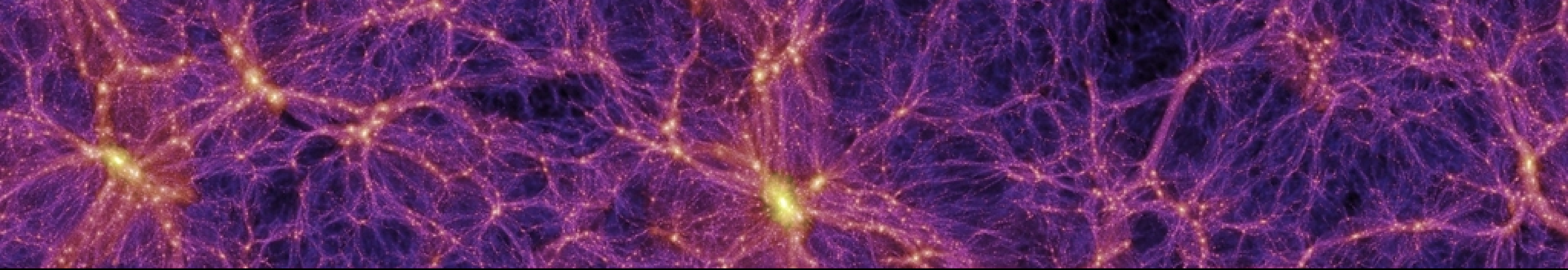


EFTofLSS meets simulation-based inference: σ_8 from biased tracers

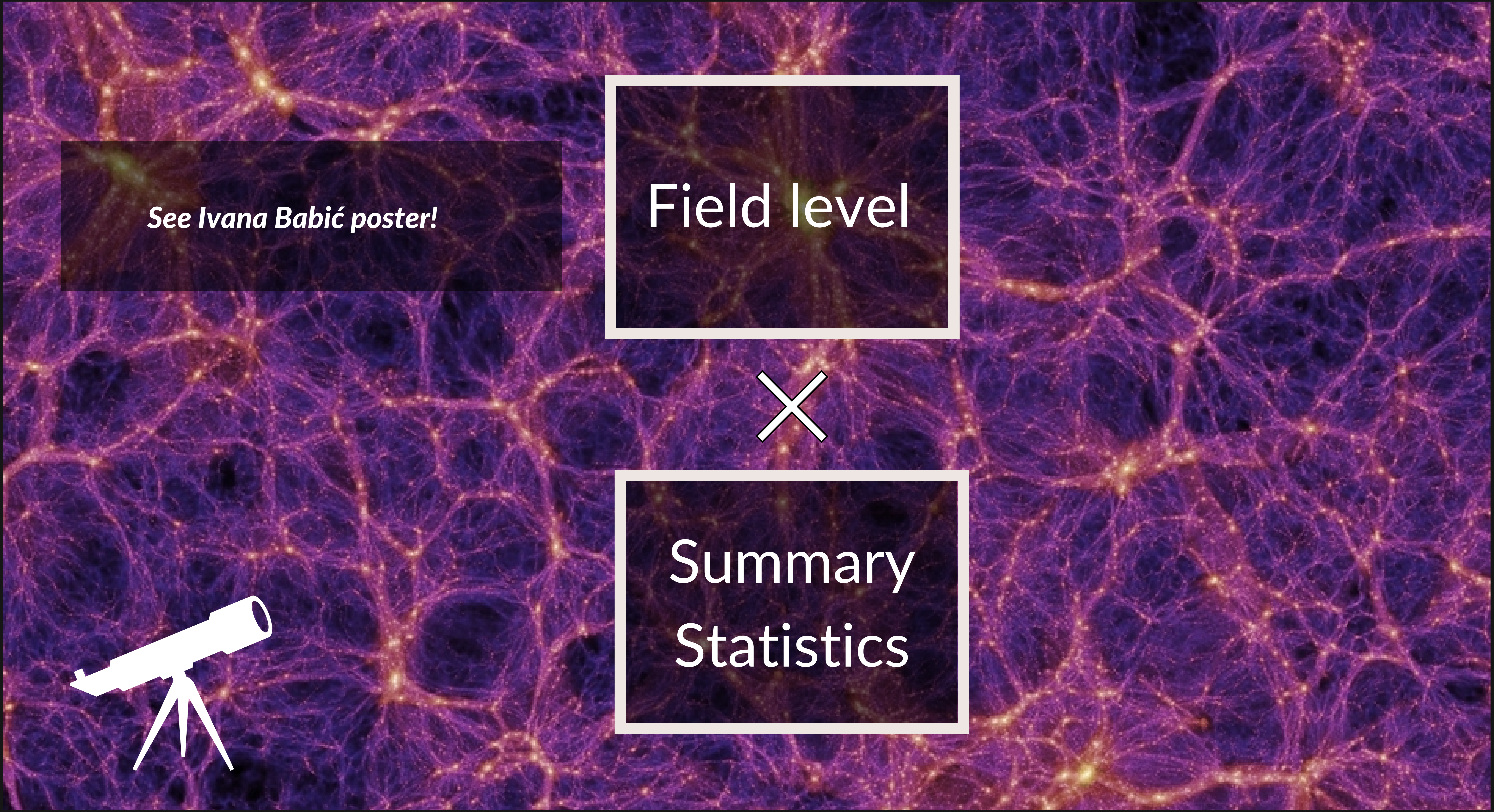
Beatriz Tucci

In collaboration with: Fabian Schmidt

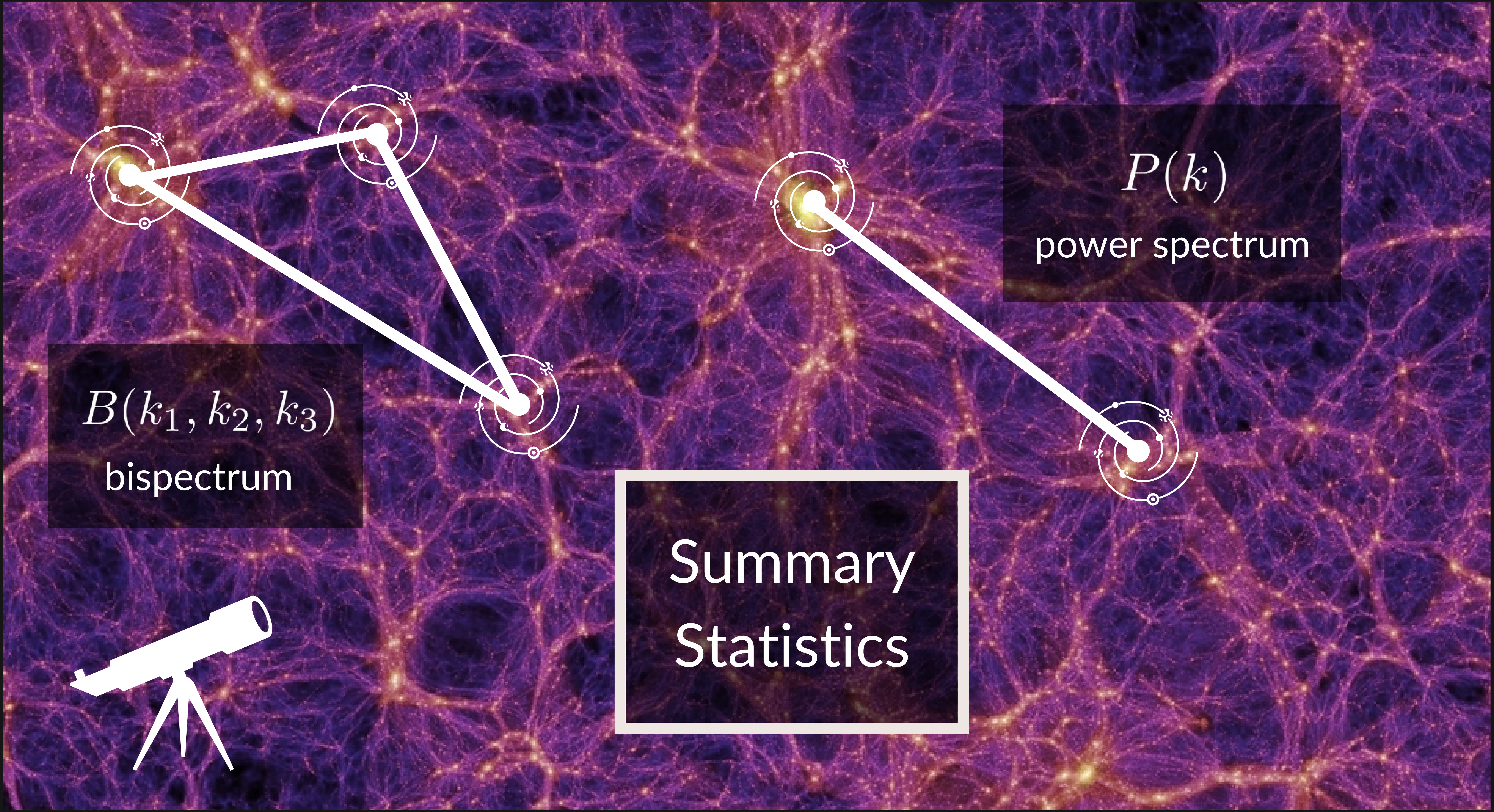
Based on arXiv:2310.03741

LEFTfield

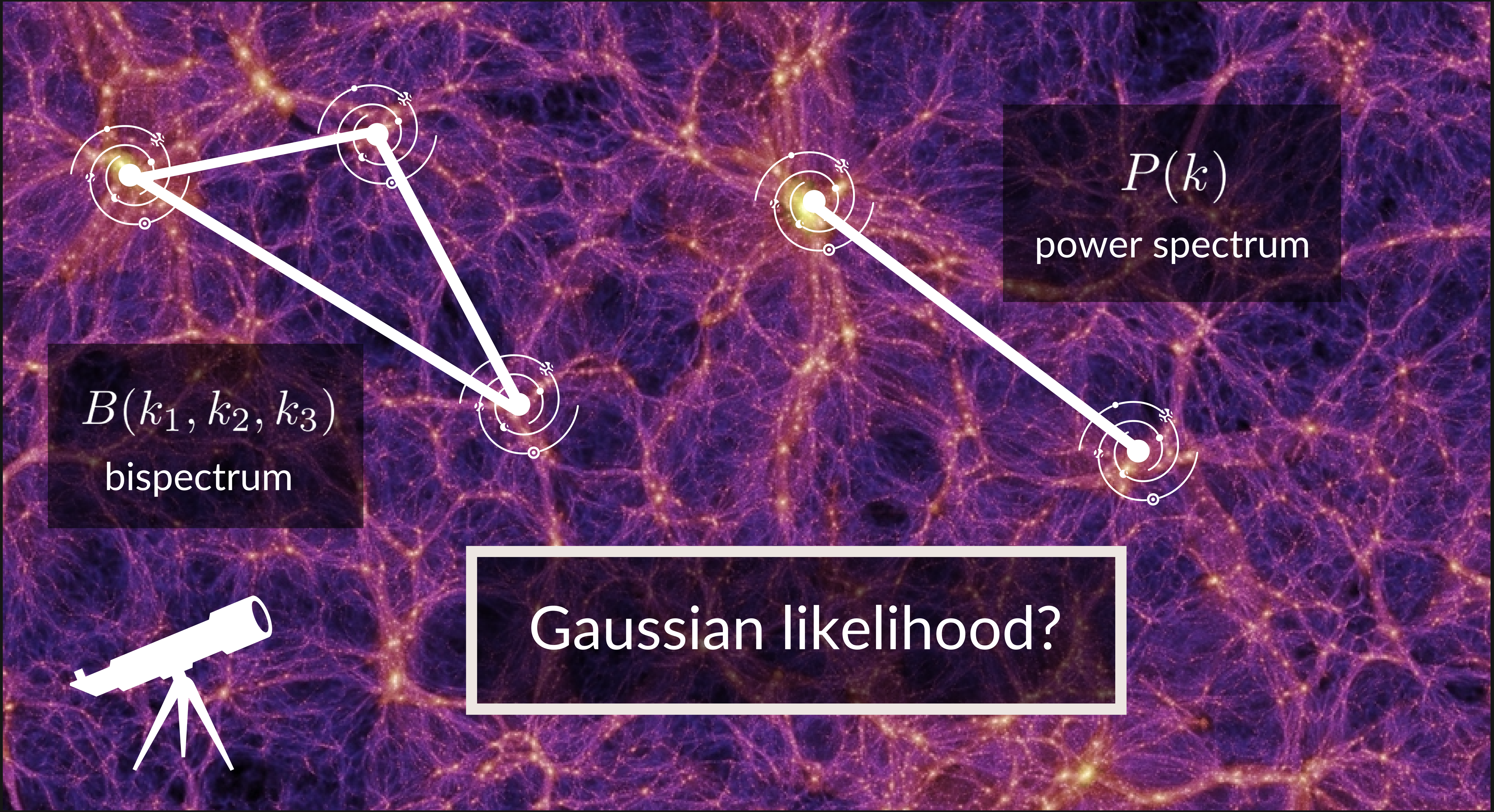
MAX-PLANCK-INSTITUT FÜR ASTROPHYSIK



The Millennium Simulation Project

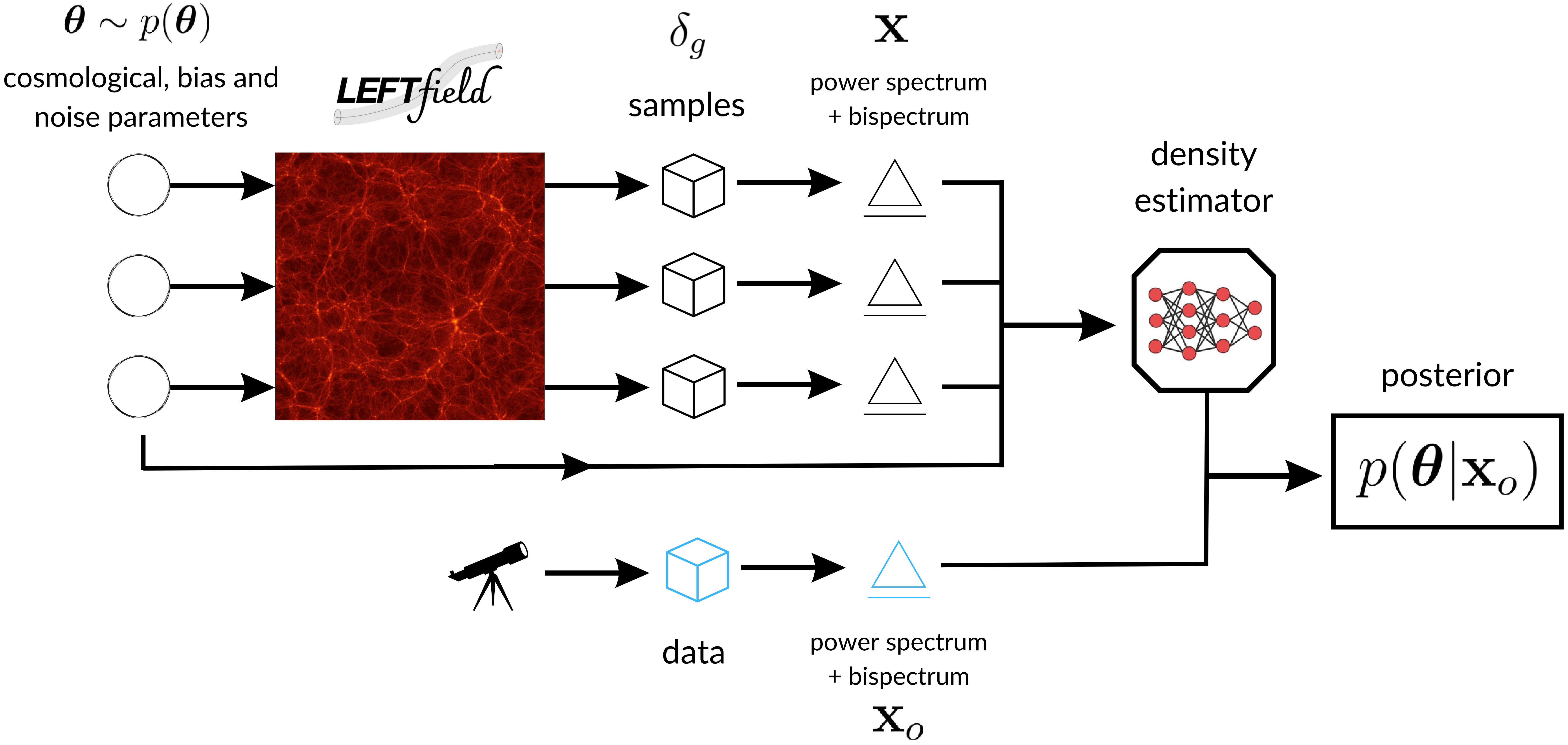


