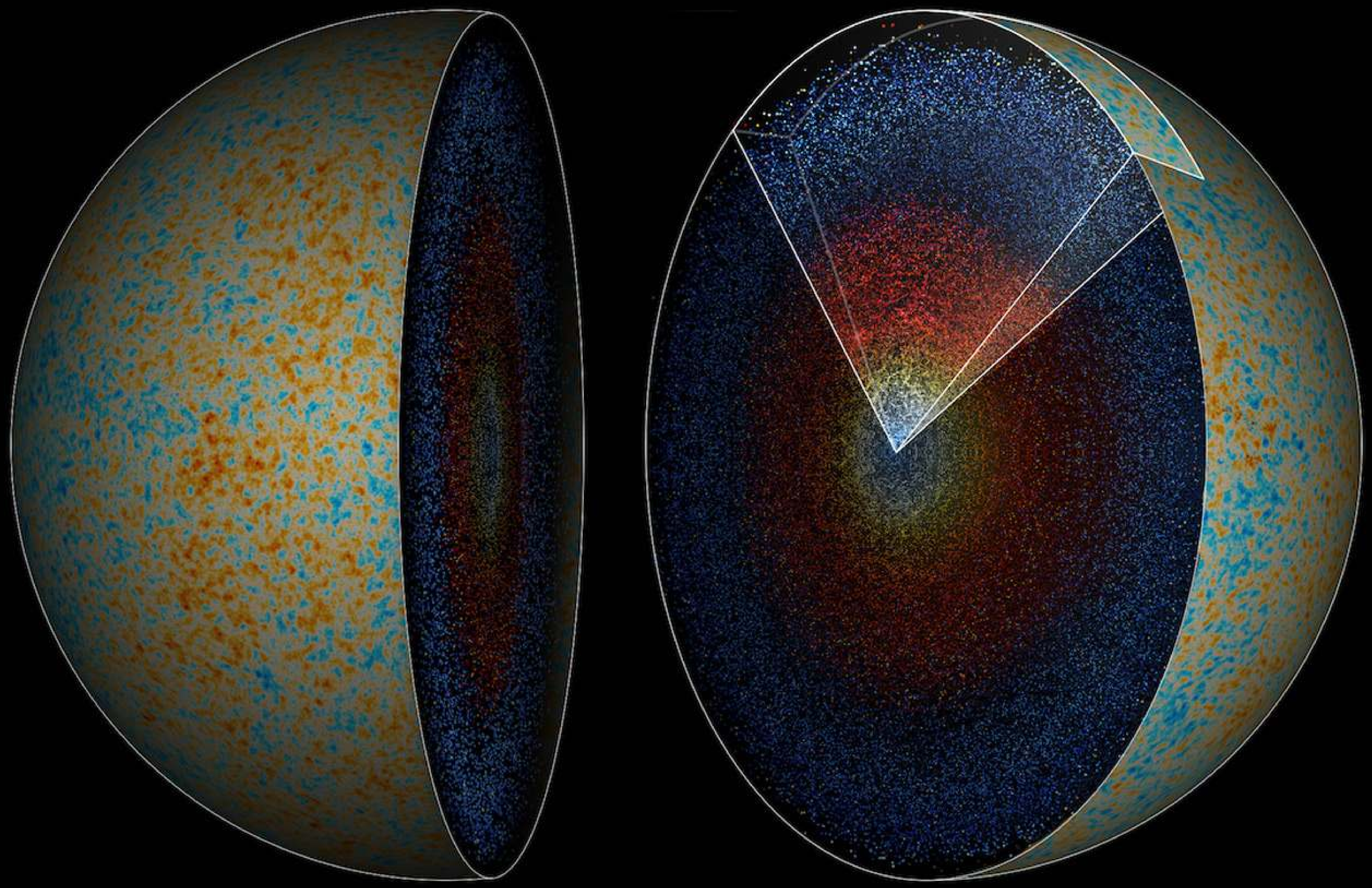
Revealing the information content of galaxy n-point functions with simulation-based inference

Beatriz Tucci

Max Planck Institute for Astrophysics (MPA)

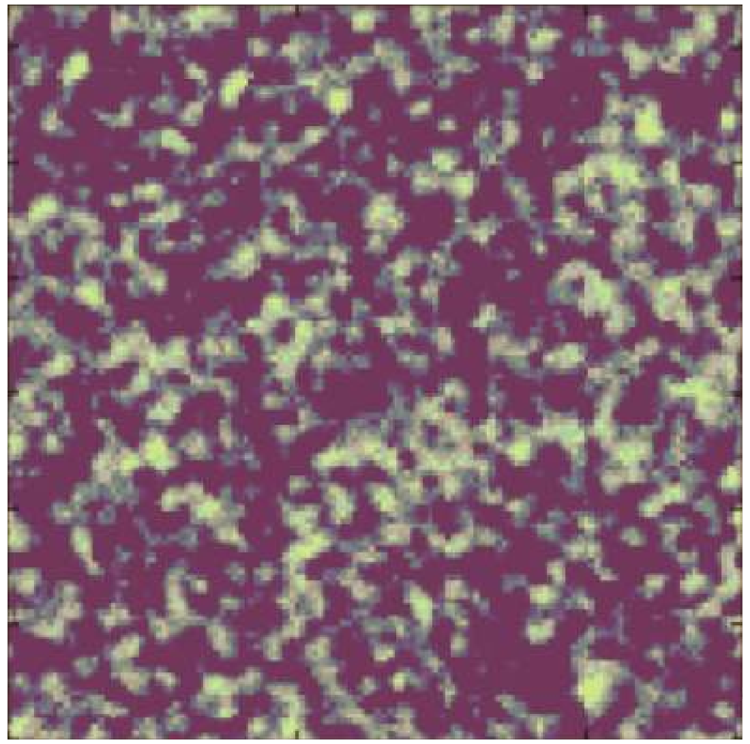
with Fabian Schmidt, Nhat-Minh Nguyen,
Ivana Babić, Andrija Kostić, Martin Reinecke

BCCP Cosmology Seminar 2024

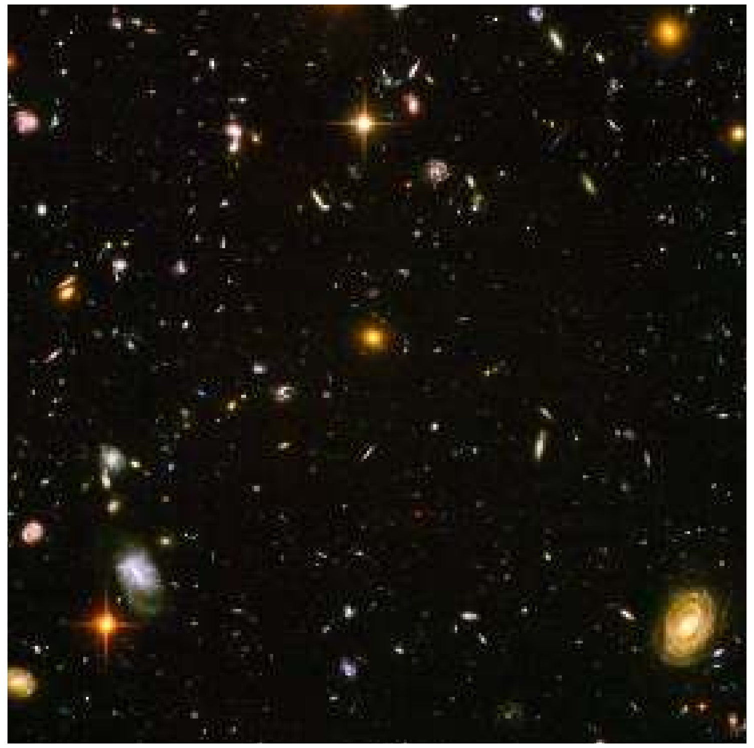


Which information does the 3D distribution of galaxies give us?

(Slice of 3D field)



time

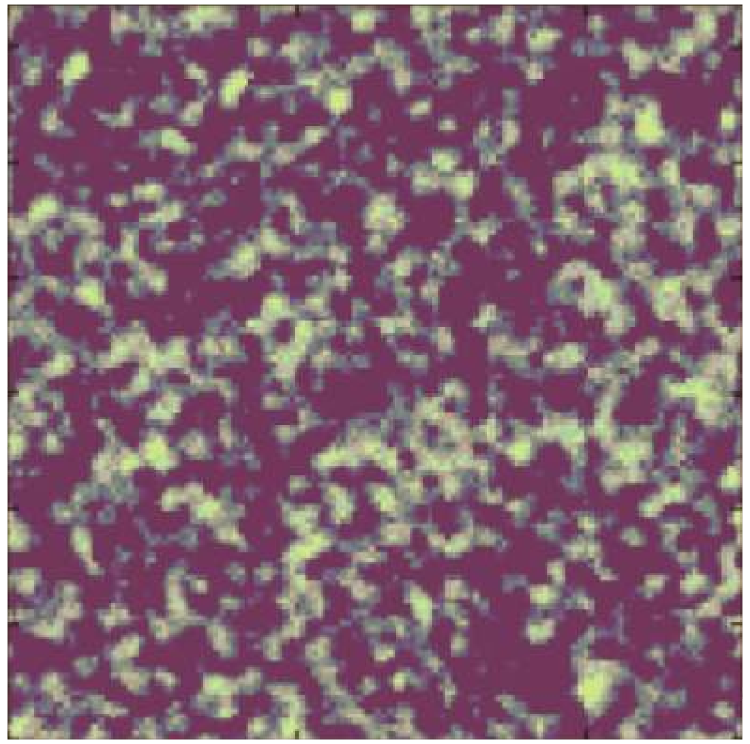


Initial conditions
of the Universe

Final distribution
of galaxies

Which information does the 3D distribution of galaxies give us?

Other particles
*during **inflation**?*

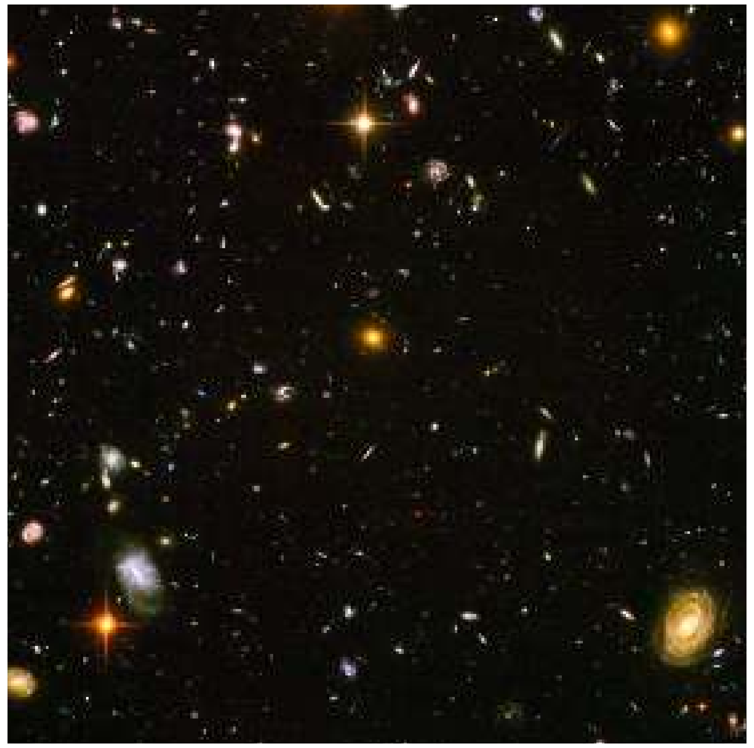


Initial conditions
of the Universe

*Nature of **dark matter***
*and **dark energy**?*

Ultimate theory
*of **gravity**?*

Hierarchy of
***Neutrino** masses?*

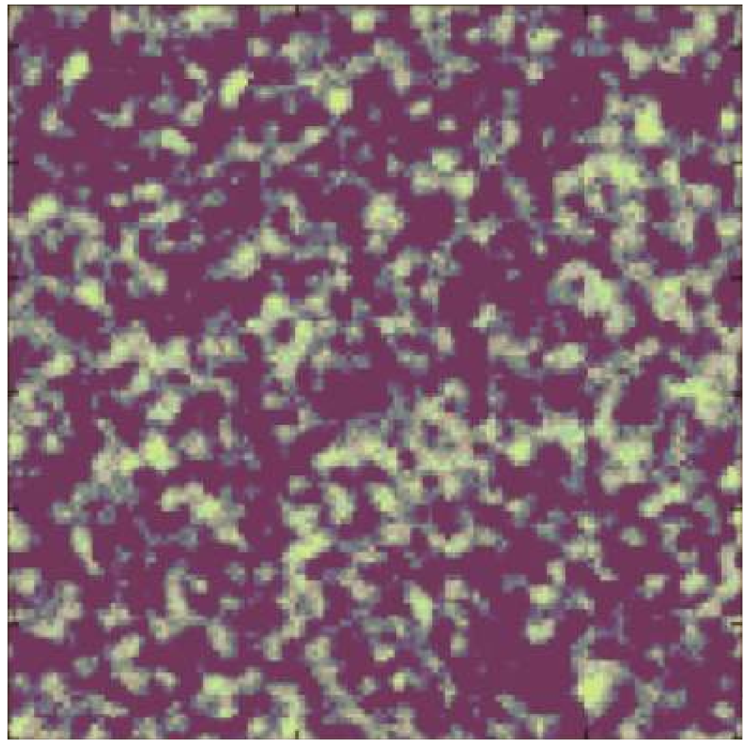


Final distribution
of galaxies

Which information does the 3D distribution of galaxies give us?

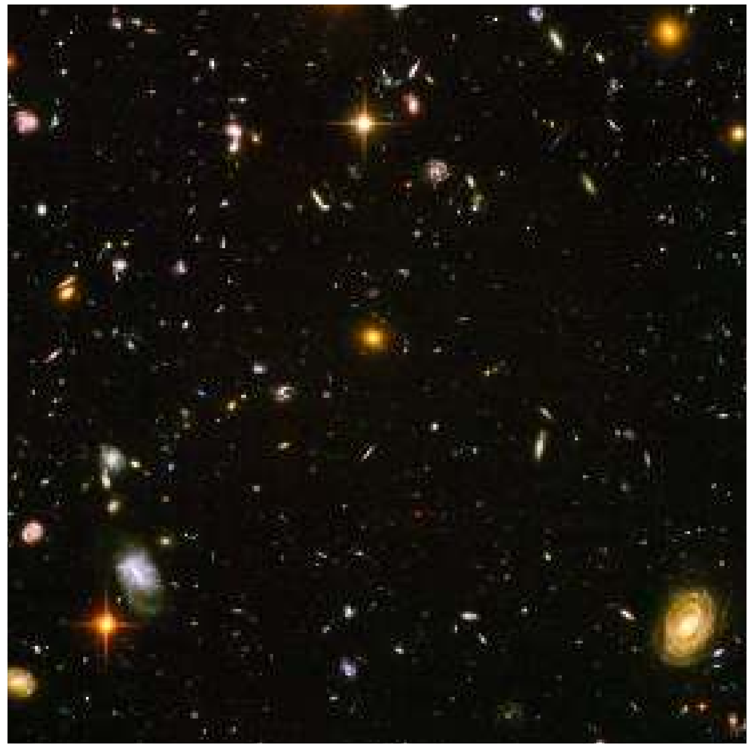
Other particles
during ***inflation****?*

Nature of ***dark matter***
and ***dark energy****?*



Ultimate theory
of ***gravity****?*

Hierarchy of
Neutrino *masses?*



Initial conditions
of the Universe

Cosmological
parameters θ

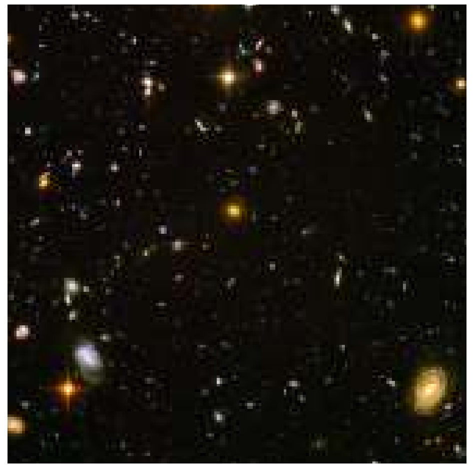
Final distribution
of galaxies

Part 0

Inferring the cosmological parameters:

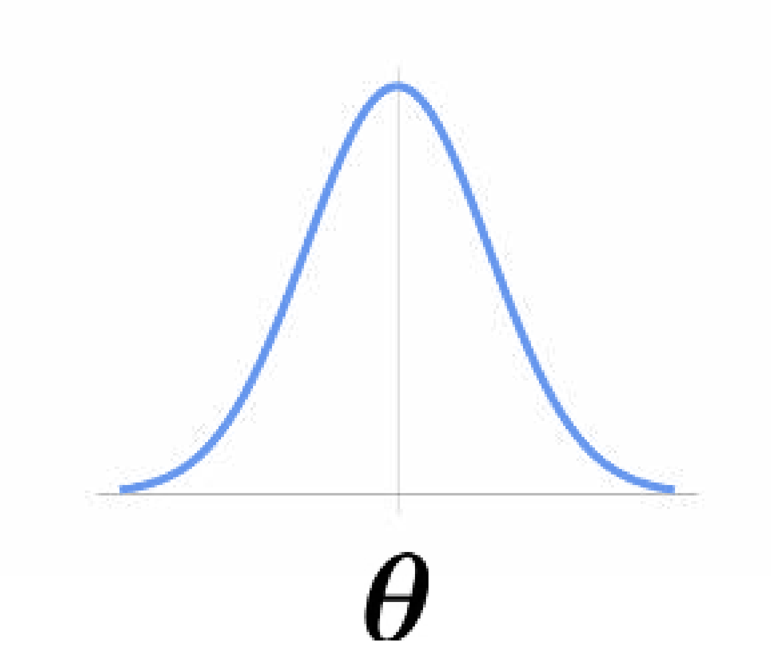
standard techniques & challenges

Observation



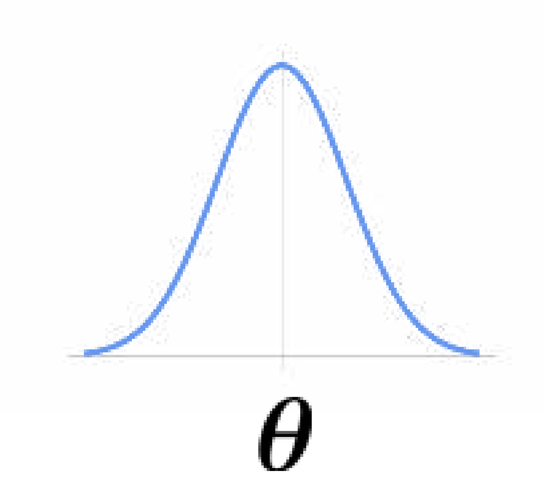
inference

theory(θ)



Posterior of cosmological parameters

Bayesian inference



Posterior Likelihood Prior

$$\boxed{\mathcal{P}(\boldsymbol{\theta}|\mathbf{D}) \propto \mathcal{L}(\mathbf{D}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}$$

Bayesian inference

Posterior Likelihood Prior

$$\boxed{\mathcal{P}(\boldsymbol{\theta}|\mathbf{D}) \propto \mathcal{L}(\mathbf{D}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}$$

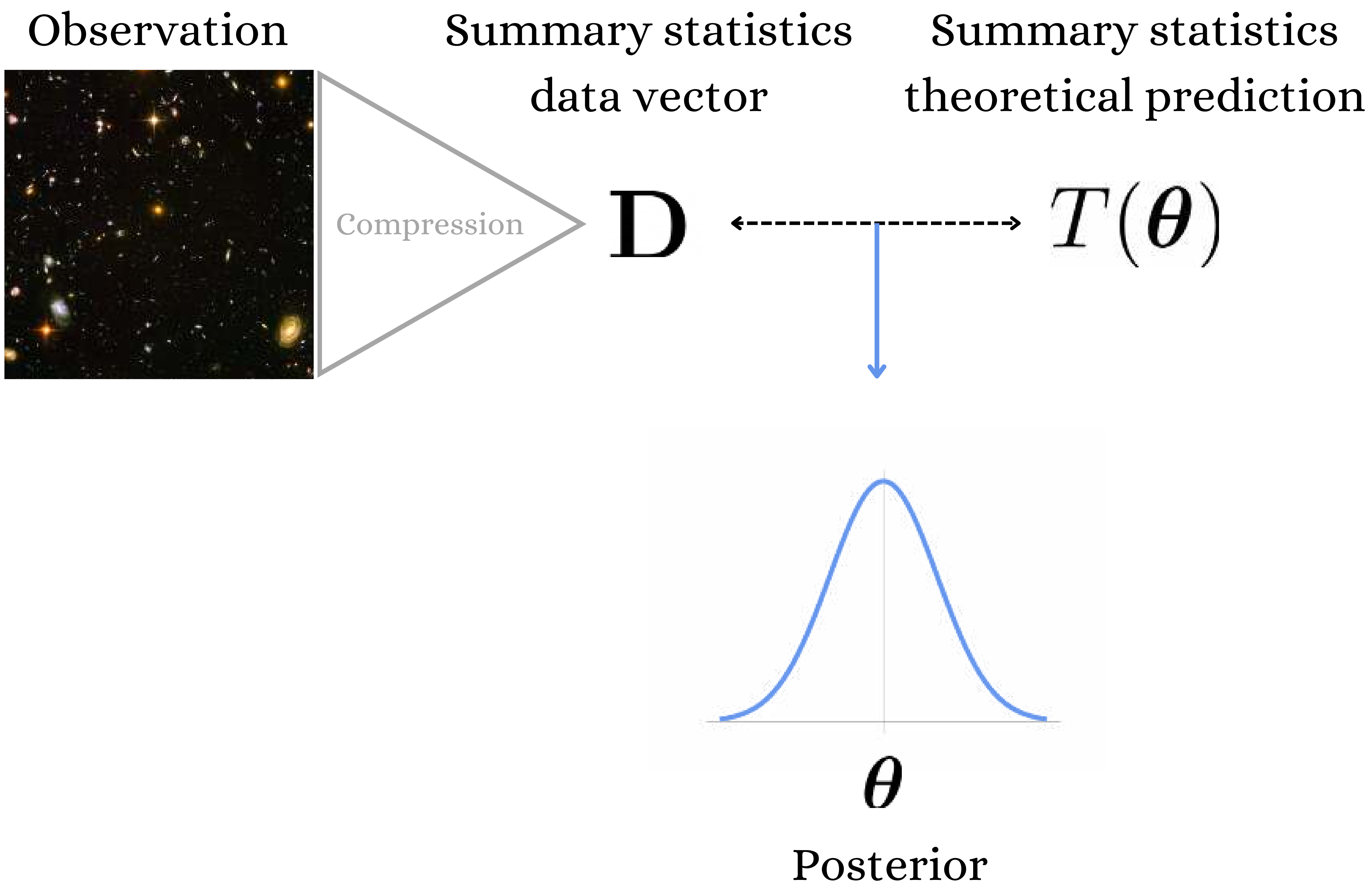
E.g., assuming that the data vector is Gaussian distributed:

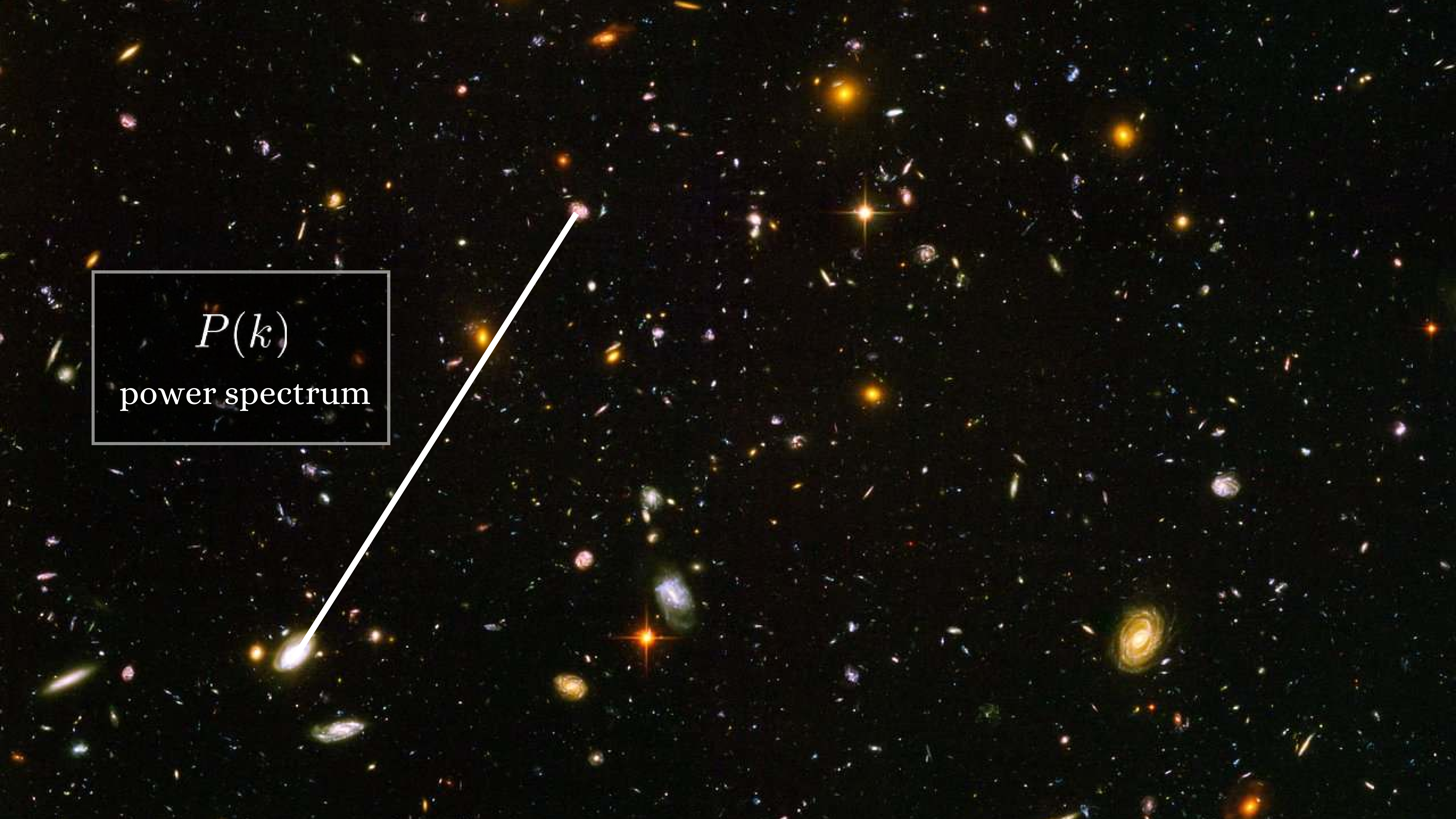
$$-2\ln\mathcal{L}(\mathbf{D}|\boldsymbol{\theta}) = (\mathrm{D} - \mathbf{T}(\boldsymbol{\theta}))\cdot\mathbf{C}^{-1}\cdot(\mathrm{D} - \mathbf{T}(\boldsymbol{\theta}))$$

