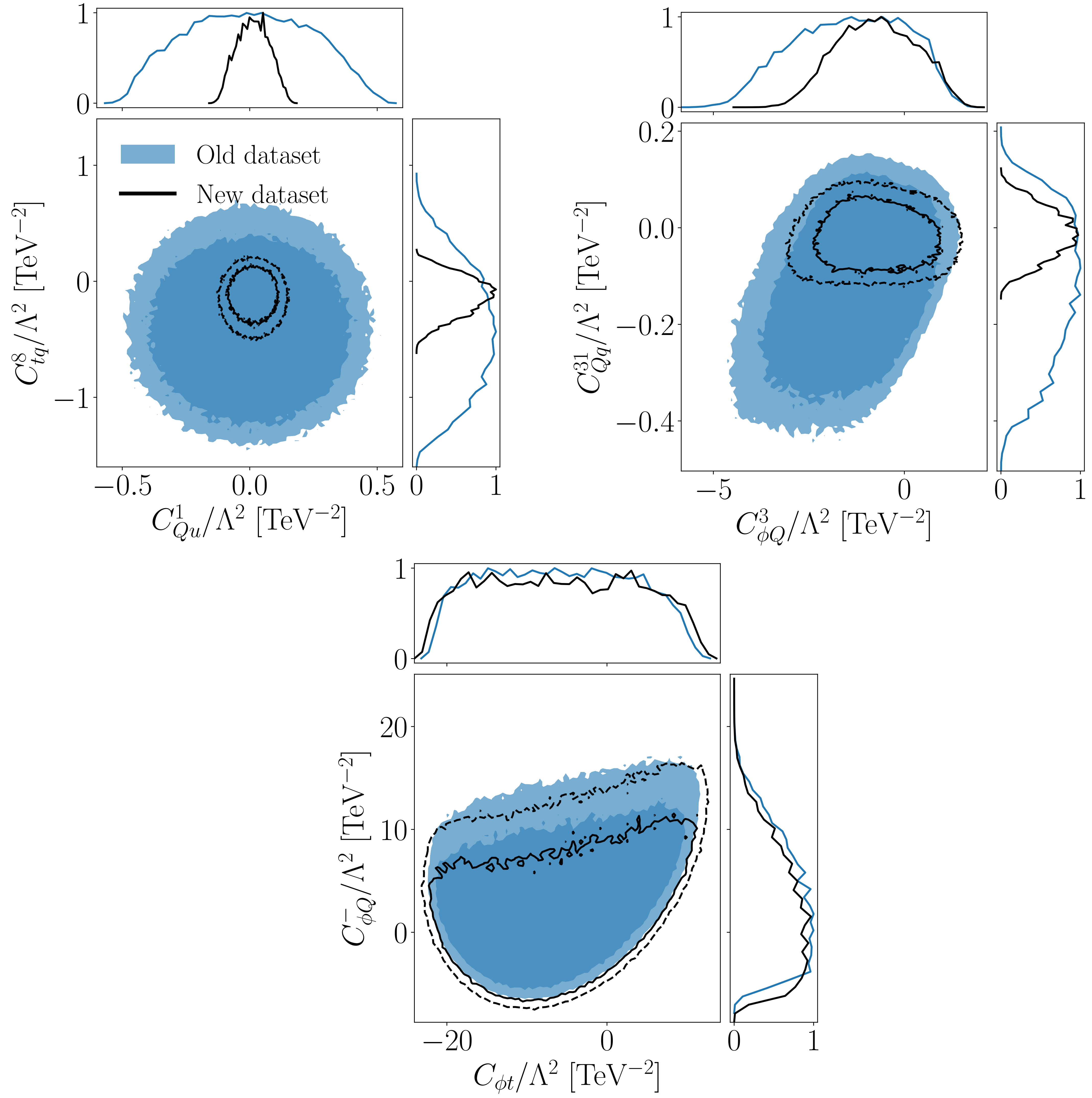


Results

Full global top fit

- All shown operators affected by one of the public likelihoods
- Visibly **stronger constraints,** especially for four fermion operators
- **However:** Strong constraints not due to measurements with likelihoods

[2312.12502]