The Millennium Simulation Project

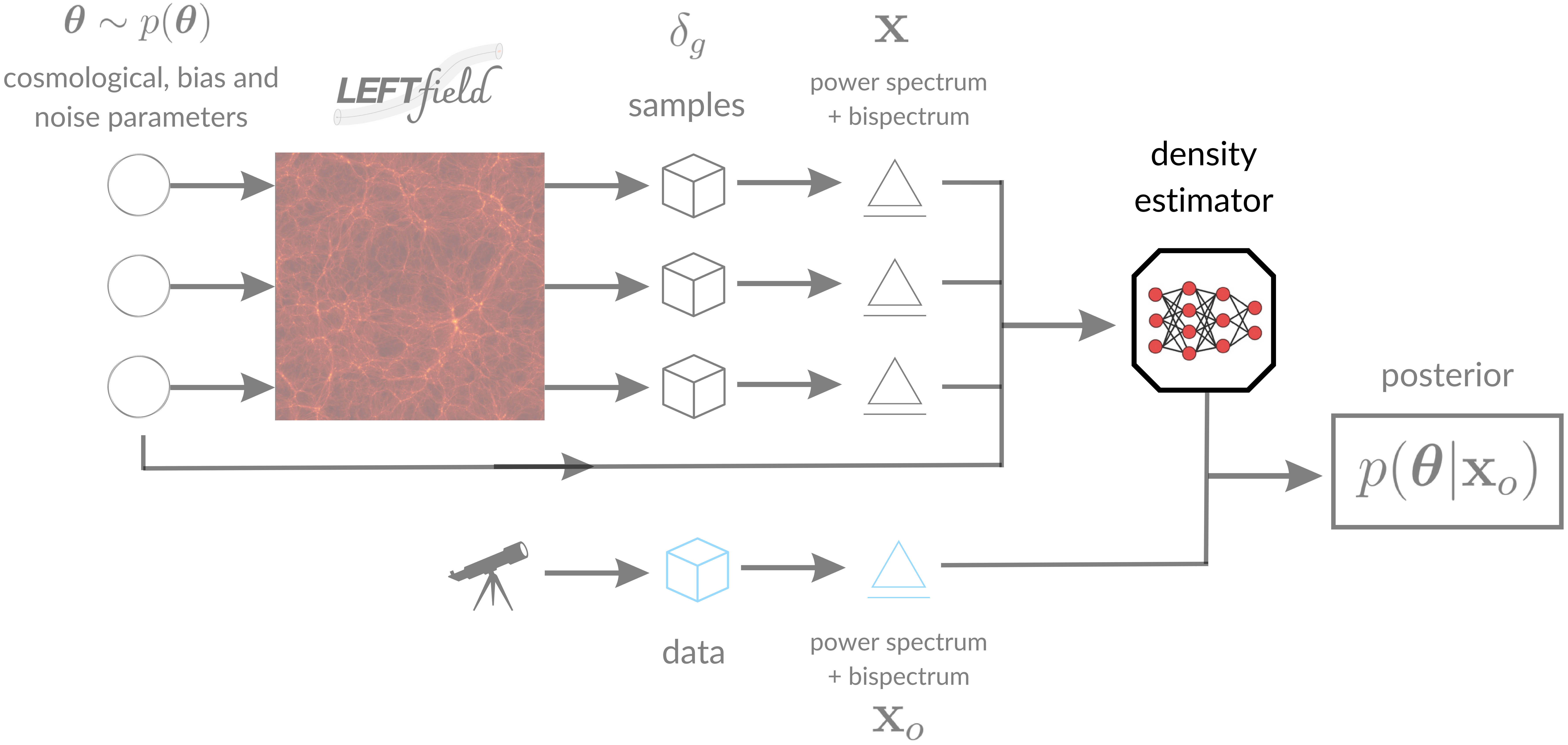


The Millennium Simulation Project

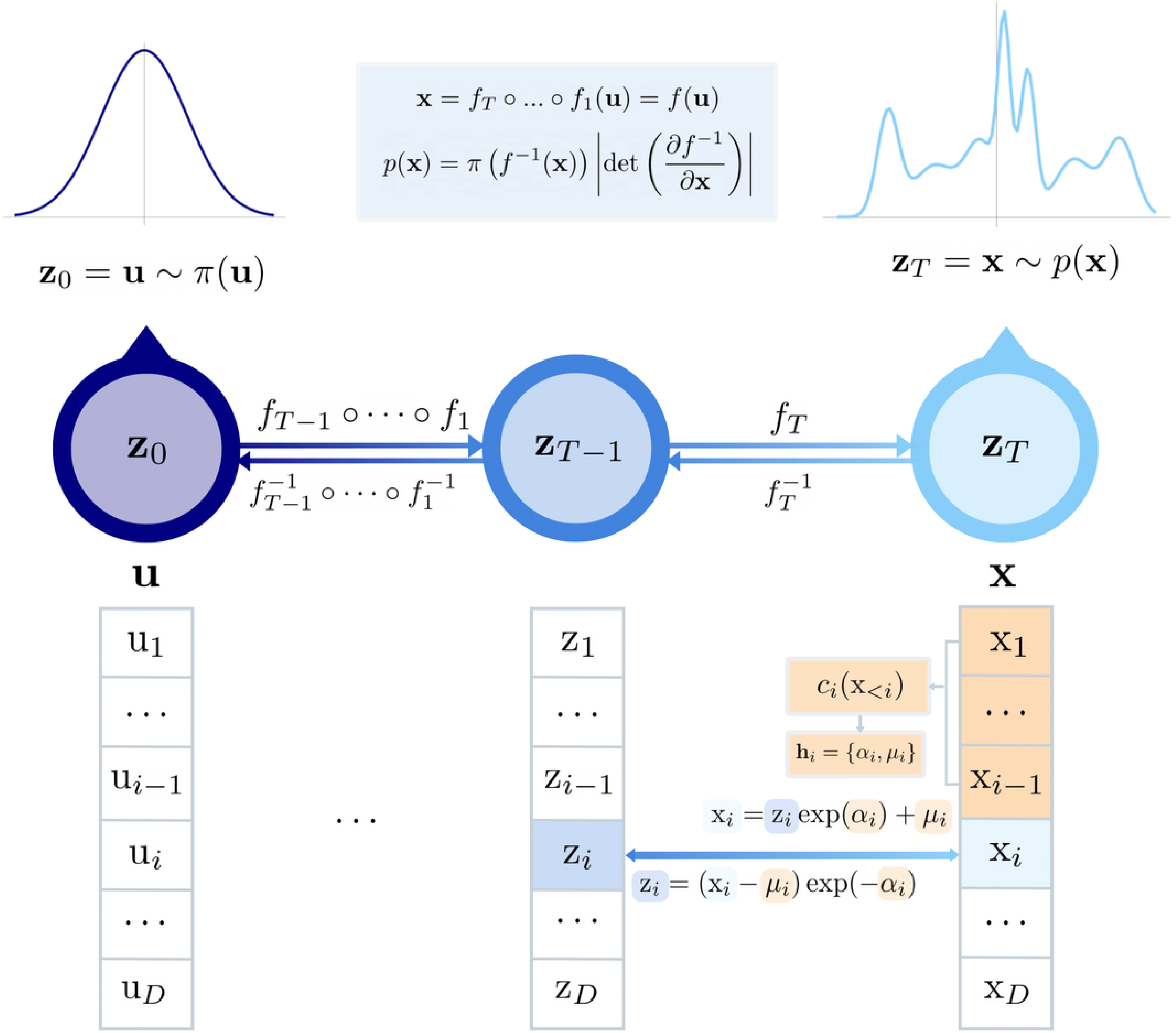
Simulation-based inference



Simulation-based inference



Normalizing Flows



Tucci, Schmidt (2023)

Neural Density Estimators:

Neural Posterior Estimation (NPE)

Neural Likelihood Estimation (NLE)

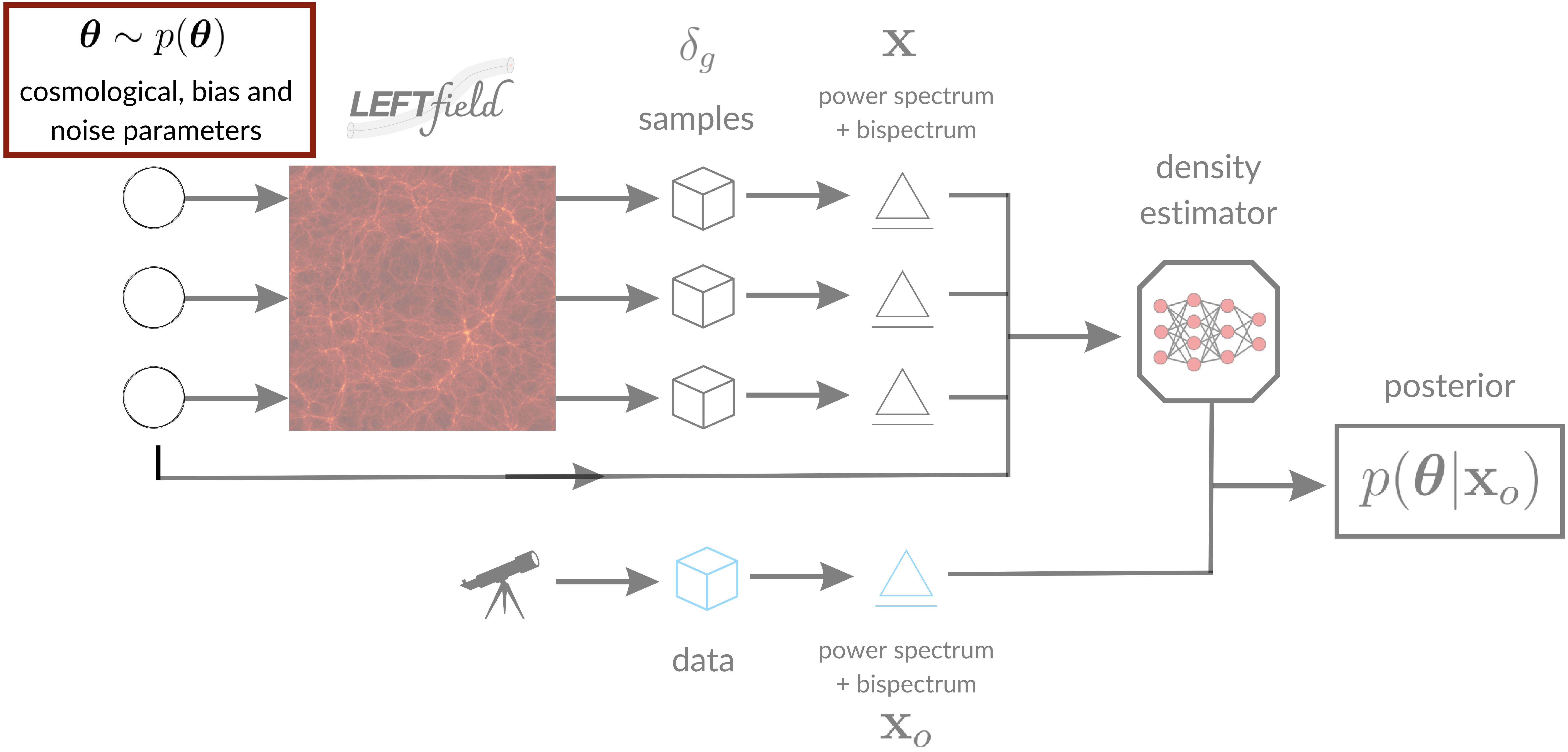
sbi: A toolkit for simulation-based inference

