Simulation Based Inference for SNe Ia cosmology

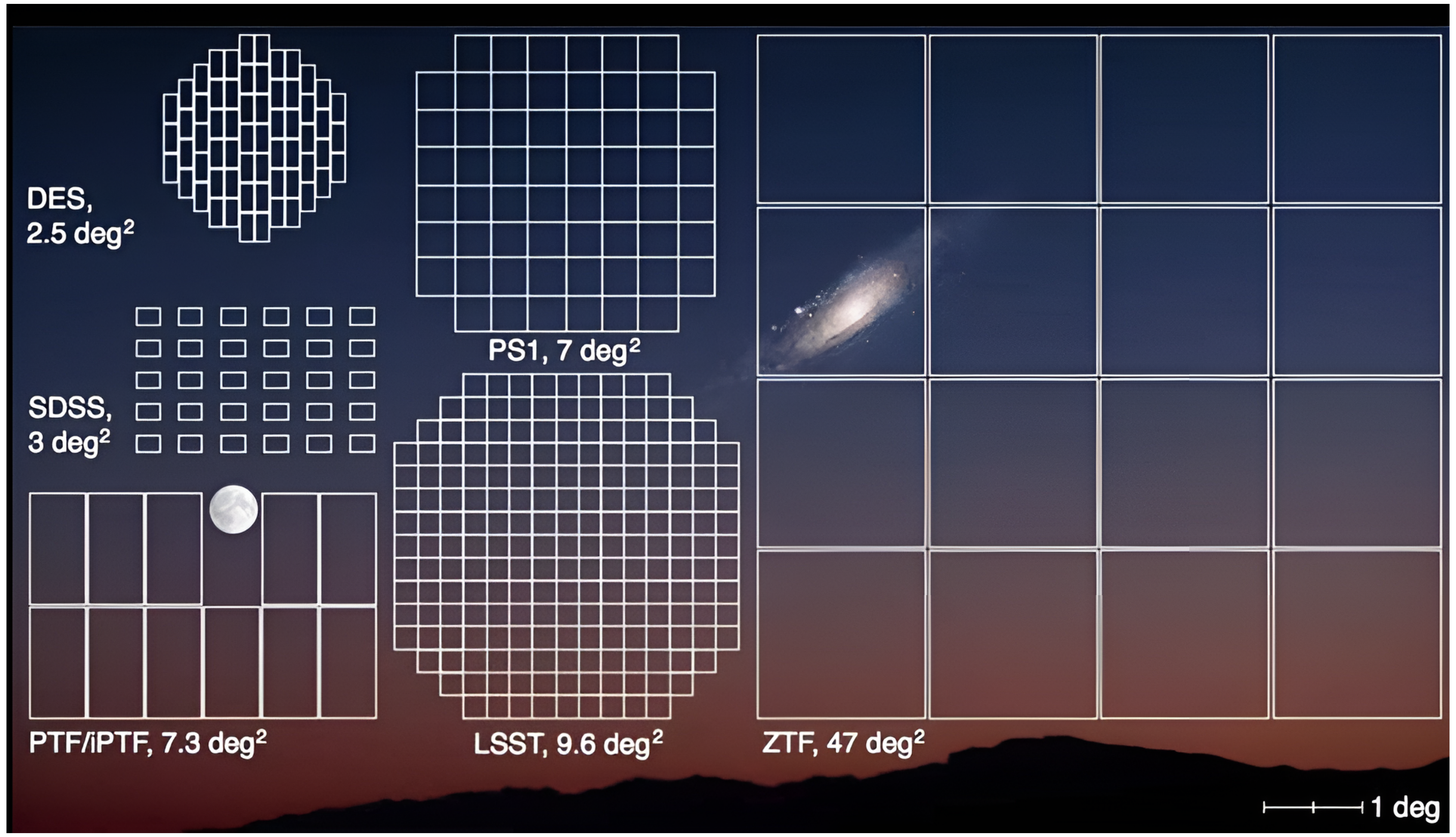
Adam Trigui

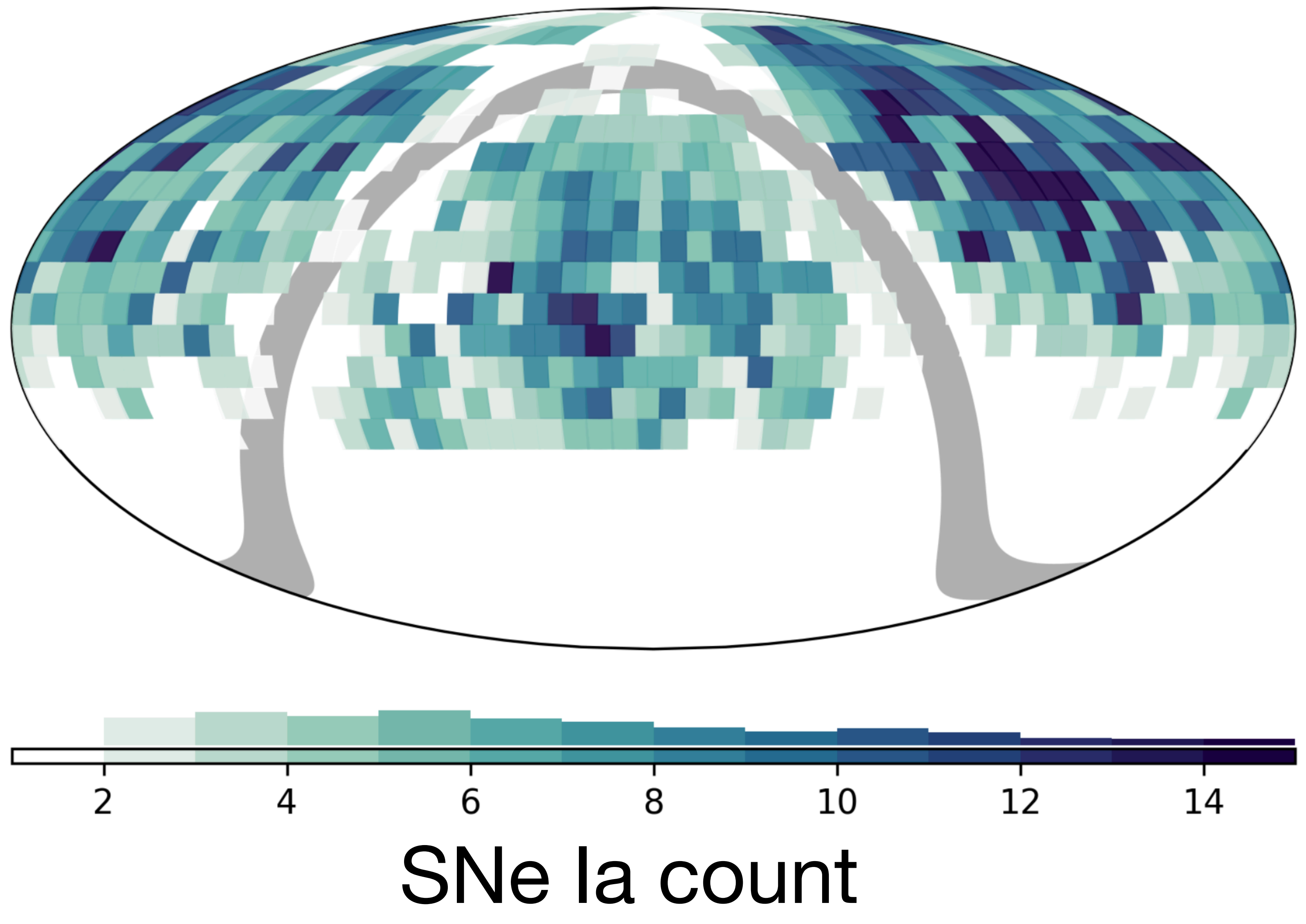
Under the supervision of Mickaël Rigault and Florian Ruppin

GDR CoPhy 2026, Clermont-Ferrand

ZTF survey

Mount Palomar, California

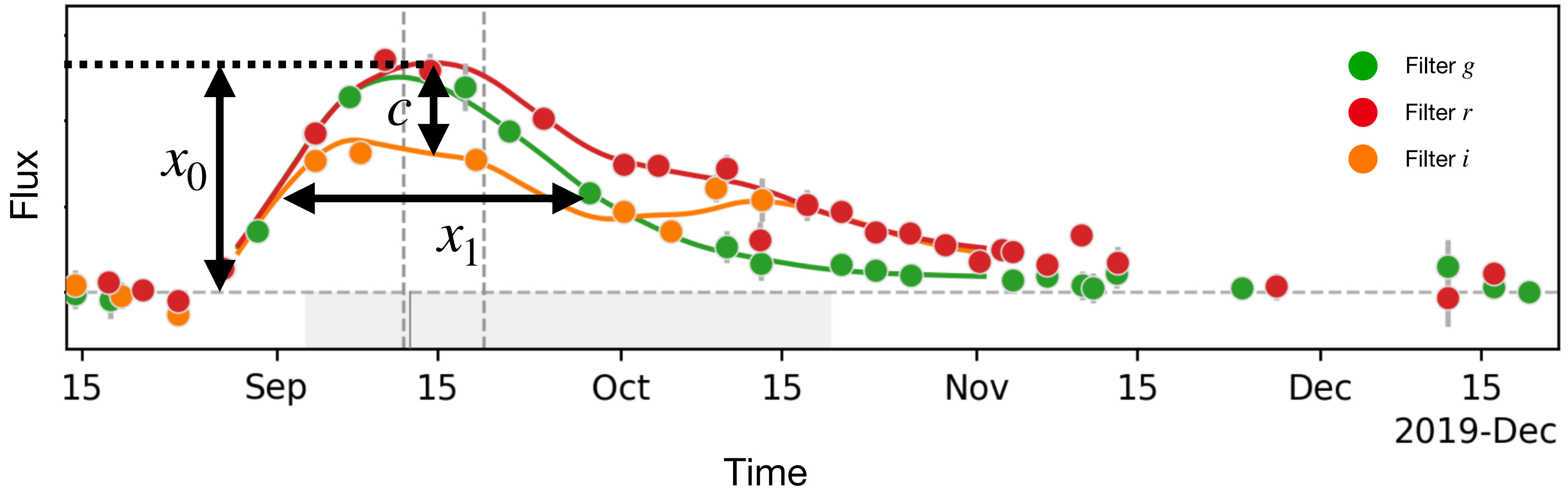




Total > 3000 (DR2)

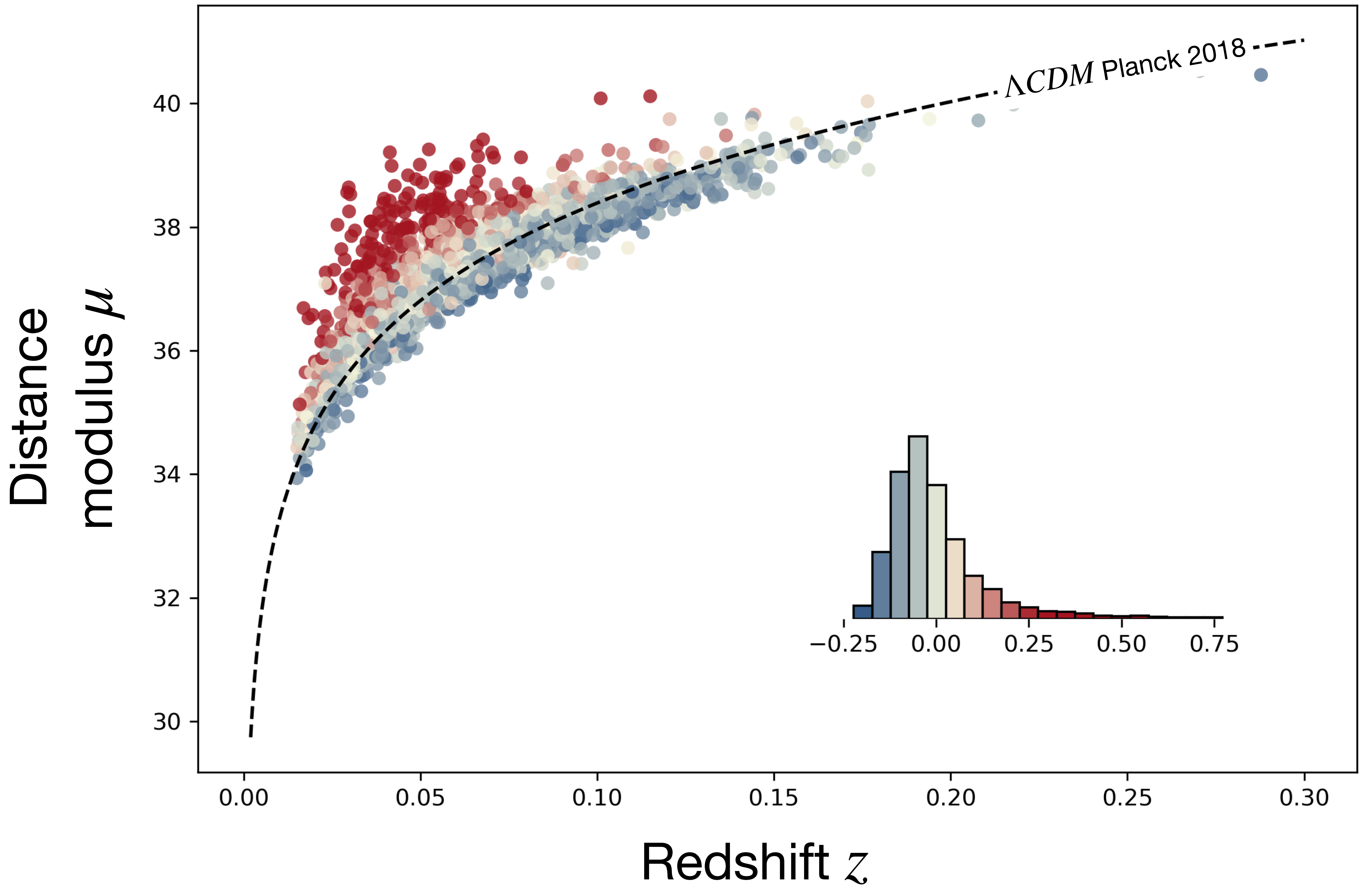
Type Ia Supernovae

ZTF DR2 : ZTF19abuonnx



- Flux x_0
  - Distance modulus $\mu \propto -log(x_0)$
- Stretch x_1
- Colour c

ZTF DR2 data
(Rigault et al. 2025)