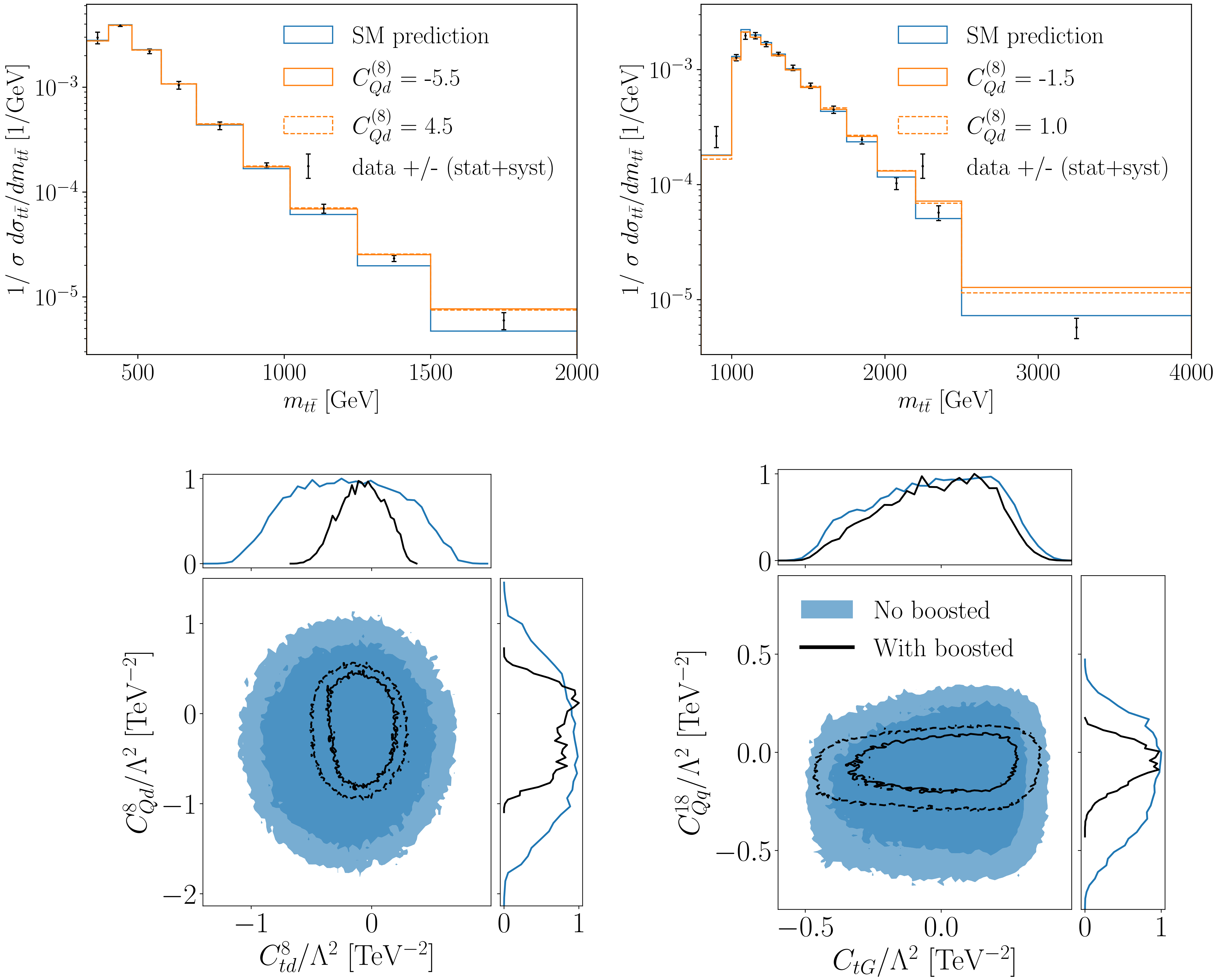


Results

Boosted top measurement

[2312.12502]

- Comparison of leptons+jets vs. all-hadronic (boosted top quarks)
- Both show clear **energy-growing** effects
- Strongest effect on constraints come from **boosted measurements**

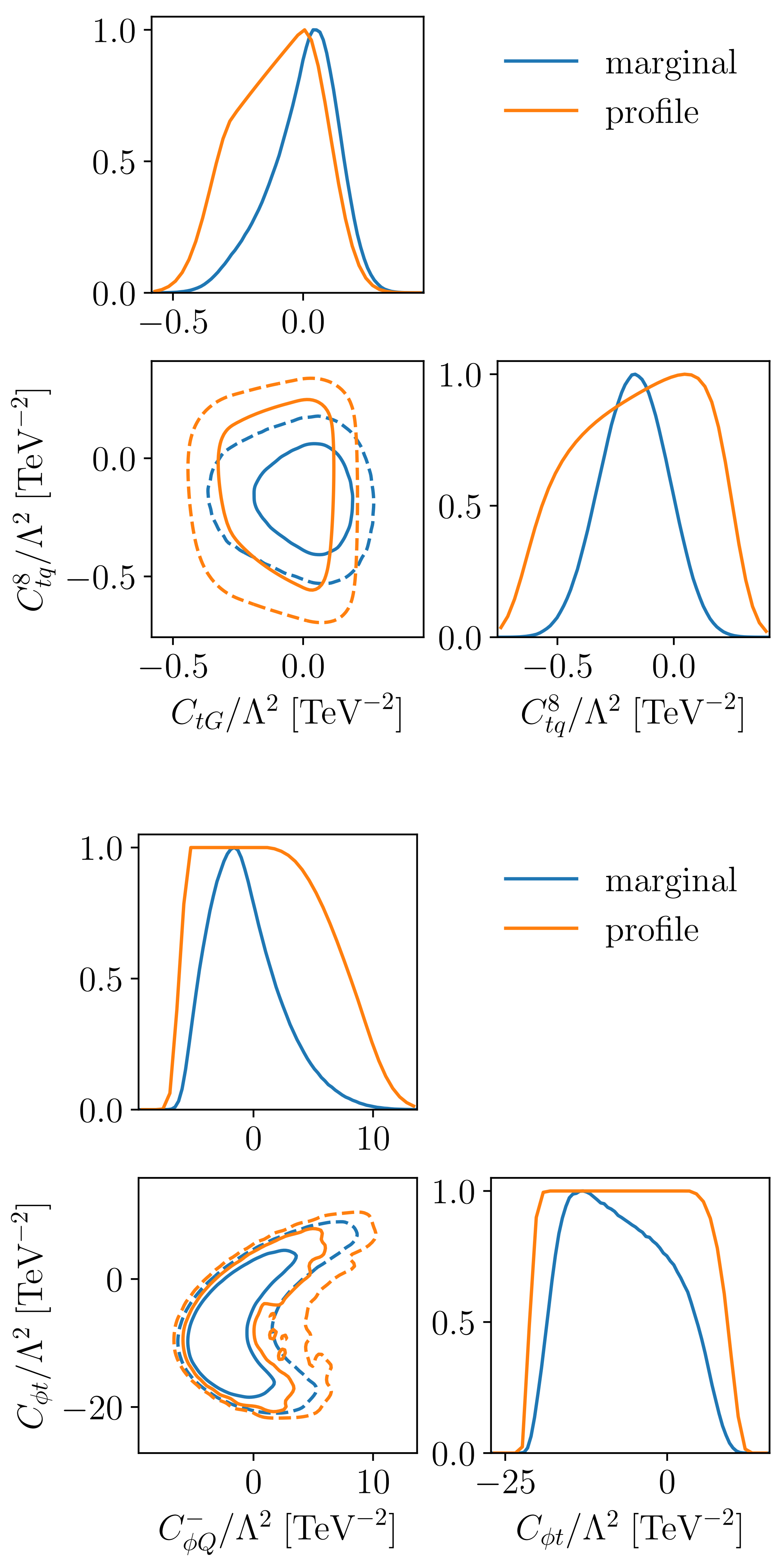


And more?

