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

Speakers: Beatriz Tucci

Collection/Series: Cosmology and Gravitation

Subject: Cosmology

Date: December 03, 2024 - 11:00 AM

URL: https://pirsa.org/24120021

Abstract:

Improving cosmological constraints from galaxy clustering presents several challenges, particularly in extracting information beyond the power spectrum due to the complexities involved in higher-order n-point function analysis. In this talk, I will introduce novel inference techniques that allow us to go beyond the state-of-the-art, not only by utilizing the galaxy trispectrum, a task that remains computationally infeasible with traditional methods, but also by accessing the full information encoded in the galaxy density field for the first time in cosmological analysis. I will present simulation-based inference (SBI), a powerful deep learning technique that enables cosmological inference directly from summary statistics in simulations, bypassing the need for explicit analytical likelihoods or covariance matrices. This is achieved using LEFTfield, a Lagrangian forward model based on the Effective Field Theory of Large Scale Structure (EFTofLSS) and the bias expansion, ensuring robustness on large scales. Furthermore, LEFTfield enables field-level Bayesian inference (FBI), where a field-level likelihood is used to directly analyze the full galaxy density field rather than relying on compressed statistics. I will conclude by exploring the question of how much cosmological information can be extracted at the field level through a comparison of σ_8 constraints obtained from FBI, which directly uses the 3D galaxy density field, and those obtained from n-point functions via SBI.

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

Beatriz Tucci

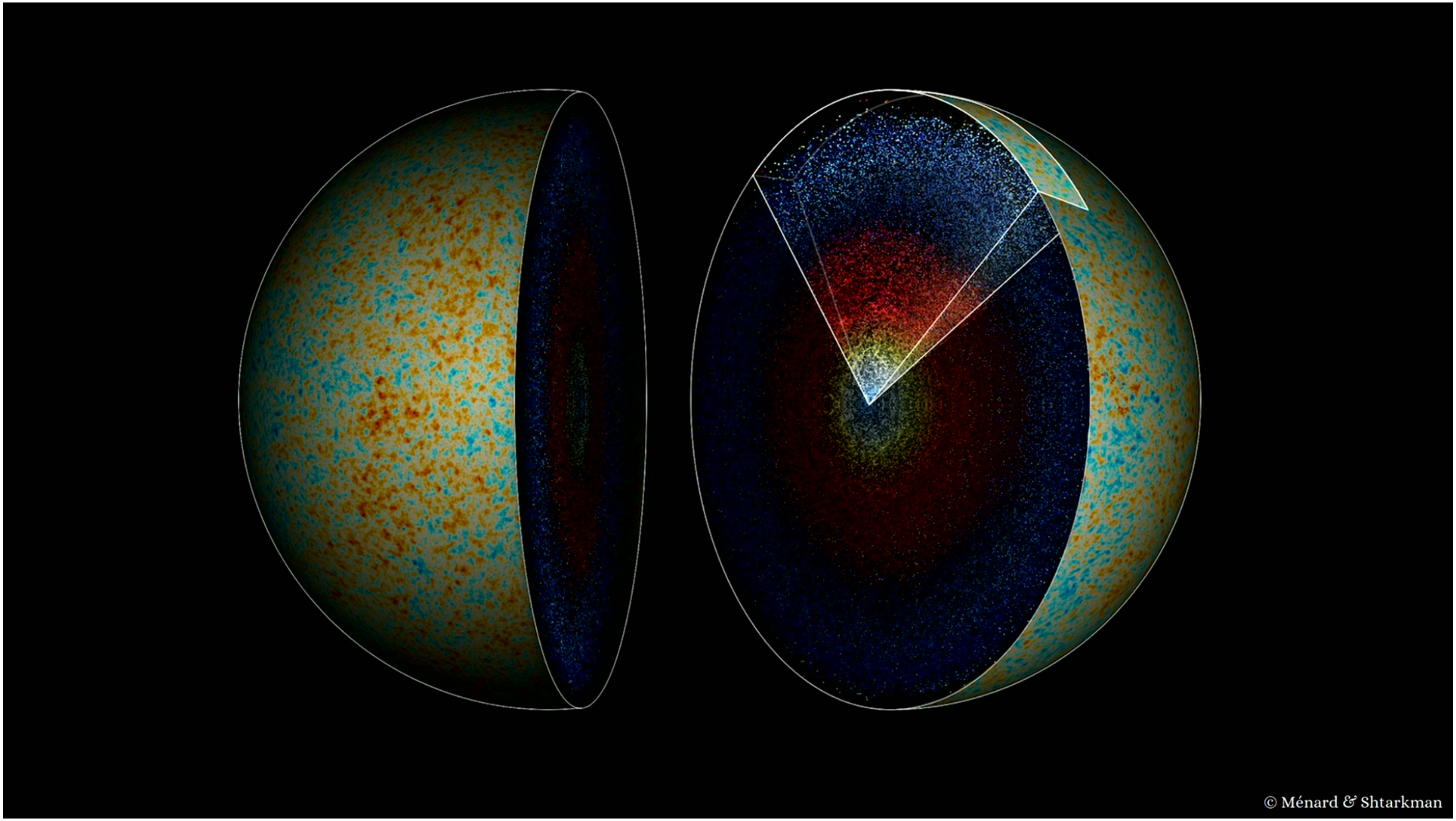
Max Planck Institute for Astrophysics (MPA)

