Simulation Based Inference for SNe Ia cosmology

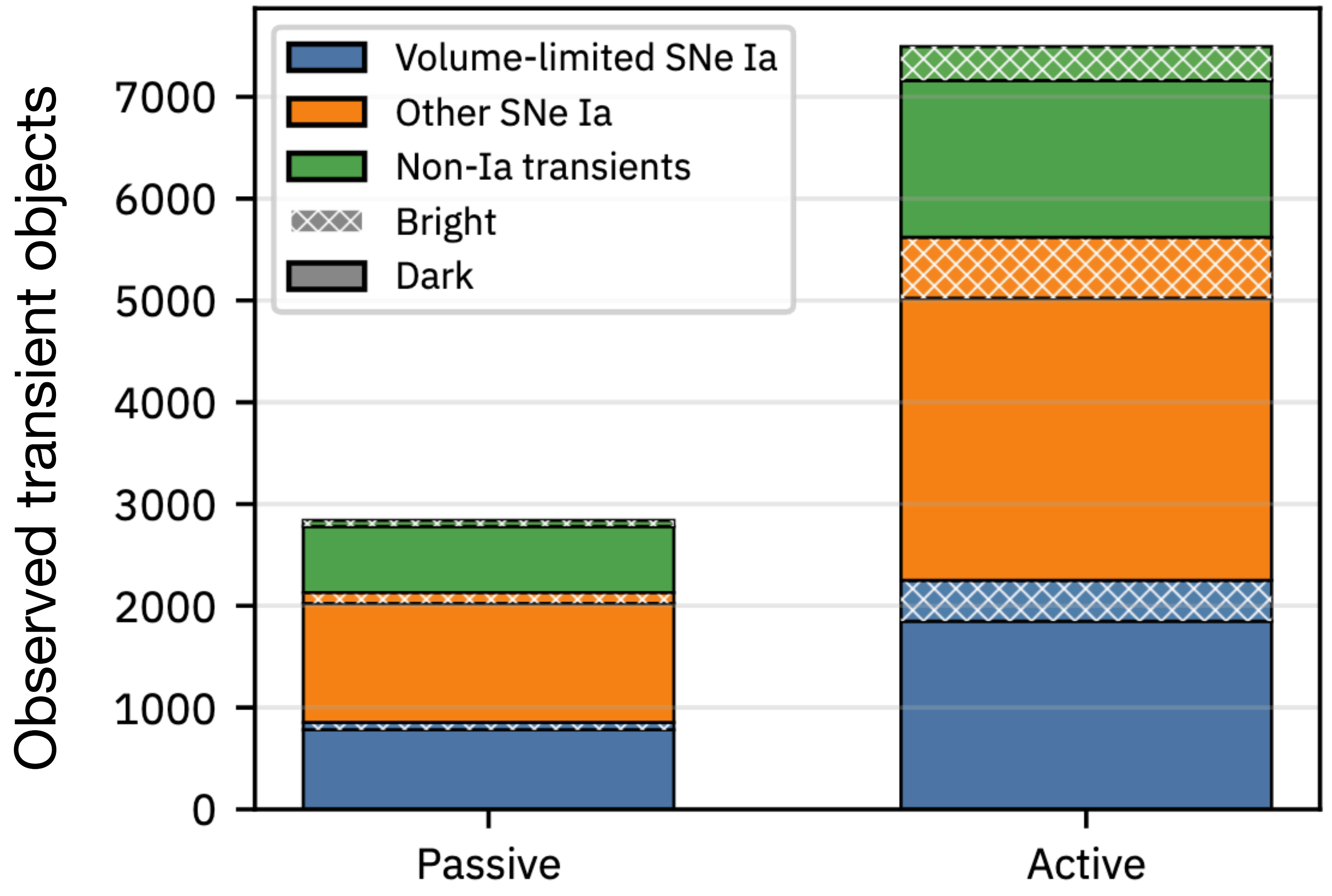
Adam Trigui

Under the supervision of Mickaël Rigault and Florian Ruppin

LSST France 2026, Annecy

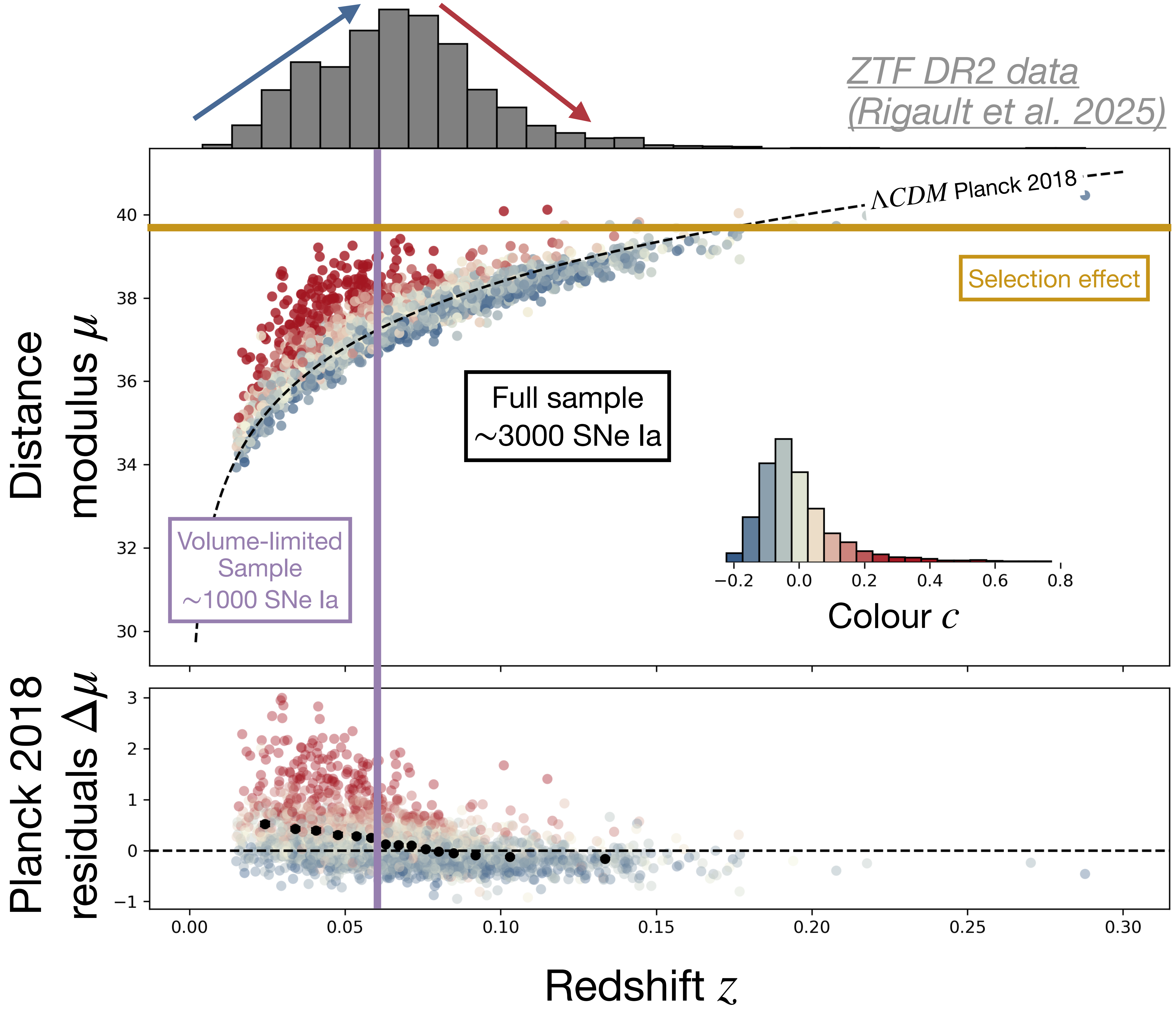
Type Ia SuperNovae

Truong et al. 2026



Up to 7000 spectro SNe Ia each year
~ 100 000 in 10 years !
→Millions of photo SNe Ia

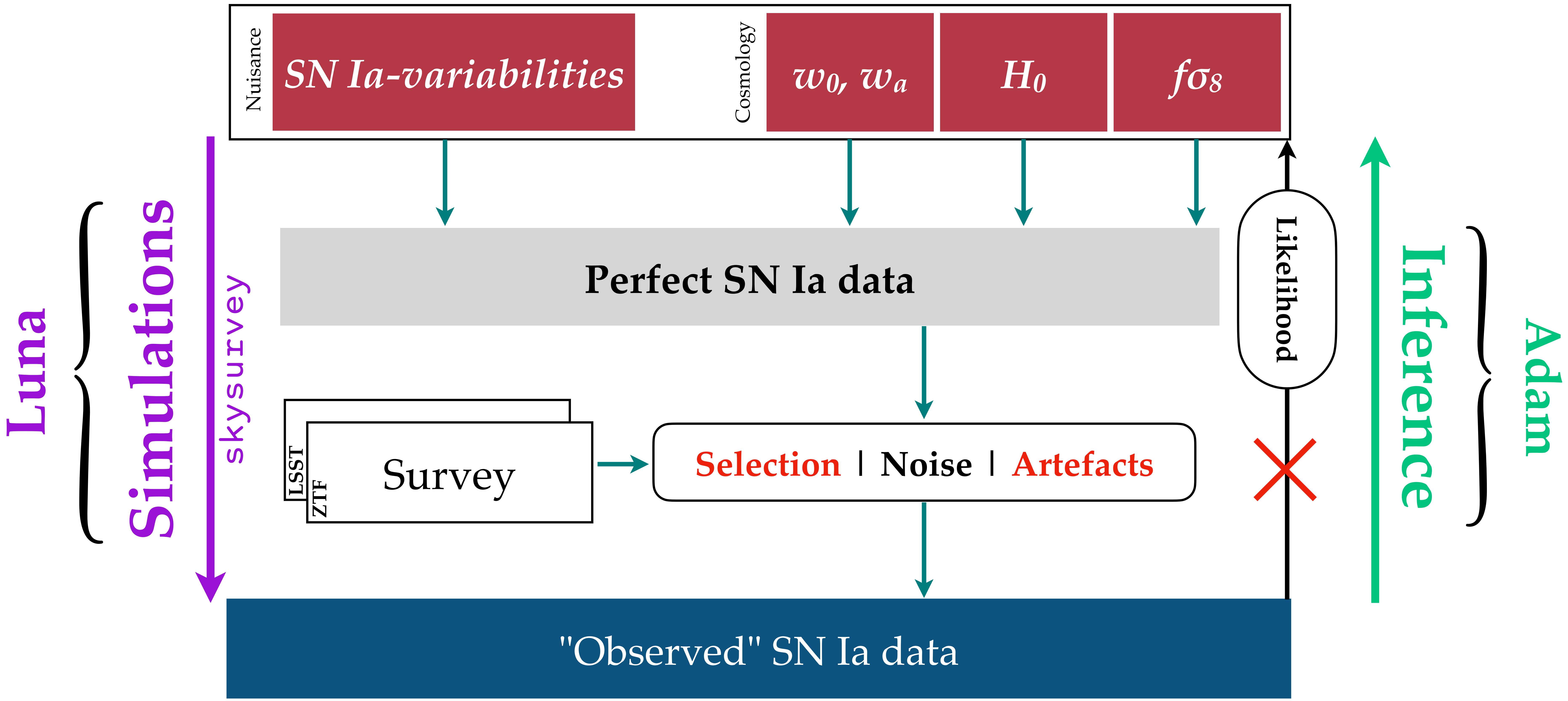
ZTF DR2 data (Rigault et al. 2025)