Tejero-Cantero et al. (2020)

Masked Autorregressive Flows (MAFs)

- Easy to evaluate: triangular Jacobian
- Expressive: composition of transforms

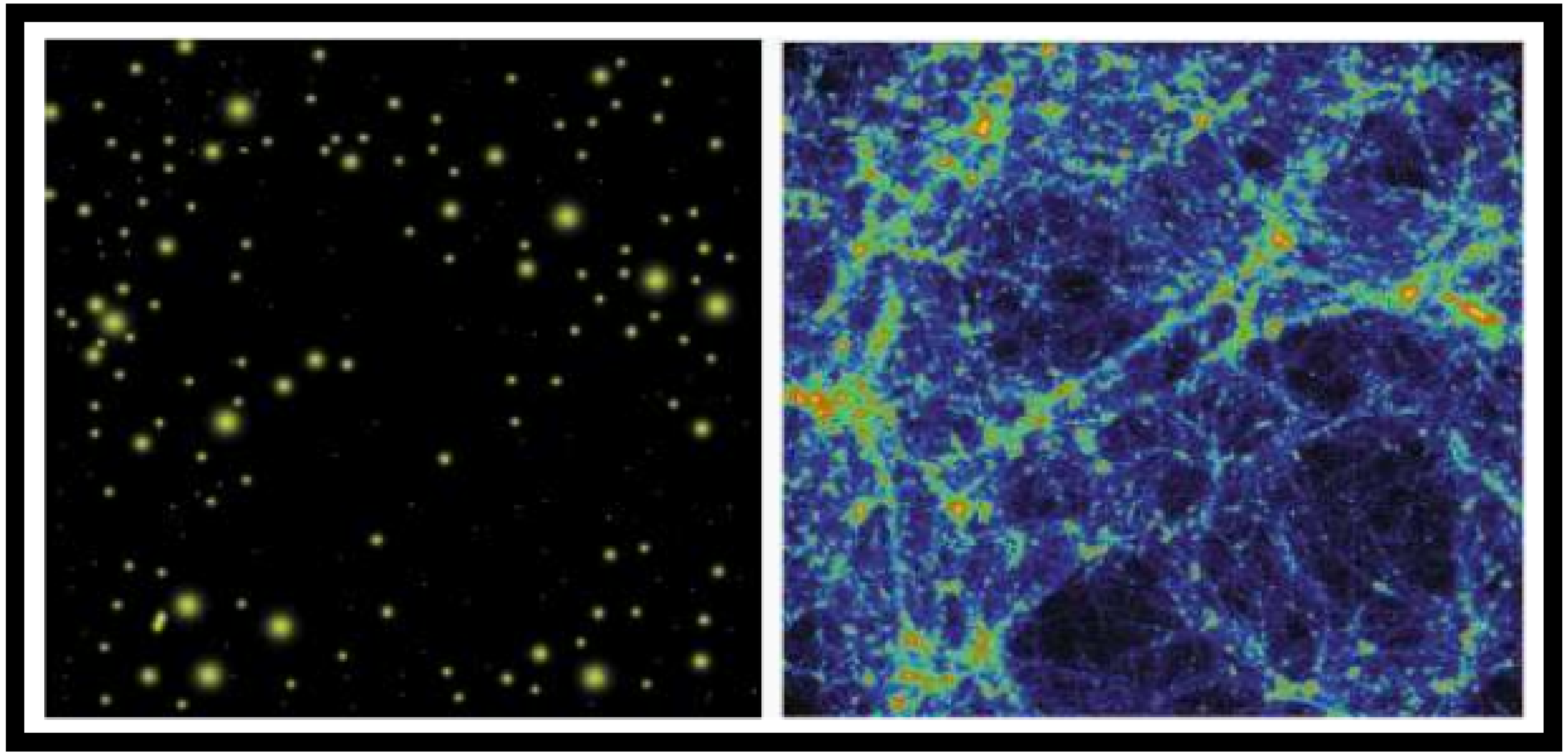
Simulation-based inference



The bias expansion

Cooray & Sheth (2002)

Cosmological tracers

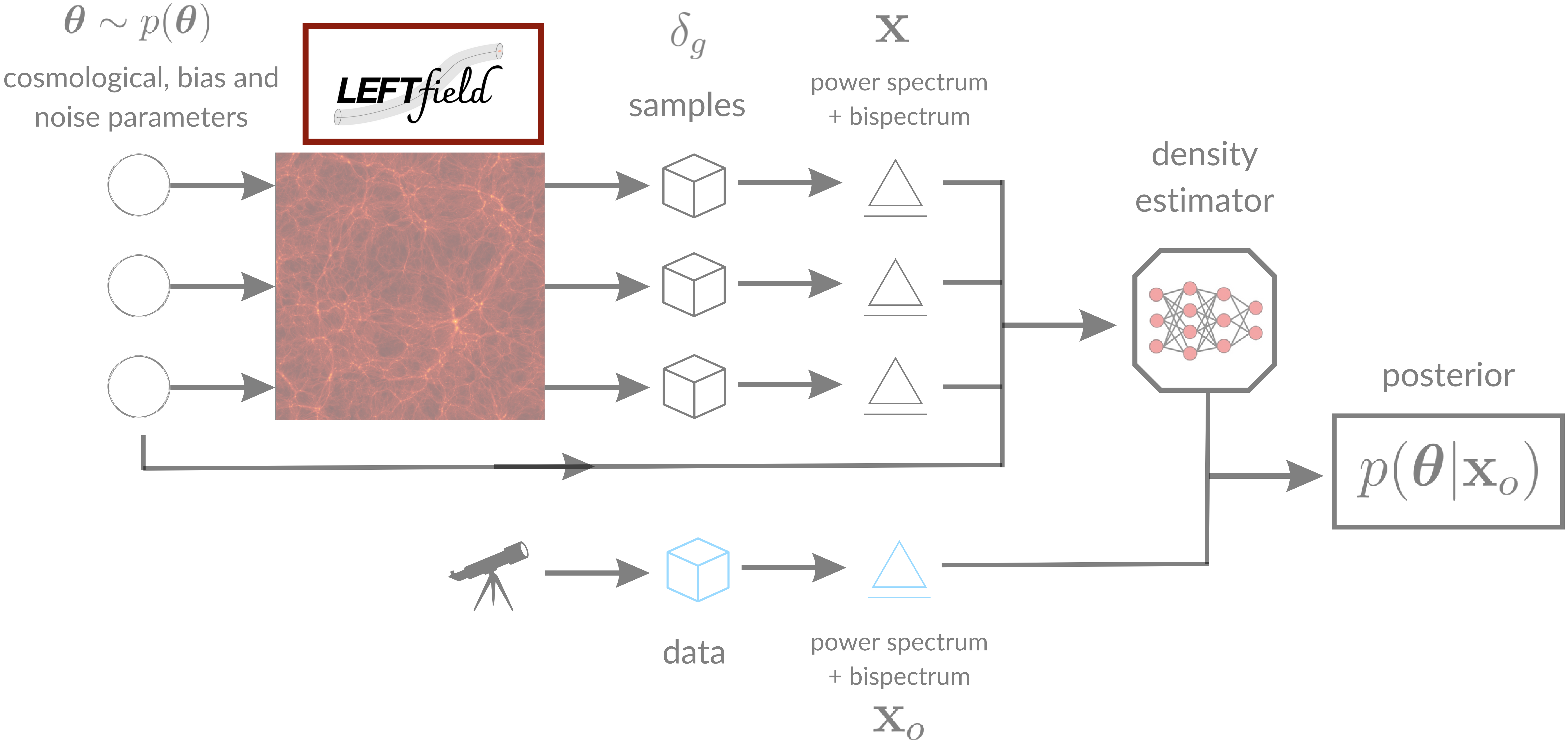


Matter distribution

$$\delta_g(\mathbf{x},\tau) = \sum_{\mathcal{O}} b_{\mathcal{O}}(\tau)\mathcal{O}(\mathbf{x},\tau) + \varepsilon(\mathbf{x},\tau) + \sum_{\mathcal{O}} \varepsilon_{\mathcal{O}}(\mathbf{x},\tau)\mathcal{O}(\mathbf{x},\tau)$$

For a review, see:
Desjacques, Jeong
& Schmidt (2016)

Simulation-based inference

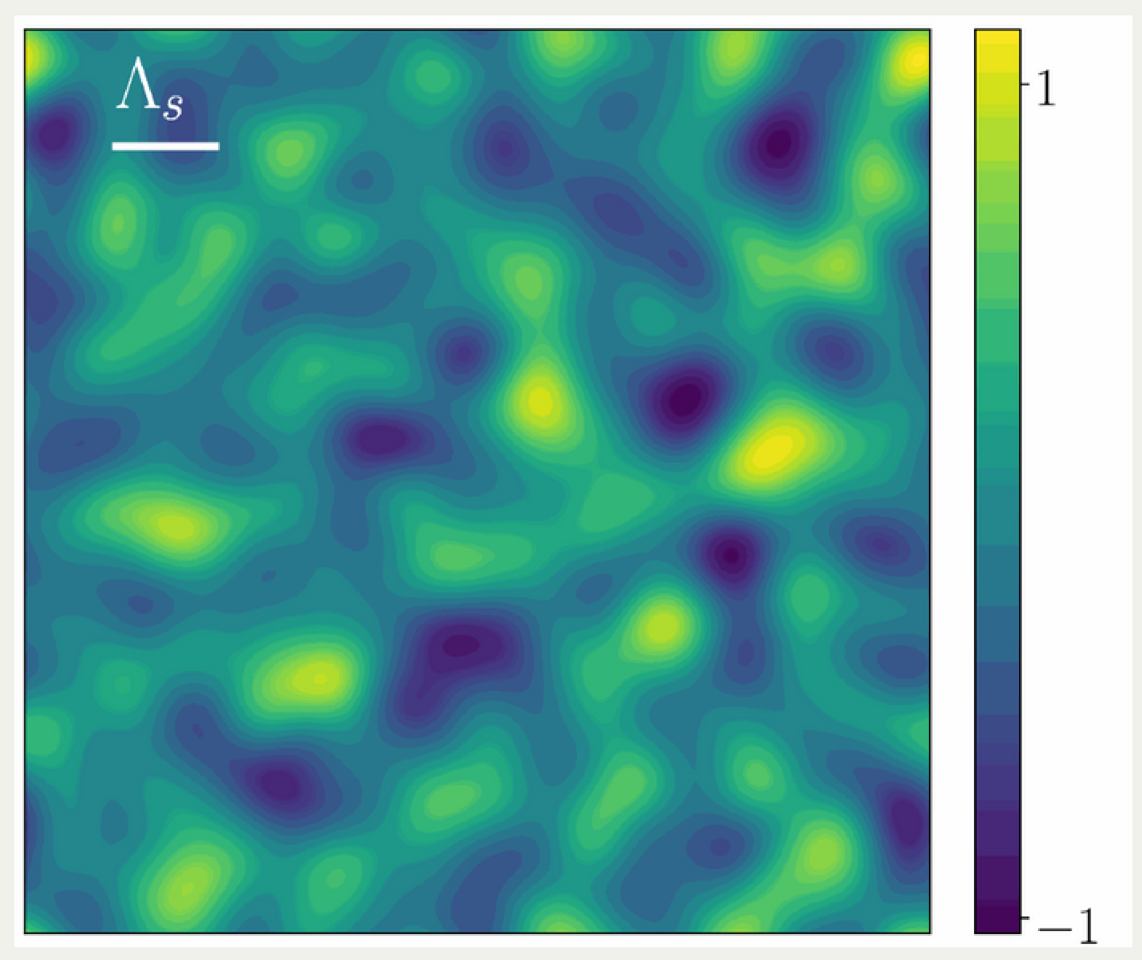


Forward model

$$\alpha \equiv \sigma_8/\sigma_8^{\mathrm{fid}} \qquad \hat{s}(\boldsymbol{x}) \sim \mathcal{N}(0,1)$$

$$\delta_\Lambda^{(1)}(\boldsymbol{k}, z) = W_\Lambda(k)\sqrt{\alpha^2 P_L(k,z)}\,\hat{s}(\boldsymbol{k})$$

An EFTofLSS based forward model



Borrowed from Pierre Zhang

"coarse-graining"

An n-th order Lagrangian Forward Model for Large-Scale Structure
Fabian Schmidt (2021)

Forward model

$\alpha \equiv \sigma_8/\sigma_8^{\text{fid}}$ $\hat{s}(\boldsymbol{x}) \sim \mathcal{N}(0,1)$

$$\delta_\Lambda^{(1)}(\boldsymbol{k}, z) = W_\Lambda(k)\sqrt{\alpha^2 P_L(k,z)}\,\hat{s}(\boldsymbol{k})$$

$1 + \delta(\boldsymbol{x},\tau) = |\mathbf{1} + \boldsymbol{M}(\boldsymbol{q},\tau)|^{-1}$ $M_{ij} \equiv \partial_i s_j$

$$\mathrm{tr}[\boldsymbol{M}_\Lambda^{(1)}] = -\delta_\Lambda^{(1)}$$

Lagrangian Bias Operators

1^{st}	$\mathrm{tr}[\boldsymbol{M}_\Lambda^{(1)}]$
2^{nd}	$\mathrm{tr}[\boldsymbol{M}_\Lambda^{(1)}\boldsymbol{M}_\Lambda^{(1)}]$, $(\mathrm{tr}[\boldsymbol{M}_\Lambda^{(1)}])^2$

