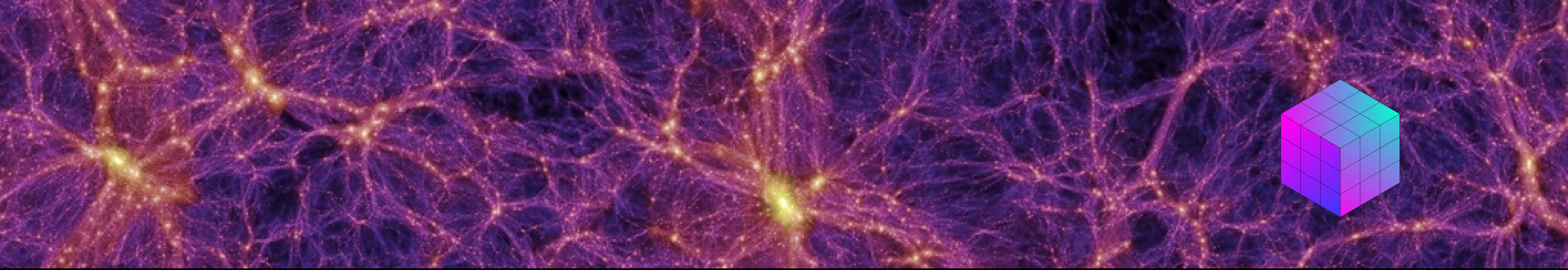


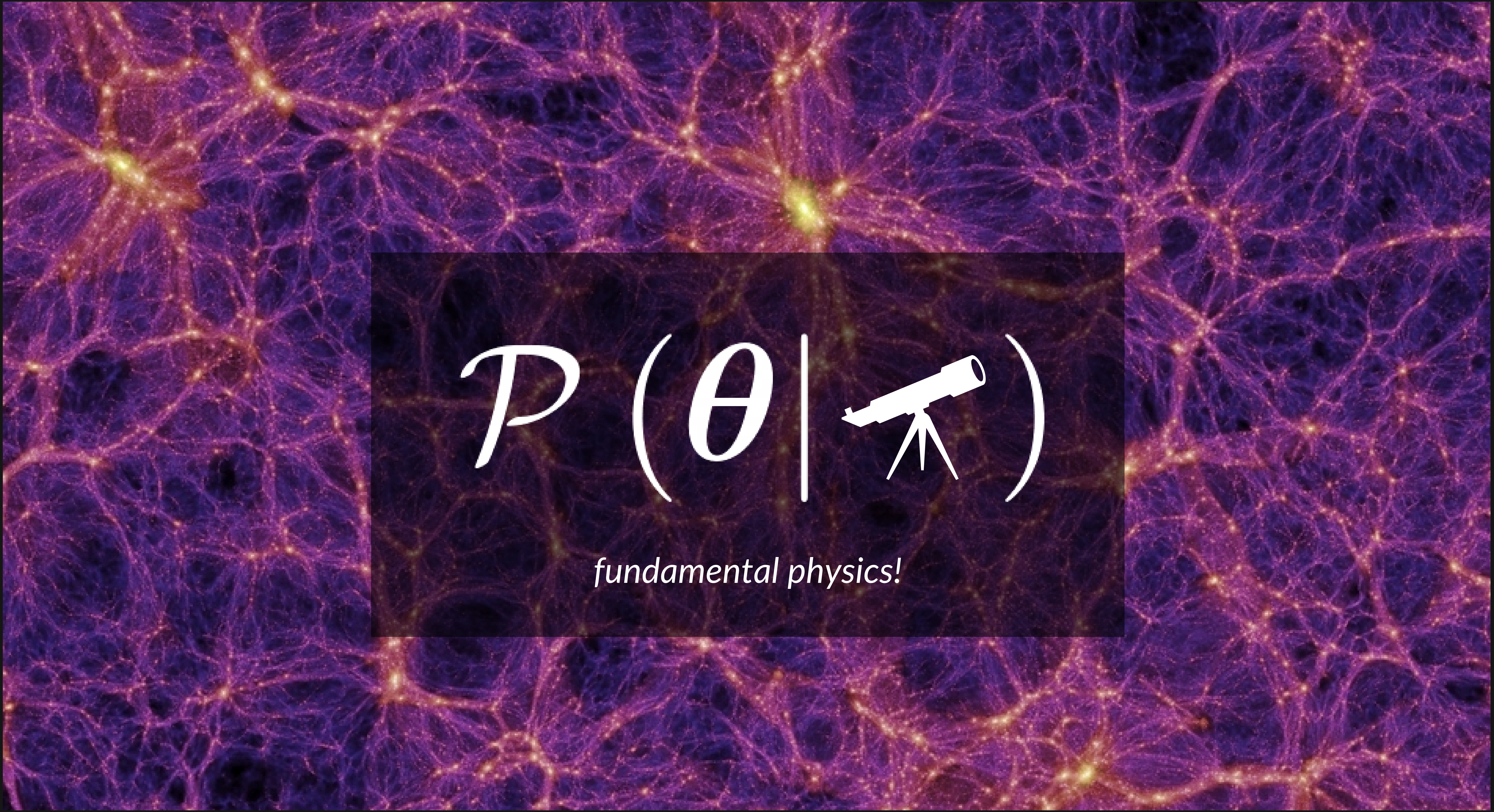
How much information can be extracted from galaxy clustering at the field level?

Beatriz Tucci

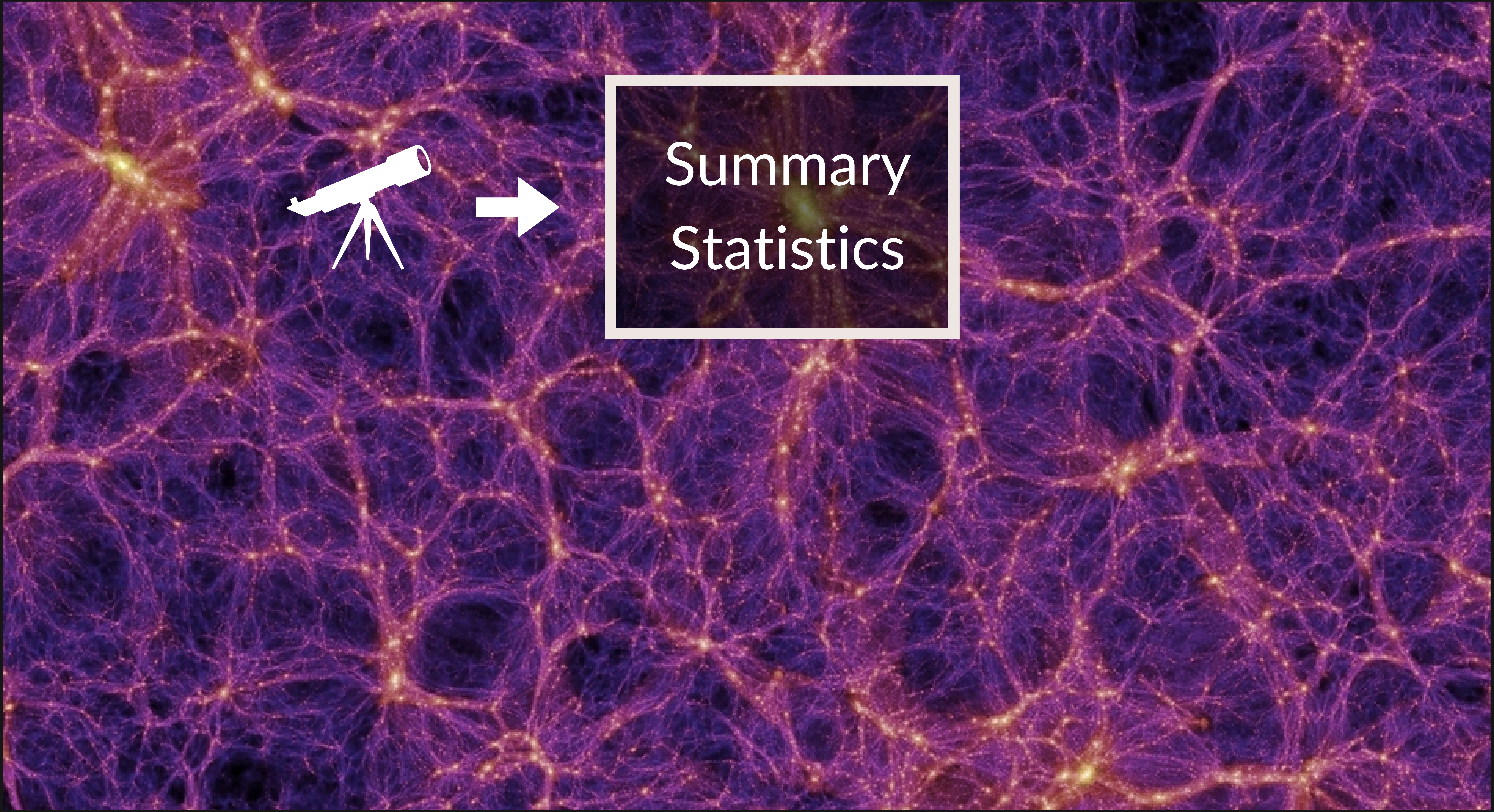
with Nhat-Minh Nguyen, Fabian Schmidt, Andrija Kostić, Martin Reinecke

