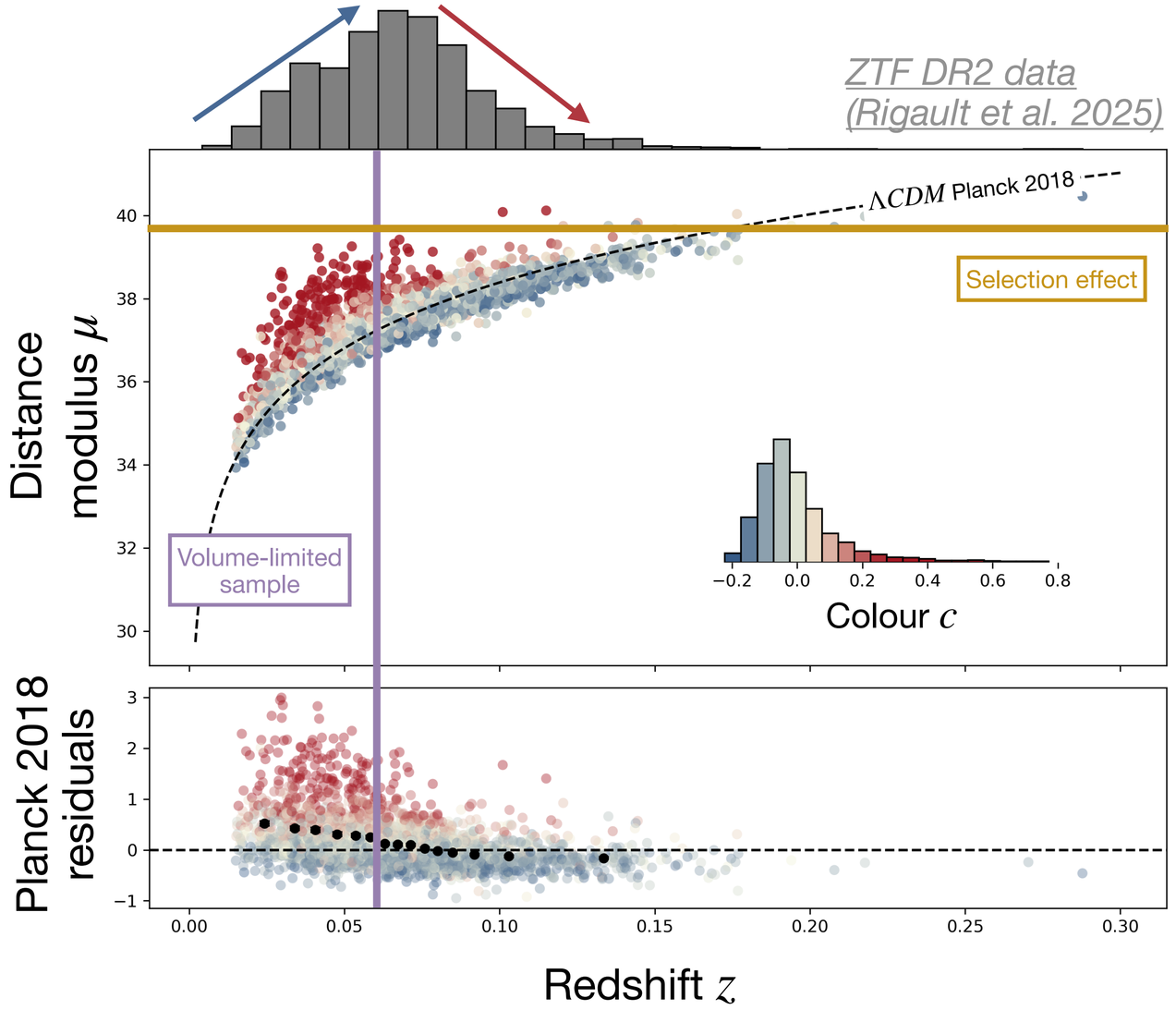
Simulation Based Inference for SNe la cosmology

Adam Trigui

PhD under the supervision of Mickaël Rigault and Florian Ruppin

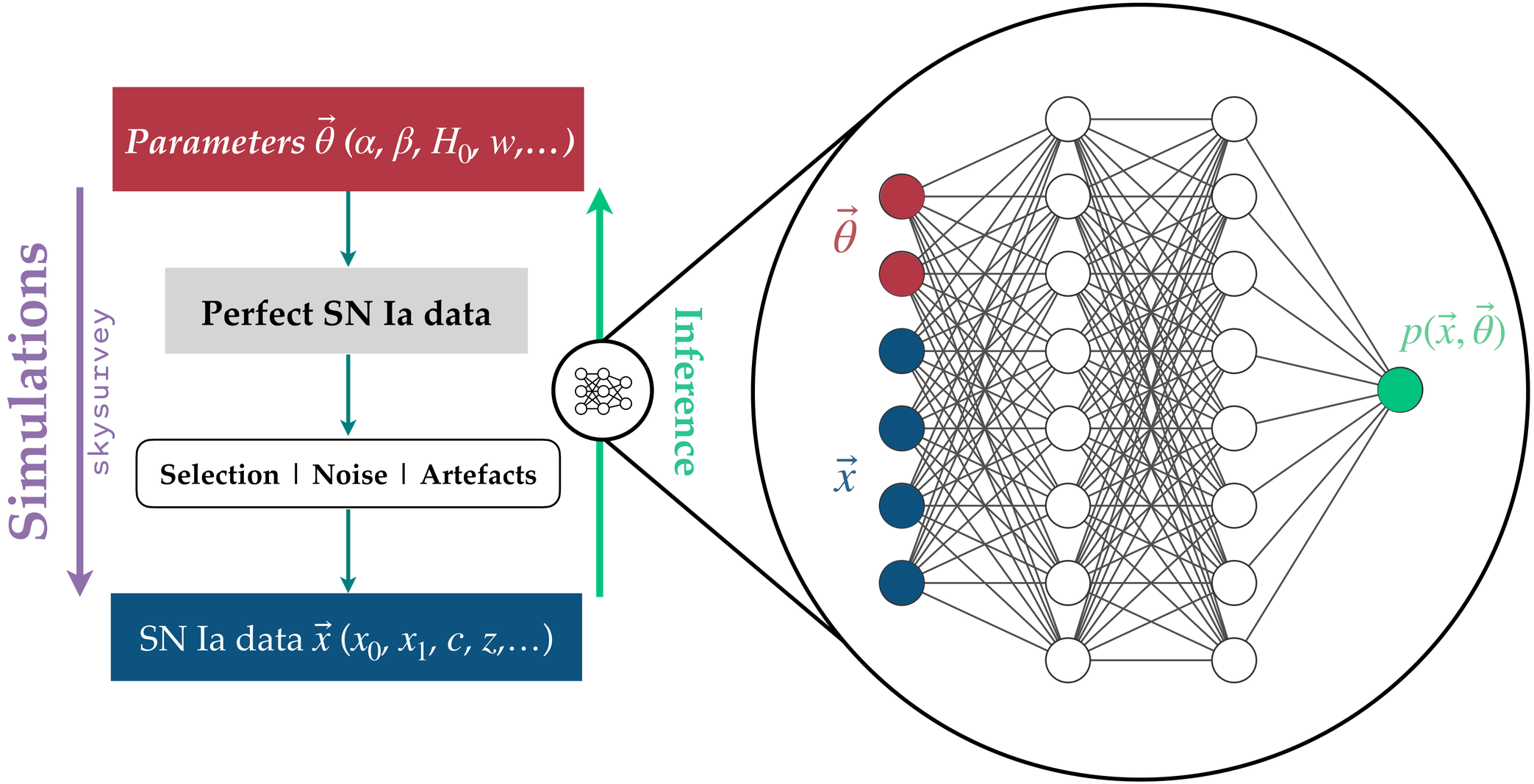
ZTF COSMO 2026, LPNHE

Motivation

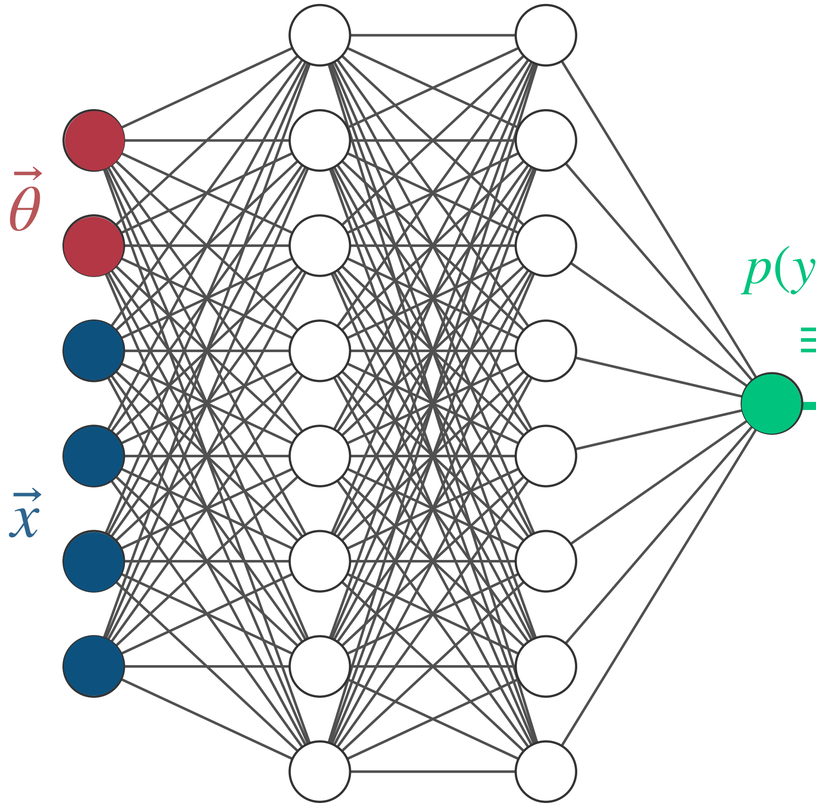

ZTF DR2 data
(Rigault et al. 2025)
ΛCDM Planck 2018
Selection effect
Volume-limited sample
Distance modulus μ
Colour c
Planck 2018 residuals
Redshift z

Method

Neural Ratio Estimator



Neural Ratio Estimator



$p(y = 1 \mid \vec{x}, \vec{\theta})$
$\equiv d(\vec{x}, \vec{\theta})$

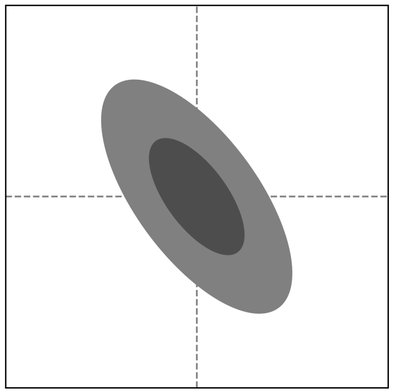
$$\frac{d(\vec{x}, \vec{\theta})}{1 - d(\vec{x}, \vec{\theta})} = \frac{p(\vec{x} \mid \vec{\theta})}{p(\vec{x})}$$