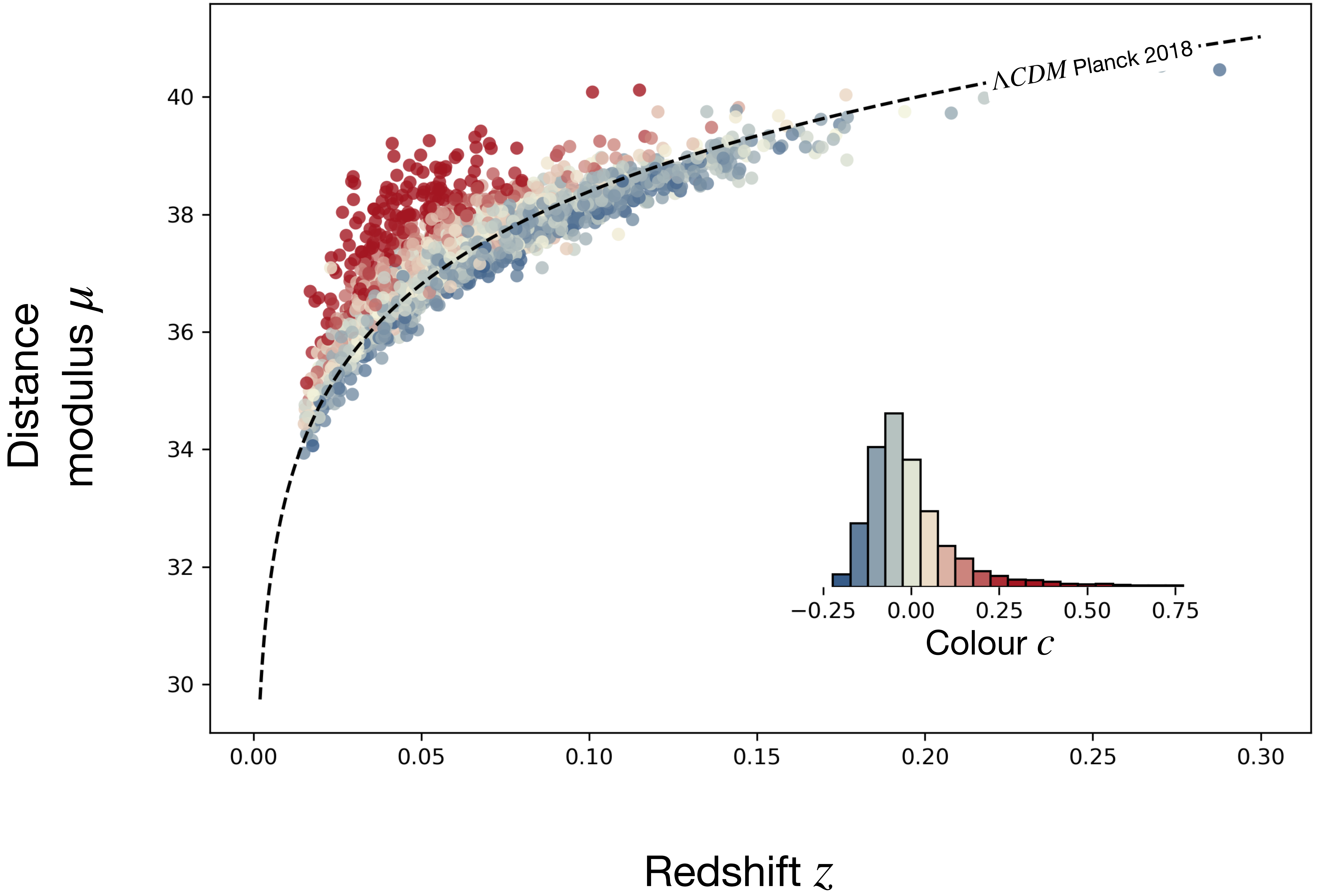


SNe Ia standardisation

Standardisation : $M = M_0 + \beta c - \alpha x_1 + \gamma p$

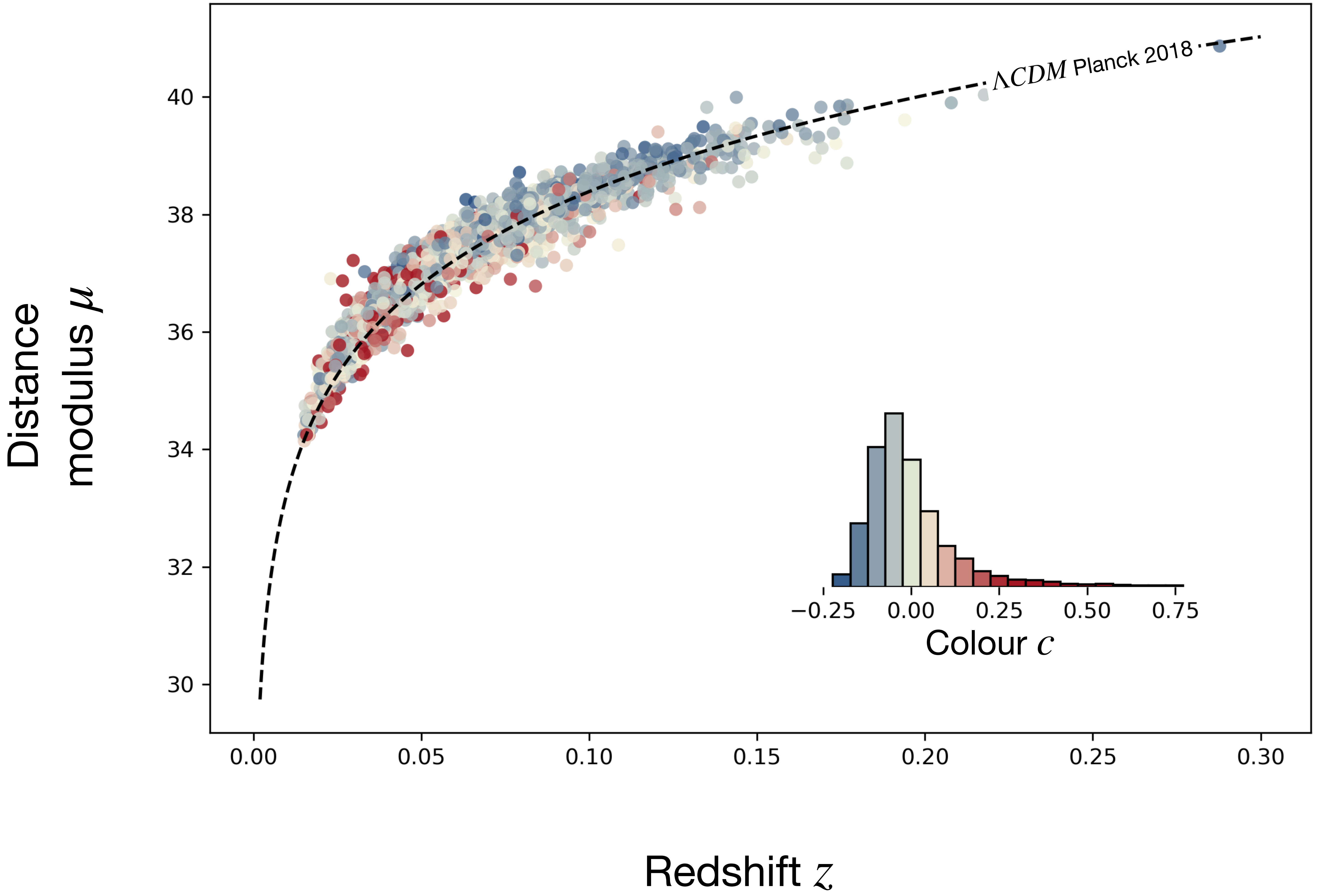
SNe Ia standardisation

Inspired by Madeleine Ginolin



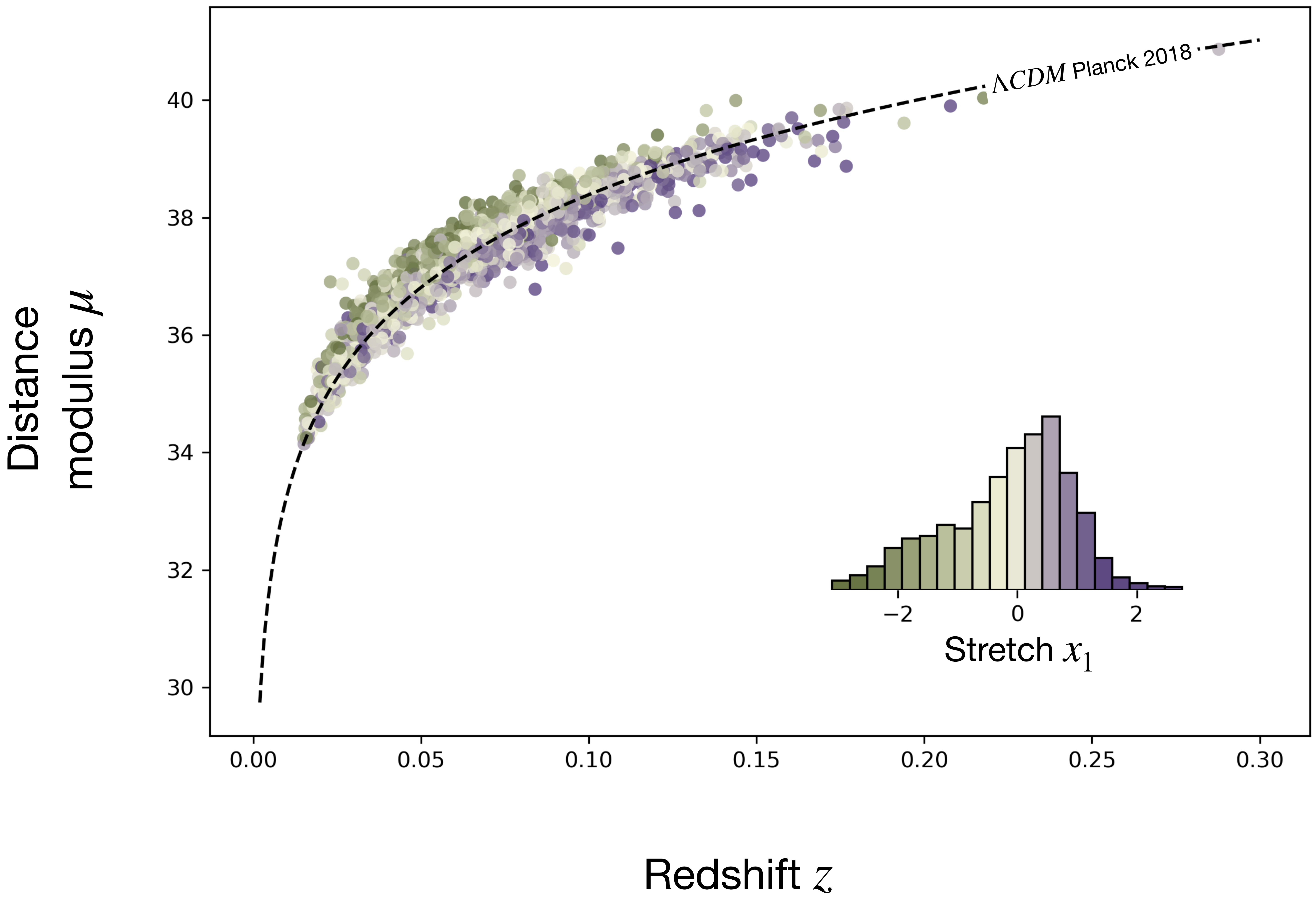
SNe Ia standardisation

Inspired by Madeleine Ginolin



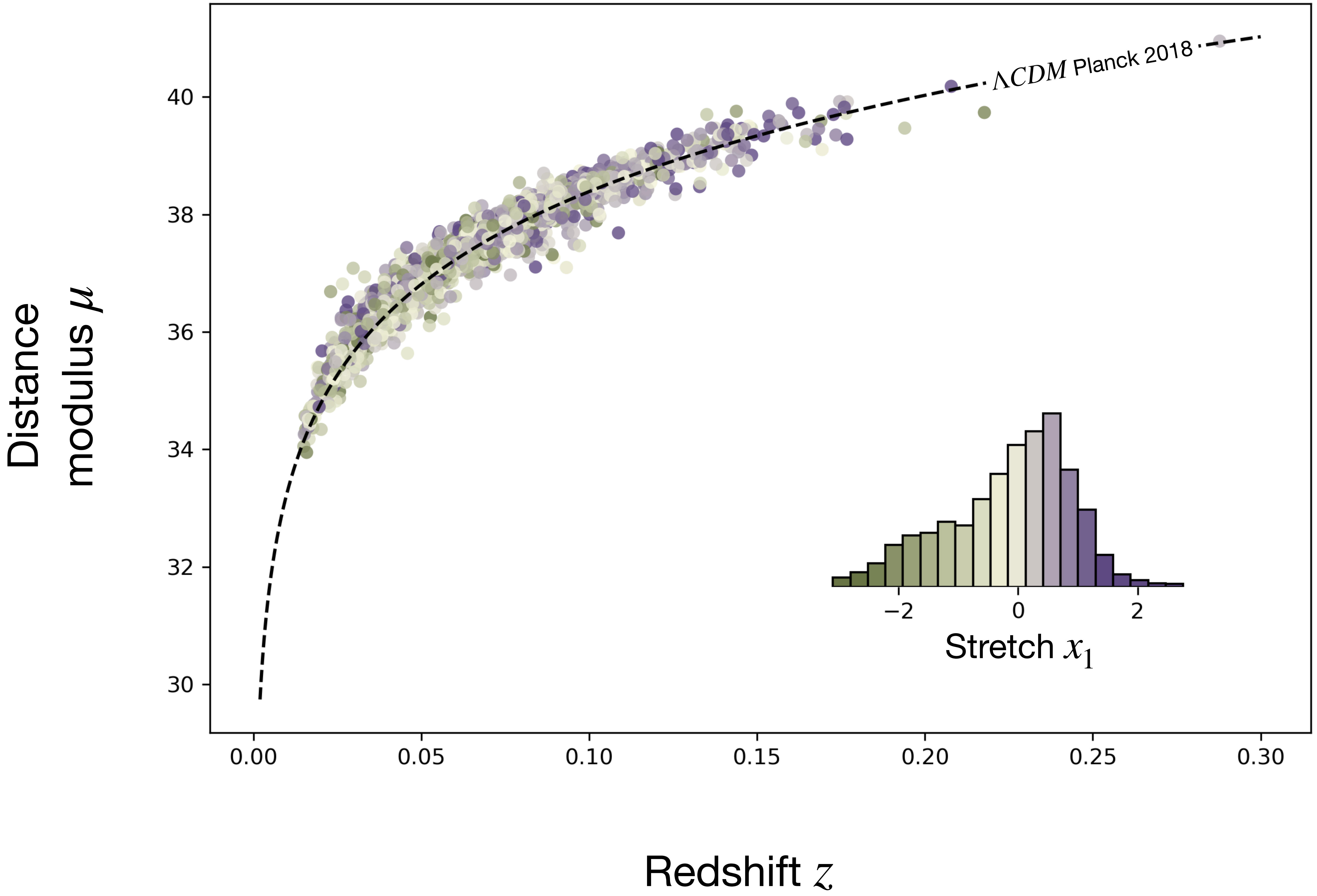
SNe Ia standardisation