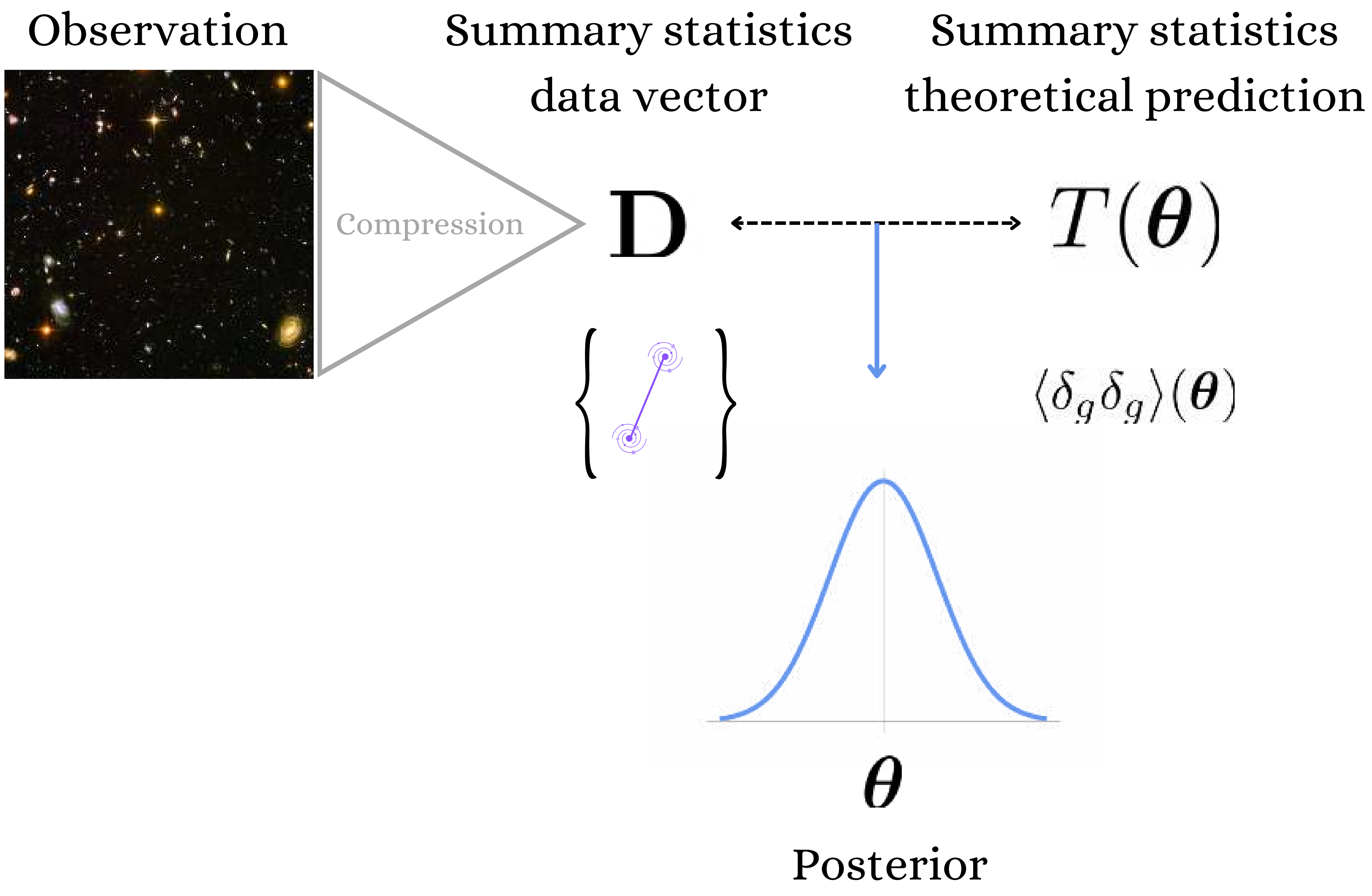
Data vector

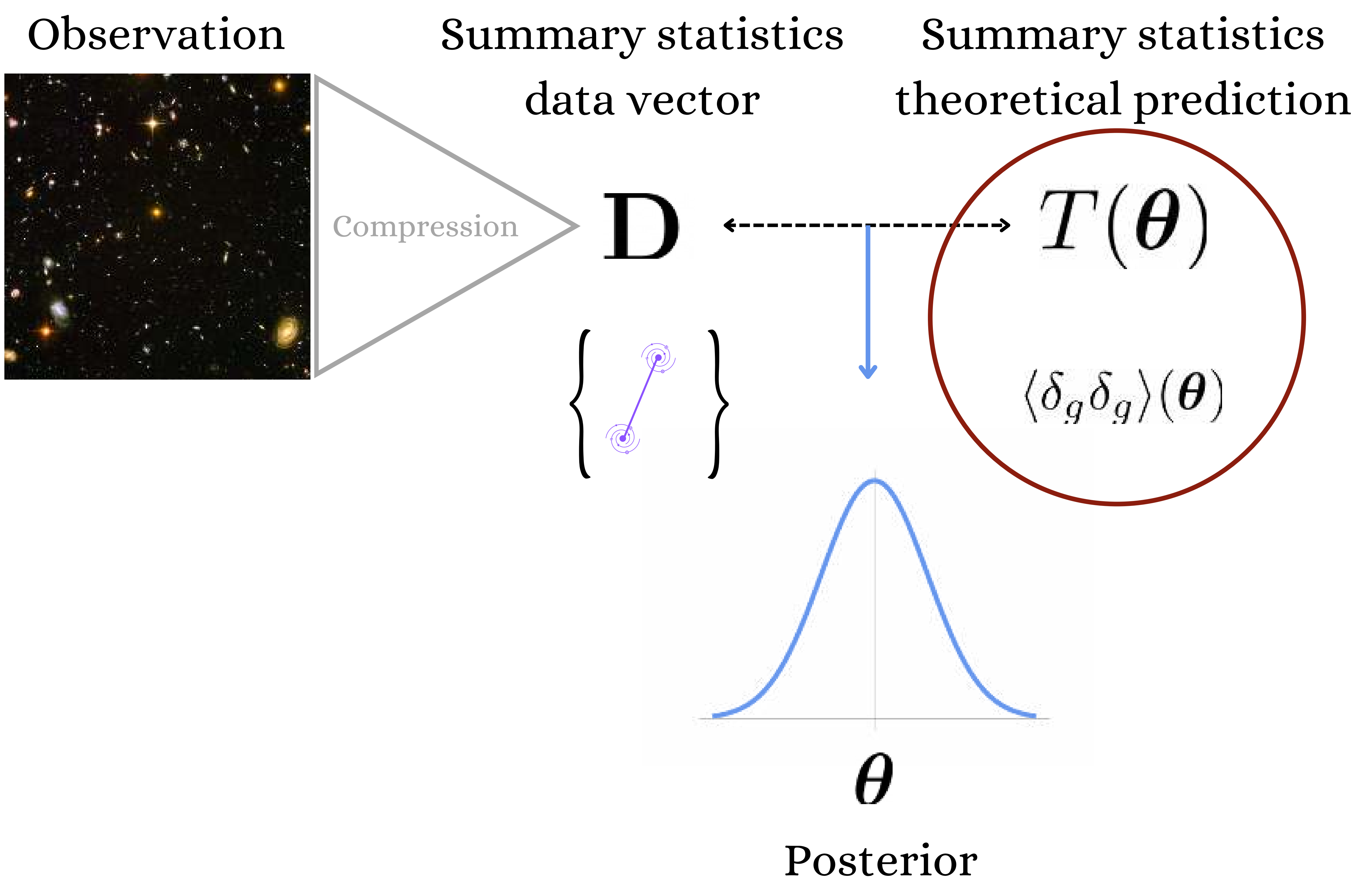
Covariance of the data vector

Theoretical prediction of data vector


Observation
Summary statistics
data vector
Summary statistics
theoretical prediction
Compression
$\mathbf{D}$
$T(\boldsymbol{\theta})$
$\boldsymbol{\theta}$
Posterior


$P(k)$
power spectrum


Observation
Summary statistics
data vector
Summary statistics
theoretical prediction
Compression
$\mathbf{D}$
$T(\boldsymbol{\theta})$
$\langle \delta_g \delta_g \rangle (\boldsymbol{\theta})$
$\boldsymbol{\theta}$
Posterior


Observation
Summary statistics
data vector
Summary statistics
theoretical prediction
Compression
$\mathbf{D}$
$T(\boldsymbol{\theta})$
$\langle\delta_g\delta_g\rangle(\boldsymbol{\theta})$
$\boldsymbol{\theta}$
Posterior

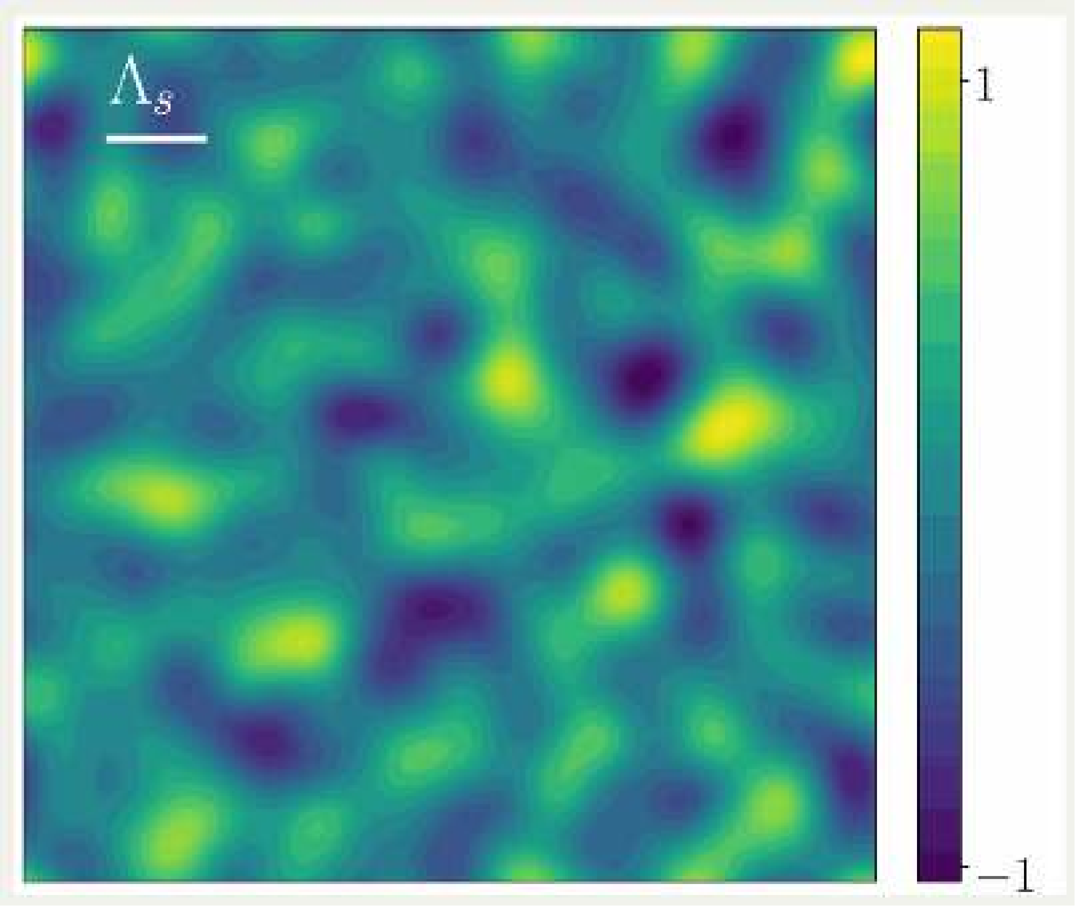
The EFTofLSS:

our theory for galaxy clustering

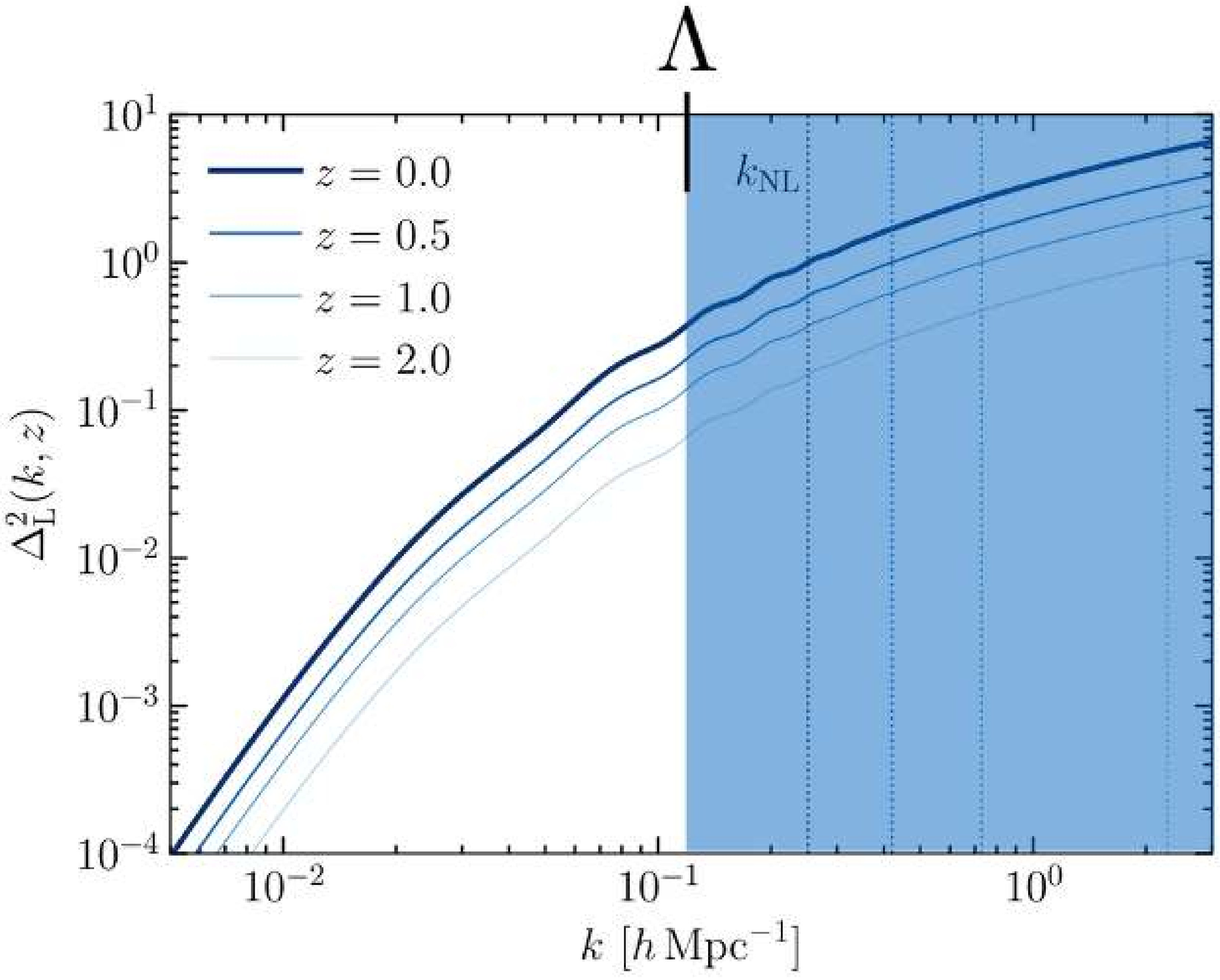
The EFTofLSS

“coarse-graining”

$$\boxed{\delta_\Lambda^{(1)}(\boldsymbol{k}) = W_\Lambda(k)\delta^{(1)}(\boldsymbol{k})}$$



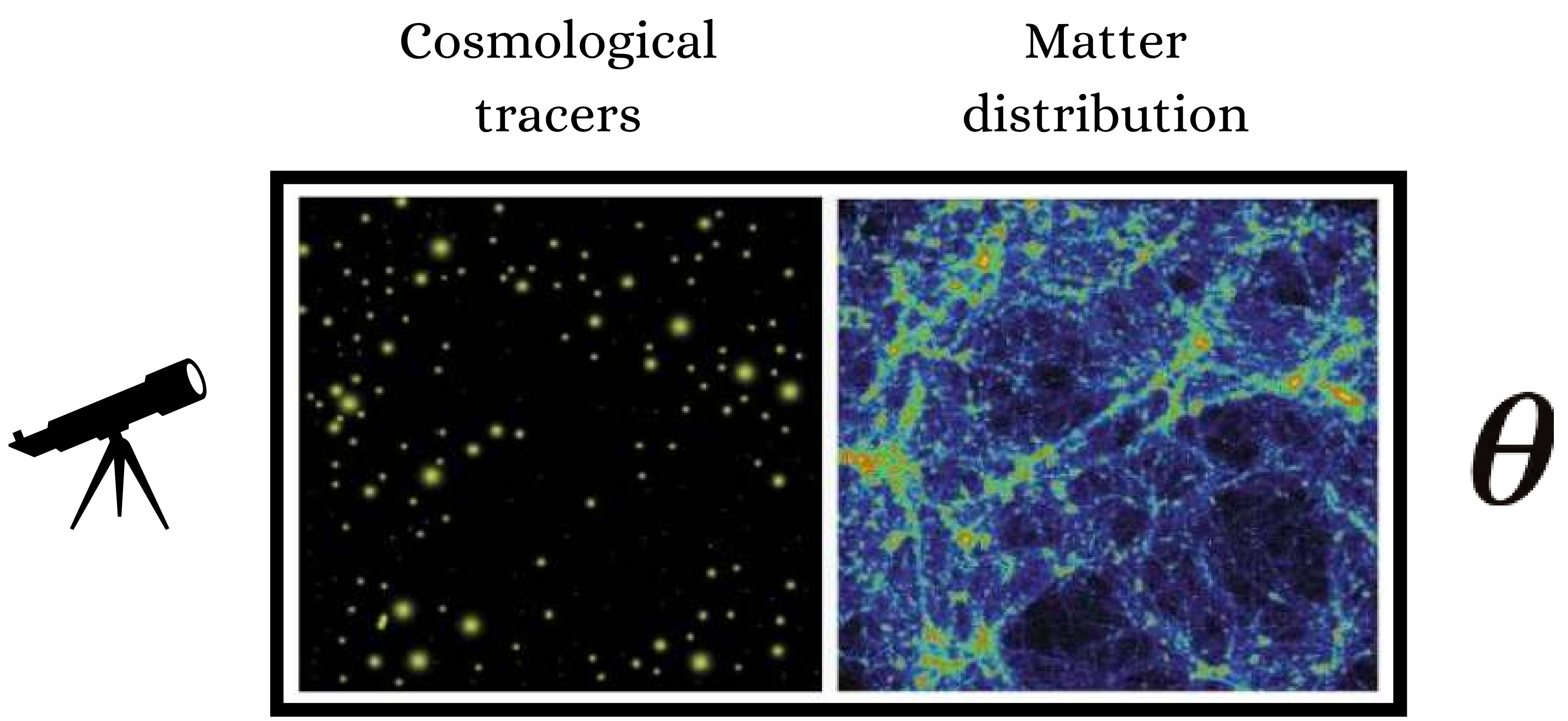
Borrowed from Pierre Zhang



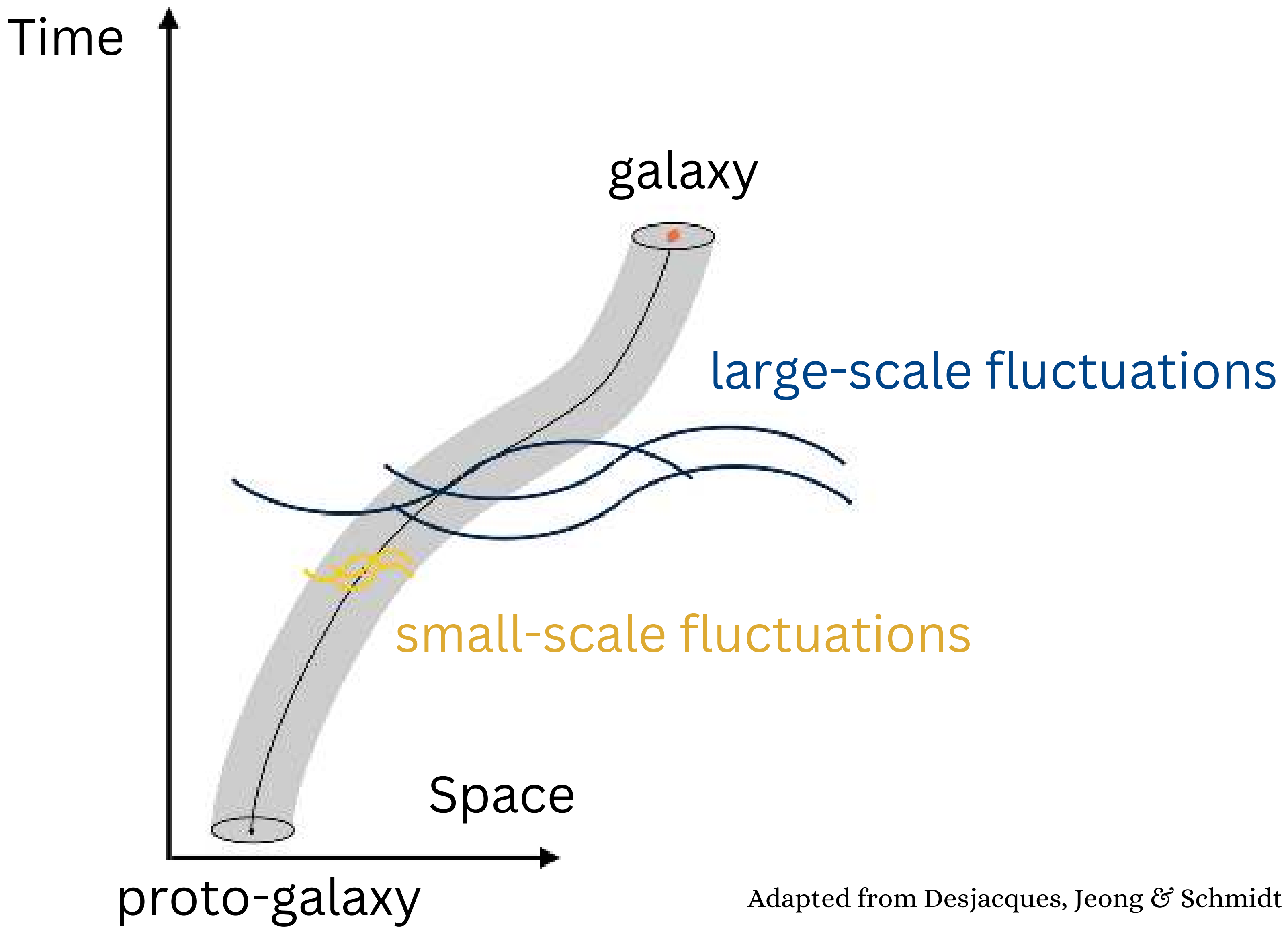
Dodelson & Schmidt
Modern Cosmology 2020

The bias expansion

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



Cooray & Sheth (2002)



Adapted from Desjacques, Jeong & Schmidt (2016)

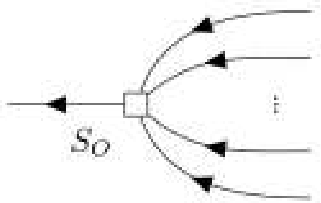
The bias expansion

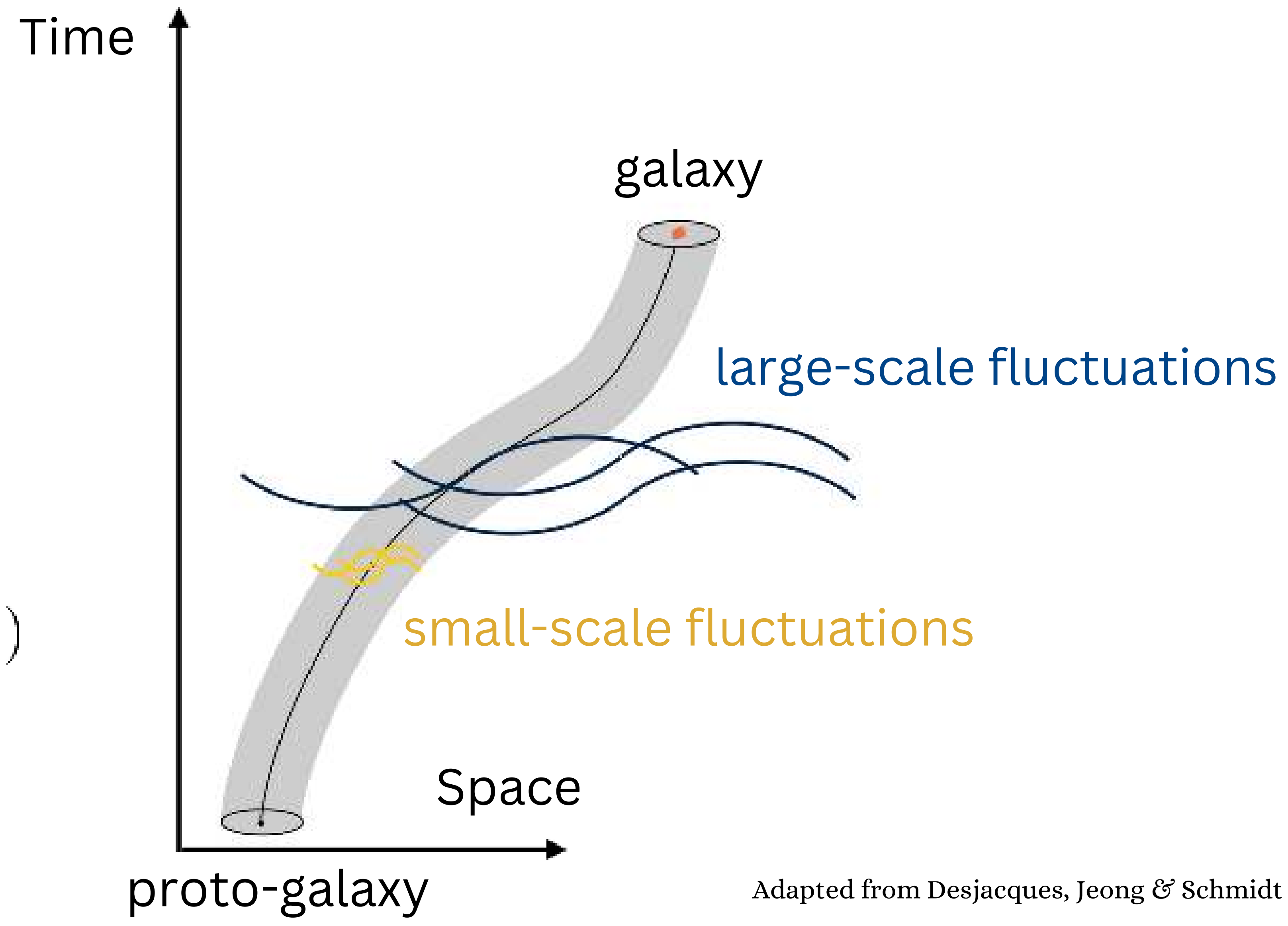
$$
\begin{aligned}
\delta_g(\boldsymbol{k}, z) &= \boxed{\delta_{g,\mathrm{det}}(\boldsymbol{k}, z)} + \delta_{g,\mathrm{stoch}}(\boldsymbol{k}, z) \\
&= \boxed{\sum_O b_O(z) O(\boldsymbol{k}, z)} + \varepsilon(\boldsymbol{k}, z)
\end{aligned}
$$