Likelihood

Evidence

Prior $p(\theta)$

MCMC



Loss function
=
binary cross entropy

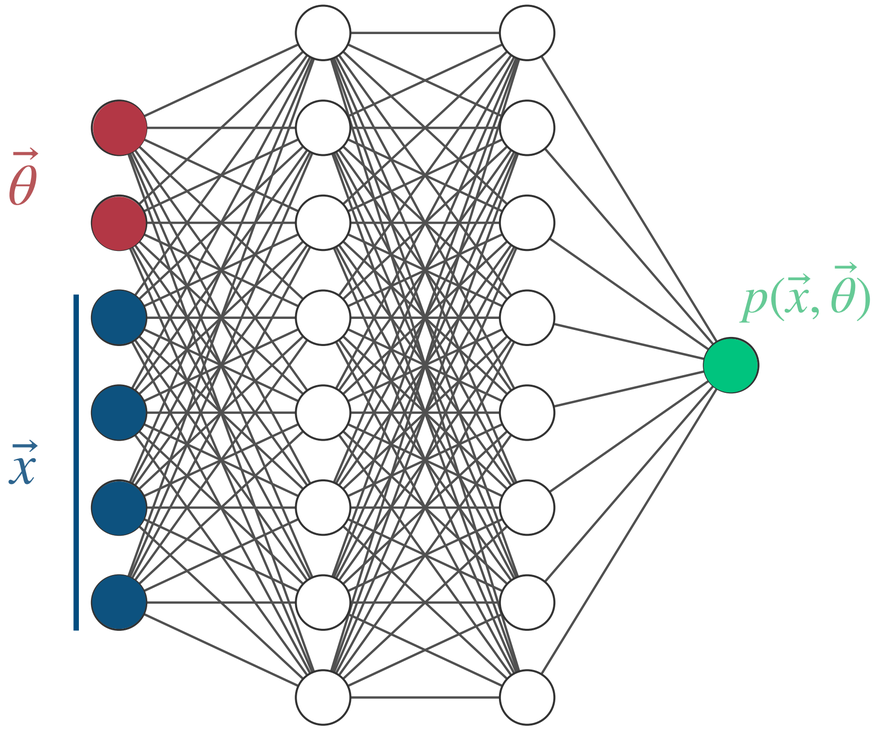
Neural Ratio Estimator

ZTF DR3 = 10 000+ SNe Ia × 10+ observables = 100 000+ input data neurons

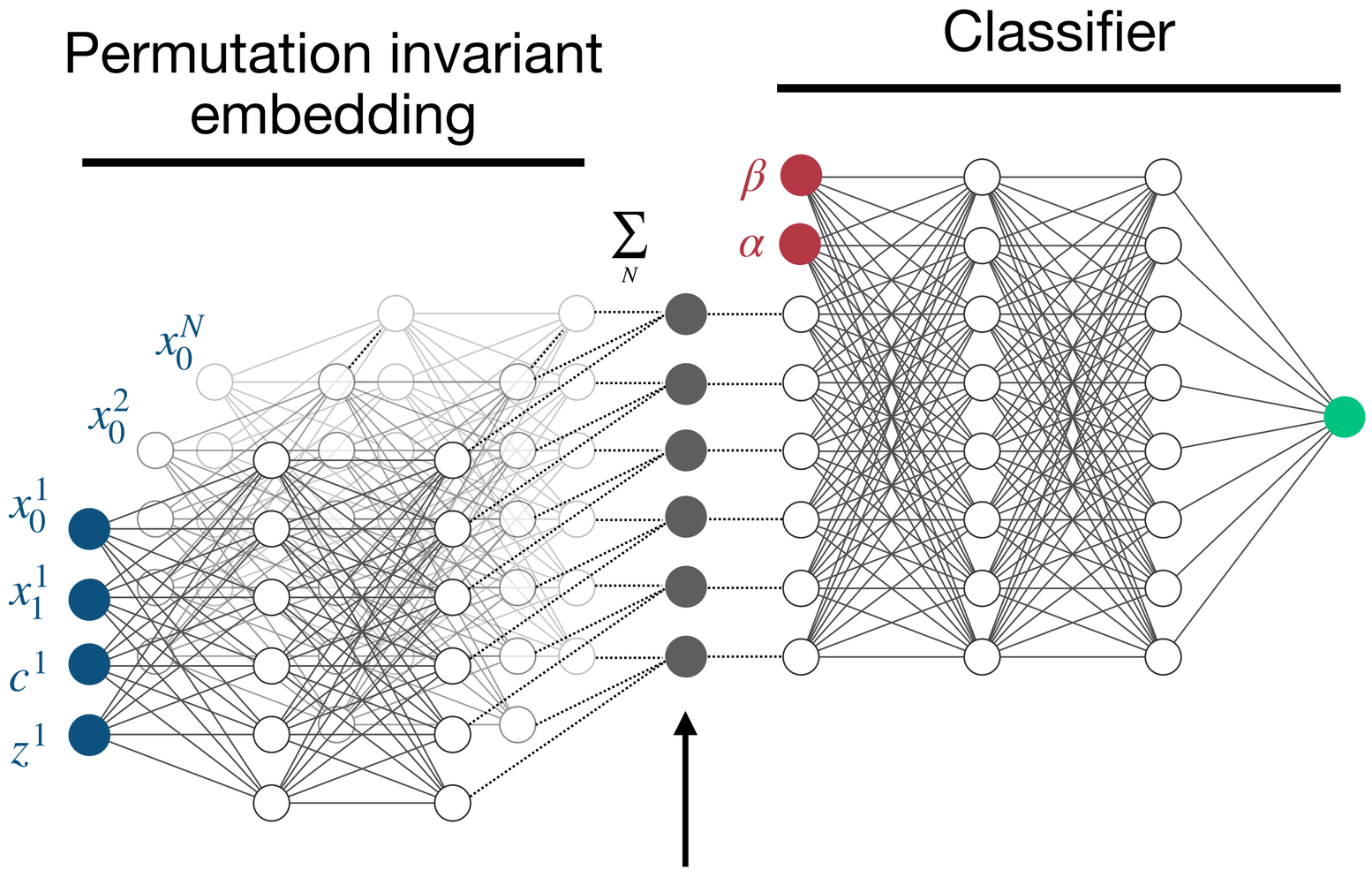
$\vec{x}$

Problem :

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



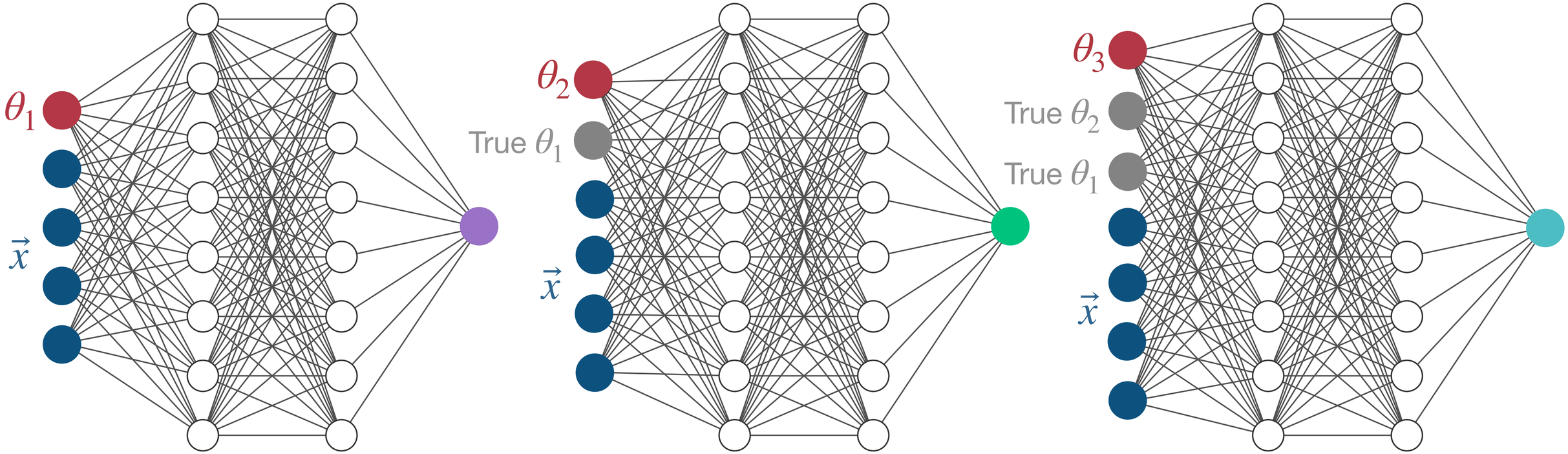
Deepset *Zaheer et al. (2018)*



Complex information encoded into a small size vector (100 000 → 256)

Autoregression *Karchev and Trotta (2024)*

Use joint posteriors to help the network learn difficult parameters.



