

# Profile Likelihoods on ML-Steroids

Accelerating global SMEFT fits using neural importance sampling

**Nikita Schmal**

27/02/2025, LHC Reinterpretation forum

Based on

[[2411.00942](#)] Theo Heimes, Tilman Plehn, [Nikita Schmal](#)



**UNIVERSITÄT  
HEIDELBERG**  
ZUKUNFT  
SEIT 1386

# A brief (SMEFT) 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 marginalisation** methods
- Mapping of likelihood using MCMC
- **Goal:** Do this, but faster and better



# SMEFT

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

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

- operators up to **dimension 6**, contributions to quadratic order
- Make SMEFT predictions **differentiable** and fast on a **GPU**

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

- ▶ Expressed as **simple bilinear operation**

# 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]

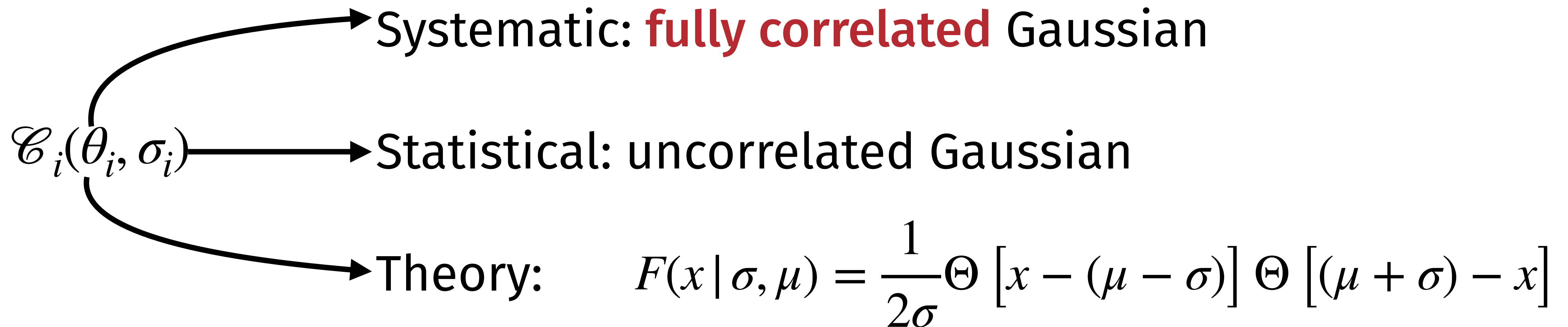
and **Higgs** sector with **20 Wilson coefficients**

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

# SFITTER likelihood

- Construct the likelihood for a single measurement

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



- Public likelihoods help handle these uncertainties properly

# SFITTER likelihood

- Remove nuisance parameters via **profiling**  $L_{\text{prof}}(x) = \max_{\theta, b} \mathcal{L}_{\text{excl}}$

## Gaussian contribution

$$\log L_{\text{Gauss}}(\tilde{s} | d, b_{CR}) = \frac{(d - b_{CR} - \tilde{s}_{\sigma})^2}{\sum_{\text{syst}} (\sigma_{d,i} - \sigma_{b,i})^2}$$

## Poisson contributions

$$\log L_{\text{pois},d}(\tilde{s} | d, b_{CR}) = d - (\tilde{s}_{\sigma} + b_{CR}) \log(\tilde{s}_{\sigma} + b_{CR}) + \log \frac{(\tilde{s}_{\sigma} + b_{CR})!}{d!}$$

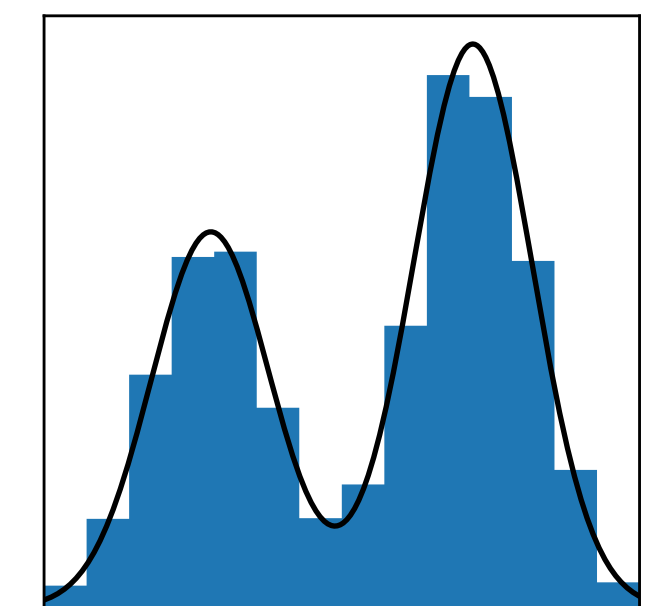
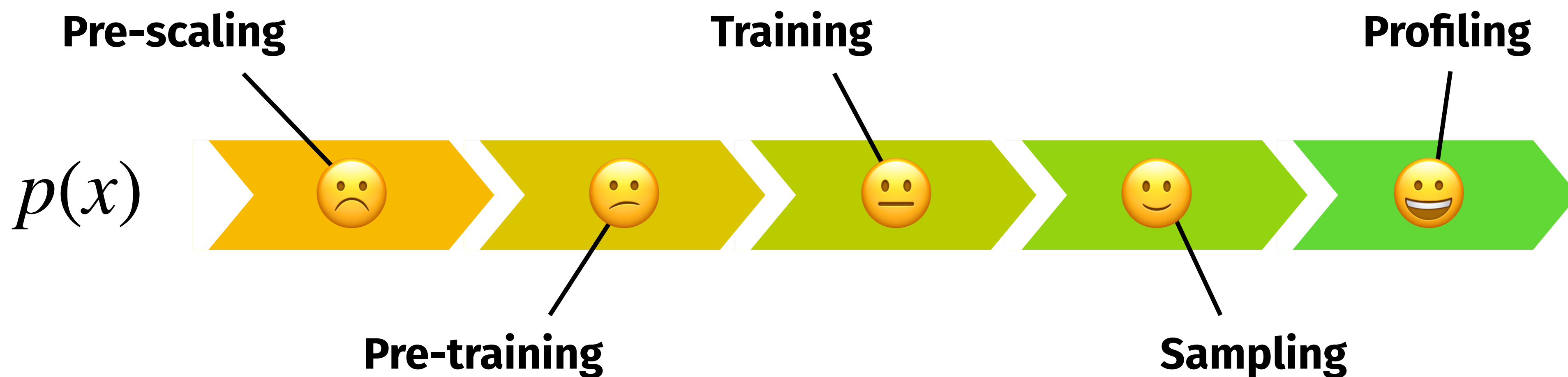
$$\log L_{\text{pois},b}(\tilde{s} | d, b_{CR}) = b_{CR} - (d - \tilde{s}_{\sigma}) \log(d - \tilde{s}_{\sigma}) + \log \frac{(d - \tilde{s}_{\sigma})!}{b_{CR}!}$$

## Full likelihood

$$\frac{1}{L_{\text{full}}} \approx \frac{1}{L_{\text{Gauss}}} + \frac{1}{L_{\text{Poiss},b}} + \frac{1}{L_{\text{Poiss},d}}$$

# The five steps to happiness

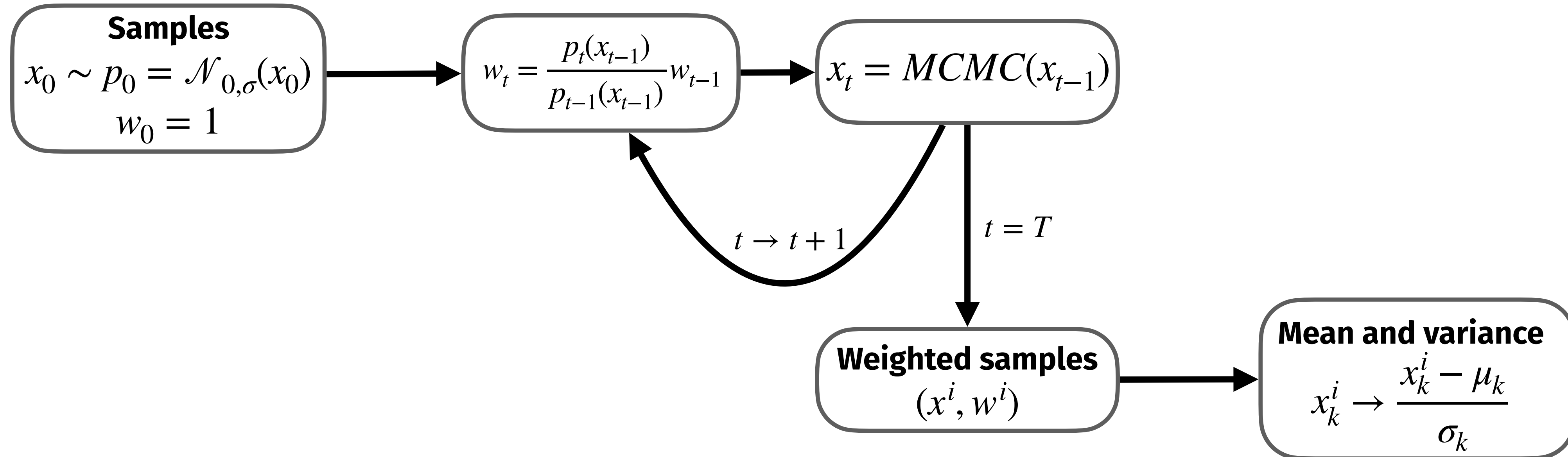
- We finally extract constraints on the WCs from this likelihood
- SFITTER makes use of MCMC, but that is too **slow**
- **Alternative:** Train a normalizing flow to speed this up
  - Many more details in arXiv: [[2411.00942](https://arxiv.org/abs/2411.00942)]