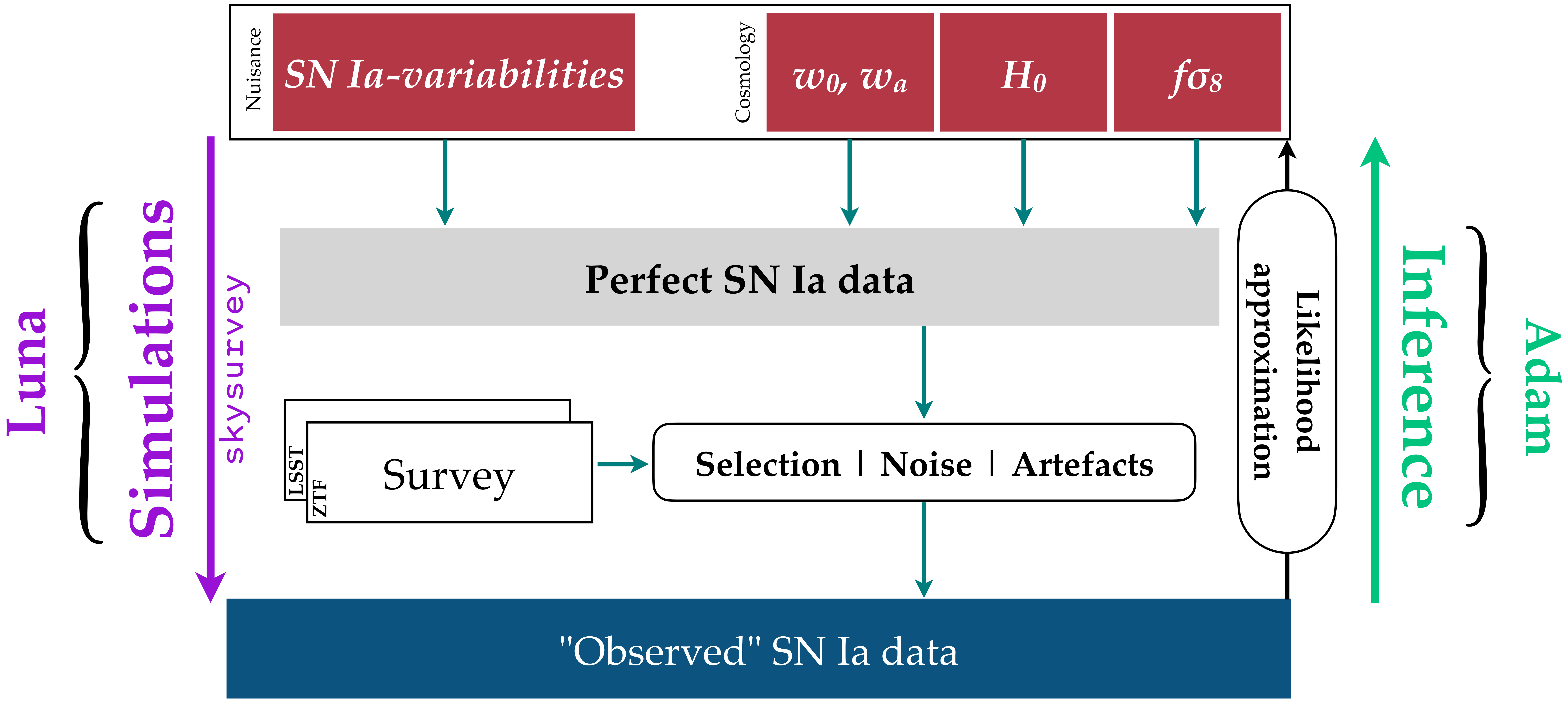
SNe Ia standardisation

Standardisation : $M = M_0 + {\color{blue}\beta} c - {\color{orange}\alpha} x_1 + {\color{red}\gamma} p$

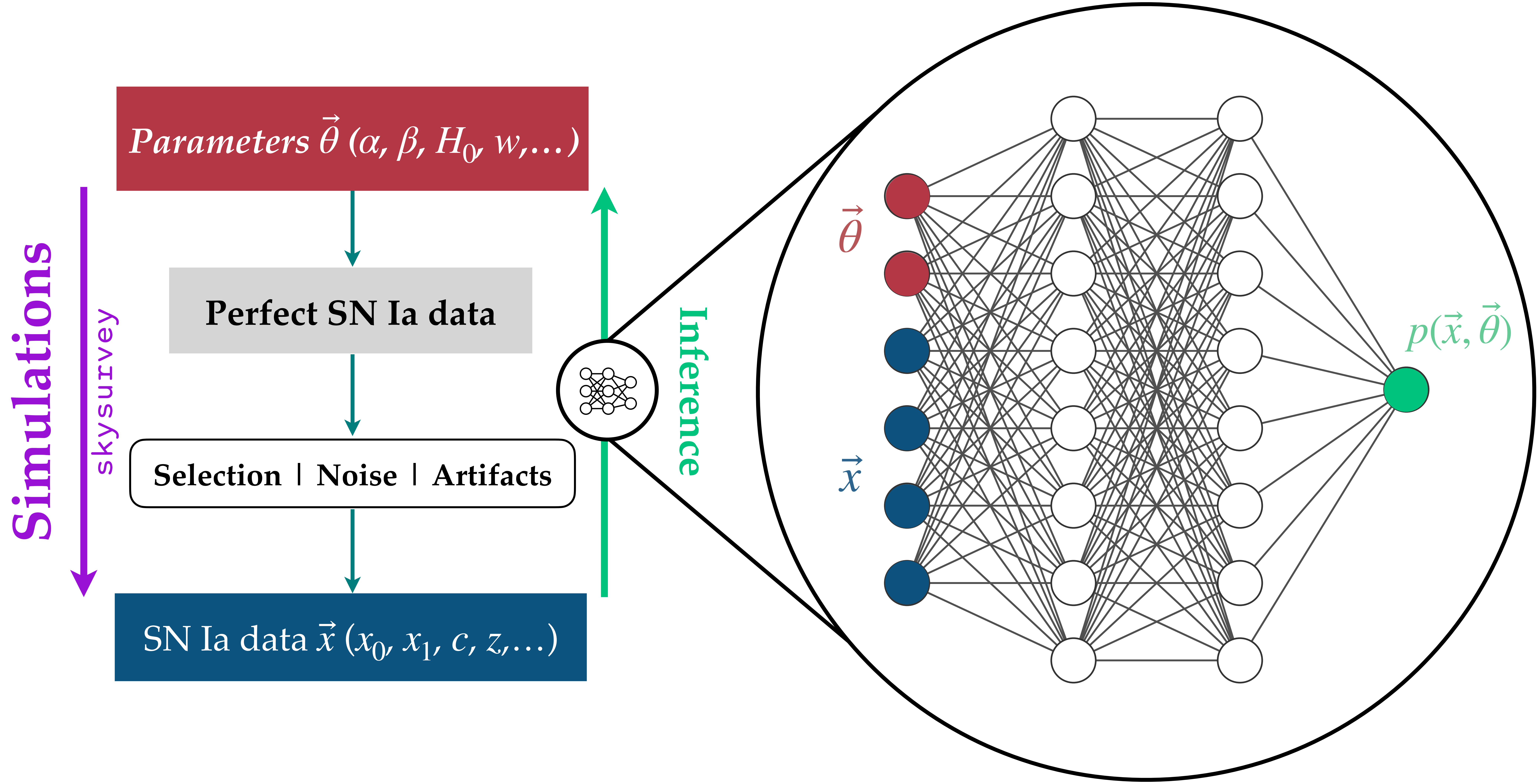
Simulation based inference (SBI)



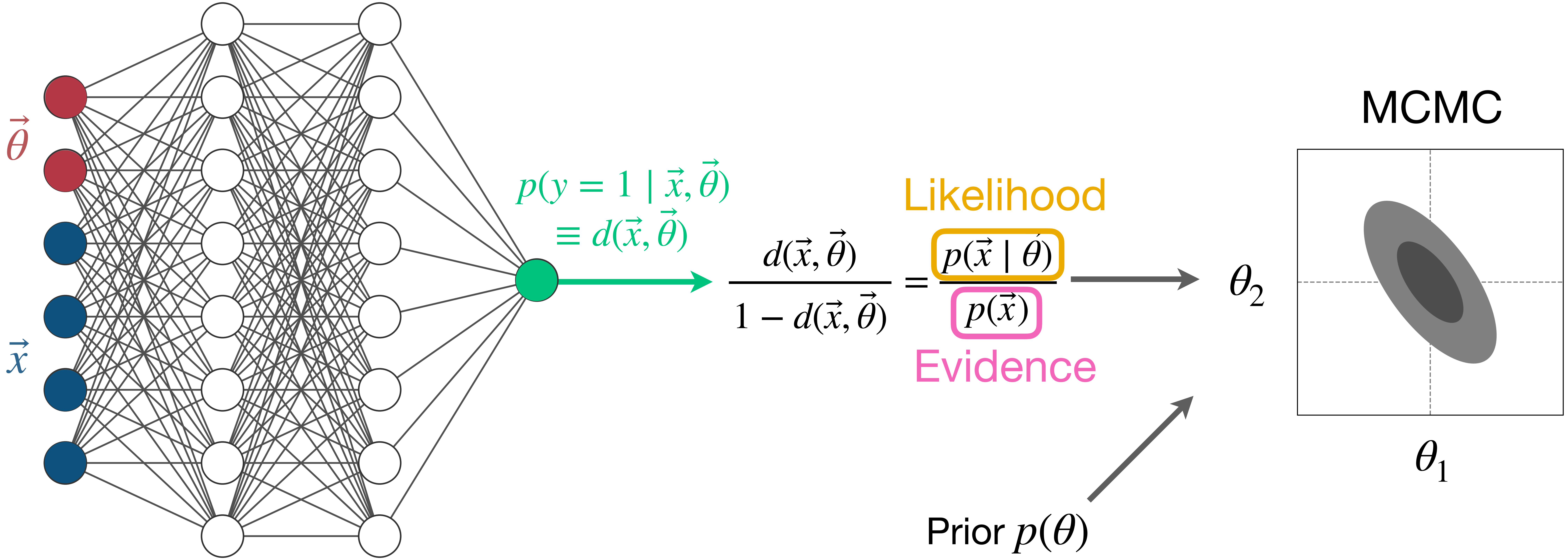
Simulation based inference (SBI)