Inspired by Madeleine Ginolin

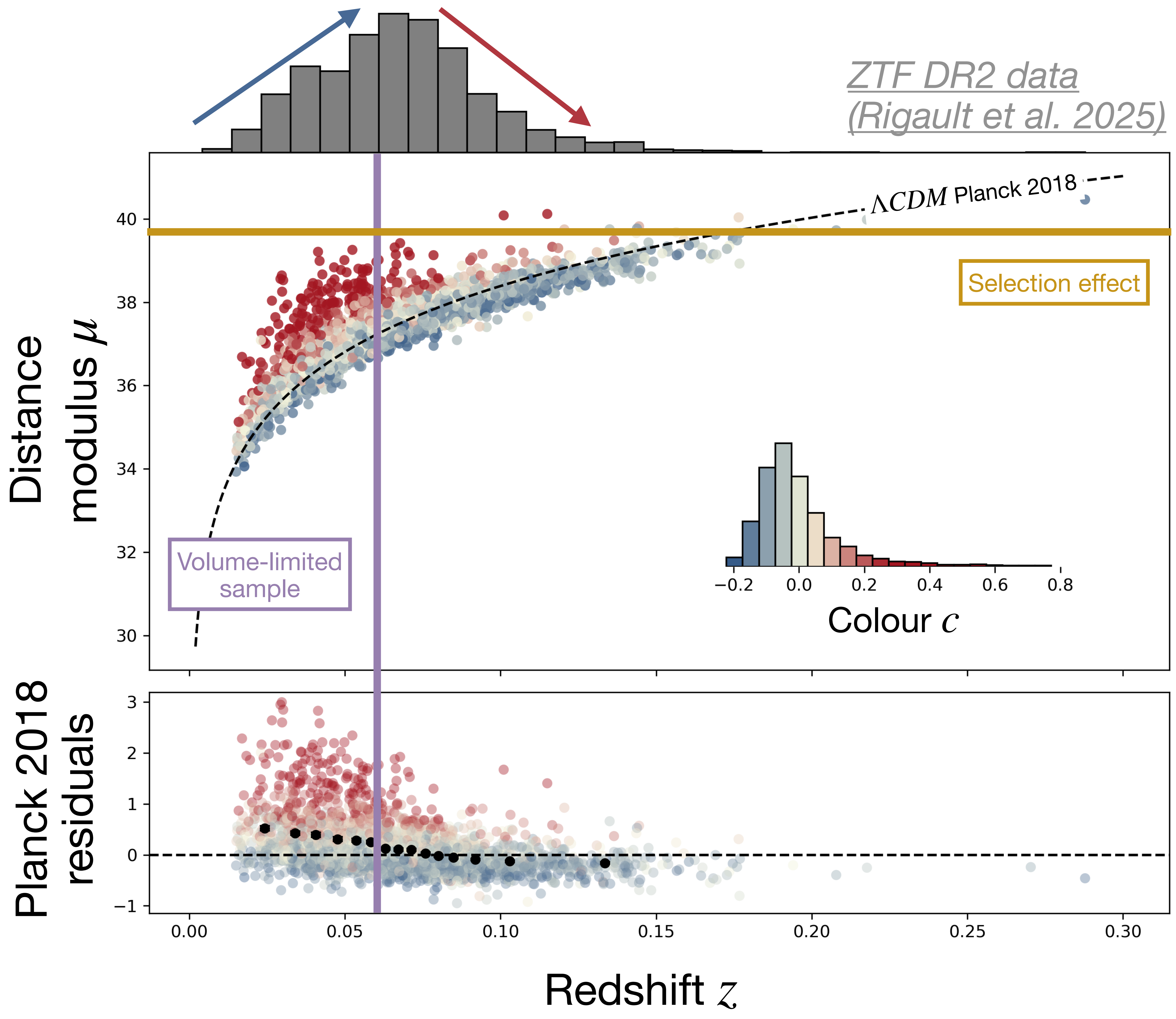


SNe Ia standardisation

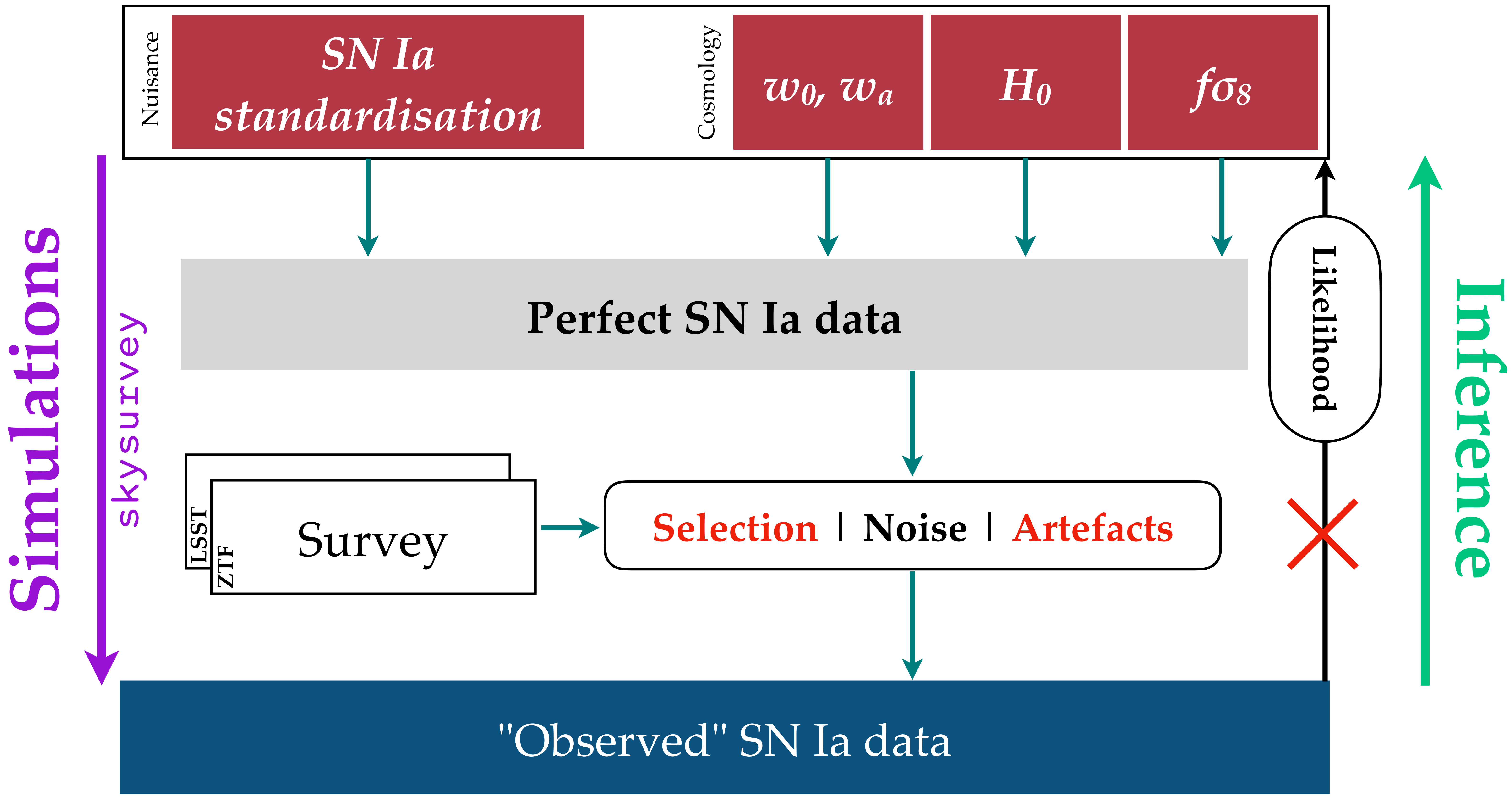
Inspired by Madeleine Ginolin



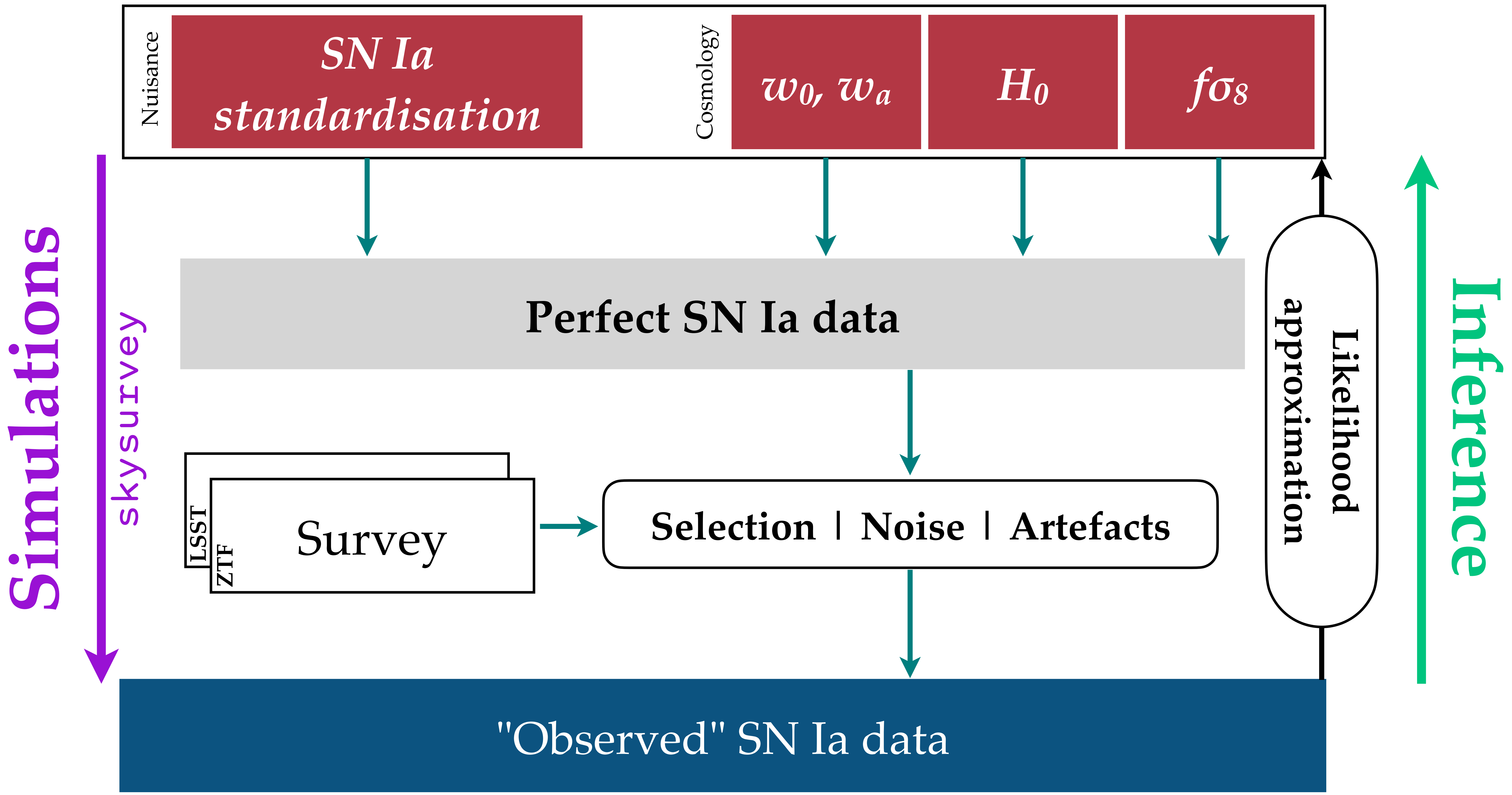
Selection effects



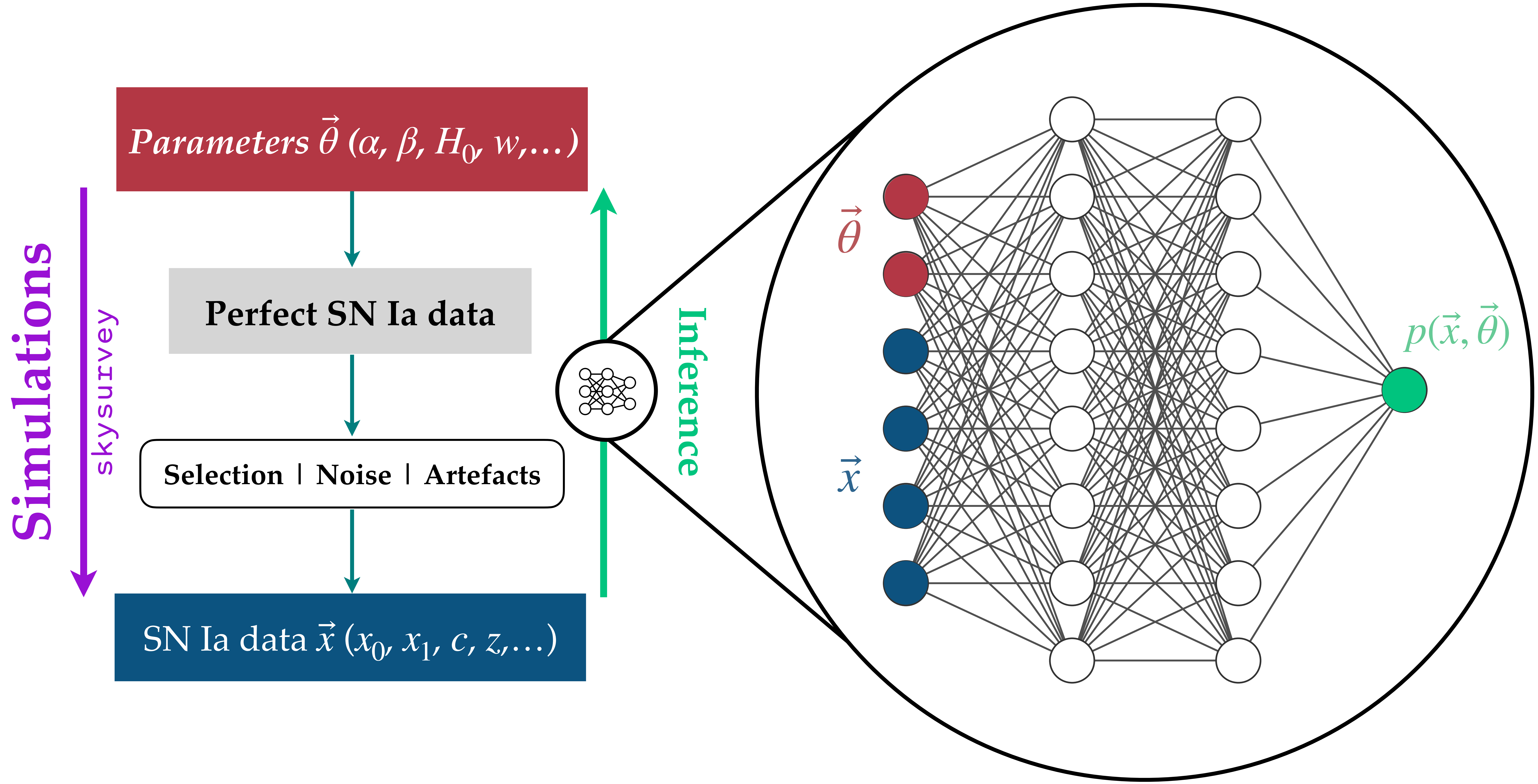
Simulation/inference asymmetry



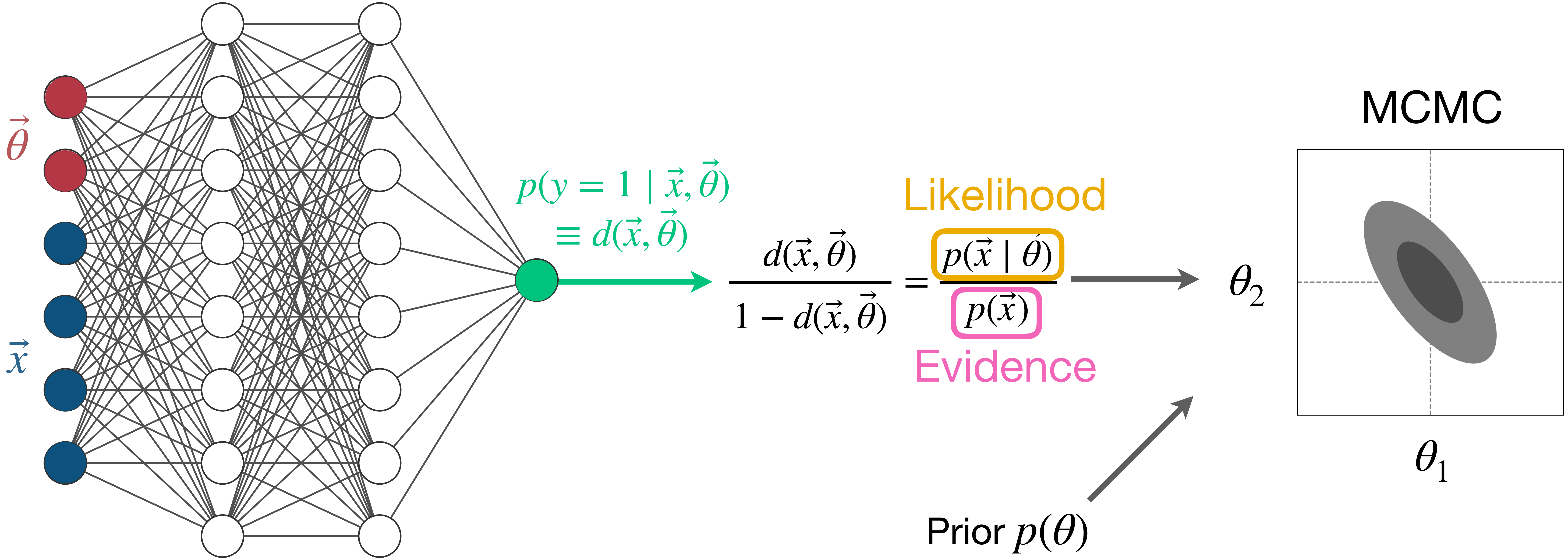
Simulation based inference (SBI)