- Scale the parameter space to be centered around 0 with unit std.
- Use annealed importance sampling (AIS) [Neal, 2001]

$$\log p_t(x) = (1 - \beta_t)\log p_0(x) + \beta_t \log p_T(x) \quad \text{with} \quad \beta_t = \frac{t}{T} \quad \text{and} \quad t = 1, \dots, T$$

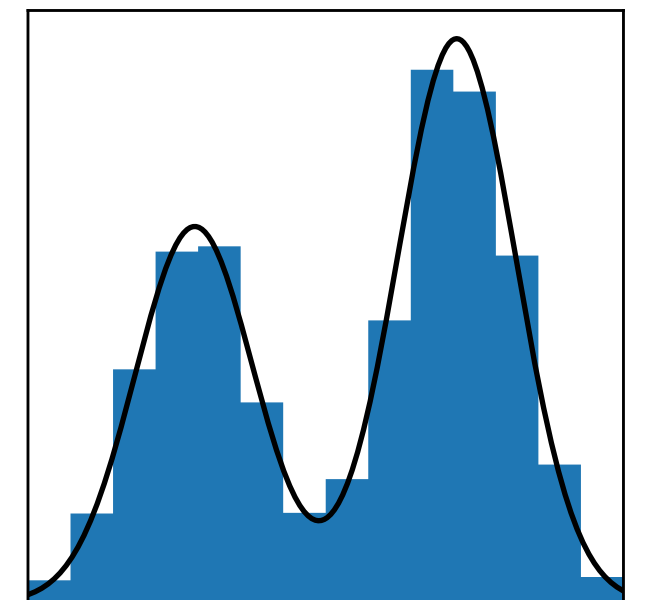


# The five steps to happiness

## Pre-scaling

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

$p(x)$



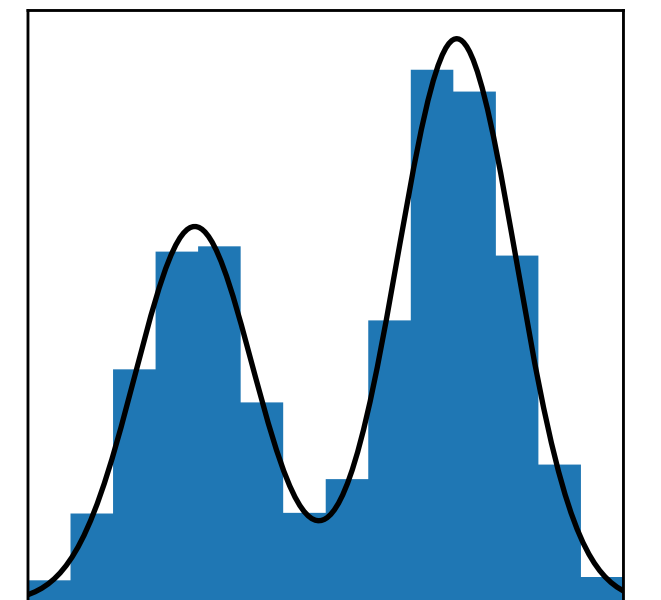


# The five steps to happiness

## Pre-scaling

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

$p(x)$



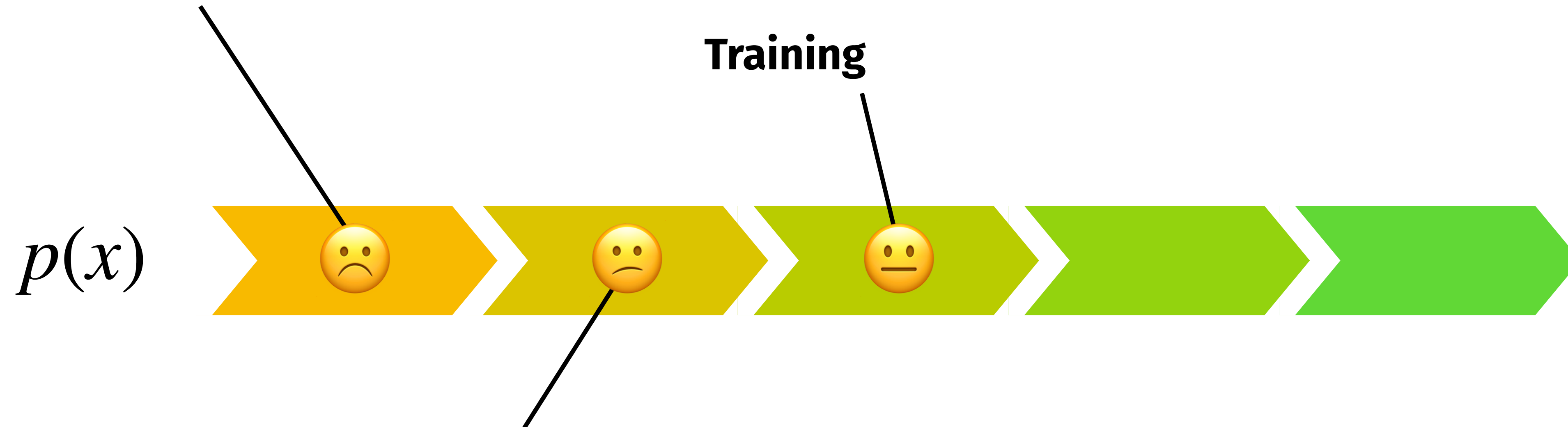