MAX-PLANCK-INSTITUT FÜR ASTROPHYSIK

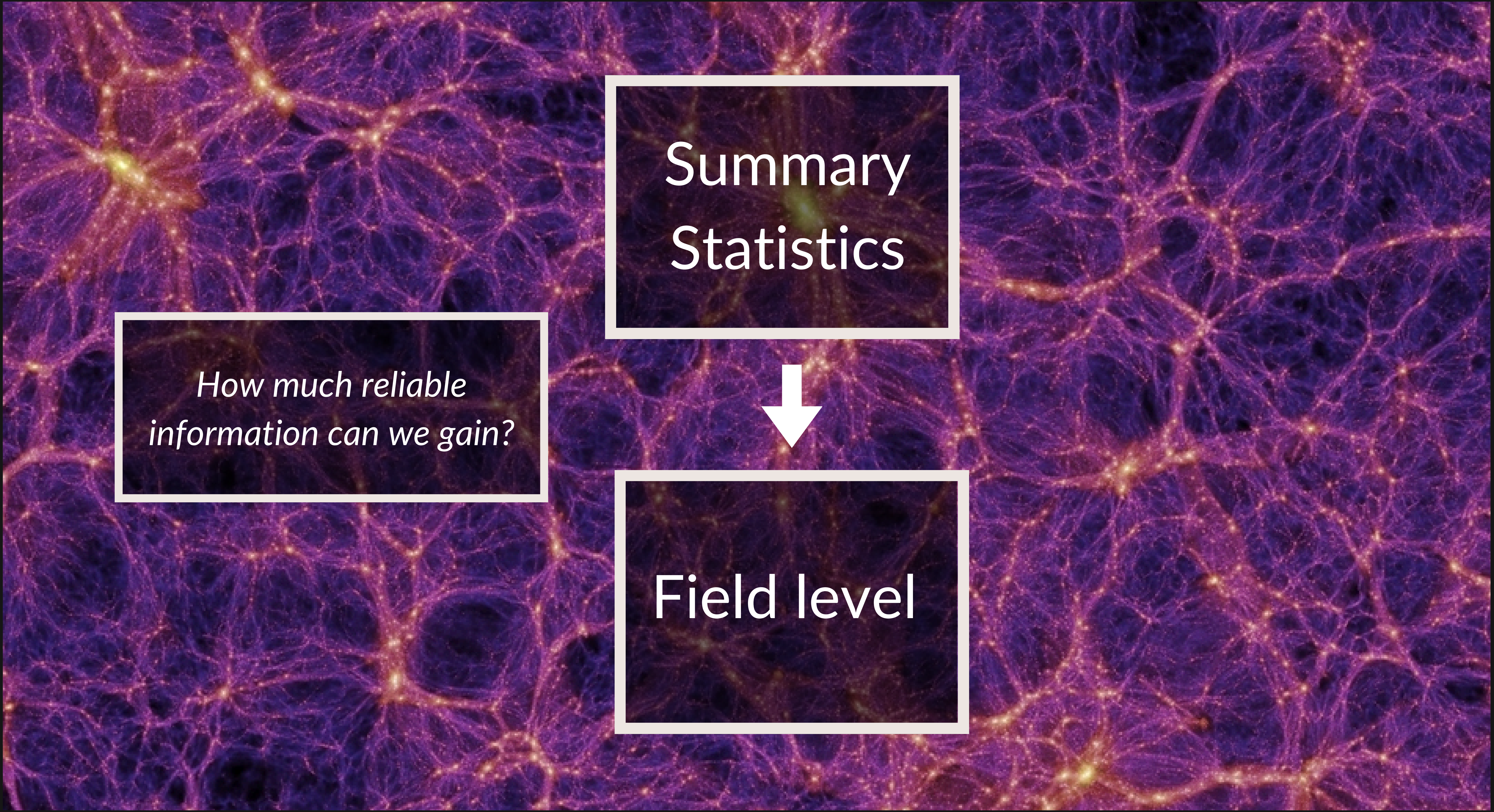
Based on arXiv:2403.03220



The Millennium Simulation Project



The Millennium Simulation Project



The Millennium Simulation Project

Field level

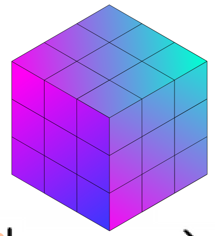
$$\mathcal{P}\left(\boldsymbol{\theta} \middle| \delta_g^{\rm obs}\right)$$

Galaxy density field

$$\delta_g(x) = \frac{n_g(x)}{\bar{n}_g} - 1$$

The Millennium Simulation Project

How to get the field level posterior?

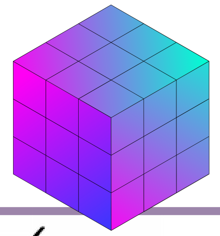


$$\mathcal{P}\left(\boldsymbol{\theta}, \{b_O\}, \{\sigma_\varepsilon\} \middle| \delta_g^{\text{obs}}\right)$$

Bias and stochastic

“nuisance” parameters

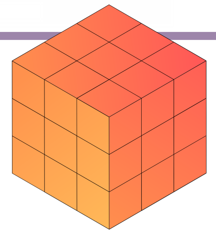
How to get the field level posterior?



Field-level Likelihood

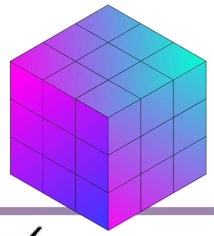
Prior

$$\mathcal{P}\left(\boldsymbol{\theta},\{b_O\},\{\sigma_\varepsilon\}\middle|\delta_g^{\rm obs}\right) \propto \int \mathcal{D}\delta_\Lambda^{(1)}\,\mathcal{P}(\delta_\Lambda^{(1)})\,\mathcal{L}\left(\delta_g^{\rm obs}\middle|\delta_{g,\rm det}[\boldsymbol{\theta},\delta_\Lambda^{(1)},\{b_O\}],\{\sigma_\varepsilon\}\right)\mathcal{P}\left(\boldsymbol{\theta},\{b_O\},\{\sigma_\varepsilon\}\right)$$



Initial conditions marginalization

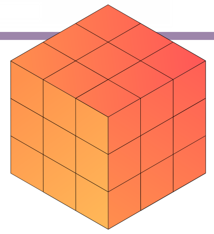
How to get the field level posterior?



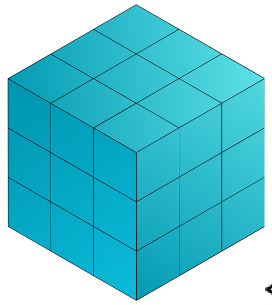
Field-level Likelihood

Prior

$$\mathcal{P}\left(\boldsymbol{\theta},\{b_O\},\{\sigma_\varepsilon\}\middle|\delta_g^{\rm obs}\right) \propto \int \mathcal{D}\delta_\Lambda^{(1)}\,\mathcal{P}(\delta_\Lambda^{(1)})\,\mathcal{L}\left(\delta_g^{\rm obs}\middle|\delta_{g,\rm det}[\boldsymbol{\theta},\delta_\Lambda^{(1)},\{b_O\}],\{\sigma_\varepsilon\}\right)\mathcal{P}\left(\boldsymbol{\theta},\{b_O\},\{\sigma_\varepsilon\}\right)$$