LPT recursion relations

$$\boldsymbol{s}^{(n)}$$

Lagrangian **P**erturbation **T**heory

$$\delta_{g,\text{det}}^L(\mathbf{q},\tau) = \sum_{\mathcal{O}^L} b_{\mathcal{O}^L}(\tau)\mathcal{O}^L(\mathbf{q},\tau)$$

$\boldsymbol{x}(\boldsymbol{q},\tau) = \boldsymbol{q} + \boldsymbol{s}(\boldsymbol{q},\tau)$

$\varepsilon \sim \mathcal{N}(0, P_\varepsilon)$

$$\delta_g(\boldsymbol{x},\tau) = \delta_{g,\text{det}}(\boldsymbol{x},\tau) + \varepsilon(\boldsymbol{x},\tau) + c_{\varepsilon\delta}(\tau)\varepsilon(\boldsymbol{x},\tau)\delta(\boldsymbol{x},\tau) + c_{\varepsilon^2}(\tau)\varepsilon^2(\boldsymbol{x},\tau)$$

An n-th order Lagrangian Forward Model for Large-Scale Structure
Fabian Schmidt (2021)

Our main goals

Considering an Euclid-like mock survey, we want to answer:

- *Does the non-Gaussianity of the power-spectrum and bispectrum distributions at low-k affect cosmological inference?*

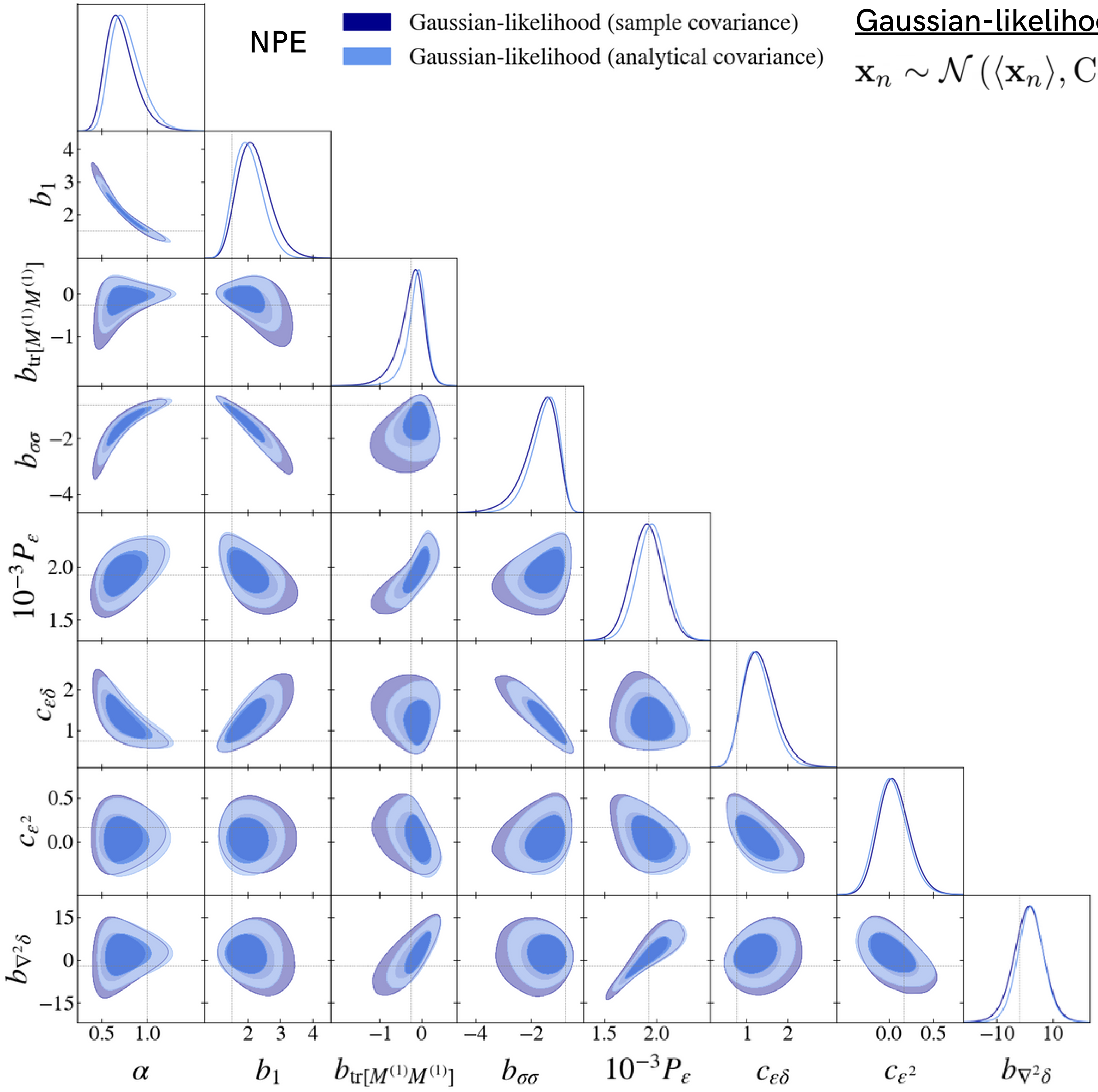
- *How many simulations are needed for posterior estimation?*

Cosmological inference | Euclid configuration

$N_{\rm sim} = 10^5$

$k_{\rm max} = \Lambda = 0.1 h{\rm Mpc}^{-1}$

$D = N_{\rm bin} + N_{\rm tri} = 33$



Gaussian-likelihood

$\mathbf{x}_n \sim \mathcal{N}(\langle \mathbf{x}_n \rangle, {\rm Cov}[\mathbf{x}_o])$

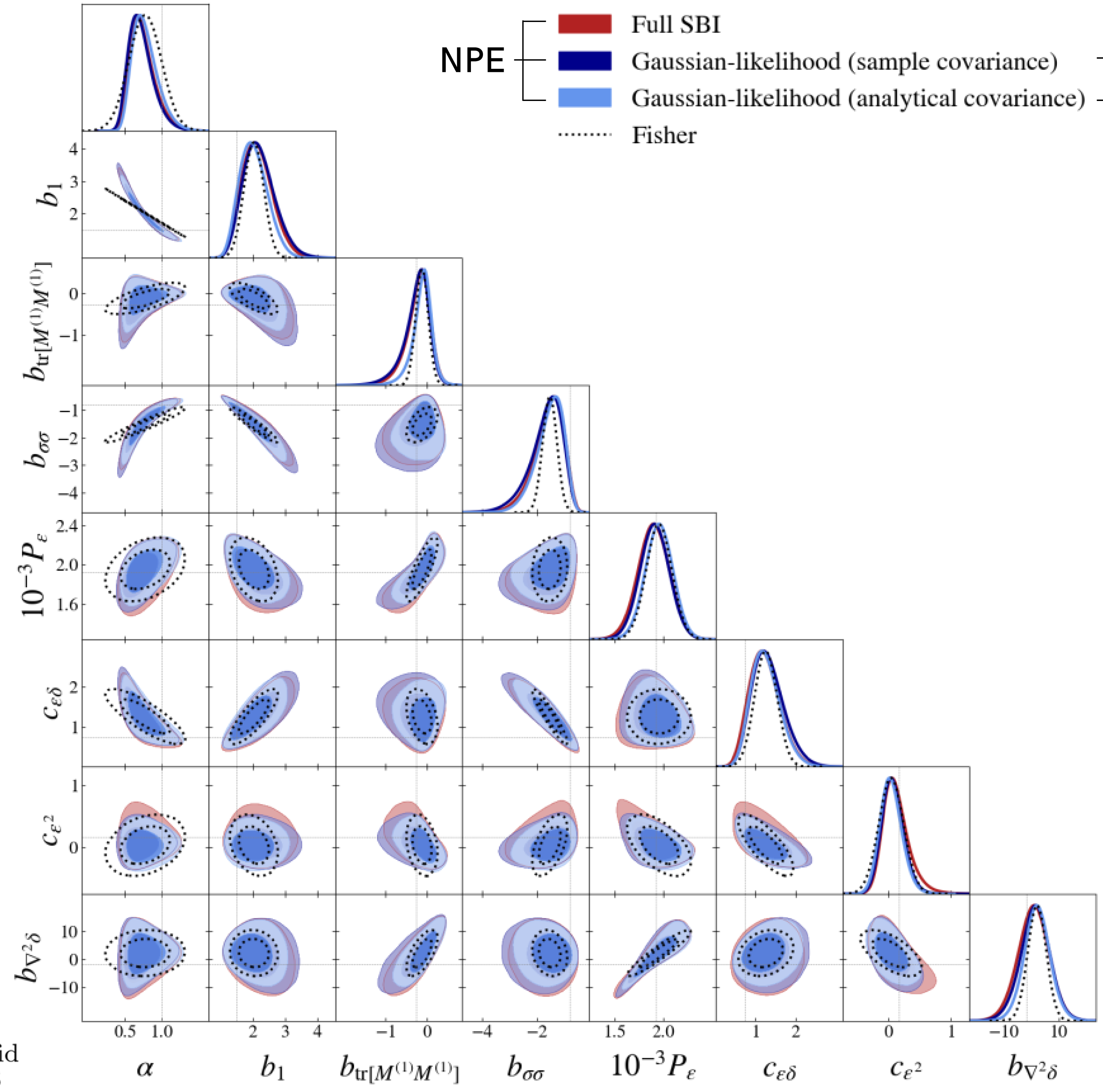
$\alpha \equiv \sigma_8/\sigma_8^{\rm fid}$

Cosmological inference | Euclid configuration

$N_{\rm sim} = 10^5$

$k_{\rm max} = \Lambda = 0.1 h{\rm Mpc}^{-1}$

$D = N_{\rm bin} + N_{\rm tri} = 33$



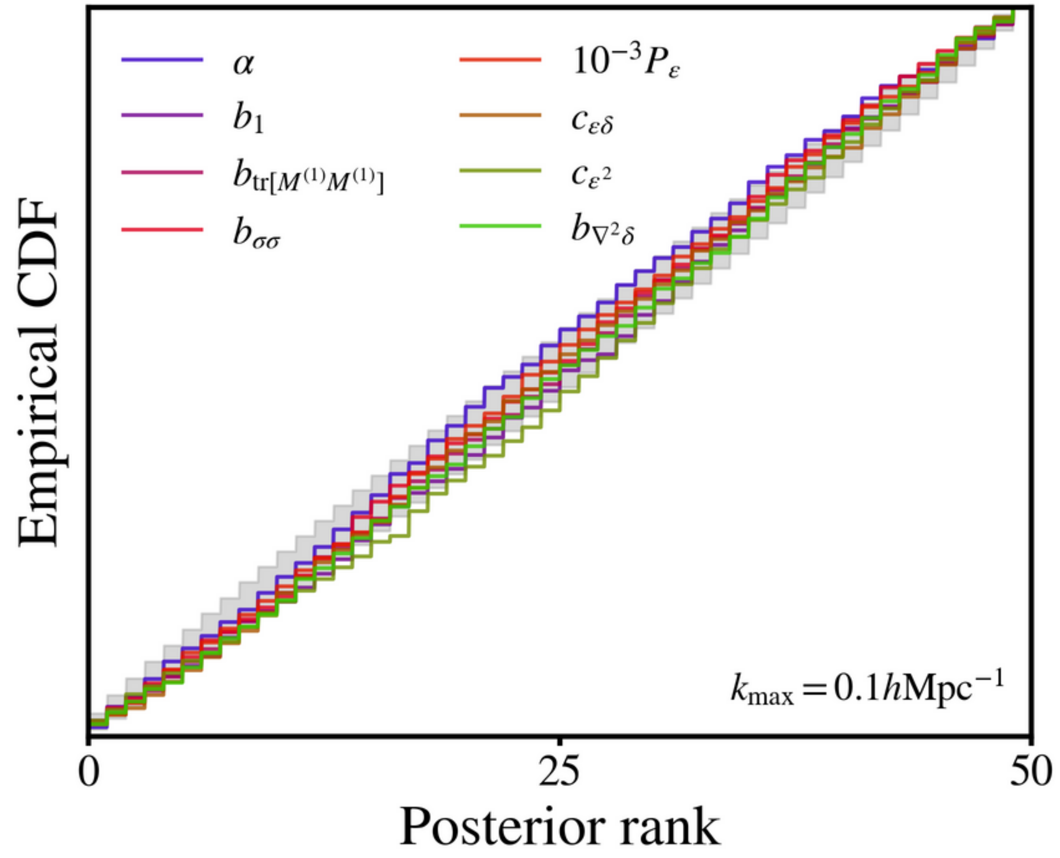
$\alpha \equiv \sigma_8/\sigma_8^{\rm fid}$