## Pre-training

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

# The five steps to happiness

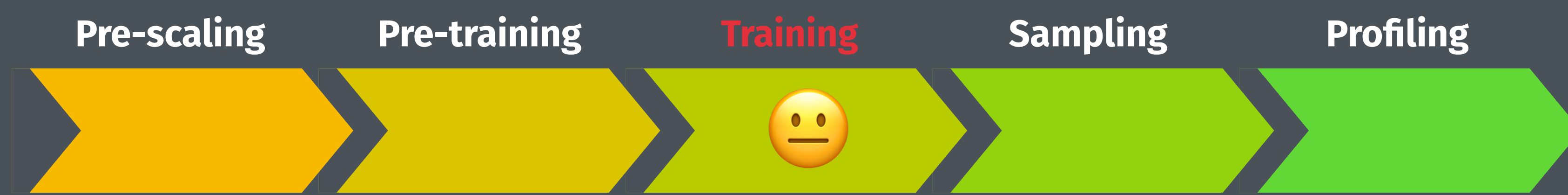
## Pre-scaling

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

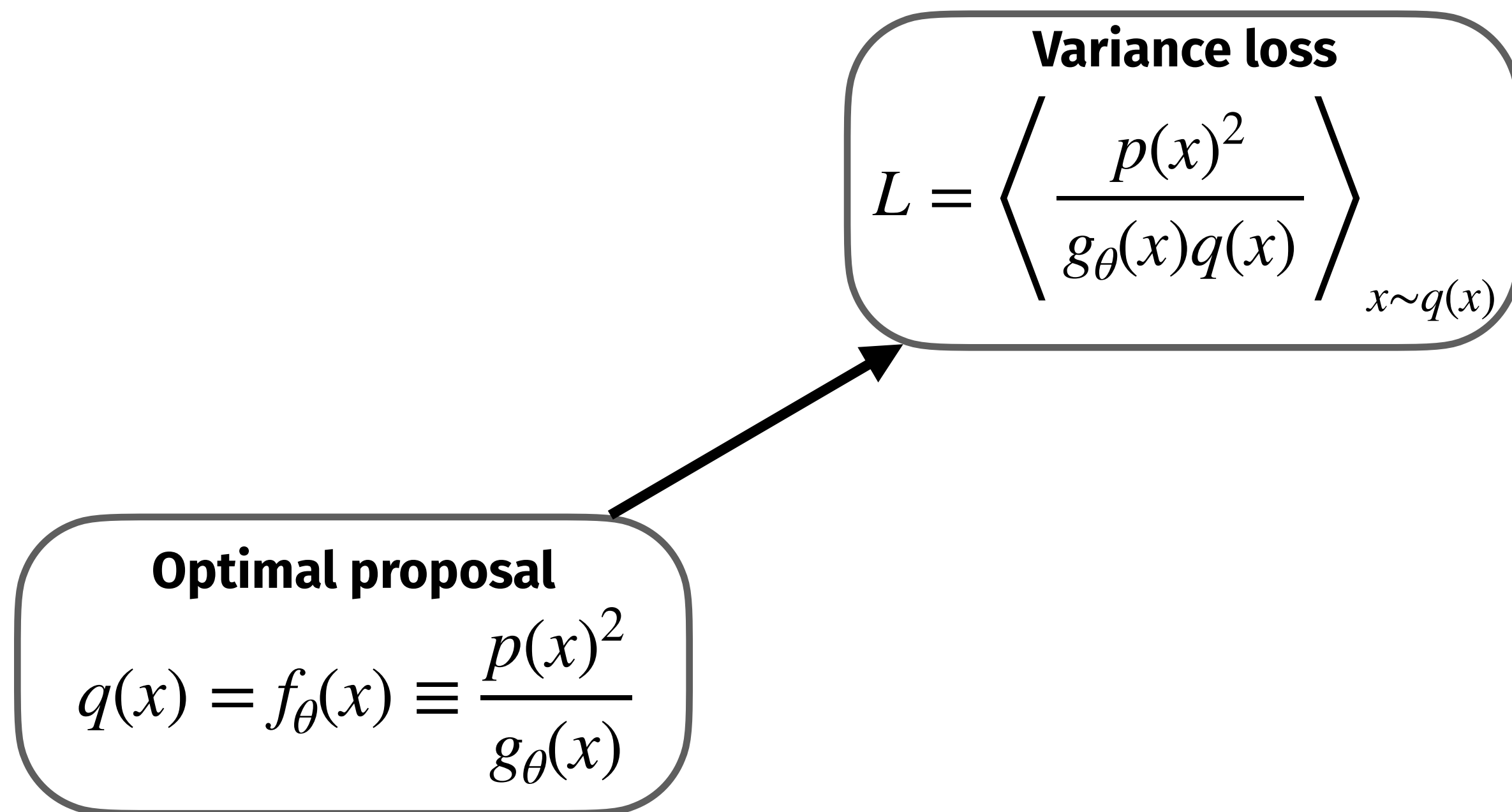


## Pre-training

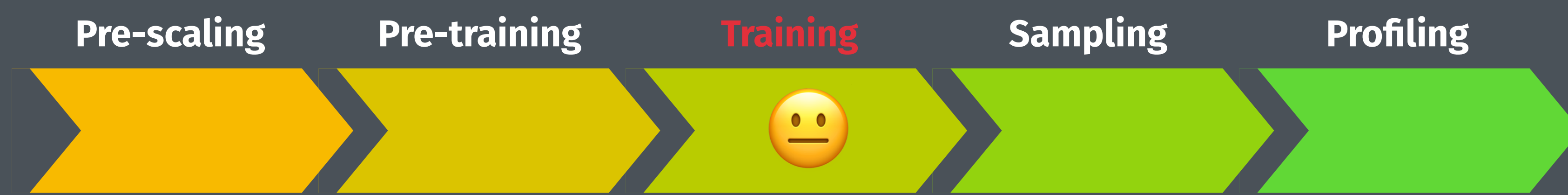
- use samples from pre-scaling to train network for a few steps
- better starting point for main training



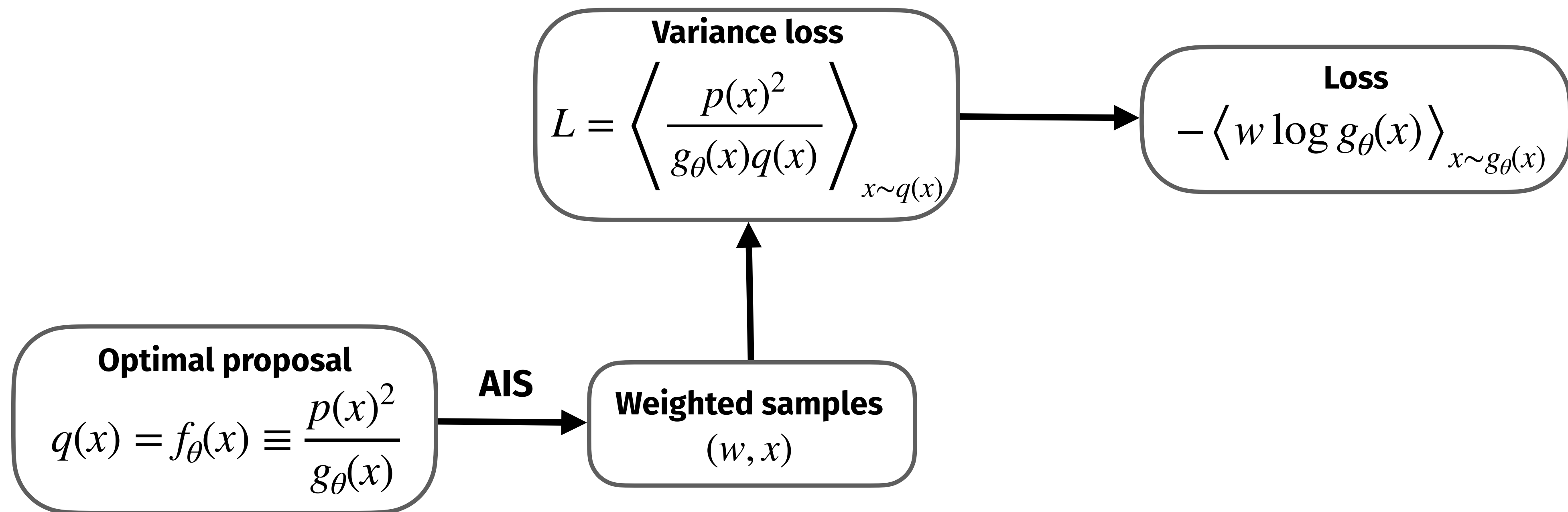
- How does the main training work? [Midgley et al. [2208.01893](#)]



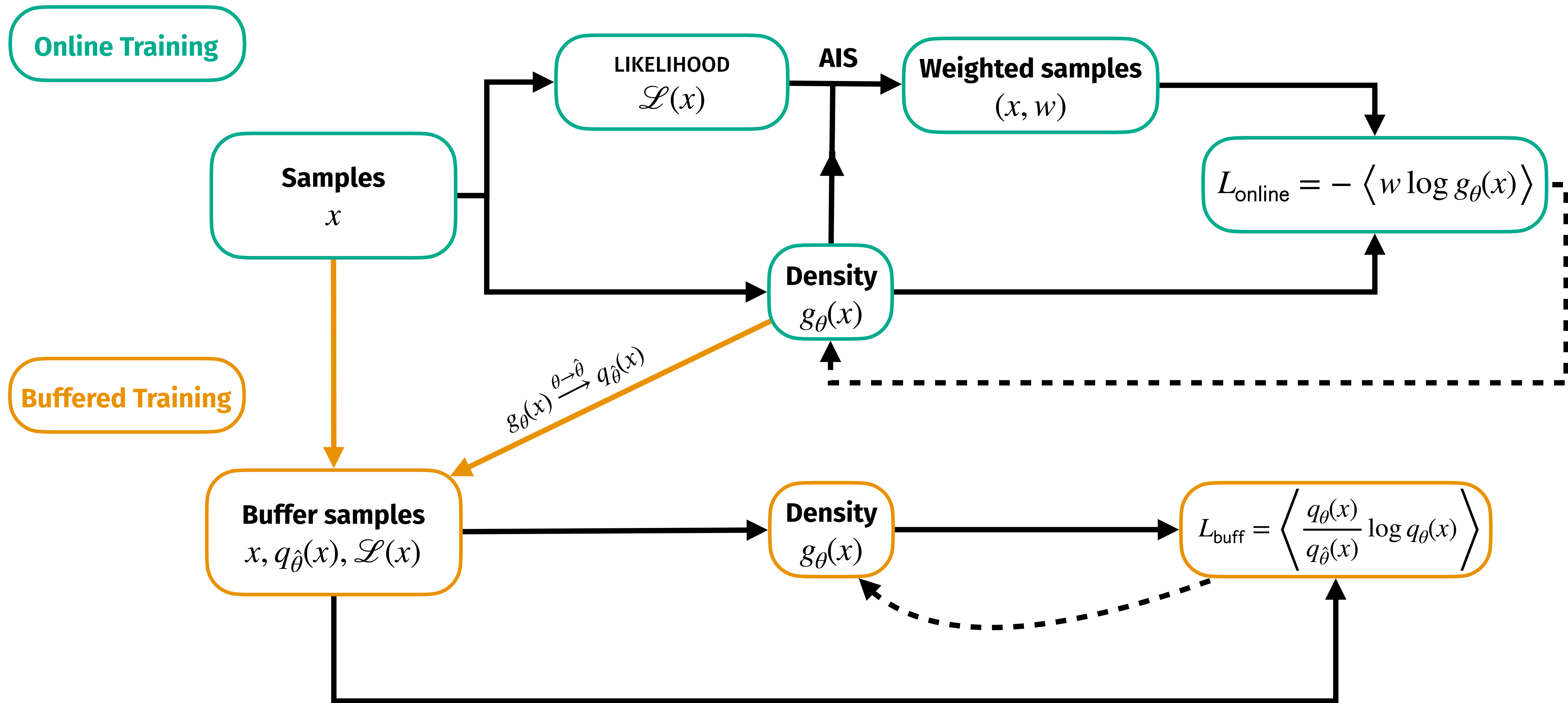
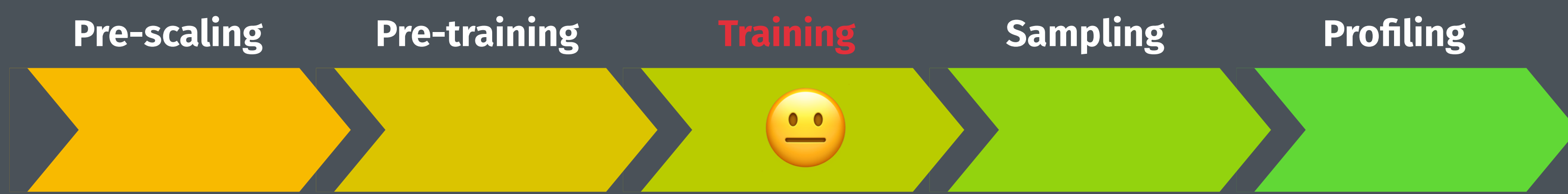
- Sampling directly from  $f_{\theta}(x)$  is **intractable**