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

$$
O[\delta](\boldsymbol{k}) = \int_{\boldsymbol{p}_1,\ldots,\boldsymbol{p}_n} \delta_{\mathrm{D}}\left(\boldsymbol{k} - \boldsymbol{p}_{1\ldots n}\right) S_O\left(\boldsymbol{p}_1, \ldots \boldsymbol{p}_n\right) \delta\left(\boldsymbol{p}_1\right) \cdots \delta\left(\boldsymbol{p}_n\right)
$$

operator “convolution”





Adapted from Desjacques, Jeong & Schmidt (2016)

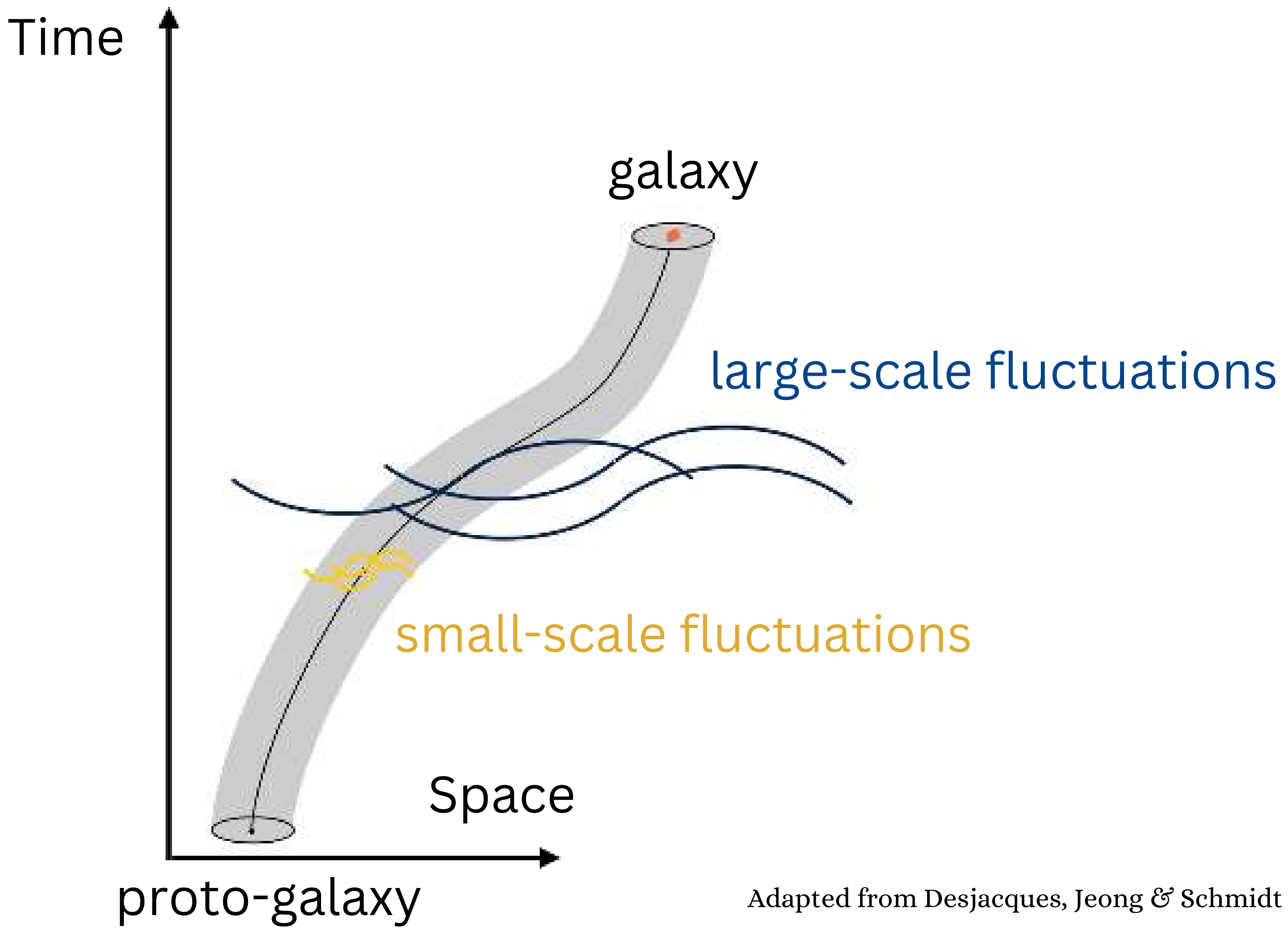
The bias expansion

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

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

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

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

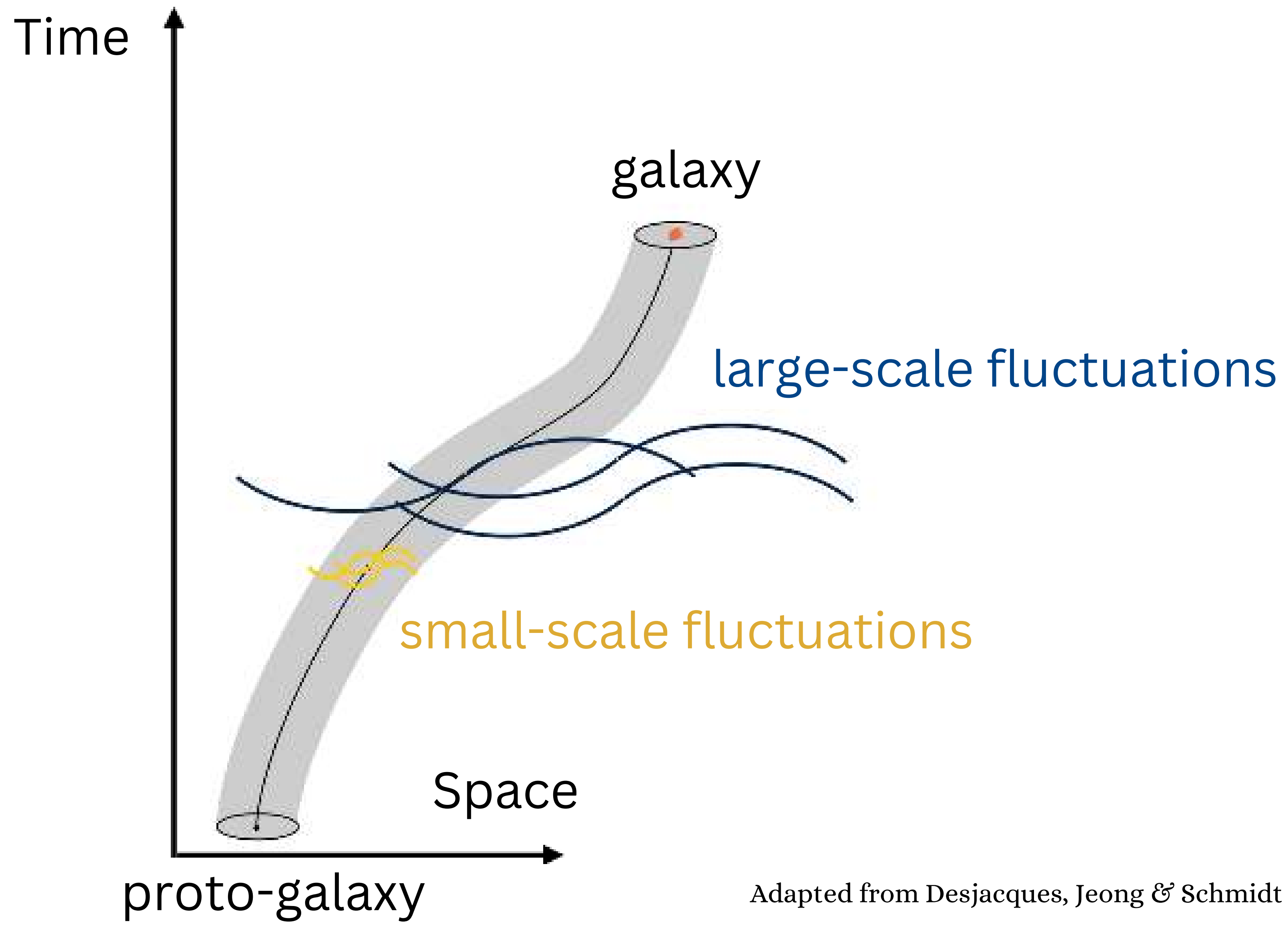


Adapted from Desjacques, Jeong & Schmidt (2016)

The bias expansion

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

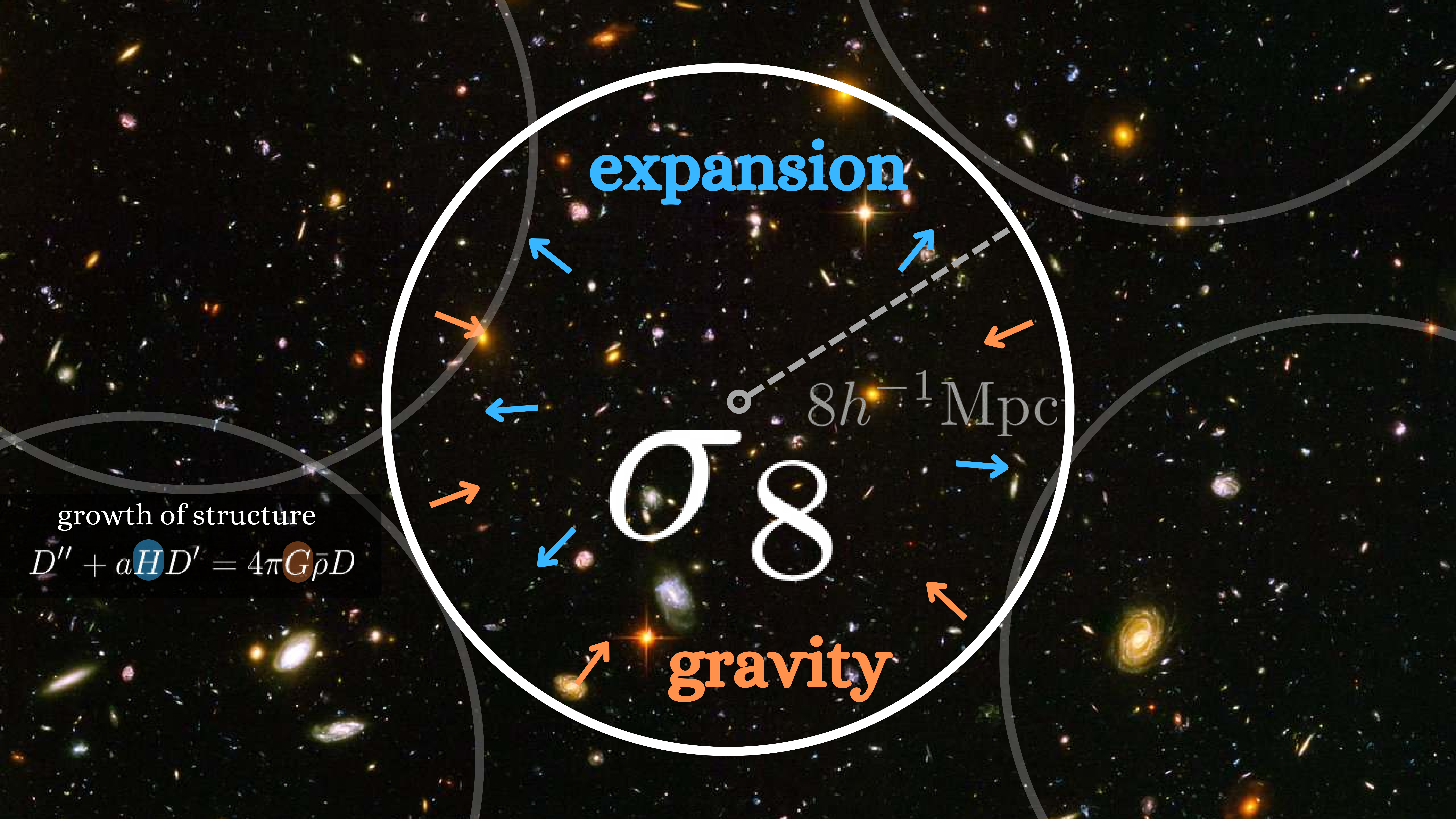
$$\boxed{\{\boldsymbol{\theta}, \{b_O\}, \{\sigma_\varepsilon\}\}}$$



Adapted from Desjacques, Jeong & Schmidt (2016)

Case study:

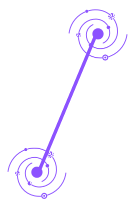
inferring the cosmological parameter σ_8


expansion
$8h^{-1}\mathrm{Mpc}$
σ_8
gravity
growth of structure
$D'' + aHD' = 4\pi G\bar{\rho}D$

Inferring σ_8 with the power-spectrum

$$T(\boldsymbol{\theta})$$

$$P_g(k) = \langle \delta_g(\boldsymbol{k}) \delta_g(\boldsymbol{k}') \rangle'$$

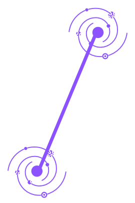


Inferring σ_8 with the power-spectrum

$$T(\boldsymbol{\theta})$$

$$\delta_g(\boldsymbol{k}) = b_1 \delta(\boldsymbol{k}) + \varepsilon(\boldsymbol{k})$$

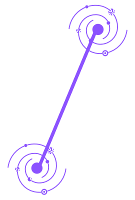
$$P_g(k) = \langle \delta_g(\boldsymbol{k}) \delta_g(\boldsymbol{k}') \rangle'$$



Inferring σ_8 with the power-spectrum

$$T(\boldsymbol{\theta})$$

$$P_g(k) = \langle \delta_g(\boldsymbol{k}) \delta_g(\boldsymbol{k}') \rangle'$$



$$\delta_g(\boldsymbol{k}) = b_1 \delta(\boldsymbol{k}) + \varepsilon(\boldsymbol{k})$$

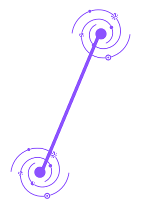
$$P_g^{\mathrm{tree}}(k) = b_1^2 P_L(k) + P_\varepsilon$$

$$P_L(k) = \langle \delta^{(1)}(\boldsymbol{k}) \delta^{(1)}(\boldsymbol{k}') \rangle'$$

Inferring σ_8 with the power-spectrum

$$T(\boldsymbol{\theta})$$

$$P_g(k) = \langle \delta_g(\boldsymbol{k}) \delta_g(\boldsymbol{k}') \rangle'$$



$$\delta_g(\boldsymbol{k}) = b_1 \delta(\boldsymbol{k}) + \varepsilon(\boldsymbol{k})$$

$$P_g^{\mathrm{tree}}(k) = b_1^2 P_L(k) + P_\varepsilon$$

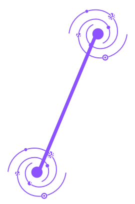
$$P_L(k) = \langle \delta^{(1)}(\boldsymbol{k}) \delta^{(1)}(\boldsymbol{k}') \rangle'$$

$$\propto \sigma_8^2$$

Inferring σ_8 with the power-spectrum

$$T(\boldsymbol{\theta})$$

$$P_g(k) = \langle \delta_g(\boldsymbol{k}) \delta_g(\boldsymbol{k}') \rangle'$$



$$\delta_g(\boldsymbol{k}) = b_1 \delta(\boldsymbol{k}) + \varepsilon(\boldsymbol{k})$$

$$P_g^{\mathrm{tree}}(k) = b_1^2 P_L(k) + P_\varepsilon$$

$$P_L(k) = \langle \delta^{(1)}(\boldsymbol{k}) \delta^{(1)}(\boldsymbol{k}') \rangle'$$

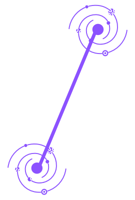
$$\propto \sigma_8^2$$

Bias parameter and σ_8 are degenerated in the tree-level galaxy power-spectrum

Inferring σ_8 with the power-spectrum

$$T(\boldsymbol{\theta})$$

$$P_g(k) = \langle \delta_g(\boldsymbol{k}) \delta_g(\boldsymbol{k}') \rangle'$$



$$\delta_g(\boldsymbol{k}) = b_1 \delta(\boldsymbol{k}) + \varepsilon(\boldsymbol{k})$$

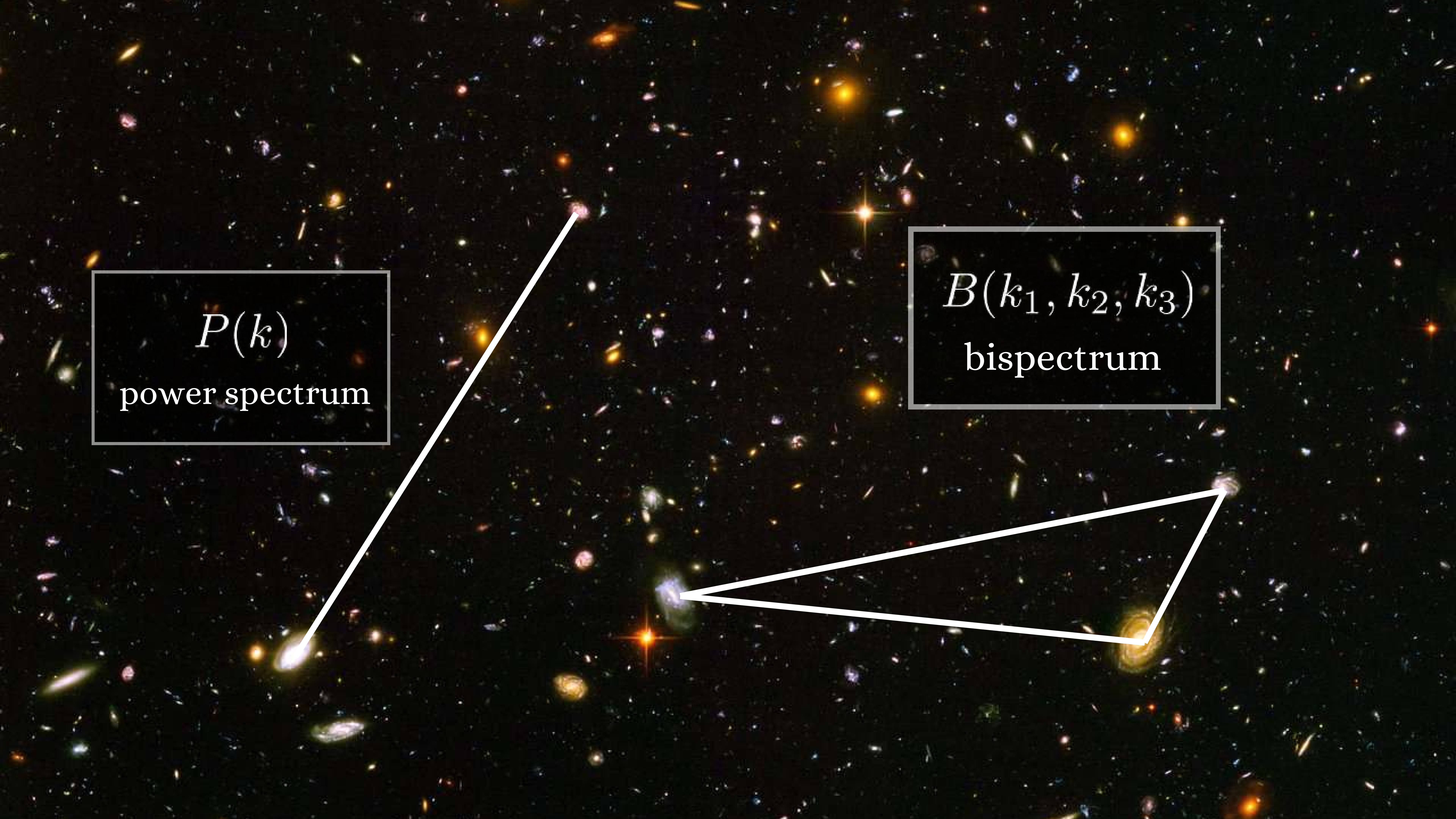
$$P_g^{\mathrm{tree}}(k) = b_1^2 P_L(k) + P_\varepsilon$$

$$P_L(k) = \langle \delta^{(1)}(\boldsymbol{k}) \delta^{(1)}(\boldsymbol{k}') \rangle'$$

$$\propto \sigma_8^2$$

Bias parameter and σ_8 are degenerated in the tree-level galaxy power-spectrum

How to break this degeneracy?


$P(k)$
power spectrum
$B(k_1, k_2, k_3)$
bispectrum

Degeneracy breaking with bispectrum

$$B_g^{\rm tree}(k_1, k_2, k_3) \supset b_1^2 \left[b_2 B_{\delta^2}(k_1, k_2, k_3) + 2 b_{K^2} B_{K^2}(k_1, k_2, k_3) \right]$$

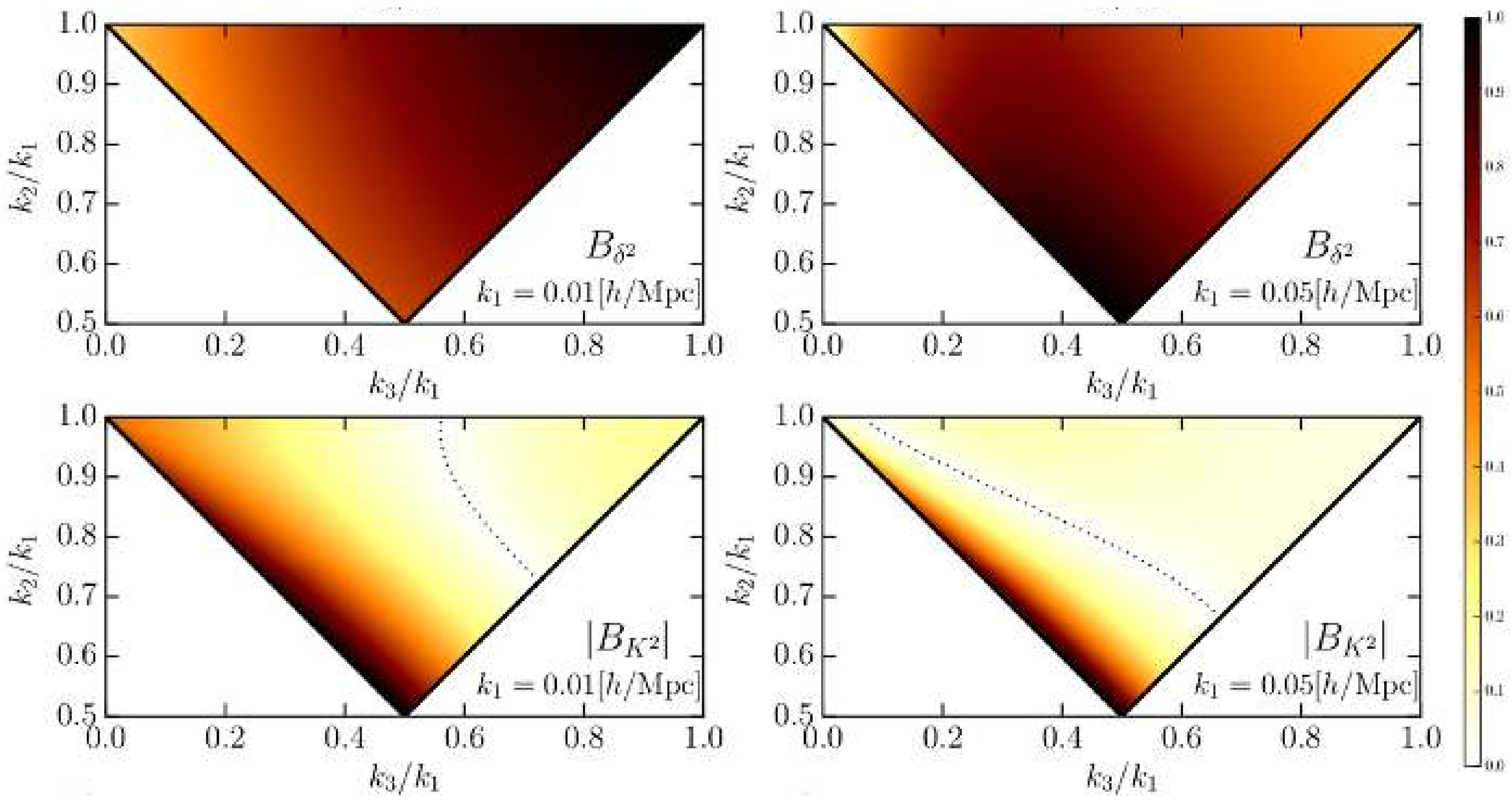
Degeneracy breaking with bispectrum

$$B_g^{\text{tree}}(k_1, k_2, k_3) \supset b_1^2 \left[b_2 B_{\delta^2}(k_1, k_2, k_3) + 2 b_{K^2} B_{K^2}(k_1, k_2, k_3) \right]$$

$$B_{\delta^2}(k_1, k_2, k_3) = P_L(k_1) P_L(k_2) + 2 \text{ perm.}$$

$$B_{K^2}(k_1, k_2, k_3) = \left([\hat{\boldsymbol{k}}_1 \cdot \hat{\boldsymbol{k}}_2]^2 - \frac{1}{3} \right) P_L(k_1) P_L(k_2) + 2 \text{ perm.}$$

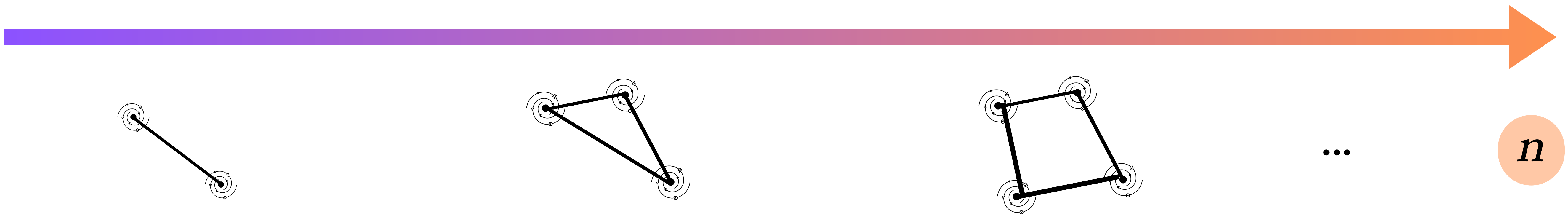
$$P_L(k) = \langle \delta^{(1)}(\boldsymbol{k}) \delta^{(1)}(\boldsymbol{k}') \rangle' \\ \propto \sigma_8^2$$