<u>Gaussian-likelihood</u>

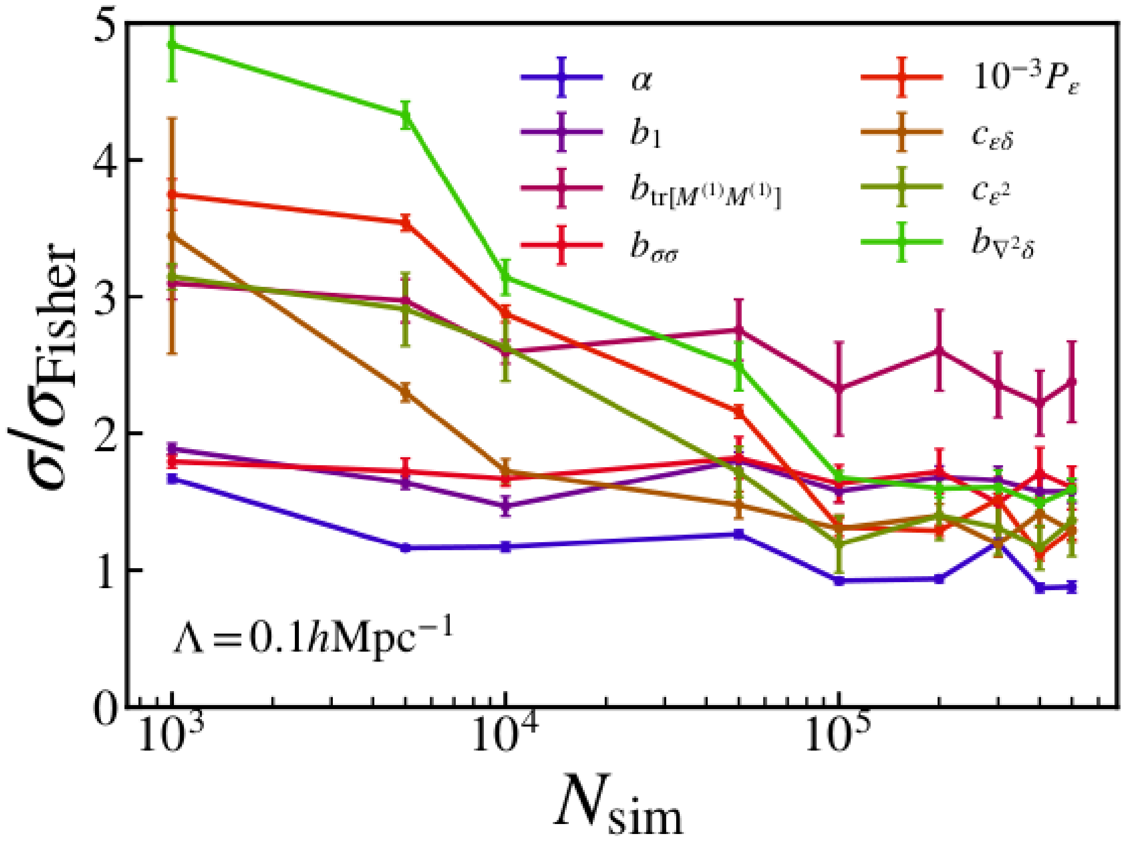
$\mathbf{x}_n \sim \mathcal{N}\left(\langle \mathbf{x}_n \rangle, {\rm Cov}[\mathbf{x}_o]\right)$

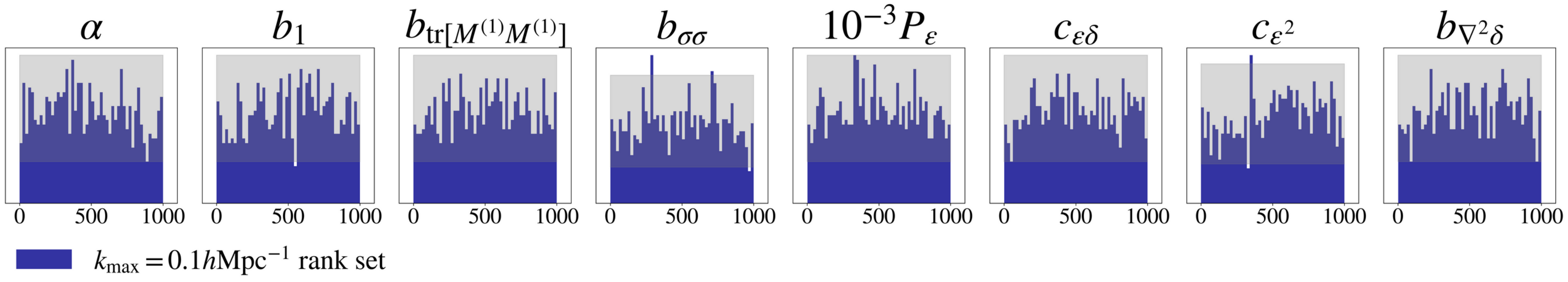
Posterior diagnostics

Simulation-based calibration



Convergence





Conclusion & Next Steps

- Simulation-based inference has proven to be a powerful tool for galaxy clustering analysis and offers several advantages over the likelihood-based approach;

- In the future, we plan to sample more cosmological parameters, add more summary statistics and improve observational aspects of the forward model (masks, systematic effects, etc);

- Comparison of field-level inference with HMC and SBI with summary statistics (P+B).

Beatriz Tucci
tucci@mpa-garching.mpg.de

Thank you!

Beatriz Tucci
tucci@mpa-garching.mpg.de

Simulation-based calibration (SBC)

How to check if the obtained posterior uncertainty is correct?

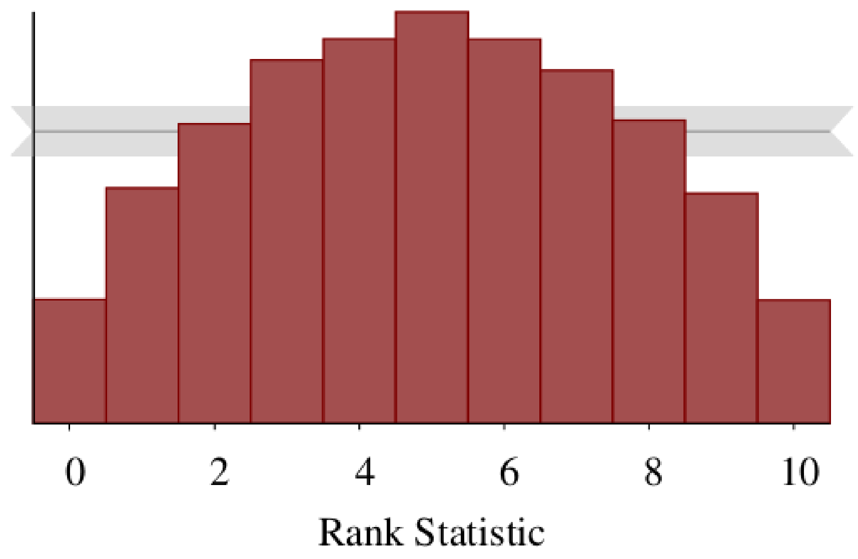
$$\mathbf{x}_o^i = \mathrm{simulator}(\boldsymbol{\theta}_o^i)$$

$$\{\hat{\boldsymbol{\theta}}\}_i \sim \hat{p}(\boldsymbol{\theta}|\mathbf{x}_o^i)$$

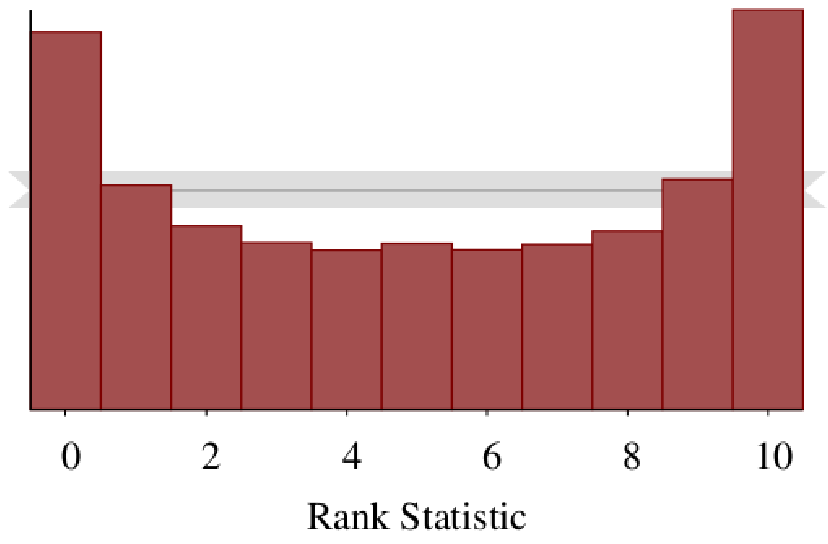
$$\hat{\theta}_1 < \hat{\theta}_2 < \ldots < \hat{\theta}_{\mathrm{rank}} < \theta_o^i < \ldots < \hat{\theta}_{\mathrm{Nsamples}}$$