Neural Ratio Estimator



Neural Ratio Estimator

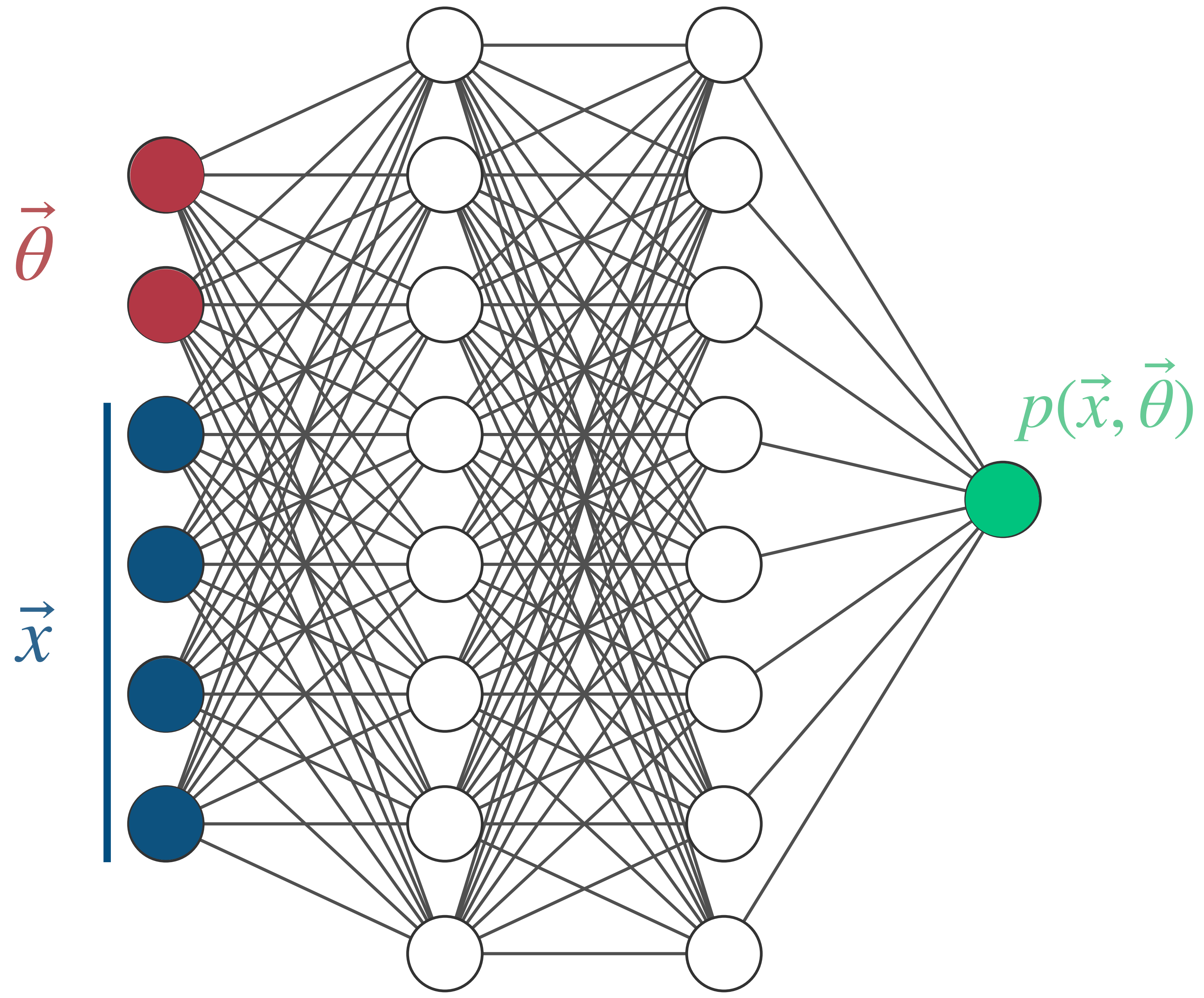


Neural Ratio Estimator

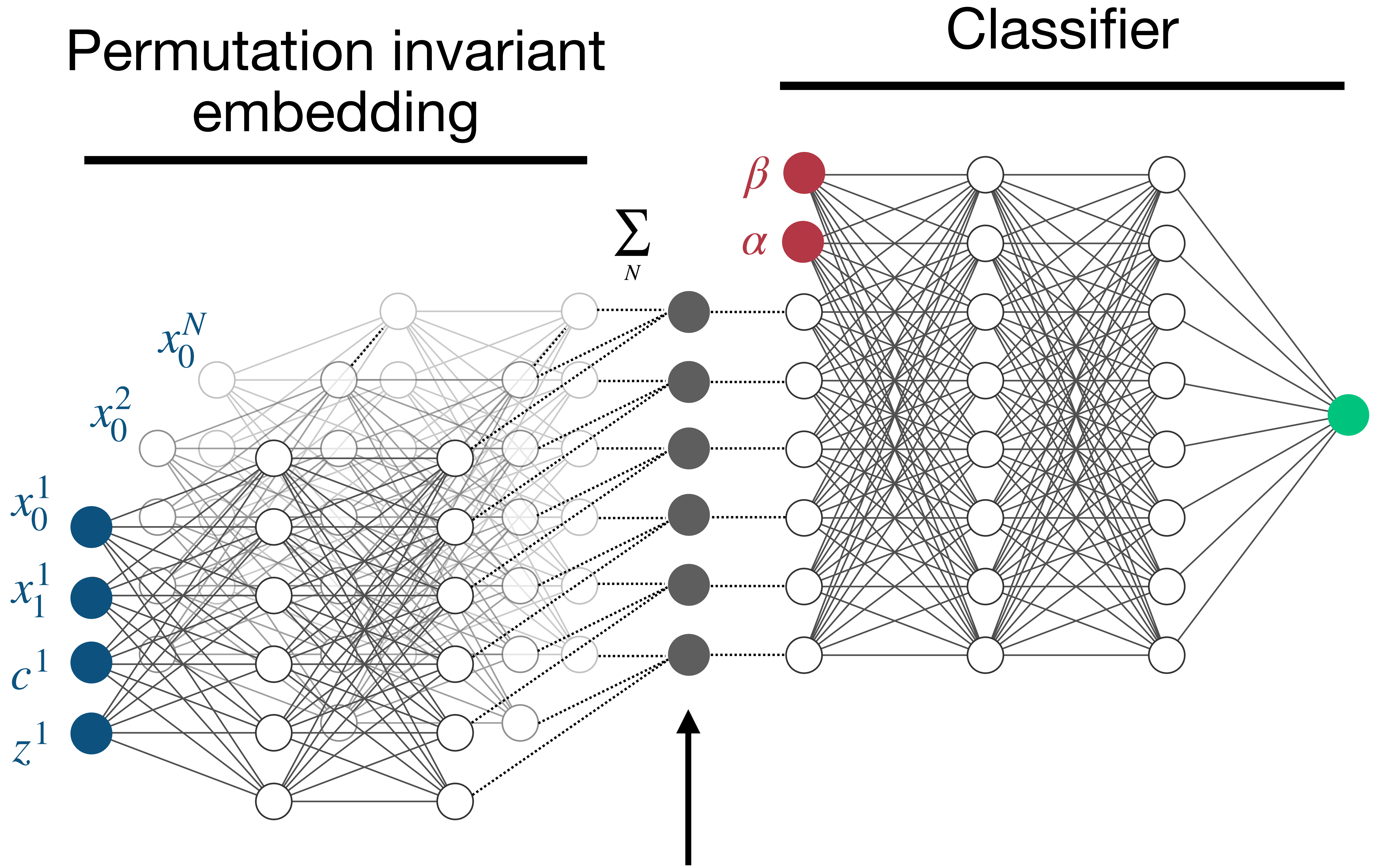
ZTF DR3 = 10 000+ SNe Ia $\times$ 10+ observables $=$ 100 000+ input data neurons $\vec{x}$

Problem :

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



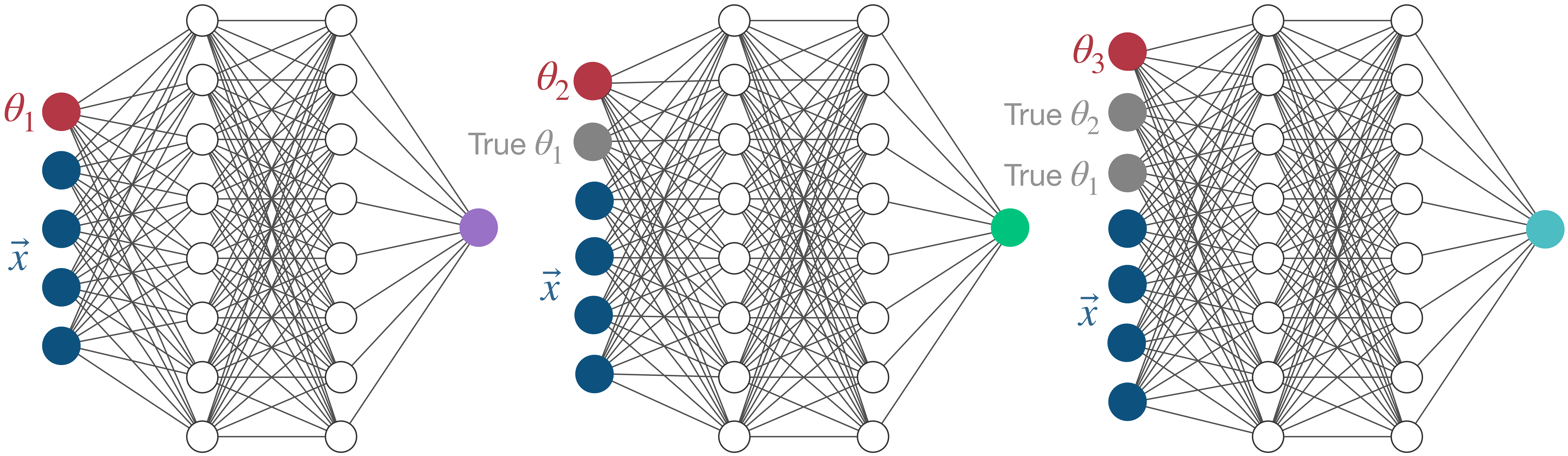
Deepset Zaheer et al. (2018)