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

Supernova standardisation

Validation simulations

Only α and β

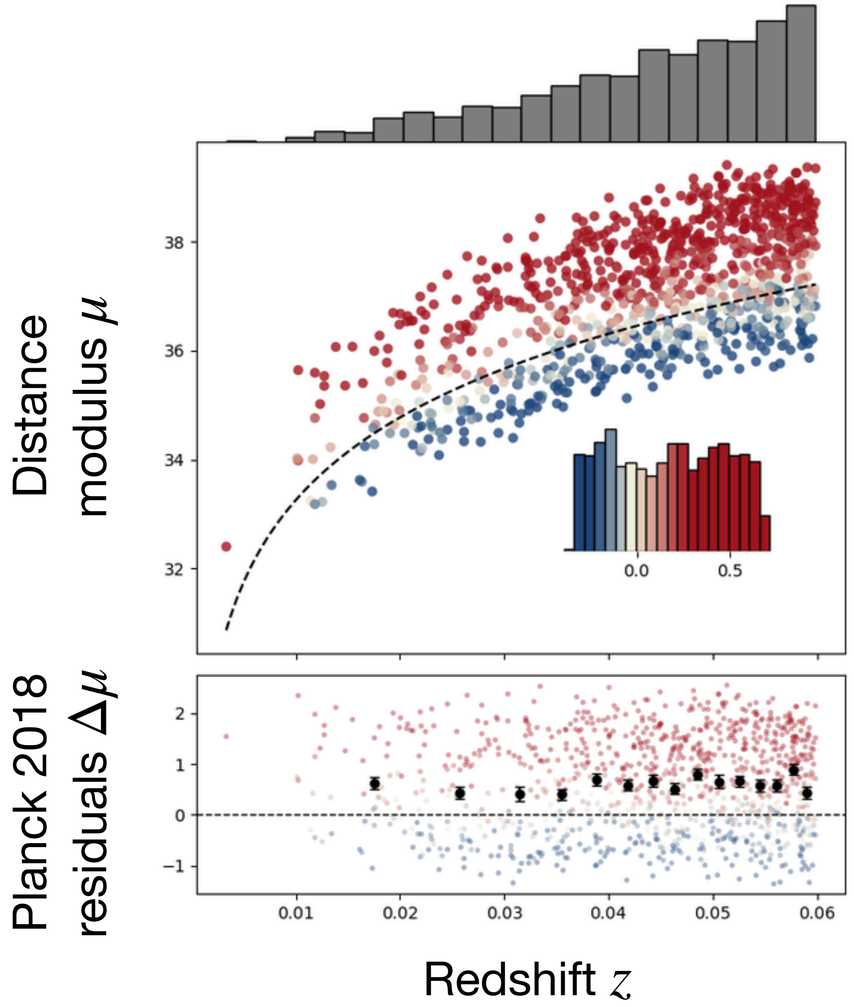
Uniform distributions

No quality cuts

$\alpha = -0.161$

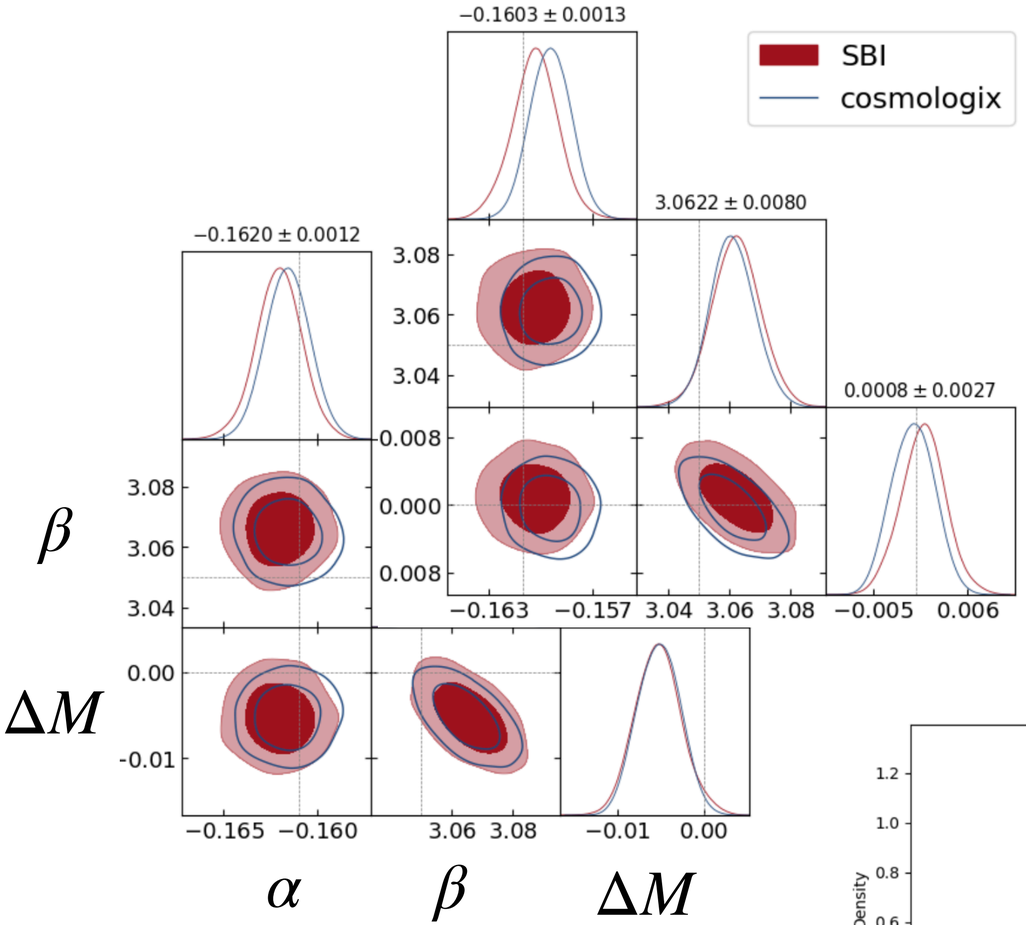
$\beta = 3.05$

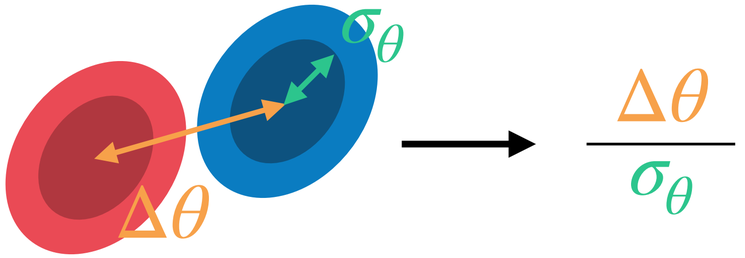
$\sigma_{int} = 0$

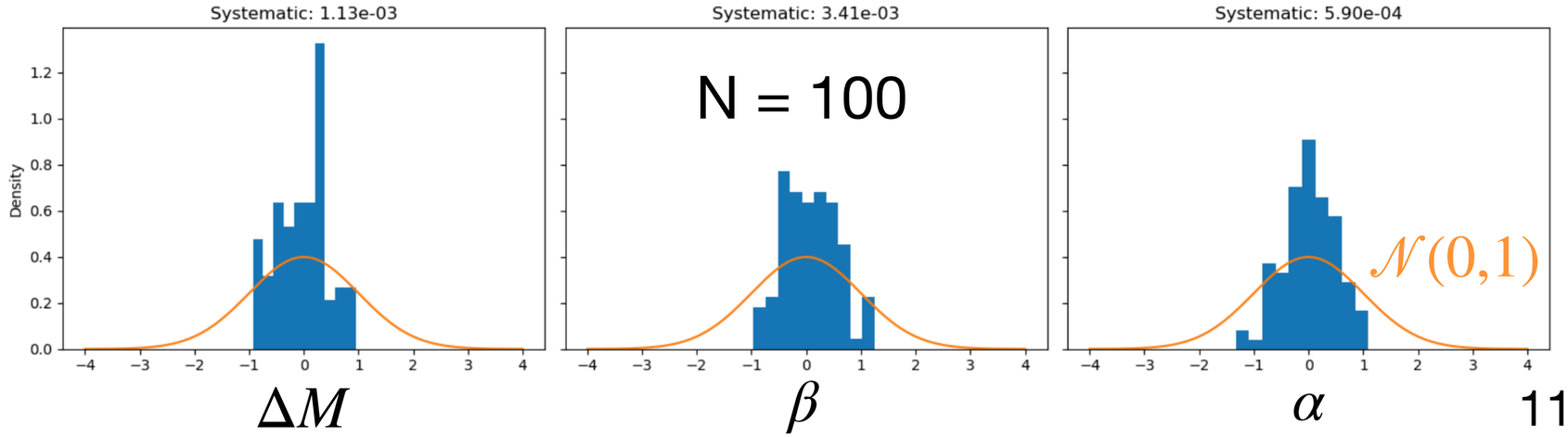


Noise always assumed Gaussian

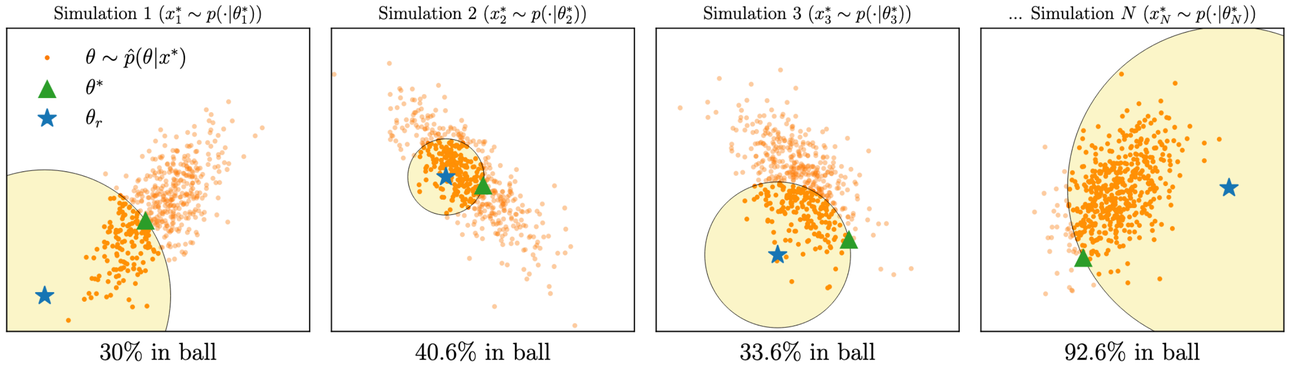
Results and systematic errors



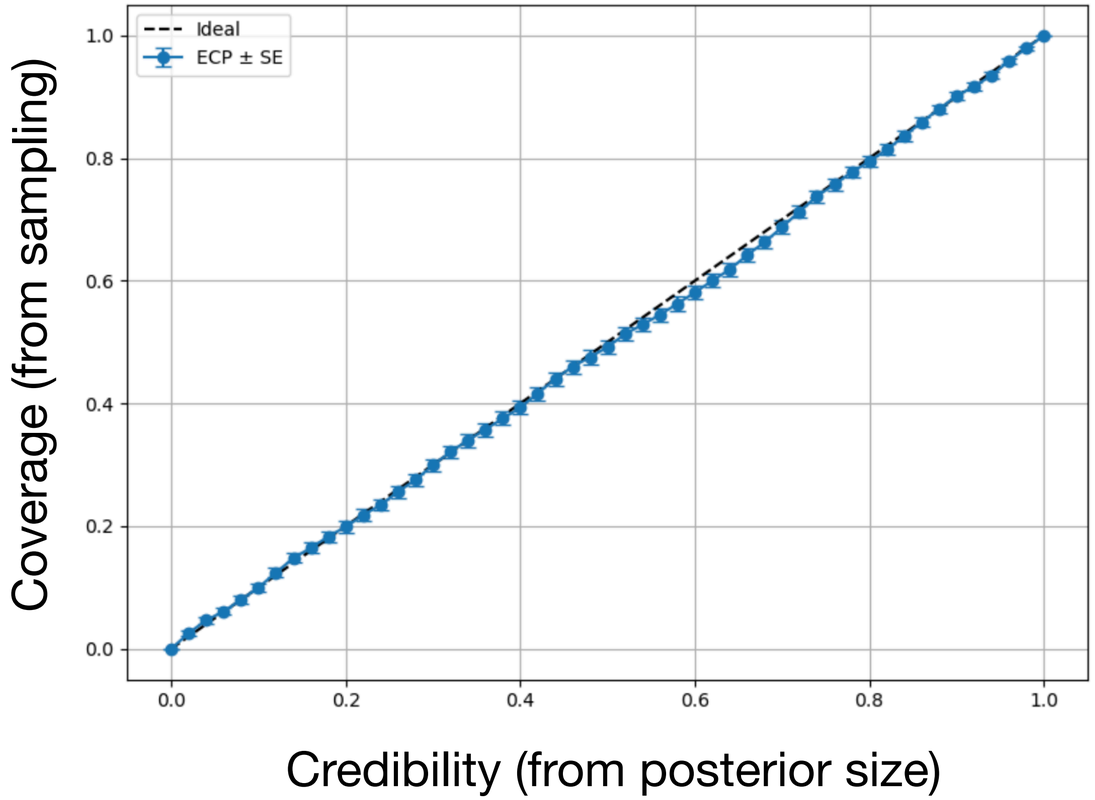




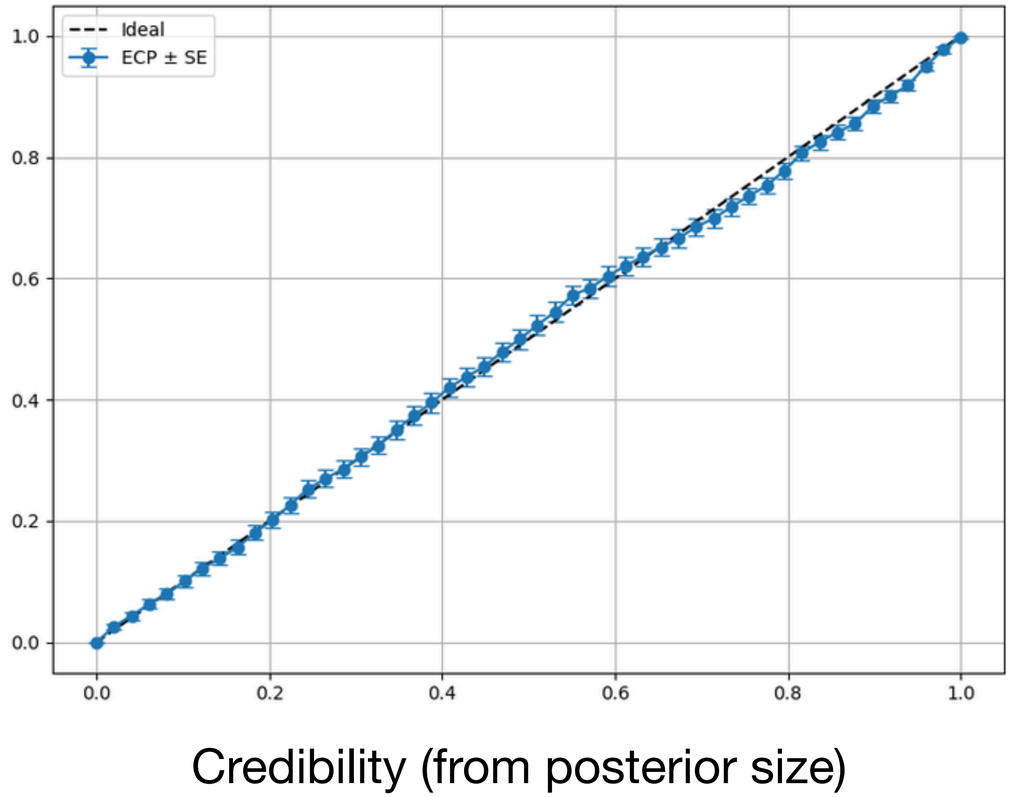
Coverage test *Lemos et al. (2023)*



Cosmologix χ^2



SBI NRE



Both methods have a good coverage !

Realistic simulations

DR2 observable distributions

DR2 error distributions

Quality cuts $\sigma_{x1} < 1$ and $\sigma_c < 0.1$

BUT

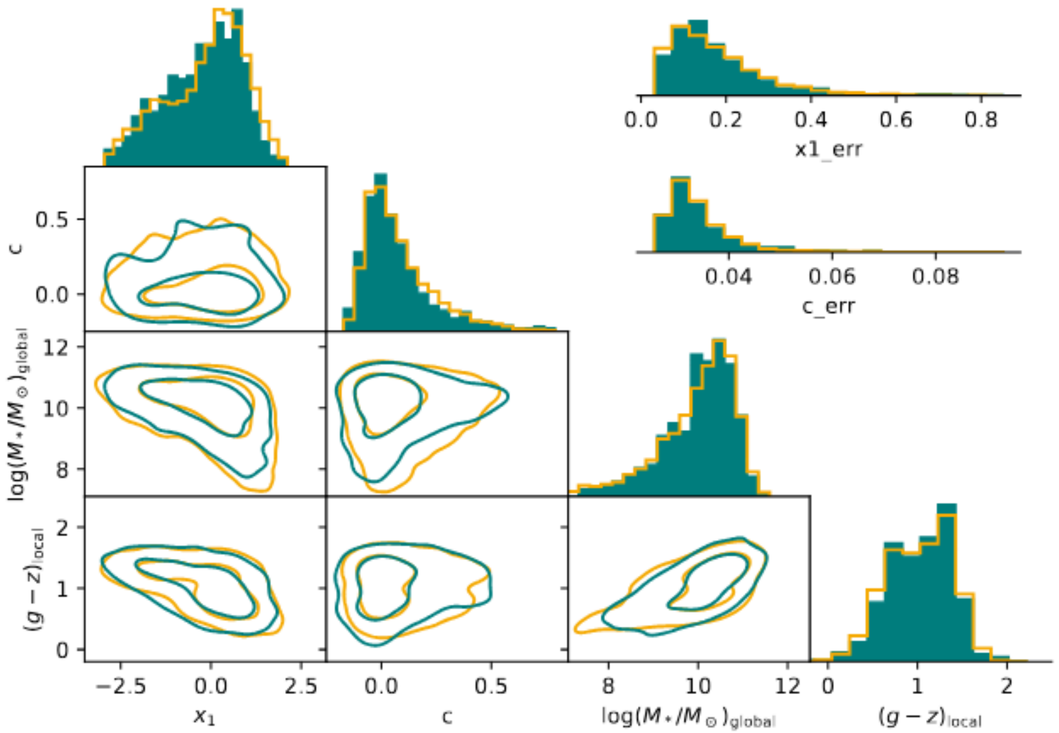
Constant rate

No use of light curves:

→ no covariance on x1 and c

→ no cut on fit quality

Ginolin et al. (2025b)