Ranks should be uniformly distributed if the posterior is well calibrated

Underconfident posterior

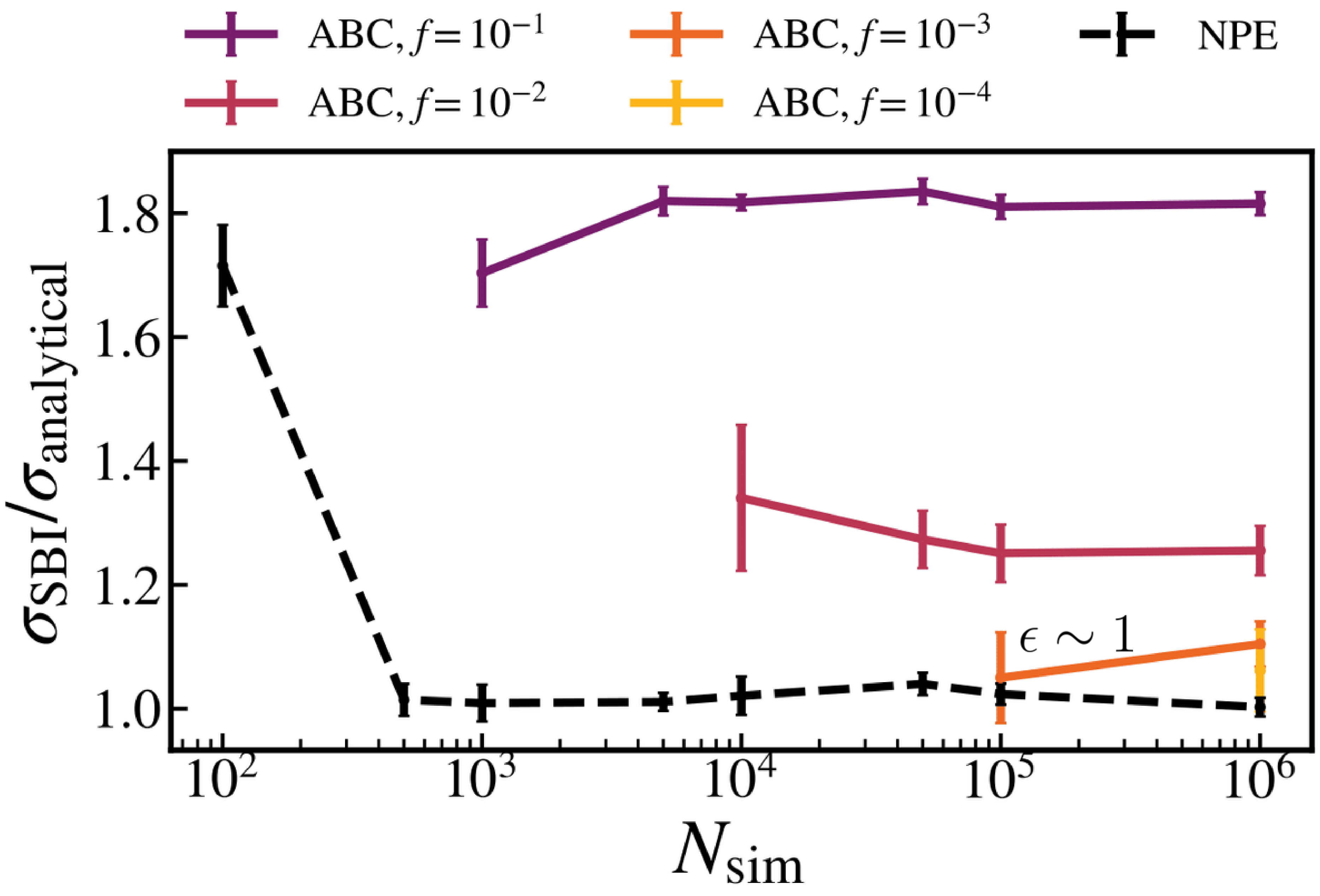


Superconfident posterior



Interlude: ABC in a symplified scenario

f: fraction of accepted samples



Reduced chi-squared as ABC metric

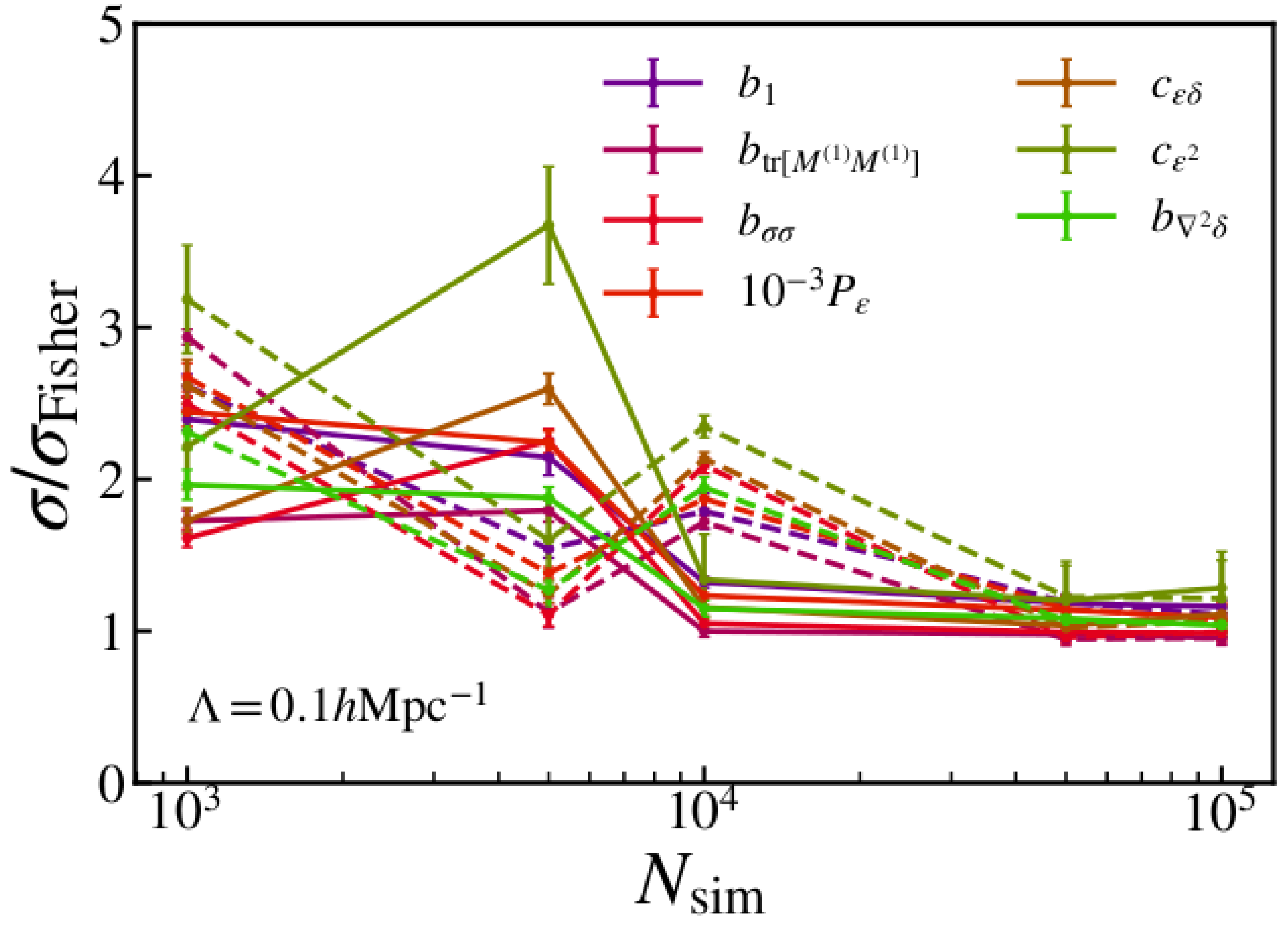
$$\rho(\mathbf{x}, \mathbf{x}_o) = \frac{1}{(D - N_\theta - 1)} \sum_{i=1}^{D} \frac{(\mathbf{x}^i - \mathbf{x}_o^i)^2}{\mathrm{Cov}[\bar{\mathbf{x}}^i]}$$

$$\boldsymbol{\theta} = \{b_1,\ 10^{-3} P_\varepsilon\}$$

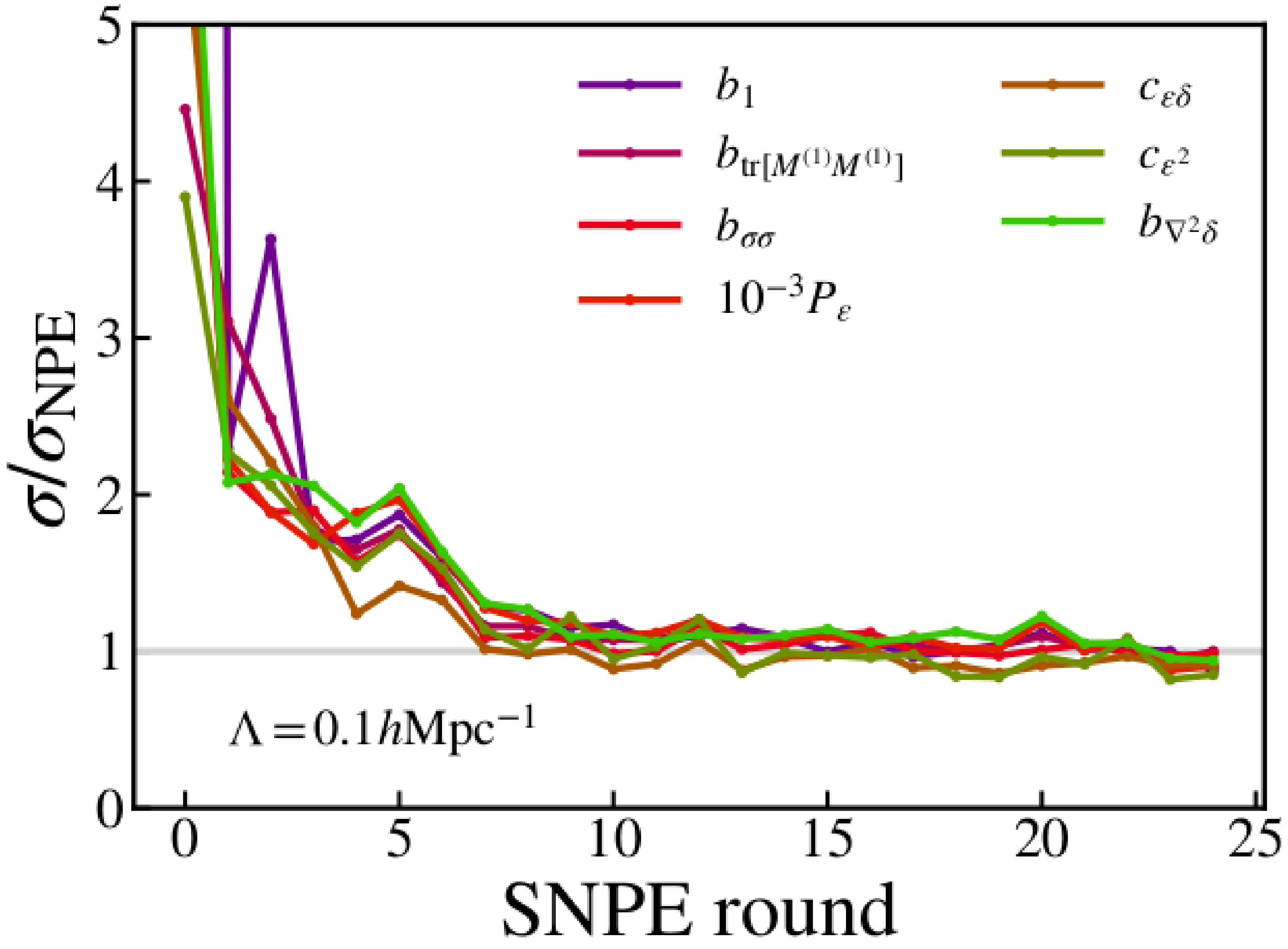
$$\mathbf{x} = \{P(k)\} = \{P(k_1), P(k_2), \ldots, P(k_D)\}$$