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

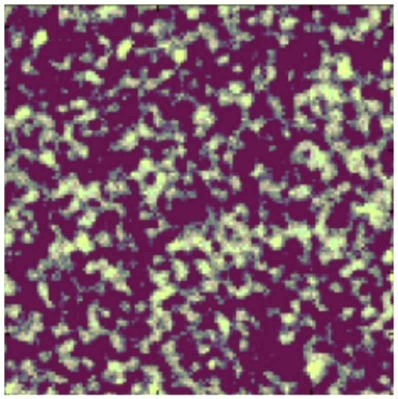
LEFTfield

MAX-PLANCK-INSTITUT
FÜR ASTROPHYSIK

Perimeter Institute | Cosmology Seminar 2024


© Ménard & Shtarkman

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



Initial conditions of the Universe

time

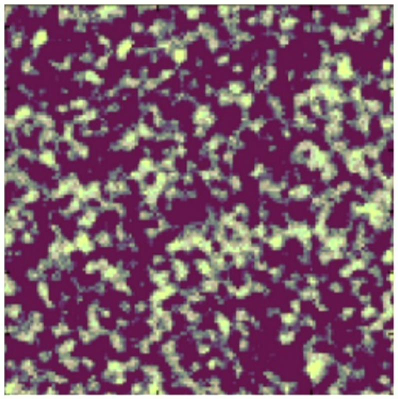
(Slice of 3D field)

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

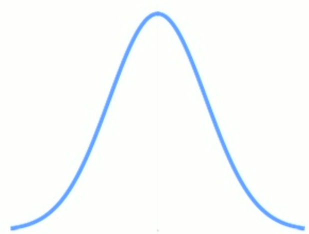
Cosmological parameters $\boldsymbol{\theta}$

Final distribution of galaxies

Observation

inference

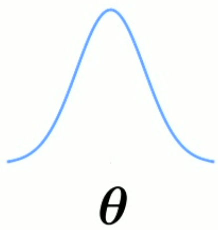
theory(θ)