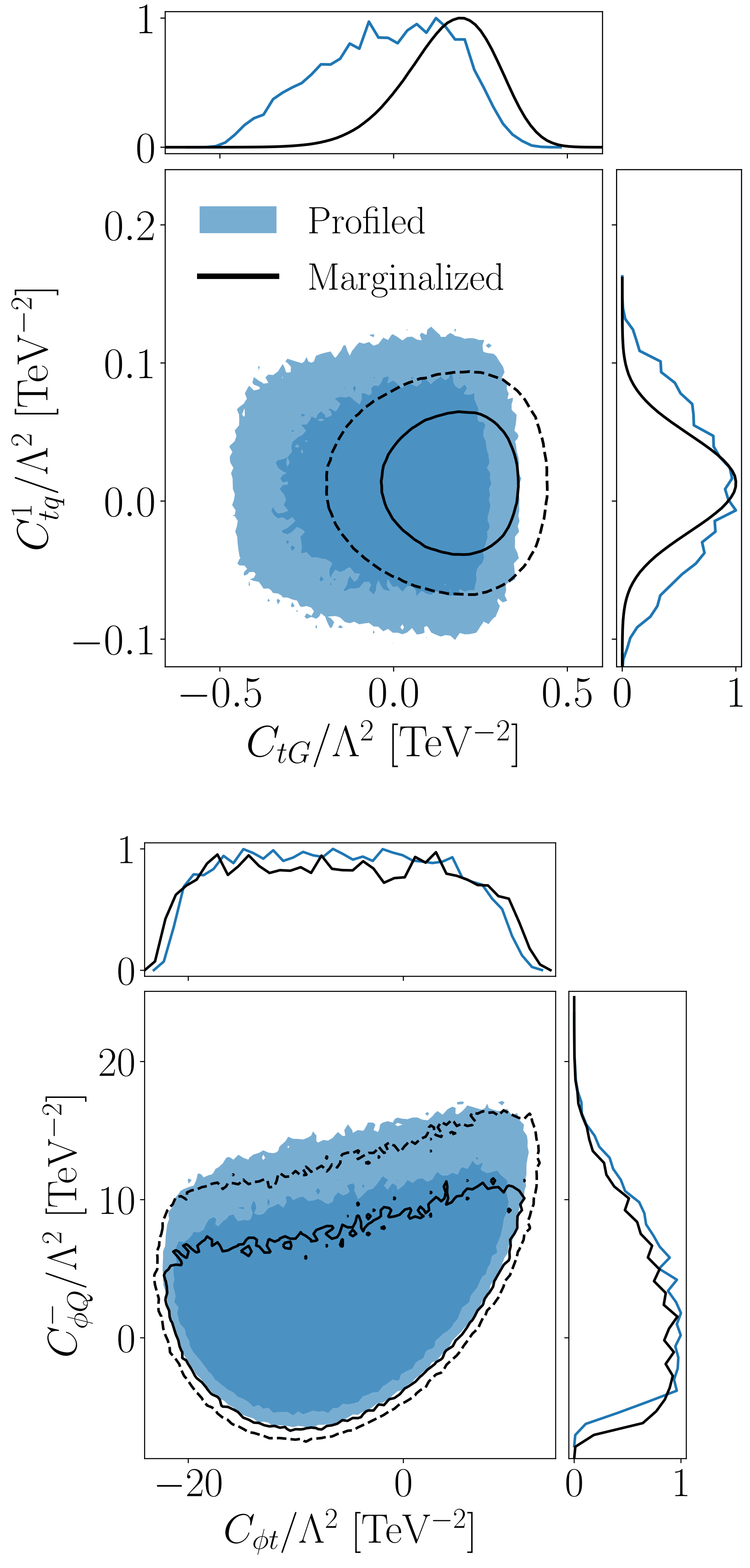
Improving our fit

- Relatively simple top likelihood, mostly unfolded data
- Large effect from **theory uncertainties**
- **Smooth results** for both profiling and marginalization
- Similar improvements for more **complicated Higgs** likelihood

[2411.00942]



[2312.12502]