Adapted from Desjacques, Jeong & Schmidt (2016)

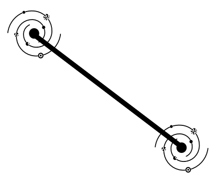
Inferring the cosmological parameters: **challenges**

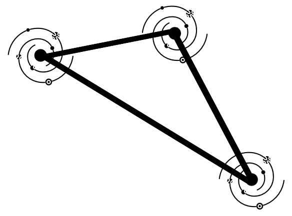
power-spectrum bispectrum trispectrum **Increasing complexity**

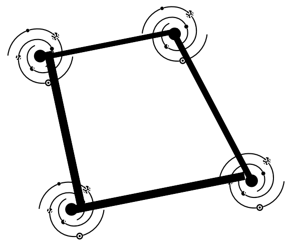


Inferring the cosmological parameters: **challenges**

power-spectrum | bispectrum | trispectrum | **Increasing complexity**







... n

$D = 18$

$D = 714$

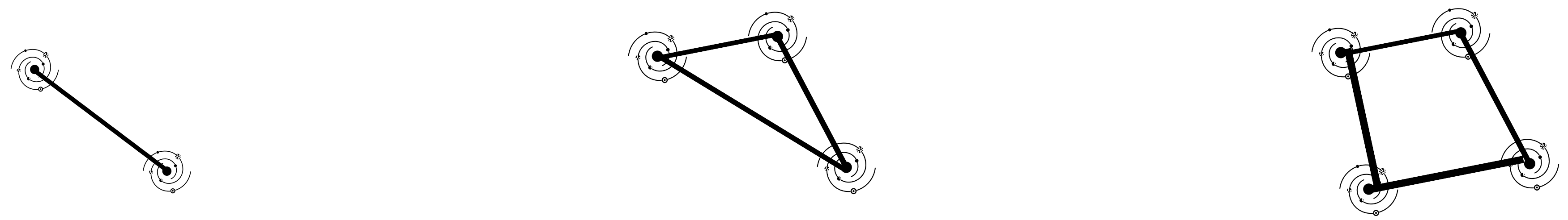
$k_{\max} = 0.12h\,\mathrm{Mpc}^{-1}$

$\Delta k = 2k_f$

$L = 2000h^{-1}\mathrm{Mpc}$

Inferring the cosmological parameters: **challenges**

power-spectrum | bispectrum | trispectrum | Increasing complexity



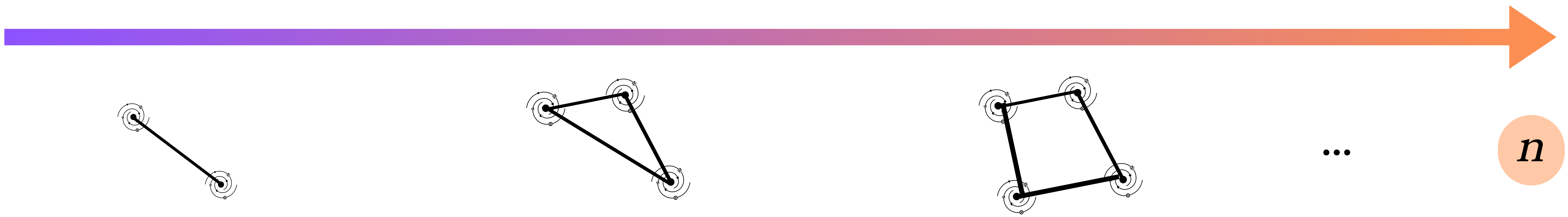
... n

$D = 18$ $\quad D = 714$ $\quad D = 15\,093$

$$k_{\max} = 0.12h\,\mathrm{Mpc}^{-1}$$
$$\Delta k = 2k_f$$
$$L = 2000h^{-1}\mathrm{Mpc}$$

Inferring the cosmological parameters: **challenges**

power-spectrum bispectrum trispectrum **Increasing complexity**



Analytical approximations Estimation Modelling

$$-2\ln\mathcal{L}(\mathbf{D}|\boldsymbol{\theta}) = \left(\mathbf{D}-\mathbf{T}(\boldsymbol{\theta})\right)\cdot\mathbf{C}^{-1}\cdot\left(\mathbf{D}-\mathbf{T}(\boldsymbol{\theta})\right)$$

Measurements

Part I

Simulation-based inference

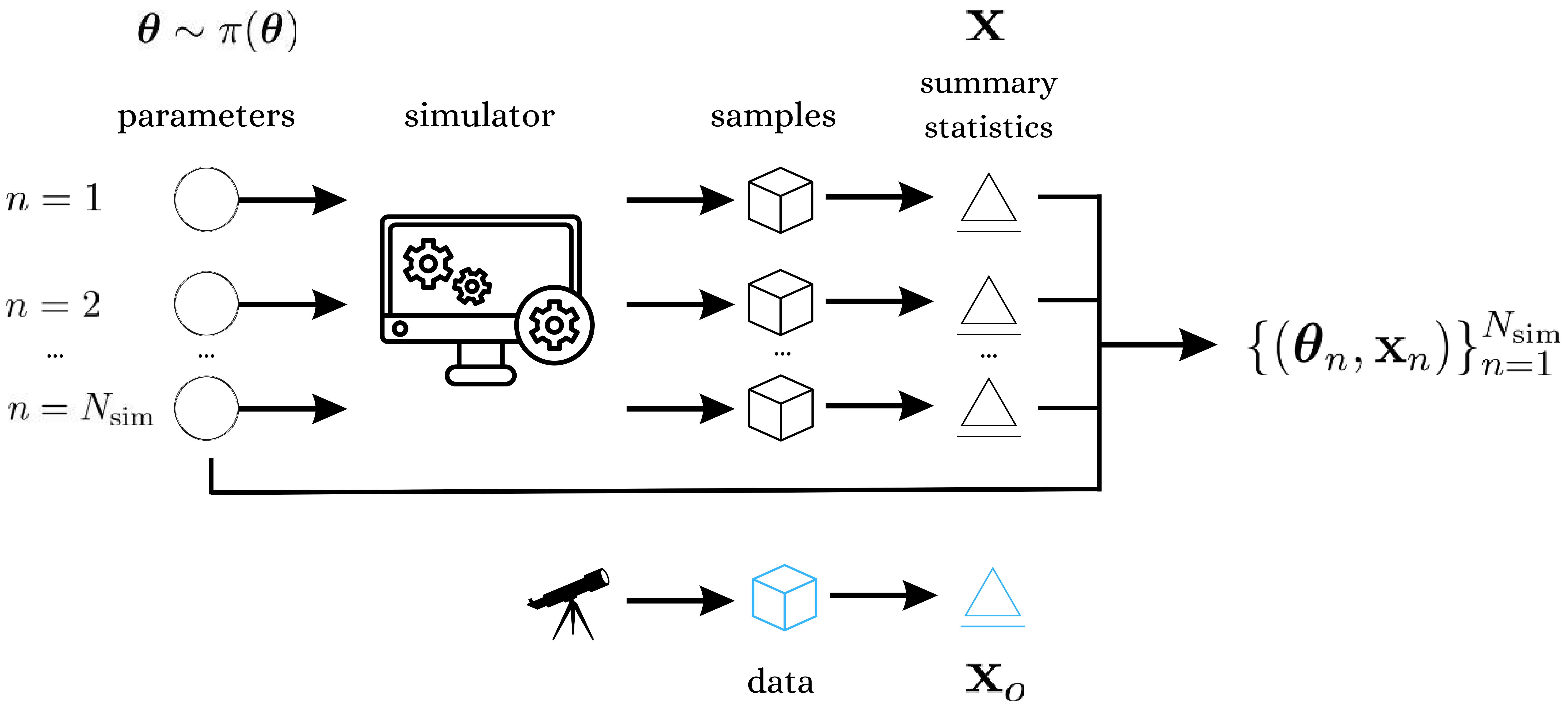
(SBI)

SBI: the main idea

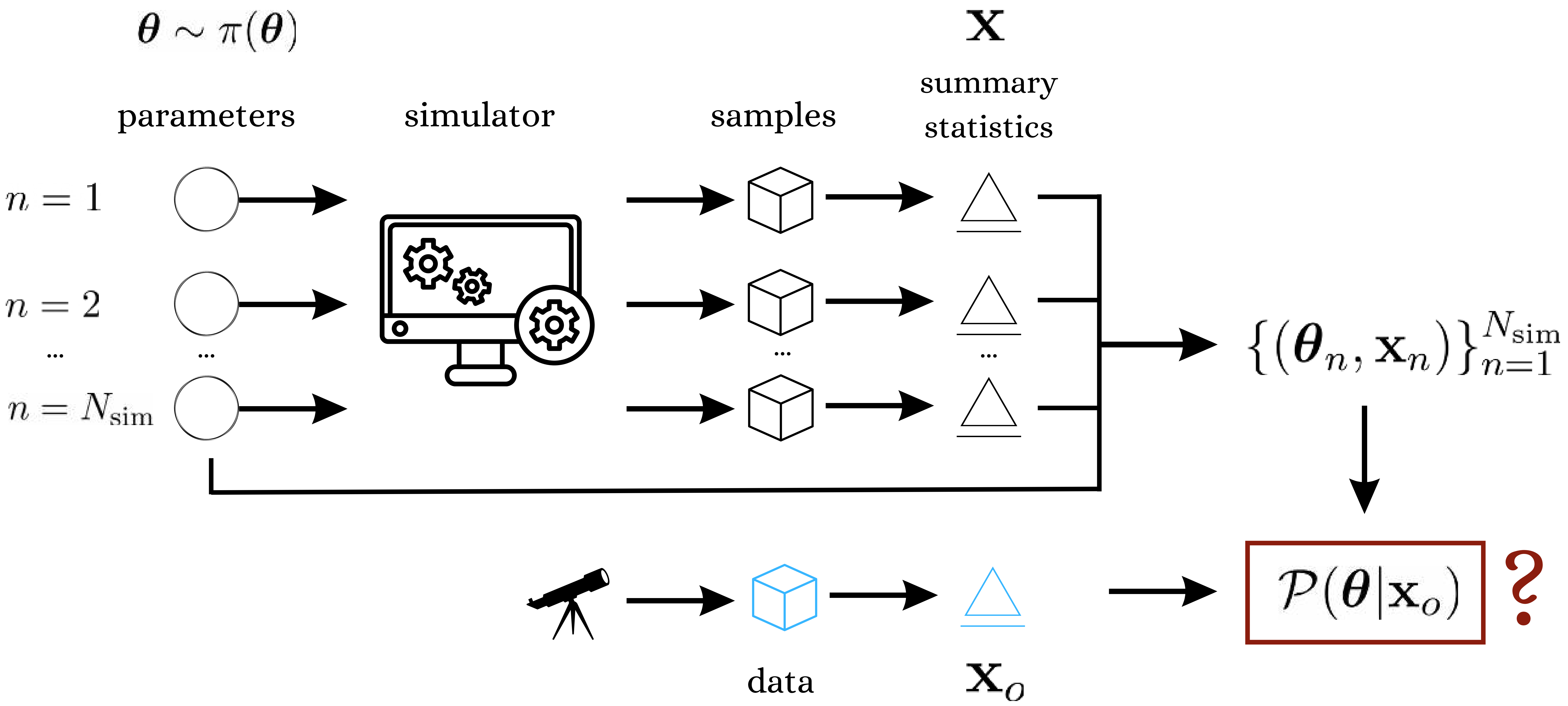
$$\mathcal{P}(\boldsymbol{\theta}|\mathrm{D}) \propto \cancel{\mathcal{L}(\mathrm{D}|\boldsymbol{\theta})}\pi(\boldsymbol{\theta})$$

$$\boxed{\mathbf{T}(\boldsymbol{\theta}) \sim \mathrm{simulator}(\boldsymbol{\theta})}$$

Simulation-based inference

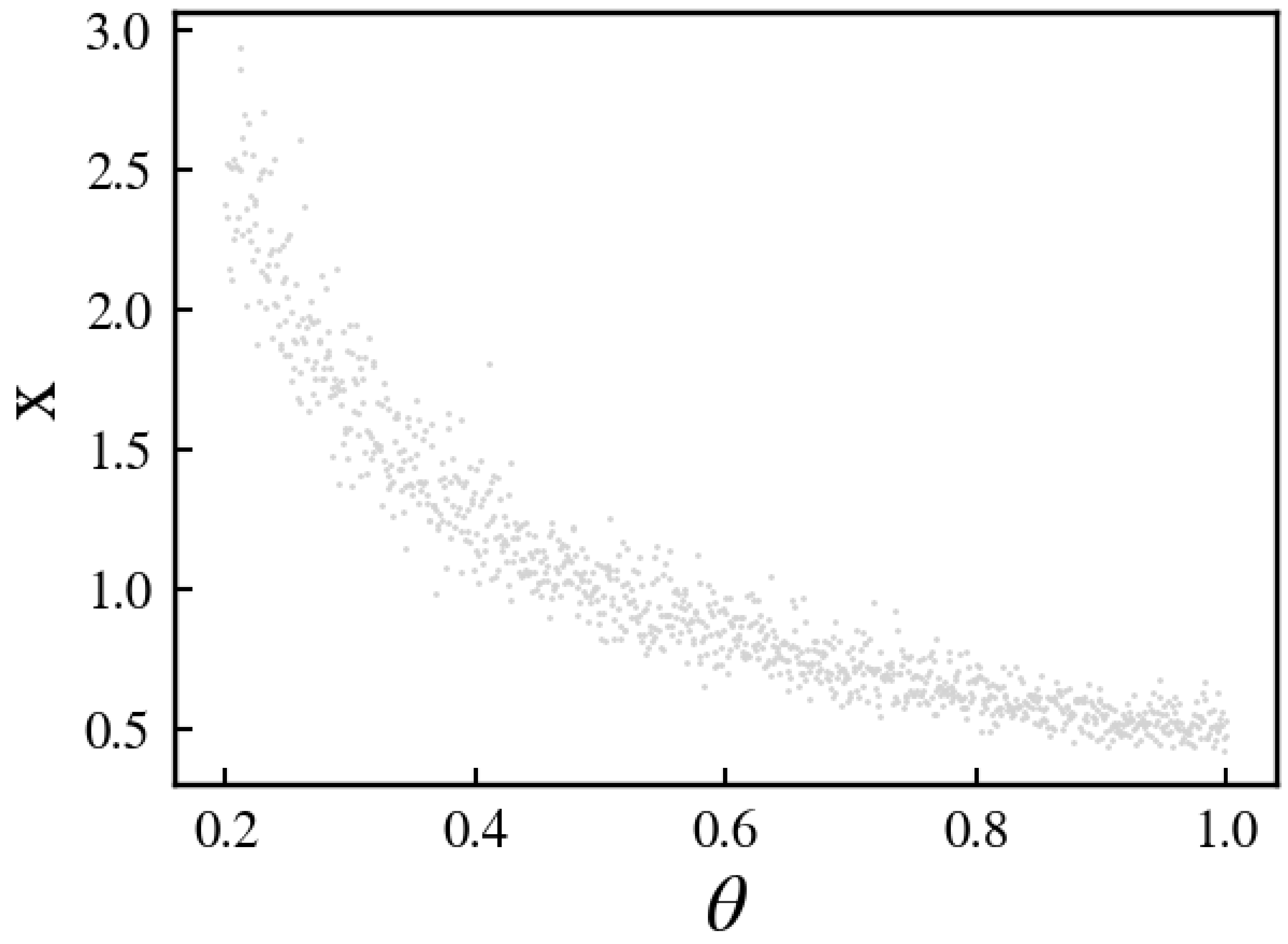


Simulation-based inference

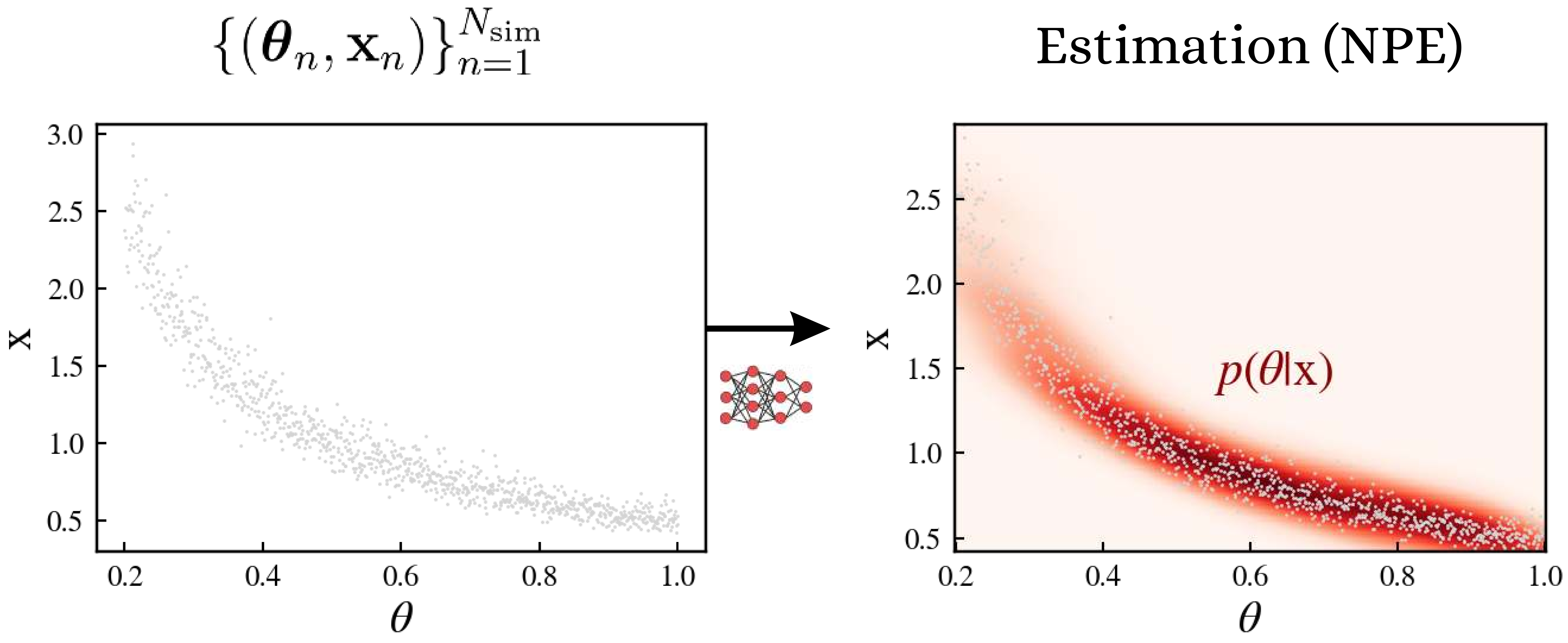


Simulation-based inference

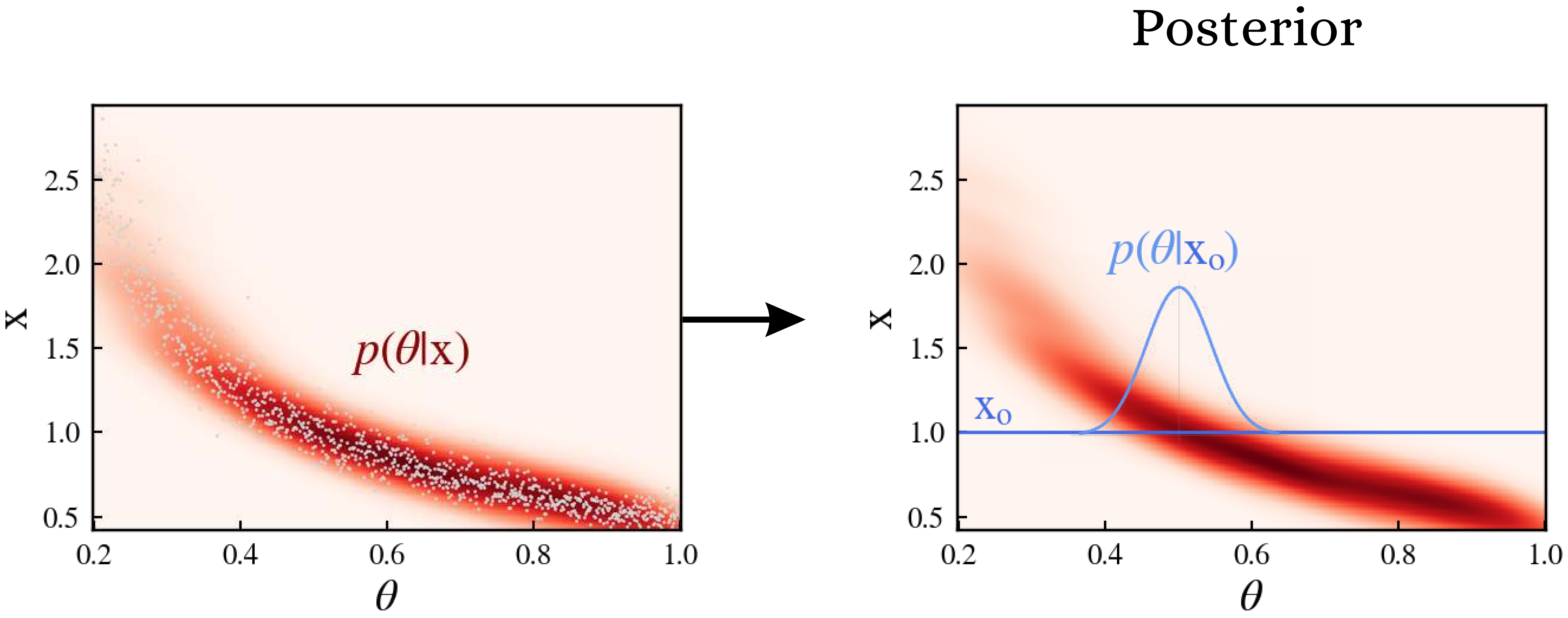
$$\{(\boldsymbol{\theta}_n, \mathbf{x}_n)\}_{n=1}^{N_{\text{sim}}}$$



Simulation-based inference



Simulation-based inference



Neural Density Estimation

$$\{(\boldsymbol{\theta}_n, \mathbf{x}_n)\}_{n=1}^{N_{\text{sim}}} \begin{cases} \textbf{Neural Posterior Estimation (NPE)} \\ q_\phi(\boldsymbol{\theta}|\mathbf{x}) \longrightarrow \hat{p}(\boldsymbol{\theta}|\mathbf{x}_o) = q_\phi(\boldsymbol{\theta}|\mathbf{x}_o) \\ \textbf{Neural Likelihood Estimation (NLE)} \\ q_\phi(\mathbf{x}|\boldsymbol{\theta}) \longrightarrow \hat{p}(\boldsymbol{\theta}|\mathbf{x}_o) \propto q_\phi(\mathbf{x}_o|\boldsymbol{\theta})p(\boldsymbol{\theta}) \end{cases}$$

Neural Density Estimation

$$\{(\boldsymbol{\theta}_n, \mathbf{x}_n)\}_{n=1}^{N_{\mathrm{sim}}}$$

Neural Posterior Estimation (NPE)

$$q_\phi(\boldsymbol{\theta}|\mathbf{x}) \longrightarrow \hat{p}(\boldsymbol{\theta}|\mathbf{x}_o) = q_\phi(\boldsymbol{\theta}|\mathbf{x}_o)$$

Neural Likelihood Estimation (NLE)

$$q_\phi(\mathbf{x}|\boldsymbol{\theta}) \longrightarrow \hat{p}(\boldsymbol{\theta}|\mathbf{x}_o) \propto q_\phi(\mathbf{x}_o|\boldsymbol{\theta})p(\boldsymbol{\theta})$$

neural network
trainable parameters

Neural Density Estimation

How to train the model? (For example, NLE)

$$\mathcal{L} = \mathbb{E}_{p(\boldsymbol{\theta})}\left[D_{\mathrm{KL}}\left[p(\mathbf{x}|\boldsymbol{\theta})\,||\,q_{\phi}(\mathbf{x}|\boldsymbol{\theta})\right]\right]$$

loss function

target density

?

neural network
trainable parameters

Neural Density Estimation

How to train the model? (For example, NLE)

$$\mathbb{E}_{p(\boldsymbol{\theta})}\left[D_{\mathrm{KL}}\left[\underbrace{p(\mathbf{x}|\boldsymbol{\theta})}_{\text{target density}} \,||\, q_{\phi}(\mathbf{x}|\boldsymbol{\theta})\right]\right] = \int d\boldsymbol{\theta}\, p(\boldsymbol{\theta}) \int d\mathbf{x}\, p(\mathbf{x}|\boldsymbol{\theta})\, \log\left(\frac{p(\mathbf{x}|\boldsymbol{\theta})}{q_{\phi}(\mathbf{x}|\boldsymbol{\theta})}\right)$$

target density

?

neural network
trainable parameters

Neural Density Estimation

How to train the model? (For example, NLE)

$$\mathbb{E}_{p(\boldsymbol{\theta})}\left[D_{\mathrm{KL}}\left[p(\mathbf{x}|\boldsymbol{\theta})\,||\,q_{\phi}(\mathbf{x}|\boldsymbol{\theta})\right]\right] = \int d\boldsymbol{\theta}\, p(\boldsymbol{\theta}) \int d\mathbf{x}\, p(\mathbf{x}|\boldsymbol{\theta}) \log\left(\frac{p(\mathbf{x}|\boldsymbol{\theta})}{q_{\phi}(\mathbf{x}|\boldsymbol{\theta})}\right)$$

target density

?

neural network
trainable parameters

$$= \int d\boldsymbol{\theta}\, d\mathbf{x}\, p(\boldsymbol{\theta}, \mathbf{x}) \log\left(\frac{p(\mathbf{x}|\boldsymbol{\theta})}{q_{\phi}(\mathbf{x}|\boldsymbol{\theta})}\right)$$

$$\boxed{p(\boldsymbol{\theta}, \mathbf{x}) = p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})}$$

Neural Density Estimation

How to train the model? (For example, NLE)

$$\mathbb{E}_{p(\boldsymbol{\theta})}\left[D_{\mathrm{KL}}\left[\underbrace{p(\mathbf{x}|\boldsymbol{\theta})}_{\text{target density}}\,||\,q_{\phi}(\mathbf{x}|\boldsymbol{\theta})\right]\right] = \int d\boldsymbol{\theta}\, p(\boldsymbol{\theta}) \int d\mathbf{x}\, p(\mathbf{x}|\boldsymbol{\theta}) \log\left(\frac{p(\mathbf{x}|\boldsymbol{\theta})}{q_{\phi}(\mathbf{x}|\boldsymbol{\theta})}\right)$$

$$= \int d\boldsymbol{\theta}\, d\mathbf{x}\, p(\boldsymbol{\theta}, \mathbf{x}) \log\left(\frac{p(\mathbf{x}|\boldsymbol{\theta})}{q_{\phi}(\mathbf{x}|\boldsymbol{\theta})}\right)$$

$$= -\mathbb{E}_{p(\boldsymbol{\theta},\mathbf{x})}\left[\log q_{\phi}(\mathbf{x}|\boldsymbol{\theta})\right] + \text{const.}$$

target density

?

neural network trainable parameters

the loss function we wish to minimize is independent of the target density form!