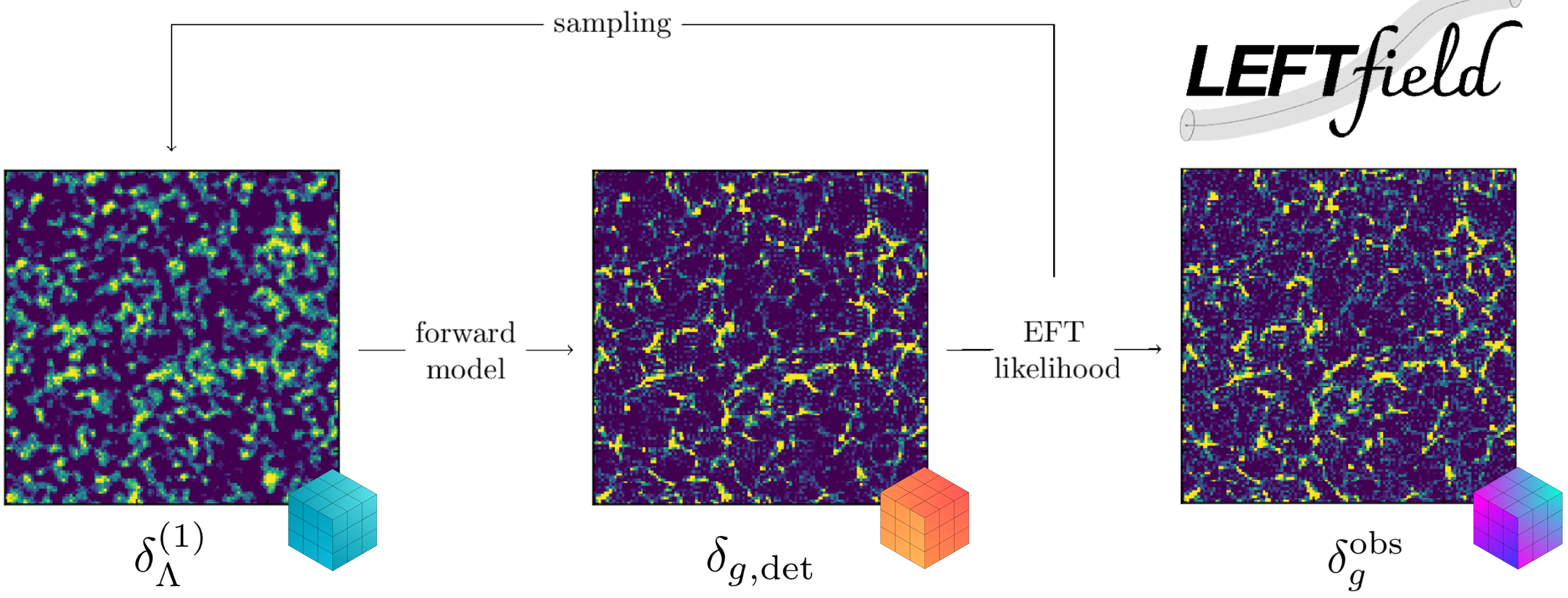
Initial conditions marginalization



$$\left\{\delta_{\Lambda,i}^{(1)}\right\}_{i=1}^{N_g^\Lambda}$$

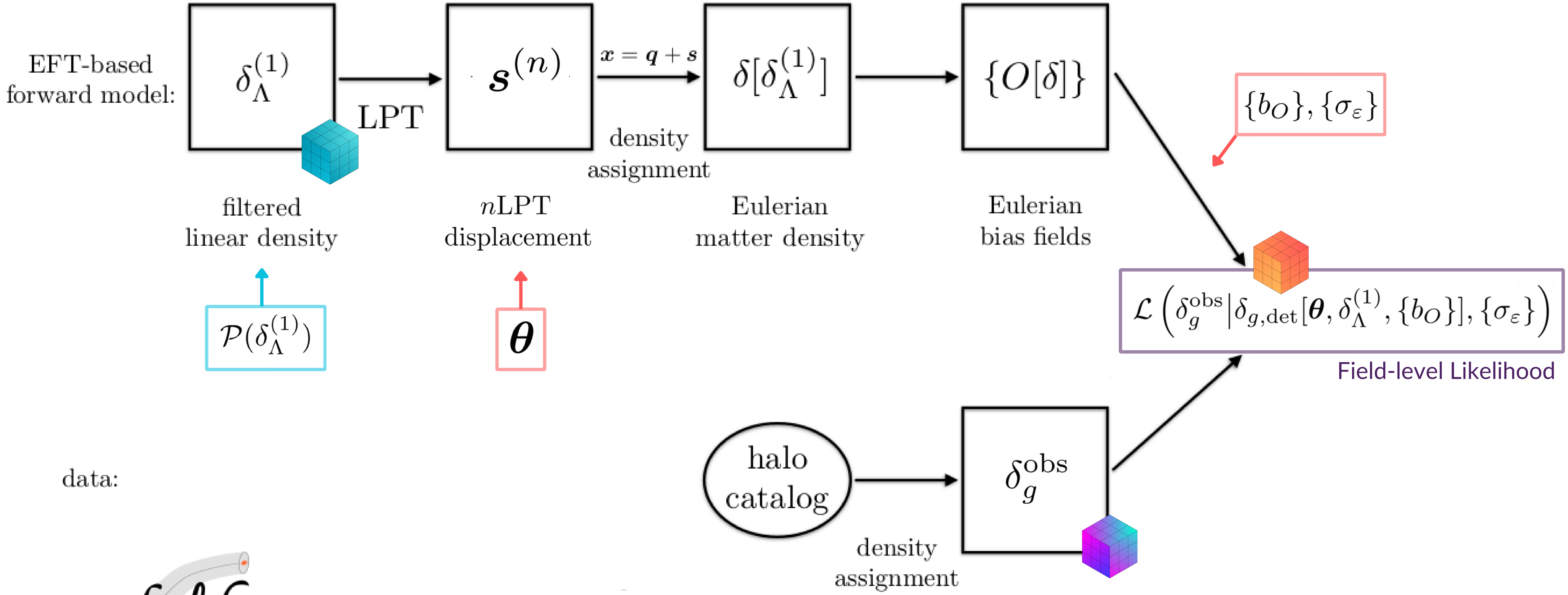
Field-level inference

Schmidt et al (2018, 2020)
Cabass & Schmidt (2019, 2020)
Babic, Schmidt, Tucci (2022)
Kostic et al (2020)
Stadler et al (2023)



Credits: Julia Stadler

Forward model

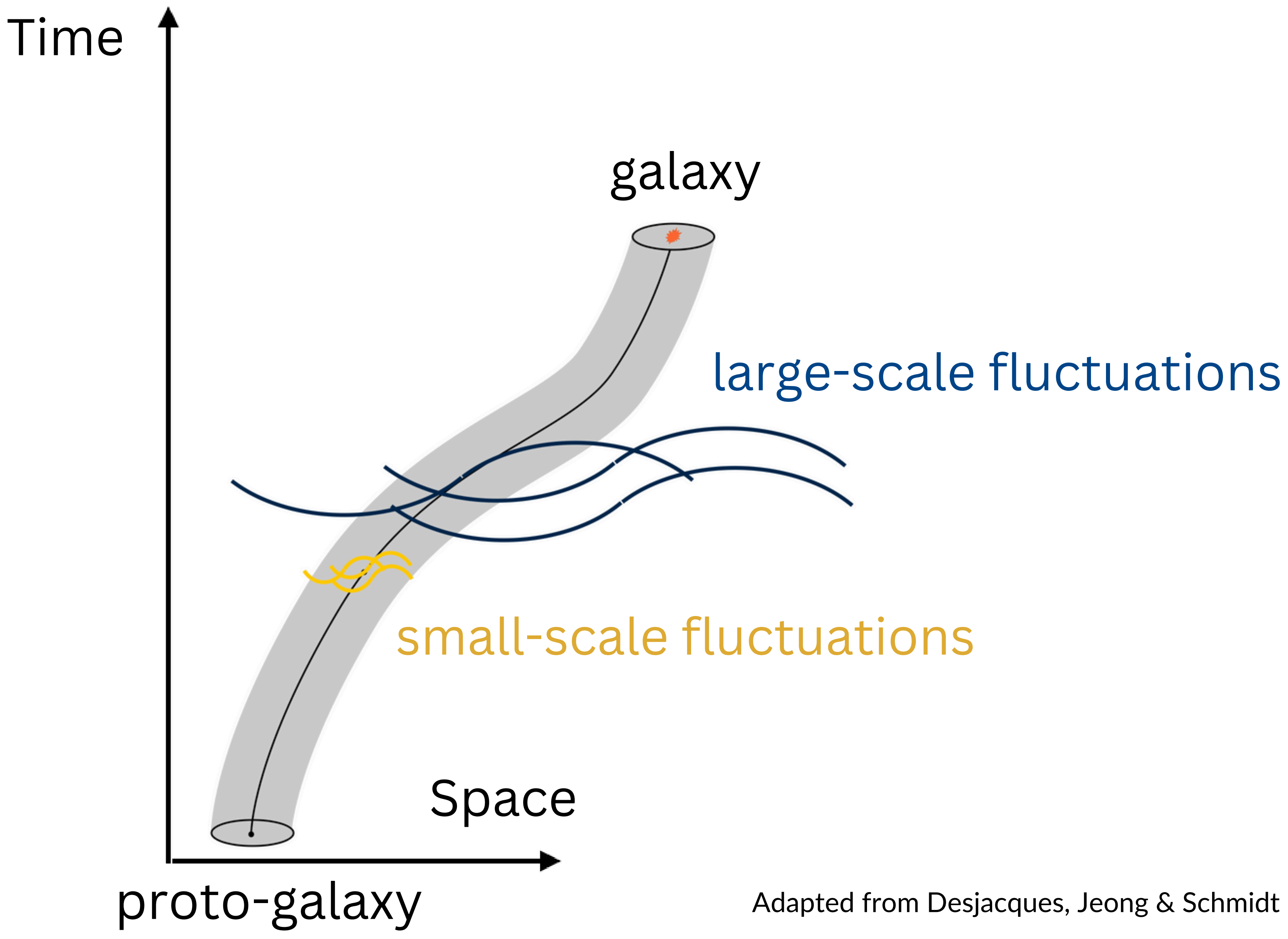


An n-th order Lagrangian Forward Model for Large-Scale Structure
Schmidt (2021)

Bias expansion

c.f. Pierre's talk

$$\delta_g(\boldsymbol{k}, z) = \delta_{g,\mathrm{det}}(\boldsymbol{k}, z) + \delta_{g,\mathrm{stoch}}(\boldsymbol{k}, z)$$



Adapted from Desjacques, Jeong & Schmidt (2016)

Bias expansion

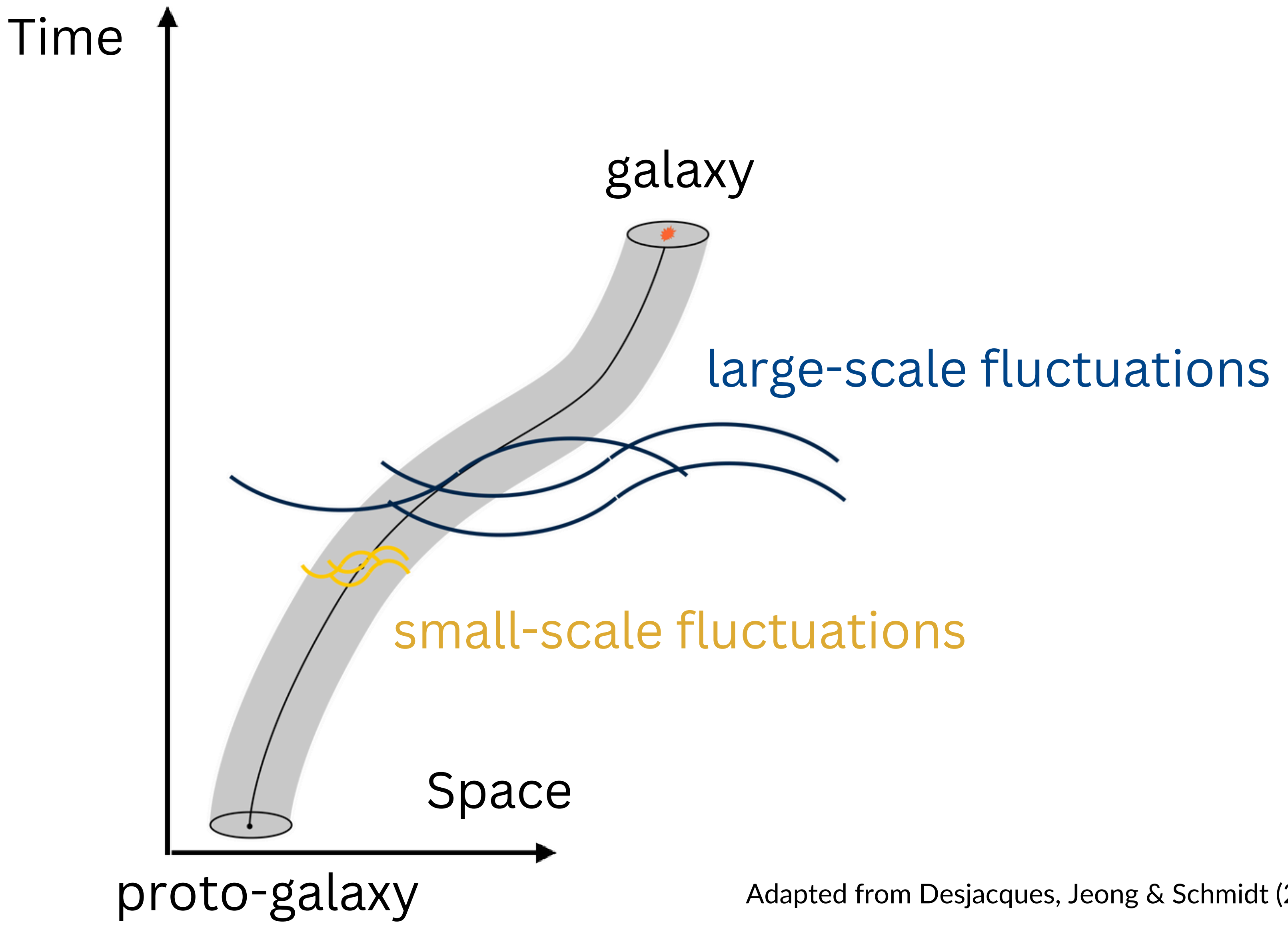
$$\delta_g(\boldsymbol{k}, z) = \delta_{g,\mathrm{det}}(\boldsymbol{k}, z) + \delta_{g,\mathrm{stoch}}(\boldsymbol{k}, z)$$
$$= \sum_O b_O(z) O(\boldsymbol{k}, z) + \varepsilon(\boldsymbol{k}, z)$$