$\boldsymbol{\theta}$

Posterior of cosmological parameters

Bayesian inference



Posterior Likelihood Prior

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

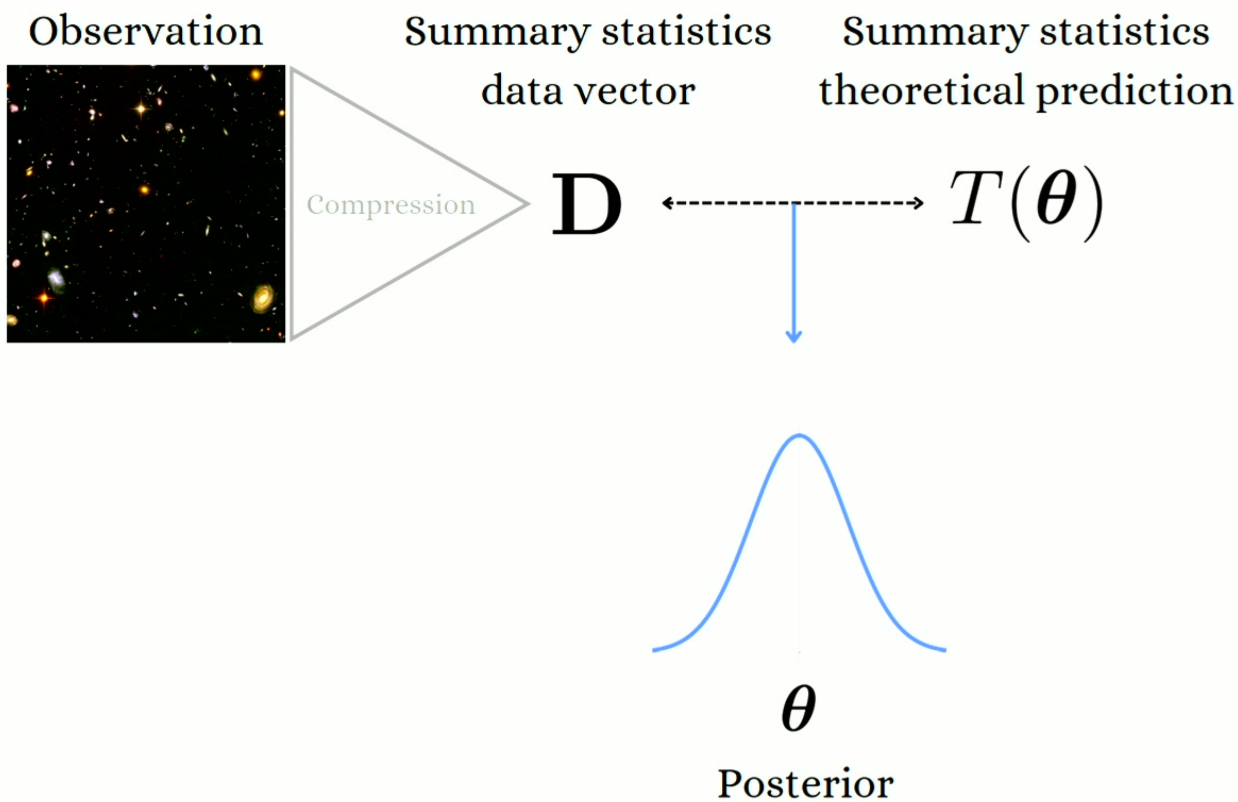
E.g., assuming that the data vector is Gaussian distributed:

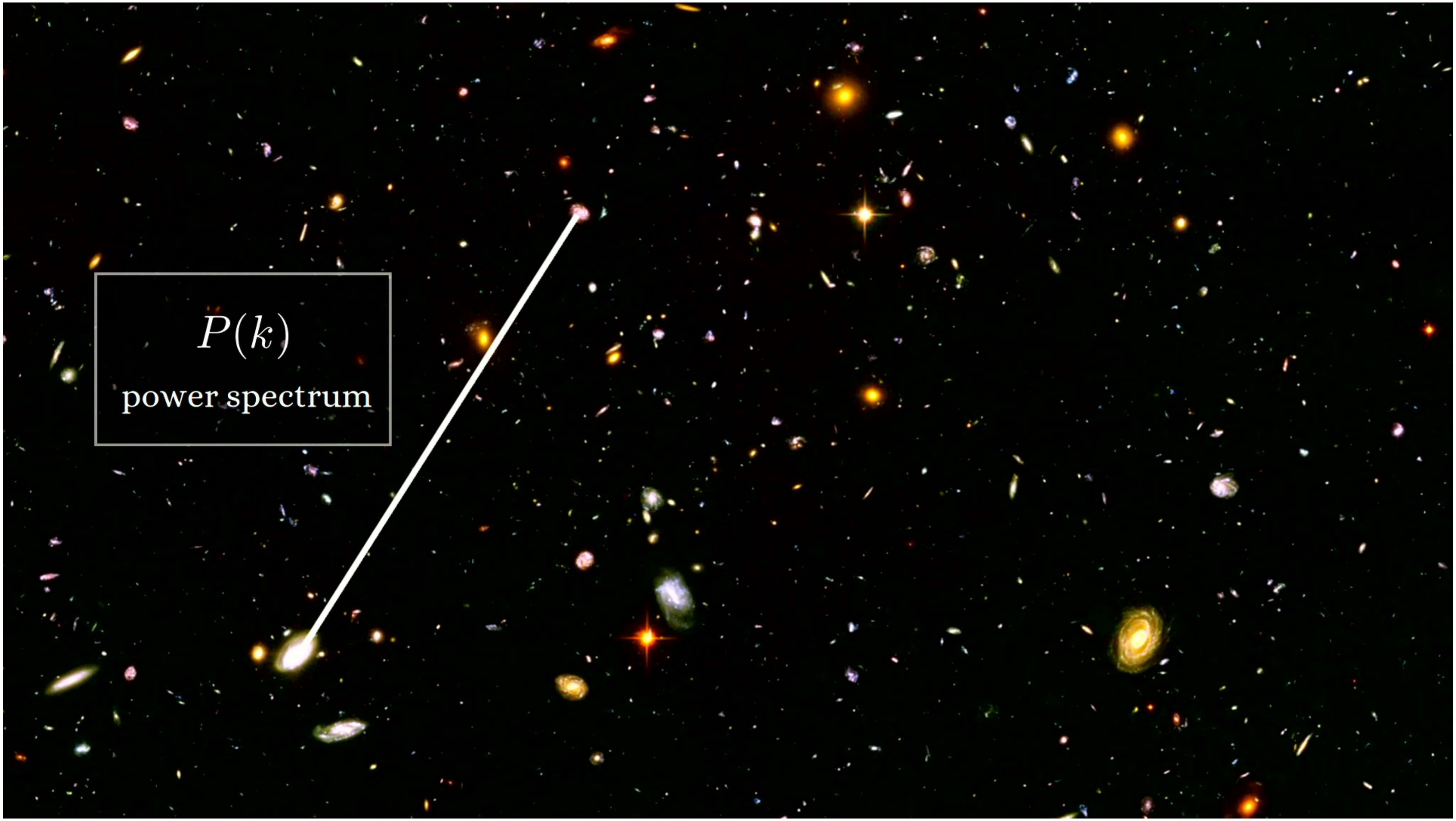
$$-2\ln\mathcal{L}(\mathbf{D}|\boldsymbol{\theta}) = \big(\mathbf{D} - \mathbf{T}(\boldsymbol{\theta})\big)\cdot\mathbf{C}^{-1}\cdot\big(\mathbf{D} - \mathbf{T}(\boldsymbol{\theta})\big)$$

Data vector

Covariance of the data vector

Theoretical prediction of data vector

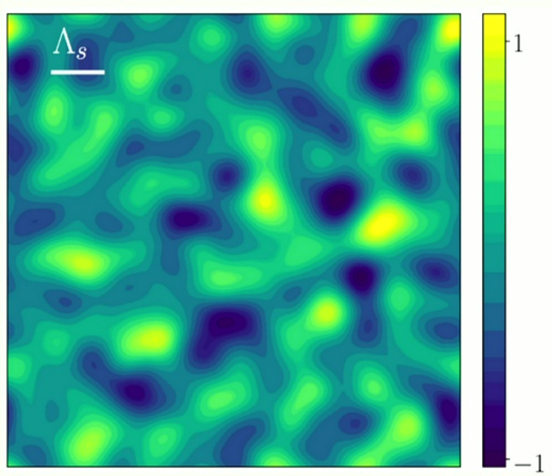

Observation
Summary statistics
data vector
Summary statistics
theoretical prediction
Compression
$\mathbf{D}$
$T(\boldsymbol{\theta})$
$\boldsymbol{\theta}$
Posterior