Neural Density Estimation

How to train the model? (For example, NLE)

$$\mathbb{E}_{p(\boldsymbol{\theta})}\left[D_{\mathrm{KL}}\left[p(\mathbf{x}|\boldsymbol{\theta})\,||\,q_{\phi}(\mathbf{x}|\boldsymbol{\theta})\right]\right] = \int d\boldsymbol{\theta}\,p(\boldsymbol{\theta}) \int d\mathbf{x}\,p(\mathbf{x}|\boldsymbol{\theta})\,\log\left(\frac{p(\mathbf{x}|\boldsymbol{\theta})}{q_{\phi}(\mathbf{x}|\boldsymbol{\theta})}\right)$$

target density — $p(\mathbf{x}|\boldsymbol{\theta})$

neural network trainable parameters — ϕ

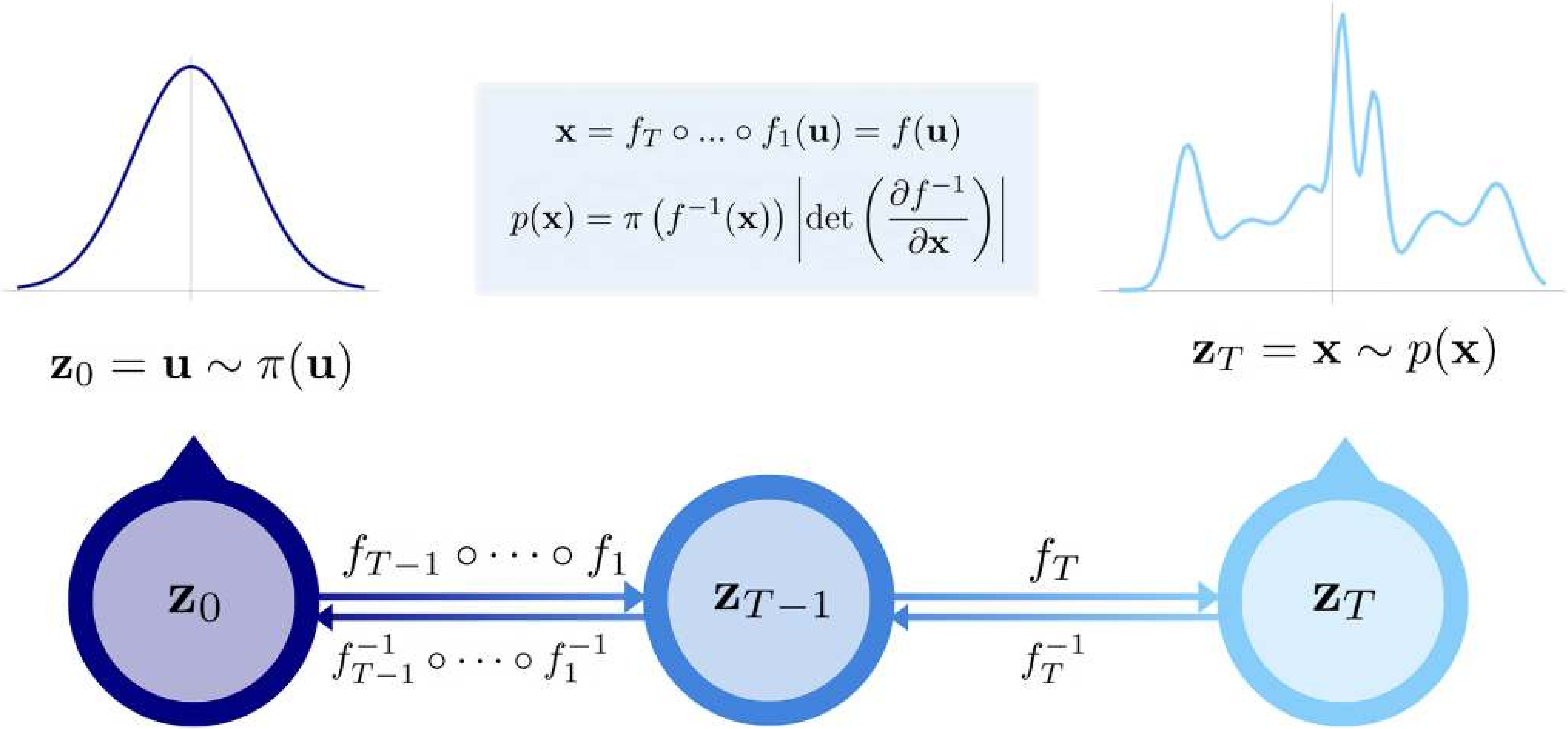
$$= \int d\boldsymbol{\theta}\,d\mathbf{x}\,p(\boldsymbol{\theta},\mathbf{x})\,\log\left(\frac{p(\mathbf{x}|\boldsymbol{\theta})}{q_{\phi}(\mathbf{x}|\boldsymbol{\theta})}\right)$$

$$= -\mathbb{E}_{p(\boldsymbol{\theta},\mathbf{x})}\left[\log q_{\phi}(\mathbf{x}|\boldsymbol{\theta})\right] + \text{const.}$$

$$\approx -\frac{1}{N_{\text{sim}}}\sum_{n=1}^{N_{\text{sim}}} \log q_{\phi}(\mathbf{x}_n|\boldsymbol{\theta}_n) + \text{const.}\,,$$

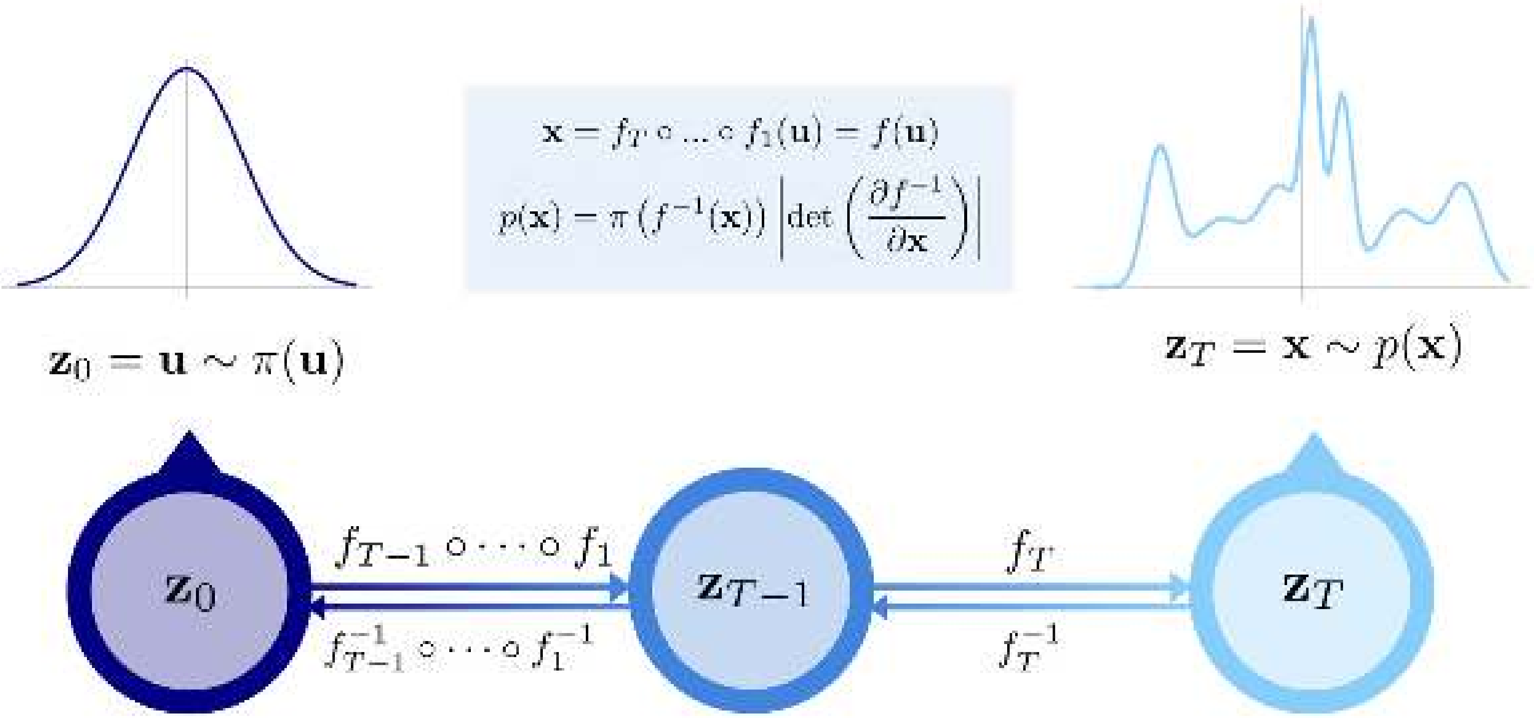
$$\boxed{\{(\boldsymbol{\theta}_n,\mathbf{x}_n)\}_{n=1}^{N_{\text{sim}}}}$$

Normalizing Flows



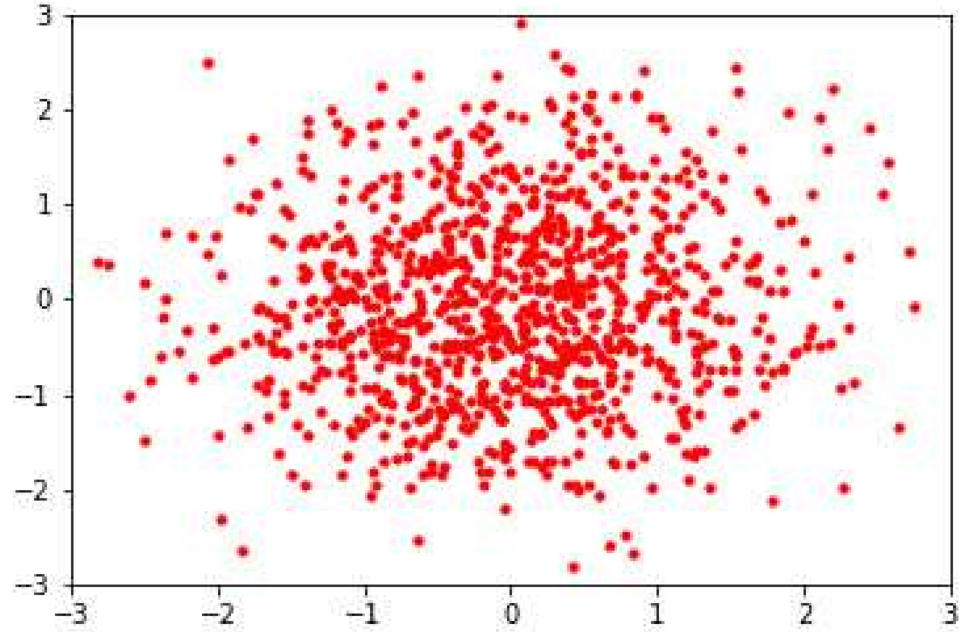
Tucci, Schmidt (2023)

Normalizing Flows



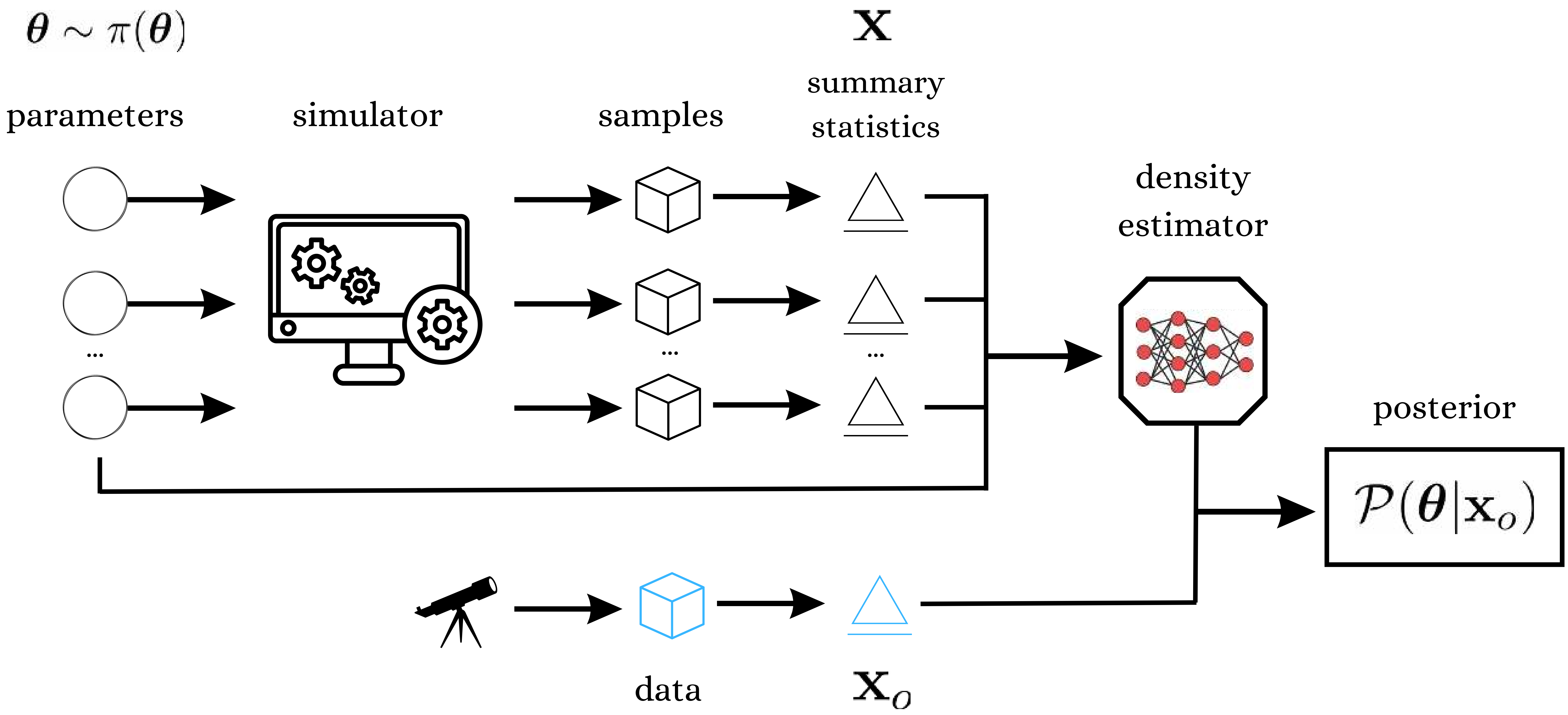
Tucci, Schmidt (2023)

$$\boxed{q_\phi(\mathbf{x}|\boldsymbol{\theta}) = \mathcal{N}(\mathbf{z}_0|\mathbf{0}, \mathbf{I}) \prod_{t=1}^{T} \left|\det\left(\frac{\partial f_t}{\partial \mathbf{z}_{t-1}}\right)\right|^{-1}}$$

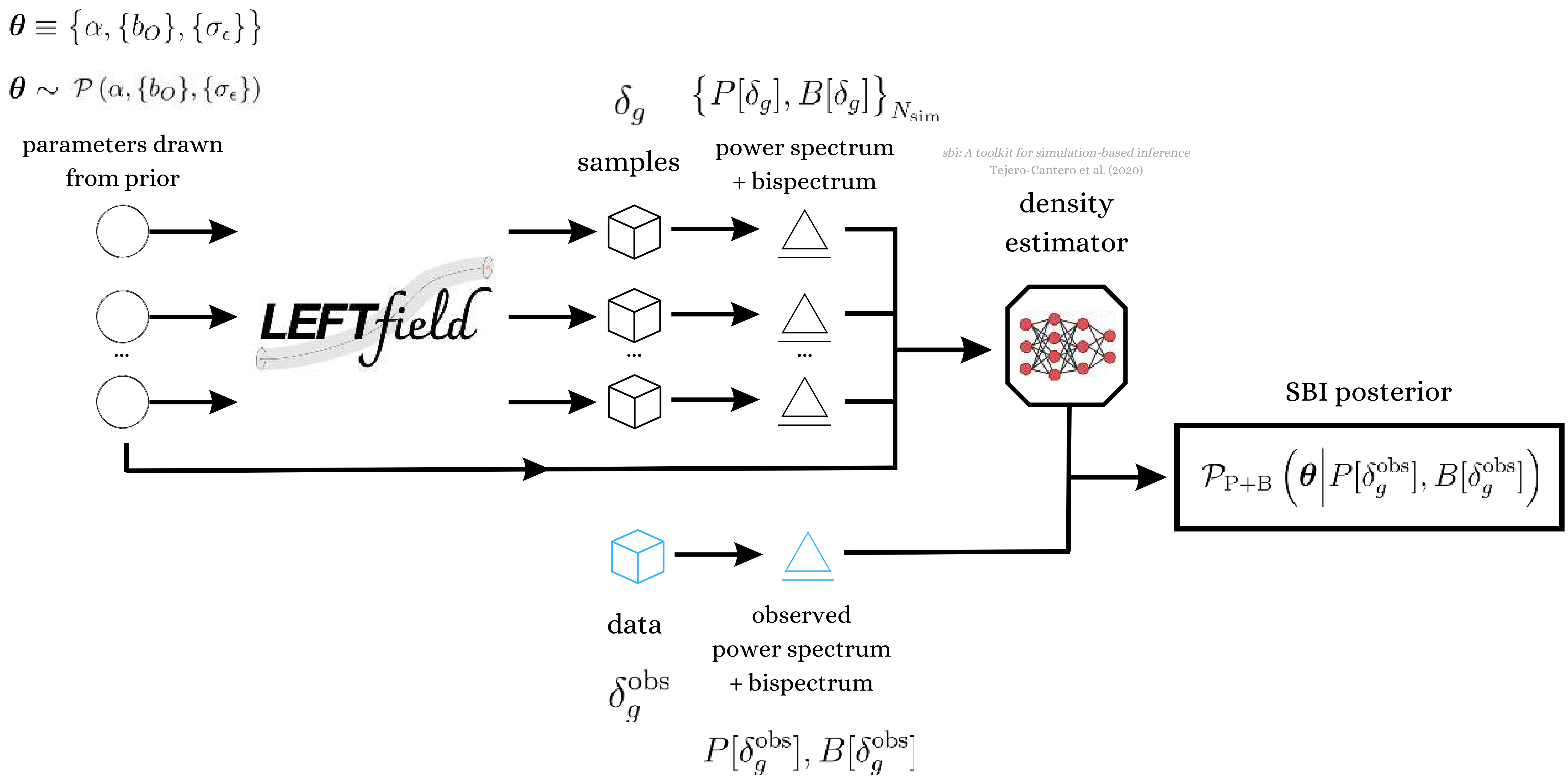


Credits: Miles Cranmer

Simulation-based inference

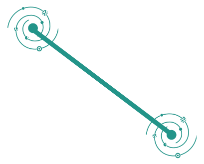


Simulation-based inference for galaxy clustering

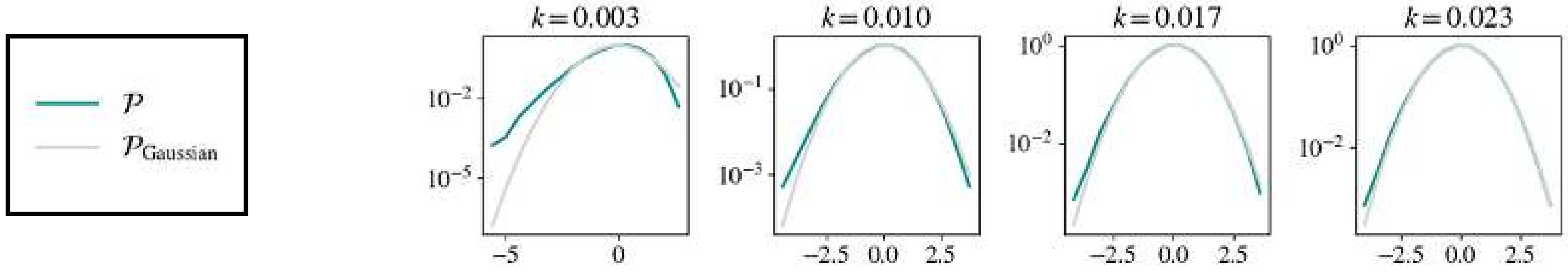


On the Gaussianity assumption of the n-point functions

power-spectrum

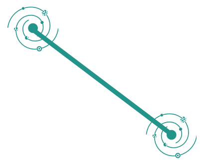


large scales

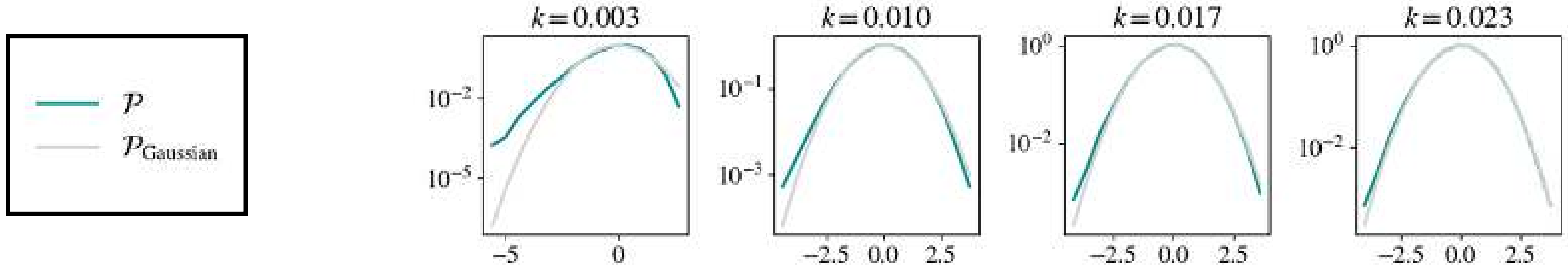


On the Gaussianity assumption of the n-point functions

power-spectrum



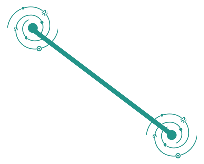
large scales

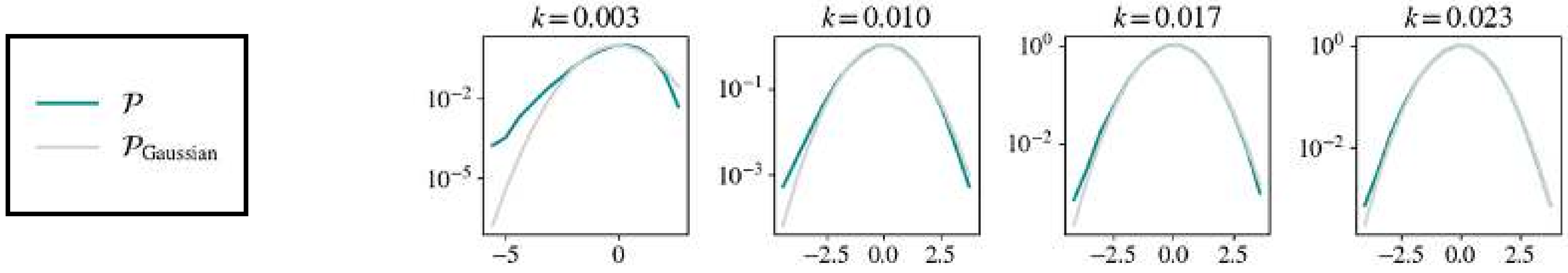


breaking of central limit theorem on large scales induces deviation from the Gaussian likelihood assumption!