$\{b_O\}$ **Free bias parameters**

$$O \in [\delta, \delta^2, K^2, \delta^3, K^3, \delta K^2, O_{\mathrm{td}}, \nabla^2\delta]$$

3rd order Eulerian bias

Desjacques, Jeong & Schmidt (2016)
Mirbabayi, Schmidt, Zaldarriaga (2015)



Adapted from Desjacques, Jeong & Schmidt (2016)

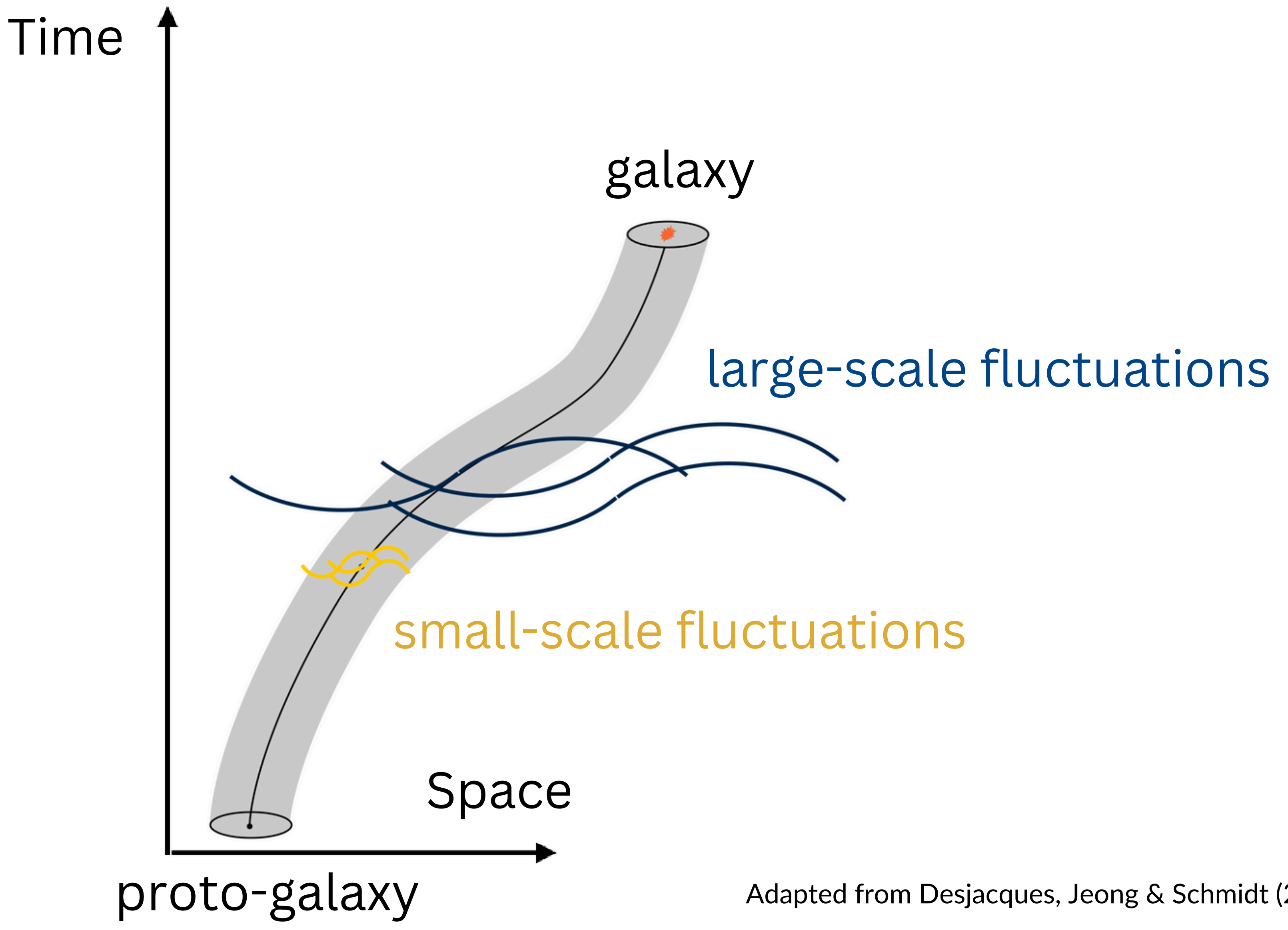
Bias expansion

$$\delta_g(\boldsymbol{k}, z) = \delta_{g,\mathrm{det}}(\boldsymbol{k}, z) + \boxed{\delta_{g,\mathrm{stoch}}(\boldsymbol{k}, z)}$$
$$= \sum_O b_O(z) O(\boldsymbol{k}, z) + \boxed{\varepsilon(\boldsymbol{k}, z)}$$

$$\langle \varepsilon(\boldsymbol{k}, z)\varepsilon(\boldsymbol{k}', z)\rangle' \propto \sigma_\varepsilon^2(k)$$

$$\sigma_\varepsilon(k) = \boxed{\sigma_{\varepsilon,0}}\left[1 + \boxed{\sigma_{\varepsilon,k^2}} k^2\right]$$

Free stochastic parameters $\boxed{\{\sigma_\varepsilon\}}$



Adapted from Desjacques, Jeong & Schmidt (2016)

Field level Likelihood

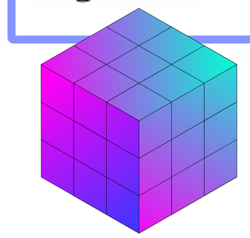
$$k_{\mathrm{max}} \leq \Lambda$$

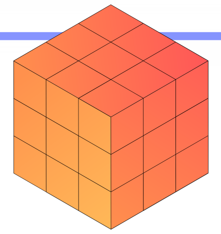
$$\ln \mathcal{L}\left(\delta_g^{\mathrm{obs}} \middle| \delta_{g,\mathrm{det}}[\boldsymbol{\theta}, \delta_\Lambda^{(1)}, \{b_O\}], \{\sigma_\varepsilon\}\right) = -\frac{1}{2} \sum_{k<k_{\mathrm{max}}} \left[\frac{1}{\sigma_\varepsilon^2(k)} \left|\delta_g^{\mathrm{obs}}(\boldsymbol{k}) - \delta_{g,\mathrm{det}}[\boldsymbol{\theta}, \delta_\Lambda^{(1)}, \{b_O\}](\boldsymbol{k})\right|^2 + \ln[2\pi\sigma_\varepsilon^2(k)]\right]$$

Field level Likelihood

Mode by mode
data and theory
comparison!

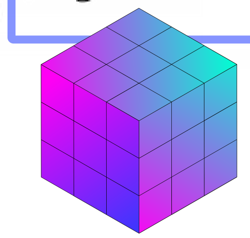
$$\ln \mathcal{L}\left(\delta_g^{\rm obs} \middle| \delta_{g,\rm det}[\boldsymbol{\theta}, \delta_\Lambda^{(1)}, \{b_O\}], \{\sigma_\varepsilon\}\right) = -\frac{1}{2}\sum_{k<k_{\max}} \left[\frac{1}{\sigma_\varepsilon^2(k)}\left|\delta_g^{\rm obs}(\boldsymbol{k}) - \delta_{g,\rm det}[\boldsymbol{\theta},\delta_\Lambda^{(1)},\{b_O\}](\boldsymbol{k})\right|^2 + \ln[2\pi\sigma_\varepsilon^2(k)]\right]$$

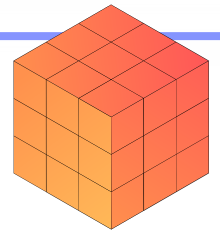




Field level Likelihood

$$\ln \mathcal{L}\left(\delta_g^{\mathrm{obs}} \middle| \delta_{g,\mathrm{det}}[\boldsymbol{\theta}, \delta_\Lambda^{(1)}, \{b_O\}], \{\sigma_\varepsilon\}\right) = -\frac{1}{2} \sum_{k<k_{\mathrm{max}}} \left[\frac{1}{\sigma_\varepsilon^2(k)} \left| \delta_g^{\mathrm{obs}}(\boldsymbol{k}) - \delta_{g,\mathrm{det}}[\boldsymbol{\theta}, \delta_\Lambda^{(1)}, \{b_O\}](\boldsymbol{k}) \right|^2 + \ln[2\pi\sigma_\varepsilon^2(k)] \right]$$

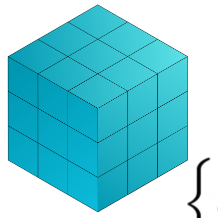




↓ **HMC**

$$\mathcal{P}\left(\boldsymbol{\theta}, \delta_\Lambda^{(1)}, \{b_O\}, \{\sigma_\varepsilon\} \middle| \delta_g^{\mathrm{obs}}\right)$$

Full posterior including initial conditions!