Inference tests

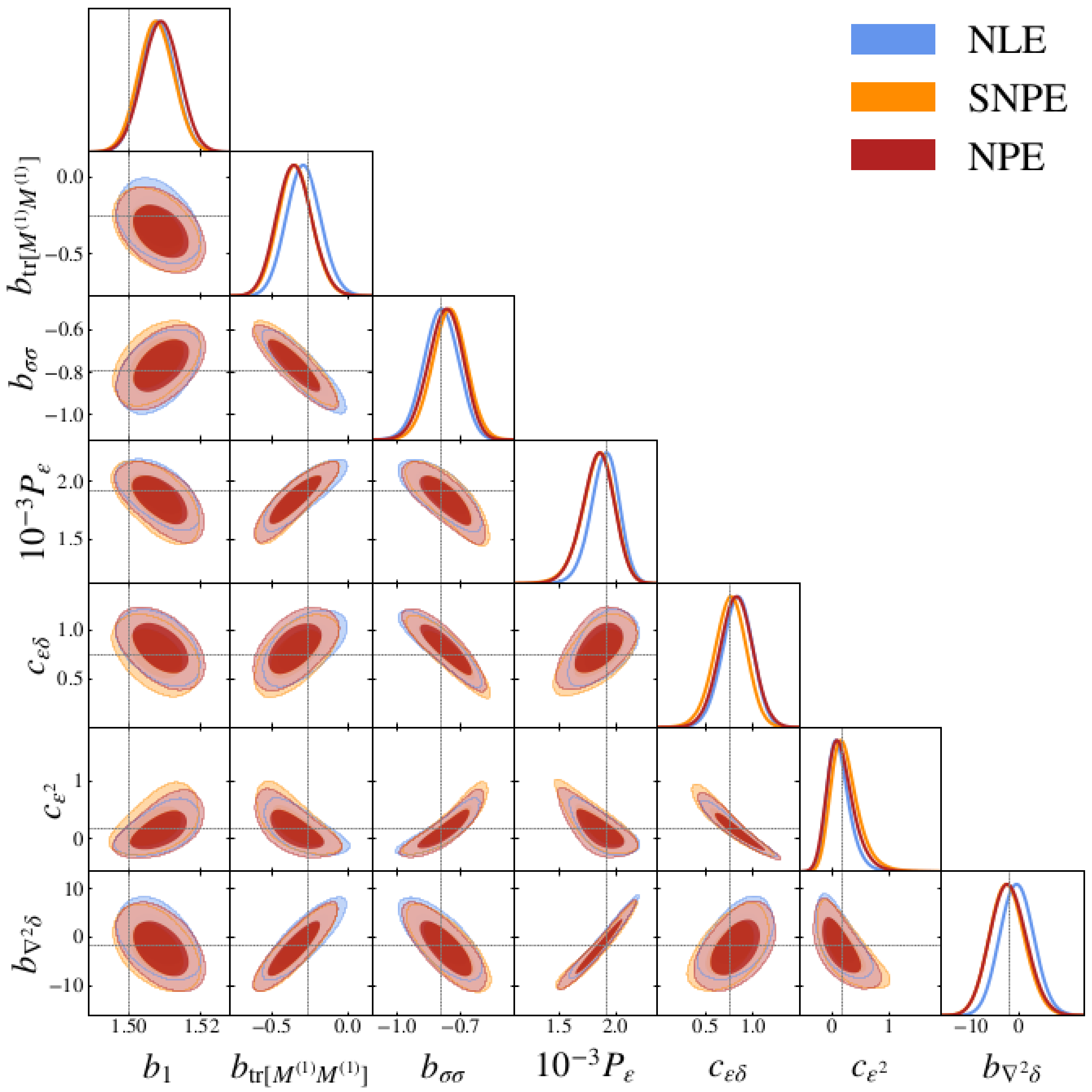
PB normalization

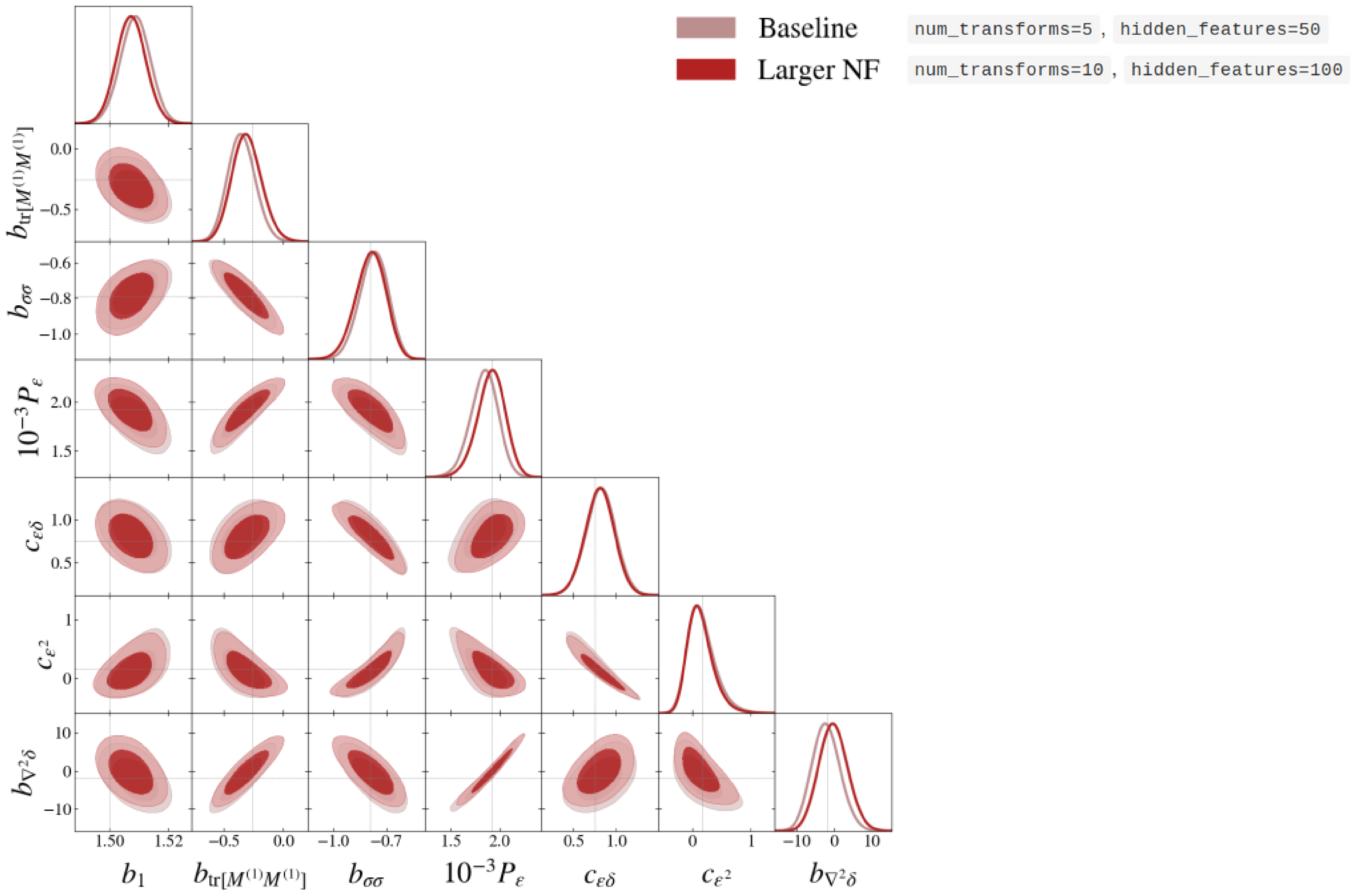


NPE vs SNPE



Inference tests





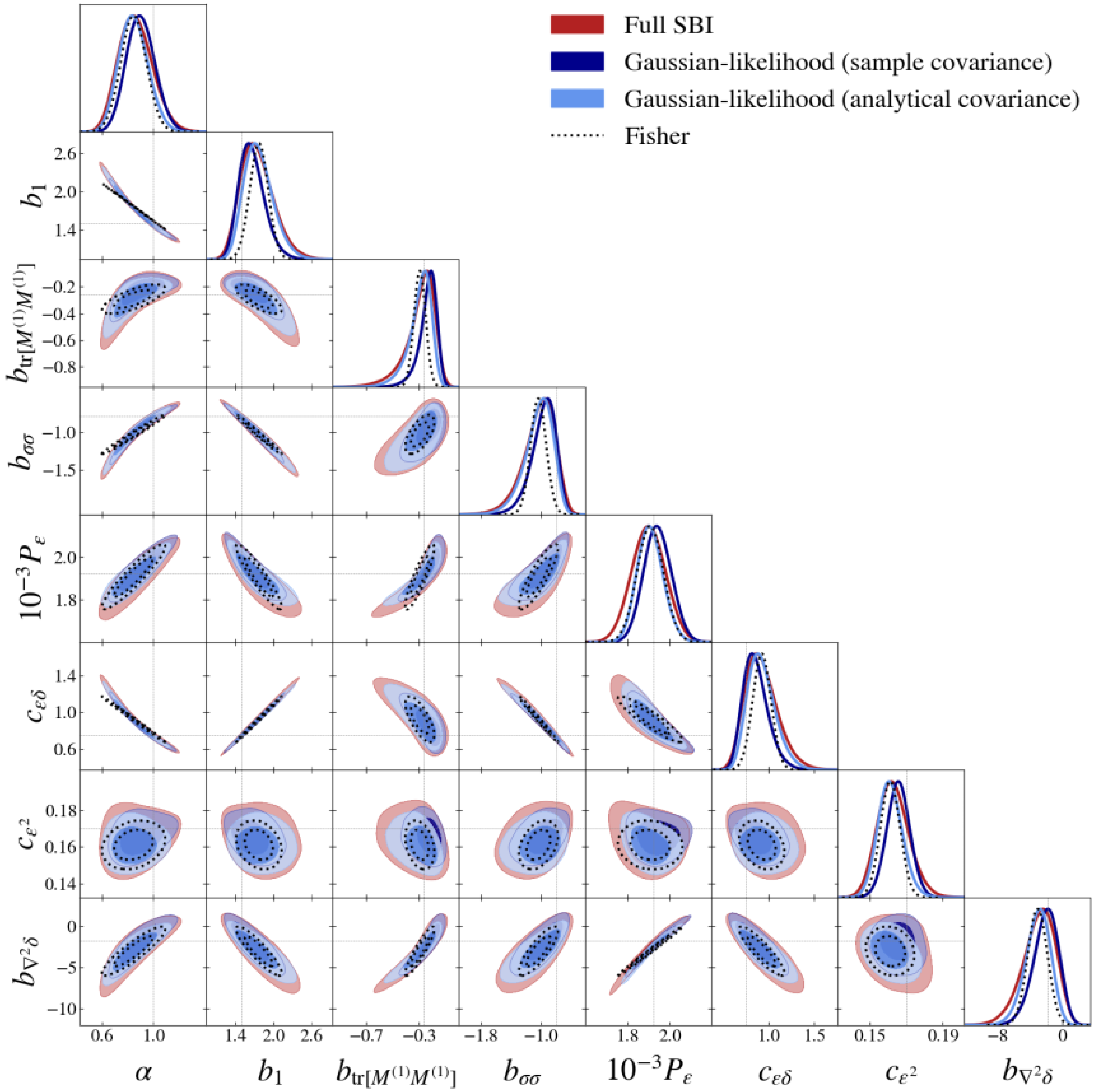
Cosmological inference | Euclid configuration

Tucci, Schmidt (2023)

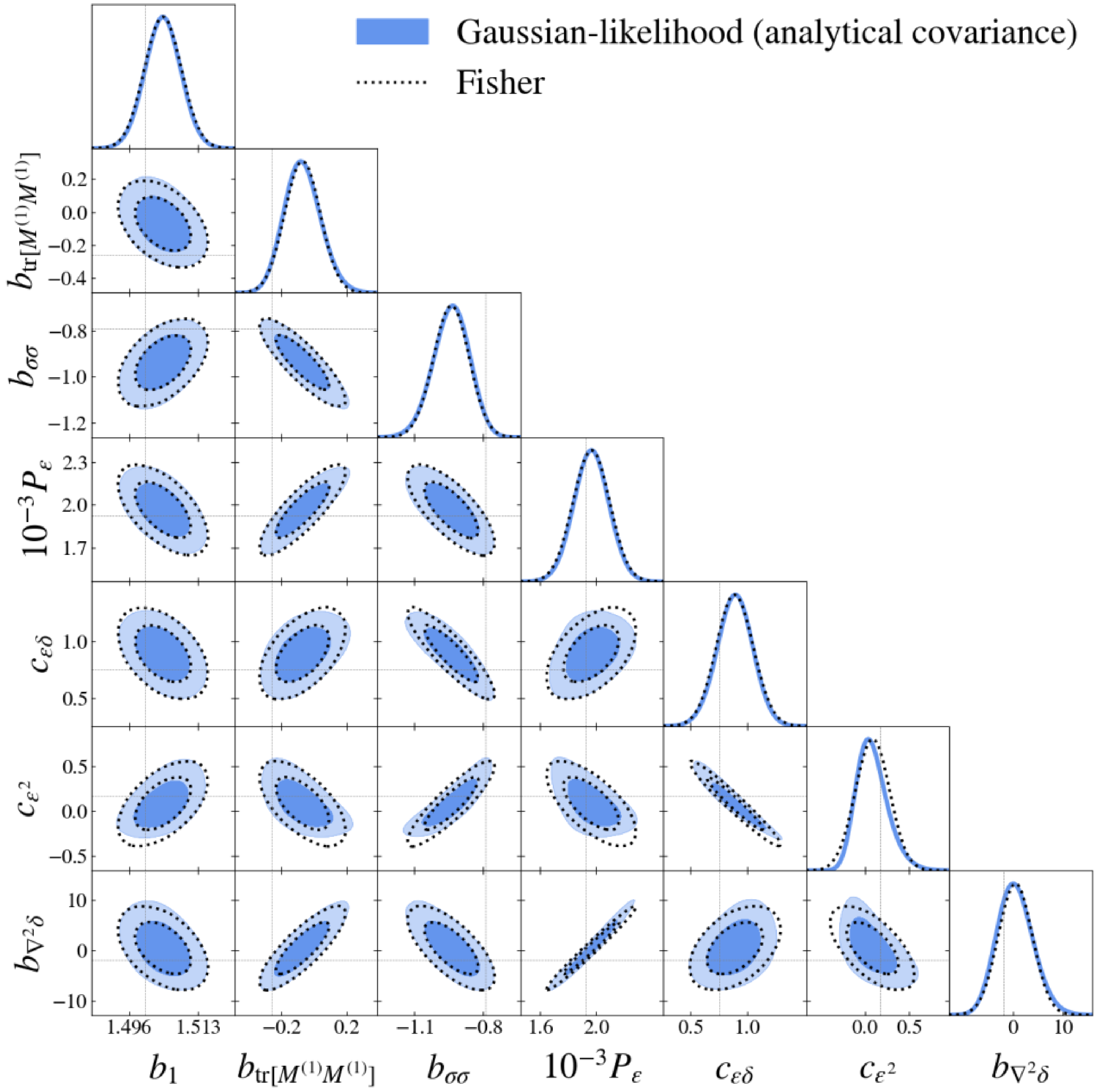
$N_{\rm sim} = 10^5$

$k_{\max} = \Lambda = 0.2h\mathrm{Mpc}^{-1}$

$D = 49$



Euclid | Gaussian-likelihood

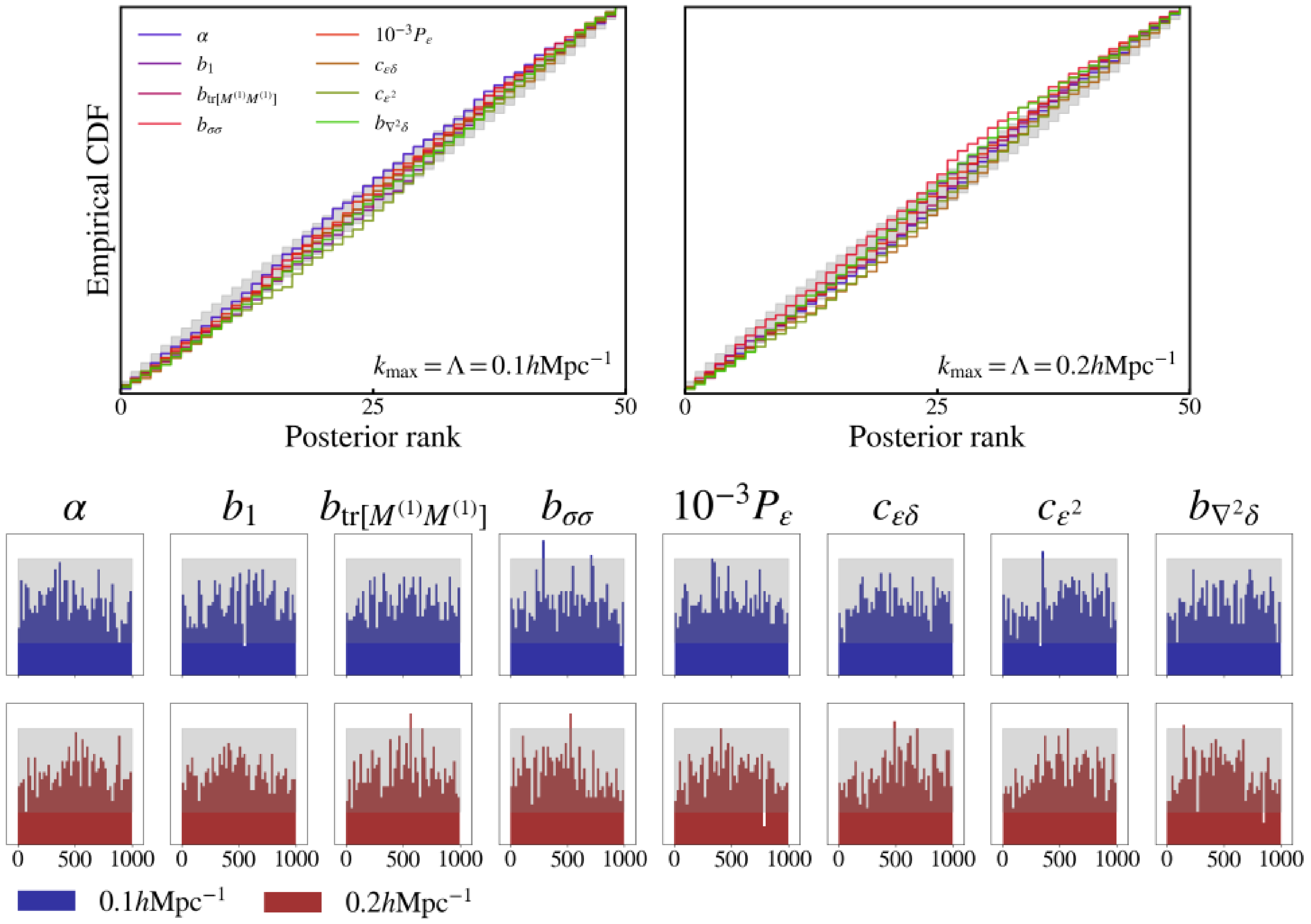


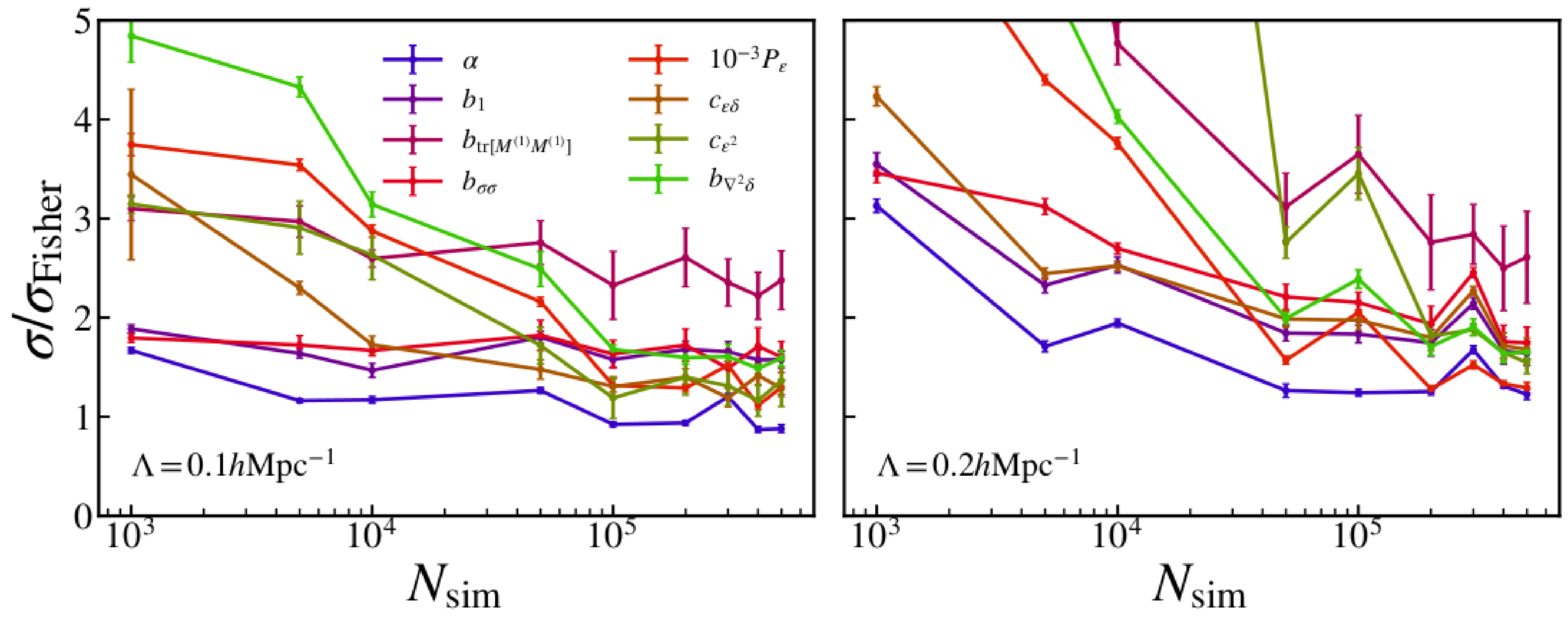
$$\mathbf{x}_n \sim \mathcal{N}\left(\langle \mathbf{x}_n \rangle, \mathrm{Cov}[\mathbf{x}_o]\right)$$