Simulations parameters

$$\alpha = -0.161$$

$$\beta = 3.05$$

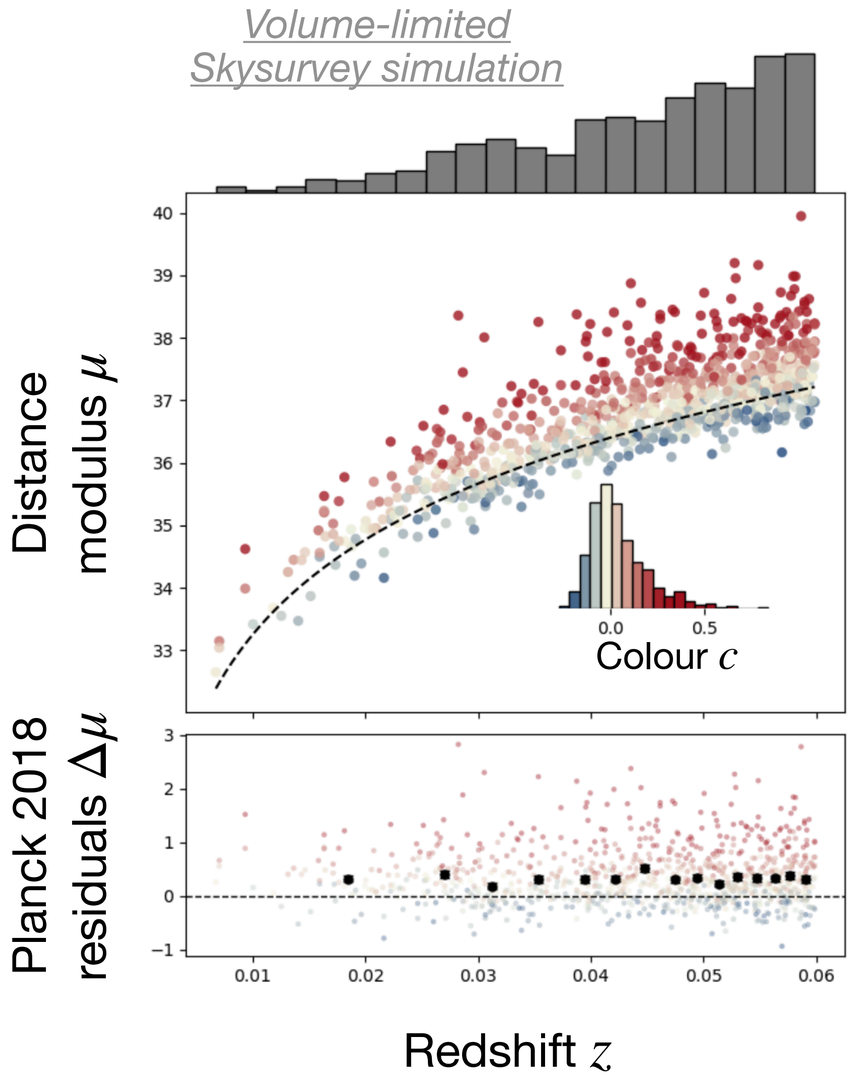
$$\gamma = 0.143$$

$$\sigma_{int} = 0.1$$

Volume-limited = ∼1000 SNe Ia

Full DR2 = ∼3000 SNe Ia

200 000 samples for SBI training



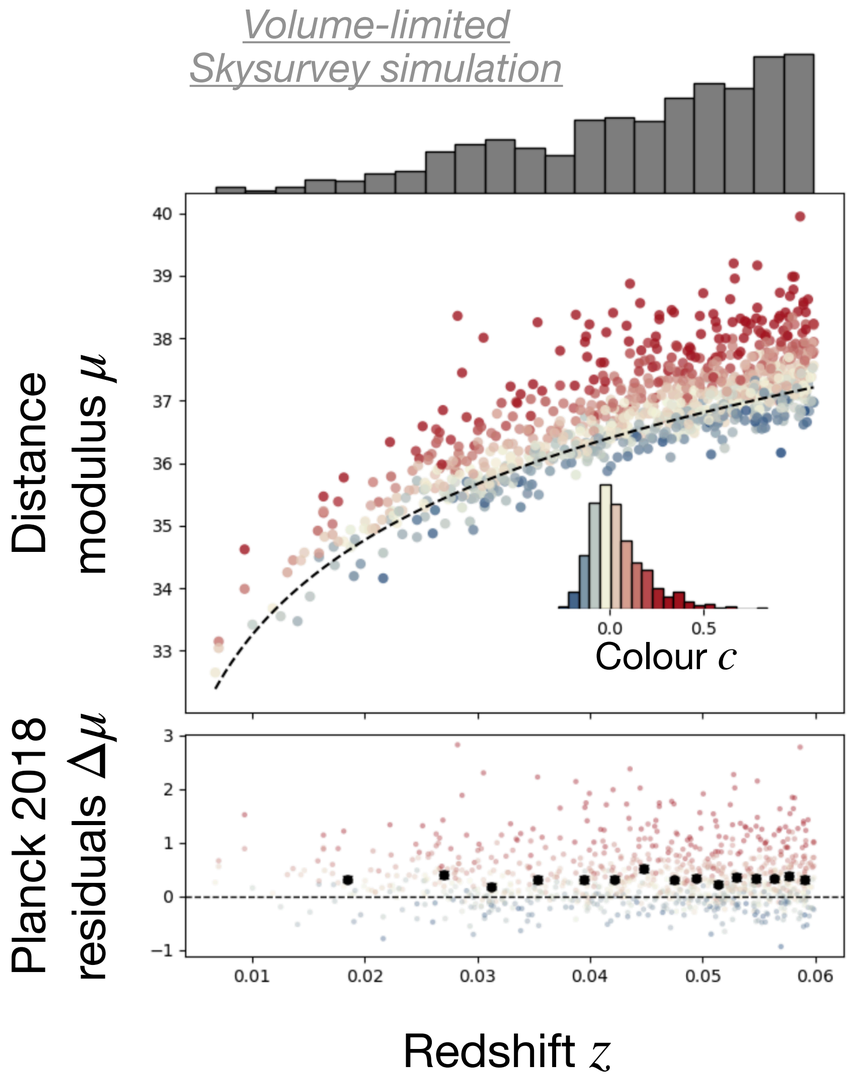
Results

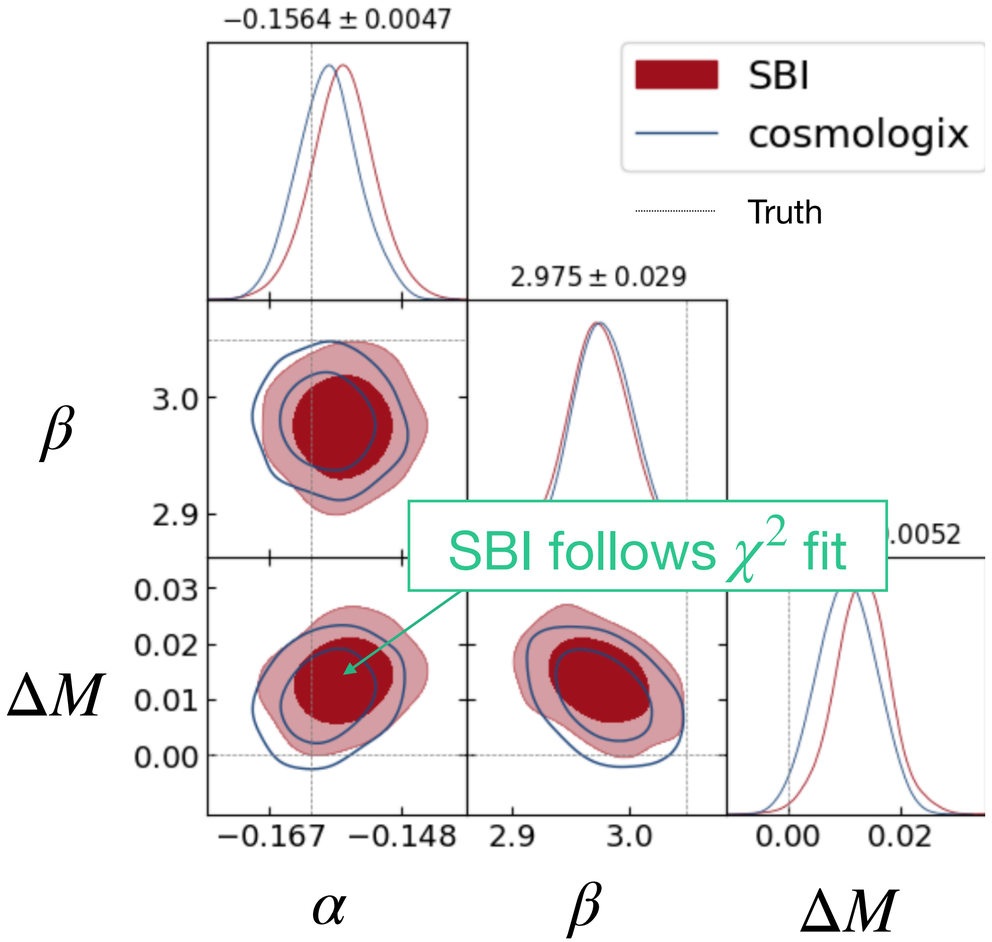
1. α and β
2. α, β and γ
3. α, β and γ with proxy (local colour)
4. α, β, γ (proxy) and σ_{int}
5. α, β, γ and σ_{int} with magnitude selection effect

Results

1. α and β
2. α, β and γ
3. α, β and γ with proxy (local colour)
4. α, β, γ (proxy) and σ_{int}
5. α, β, γ and σ_{int} with magnitude selection effect