$\left\{\delta_{\Lambda,i}^{(1)}\right\}_{i=1}^{N_g^\Lambda}$

Inference setup

Using Nbody halos as observed data, can we break the degeneracy between σ_8 and bias parameters with field level inference?

$$\alpha \equiv \sigma_8 / \sigma_8^{\mathrm{true}}$$

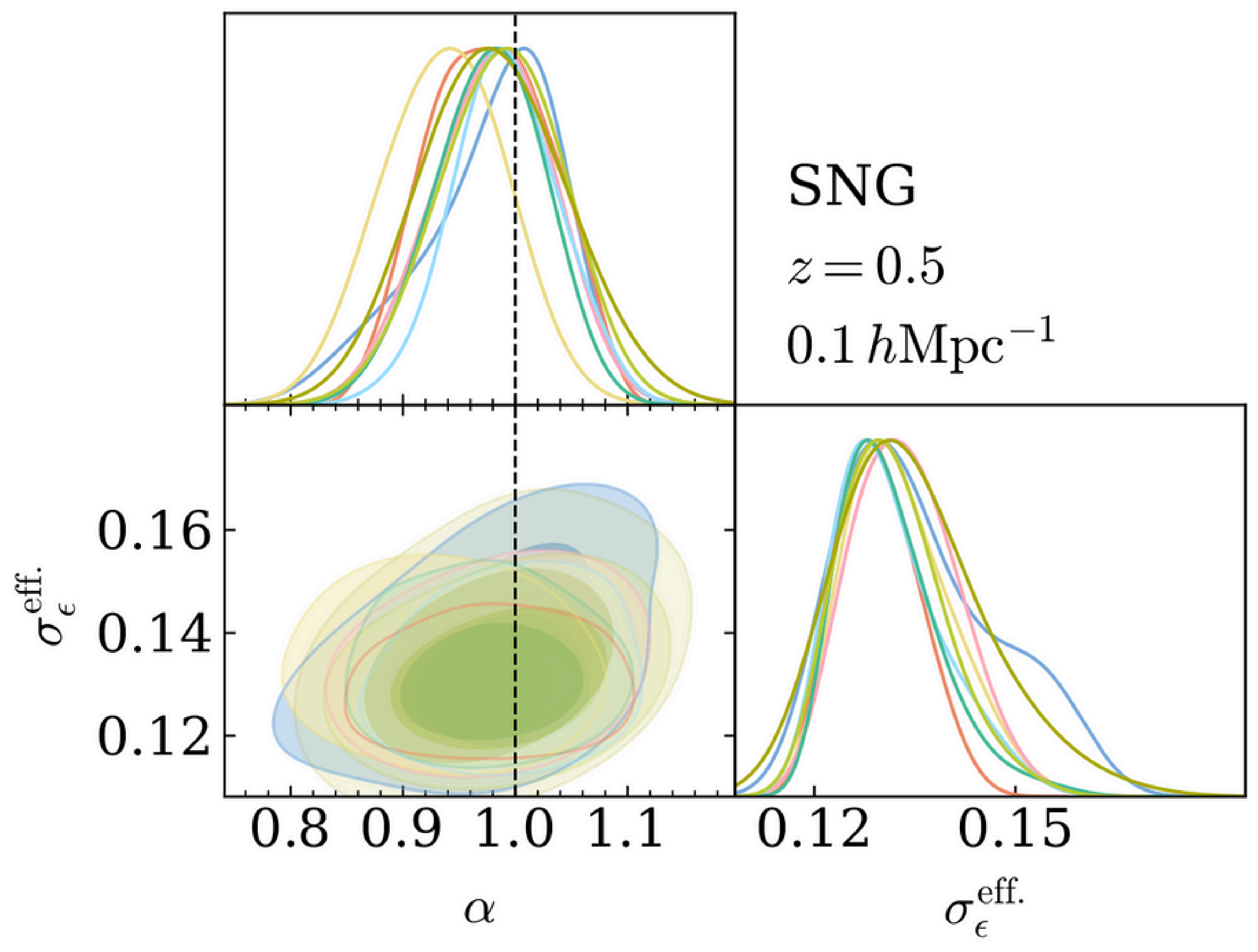
$$\boldsymbol{\theta} = \{\alpha\}$$

Inference setup: halo samples

	SNG	Uchuu
Redshift	$z = 0.50$	$z = 1.03$
$V\,[h^{-3}\mathrm{Mpc}^3]$	2000^3	2000^3
$\bar{n}_g\,[h^3\mathrm{Mpc}^{-3}]$	1.3×10^{-3}	3.6×10^{-3}

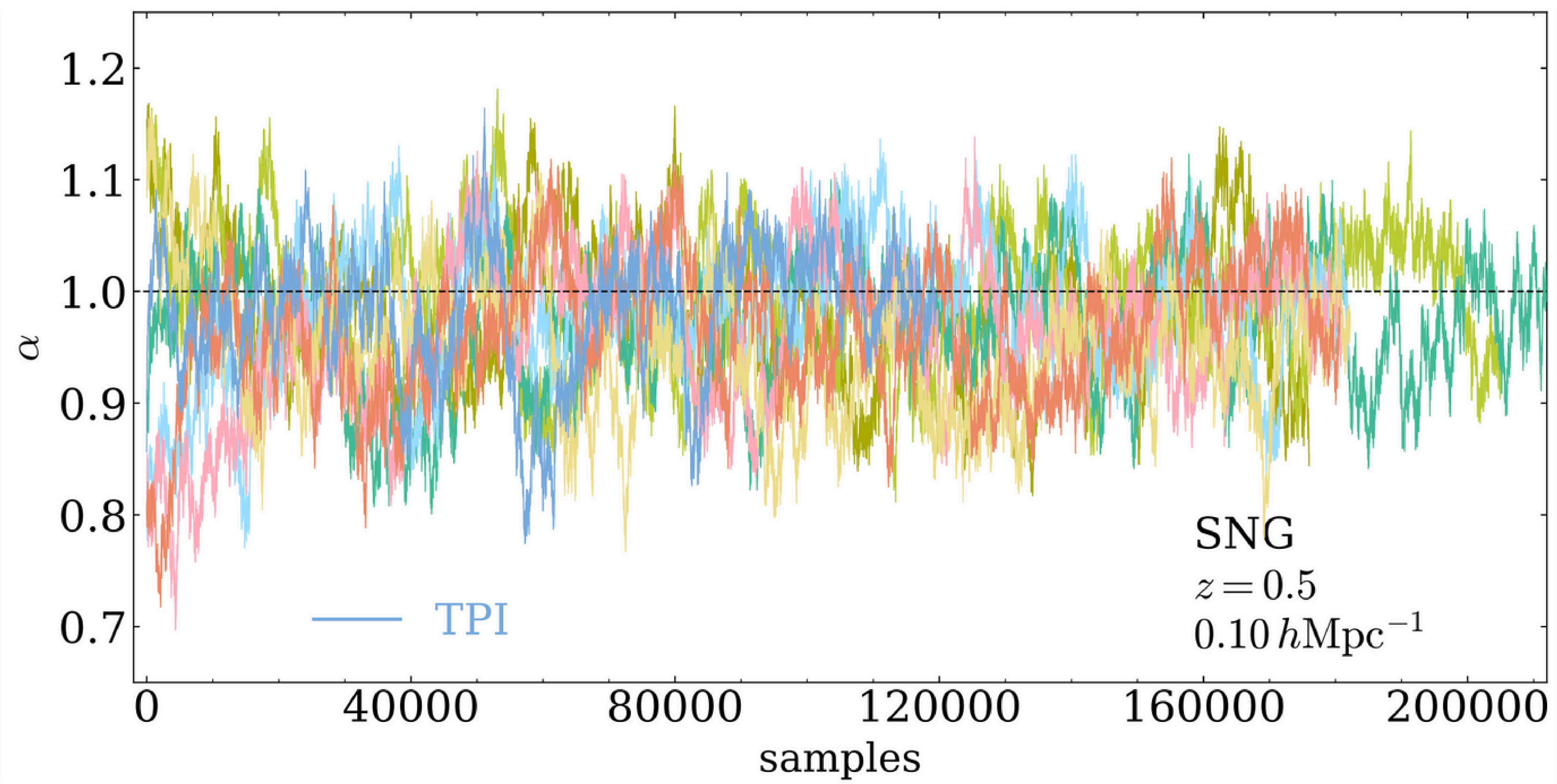
Field level inference results

	SNG	Uchuu
$k_{\rm max} = 0.10\, h{\rm Mpc}^{-1}$	$\alpha = 0.976 \pm 0.056$	$\alpha = 0.941 \pm 0.090$
$k_{\rm max} = 0.12\, h{\rm Mpc}^{-1}$	$\alpha = 1.013 \pm 0.033$	$\alpha = 0.993 \pm 0.053$