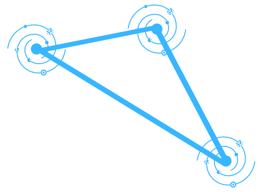
On the Gaussianity assumption of the n-point functions

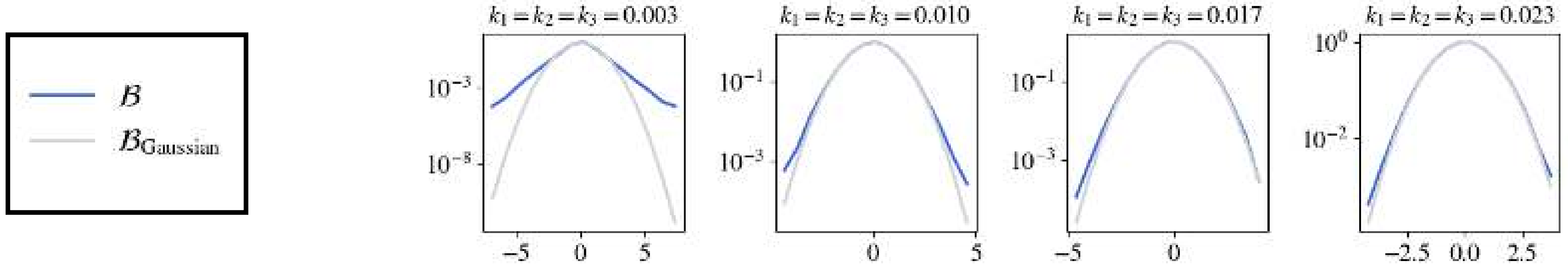
power-spectrum



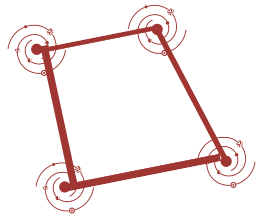


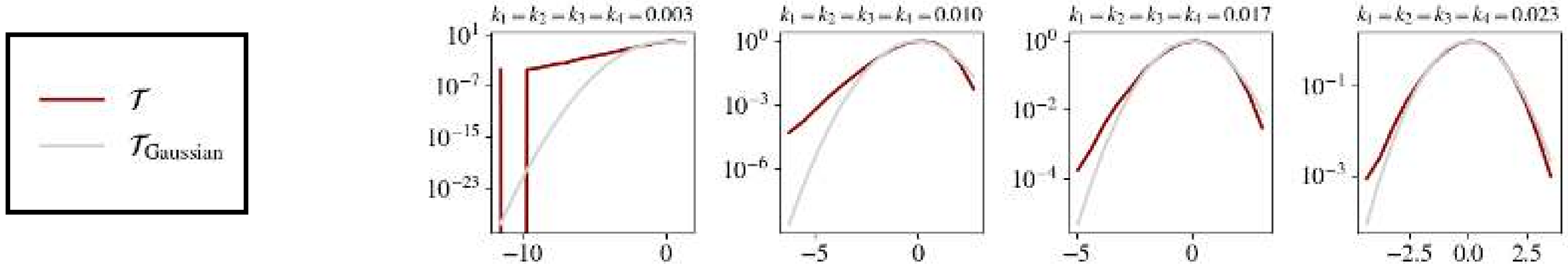
bispectrum





trispectrum



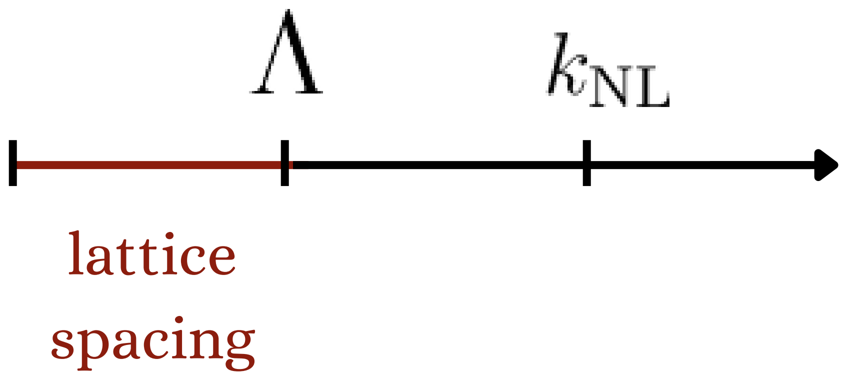


The forward model

based on the EFTofLSS & the bias expansion

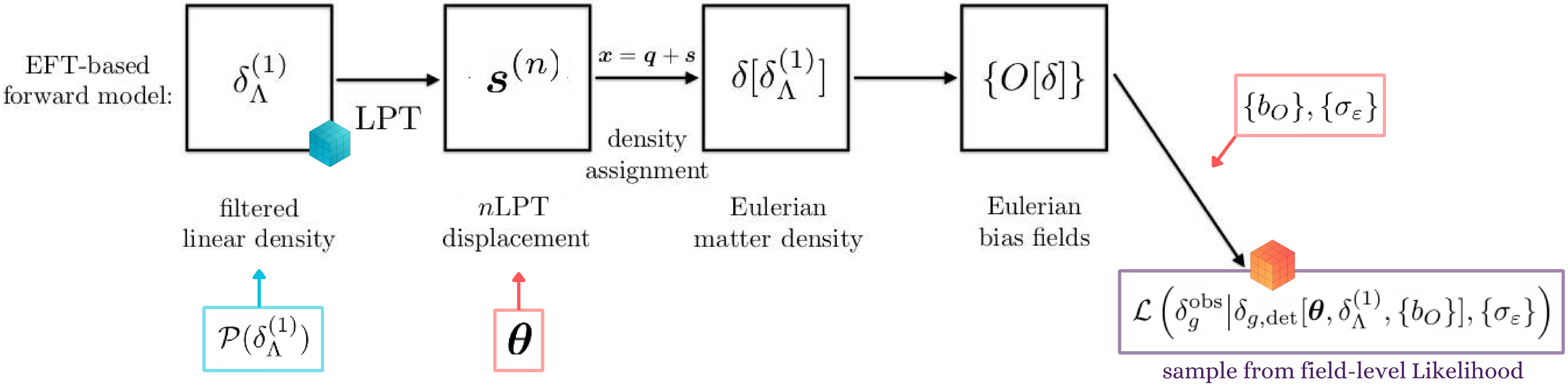
LEFTfield

- A **fast** forward model based on the EFTofLSS that solves the gravitational evolution of all modes in a lattice up to the cutoff scale
- nLPT and incorporates bias and stochastic parameters, marginalizing over reasonable models of galaxy formation
- Easier to deal with redshift space, masks and systematic effects



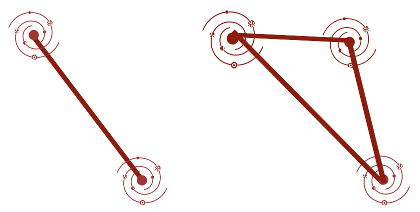
The forward model

LEFTfield



An n-th order Lagrangian Forward Model for Large-Scale Structure

Schmidt (2021)



Testing SBI on Euclid-like mock data

Breaking degeneracy between σ_8 and bias parameters with the galaxy power-spectrum and bispectrum

Tucci & Schmidt (2024)

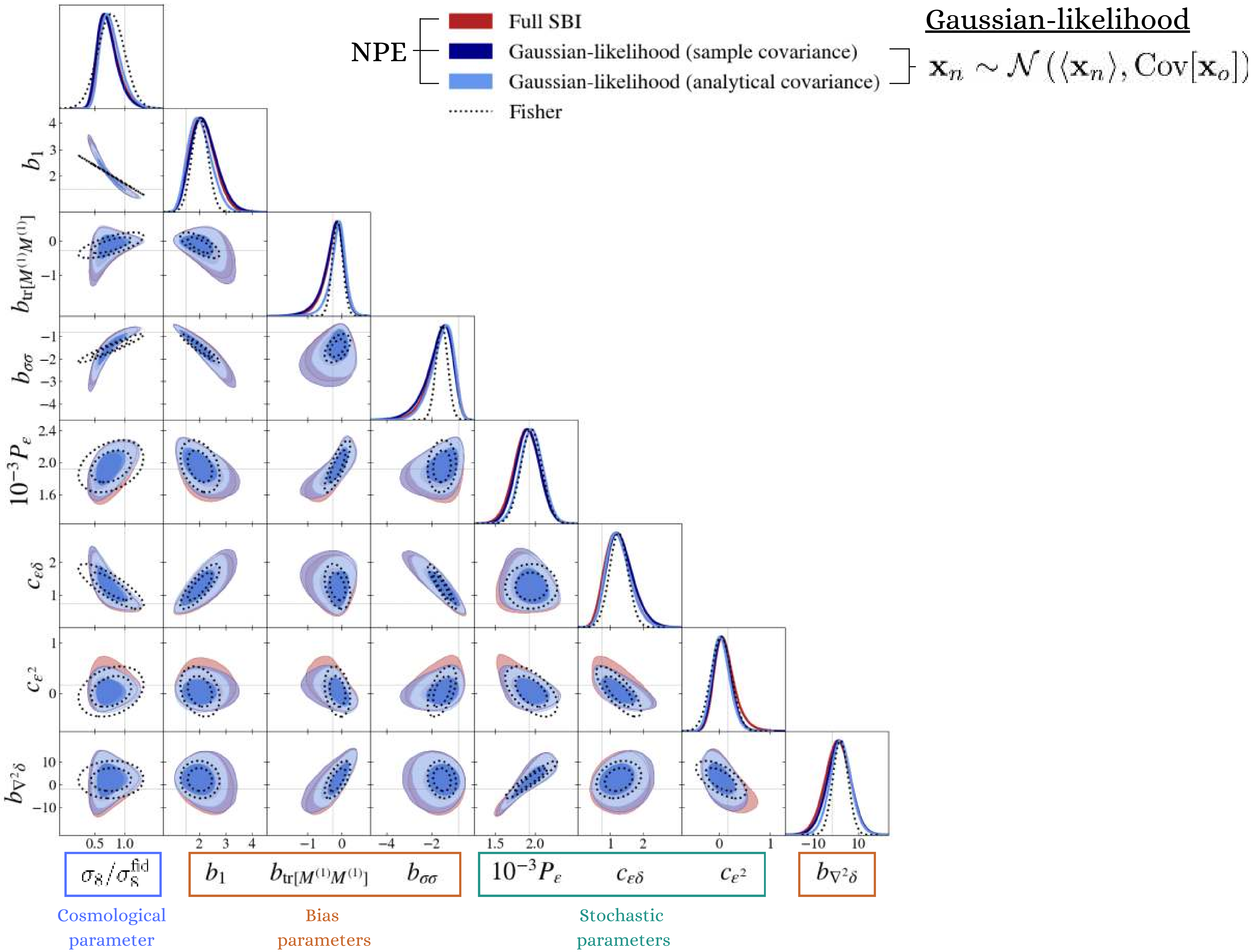
JCAP

Cosmological constraints

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

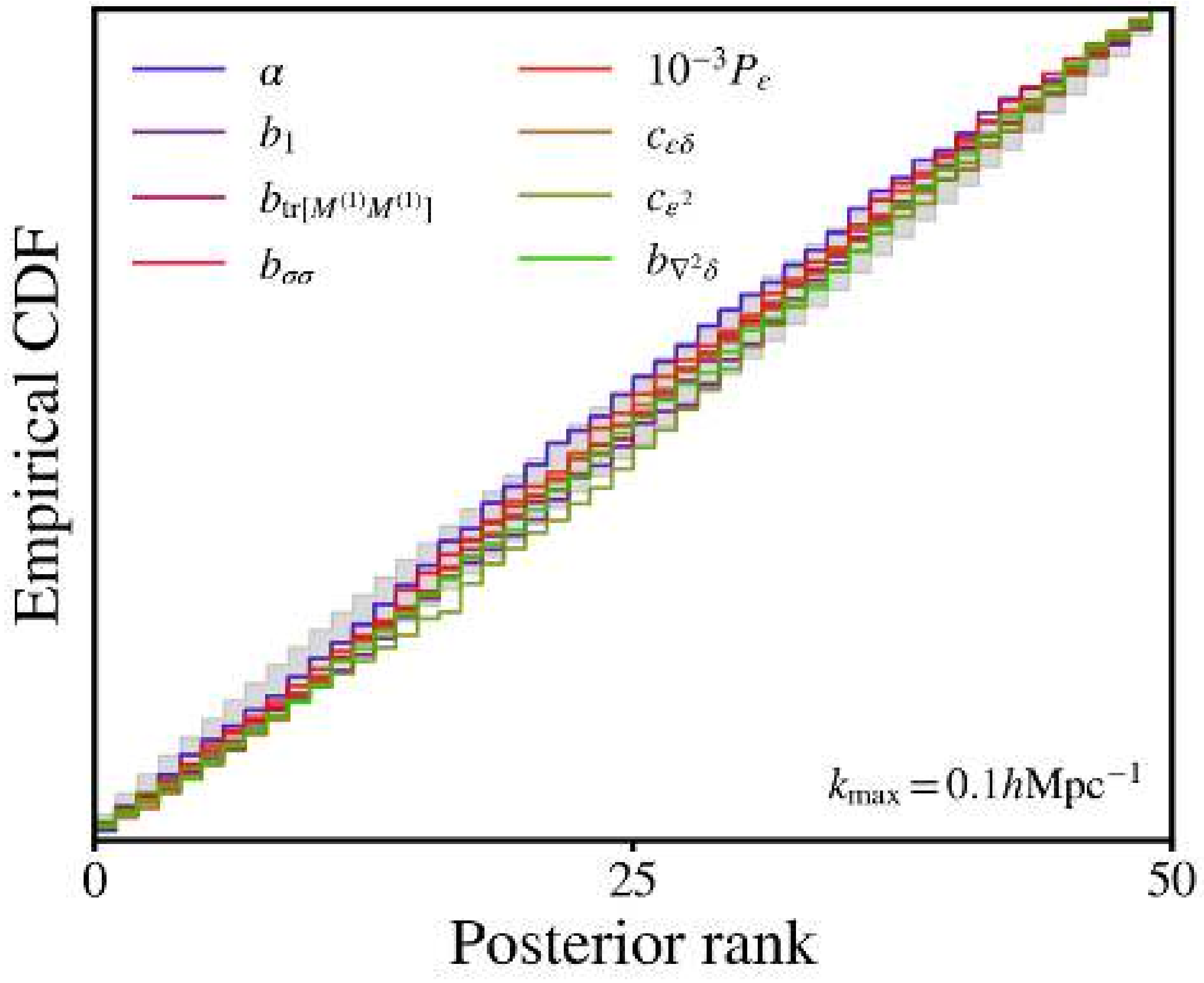
$k_{\rm max} = \Lambda = 0.1 h\mathrm{Mpc}^{-1}$

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

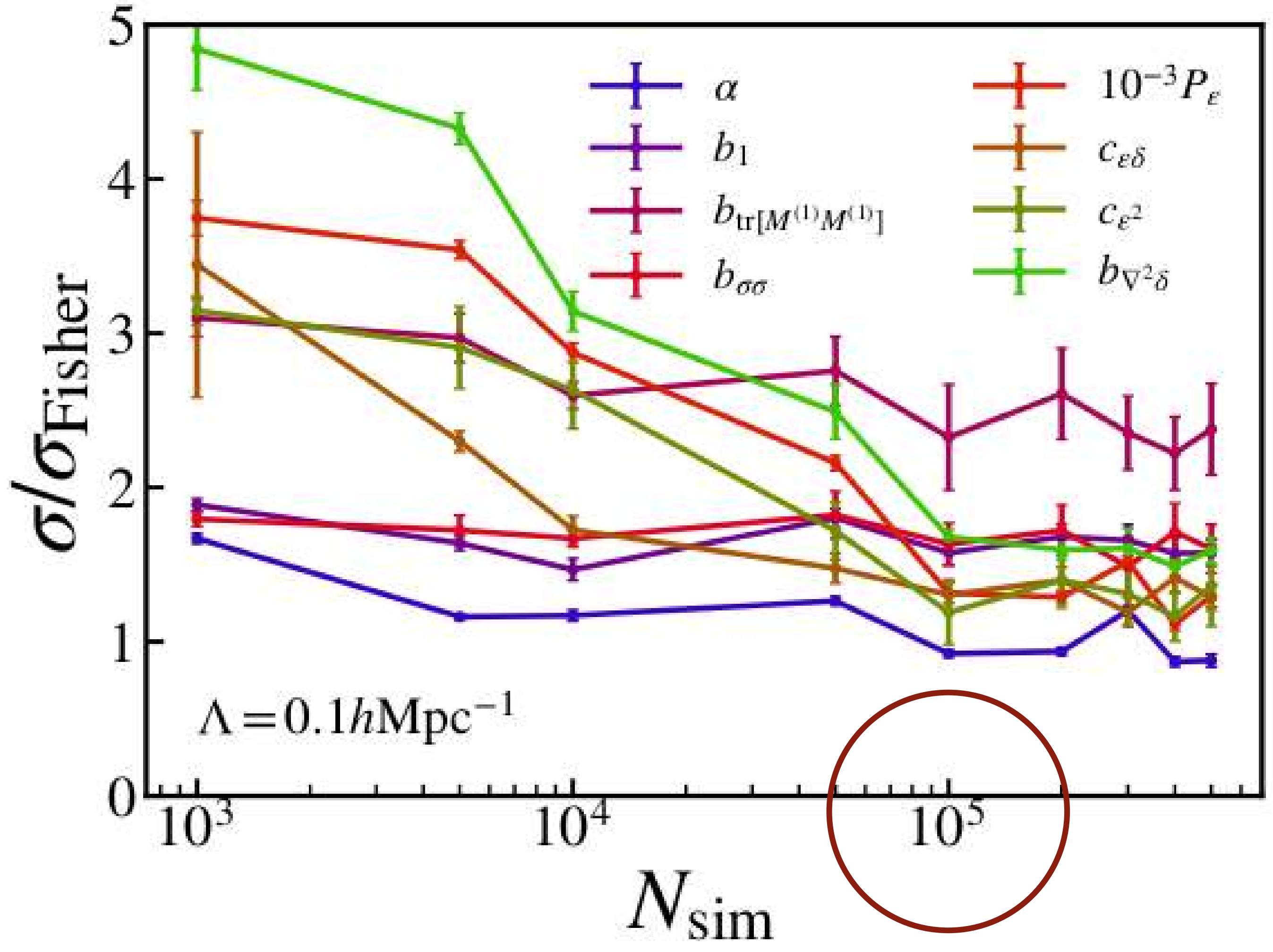


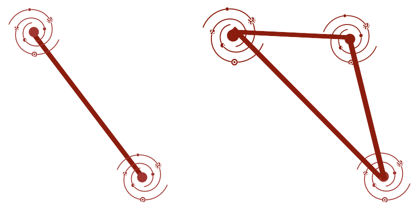
Tests of inference

Simulation-based calibration



Convergence





SBI on dark-matter halos

Breaking degeneracy between σ_8 and bias parameters with the galaxy power-spectrum and bispectrum

Nguyen, Schmidt, **Tucci** et al. (2024)

PRL (accepted)

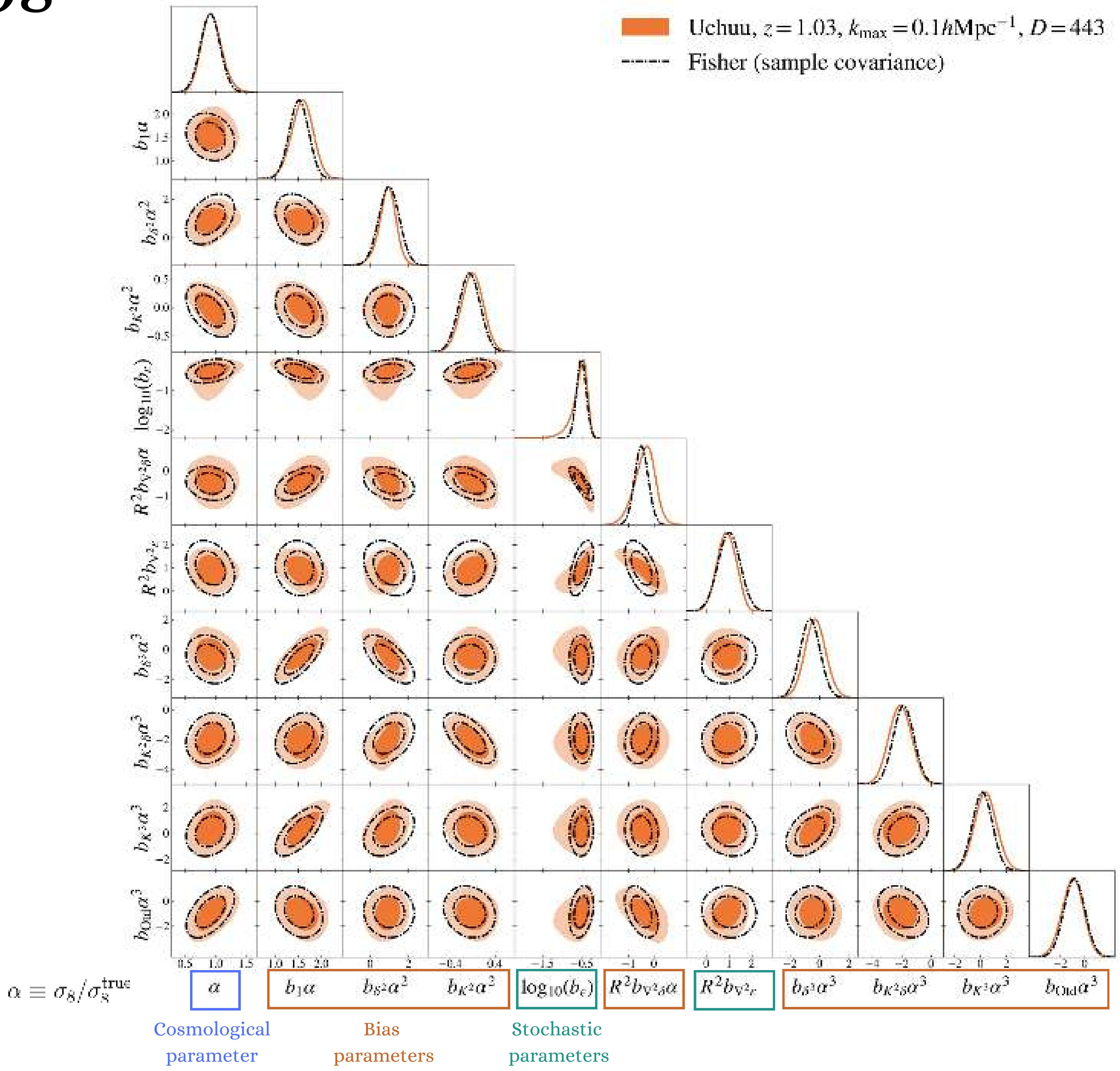
Inference setup: halo samples

	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}

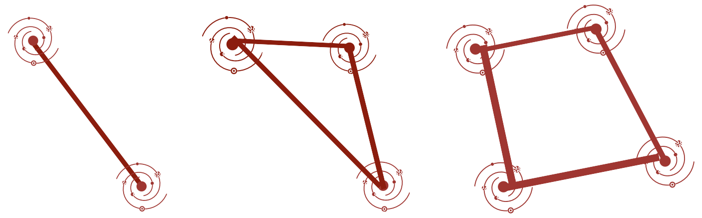
Two scale cuts:

$k_{\mathrm{max}} = 0.1 h\mathrm{Mpc}^{-1}$ & $k_{\mathrm{max}} = 0.12 h\mathrm{Mpc}^{-1}$

SBI on halos



Uchuu, $z = 1.03$, $k_{\max} = 0.1 h\mathrm{Mpc}^{-1}$, $D = 443$
Fisher (sample covariance)
$\alpha \equiv \sigma_8/\sigma_8^{\mathrm{true}}$
α
$b_1\alpha$
$b_{\delta^2}\alpha^2$
$b_{\mathcal{K}^2}\alpha^2$
$\log_{10}(b_\epsilon)$
$R^2 b_{\nabla^2\delta}\alpha$
$R^2 b_{\nabla^2\epsilon}$
$b_{\delta^3}\alpha^3$
$b_{\mathcal{K}^2\delta}\alpha^3$
$b_{\mathcal{K}^3}\alpha^3$
$b_{\mathrm{Odd}}\alpha^3$
Cosmological parameter
Bias parameters
Stochastic parameters

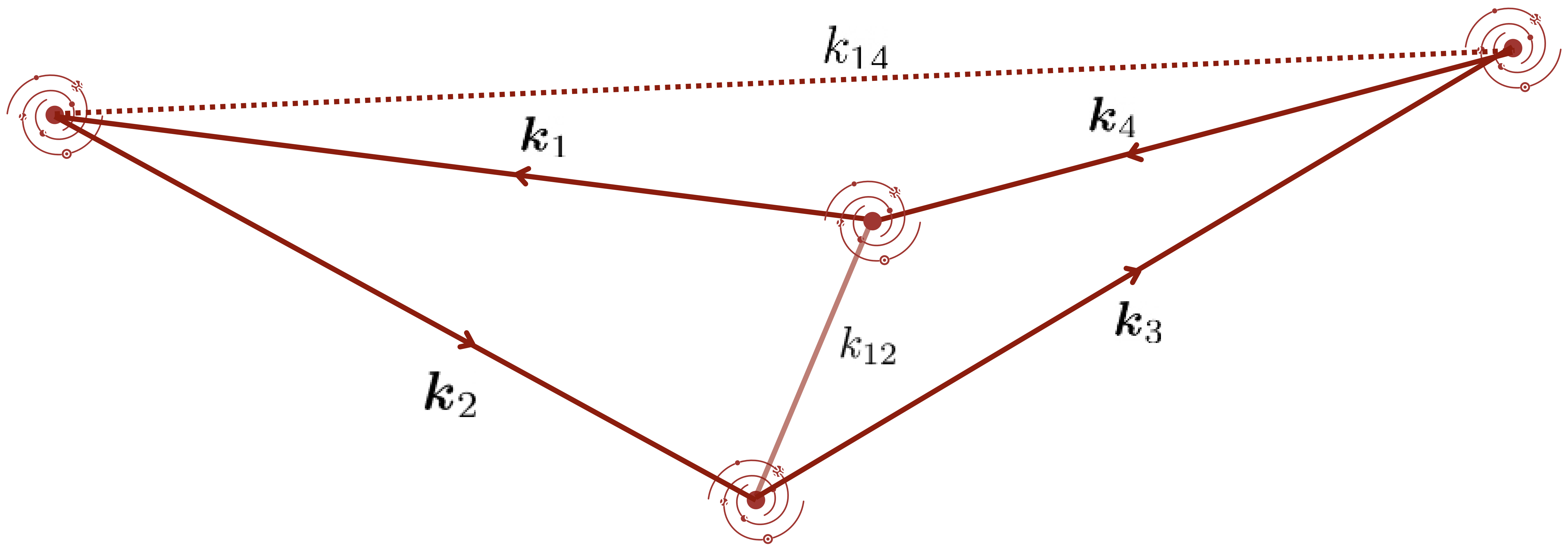


What if we add the galaxy **trispectrum**?

Breaking degeneracy between σ_8 and bias parameters with power-spectrum, bispectrum and trispectrum on dark-matter halos

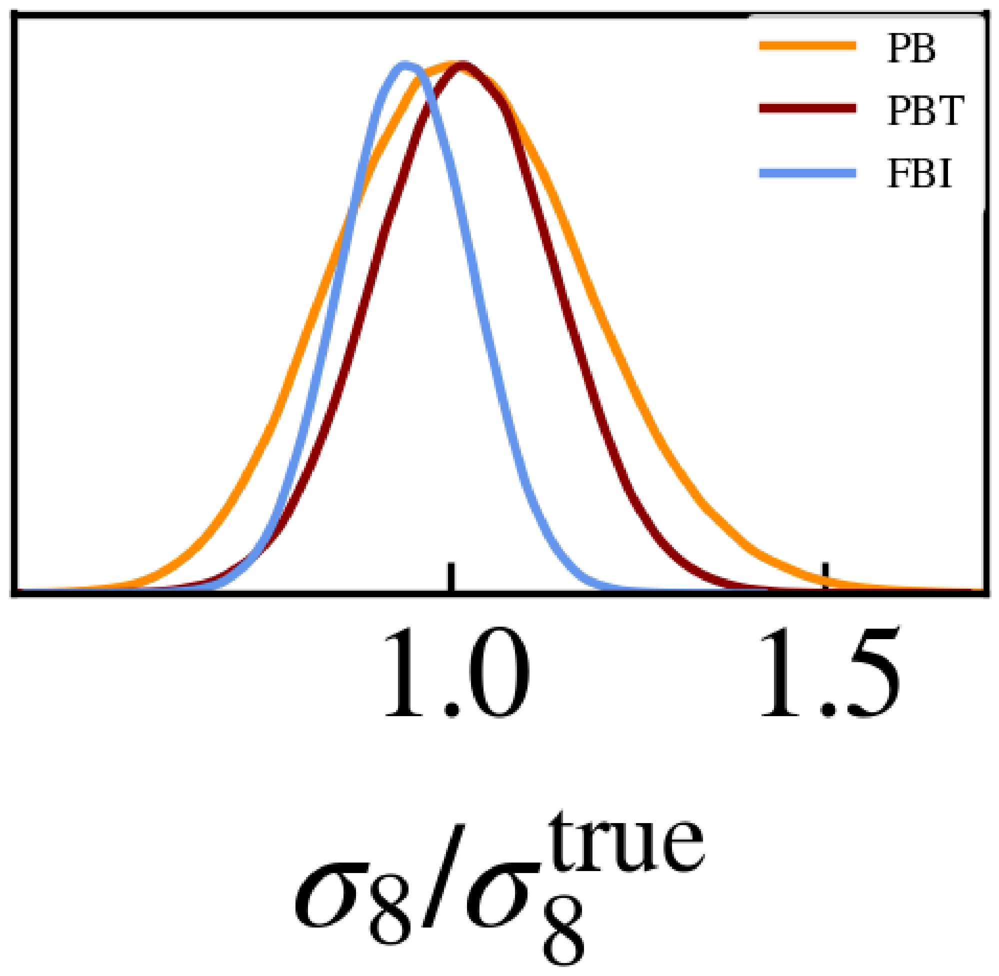
Tucci & Schmidt (in prep.)