- How does the main training work? [Midgley et al. [2208.01893](#)]



- Combining AIS and flow leads to simple loss



# The five steps to happiness

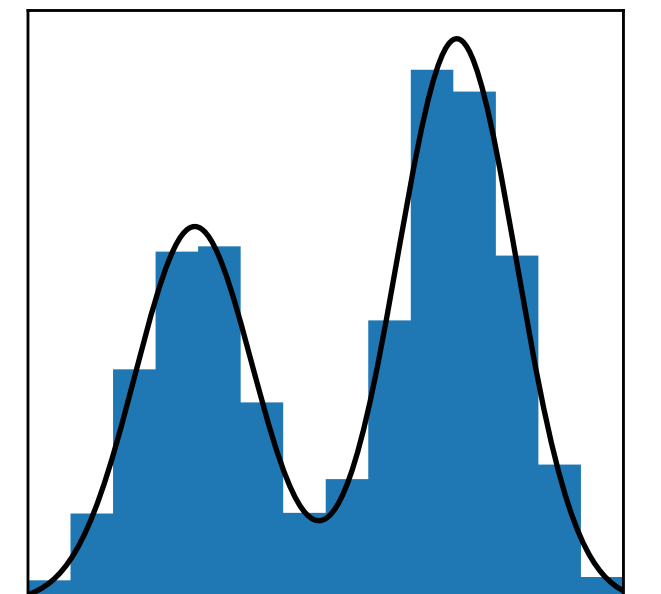
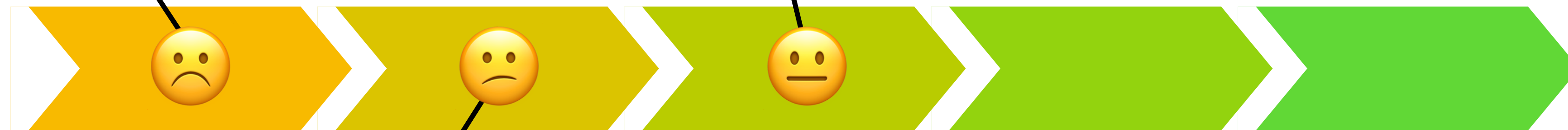
## Pre-scaling

- 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

$p(x)$



## Pre-training

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



# To profile or to marginalize

How to compute constraints  
from high-dimensional likelihood?

## Profiling

Look for the **maximum**  
in parameter space

$$\max_T p(M | T)$$

## Marginalization

**Integrate** over  
parameter space

$$\int_T p(T | M) = \int_T p(M | T) \frac{p(T)}{p(M)}$$

→ Support both in our global fits

# The final steps to happiness

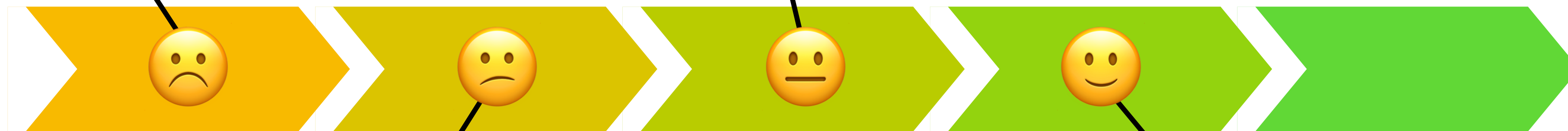
## Pre-scaling

- 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

$p(x)$

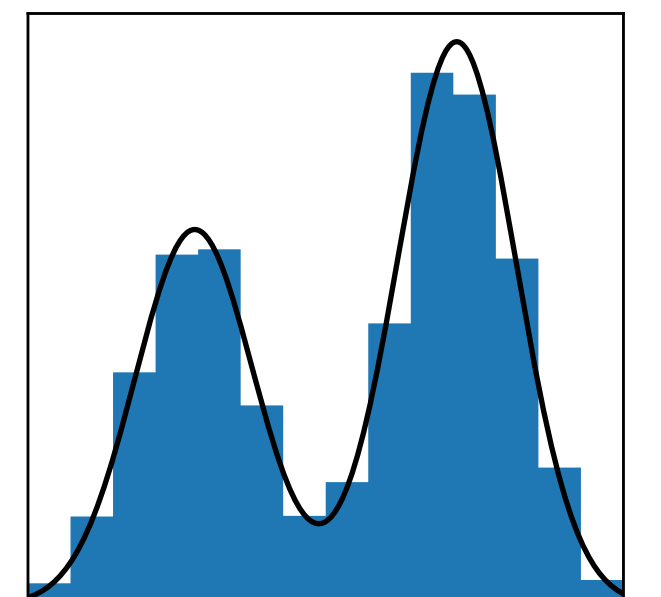


## Pre-training

- 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



# The final steps to happiness

## Pre-scaling

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

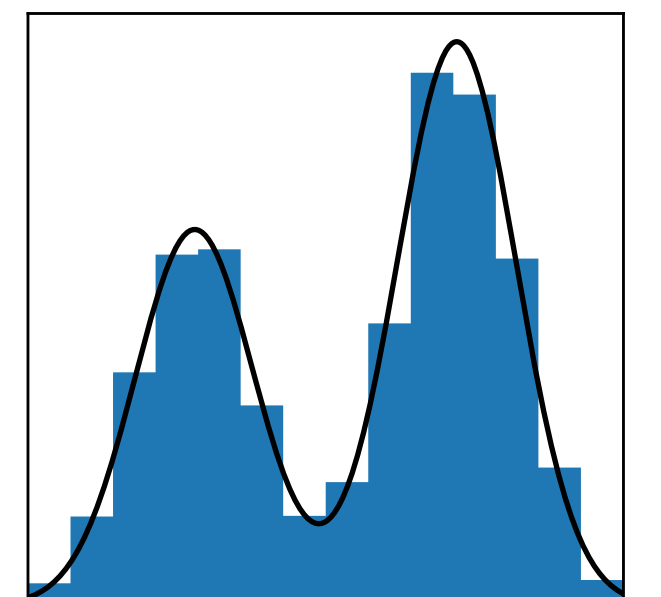
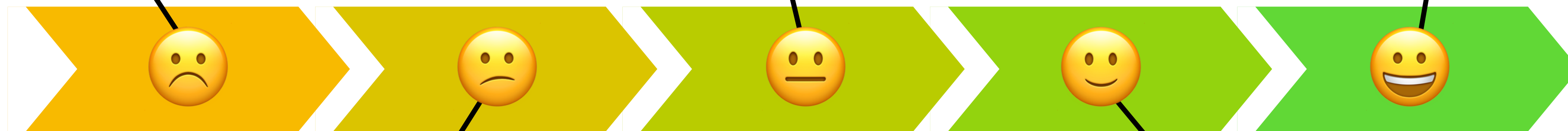
## Training