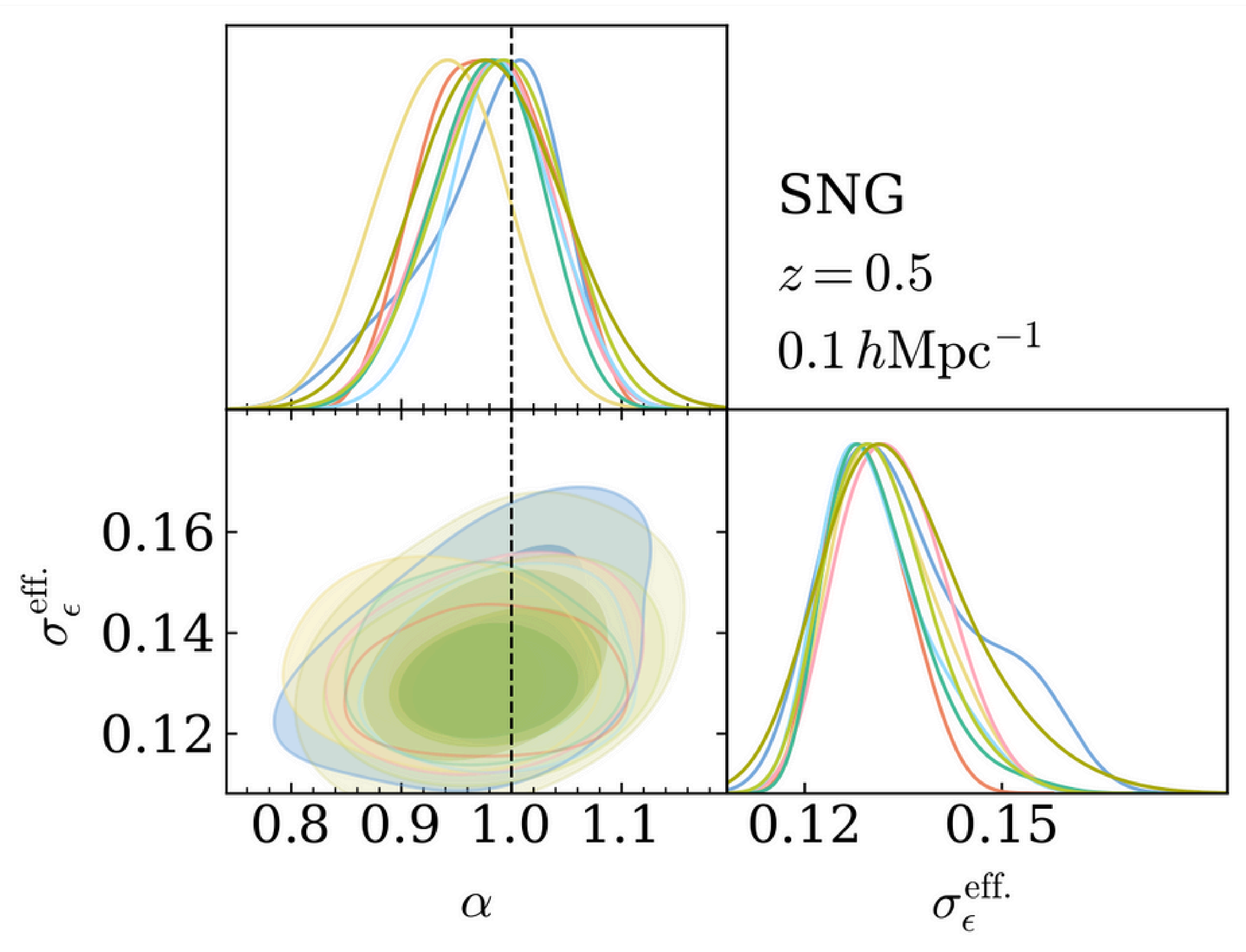


Field level inference results

Trace plot for independent HMC chains



Posterior for independent HMC chains

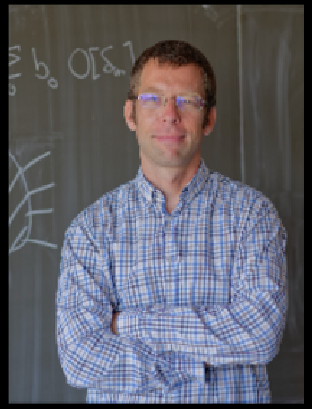


Information content?

How can we quantify the cosmological information extracted from field-level inference?

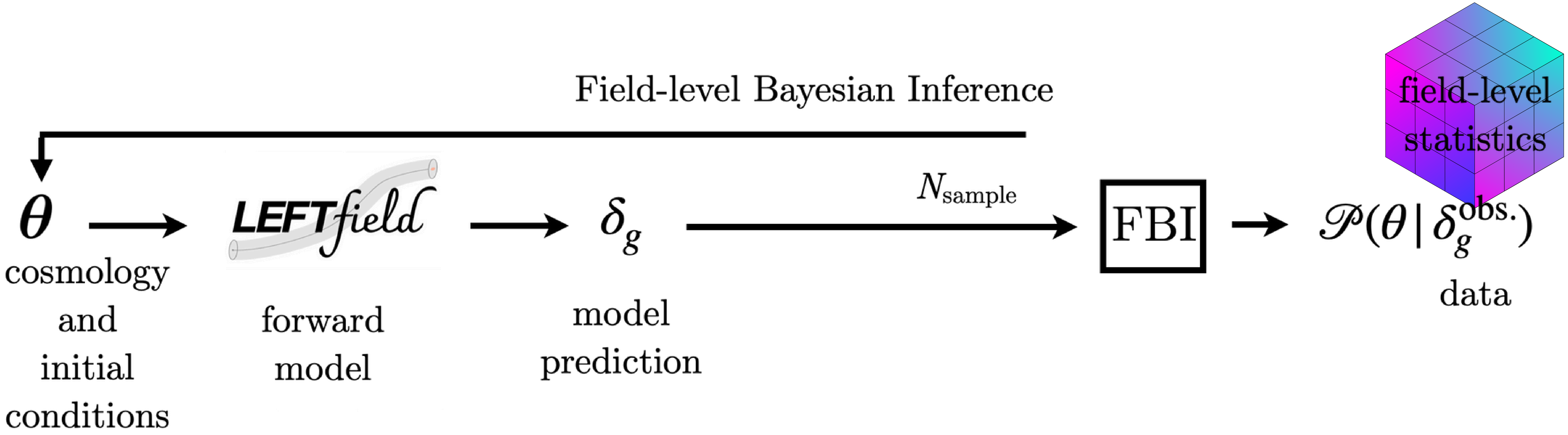
How much information can be extracted from galaxy clustering at the field level?

arXiv:2403.03220



Nhat-Minh Nguyen
(LCTP, University of Michigan)

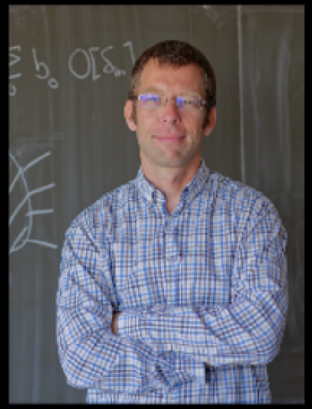
Fabian Schmidt
(MPA)



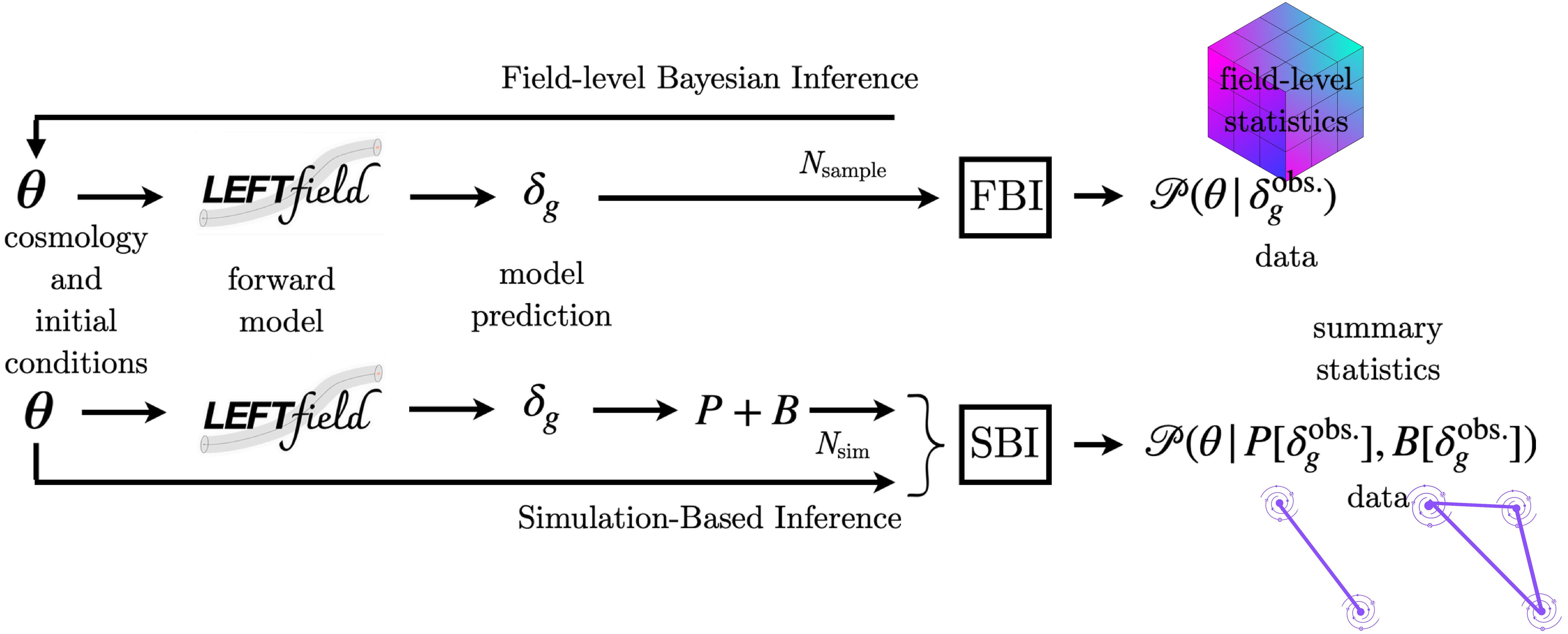
How much information can be extracted from galaxy clustering at the field level?