Neural Ratio Estimator



Neural Ratio Estimator

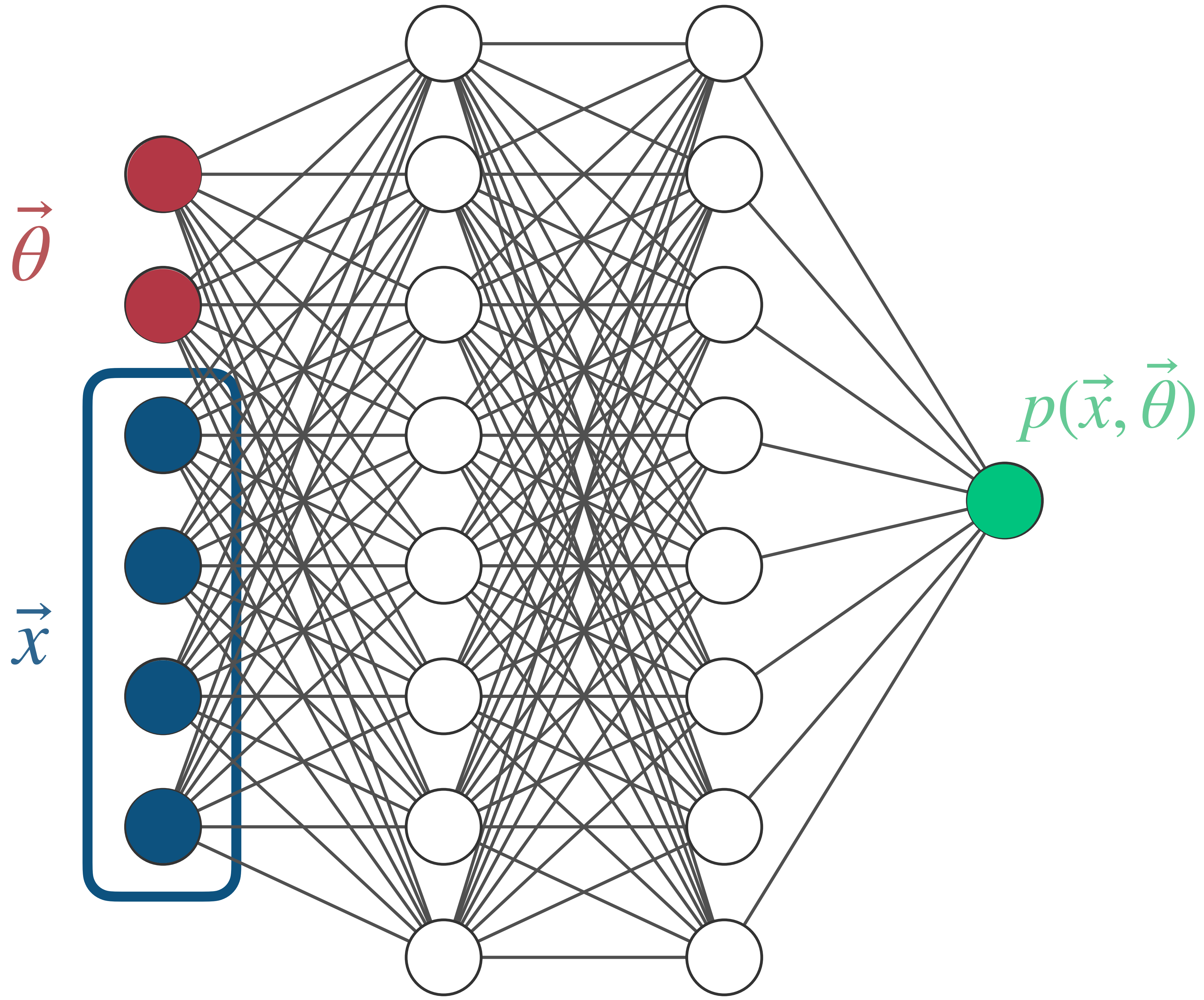


Neural Ratio Estimator

100 000+ SNe Ia

$\times 4+$ observables $(x_0, x_1, c, z, \ldots)$

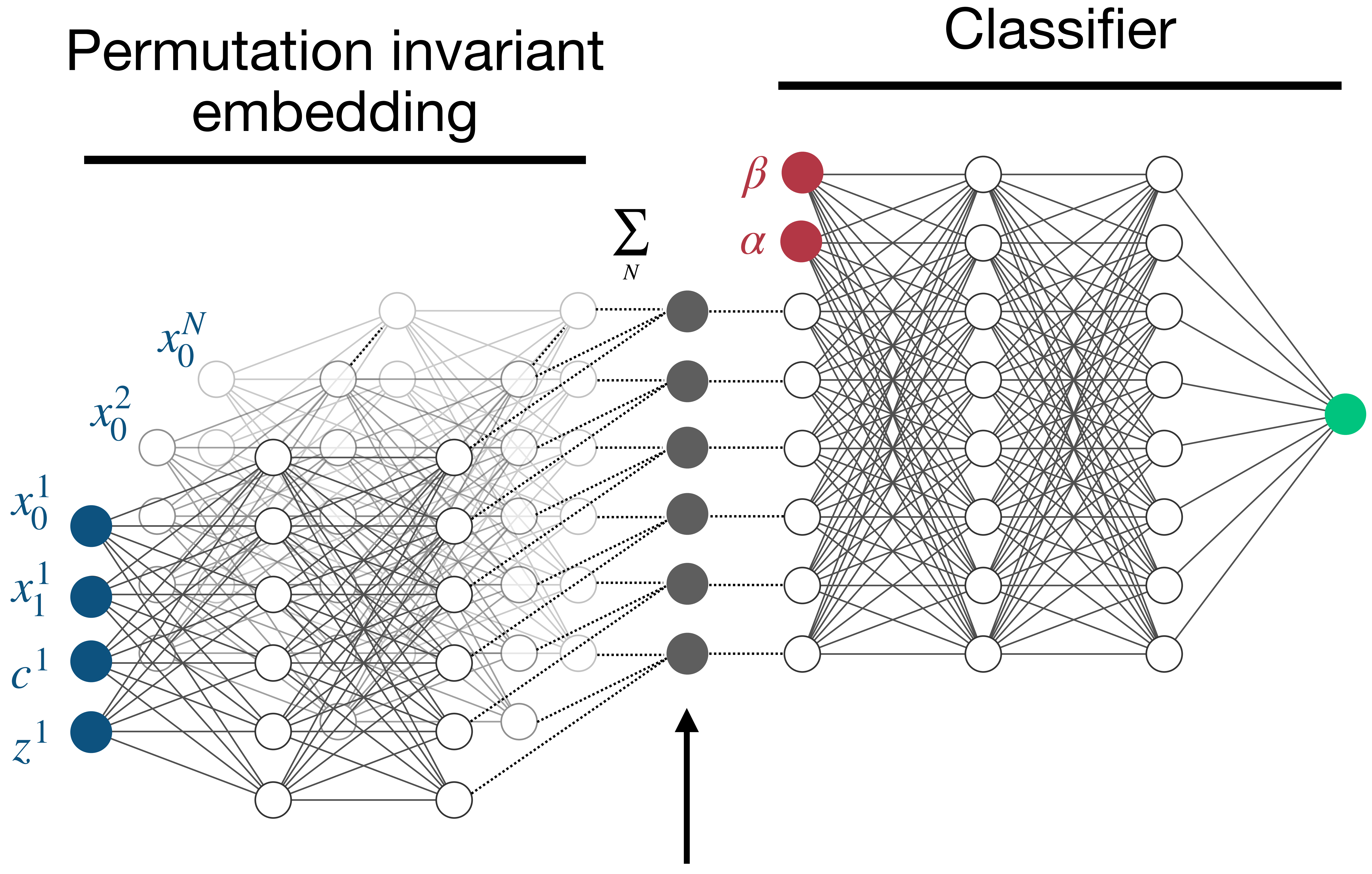
= minimum <u>400 000</u> input neurons !



But :

- Supernovae can be given in any order
- Sample size can vary

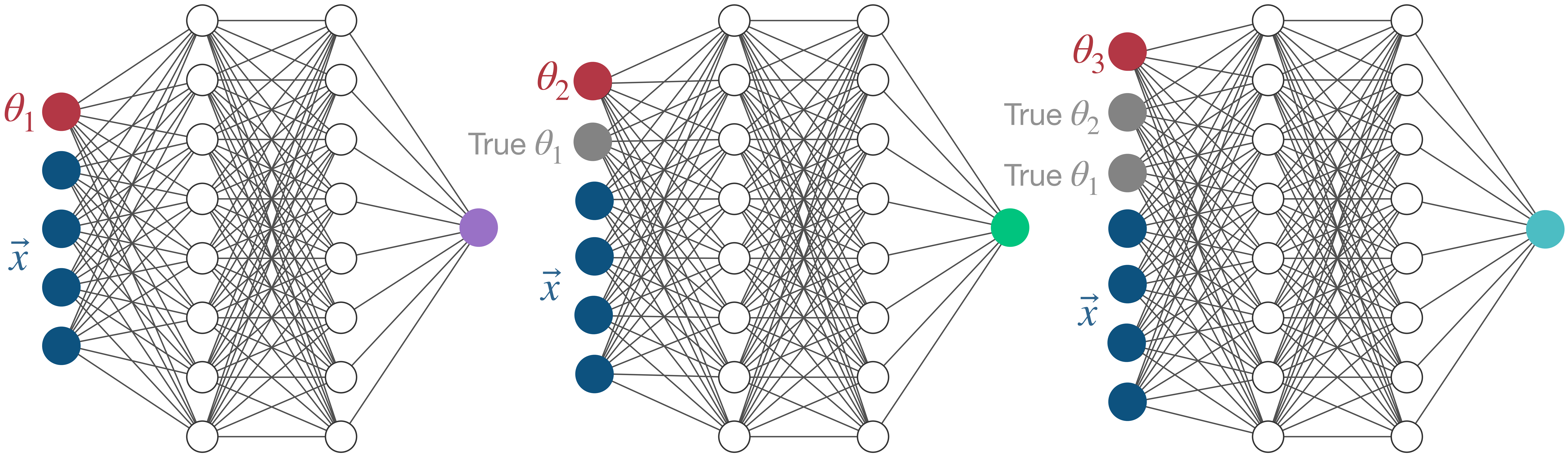
Deepset Zaheer et al. (2018)



Complex information encoded into a small size vector (400000 → 64)