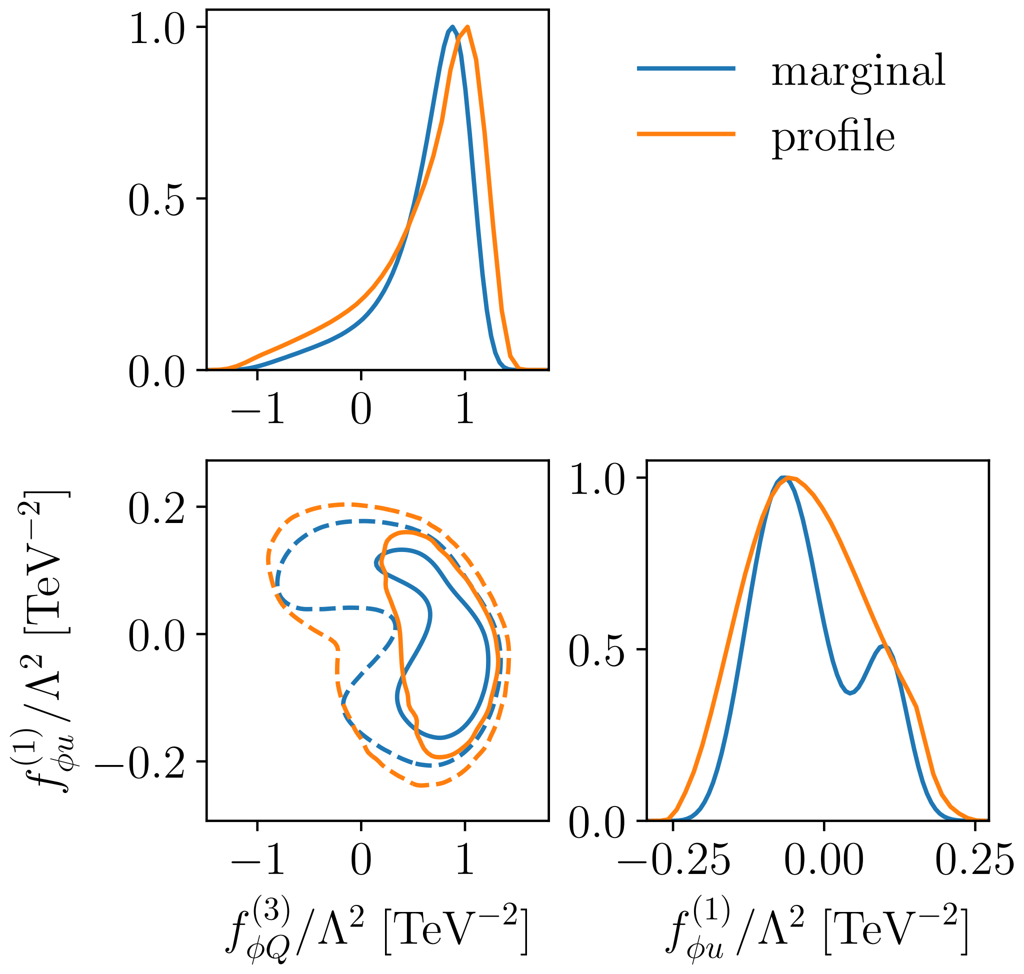


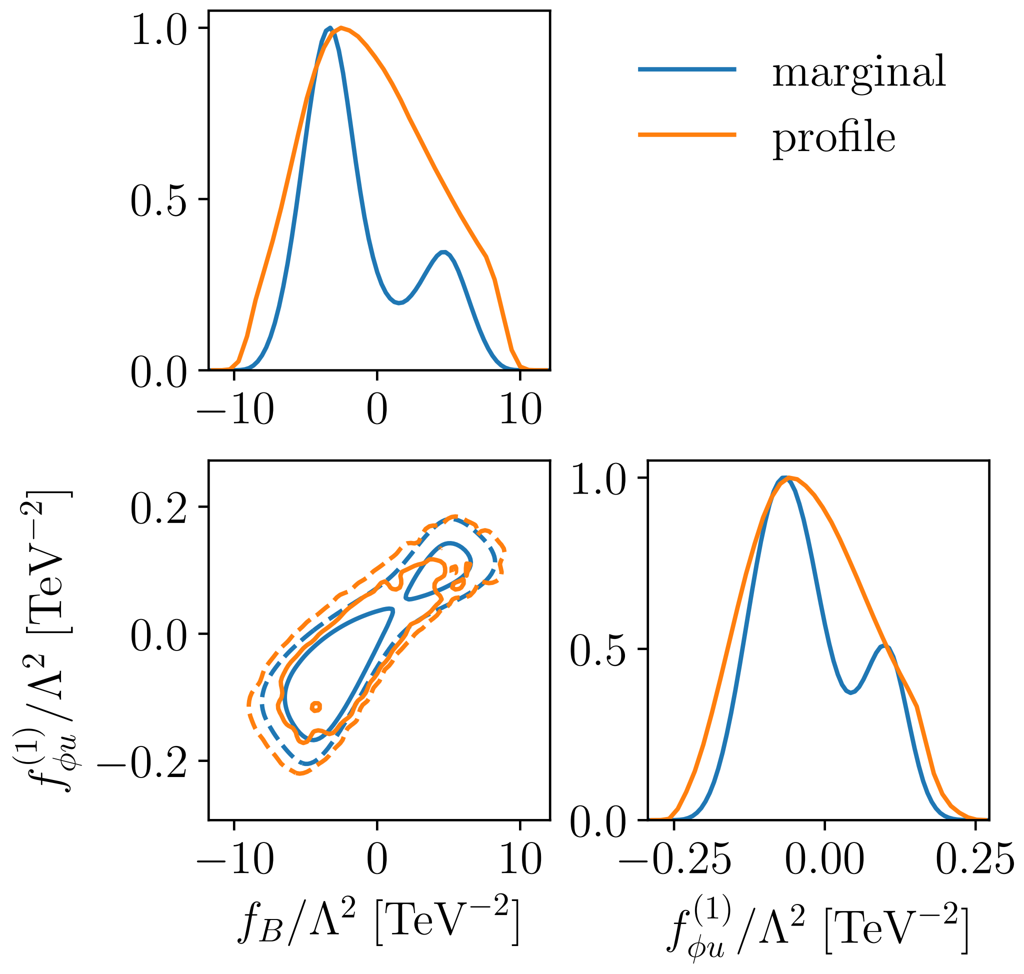
And more?

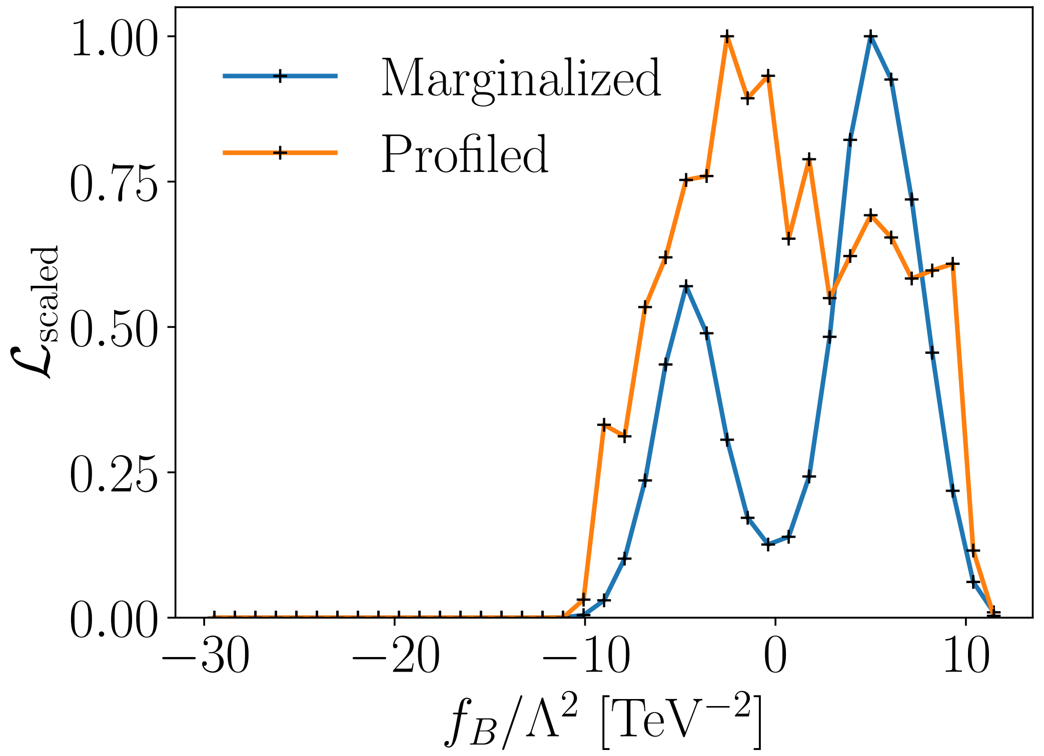
Improving our fit

[2411.00942]