Results: α and β

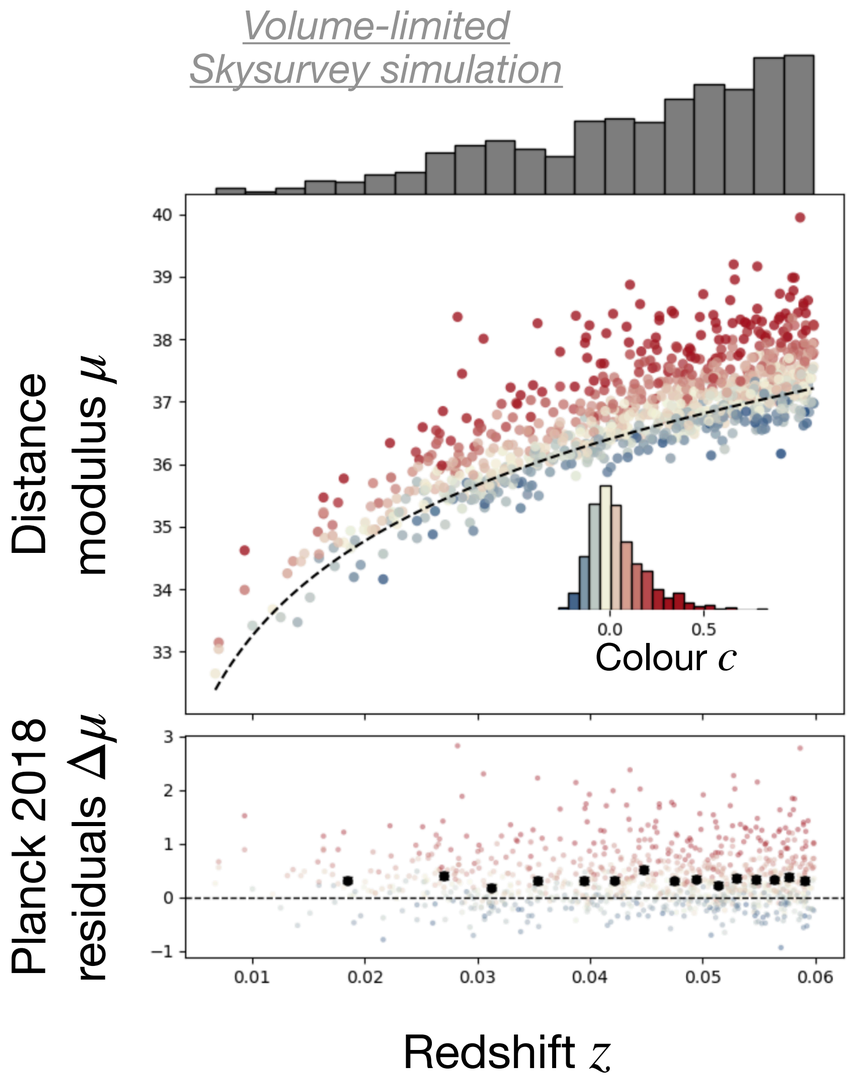


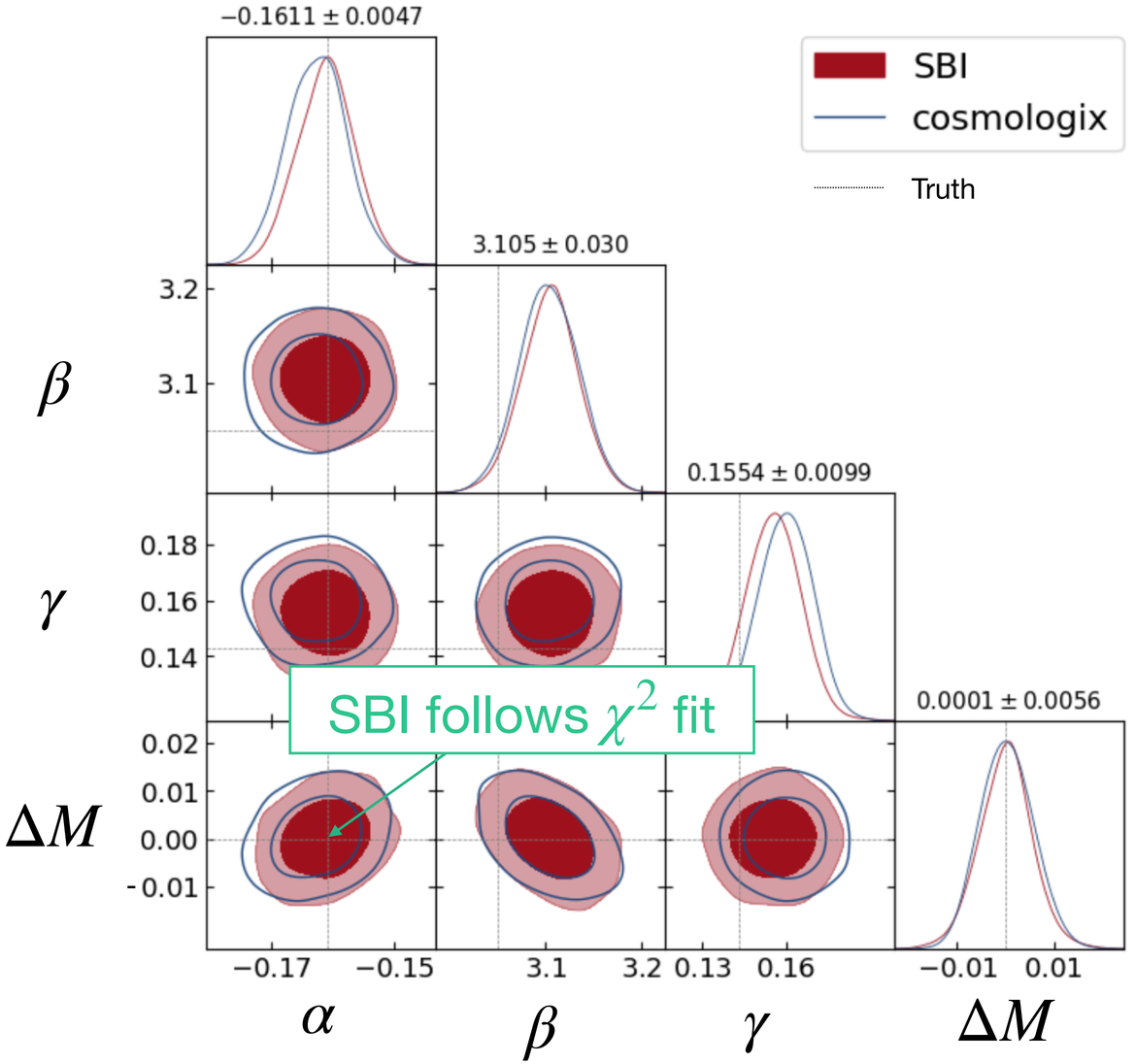


Results

1. α and β
2. α, β and γ
3. α, β and γ with proxy (local colour)
4. α, β, γ (proxy) and σ_{int}
5. α, β, γ and σ_{int} with magnitude selection effect

Results: α, β and γ

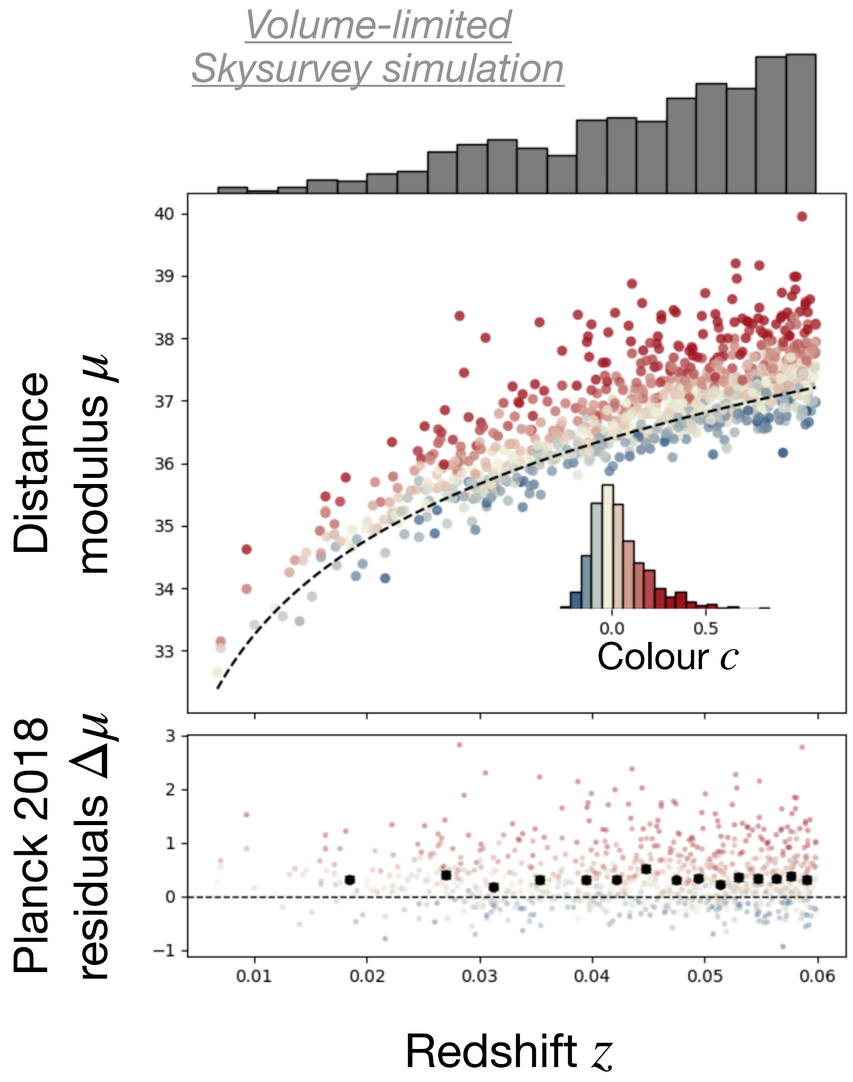


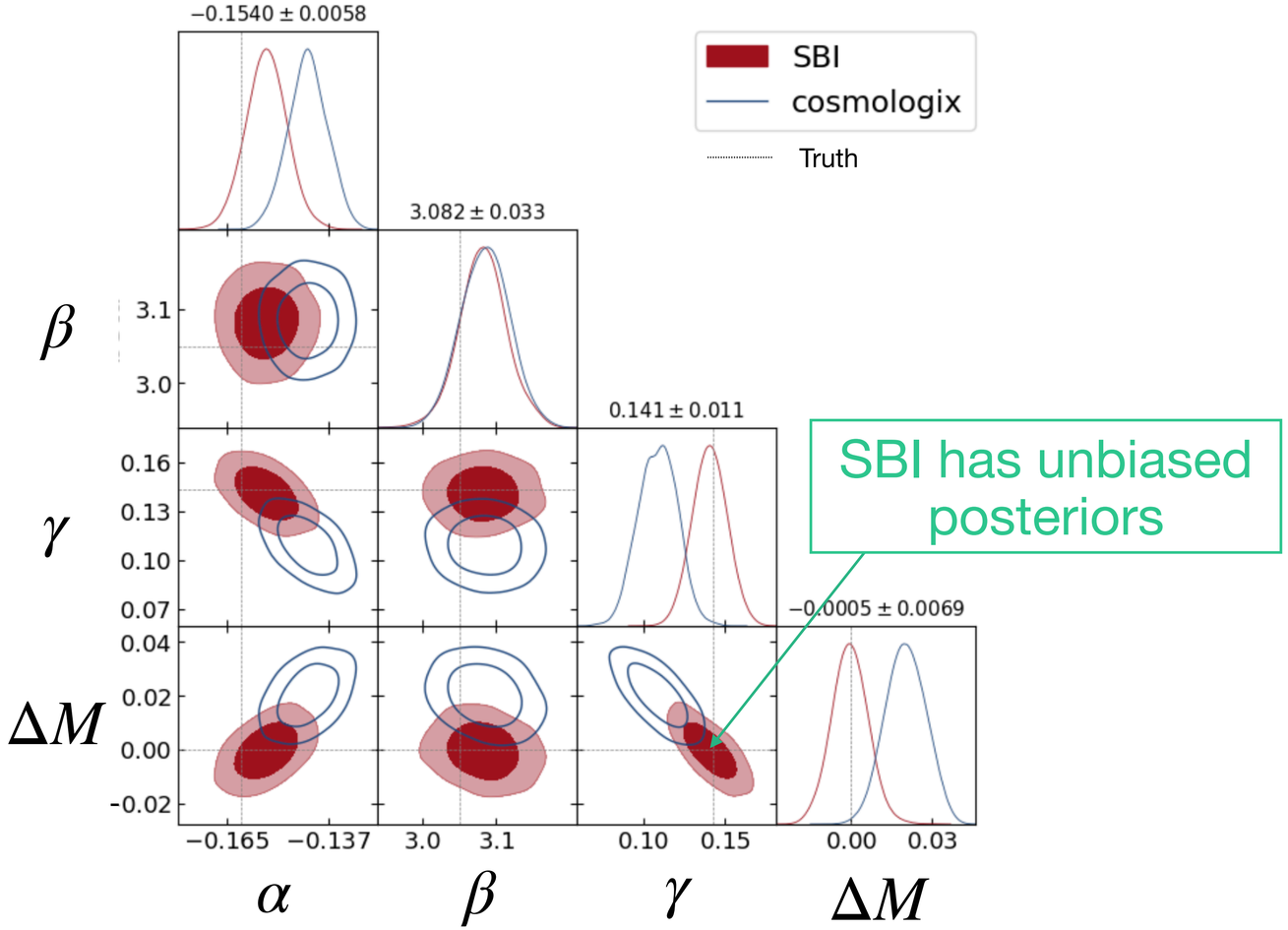


Results

1. α and β
2. α, β and γ
3. α, β and γ with proxy (local colour)
4. α, β, γ (proxy) and σ_{int}
5. α, β, γ and σ_{int} with magnitude selection effect

Results: α, β and γ with proxy (local colour)





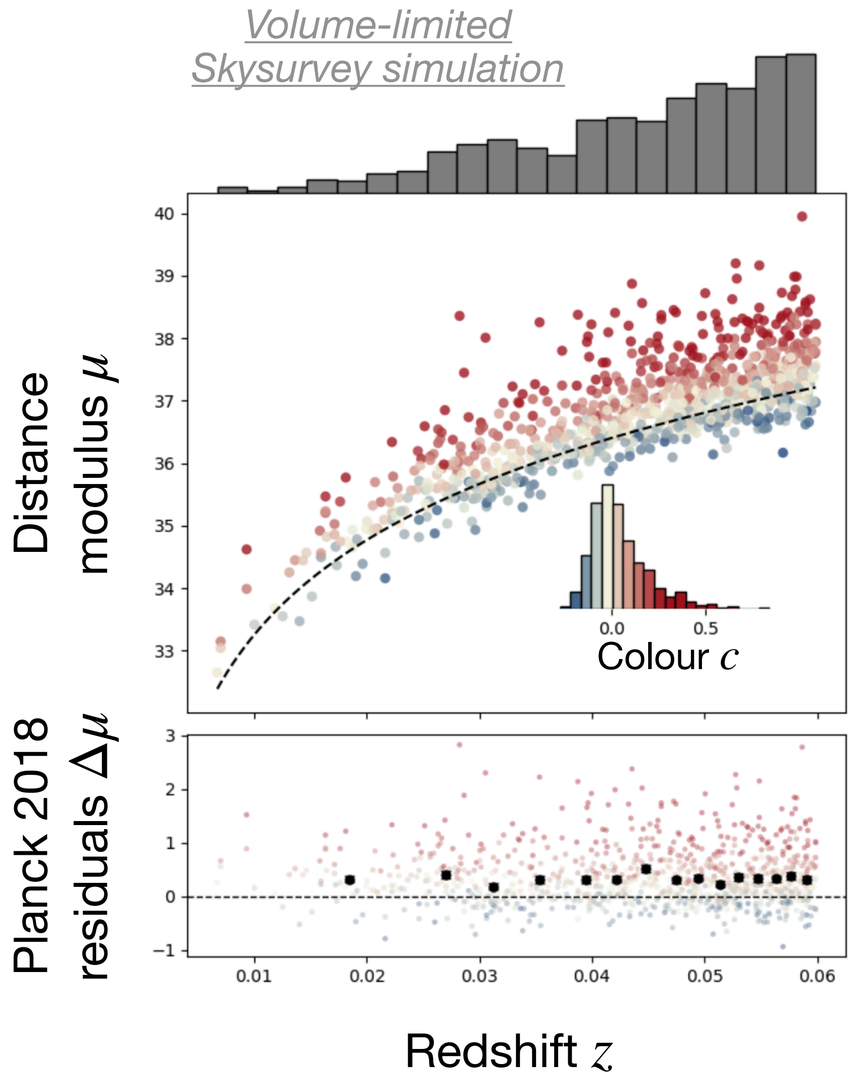
SBI can even use different proxies at once