Complex information encoded into a small size vector (100 000 → 64)

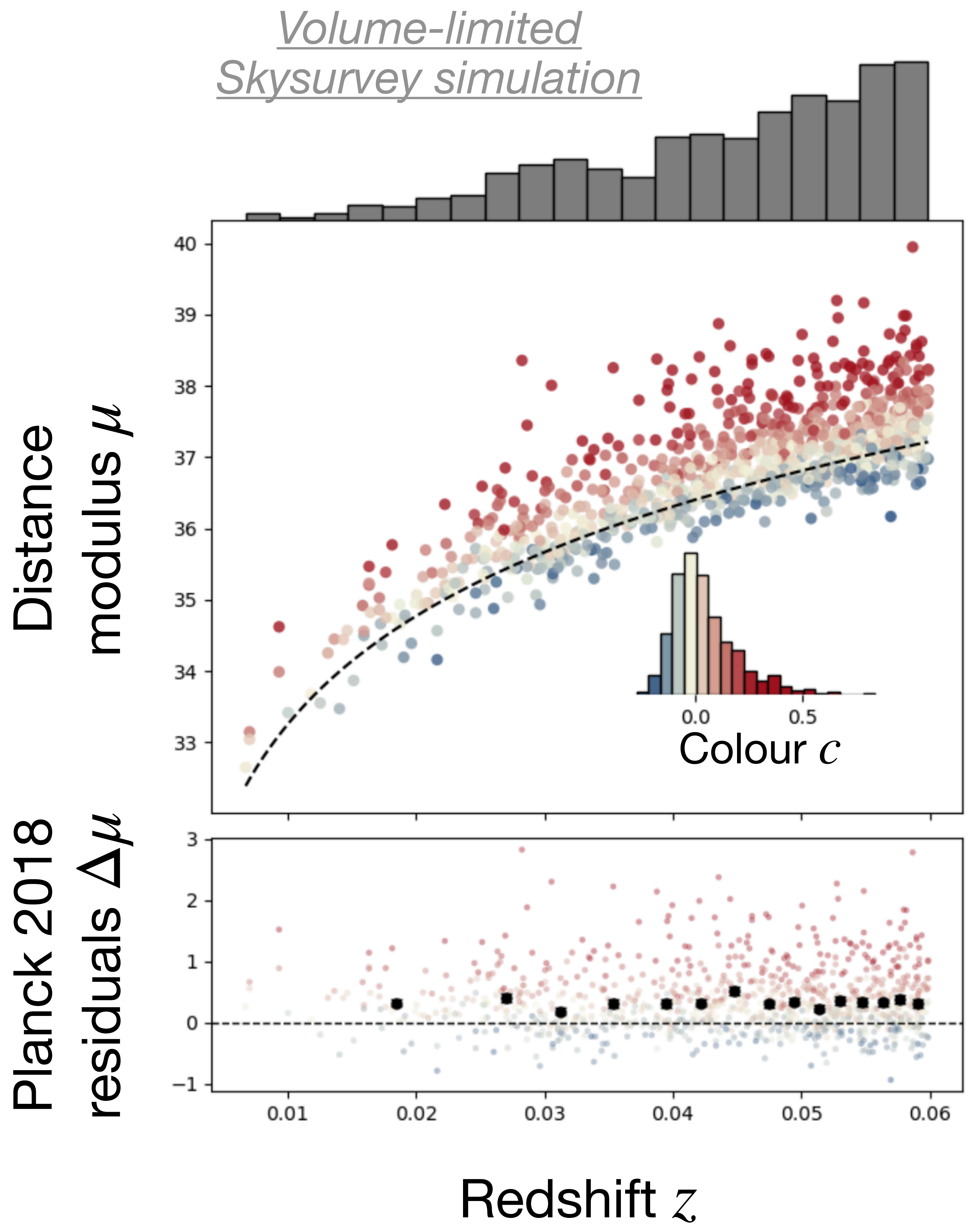
Autoregression Karchev and Trotta (2024)

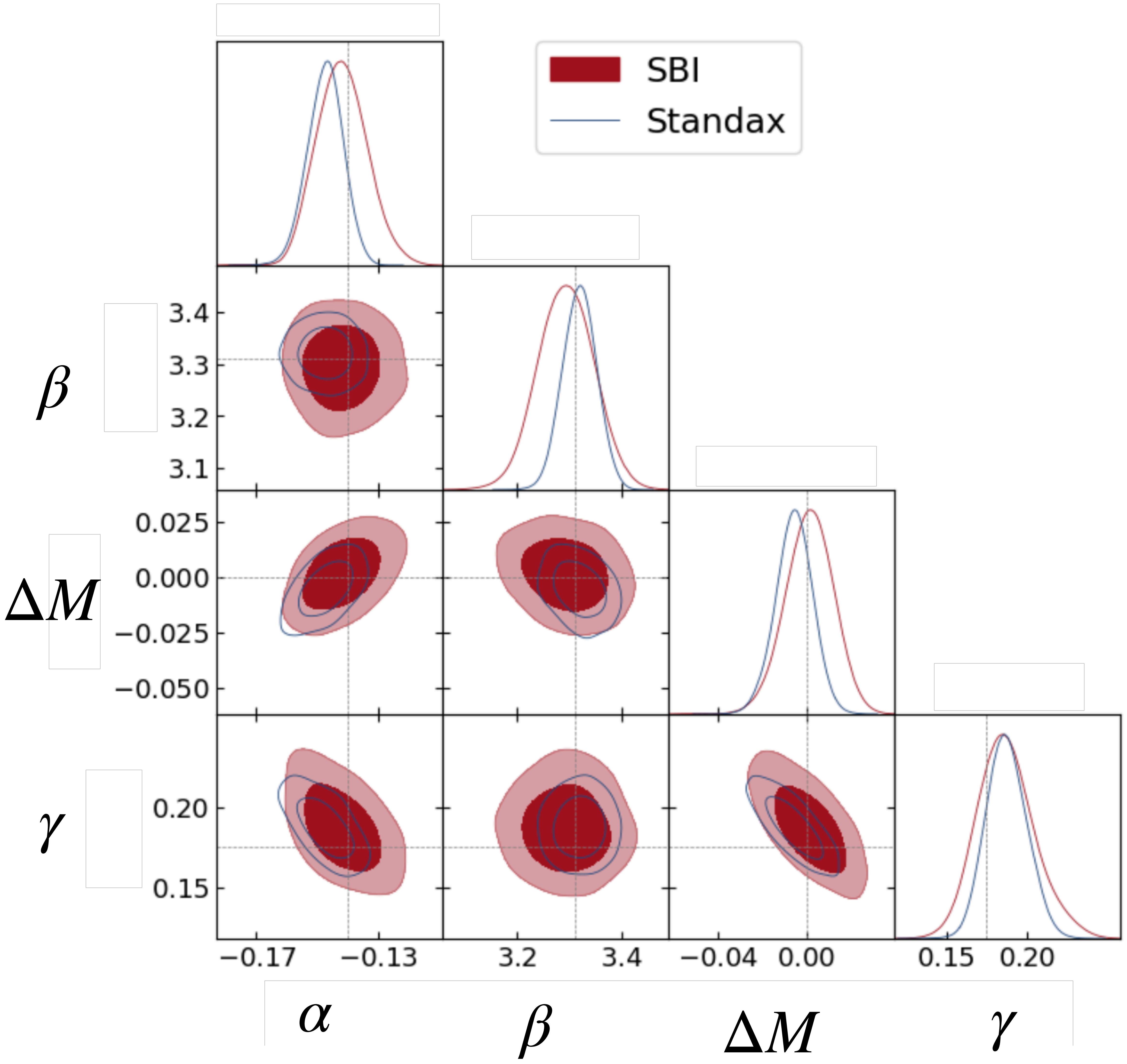
Use joint posteriors to help the network learn difficult parameters.



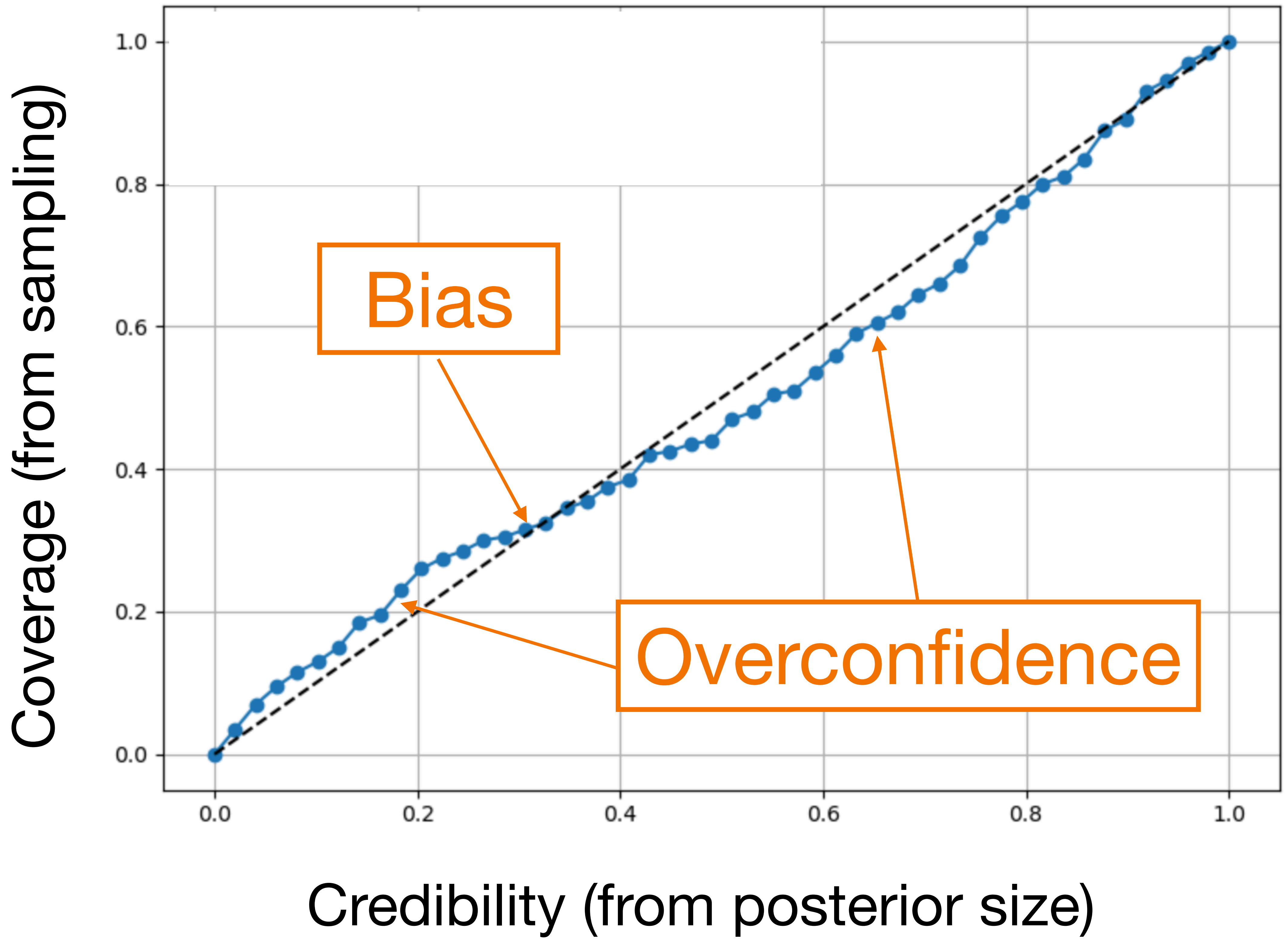
$$p(\theta \mid x) \propto p(\theta_1 \mid x) \cdot p(\theta_2 \mid x, \theta_1) \cdot p(\theta_3 \mid x, \theta_1, \theta_2)$$