- 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

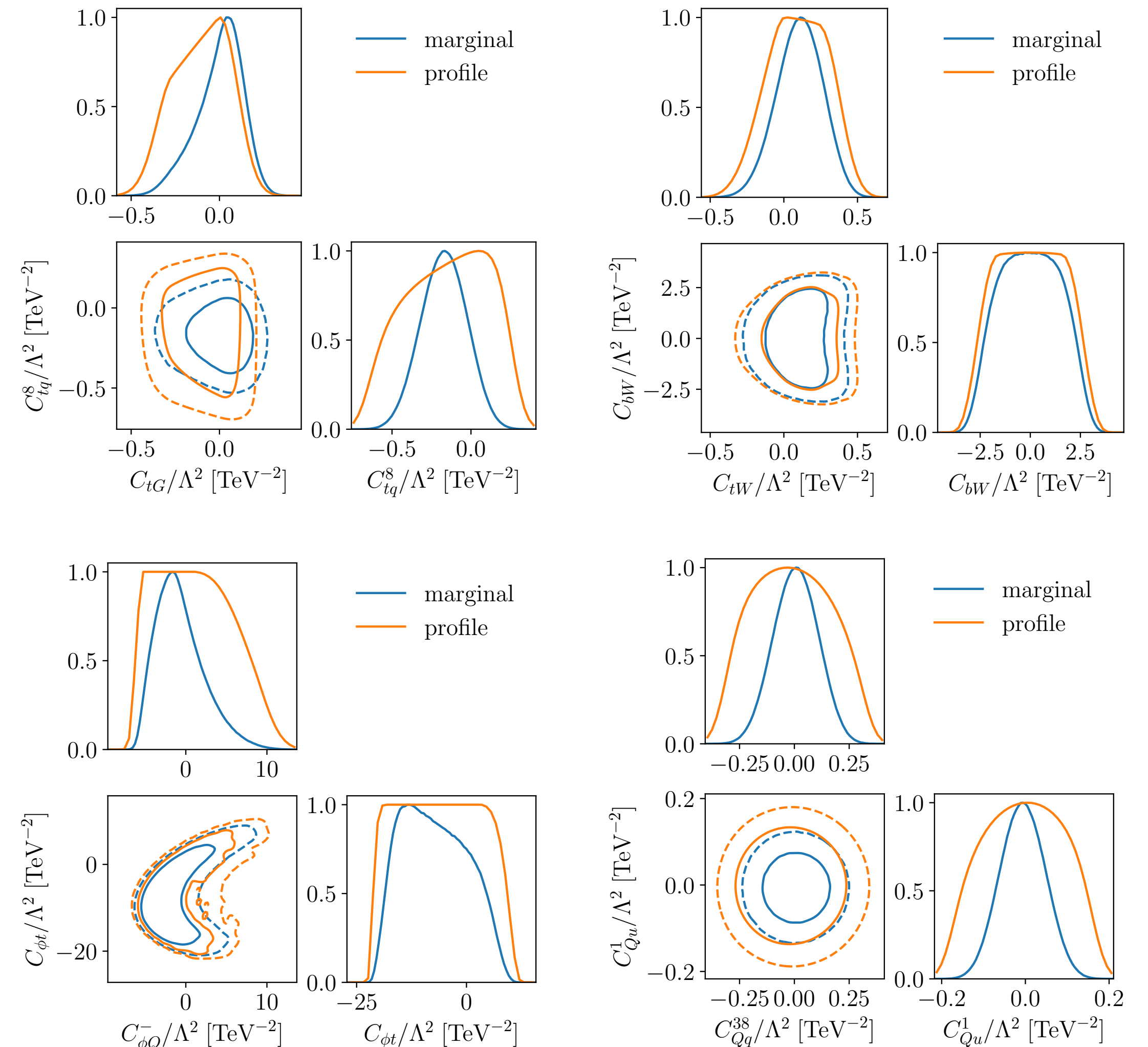
## Sampling

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

# Results

## Top likelihood

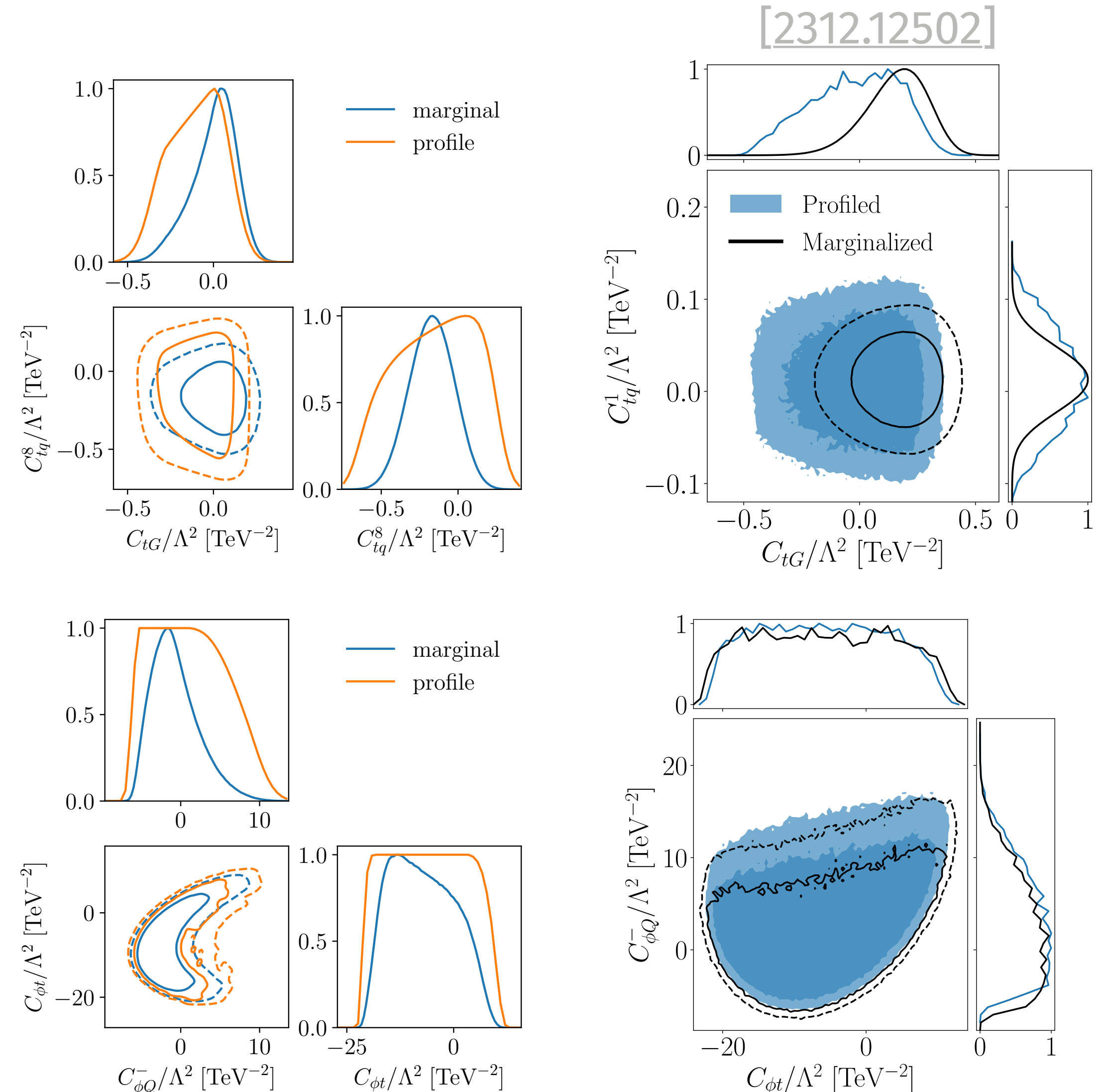
- Relatively simple likelihood, mostly unfolded data
- Large effect from **theory uncertainties**
- **Smooth results** for both profiling and marginalization



# Results

## Top likelihood

- Relatively simple likelihood, mostly unfolded data
- Large effect from **theory uncertainties**
- **Smooth results** for both profiling and marginalization





# Performance

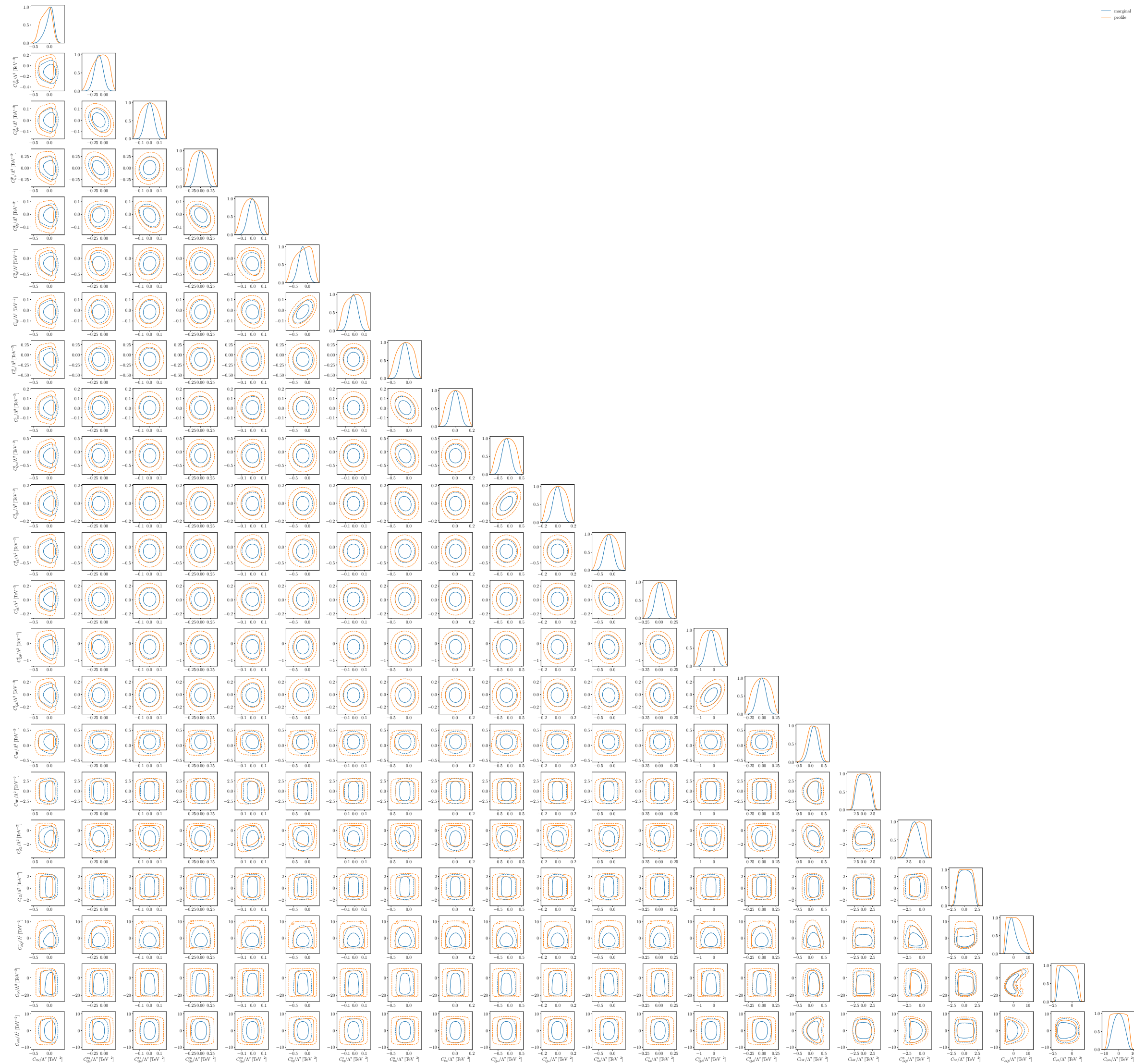
## Top likelihood

- For simple likelihoods sampling on both CPU and GPU is fast
- Training and sampling finished in **around a minute**
- Most time spent on **profiling**