arXiv:2403.03220

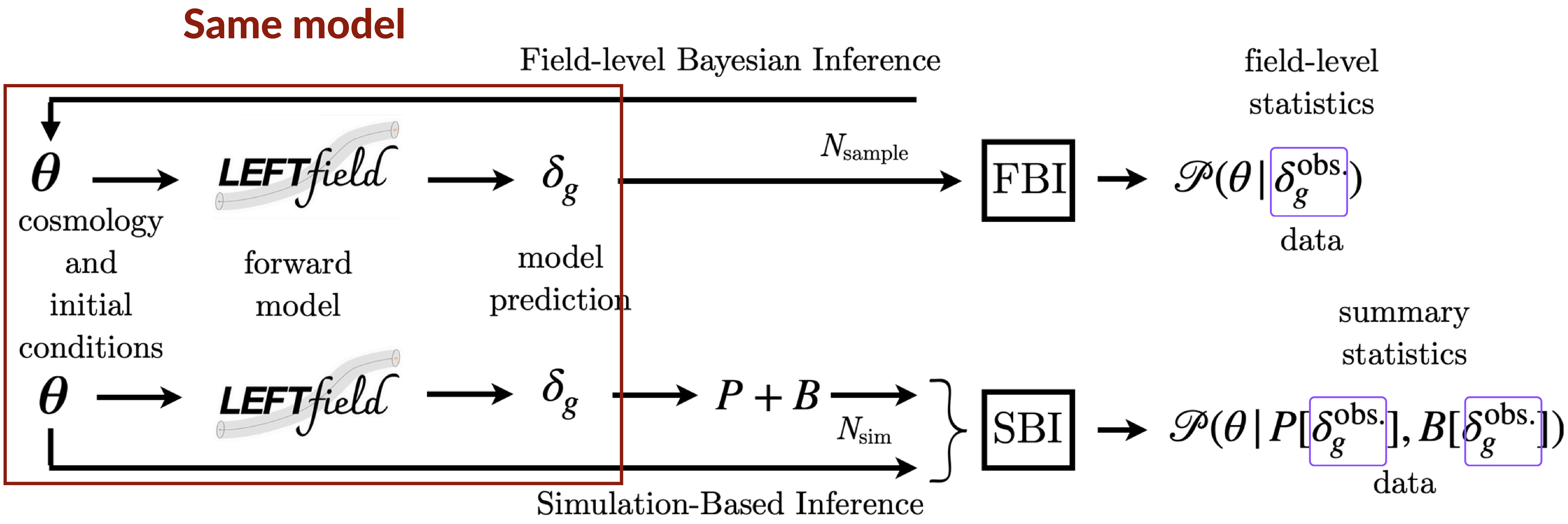
Nhat-Minh Nguyen
(LCTP, University of Michigan)



Fabian Schmidt
(MPA)



Apples-to-apples comparison

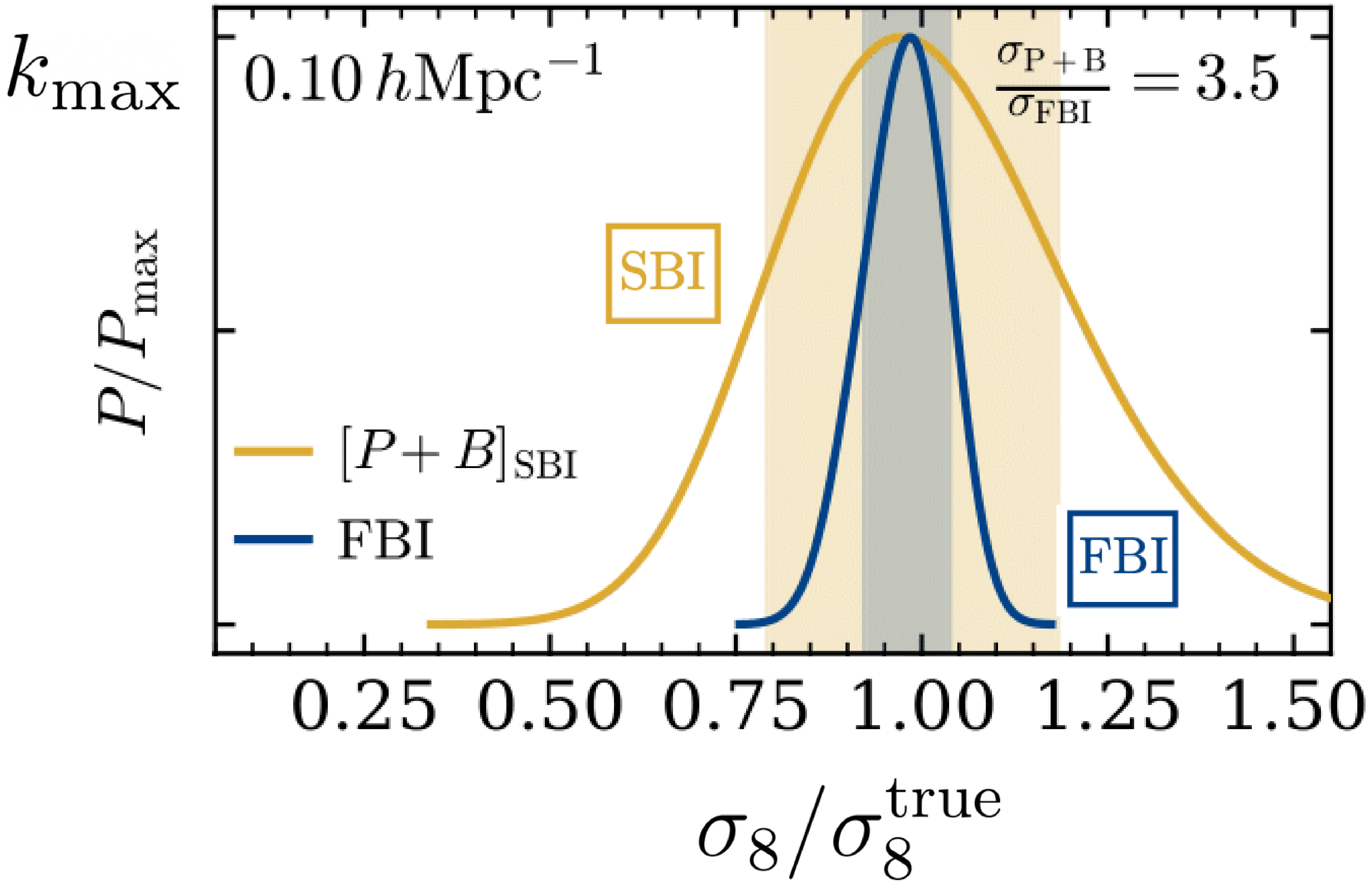


Same halos

Same scale cuts

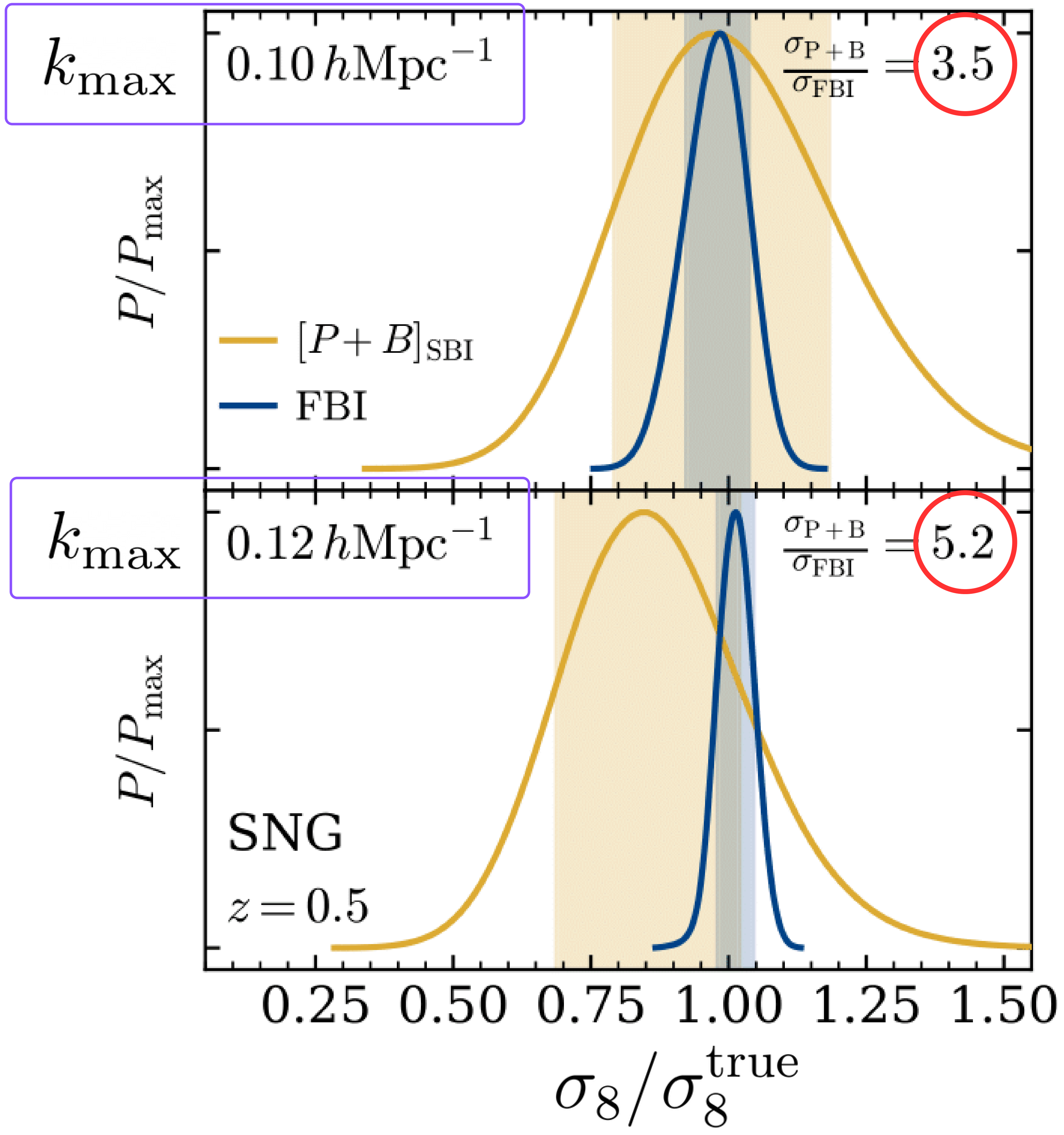
A lot of reliable information at the field-level!