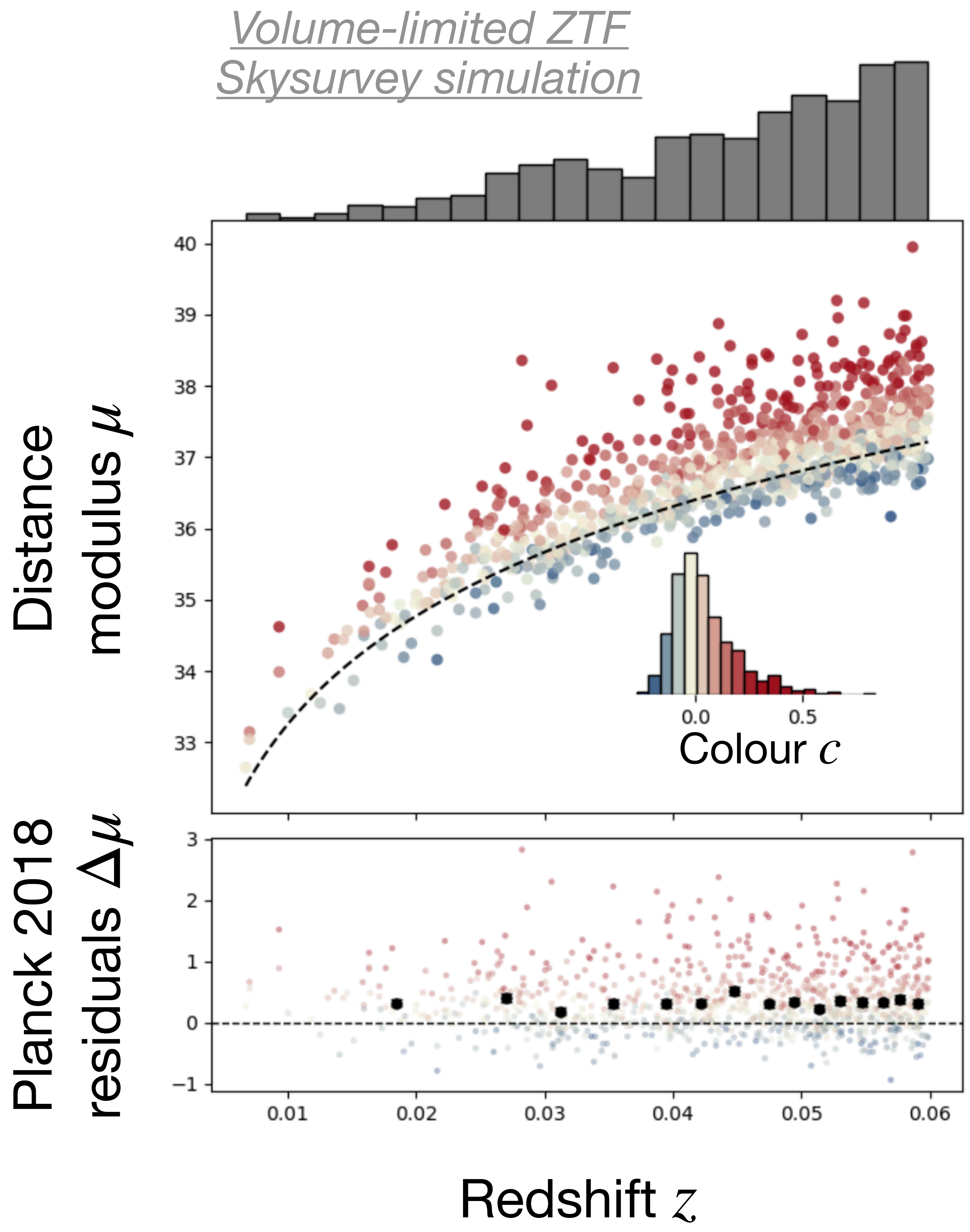
Autoregression Karchev and Trotta (2024)

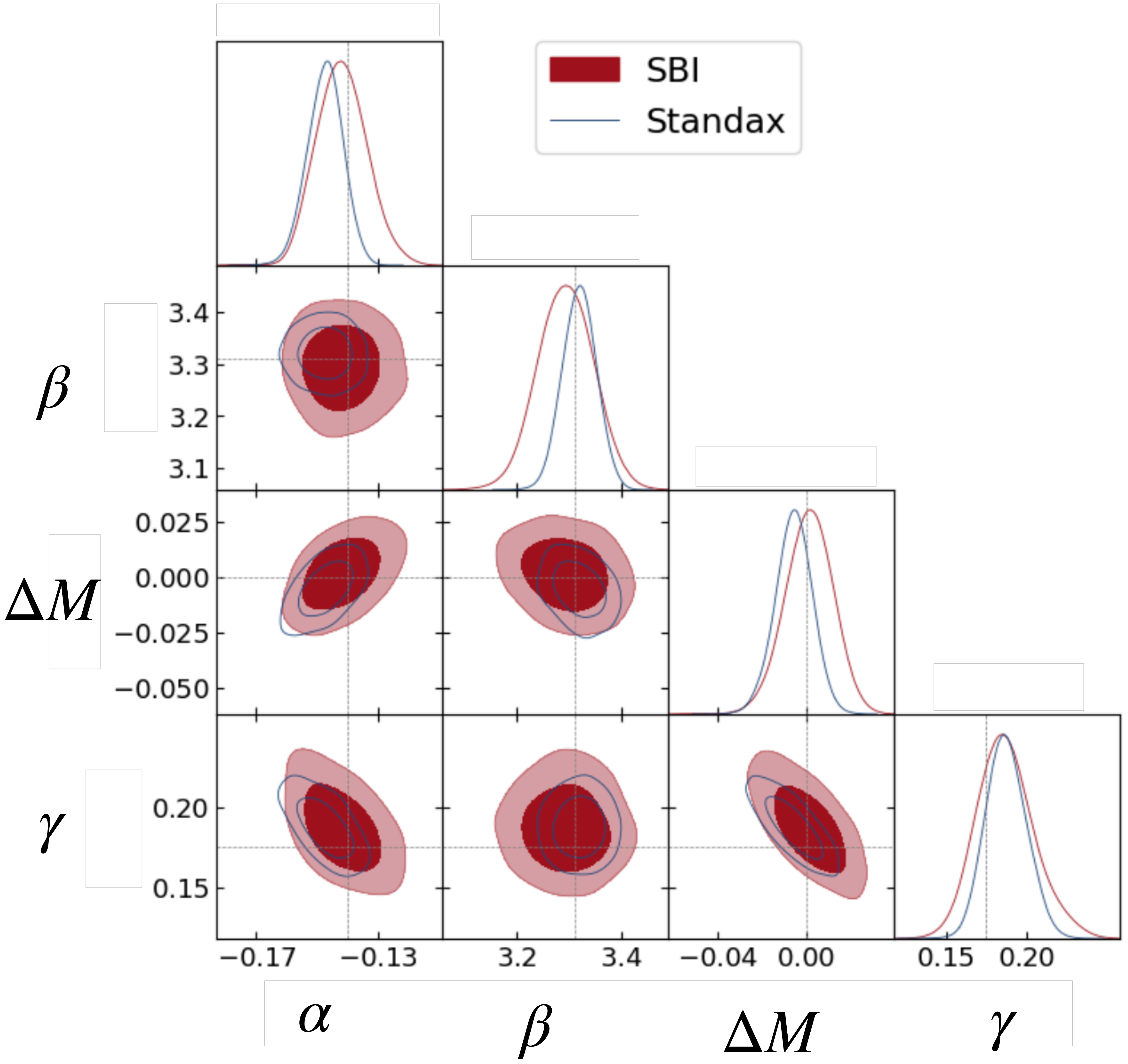
Use joint posteriors to help the network learn difficult parameters.



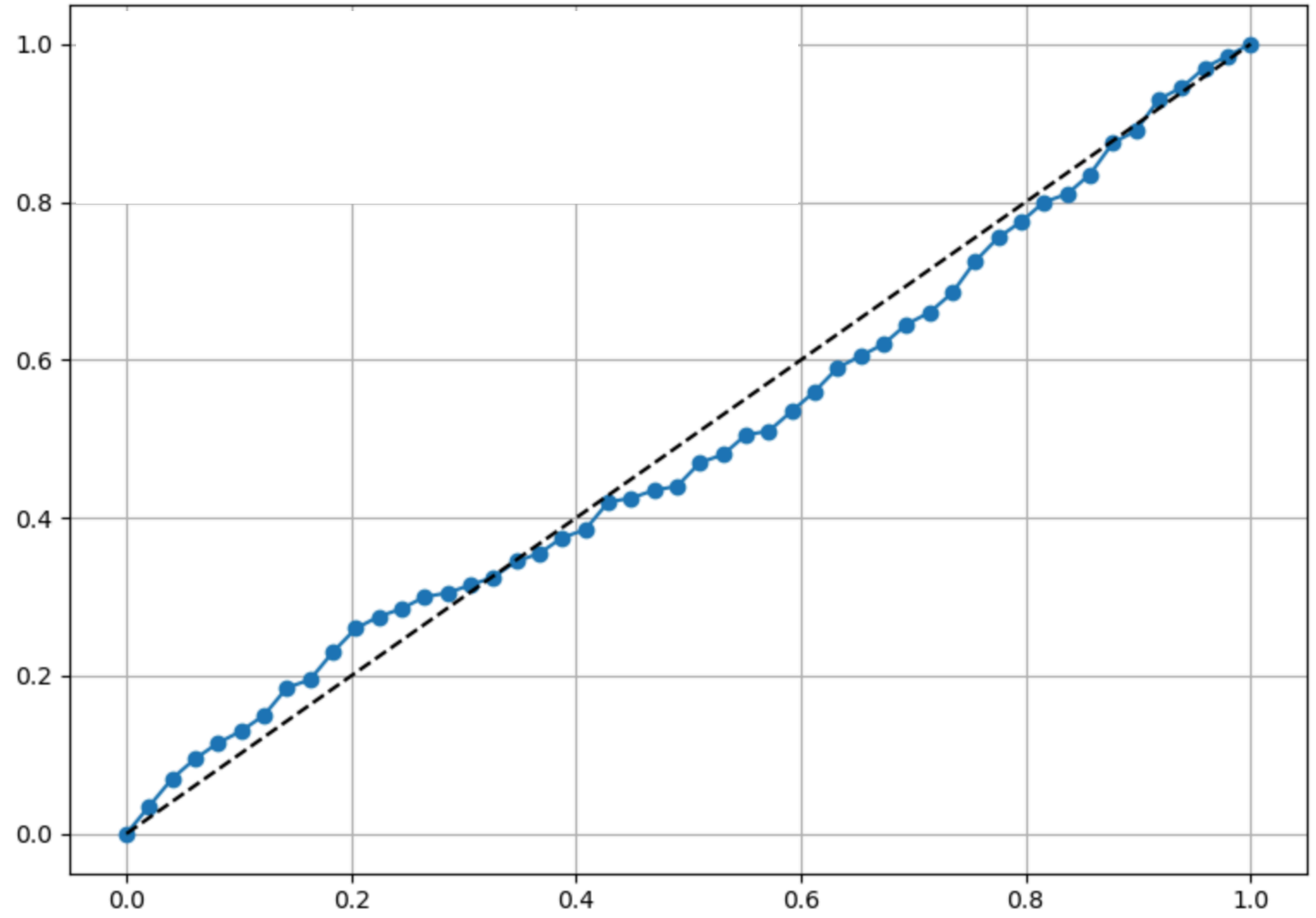
$$p(\theta \mid x) \propto p(\theta_1 \mid x) \cdot p(\theta_2 \mid x, \theta_1) \cdot p(\theta_3 \mid x, \theta_1, \theta_2)$$