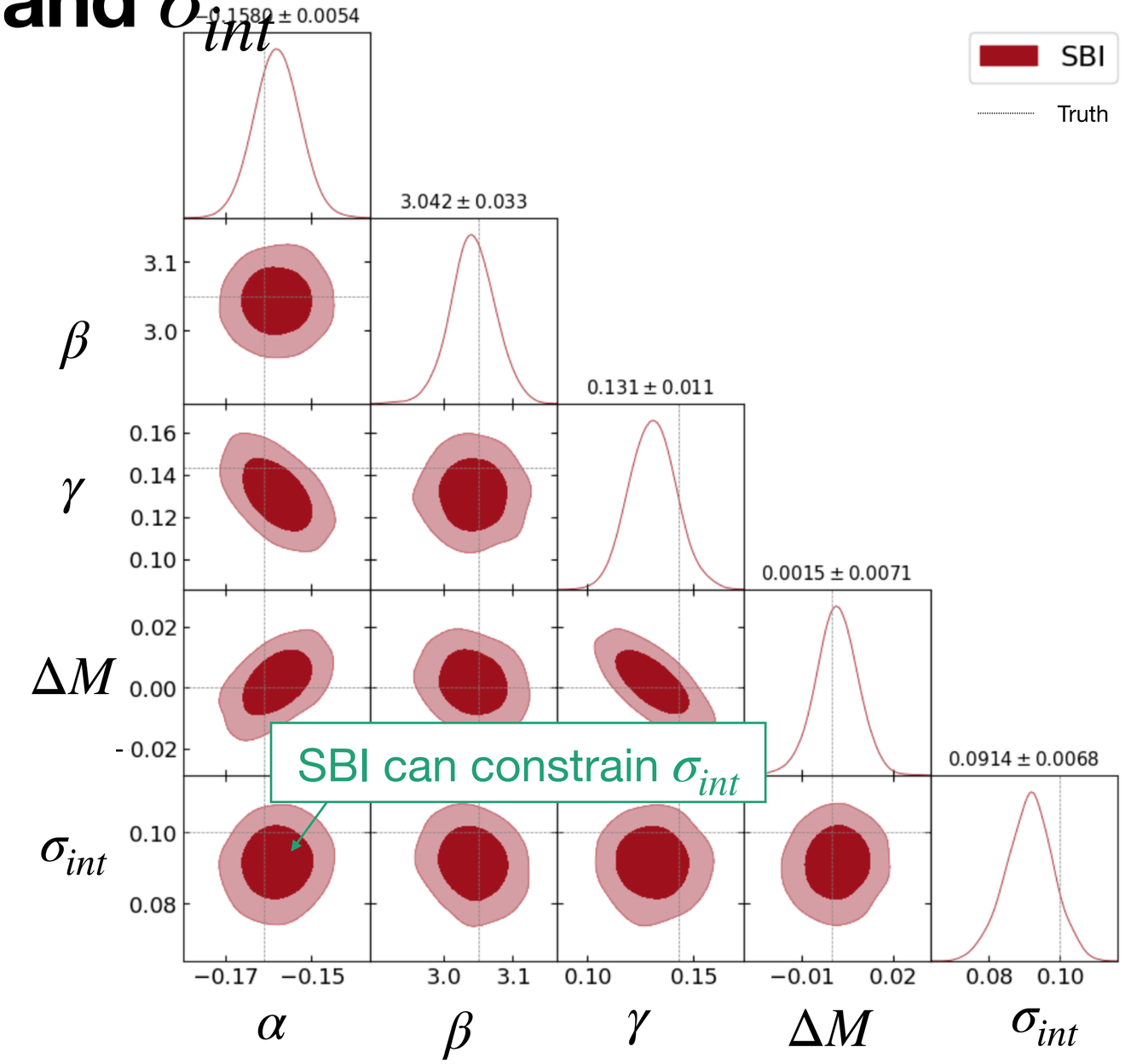
BUT the result is dependent on the true physics behind

Results

1. α and β
2. α, β and γ
3. α, β and γ with proxy (local colour)
4. α, β, γ (proxy) and σ_{int}
5. α, β, γ and σ_{int} with magnitude selection effect

Results: α, β, γ (proxy) and σ_{int}

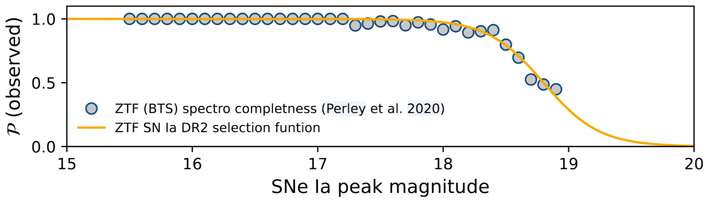




Results

1. α and β
2. α, β and γ
3. α, β and γ with proxy (local colour)
4. α, β, γ (proxy) and σ_{int}
5. α, β, γ and σ_{int} with magnitude selection effect

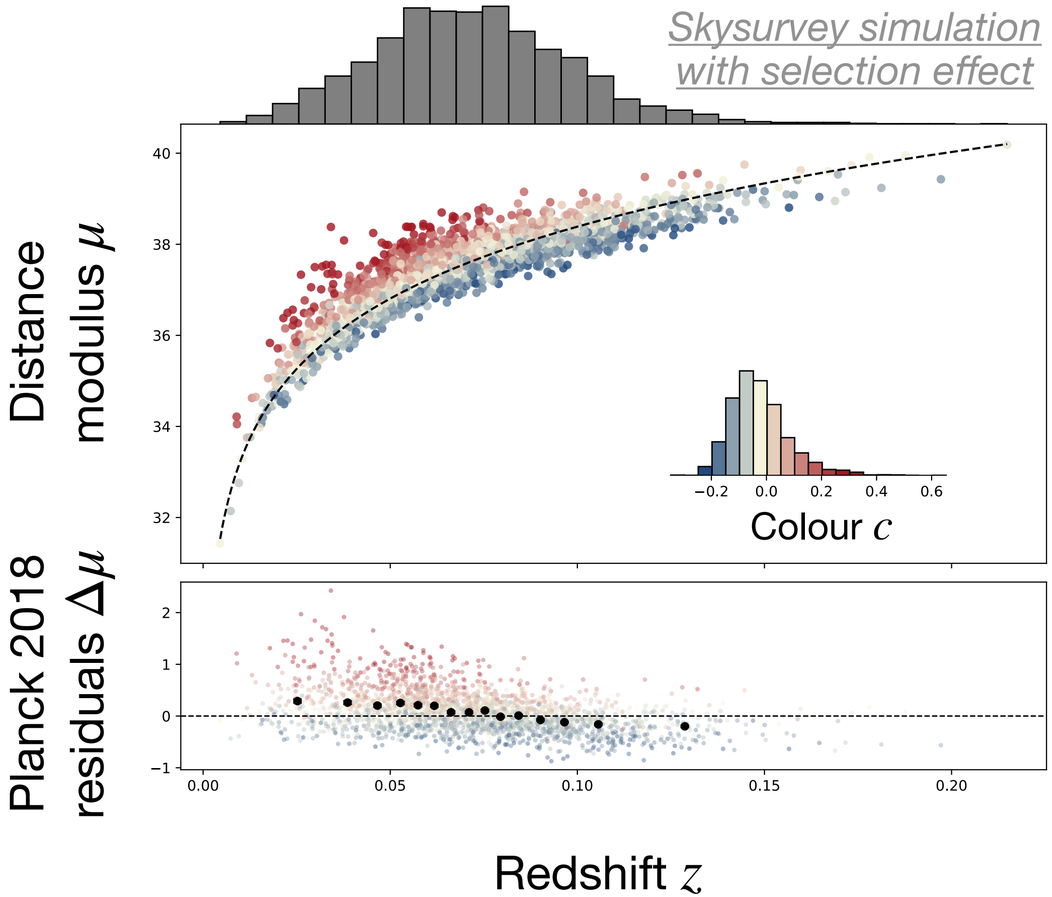
Results: α, β, γ and σ_{int} with magnitude selection effect

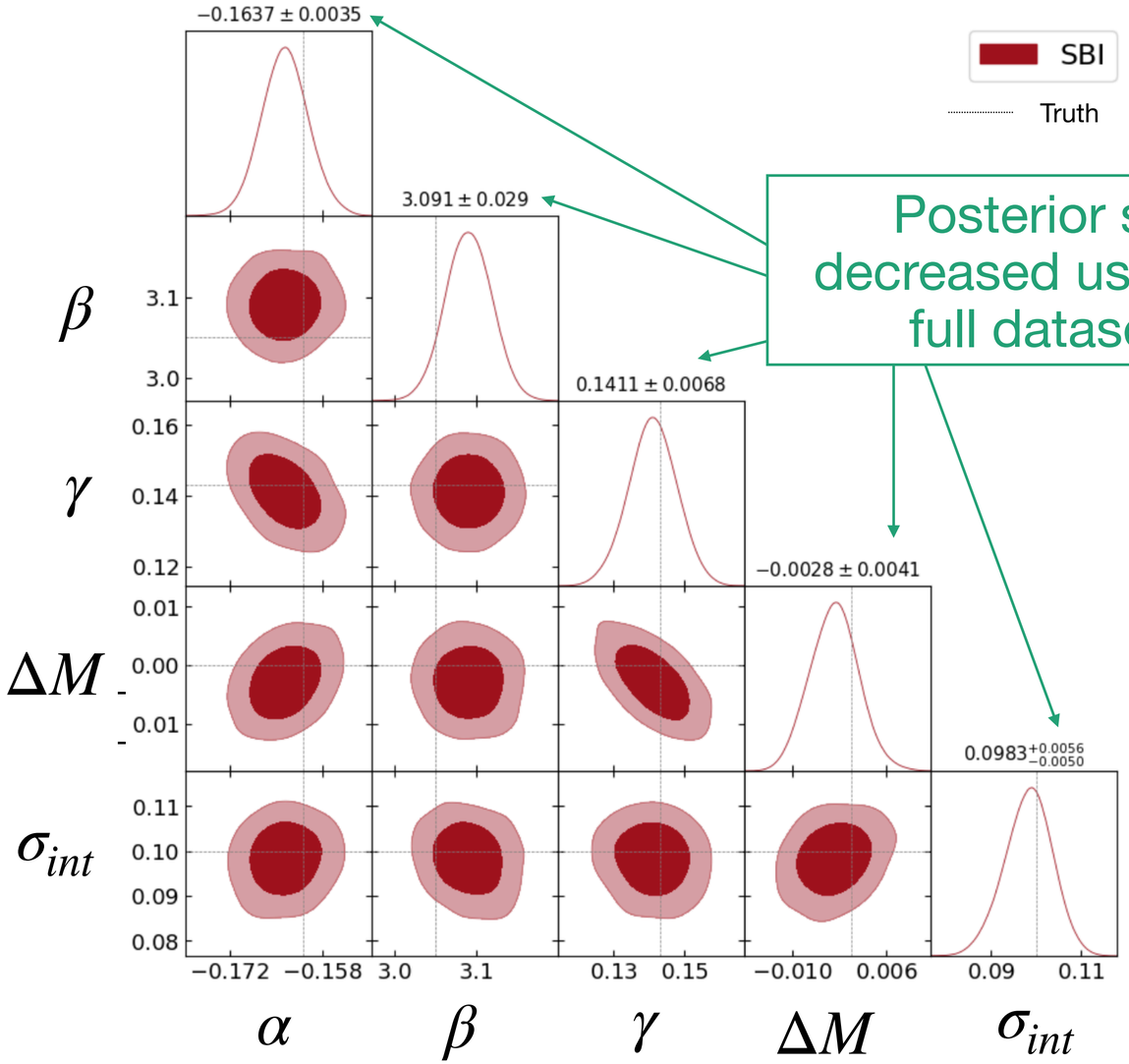


Selection function

Rigault et al. (2025)