- Relatively simple top likelihood, mostly unfolded data
- Large effect from **theory uncertainties**
- **Smooth results** for both profiling and marginalization
- Similar improvements for more **complicated Higgs** likelihood







Performance

Then vs Now

- For simple likelihoods sampling on both CPU and GPU is fast
- Most time spent on **profiling**
- Training and sampling finished in **around a minute**
- **Furthermore:** More complicated likelihoods sped up significantly
  - You should read [2411.00942]

	Top	Higgs-gauge	Combined
Dimensions	22	20	42
Training batches	100	2000	6000
Samples	10M	200M	100M
Effective sample size	7.1M	97M	21M
Pre-scaling time	7s	3.5min	5.3min
Pre-training time	18s	1.7min	2.5min
Training time	36s	17.3min	1.2h
Sampling time	26s	14.8min	17.6min
Profiling time	17.7min	24.8min	3.7h
Number of CPUs	20	80	120
Accepted samples	37M	26.4M	60M
CPU sampling time	29min 49s	3h 23min	20h 50min
CPU profiling time	4min 43s	8min 24s	N/A

Concluding

- **Summary:** Uncertainties and correlations are essential to SMEFT analyses
  - Published likelihoods provide more **flexibility** when using experimental data
    - **Validates assumptions** made in previous analyses
  - Difficult to make use of this in **global fits** due to large number of measurements without likelihood

- **However:** Currently available likelihoods do not drive SMEFT constraints
  - Publications of more e.g. **kinematic measurements** would be nice
  - For a **global fit**, likelihoods from all kinds of measurements are needed

Appendix/Backup

Top operator definitions

Operator	Definition	Operator	Definition
$\mathcal{O}_{Qq}^{1,8}$	$(\bar{Q}\gamma_\mu T^A Q)\,(\bar{q}_i\gamma^\mu T^A q_i)$	$\mathcal{O}_{tu}^{8}$	$(\bar{t}\gamma_\mu T^A t)\,(\bar{u}_i\gamma^\mu T^A u_i)$
$\mathcal{O}_{Qq}^{1,1}$	$(\bar{Q}\gamma_\mu Q)\,(\bar{q}_i\gamma^\mu q_i)$	$\mathcal{O}_{tu}^{1}$	$(\bar{t}\gamma_\mu t)\,(\bar{u}_i\gamma^\mu u_i)$
$\mathcal{O}_{Qq}^{3,8}$	$(\bar{Q}\gamma_\mu T^A\tau^I Q)\,(\bar{q}_i\gamma^\mu T^A\tau^I q_i)$	$\mathcal{O}_{td}^{8}$	$(\bar{t}\gamma^\mu T^A t)\,(\bar{d}_i\gamma_\mu T^A d_i)$
$\mathcal{O}_{Qq}^{3,1}$	$(\bar{Q}\gamma_\mu \tau^I Q)\,(\bar{q}_i\gamma^\mu \tau^I q_i)$	$\mathcal{O}_{td}^{1}$	$(\bar{t}\gamma^\mu t)\,(\bar{d}_i\gamma_\mu d_i)$
$\mathcal{O}_{Qu}^{8}$	$(\bar{Q}\gamma^\mu T^A Q)\,(\bar{u}_i\gamma_\mu T^A u_i)$	$\mathcal{O}_{Qd}^{1}$	$(\bar{Q}\gamma^\mu Q)\,(\bar{d}_i\gamma_\mu d_i)$
$\mathcal{O}_{Qu}^{1}$	$(\bar{Q}\gamma^\mu Q)\,(\bar{u}_i\gamma_\mu u_i)$	$\mathcal{O}_{tq}^{8}$	$(\bar{q}_i\gamma^\mu T^A q_i)\,(\bar{t}\gamma_\mu T^A t)$
$\mathcal{O}_{Qd}^{8}$	$(\bar{Q}\gamma^\mu T^A Q)\,(\bar{d}_i\gamma_\mu T^A d_i)$	$\mathcal{O}_{tq}^{1}$	$(\bar{q}_i\gamma^\mu q_i)\,(\bar{t}\gamma_\mu t)$
$\mathcal{O}_{\phi Q}^{1}$	$(\phi^\dagger\, i\, \overleftrightarrow{D}_\mu\, \phi)\,(\bar{Q}\gamma^\mu Q)$	$^\ddagger\mathcal{O}_{tB}$	$(\bar{Q}\sigma^{\mu\nu} t)\,\tilde{\phi}\, B_{\mu\nu}$
$\mathcal{O}_{\phi Q}^{3}$	$(\phi^\dagger\, i\, \overleftrightarrow{D}^I_\mu\, \phi)\,(\bar{Q}\gamma^\mu\tau^I Q)$	$^\ddagger\mathcal{O}_{tW}$	$(\bar{Q}\sigma^{\mu\nu} t)\,\tau^I\tilde{\phi}\, W^I_{\mu\nu}$
$\mathcal{O}_{\phi t}$	$(\phi^\dagger\, i\, \overleftrightarrow{D}_\mu\, \phi)\,(\bar{t}\gamma^\mu t)$	$^\ddagger\mathcal{O}_{bW}$	$(\bar{Q}\sigma^{\mu\nu} b)\,\tau^I\phi\, W^I_{\mu\nu}$
$^\ddagger\mathcal{O}_{\phi tb}$	$(\tilde{\phi}^\dagger i D_\mu \phi)\,(\bar{t}\gamma^\mu b)$	$^\ddagger\mathcal{O}_{tG}$	$(\bar{Q}\sigma^{\mu\nu} T^A t)\,\tilde{\phi}\, G^A_{\mu\nu}$

Higgs-gauge operator definitions (HISZ)