Preliminary results: standardisation





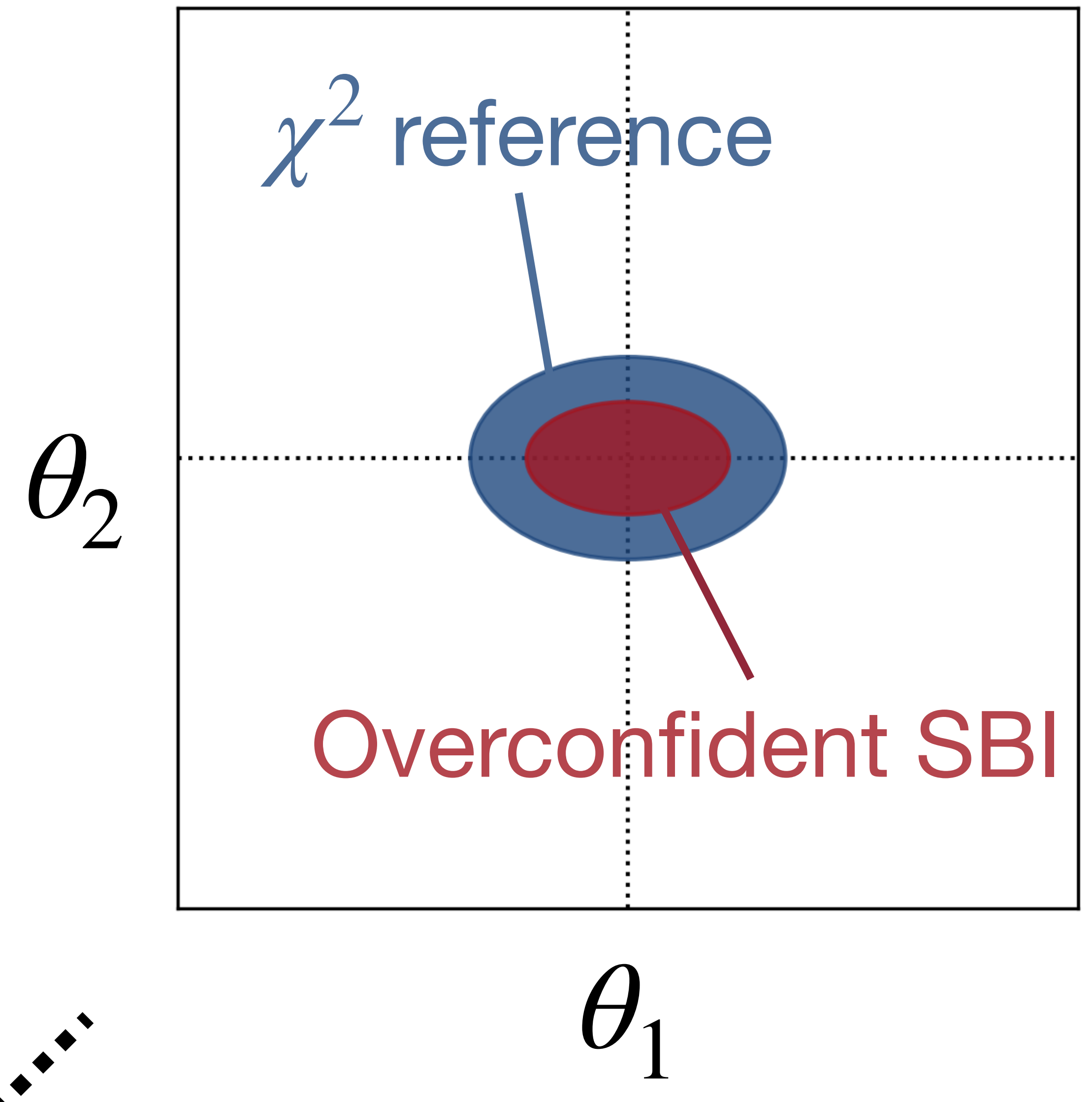
Coverage test [Lemos et al. (2023)]



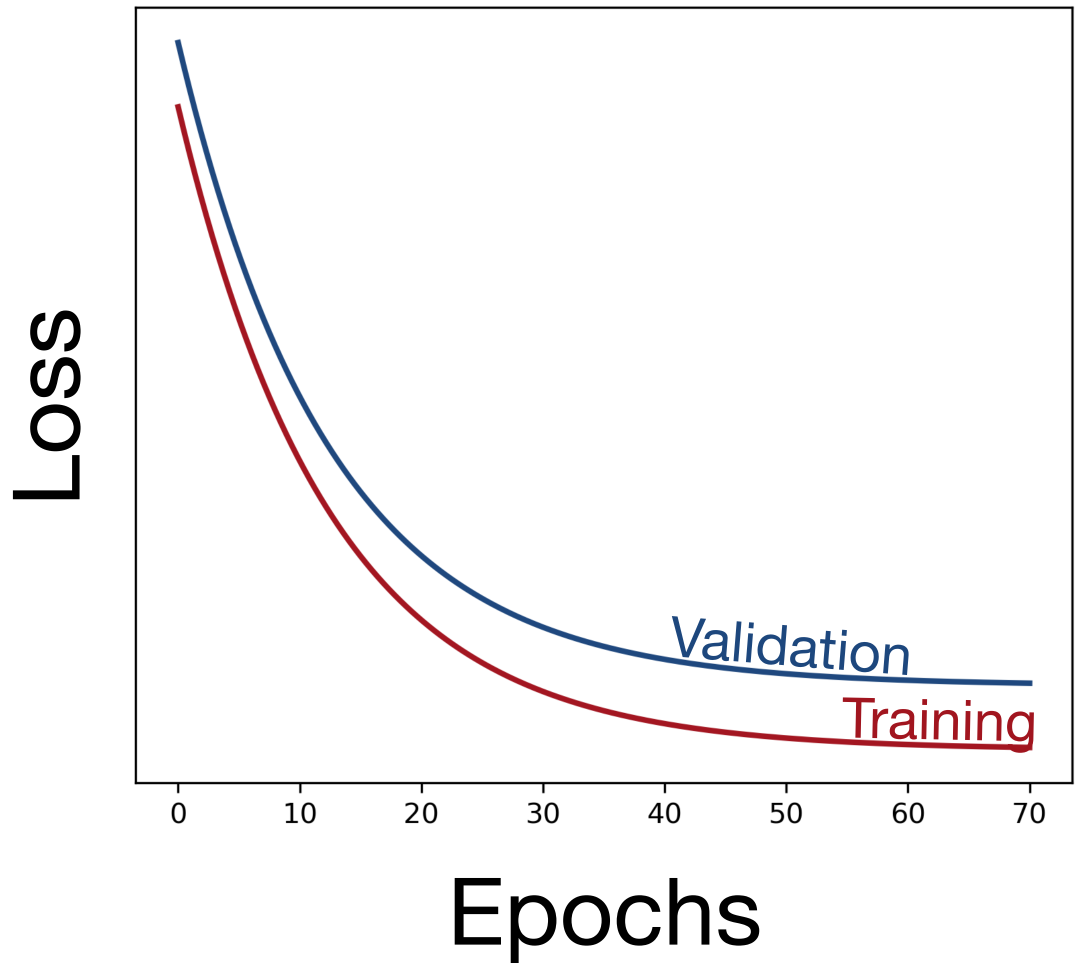
Coverage (from sampling)

Credibility (from posterior size)