Skysurvey simulation with selection effect

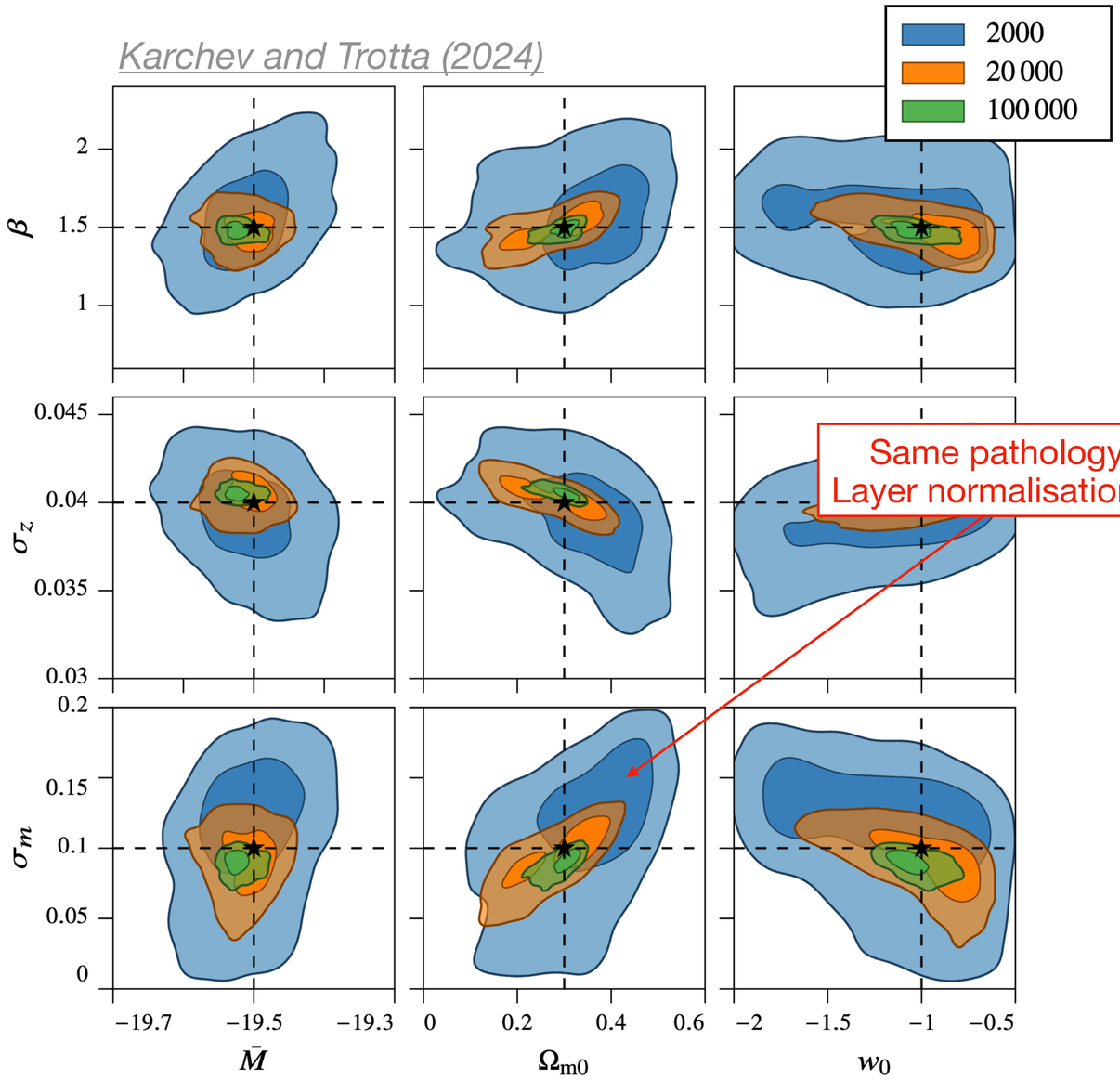




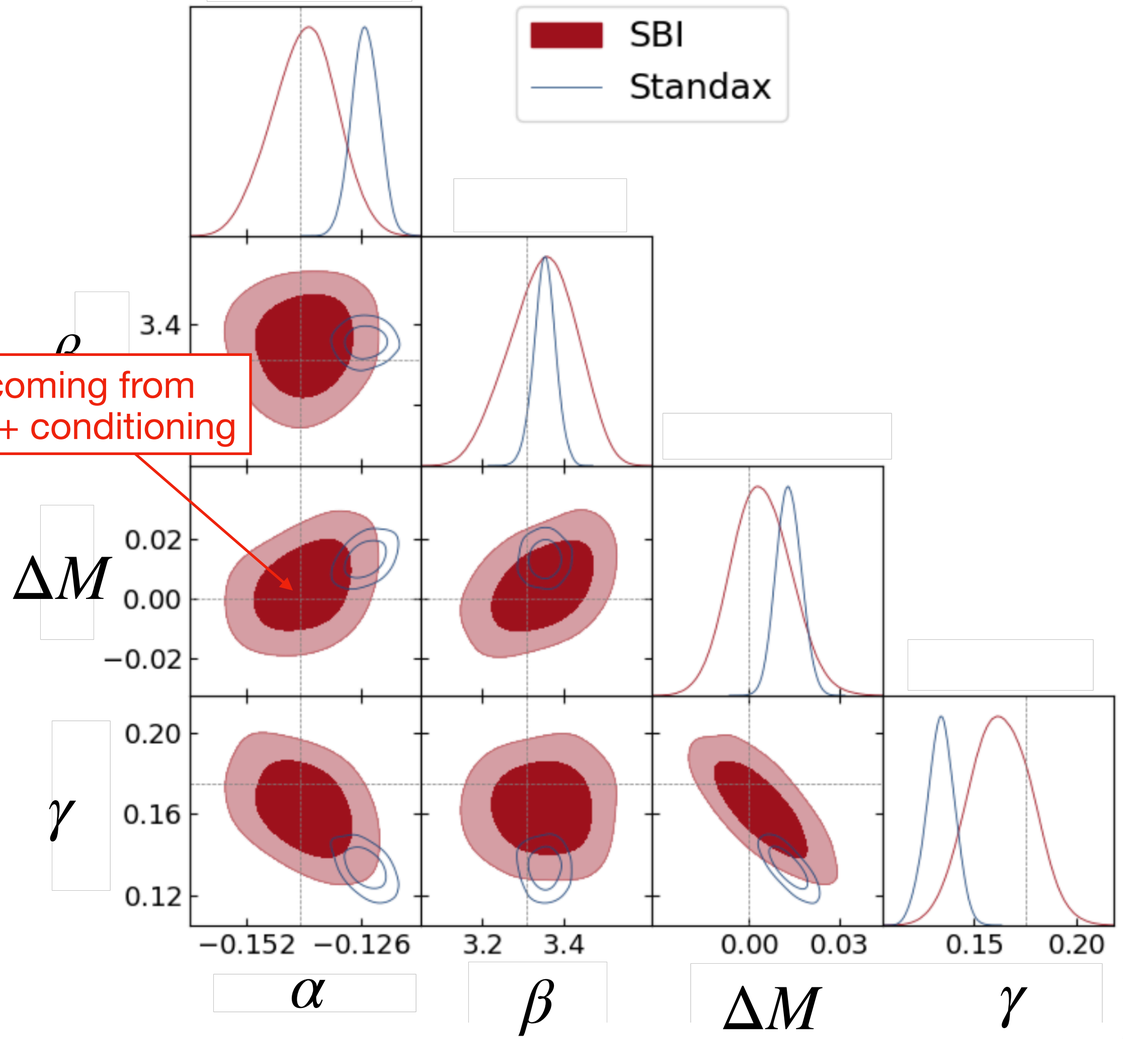
Posterior size decreased using the full dataset !

Comparison with state-of-the-art

Karchev and Trotta (2024)



Previous results



Same pathology coming from
Layer normalisation + conditioning

Conclusions and perspectives

-
- Apply standardisation on DR2 to compare with Ginolin 2025b
- New standardisation on the complete DR2 sample with selection effects
- Cosmological inference on DR2.5 to compare with EDRIS ?

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

Joint posterior proof

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

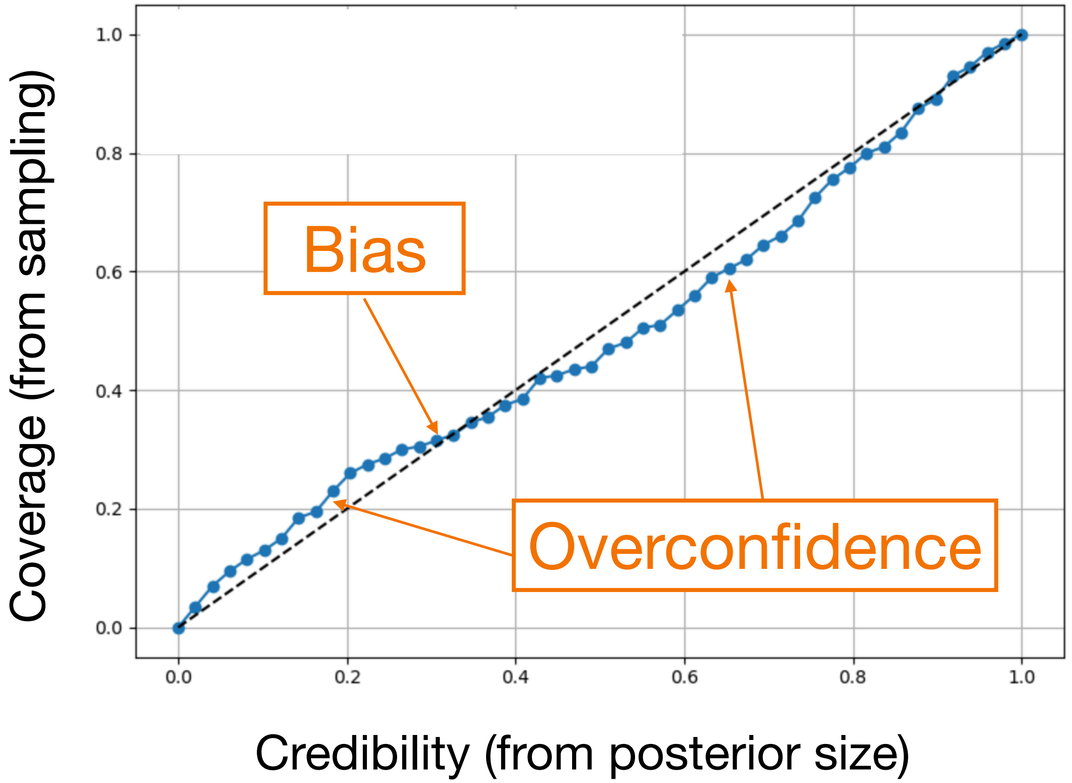
Deepset

Zaheer et al. (2018)

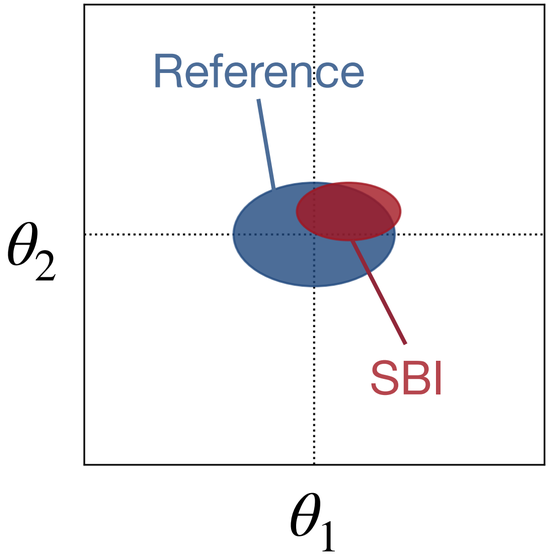
$$f(X) = \rho\left(\sum_{x \in X} \phi(X_i)\right)$$