	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



# Performance

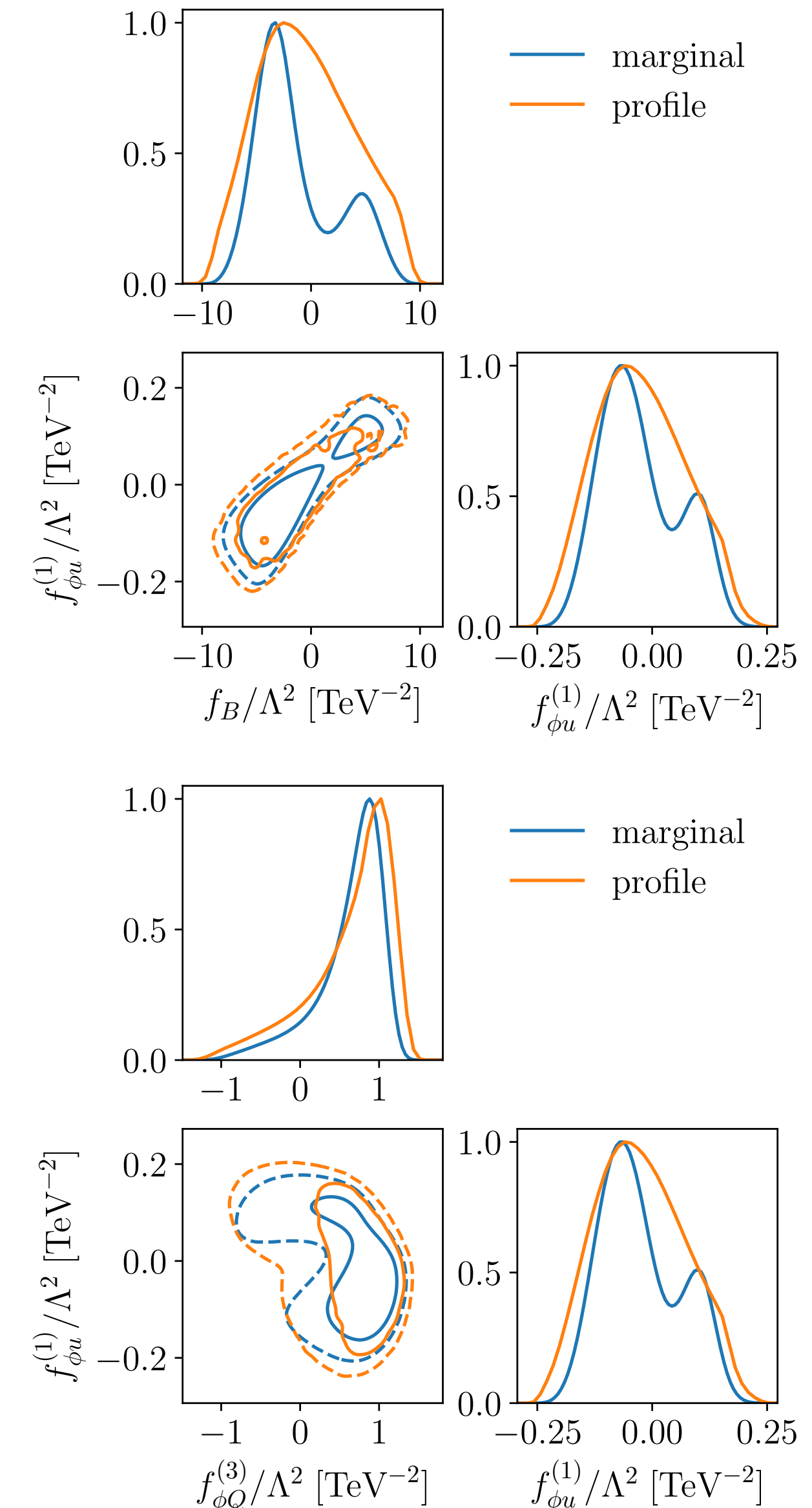


	Top	Higgs-gauge	Combined
Dimensions	22	20	42
Training batches	100	2000	6000
Samples	10M	200M	100M
Effective sample size	7.1M	97M	21M
Pre-scaling time	7s	3.5min	5.3min
Pre-training time	18s	1.7min	2.5min
Training time	36s	17.3min	1.2h
Sampling time	26s	14.8min	17.6min
Profiling time	17.7min	24.8min	3.7h
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

# Results

## Higgs likelihood

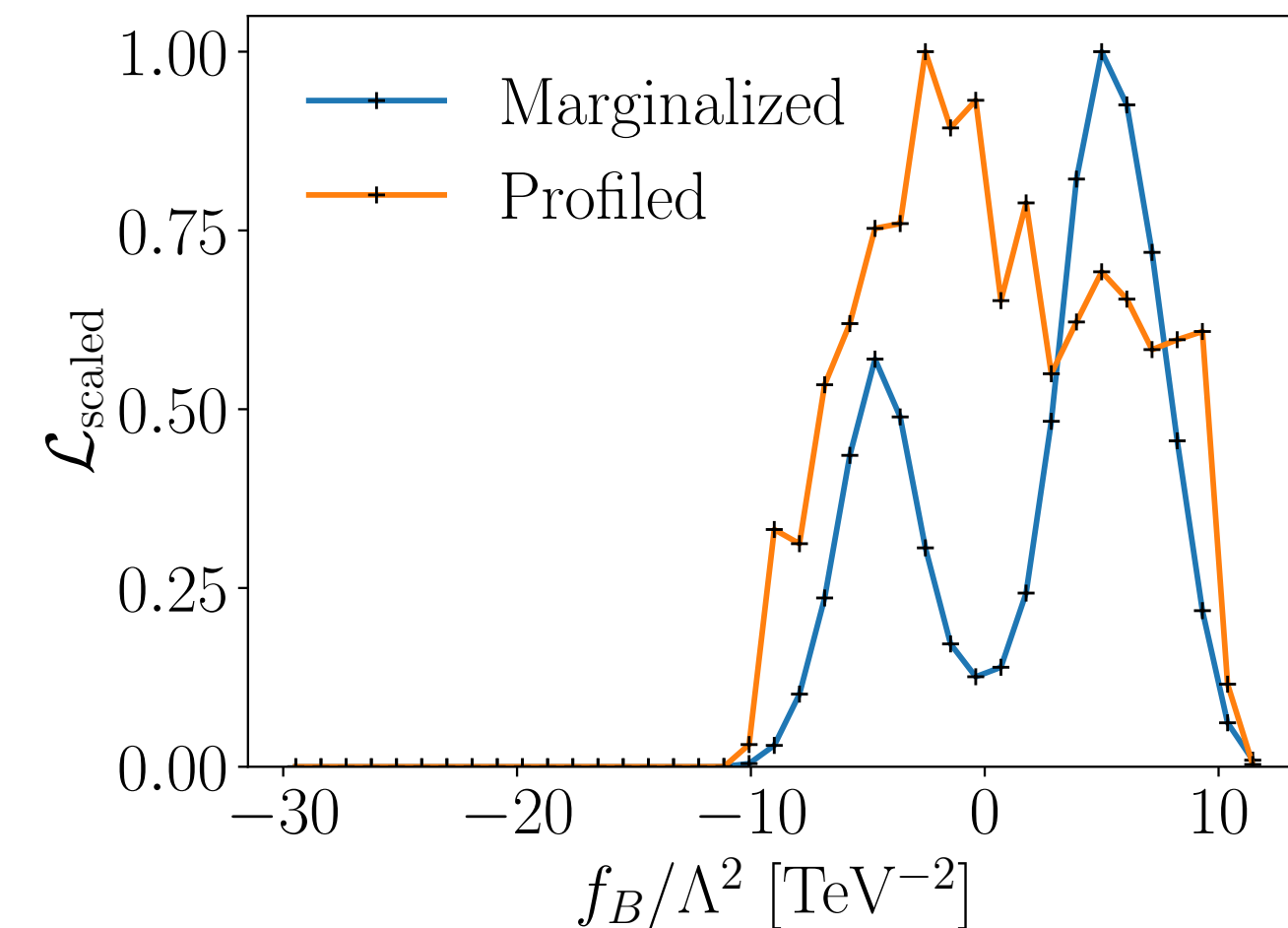
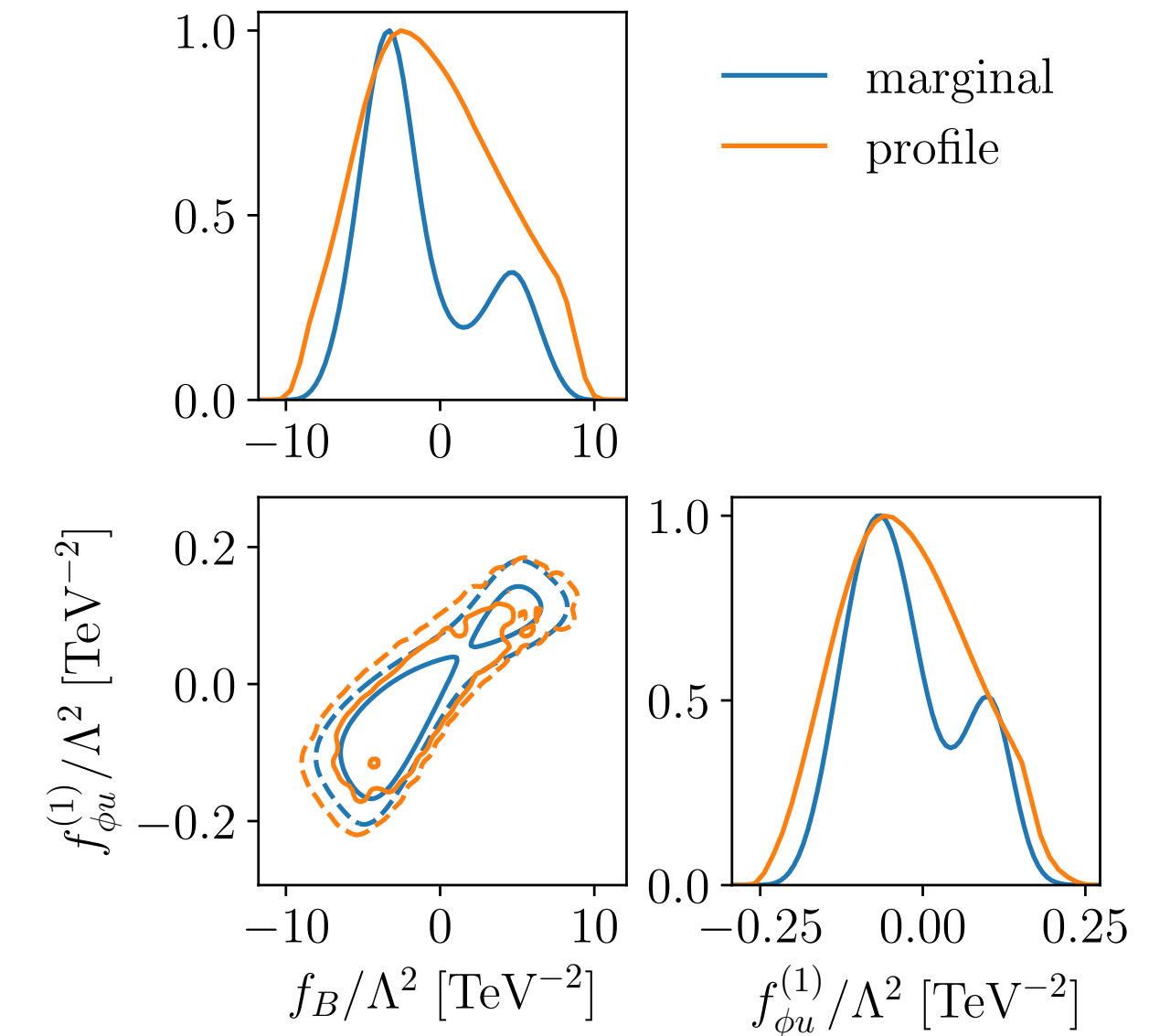
- More **complicated likelihood**
  - more measurements, less unfolded data
- Many coefficients with **multiple modes**
- Significantly **smoother results** for profiling