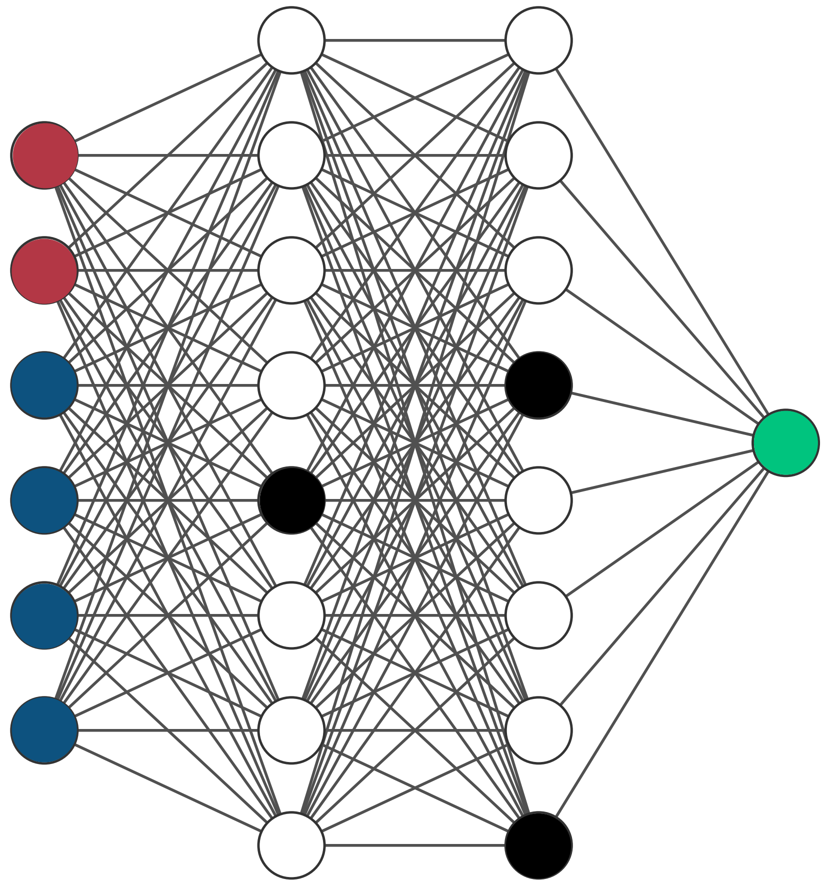
Regularisation



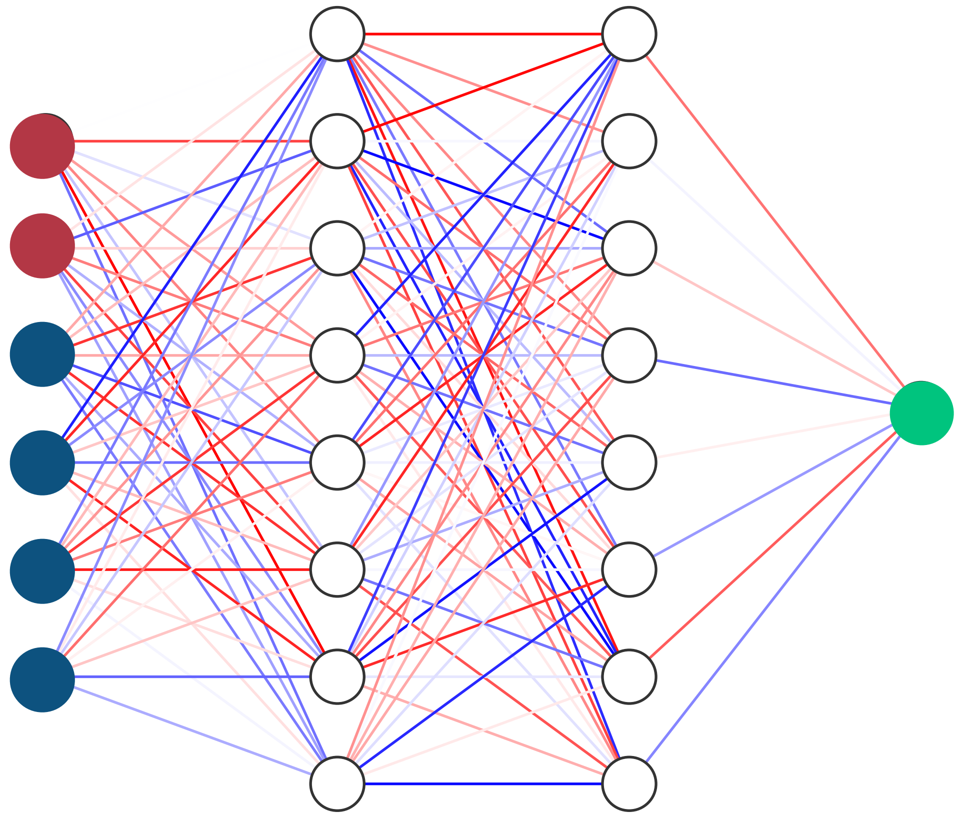
Early stopping



Dropout

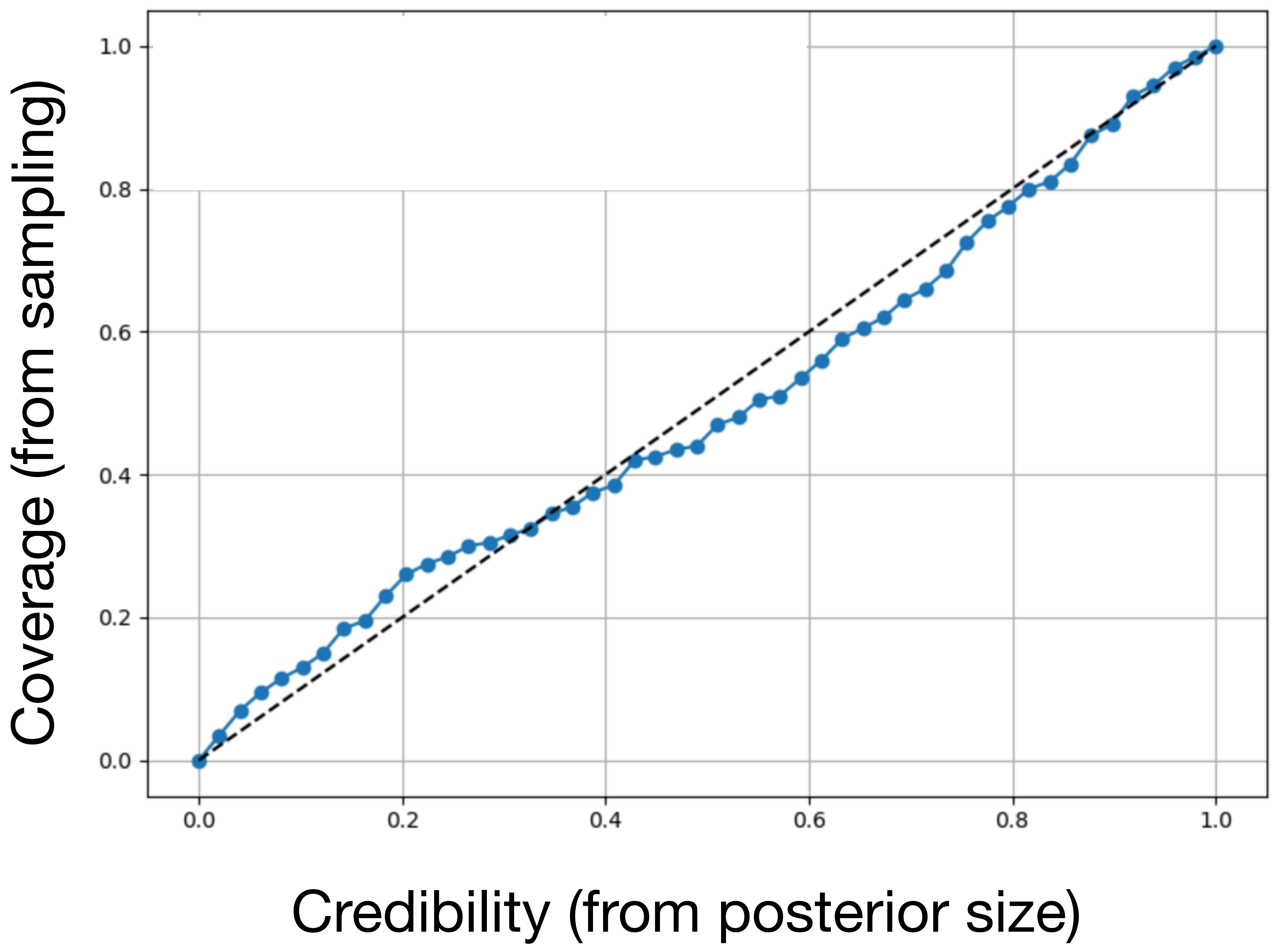


Weight decay

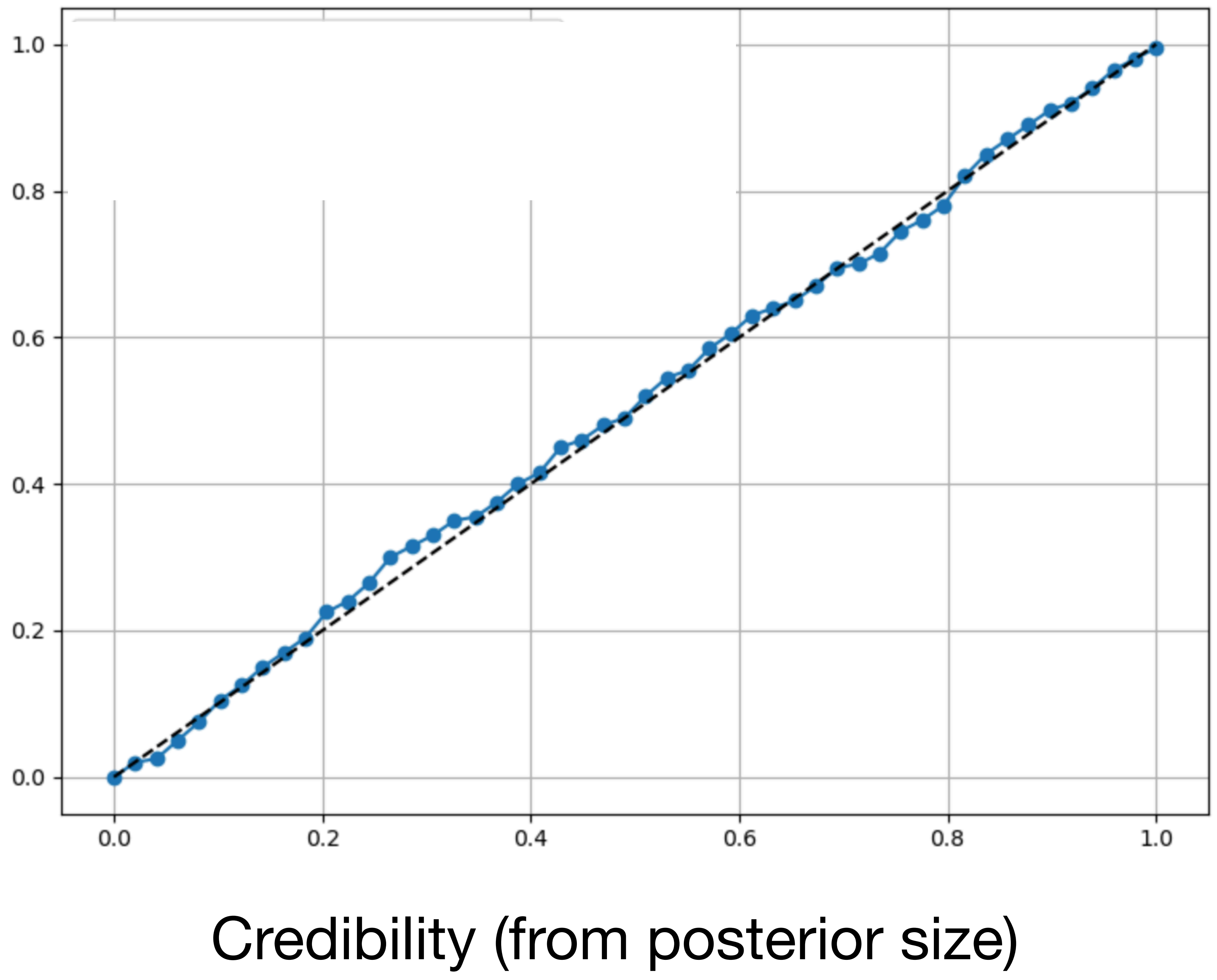


Coverage test

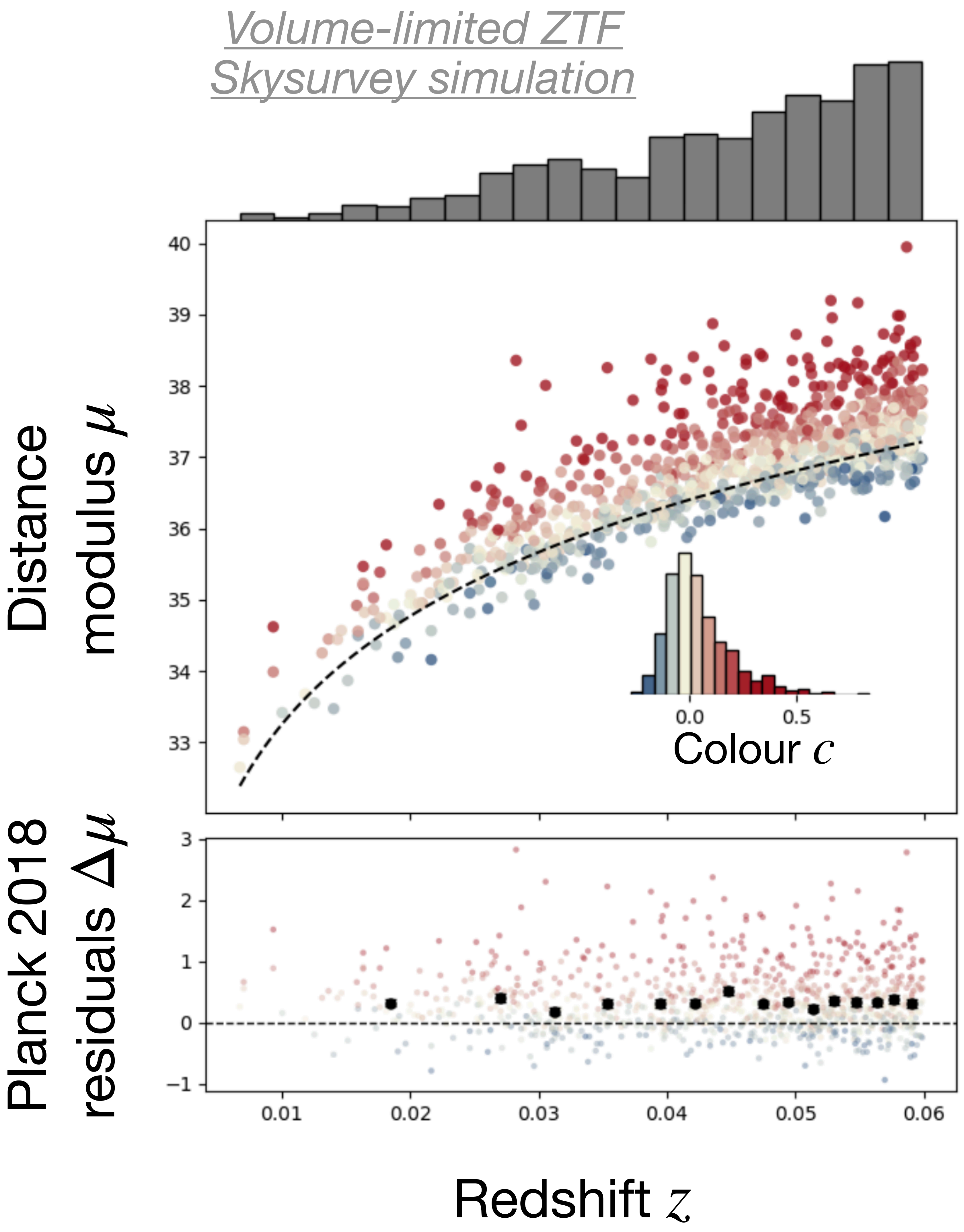
Before regularisation

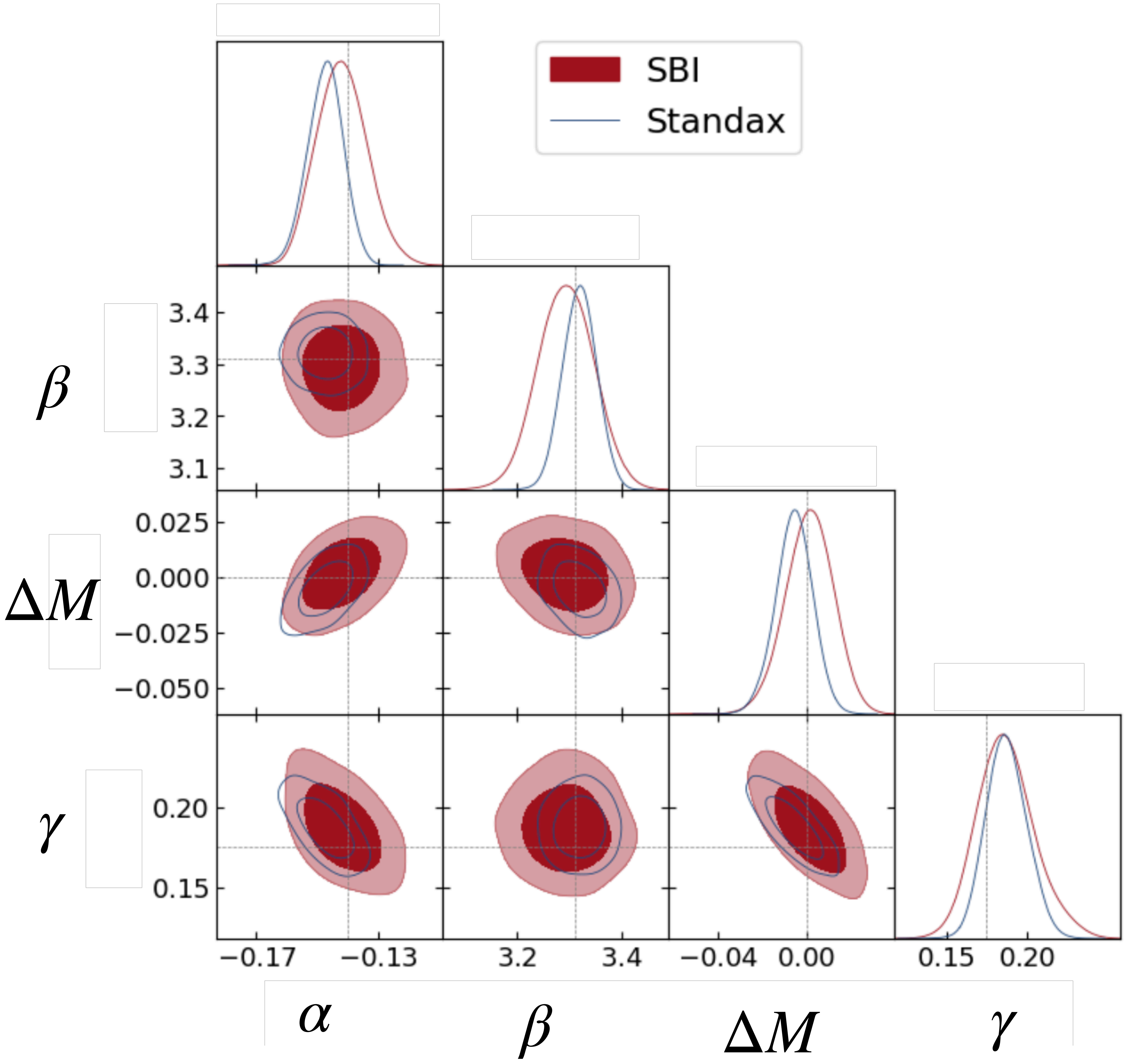


After regularisation

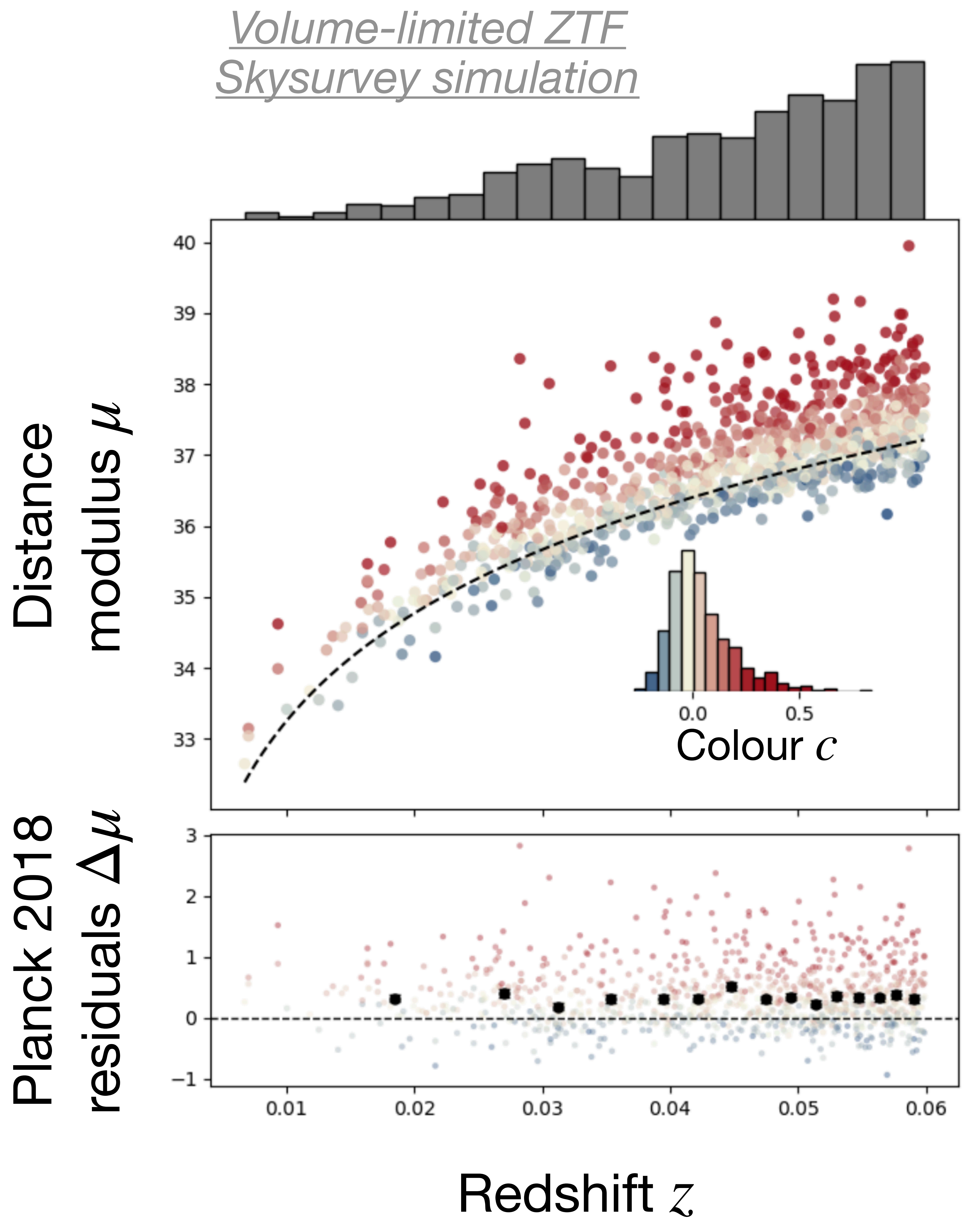


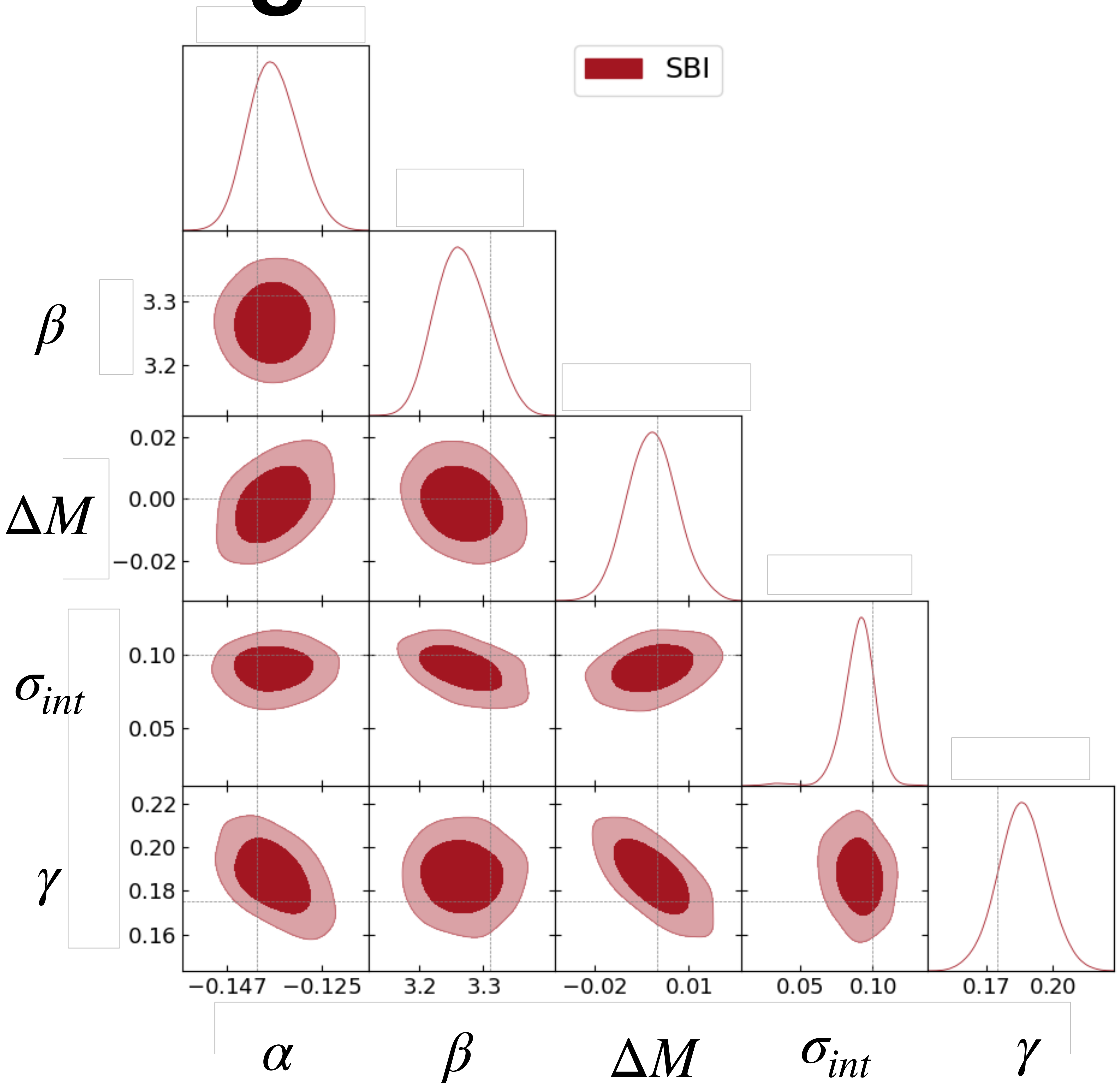
Preliminary results: standardisation



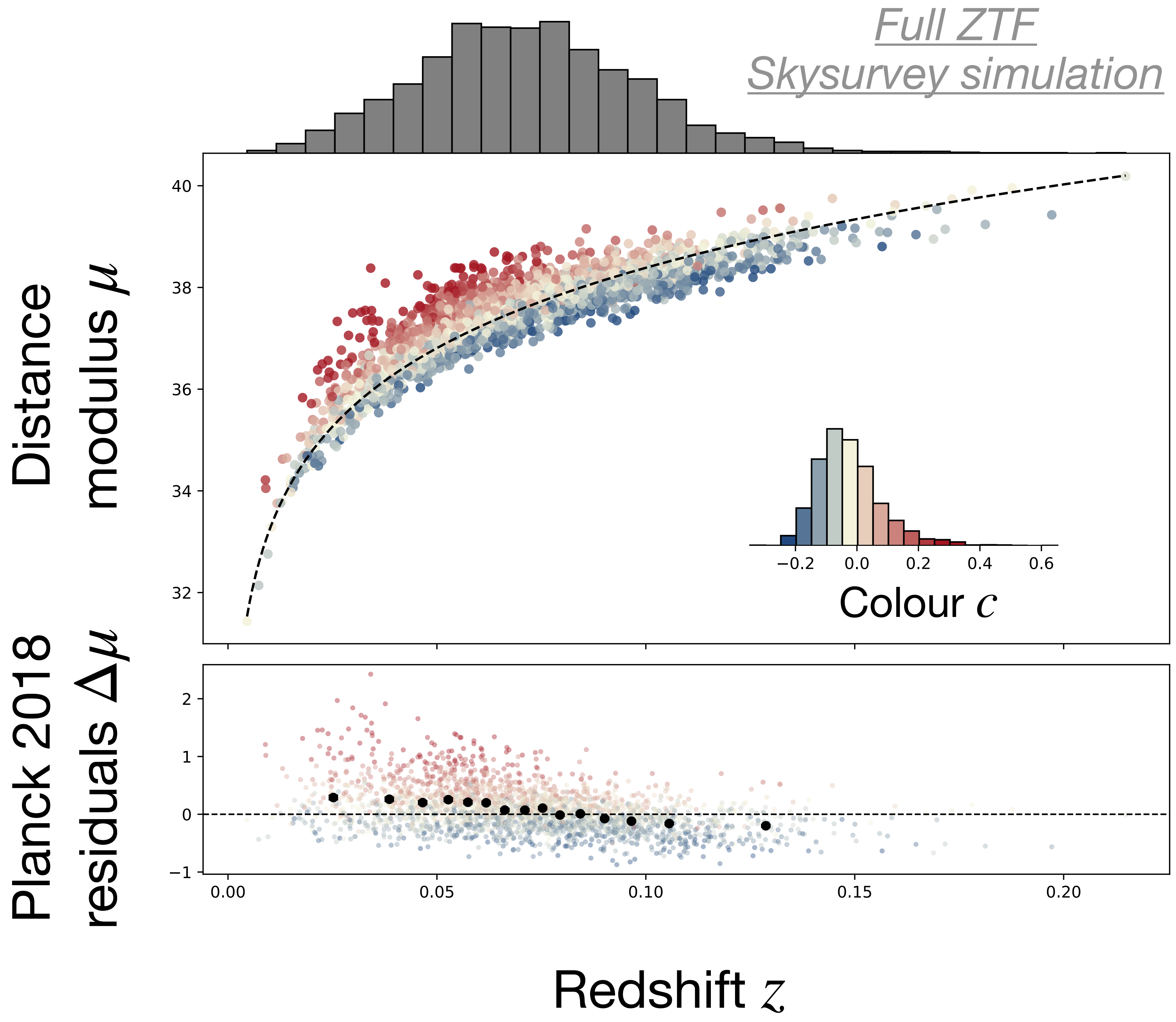


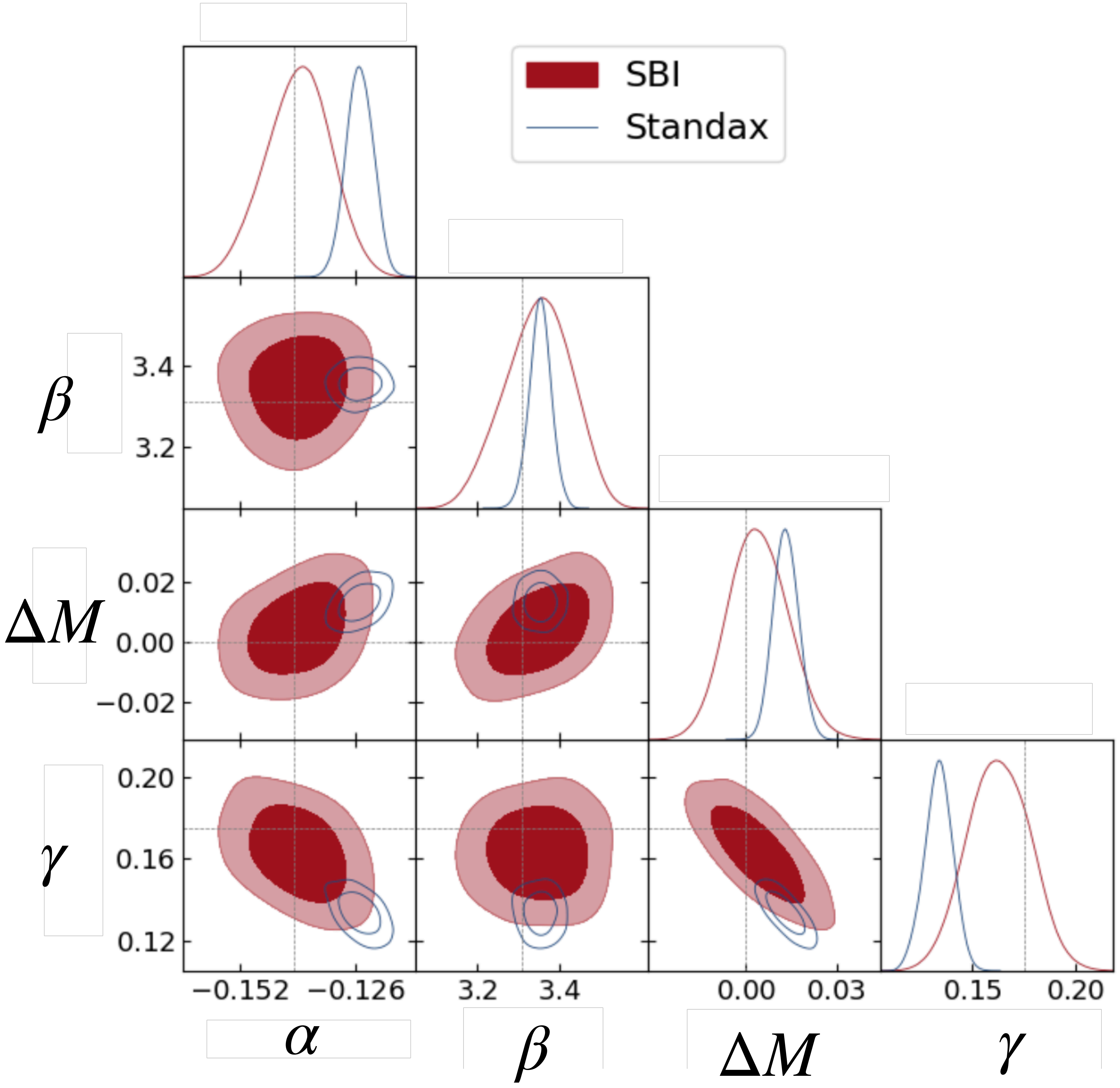
Preliminary results: fitting intrinsic scatter



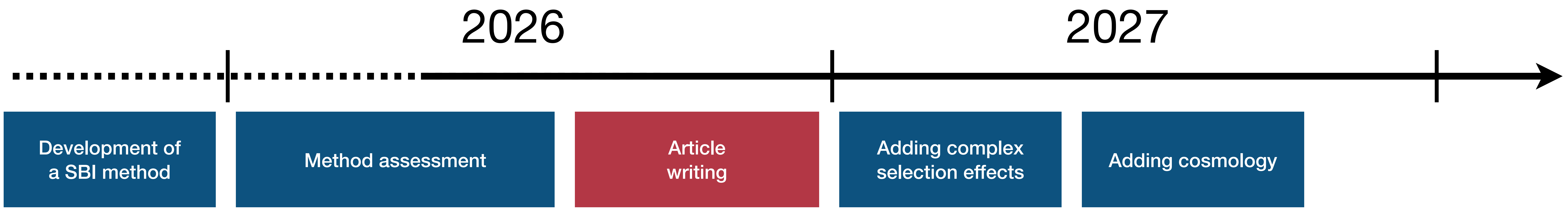


Preliminary results: adding malmquist bias