There exist ϕ and ρ such that
f is permutation invariant

Coverage test *Lemos et al. (2023)*

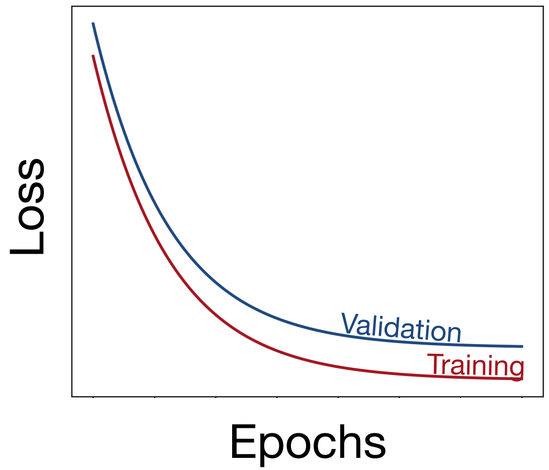


=

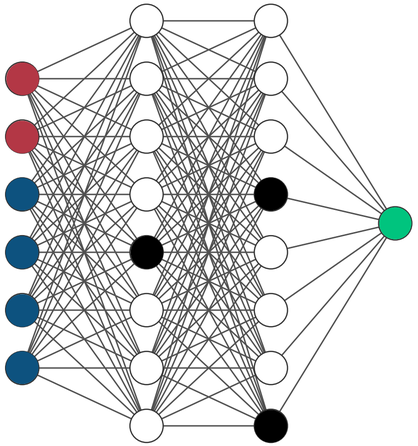


Regularisation

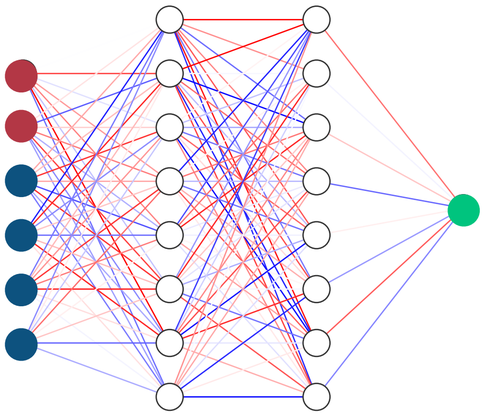
Early stopping



Dropout

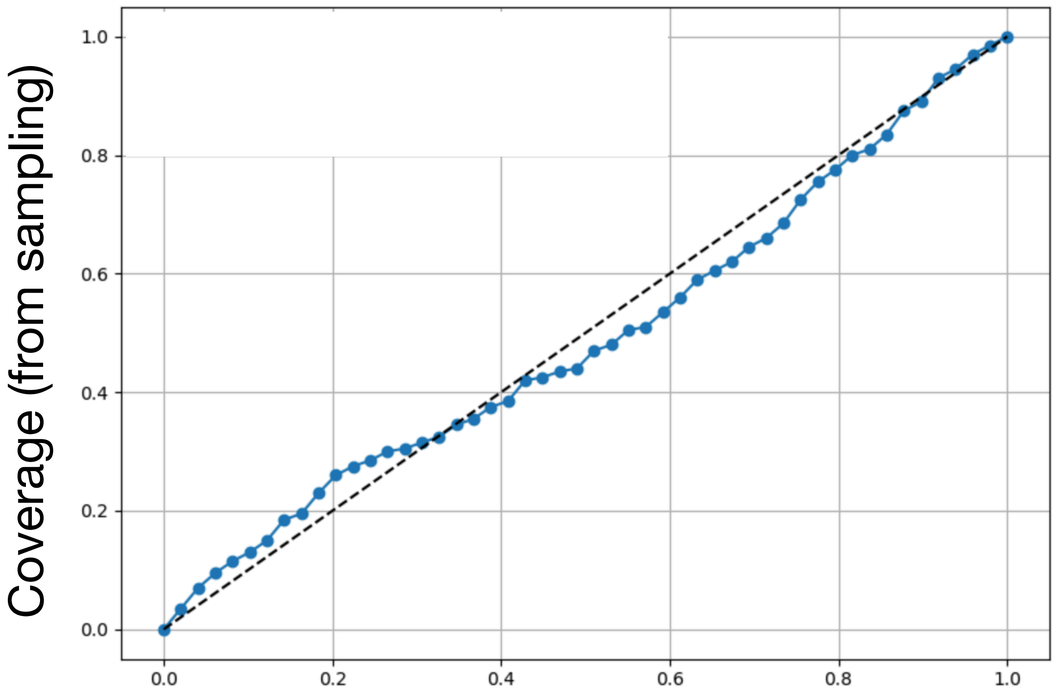


Weight decay

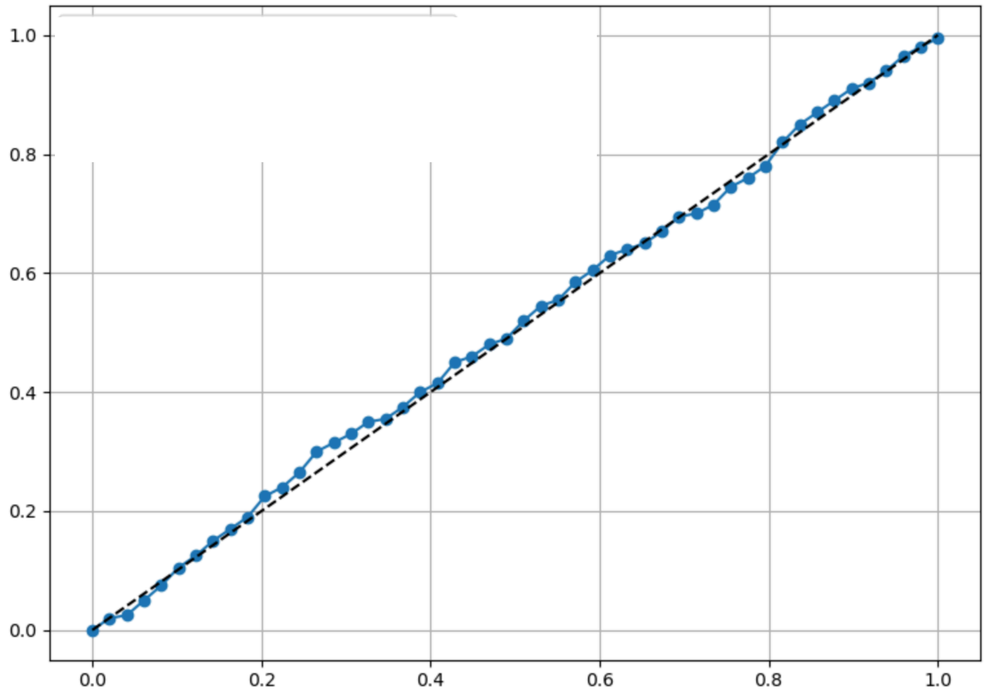


Coverage test

Before regularisation



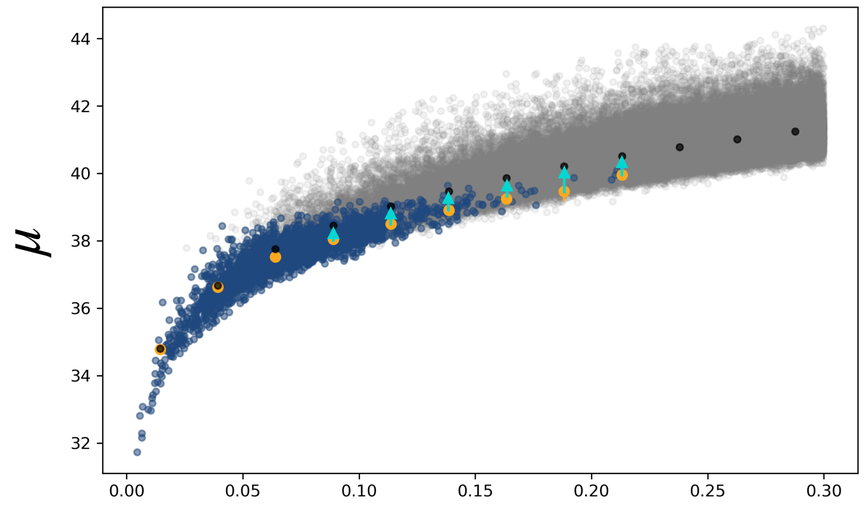
After regularisation

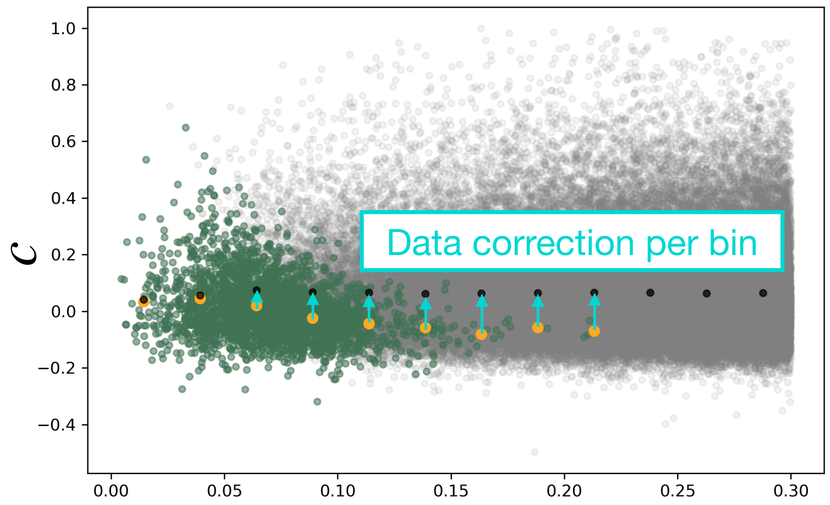


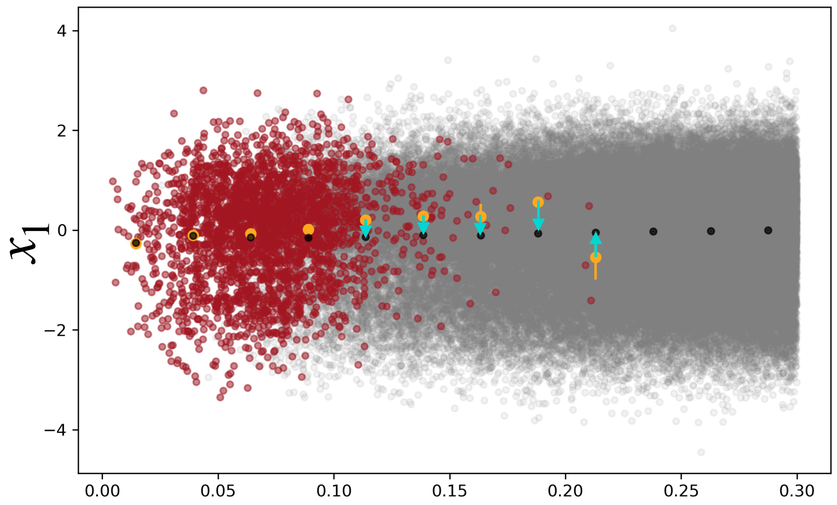
Kessler and Scolnic 2017

BEAMS with Bias Correction (BBC)

Skysurvey simulation



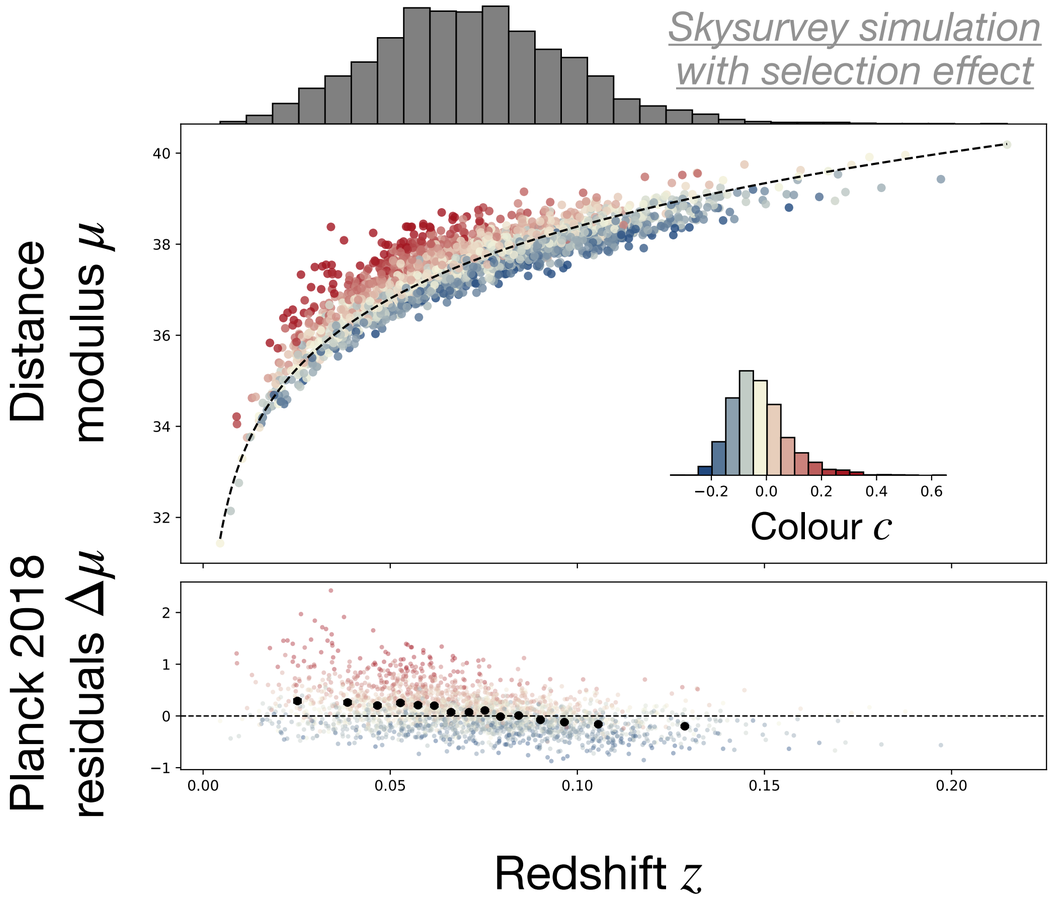




Standardisation : $M = (M_0 - \bar{\delta}_{M_0}) + \beta(c - \bar{\delta}_c) - \alpha(x_1 - \bar{\delta}_{x_1}) + \gamma p$

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

Results: α, β, γ and σ_{int} with magnitude selection effect



β

ΔM

σ_{int}

γ

SBI is not biased by magnitude selection effect

Placeholder: Corner plot

α β ΔM σ_{int} γ