Preliminary results: standardisation

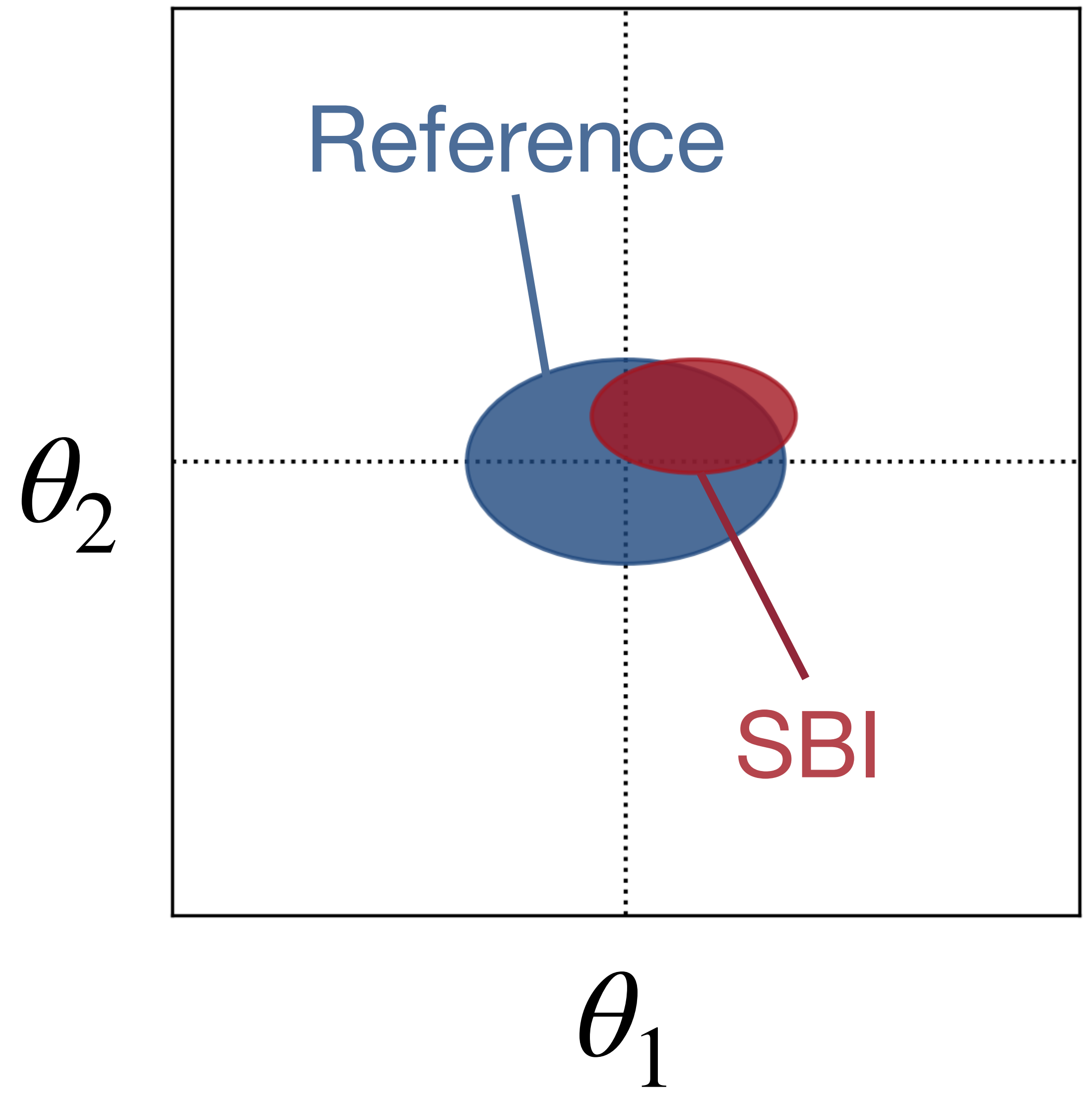




Coverage test Lemos et al. (2023)