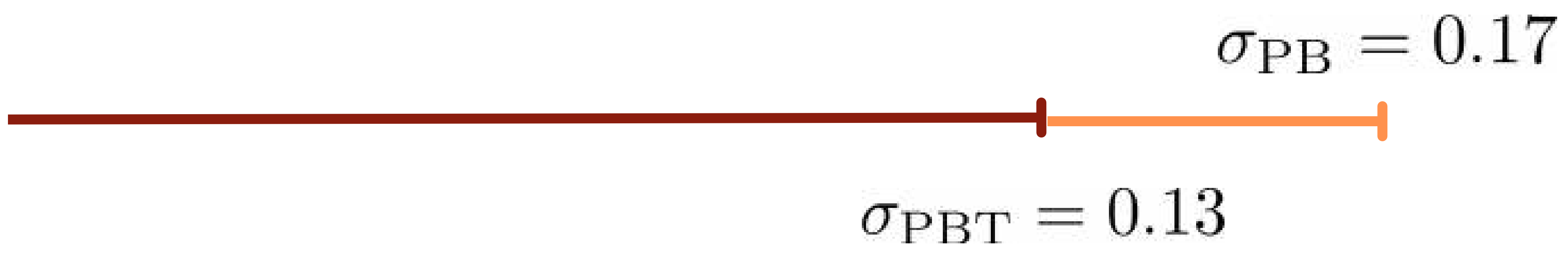
Trispectrum: the estimator



Jung+23, Coulton+23, Goldstein+24

Trispectrum: **preliminary** results





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

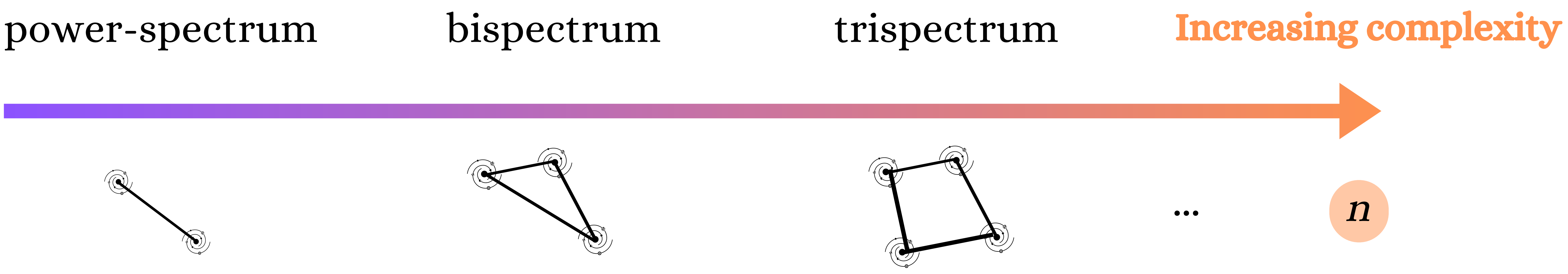
Uchuu halos at z=1

Brute force approach:
10^6 simulations

SBI with LEFTfield: Conclusions

- Robust analysis with EFTofLSS and bias expansion
- LEFTfield allows for **fast** analysis in **cosmological volumes** with **convergence** and posterior **diagnostics** tests
- Need order of 10^5 simulations for convergence (investigating how we can improve that)
- SBI allows for cosmological inference using **trispectrum**, which is **unfeasible** with standard inference techniques
- No need to assume Gaussian likelihood, explicit loop or covariance calculations

Inferring the cosmological parameters: **challenges**



Analytical approximations

Estimation

Modelling

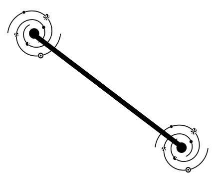
$$-2\ln \mathcal{L}(\mathbf{D}|\boldsymbol{\theta}) = (\mathbf{D}-\mathbf{T}(\boldsymbol{\theta}))\cdot \mathbf{C}^{-1}\cdot(\mathbf{D}-\mathbf{T}(\boldsymbol{\theta}))$$

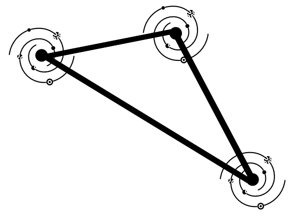
Measurements

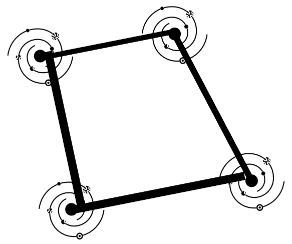
SBI

Inferring the cosmological parameters: **challenges**

power-spectrum | bispectrum | trispectrum | **Increasing complexity**







... n

where to stop?
is there a better way?

Analytical approximations | Estimation | Modelling

$$-2\ln \mathcal{L}(\mathbf{D}|\boldsymbol{\theta}) = (\mathbf{D}-\mathbf{T}(\boldsymbol{\theta}))\cdot \mathbf{C}^{-1}\cdot(\mathbf{D}-\mathbf{T}(\boldsymbol{\theta}))$$

Measurements

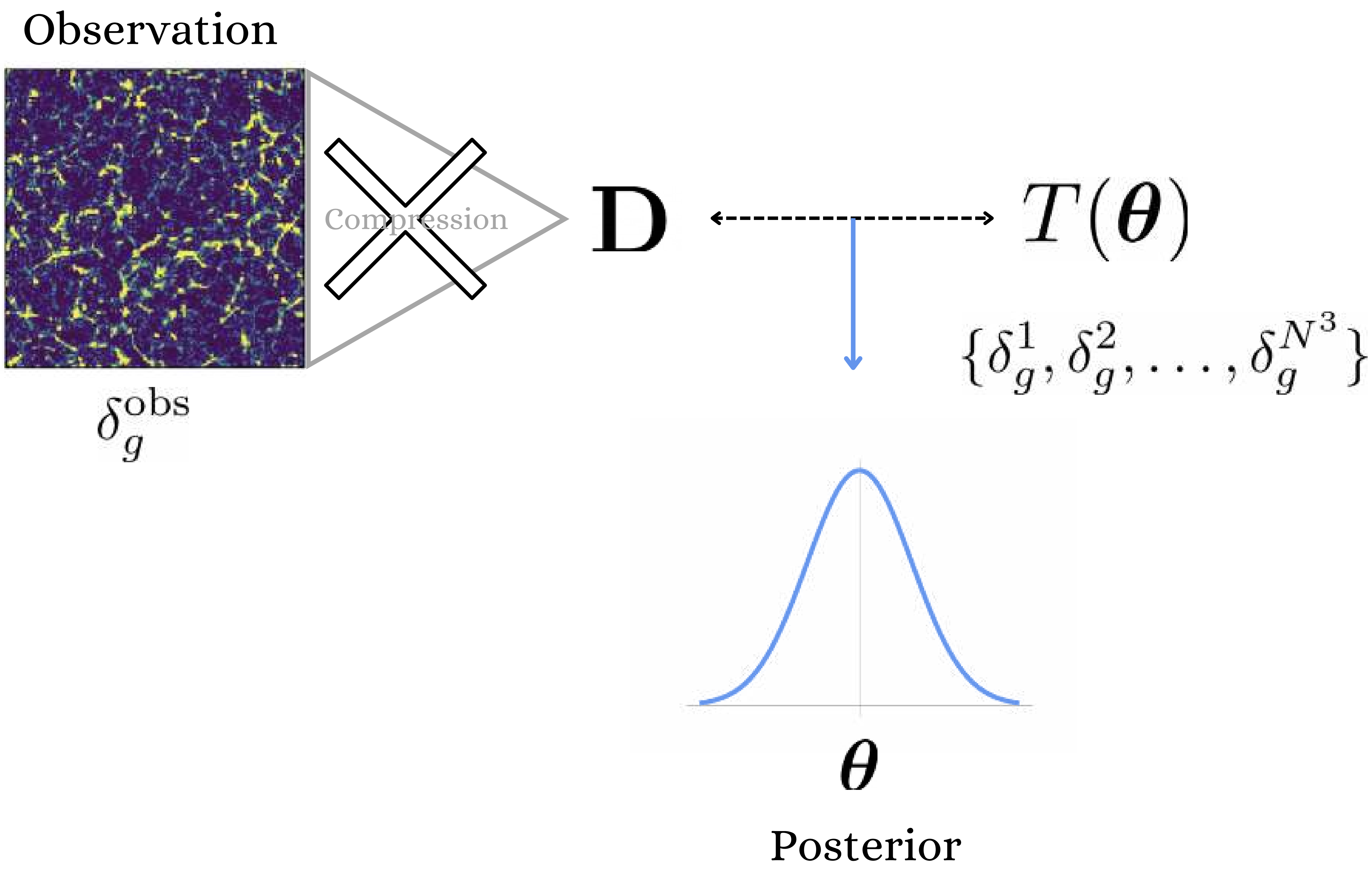
SBI

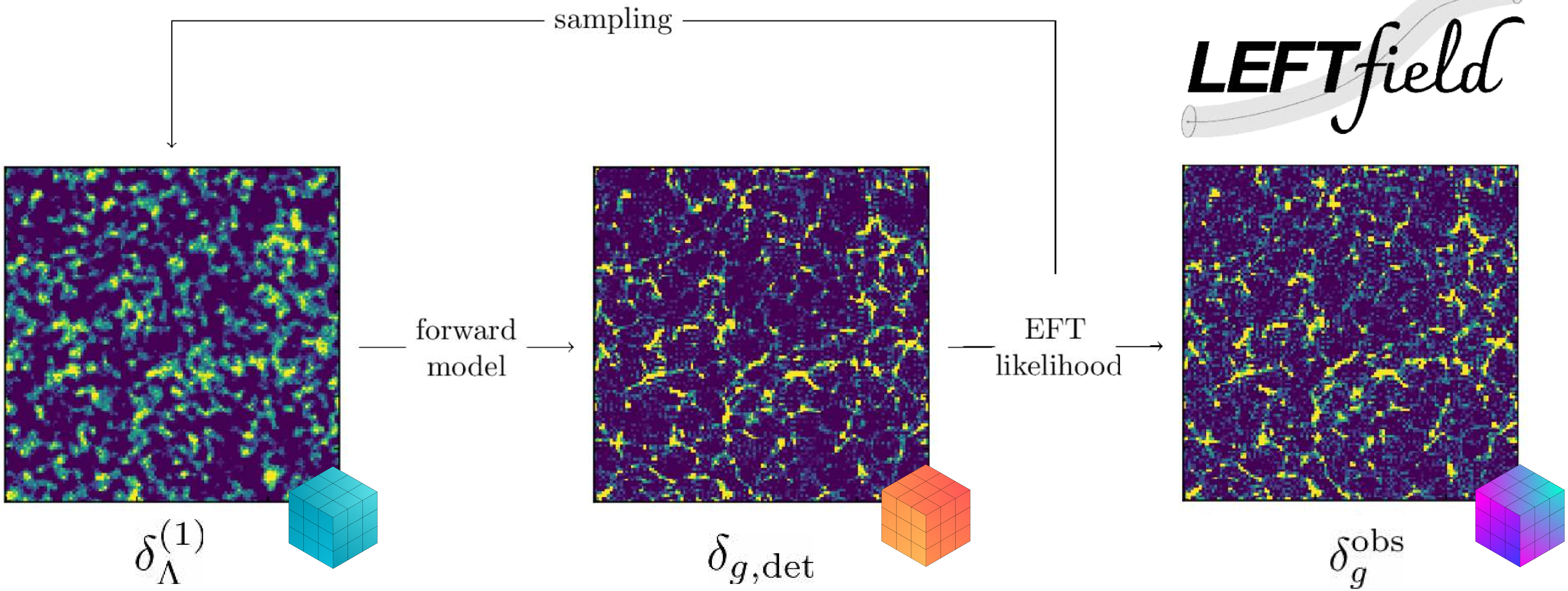
Part II

Field-level Bayesian inference

(FBI)

FBI: the main idea



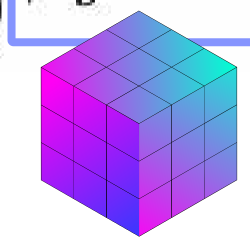


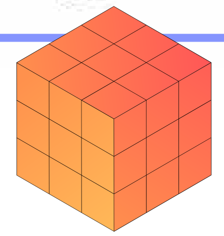
Credits: Julia Stadler

Field level Likelihood

Mode by mode
data and theory
comparison!

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

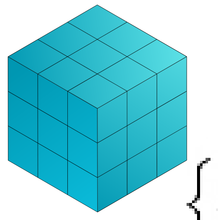




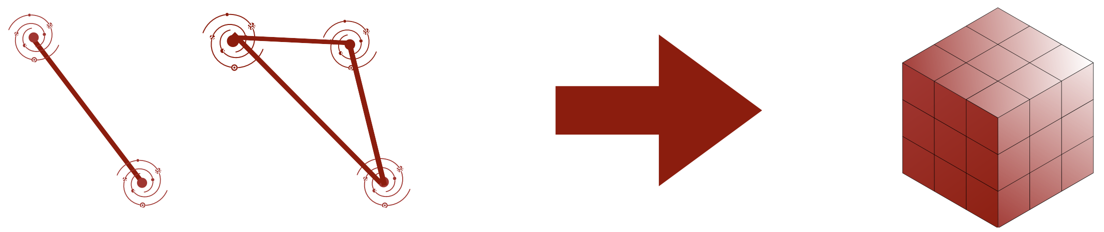
$\downarrow$ HMC

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

Full posterior
including initial
conditions!



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



How much information is retained at the galaxy density field?

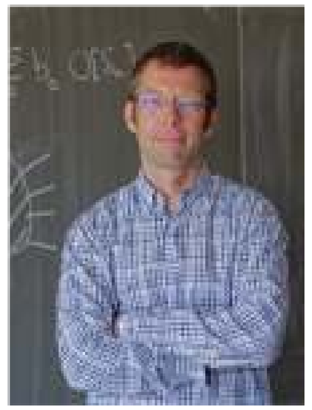
Breaking degeneracy between σ_8 and bias parameters on dark-matter halos

Nguyen, Schmidt, **Tucci** et al. (2024)
PRL (accepted)

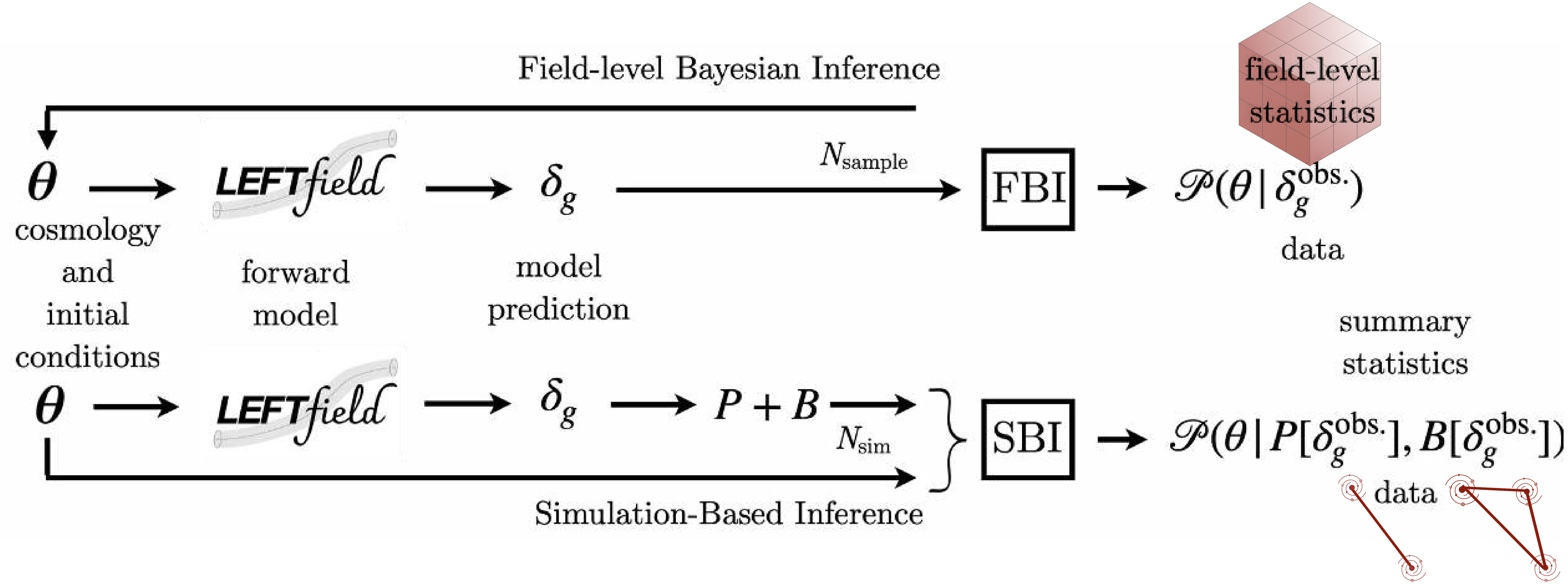
3rd order bias expansion

$$
\boxed{\begin{aligned} O_{\rm det} &\in [\delta, \delta^2, K^2, \delta^3, K^3, \delta K^2, O_{\rm td}, \nabla^2\delta] \\ O_{\rm stoch} &\in [\varepsilon, \nabla^2\varepsilon] \end{aligned}}
$$

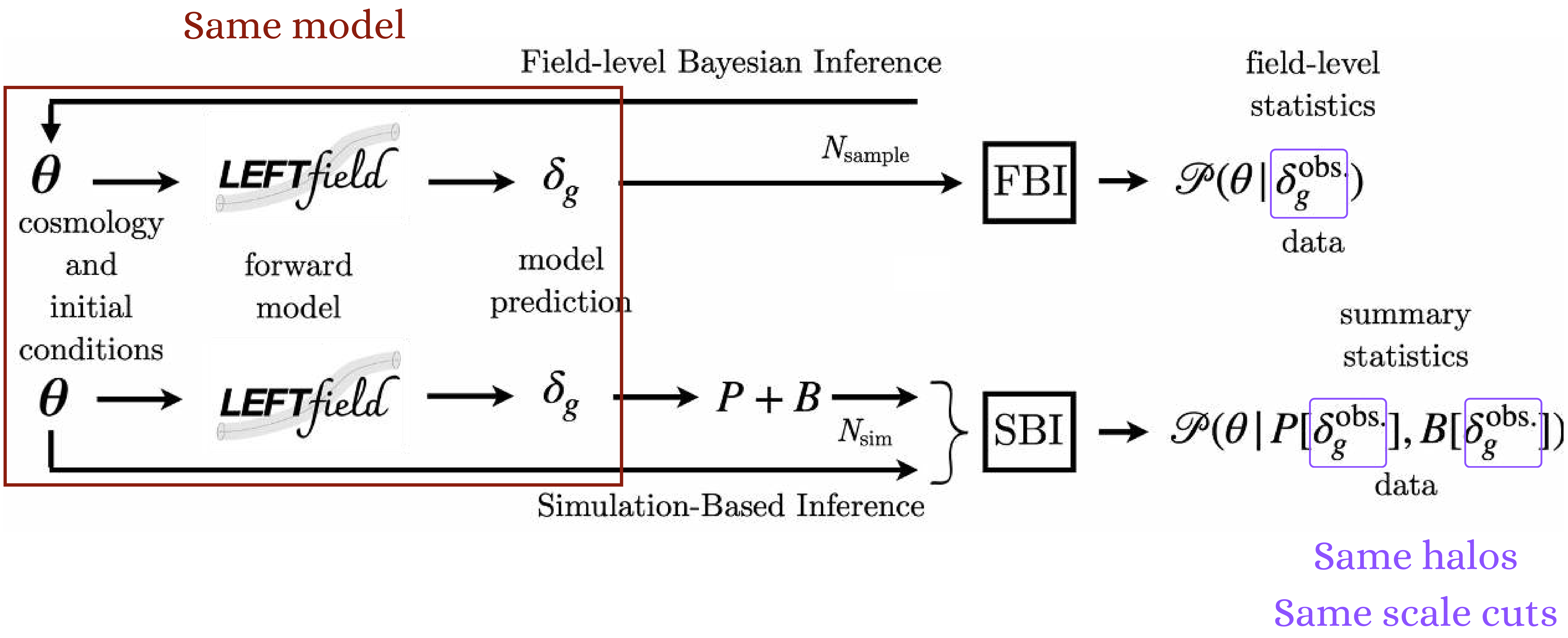
Nhat-Minh Nguyen
(IPMU)



Fabian Schmidt
(MPA)

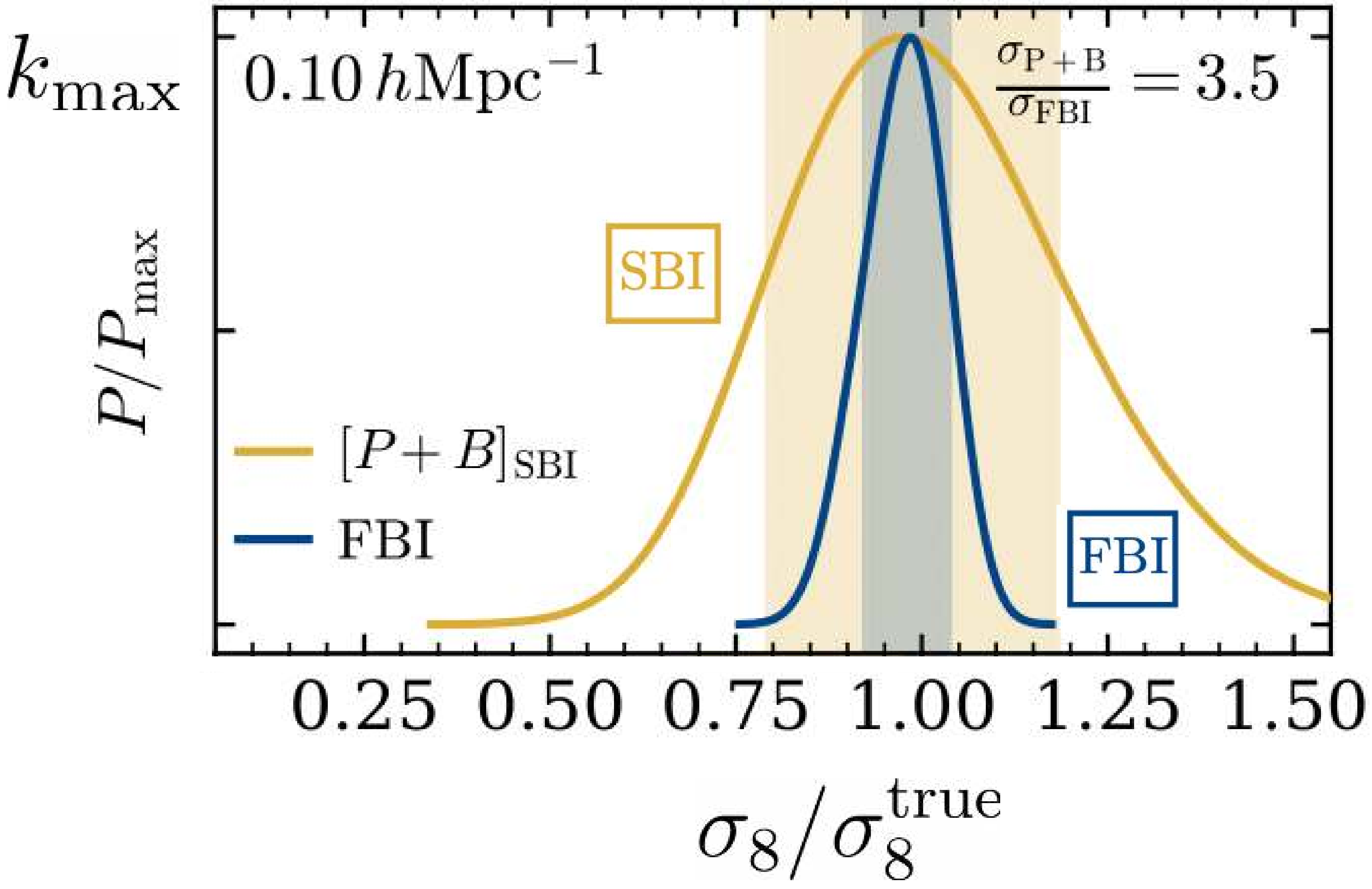


Apples-to-apples comparison



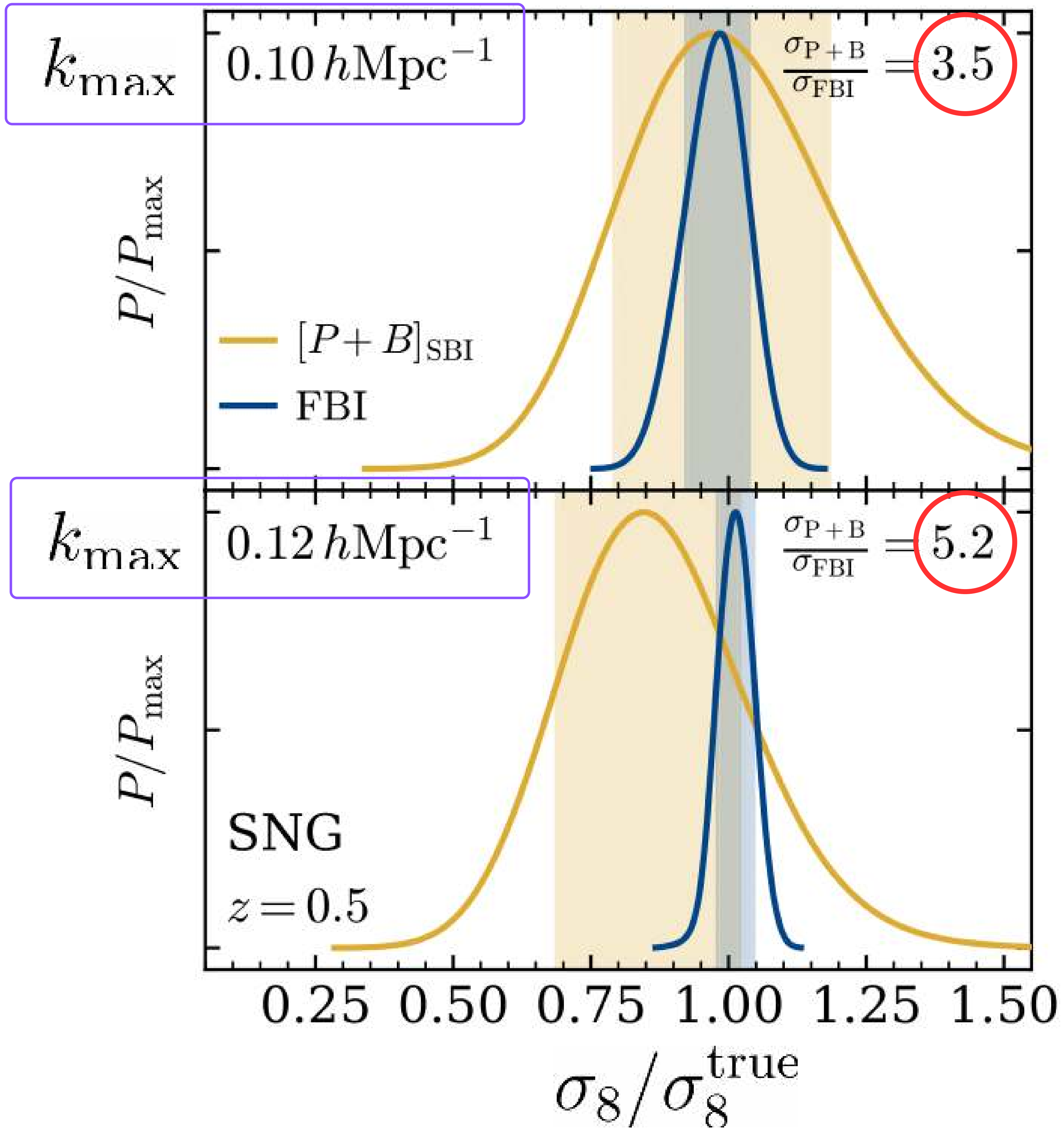
A lot of reliable information at the field-level!