Operator	Definition	Operator	Definition
$\mathcal{O}_{GG}$	$\phi^\dagger\phi\ G^a_{\mu\nu}G^{a\mu\nu}$	$\mathcal{O}_{WW}$	$\phi^\dagger\ \hat{W}_{\mu\nu}\hat{W}^{\mu\nu}\ \phi$
$\mathcal{O}_{BB}$	$\phi^\dagger\ \hat{B}_{\mu\nu}\hat{B}^{\mu\nu}\ \phi$	$\mathcal{O}_{W}$	$(D_\mu\phi)^\dagger\hat{W}^{\mu\nu}(D_\nu\phi)$
$\mathcal{O}_{B}$	$(D_\mu\phi)^\dagger\hat{B}^{\mu\nu}(D_\nu\phi)$	$\mathcal{O}_{BW}$	$\phi^\dagger\ \hat{B}_{\mu\nu}\hat{W}^{\mu\nu}\ \phi$
$\mathcal{O}_{\phi 1}$	$(D_\mu\phi)^\dagger\ \phi\phi^\dagger\ (D^\mu\phi)$	$\mathcal{O}_{\phi 2}$	$\frac{1}{2}\partial^\mu(\phi^\dagger\phi)\partial_\mu(\phi^\dagger\phi)$
$\mathcal{O}_{3W}$	$\mathrm{Tr}\left(\hat{W}_{\mu\nu}\hat{W}^{\nu\rho}\hat{W}^{\mu}_{\rho}\right)$		
$\mathcal{O}^{(1)}_{\phi u}$	$\phi^\dagger(i\overleftrightarrow{D}_\mu\phi)(\bar{u}_R\gamma^\mu u_R)$	$\mathcal{O}^{(1)}_{\phi Q}$	$\phi^\dagger(i\overleftrightarrow{D}_\mu\phi)(\bar{Q}\gamma^\mu Q)$
$\mathcal{O}^{(1)}_{\phi d}$	$\phi^\dagger(i\overleftrightarrow{D}_\mu\phi)(\bar{d}_R\gamma^\mu d_R)$	$\mathcal{O}^{(3)}_{\phi Q}$	$\phi^\dagger(i\overleftrightarrow{D}^a{}_\mu\phi)\left(\bar{Q}\gamma^\mu\frac{\sigma_a}{2}Q\right)$
$\mathcal{O}^{(1)}_{\phi e}$	$\phi^\dagger(i\overleftrightarrow{D}_\mu\phi)(\bar{e}_R\gamma^\mu e_R)$		
$\mathcal{O}_{e\phi,22}$	$\phi^\dagger\phi\ \bar{L}_2\phi e_{R,2}$	$\mathcal{O}_{e\phi,33}$	$\phi^\dagger\phi\ \bar{L}_3\phi e_{R,3}$
$\mathcal{O}_{u\phi,33}$	$\phi^\dagger\phi\ \bar{Q}_3\tilde{\phi}u_{R,3}$	$\mathcal{O}_{d\phi,33}$	$\phi^\dagger\phi\ \bar{Q}_3\phi d_{R,3}$
$\mathcal{O}_{4L}$	$(\bar{L}_1\gamma_\mu L_2)\,(\bar{L}_2\gamma^\mu L_1)$		

Higgs-gauge operator definitions (Warsaw)

Operator	Definition	Operator	Definition
$\mathcal{O}_{\phi G}$	$\phi^\dagger\phi G^A_{\mu\nu}G^{A\mu\nu}$	$\mathcal{O}_W$	$\varepsilon^{IJK}W^{I\nu}_\mu W^{J\rho}_\nu W^{K\mu}_\rho$
$\mathcal{O}_{\phi B}$	$\phi^\dagger\phi B_{\mu\nu}B^{\mu\nu}$	$\mathcal{O}_{\phi W}$	$\phi^\dagger\phi W^I_{\mu\nu}W^{I\mu\nu}$
$\mathcal{O}_{\phi WB}$	$\phi^\dagger\tau^I\phi W^I_{\mu\nu}B^{\mu\nu}$		
$\mathcal{O}_{\phi\Box}$	$(\phi^\dagger\phi)\Box(\phi^\dagger\phi)$	$\mathcal{O}_{\phi D}$	$(\phi^\dagger D^\mu\phi)^*(\phi^\dagger D^\mu\phi)$
$\mathcal{O}_{\phi e}$	$(\phi^\dagger i\overleftrightarrow{D}_\mu\phi)(\bar{e}_i\gamma^\mu e_i)$	$\mathcal{O}_{\phi b}$	$(\phi^\dagger i\overleftrightarrow{D}_\mu\phi)(\bar{b}_i\tau^I\gamma^\mu b_i)$
$\mathcal{O}_{\phi d}$	$\sum_{i=1}^2(\phi^\dagger i\overleftrightarrow{D}_\mu\phi)(\bar{d}_i\gamma^\mu d_i)$	$\mathcal{O}_{\phi u}$	$\sum_{i=1}^2(\phi^\dagger i\overleftrightarrow{D}_\mu\phi)(\bar{u}_i\gamma^\mu u_i)$
$\mathcal{O}^{(1)}_{\phi q}$	$\sum_{i=1}^2(\phi^\dagger i\overleftrightarrow{D}_\mu\phi)(\bar{q}_i\gamma^\mu q_i)$	$\mathcal{O}^{(3)}_{\phi q}$	$\sum_{i=1}^2(\phi^\dagger i\overleftrightarrow{D}_\mu\phi)(\bar{q}_i\tau^I\gamma^\mu q_i)$
$\mathcal{O}^{(1)}_{\phi l}$	$(\phi^\dagger i\overleftrightarrow{D}_\mu\phi)(\bar{l}\gamma^\mu l)$	$\mathcal{O}^{(3)}_{\phi l}$	$(\phi^\dagger i\overleftrightarrow{D}^I_\mu\phi)(\bar{l}\tau^I\gamma^\mu l)$
$\mathcal{O}_{d\phi,33}$	$(\phi^\dagger\phi)(\bar{Q}_3 b\phi)$	$\mathcal{O}_{u\phi,33}$	$(\phi^\dagger\phi)(\bar{Q}_3 t\phi)$
$\mathcal{O}_{e\phi,22}$	$(\phi^\dagger\phi)(\bar{l}_2\mu\phi)$	$\mathcal{O}_{e\phi,33}$	$(\phi^\dagger\phi)(\bar{l}_3\tau\phi)$
$\mathcal{O}_{ll}$	$(\bar{l}\gamma_\mu l)(\bar{l}\gamma^\mu l)$		