# Results

## Higgs likelihood

- More **complicated likelihood**
  - more measurements, less unfolded data
- Many coefficients with **multiple modes**
- Significantly **smoother results** for profiling





# Performance

## Higgs likelihood

- Already requires **4x CPUs** for decent results
- Slightly longer training required, still **significantly faster** even for complicated likelihoods

	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

# Performance

## Combined likelihood

- Full combined likelihood requires serious training
- Sampling **orders of magnitude** faster than before
- **Reliable profiled results**, previously out of reach for CPU implementation

	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



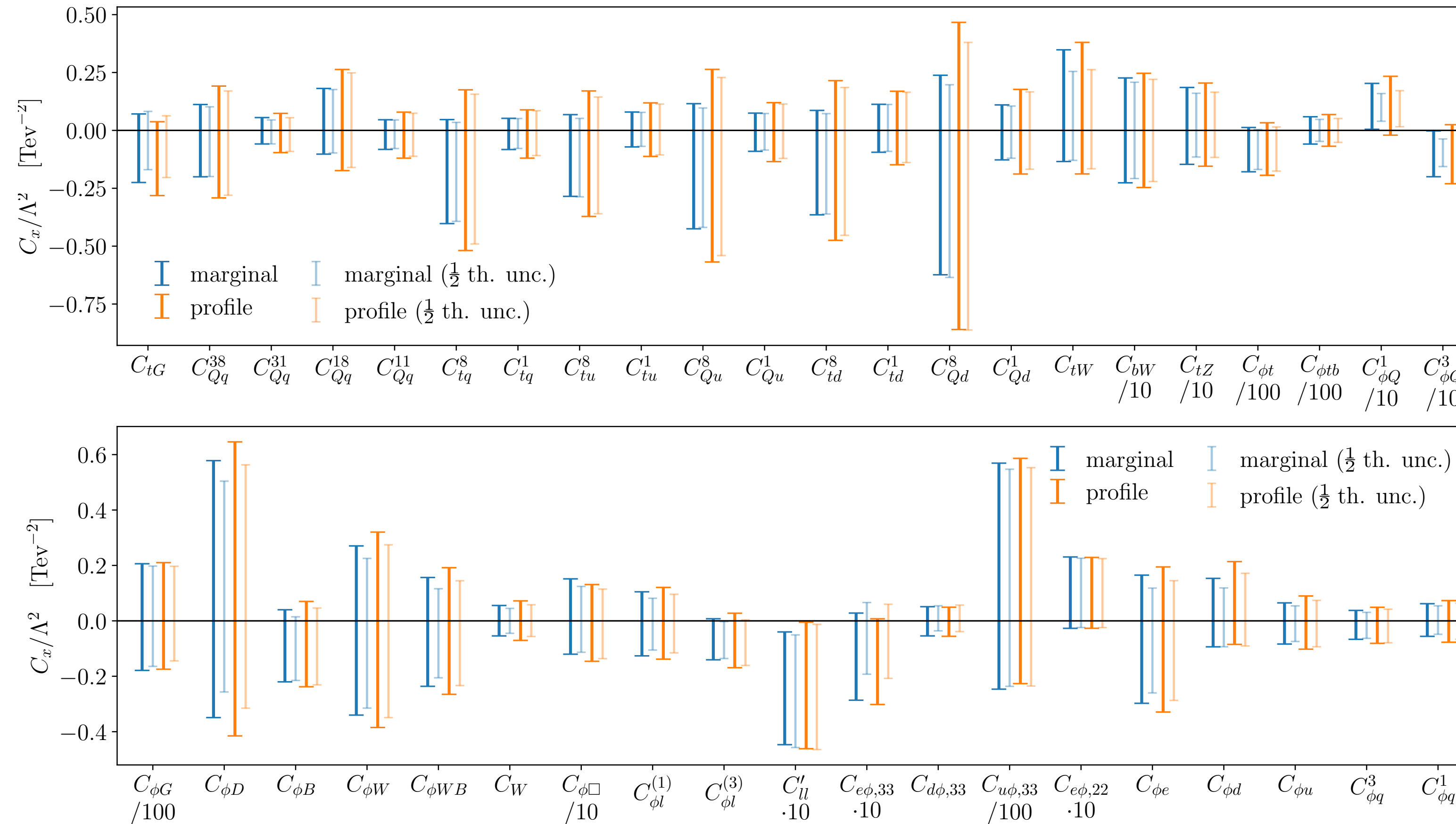
# Performance



	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



# Final results



- Previously: Takes (at least) a day to compute these constraints
- **Now:** Can find signs for new physics **during lunch break** (+coffee)

# Summary

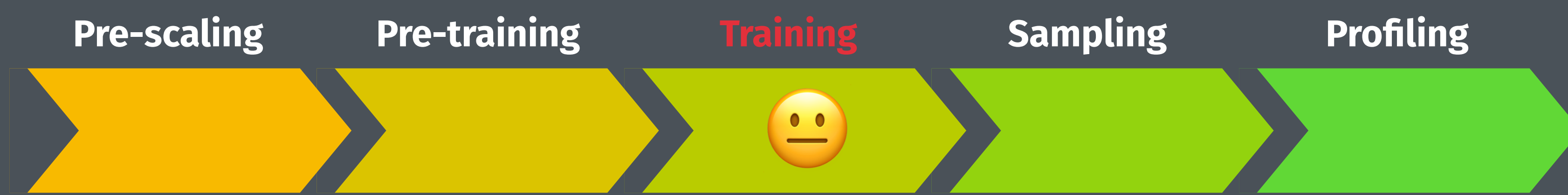
- NIS: powerful tool to **accelerate sampling**
- Significantly **improves sampling** of complex likelihoods while simultaneously providing **smoother** results
- The five steps to happiness can help fundamentally improve future SFITTER studies
- Perfect time to implement new kinds of data formats

Thank you, that's it!