SNG halos



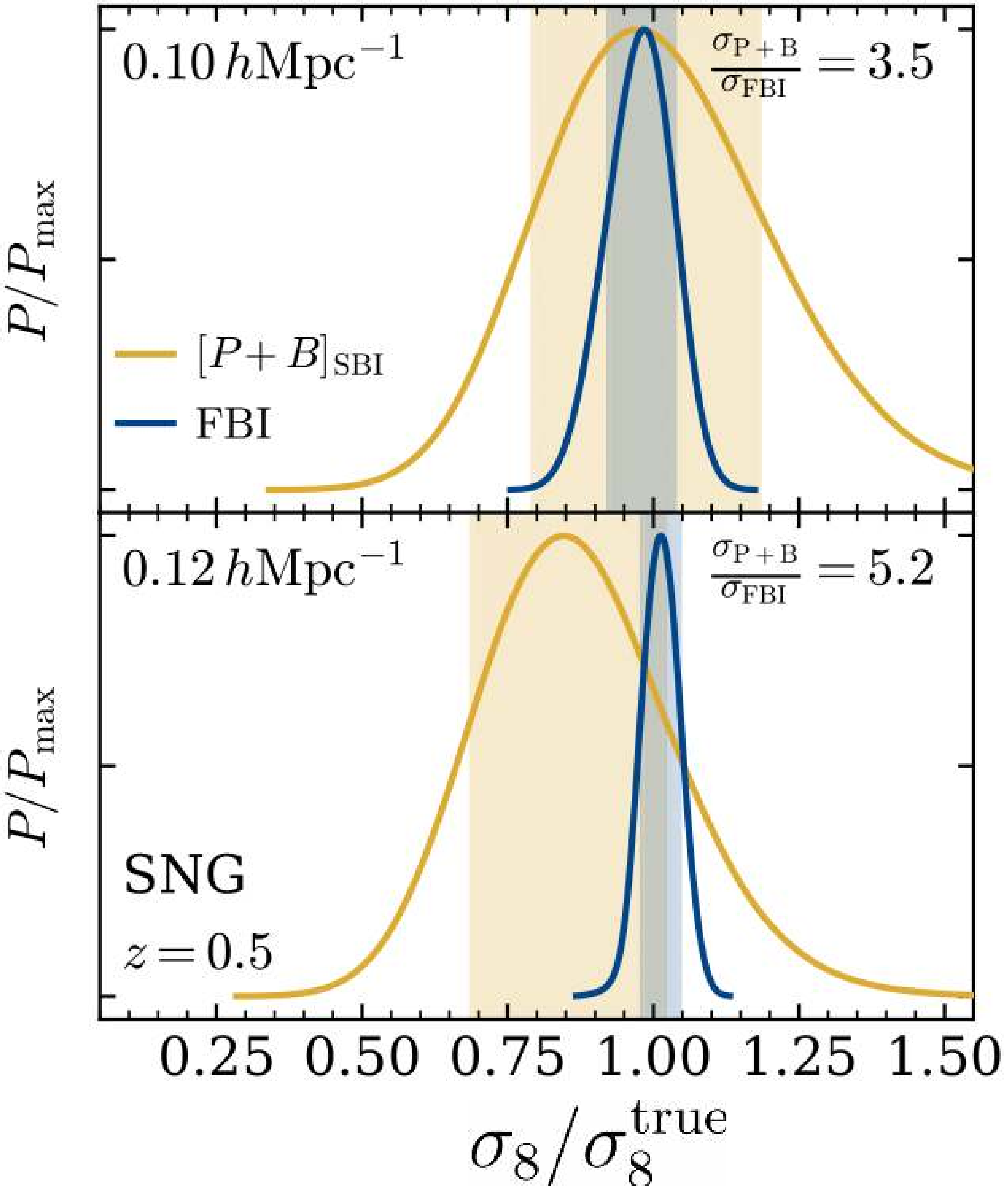
3.5 improvement factor!

SNG



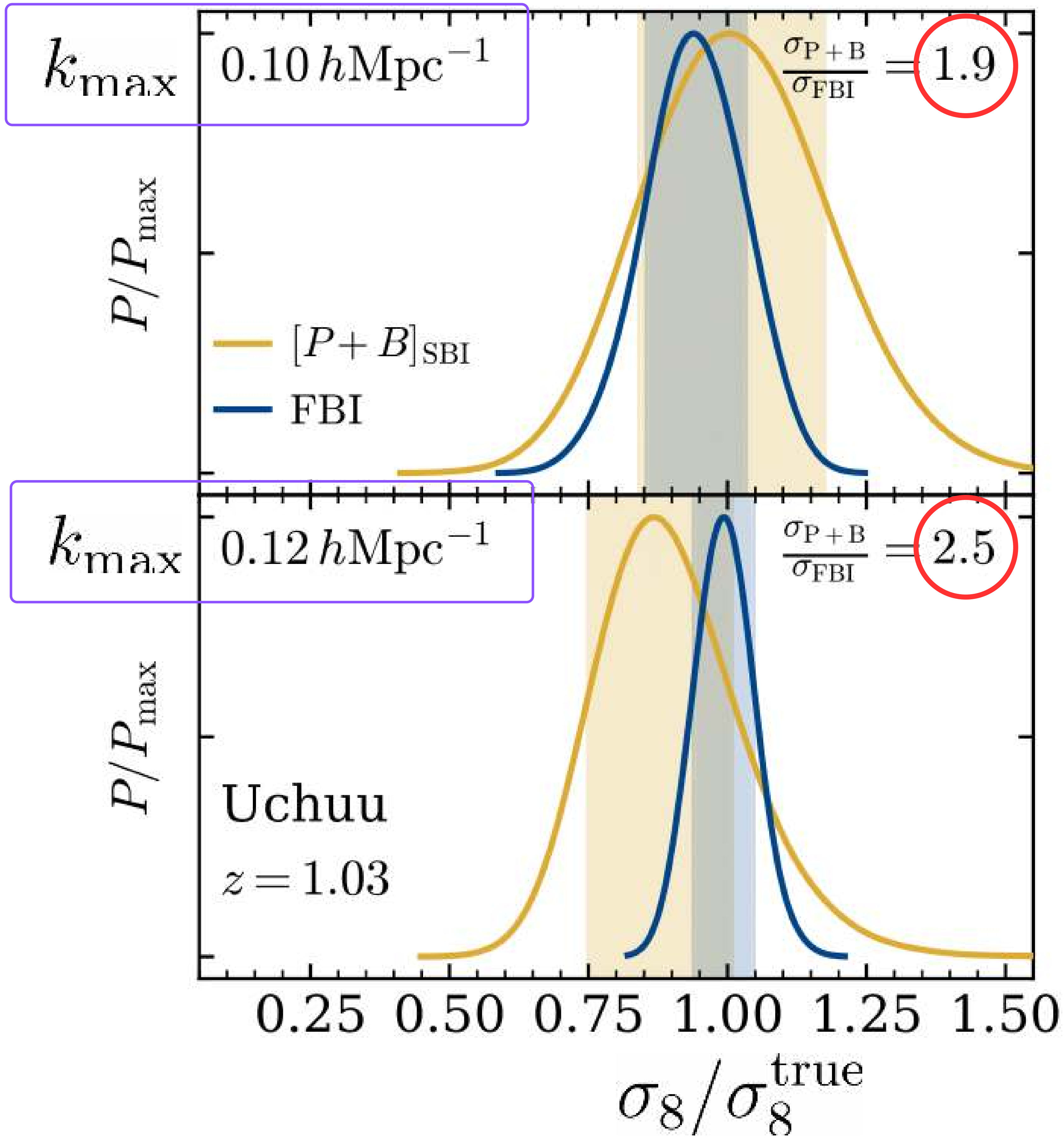
$k_{\max}$
$0.10\,h\mathrm{Mpc}^{-1}$
$\frac{\sigma_{\mathrm{P+B}}}{\sigma_{\mathrm{FBI}}} = 3.5$
$P/P_{\max}$
$[P+B]_{\mathrm{SBI}}$
FBI
$k_{\max}$
$0.12\,h\mathrm{Mpc}^{-1}$
$\frac{\sigma_{\mathrm{P+B}}}{\sigma_{\mathrm{FBI}}} = 5.2$
$P/P_{\max}$
SNG
$z = 0.5$
0.25 0.50 0.75 1.00 1.25 1.50
$\sigma_8/\sigma_8^{\mathrm{true}}$

SNG



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

Uchuu



$k_{\max}$
0.10 hMpc^{-1}
$\frac{\sigma_{\mathrm{P+B}}}{\sigma_{\mathrm{FBI}}} = 1.9$
$P/P_{\max}$
$[P+B]_{\mathrm{SBI}}$
FBI
$k_{\max}$
0.12 hMpc^{-1}
$\frac{\sigma_{\mathrm{P+B}}}{\sigma_{\mathrm{FBI}}} = 2.5$
$P/P_{\max}$
Uchuu
$z = 1.03$
0.25 0.50 0.75 1.00 1.25 1.50
$\sigma_8/\sigma_8^{\mathrm{true}}$

On the Bispectrum stochasticity

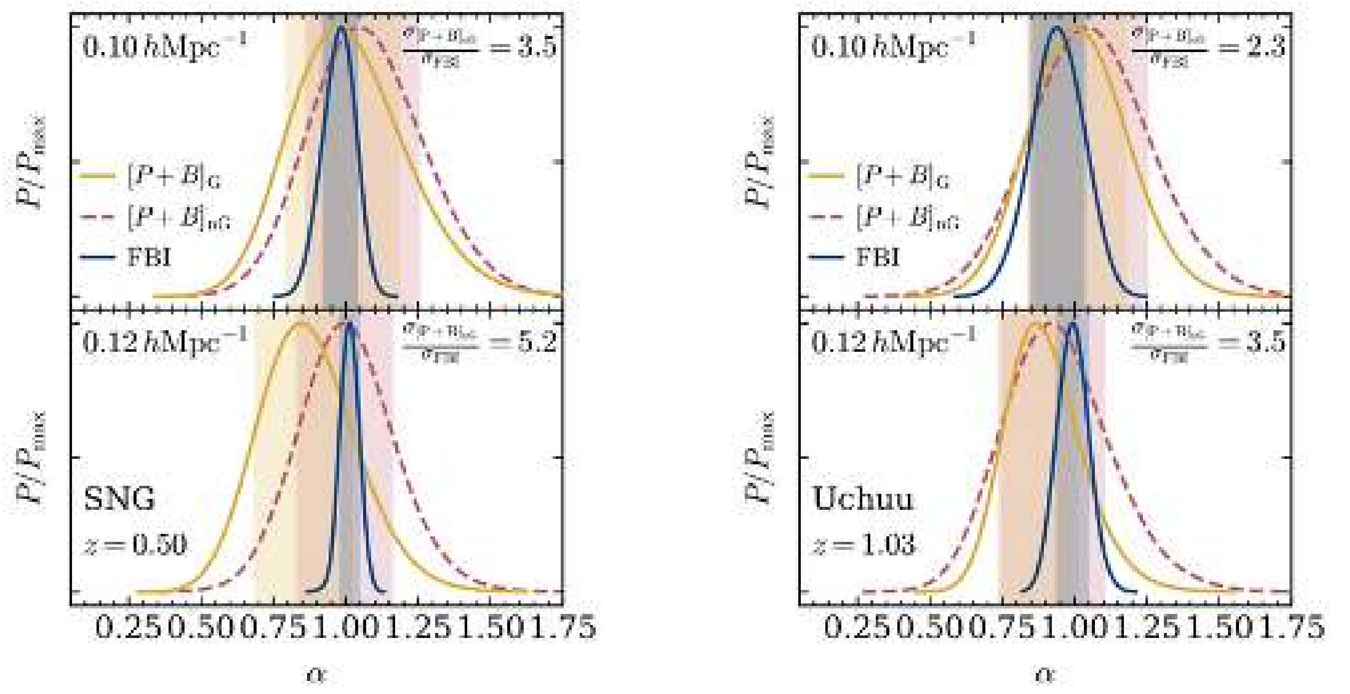
Perturbation Theory

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

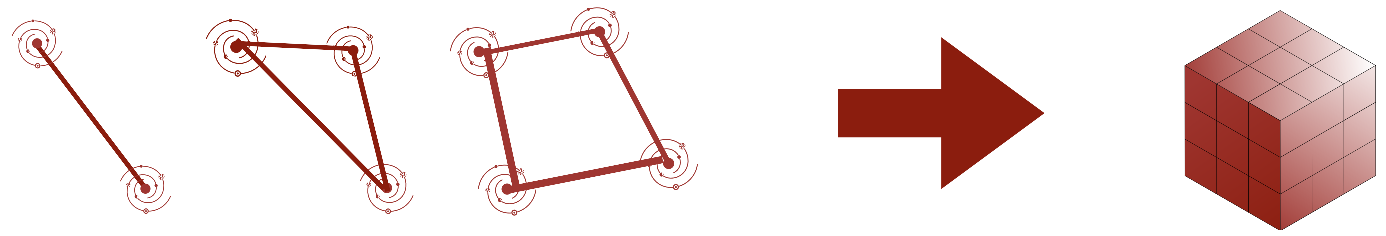
Forward Model with Non-Gaussian Noise

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

$$\boxed{\delta_g(\mathbf{x},\tau) = \delta_{g,\text{det}}(\mathbf{x},\tau) + \varepsilon(\mathbf{x},\tau) + \sigma_{\varepsilon\delta}(\tau)\varepsilon(\mathbf{x},\tau)\delta(\mathbf{x},\tau) + c^{\text{NG}}_\varepsilon(\tau)\varepsilon^2(\mathbf{x},\tau)} \qquad \varepsilon \sim \mathcal{N}(0,\sigma^2_\varepsilon)$$



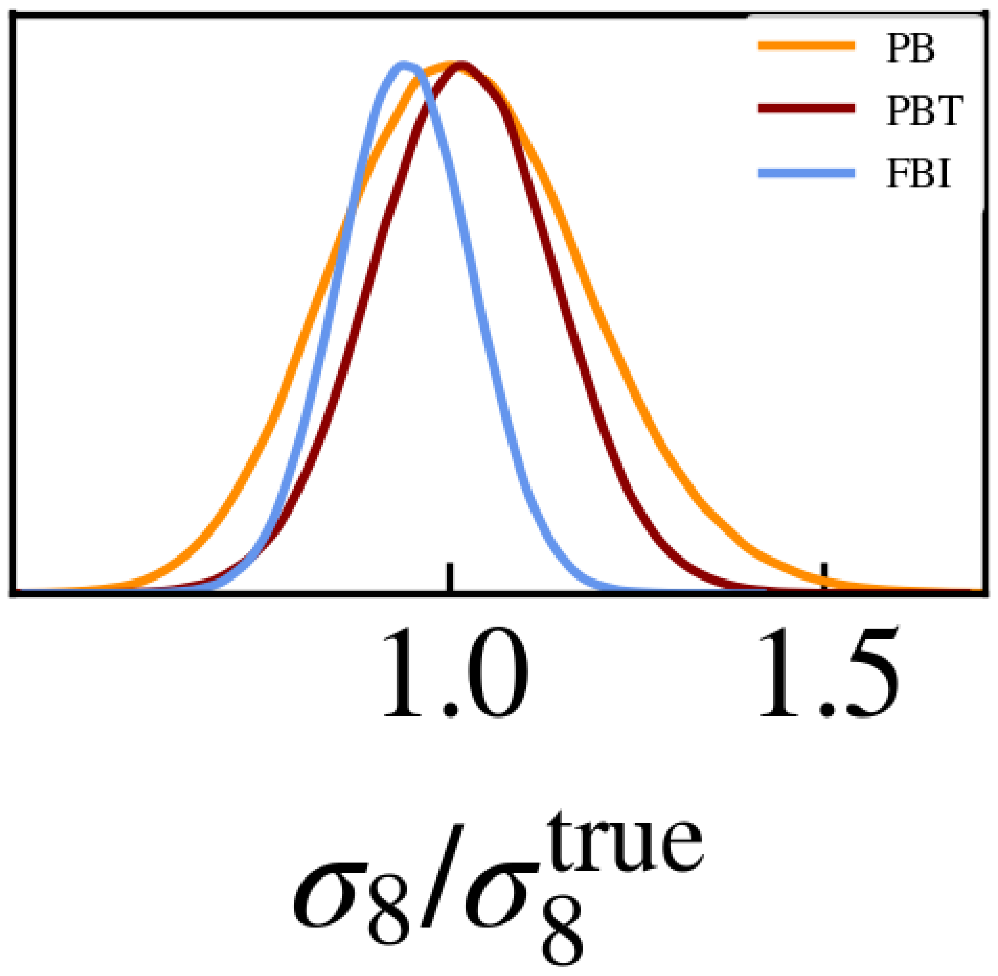
$$O_{\text{det}} \in [\delta, \delta^2, K^2, O_{\text{td}}, \nabla^2\delta]$$
$$O_{\text{stoch}} \in [\varepsilon, \varepsilon\delta, \varepsilon^2, \nabla^2\varepsilon]$$

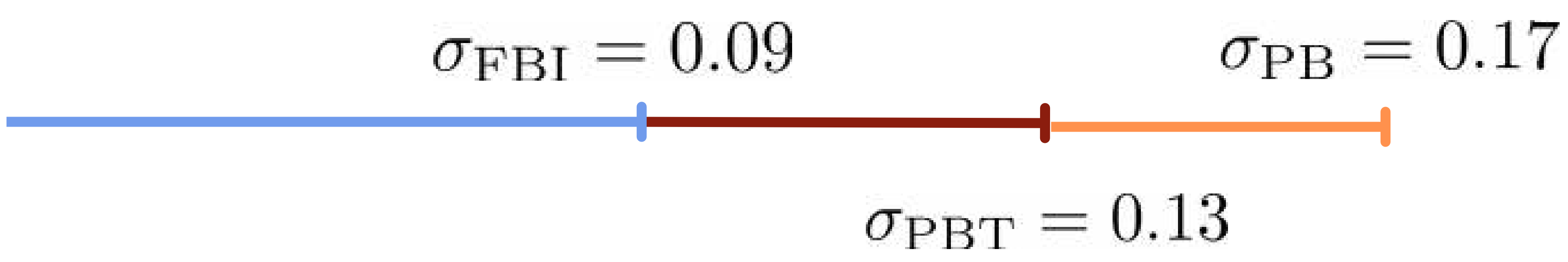


What if we add the **trispectrum**?

Tucci & Schmidt (in prep.)

Trispectrum: **preliminary** results

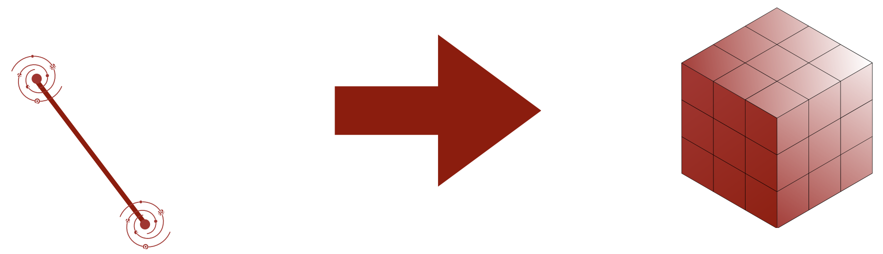




$k_{\mathrm{max}} = 0.1 h\,\mathrm{Mpc}^{-1}$

Uchuu halos at z=1

Brute force approach:
10^6 simulations



Field-level inference of BAO scale

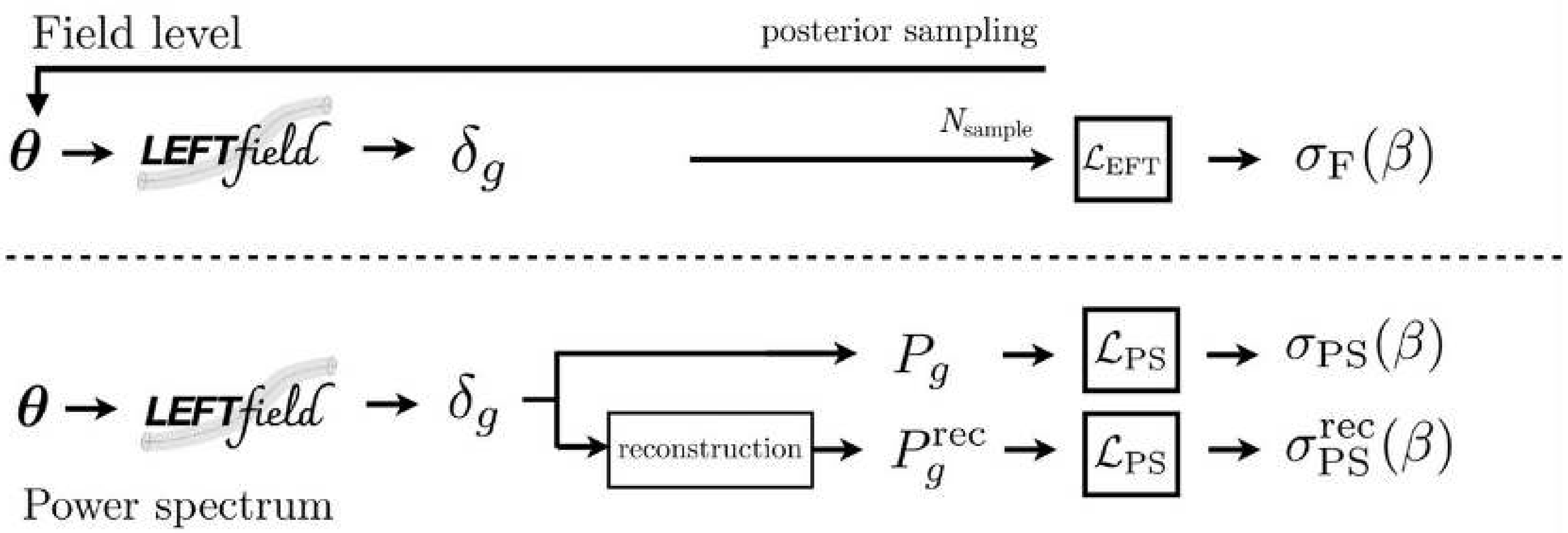
Constraining the BAO scale with FBI

Babić, Schmidt & **Tucci** (2022)

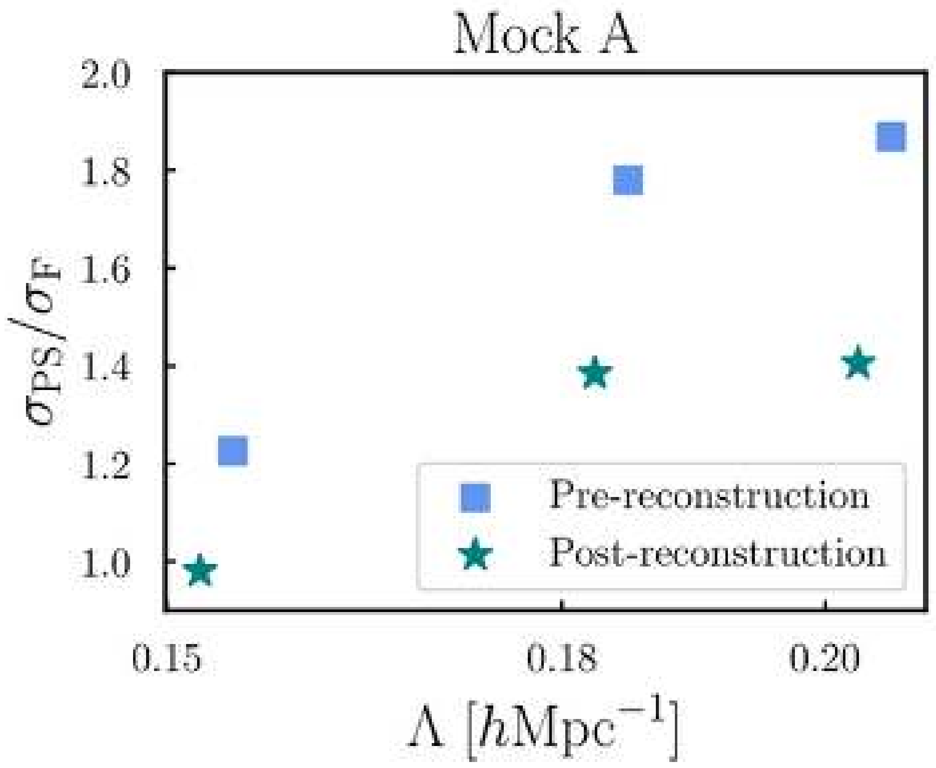
Babić, Schmidt & **Tucci** (2024)

Can we constrain the BAO scale with FBI?

Ivana Babić
(MPA)



$$\beta \equiv r_s / r_{\rm fid}$$



Babić, Schmidt & Tucci (2022)
Babić, Schmidt & Tucci (2024)

Conclusion & Next Steps

- We demonstrated to have **unbiased** and **accurate** results from halo catalogs using LEFTfield for SBI and FBI
- **Apple-to-apple comparison** of field-level inference and SBI shows that there is a lot of **reliable** information beyond 2+3(+4)-point functions in the 3D maps of galaxies

Next steps to connect with observations:

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

MAX-PLANCK-INSTITUT FÜR ASTROPHYSIK

LEFTfield

Beatriz Tucci

tucci@mpa-garching.mpg.de