Full top dataset

Experiment	Energy [TeV]	$\mathcal{L}$ [fb^{-1}]	Channel	Observable	# Bins	New	Likelihood	QCD k-factor
CMS [79]	8	19.7	$e\mu$	$\sigma_{t\bar{t}}$				[80]
ATLAS [81]	8	20.2	lj	$\sigma_{t\bar{t}}$				[80]
CMS [82]	13	137	lj	$\sigma_{t\bar{t}}$		✓		[80]
CMS [83]	13	35.9	ll	$\sigma_{t\bar{t}}$				[80]
ATLAS [84]	13	36.1	ll	$\sigma_{t\bar{t}}$		✓		[80]
ATLAS [85]	13	36.1	aj	$\sigma_{t\bar{t}}$		✓		[80]
ATLAS [47]	13	139	lj	$\sigma_{t\bar{t}}$		✓	✓	[80]
CMS [86]	13.6	1.21	ll, lj	$\sigma_{t\bar{t}}$		✓		[86]
CMS [87]	8	19.7	lj	$\frac{1}{\sigma}\frac{d\sigma}{dp_T^t}$	7			[88–90]
CMS [87]	8	19.7	ll	$\frac{1}{\sigma}\frac{d\sigma}{dp_T^t}$	5			[88–90]
ATLAS [91]	8	20.3	lj	$\frac{1}{\sigma}\frac{d\sigma}{dm_{t\bar{t}}}$	7			[88–90]
CMS [82]	13	137	lj	$\frac{1}{\sigma}\frac{d\sigma}{dm_{t\bar{t}}}$	15	✓		[45]
CMS [92]	13	35.9	ll	$\frac{1}{\sigma}\frac{d\sigma}{d\Delta y_{t\bar{t}}}$	8			[88–90]
ATLAS [93]	13	36	lj	$\frac{1}{\sigma}\frac{d\sigma}{dm_{t\bar{t}}}$	9	✓		[45]
ATLAS [94]	13	139	aj, high-p_T	$\frac{1}{\sigma}\frac{d\sigma}{dm_{t\bar{t}}}$	13	✓		
CMS [95]	8	19.7	lj	A_C				[96]
CMS [97]	8	19.5	ll	A_C				[96]
ATLAS [98]	8	20.3	lj	A_C				[96]
ATLAS [99]	8	20.3	ll	A_C				[96]
CMS [100]	13	138	lj	A_C		✓		[96]
ATLAS [101]	13	139	lj	A_C		✓		[96]
ATLAS [48]	13	139		$\sigma_{t\bar{t}Z}$		✓	✓	[102]
CMS [103]	13	77.5		$\sigma_{t\bar{t}Z}$				[102]
CMS [104]	13	35.9		$\sigma_{t\bar{t}W}$				[102]
ATLAS [105]	13	36.1		$\sigma_{t\bar{t}W}$		✓		[102]
CMS [106]	8	19.7		$\sigma_{t\bar{t}\gamma}$		✓		
ATLAS [107]	8	20.2		$\sigma_{t\bar{t}\gamma}$		✓		

Exp.	$\sqrt{s}$ [TeV]	$\mathcal{L}$ [fb^{-1}]	Channel	Observable	# Bins	New	Likelihood	QCD k-factor
ATLAS [108]	7	4.59	t-ch	$\sigma_{tq+\bar{t}q}$				
CMS [109]	7	1.17 (e), 1.56 (μ)	t-ch	$\sigma_{tq+\bar{t}q}$				
ATLAS [110]	8	20.2	t-ch	σ_{tq}, $\sigma_{\bar{t}q}$				
CMS [111]	8	19.7	t-ch	σ_{tq}, $\sigma_{\bar{t}q}$				
ATLAS [112]	13	3.2	t-ch	σ_{tq}, $\sigma_{\bar{t}q}$				[113]
CMS [114]	13	2.2	t-ch	σ_{tq}, $\sigma_{\bar{t}q}$				[113]
CMS [115]	13	35.9	t-ch	$\frac{1}{\sigma}\frac{d\sigma}{d\lvert p_{T,t}\rvert}$	5	✓		
CMS [116]	7	5.1	s-ch	$\sigma_{t\bar{b}+\bar{t}b}$				
CMS [116]	8	19.7	s-ch	$\sigma_{t\bar{b}+\bar{t}b}$				
ATLAS [117]	8	20.3	s-ch	$\sigma_{t\bar{b}+\bar{t}b}$				
ATLAS [49]	13	139	s-ch	$\sigma_{t\bar{b}+\bar{t}b}$		✓	✓	
ATLAS [118]	7	2.05	tW ($2l$)	$\sigma_{tW+\bar{t}W}$				
CMS [119]	7	4.9	tW ($2l$)	$\sigma_{tW+\bar{t}W}$				
ATLAS [120]	8	20.3	tW ($2l$)	$\sigma_{tW+\bar{t}W}$				
ATLAS [121]	8	20.2	tW ($1l$)	$\sigma_{tW+\bar{t}W}$		✓		
CMS [122]	8	12.2	tW ($2l$)	$\sigma_{tW+\bar{t}W}$				
ATLAS [123]	13	3.2	tW ($1l$)	$\sigma_{tW+\bar{t}W}$				
CMS [124]	13	35.9	tW ($e\mu j$)	$\sigma_{tW+\bar{t}W}$				
CMS [125]	13	36	tW ($2l$)	$\sigma_{tW+\bar{t}W}$		✓		
ATLAS [126]	13	36.1	tZ	σ_{tZq}				
ATLAS [127]	7	1.04		F_0,F_L				
CMS [128]	7	5		F_0,F_L				
ATLAS [129]	8	20.2		F_0,F_L				
CMS [130]	8	19.8		F_0,F_L				
ATLAS [131]	13	139		F_0,F_L		✓		

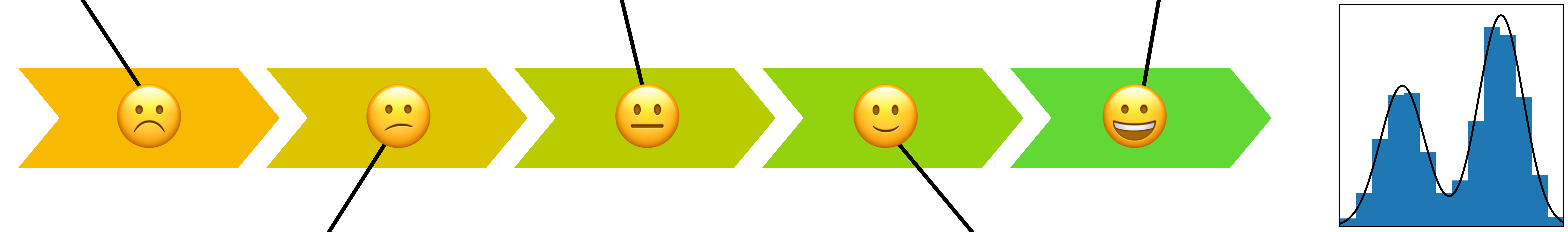
The five steps to happiness

Pre-scaling
- run many parallel MCs to determine mean std.
- normalize distribution