$$N_{\mathrm{sim}} = 10^5$$

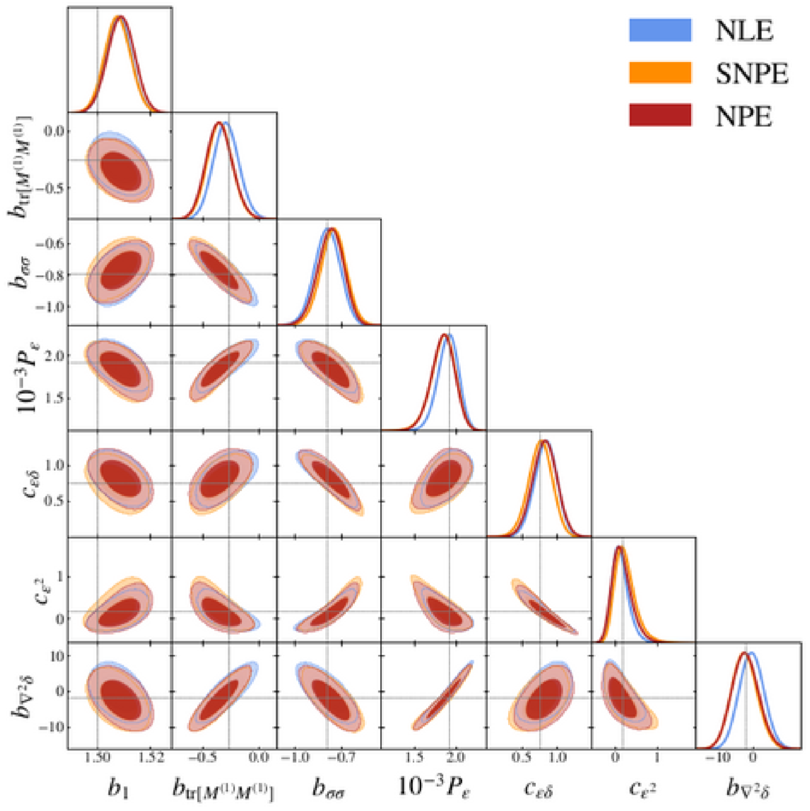
$$k_{\mathrm{max}} = \Lambda = 0.1h\mathrm{Mpc}^{-1}$$

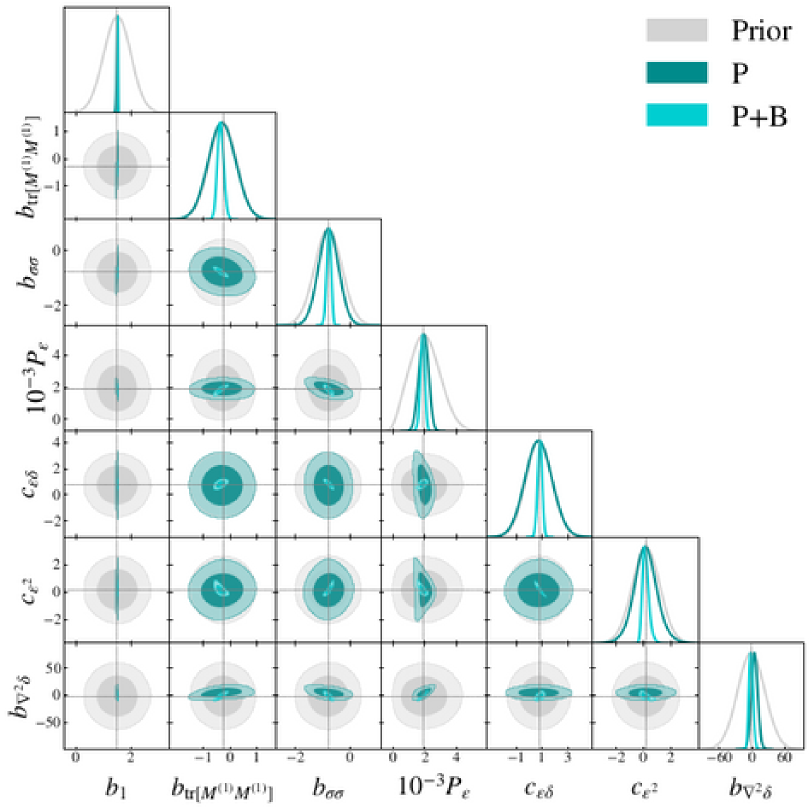
Posterior diagnostics

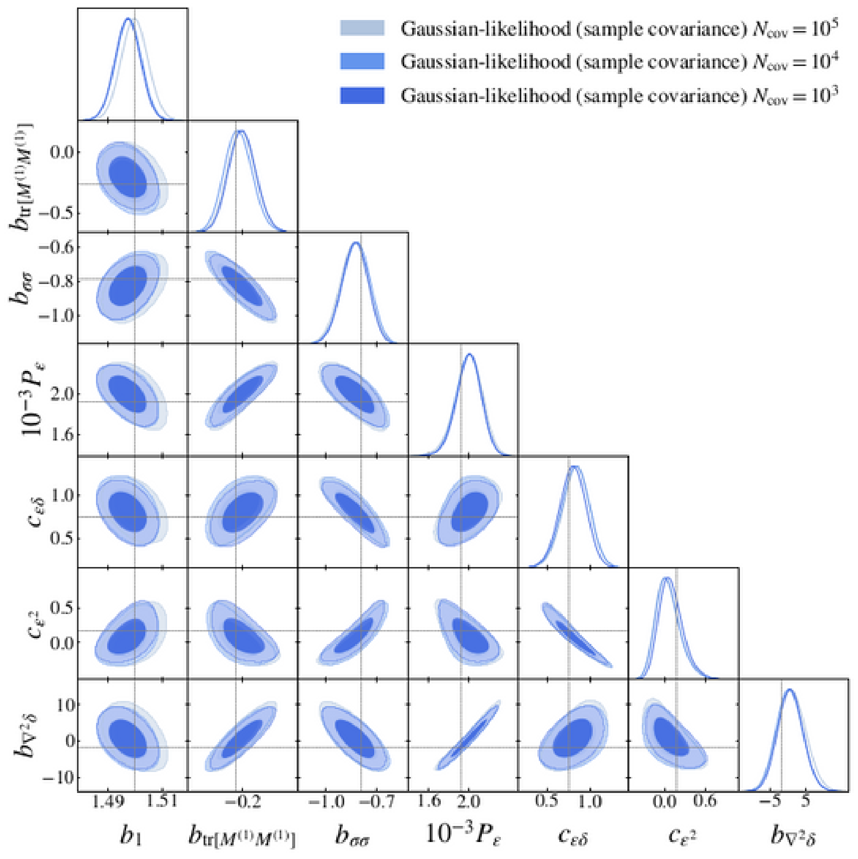


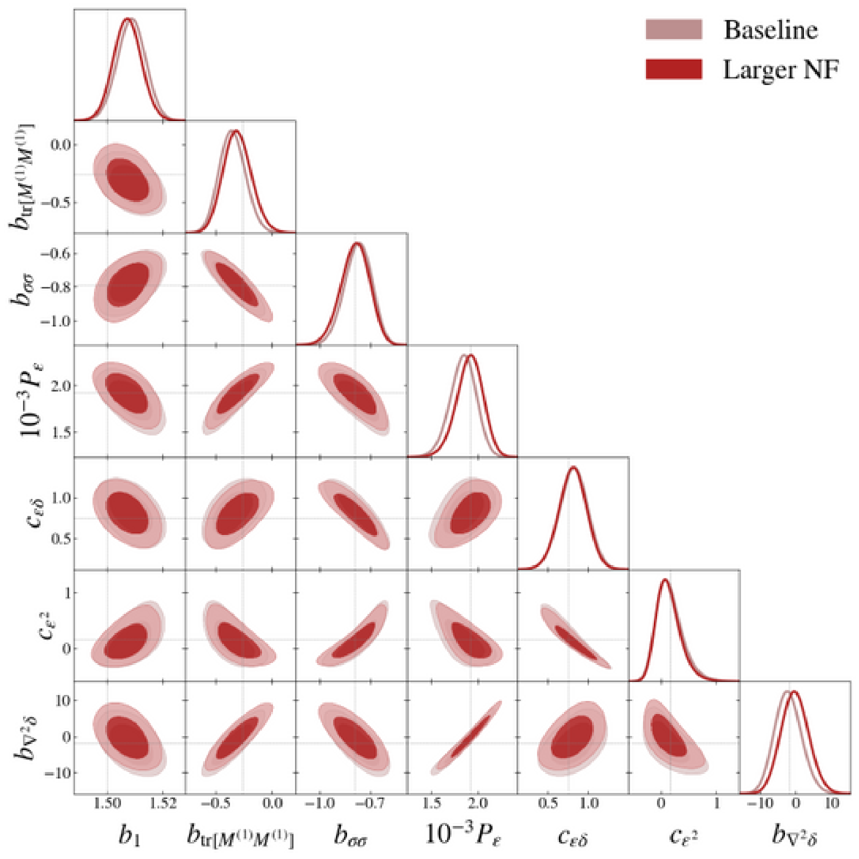


Inference tests



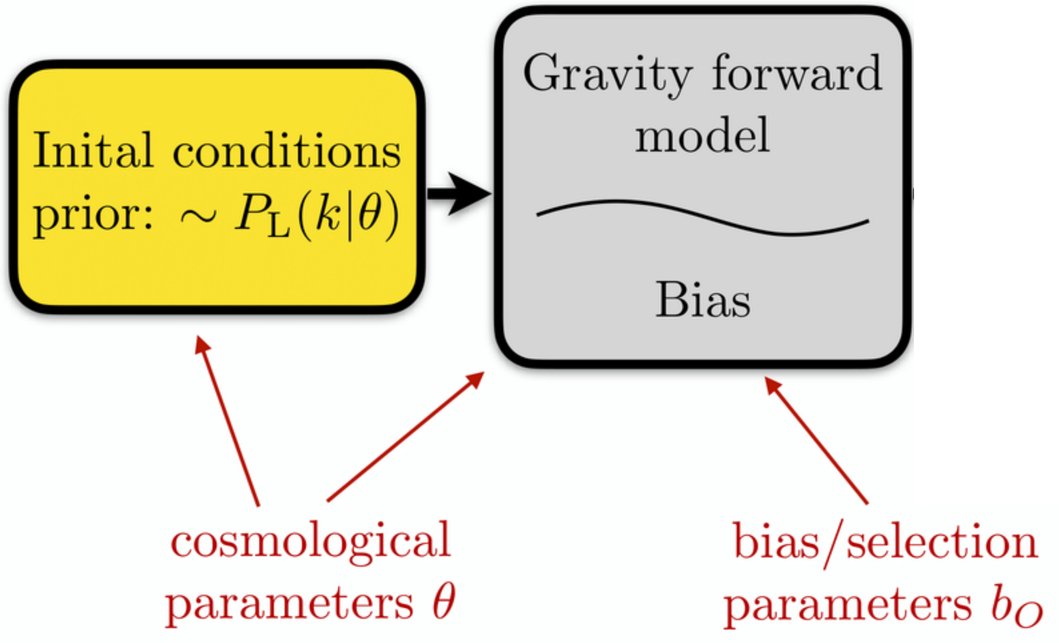






LEFTfield | forward model

EFTofLSS based approach



An n-th order Lagrangian Forward Model for Large-Scale Structure
Fabian Schmidt (2021)

$$\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon)$$

Perturbation Theory

$$\langle \delta_g(k_1)\delta_g(k_2)\delta_g(k_3)\rangle'^{\mathrm{LO}}_{\mathrm{stoch}} = B_\varepsilon + 2b_1 P_{\varepsilon\varepsilon_\delta}(P_{\mathrm{m}}(k_1) + 2 \text{ perm.})$$

Forward Model

$$\langle \delta_g(k_1)\delta_g(k_2)\delta_g(k_3)\rangle'^{\mathrm{LO}}_{\mathrm{stoch}} = 6c_\varepsilon^{\mathrm{NG}} P_\varepsilon^2 + 2b_1 P_\varepsilon \sigma_{\varepsilon\delta}(P_{\mathrm{m}}(k_1) + 2 \text{ perm.})$$

$$\delta_g(\mathbf{x},\tau) = \delta_{g,\mathrm{det}}(\mathbf{x},\tau) + \varepsilon(\mathbf{x},\tau) + \sigma_{\varepsilon\delta}(\tau)\varepsilon(\mathbf{x},\tau)\delta(\mathbf{x},\tau) + c_\varepsilon^{\mathrm{NG}}(\tau)\varepsilon^2(\mathbf{x},\tau)$$