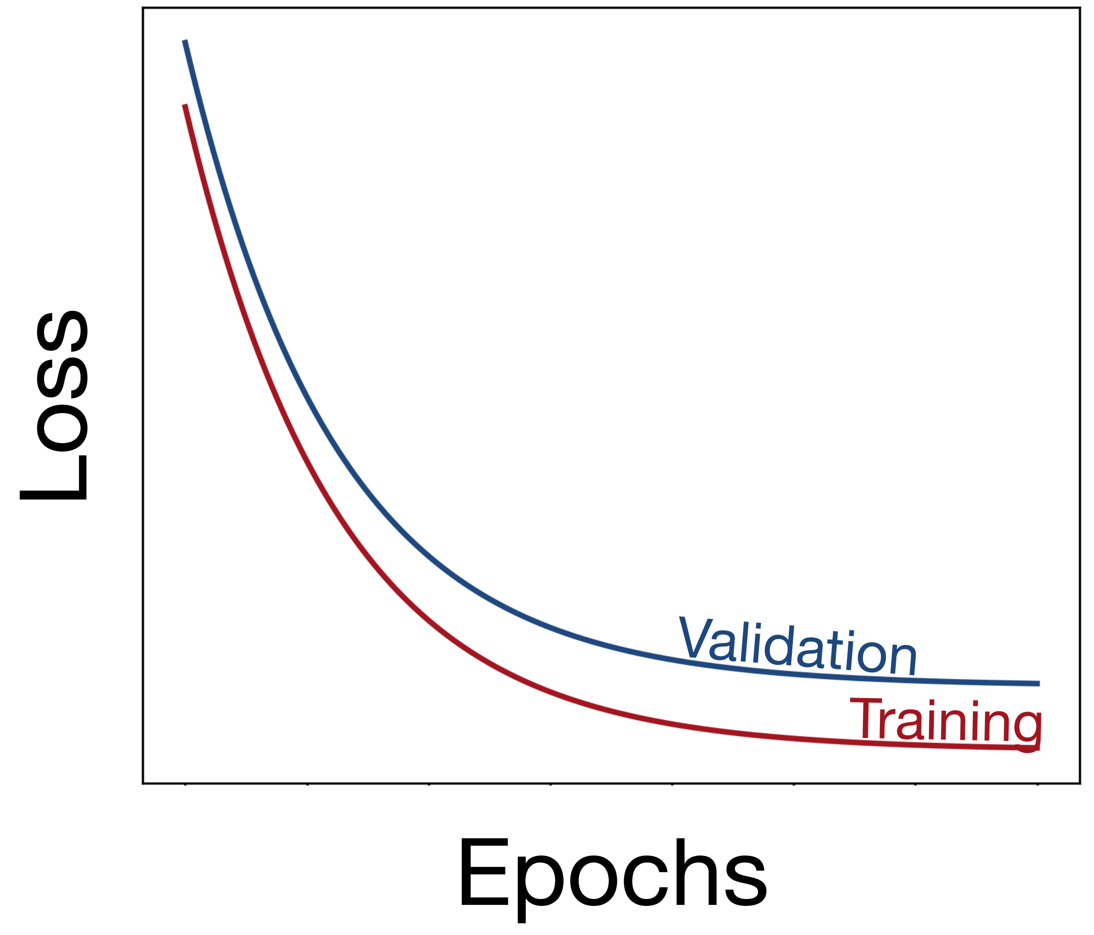


=

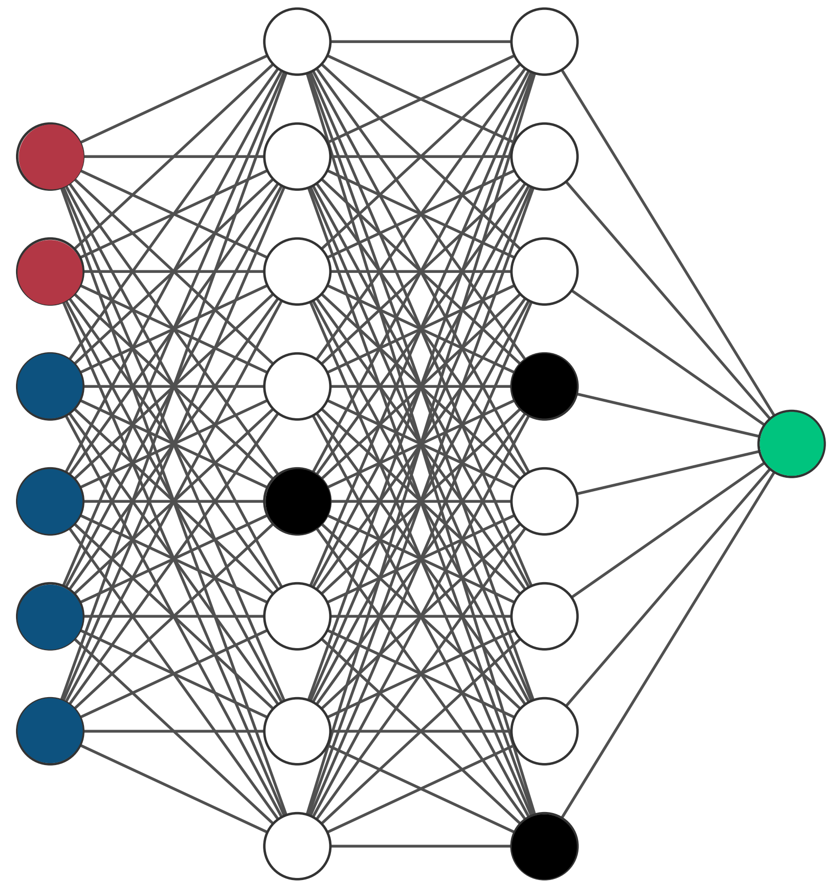


Regularisation

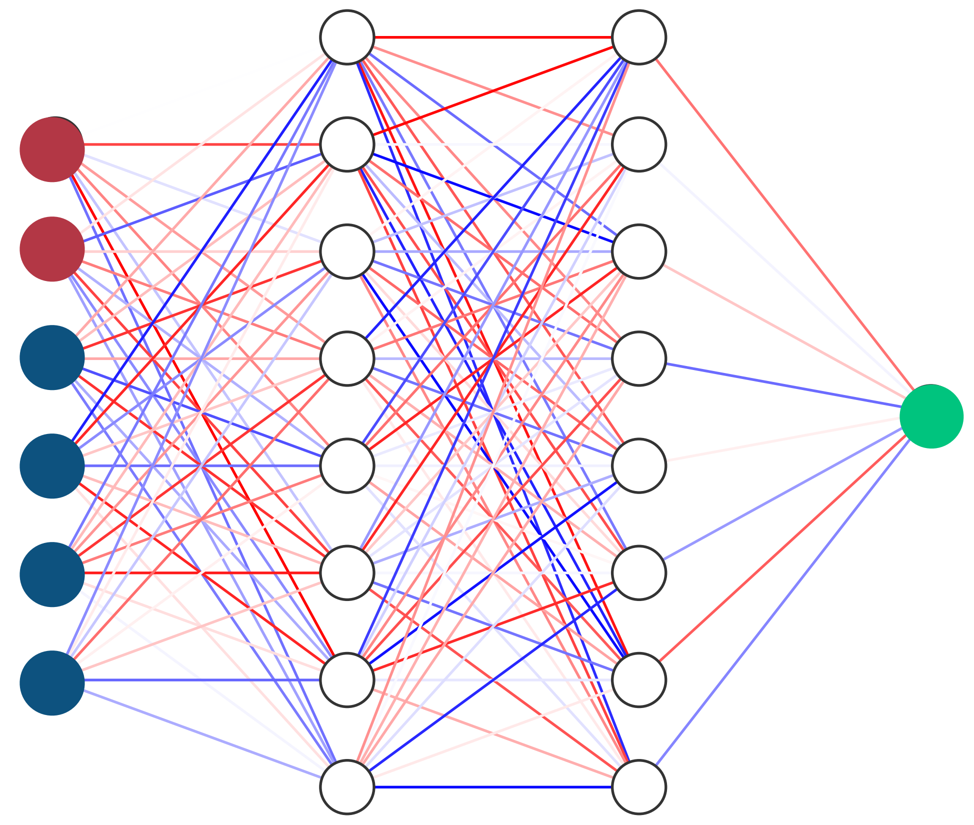
Early stopping



Dropout

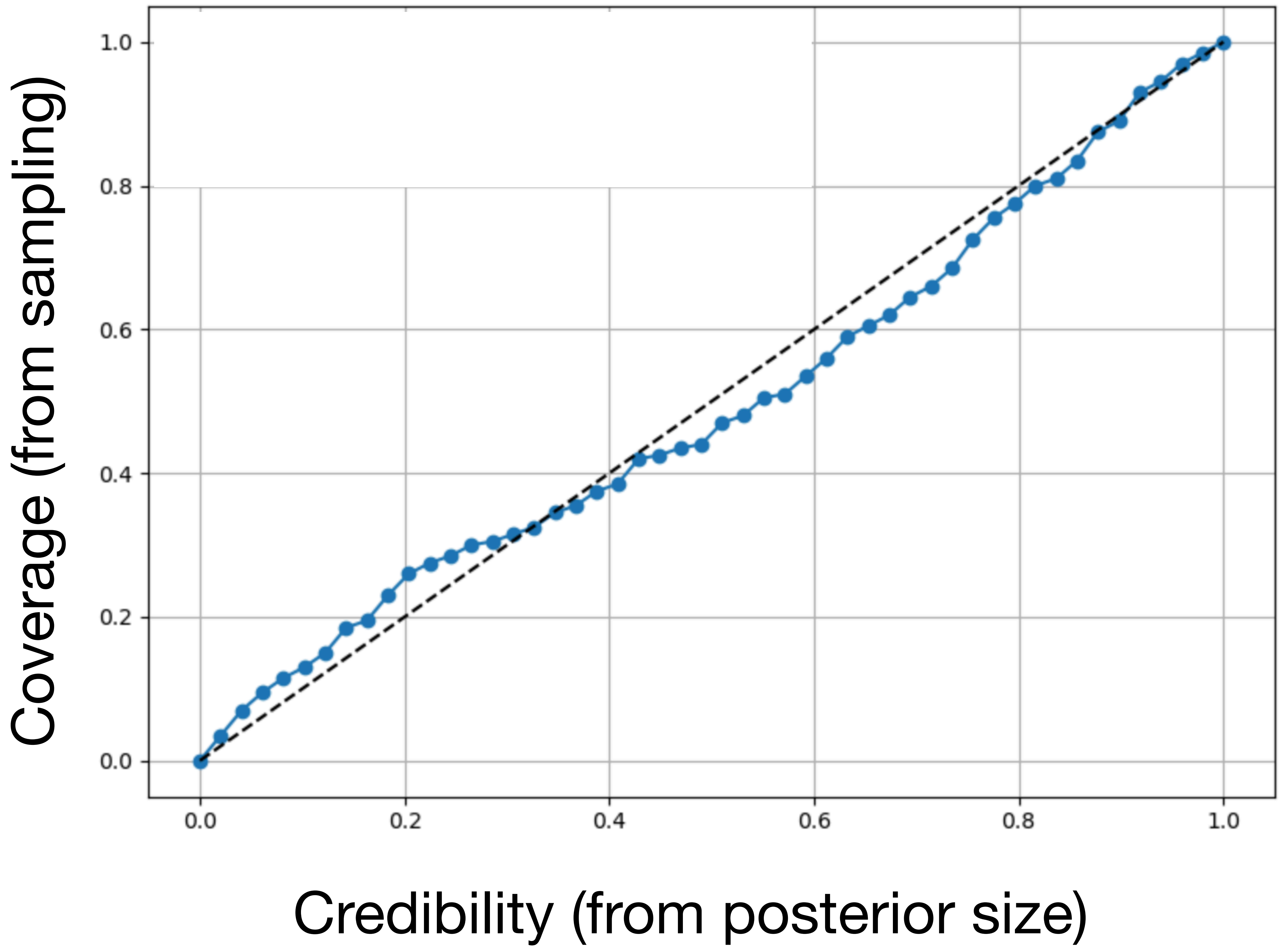


Weight decay

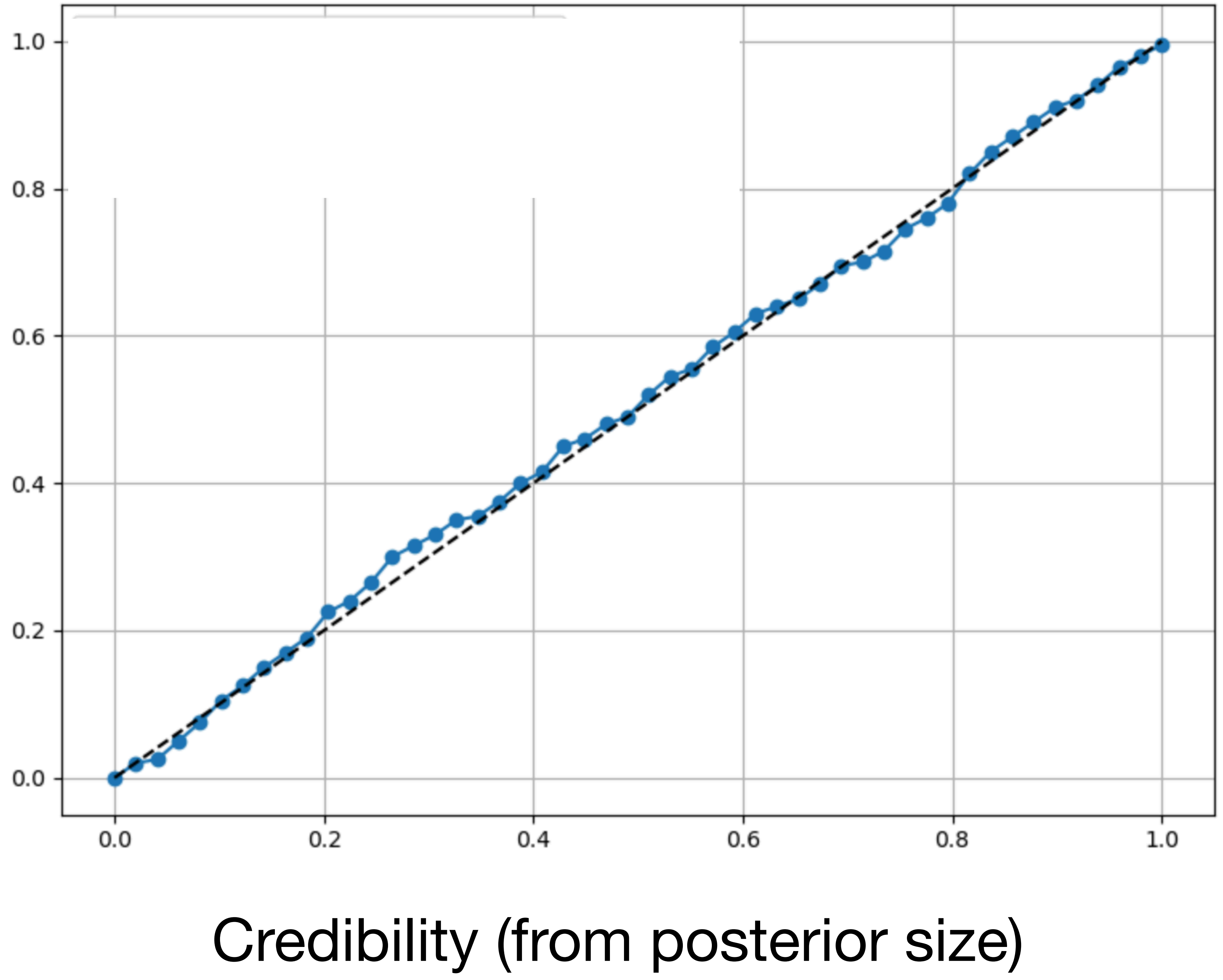


Coverage test

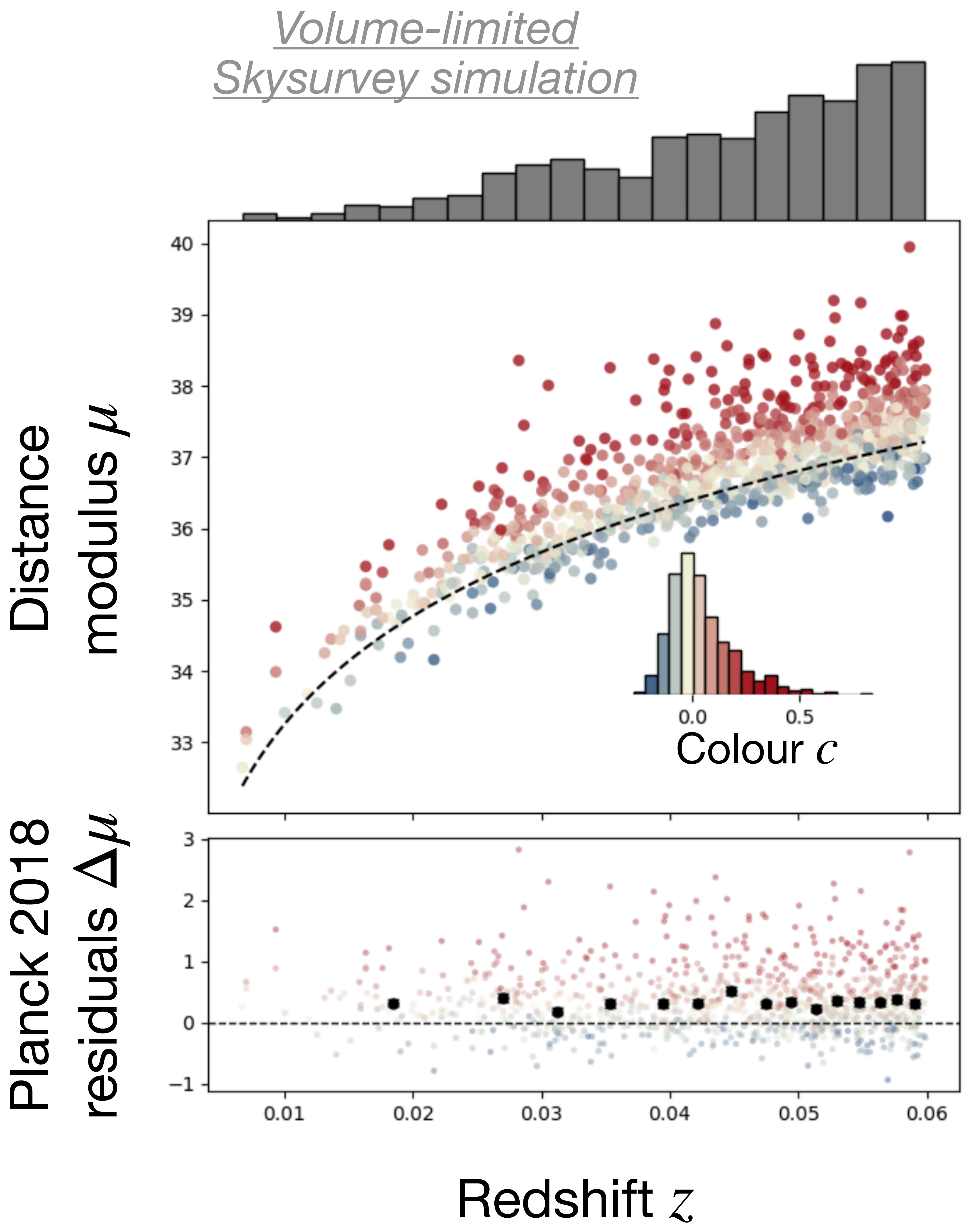
Before regularisation

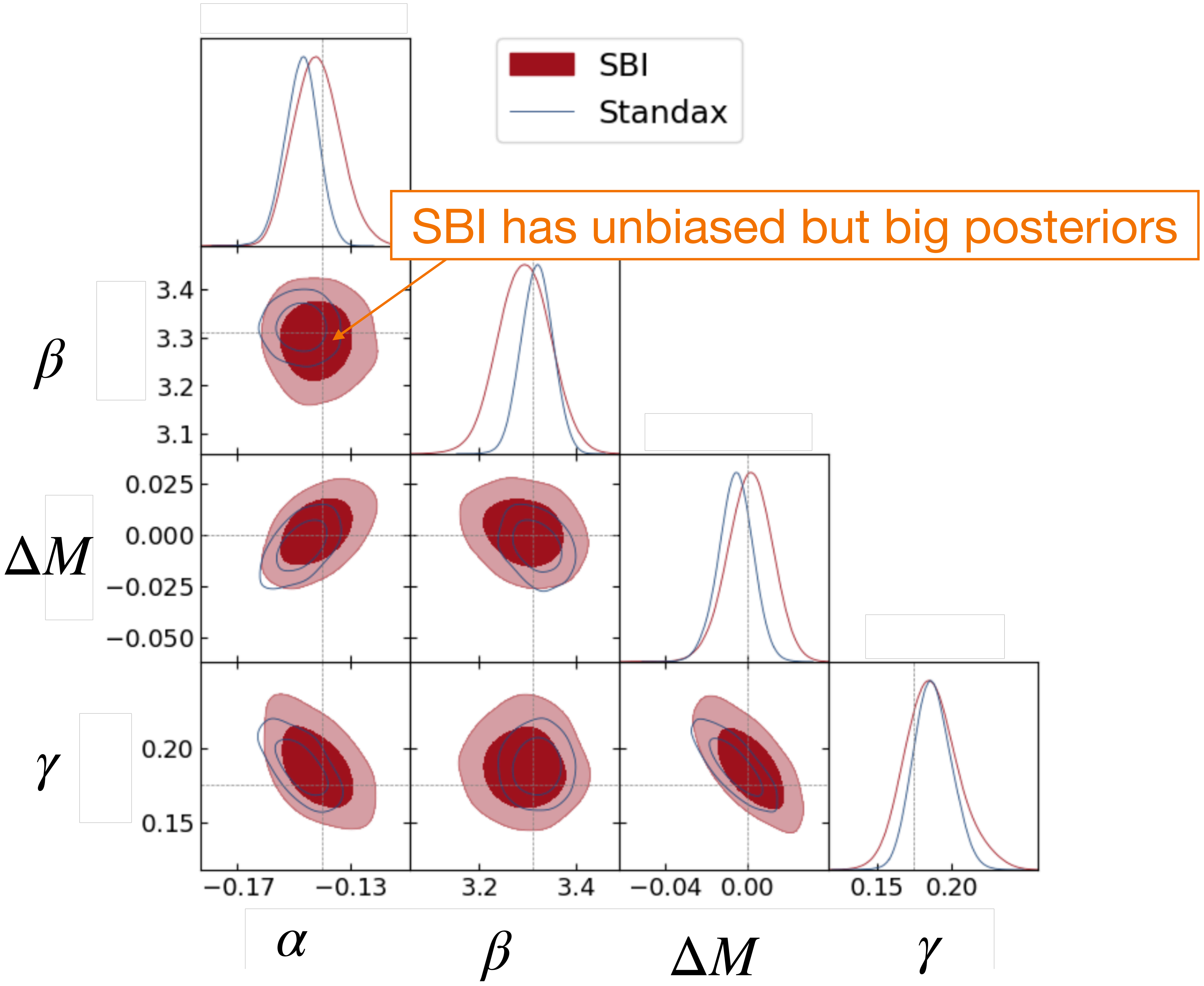


After regularisation

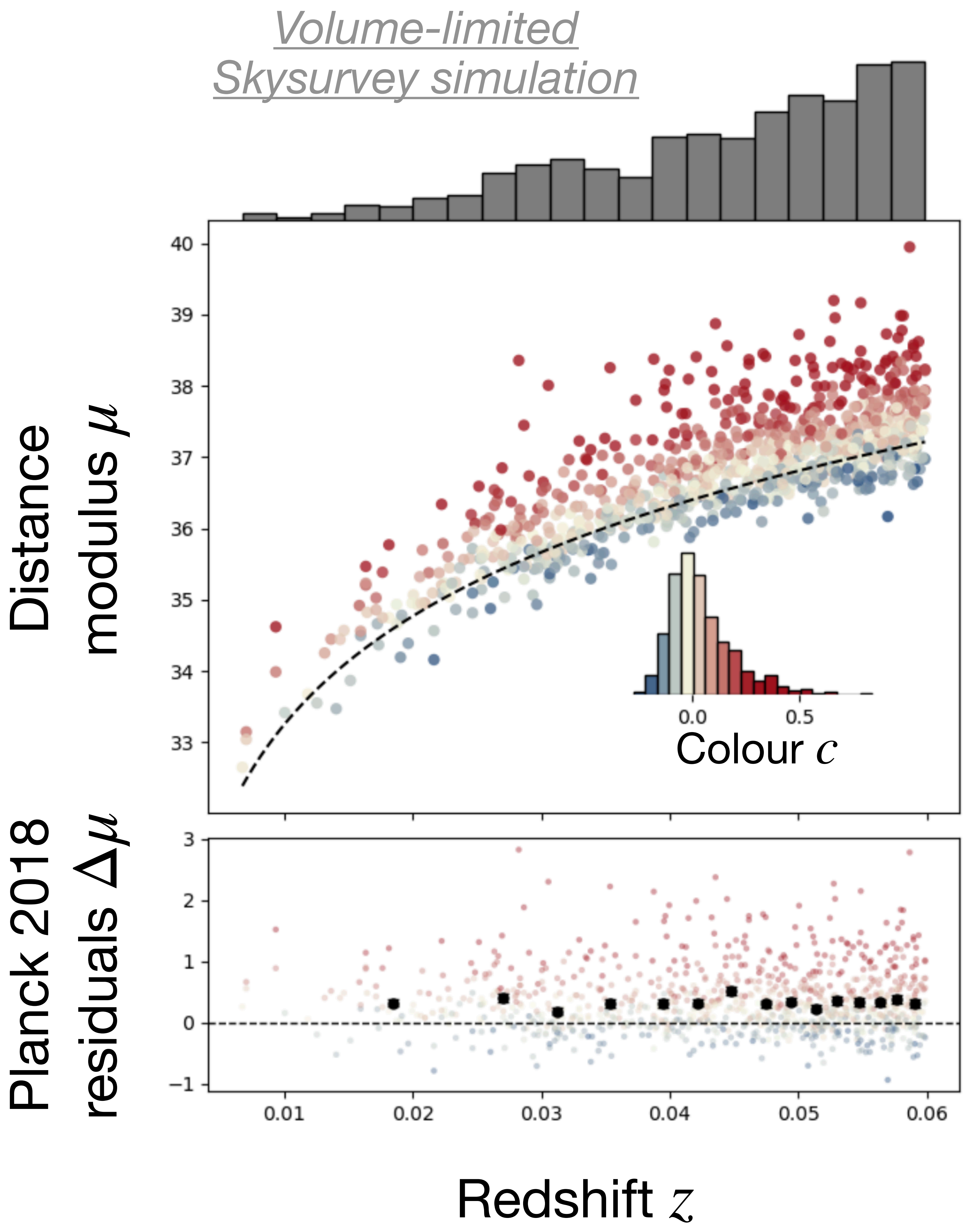


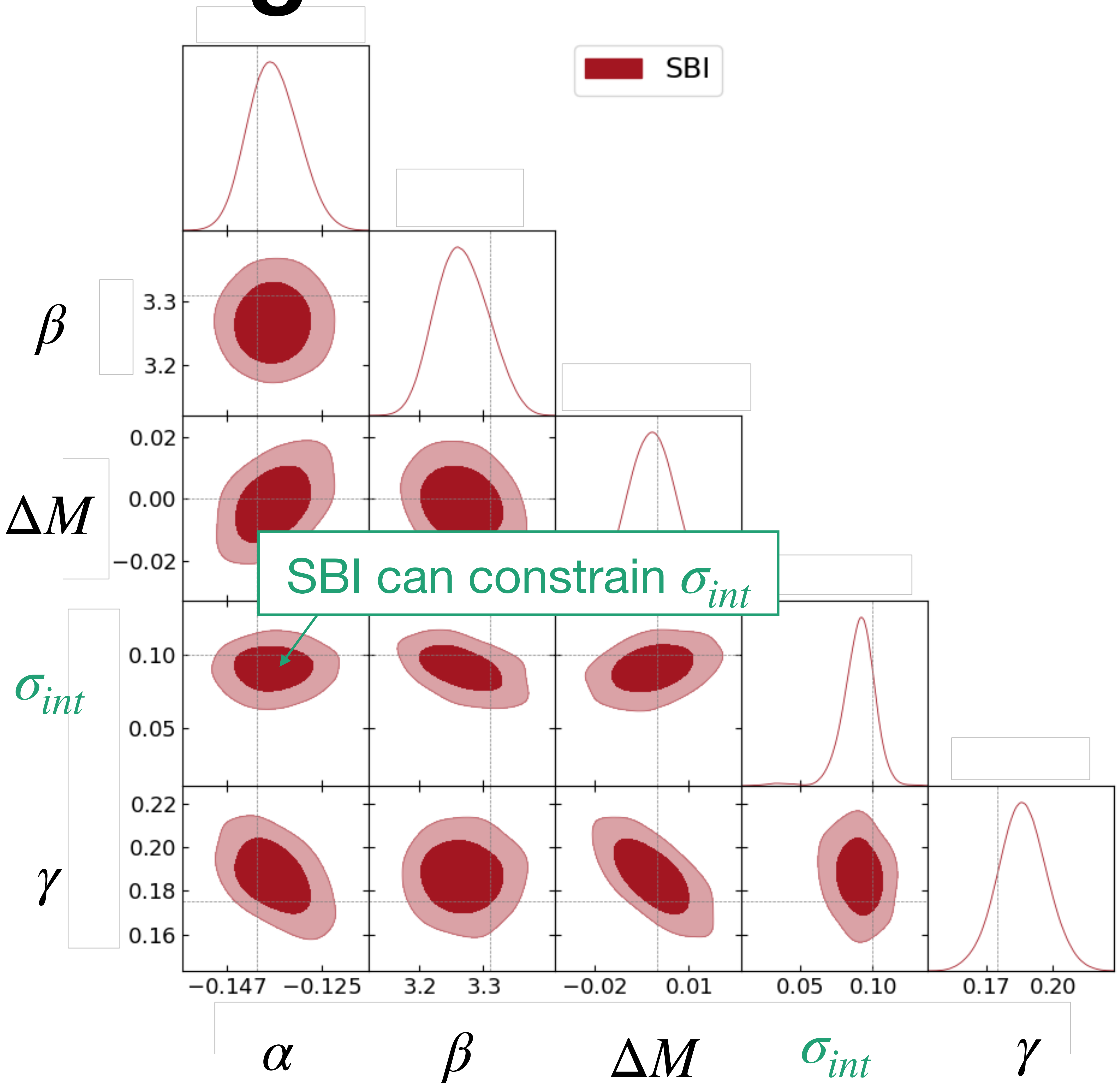
Preliminary results: standardisation



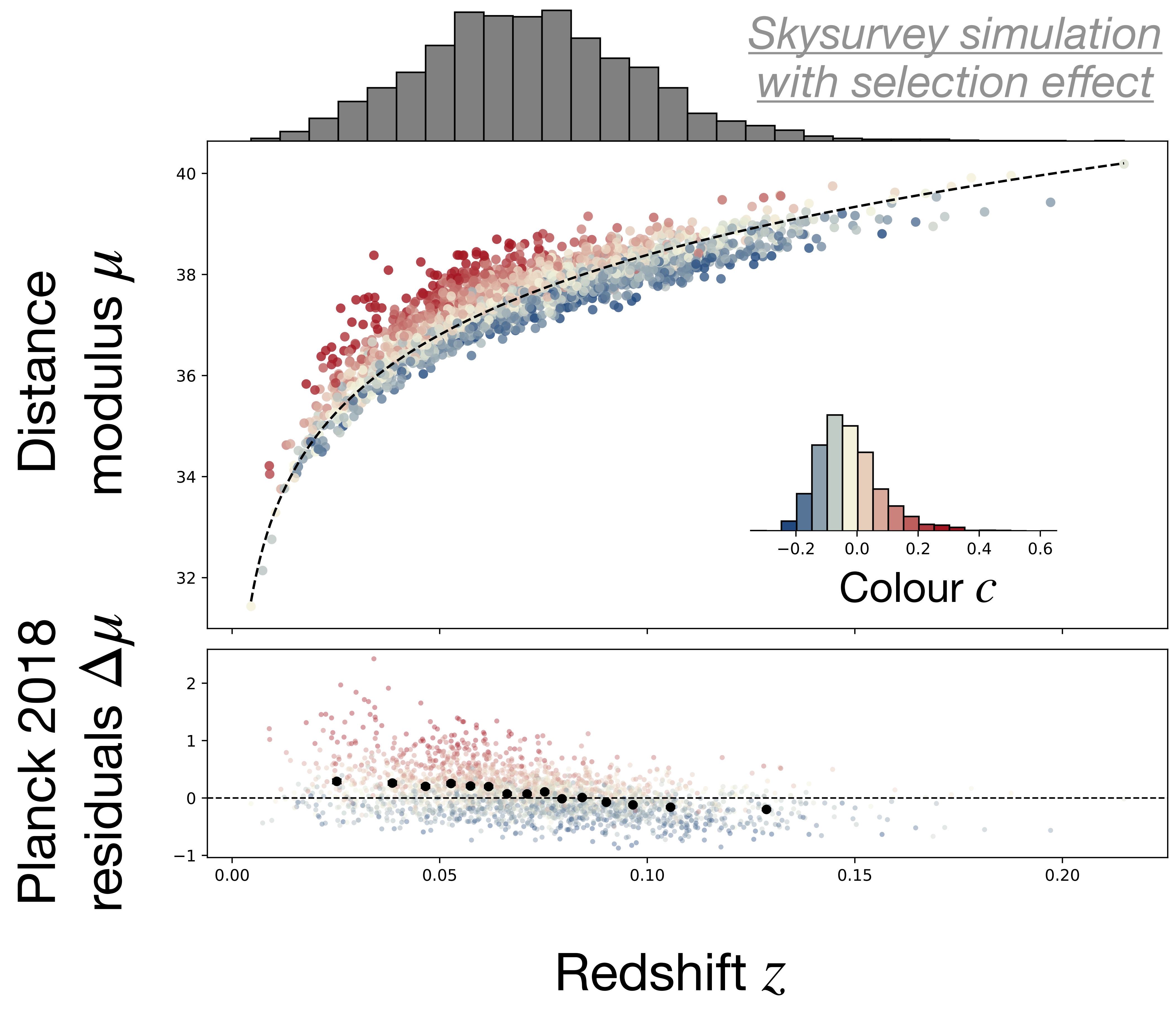


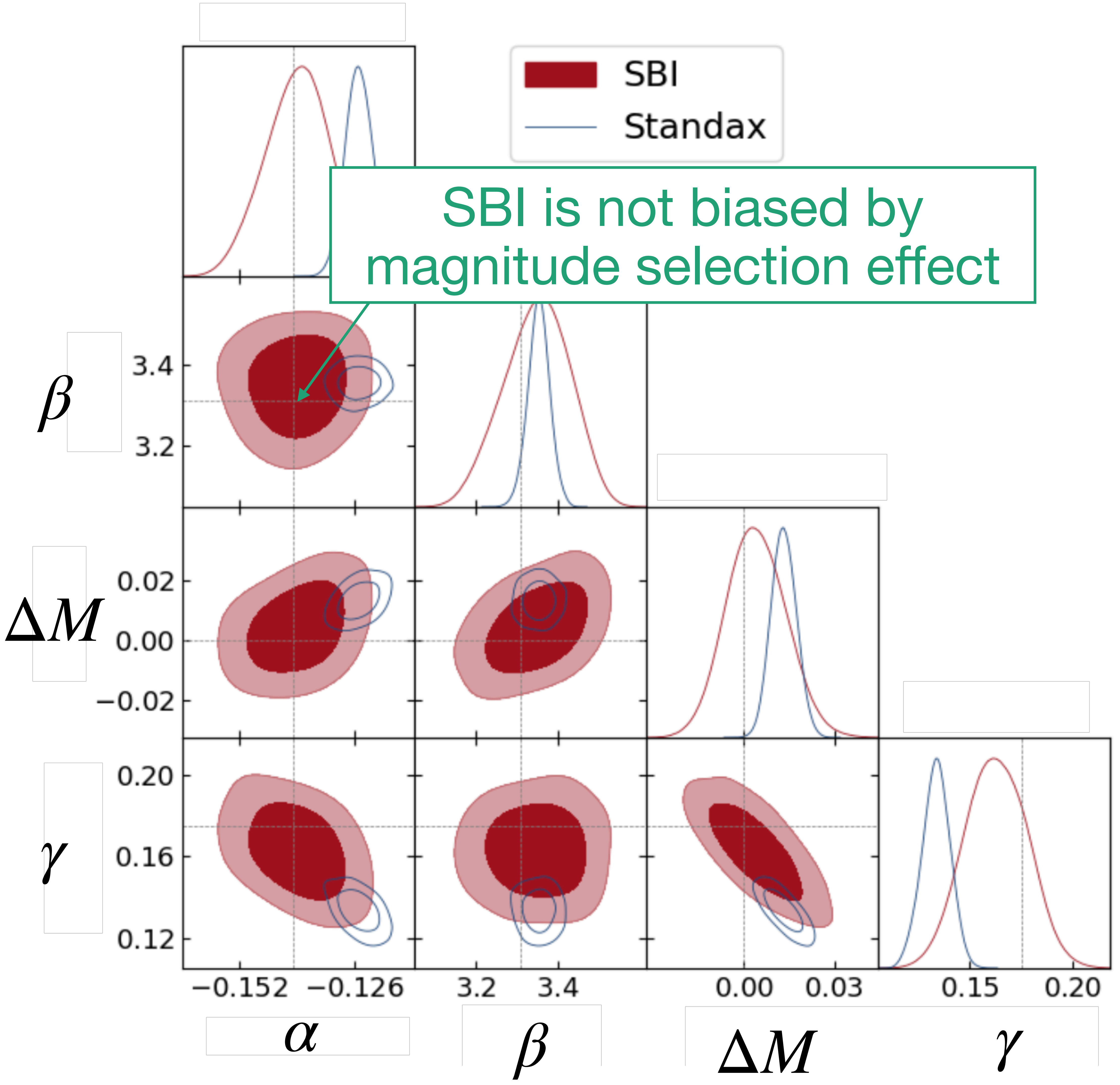
Preliminary results: fitting intrinsic scatter