Training
- similar to MadNIS
- online + buffered training
- refine samples using small number of MCMC steps

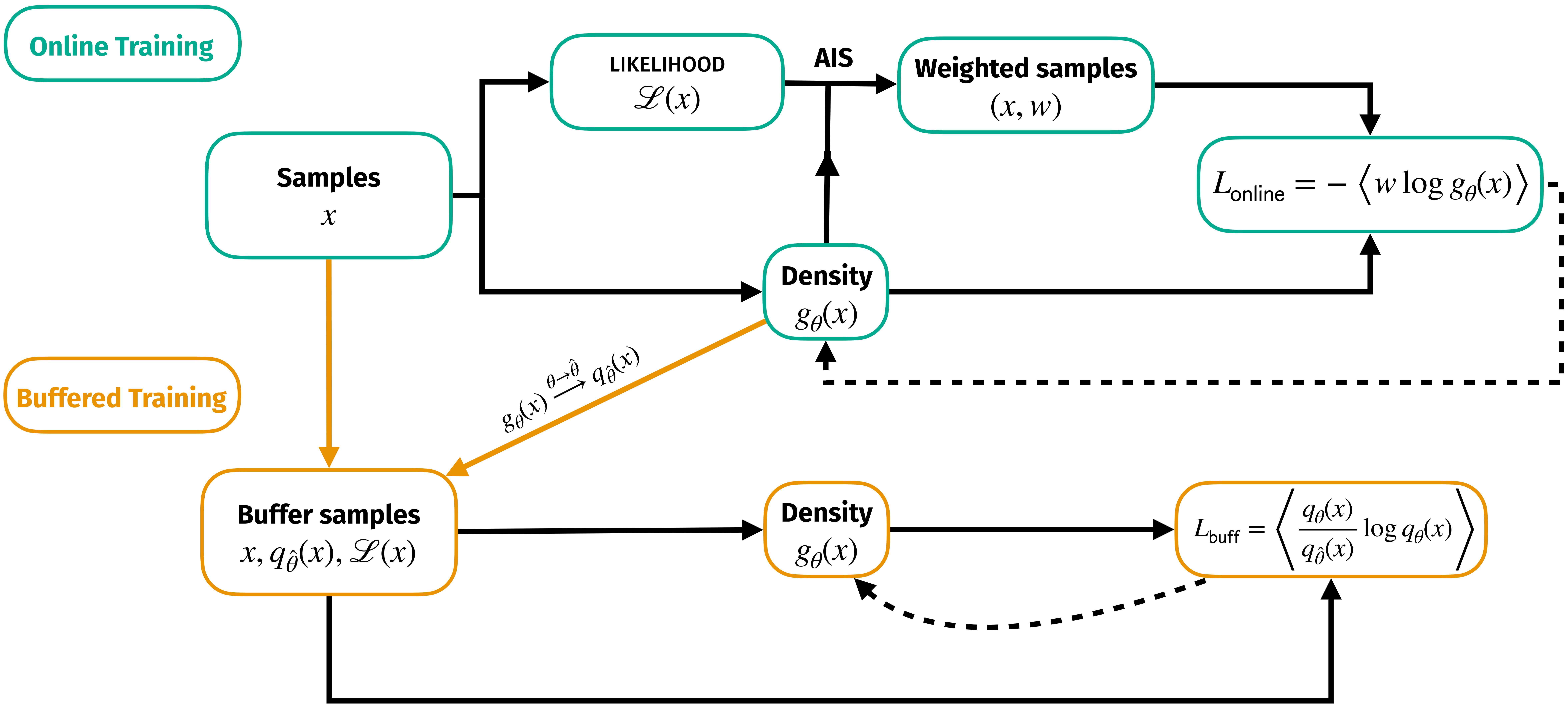
Profiling
- run maximization algorithm for each bin (L-BFGS)
- use gradient information

$p(x)$



Pre-training
- use samples from pre-scaling to train network for a few steps
- better starting point for main training

Sampling
- generate weighted samples
- keep track of points with highest likelihood in each bin


Online Training
LIKELIHOOD
$\mathscr{L}(x)$
AIS
Weighted samples
(x, w)
Samples
x
$L_{\text{online}} = -\langle w \log g_\theta(x) \rangle$
Density
$g_\theta(x)$
$g_\theta(x) \xrightarrow{\theta \to \hat{\theta}} q_{\hat{\theta}}(x)$
Buffered Training
Buffer samples
$x, q_{\hat{\theta}}(x), \mathscr{L}(x)$
Density
$g_\theta(x)$
$L_{\text{buff}} = \left\langle \frac{q_\theta(x)}{q_{\hat{\theta}}(x)} \log q_\theta(x) \right\rangle$

Hyperparameters

		Top	Higgs-gauge	Combined
Architecture	Coupling blocks	RQ splines		
	Spline bins	16		
	Subnet layers	3		
	Hidden layers	64		
Pre-scaling	Number of samples	10240	40960	40960
	AIS steps	1500	5500	5500
	Target acceptance	0.33		
Pre-training	Batch size	1024		
	Epochs	15	6	6
	MCMC steps between batches	20	10	10
Training	Learning rate	0.001		
	Batch size	1024		
	Batches	100	2000	6000
	AIS steps	4	4	8
	Buffer capacity	262k		
	Ratio buffered/online steps	6		
Sampling	Batches	100	2000	1000
	Batch size	100k		
	Marginalization bins, 1D	80		
	Marginalization bins, 2D	40		
	Profiling bins, 1D	40		
	Profiling bins, 2D	30	30	20
Profiling	Batch size	100k		
	Optimizer	LBFGS		
	Optimization steps	200		