Conclusions and perspectives



Key Project : DESC SNIa pipeline validation

SLACK : #desc-snia-cosmology-collab

Annexes

Likelihood proof

We have a classification problem with two classes :

$y = 1$: the parameters in input correspond to the data also in input, sampled from $p(x, \theta)$

$y = 0$: the parameters in input do not correspond to the data in input sampled from $p(x)p(\theta)$

The neural network outputs $p(y = 1 \mid x, \theta)$, the probability that the sample corresponds to the parameters in input, and if the priors on the two classes are equal $p(y = 0) = p(y = 1) = 1/2$ we can write the Bayes optimal classifier of binary cross-entropy as :

$$d(x, \theta) = p(y = 1 \mid x, \theta) = \frac{p(x, \theta)}{p(x, \theta) + p(x)p(\theta)}$$

By taking $\dfrac{d(x, \theta)}{1 - d(x, \theta)} = \dfrac{\frac{p(x, \theta)}{p(x, \theta) + p(x)p(\theta)}}{1 - \frac{p(x, \theta)}{p(x, \theta) + p(x)p(\theta)}}$ we get $\dfrac{\frac{p(x, \theta)}{p(x, \theta) + p(x)p(\theta)}}{\frac{p(x)p(\theta)}{p(x, \theta) + p(x)p(\theta)}} = \dfrac{p(x, \theta)}{p(x)p(\theta)} = \dfrac{p(x \mid \theta)}{p(x)} = r(x \mid \theta)$

Joint posterior proof

$$\frac{p(x \mid \theta)}{p(x)} = \frac{p(\theta \mid x)}{p(\theta)} \longrightarrow \frac{p(\theta \mid x)}{p(\theta)} = \frac{p(\theta_1 \mid x)}{p(\theta_1)} \cdot \frac{p(\theta_2 \mid x, \theta_1)}{p(\theta_2)} \cdot \frac{p(\theta_3 \mid x, \theta_1, \theta_2)}{p(\theta_3)}$$