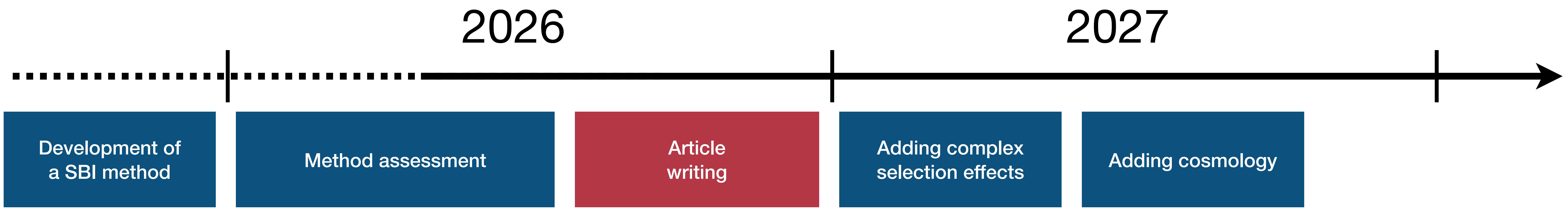


Preliminary results: adding selection effect





Conclusions and perspectives



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

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

Deepset

Zaheer et al. (2018)

$$f(X) = \rho\left(\sum_{x \in X} \phi(X_i)\right)$$

There exist ϕ and ρ such that

f is permutation invariant

Kessler and Scolnic 2017

BEAMS with Bias Correction (BBC)

Skysurvey simulation

$$\text{Standardisation}: M = (M_0 - \bar{\delta}_{M_0}) + \beta(c - \bar{\delta}_c) - \alpha(x_1 - \bar{\delta}_{x_1}) + \gamma p$$

Problem : the inference in two steps loses part of the covariance between selection effects and cosmological parameters