SNG halos



3.5 improvement factor!

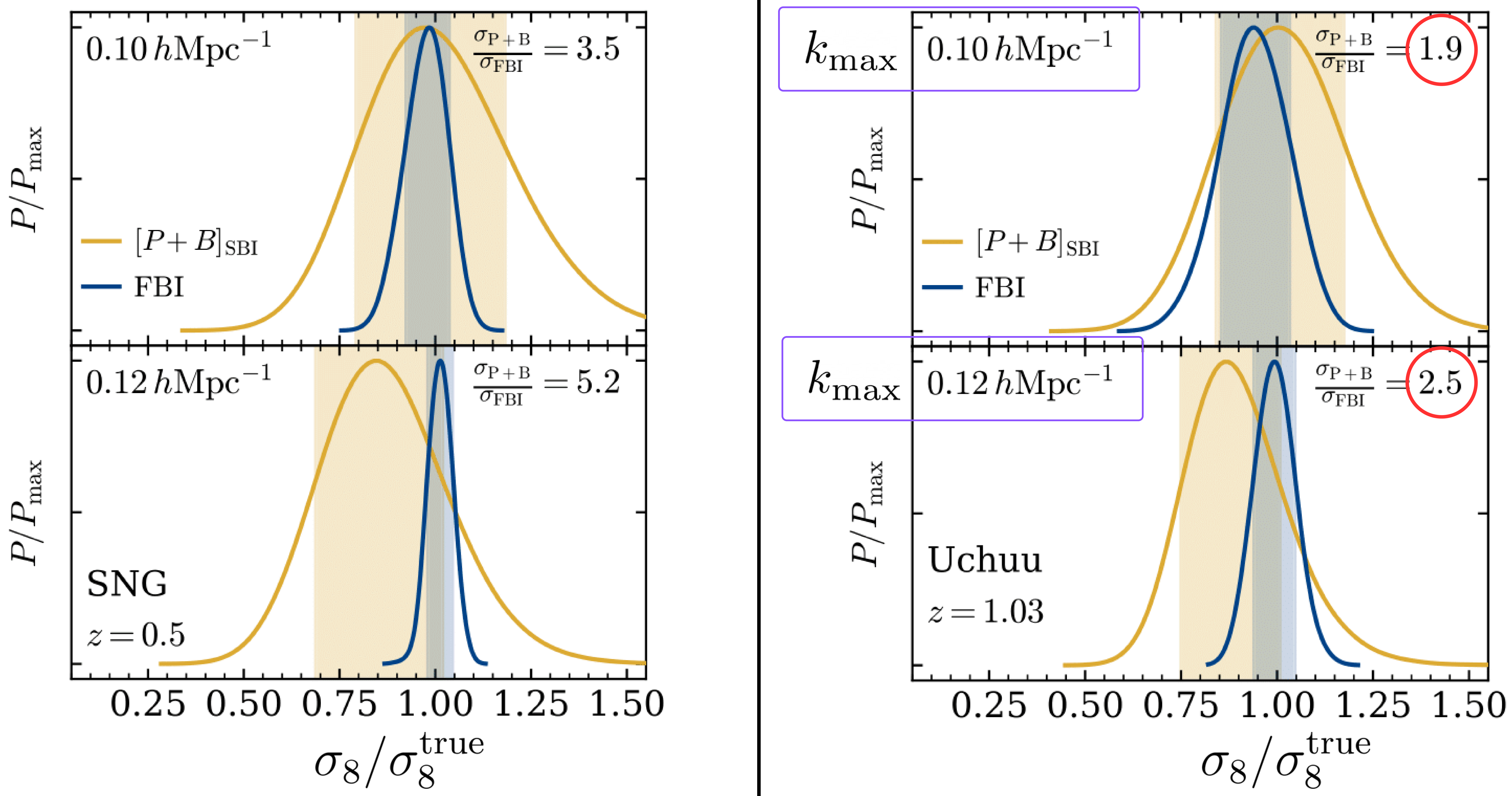
SNG



$k_{\max}$
$0.10\,h\mathrm{Mpc}^{-1}$
$\frac{\sigma_{\mathrm{P+B}}}{\sigma_{\mathrm{FBI}}} = 3.5$
$[P+B]_{\mathrm{SBI}}$
FBI
$k_{\max}$
$0.12\,h\mathrm{Mpc}^{-1}$
$\frac{\sigma_{\mathrm{P+B}}}{\sigma_{\mathrm{FBI}}} = 5.2$
SNG
$z = 0.5$
$P/P_{\max}$
0.25 0.50 0.75 1.00 1.25 1.50
$\sigma_8/\sigma_8^{\mathrm{true}}$

SNG

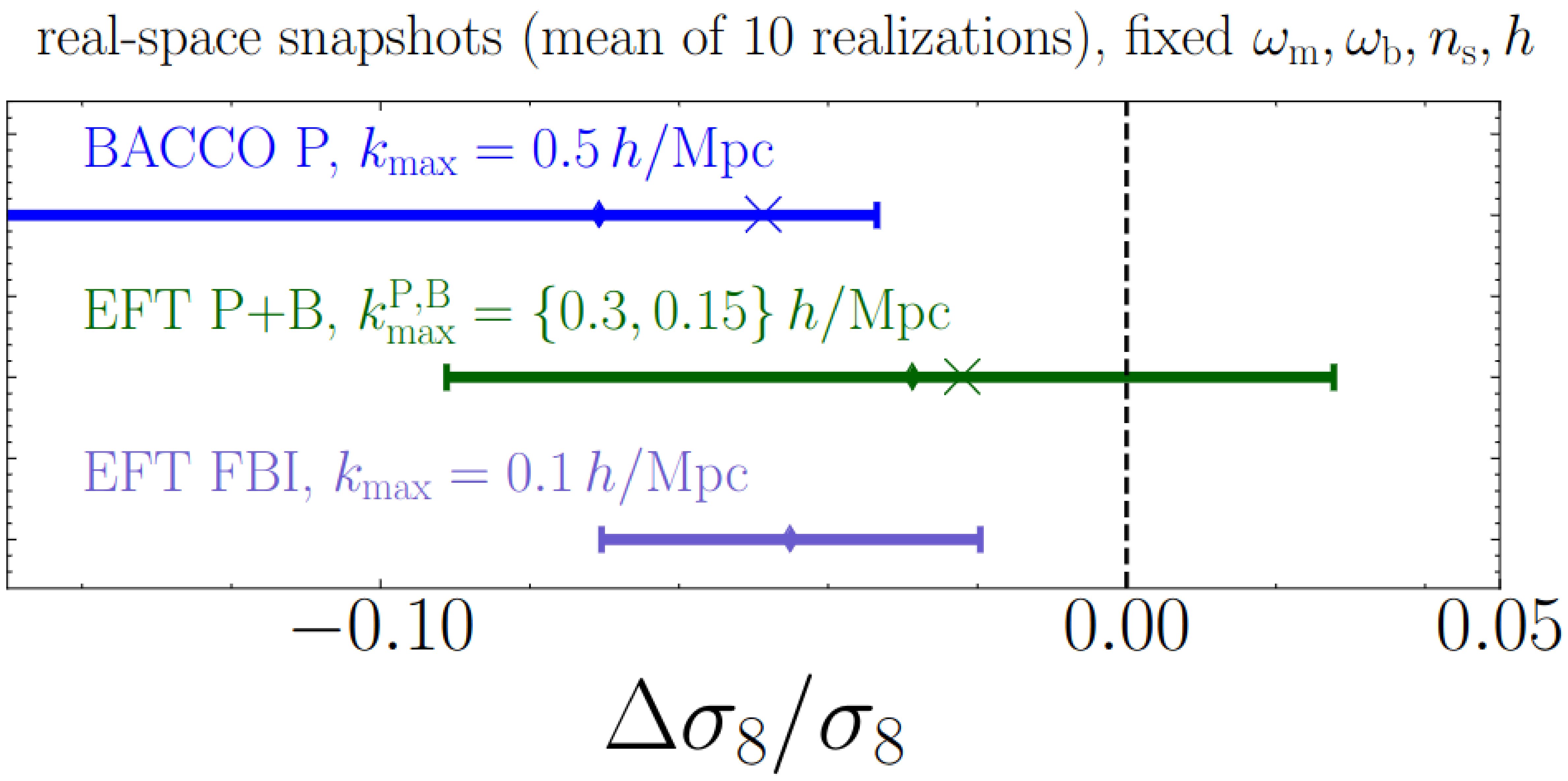
Uchuu



k_{max}
$0.10\,h\mathrm{Mpc}^{-1}$
$\frac{\sigma_{P+B}}{\sigma_{FBI}} = 3.5$
$\frac{\sigma_{P+B}}{\sigma_{FBI}} = 1.9$
$[P+B]_{SBI}$
FBI
k_{max}
$0.12\,h\mathrm{Mpc}^{-1}$
$\frac{\sigma_{P+B}}{\sigma_{FBI}} = 5.2$
$\frac{\sigma_{P+B}}{\sigma_{FBI}} = 2.5$
P/P_{max}
SNG
$z = 0.5$
Uchuu
$z = 1.03$
0.25 0.50 0.75 1.00 1.25 1.50
σ_8/σ_8^{true}

Beyond 2-point mock data challenge

Krause, ..., Nguyen, Schmidt+ (2024)
arXiv:2405.02252



Conclusion & Next Steps

- **Field-level inference** is a very **challenging** task, but very **promising!**
- We demonstrated to have **unbiased** and **accurate** results from halo catalogs.
- **Apple-to-apple comparison** of FBI and SBI P+B shows that there is a lot of **reliable** information beyond 2+3-point functions in the 3D maps of galaxies.

- See **Ivana Babić** talk on Friday on full **field-level BAO scale inference** with LEFTfield!

Next step to connect with observations

- Include more observational effects
- Expand the cosmological parameter space
- More efficient sampling
- Explore summaries in SBI

LEFTfield

Beatriz Tucci
tucci@mpa-garching.mpg.de

arXiv:2403.03220