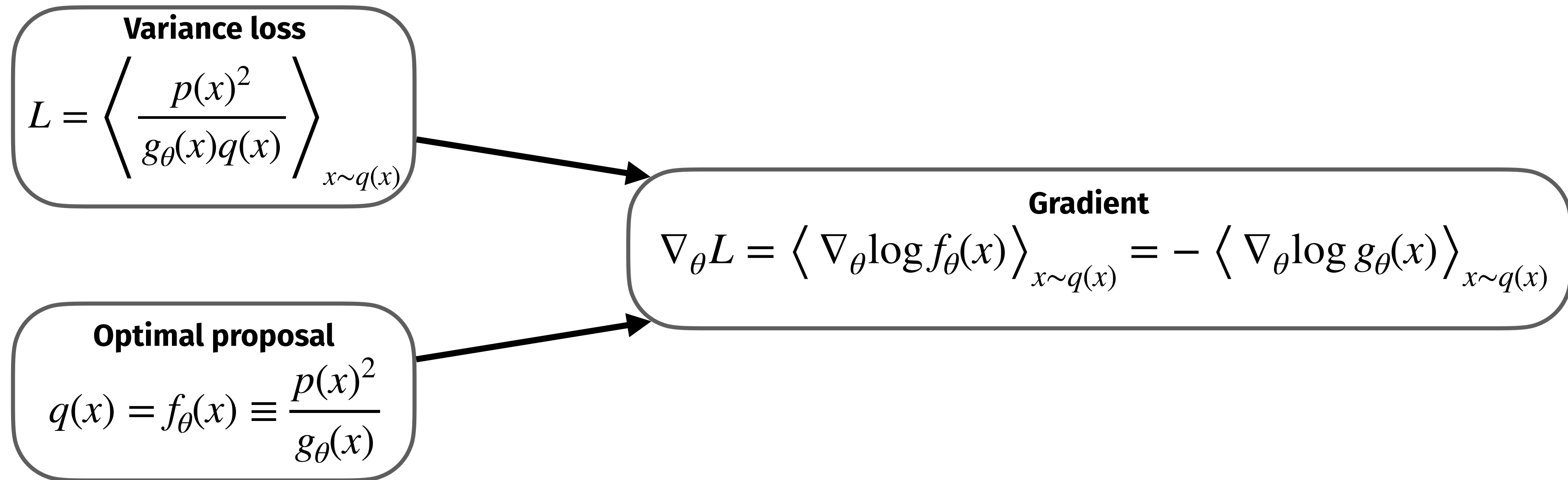
# Appendix / Backup

# 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		

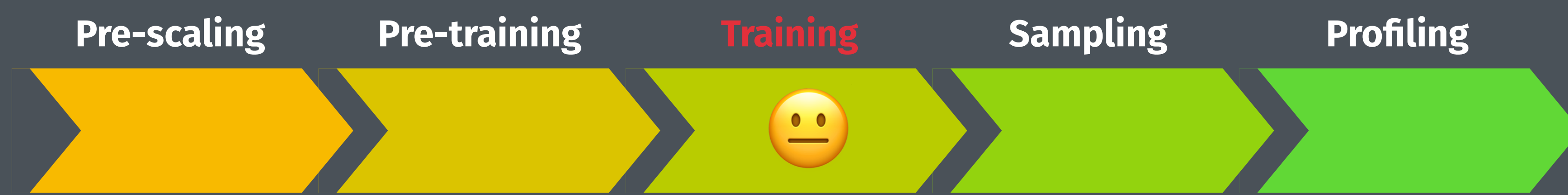


- How does the main training work?

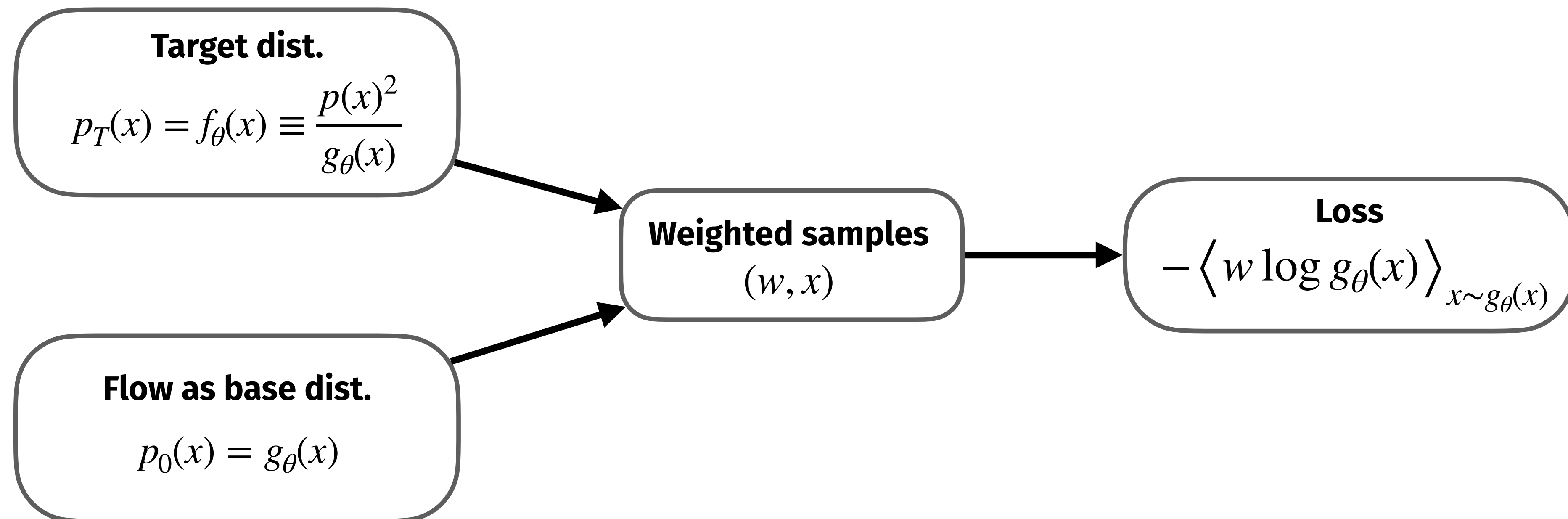


- Sampling directly from  $f_\theta(x)$  is **intractable** [Midgley et al. 2208.01893]





- Again annealed importance sampling can help [Midgley et al. [2208.01893](#)]



- Simple loss after combining AIS and our flow

# Top operator definitions

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

# Higgs-gauge operator definitions (HISZ)

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



# Higgs-gauge operator definitions (Warsaw)

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

# Full top dataset

Experiment	Energy [TeV]	$\mathcal{L}$ [fb <sup>-1</sup> ]	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	$lj$	$\frac{1}{\sigma} \frac{d\sigma}{dp_T^i}$	7		[88–90]
CMS [87]	8	19.7	$ll$	$\frac{1}{\sigma} \frac{d\sigma}{dp_T^i}$	5		[88–90]
ATLAS [91]	8	20.3	$lj$	$\frac{1}{\sigma} \frac{d\sigma}{dm_{i\bar{i}}}$	7		[88–90]
CMS [82]	13	137	$lj$	$\frac{1}{\sigma} \frac{d\sigma}{dm_{i\bar{i}}}$	15	✓	[45]
CMS [92]	13	35.9	$ll$	$\frac{1}{\sigma} \frac{d\sigma}{d\Delta y_{i\bar{i}}}$	8		[88–90]
ATLAS [93]	13	36	$lj$	$\frac{1}{\sigma} \frac{d\sigma}{dm_{i\bar{i}}}$	9	✓	[45]
ATLAS [94]	13	139	$aj, \text{high-}p_T$	$\frac{1}{\sigma} \frac{d\sigma}{dm_{i\bar{i}}}$	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 <sup>-1</sup> ]	Channel	Observable	# Bins	New Likelihood	QCD k-factor
ATLAS [108]	7	4.59	$t\text{-ch}$	$\sigma_{tq+\bar{t}q}$			
CMS [109]	7	1.17 (e), 1.56 ( $\mu$ )	$t\text{-ch}$	$\sigma_{tq+\bar{t}q}$			
ATLAS [110]	8	20.2	$t\text{-ch}$	$\sigma_{tq}, \sigma_{\bar{t}q}$			
CMS [111]	8	19.7	$t\text{-ch}$	$\sigma_{tq}, \sigma_{\bar{t}q}$			
ATLAS [112]	13	3.2	$t\text{-ch}$	$\sigma_{tq}, \sigma_{\bar{t}q}$			[113]
CMS [114]	13	2.2	$t\text{-ch}$	$\sigma_{tq}, \sigma_{\bar{t}q}$			[113]
CMS [115]	13	35.9	$t\text{-ch}$	$\frac{1}{\sigma} \frac{d\sigma}{d p_{T,t} }$	5	✓	
CMS [116]	7	5.1	$s\text{-ch}$	$\sigma_{t\bar{b}+\bar{t}b}$			
CMS [116]	8	19.7	$s\text{-ch}$	$\sigma_{t\bar{b}+\bar{t}b}$			
ATLAS [117]	8	20.3	$s\text{-ch}$	$\sigma_{t\bar{b}+\bar{t}b}$			
ATLAS [49]	13	139	$s\text{-ch}$	$\sigma_{t\bar{b}+\bar{t}b}$		✓	✓
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	$tZ$	$\sigma_{tZq}$			
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		$F_0, F_L$		✓	