$P(k)$
power spectrum

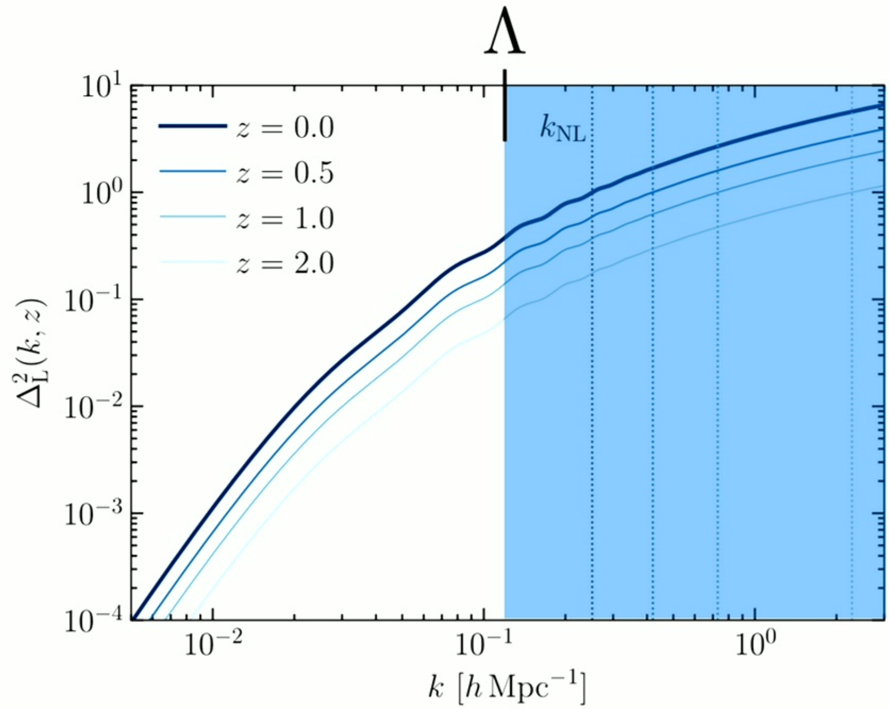
The EFTofLSS

"coarse-graining"

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



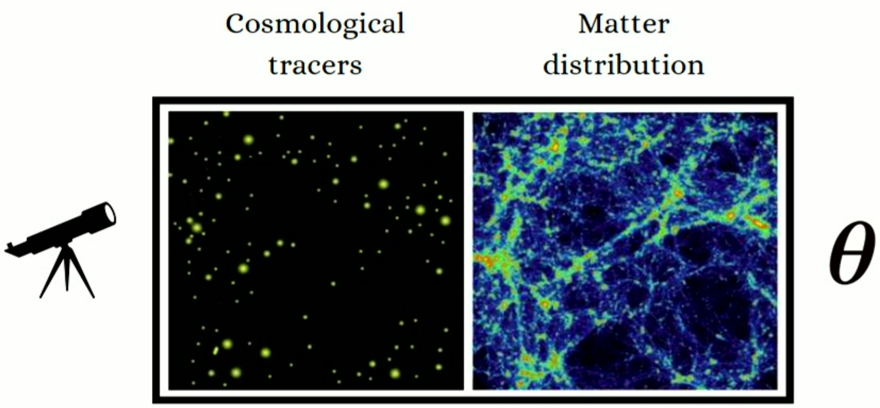
Borrowed from Pierre Zhang



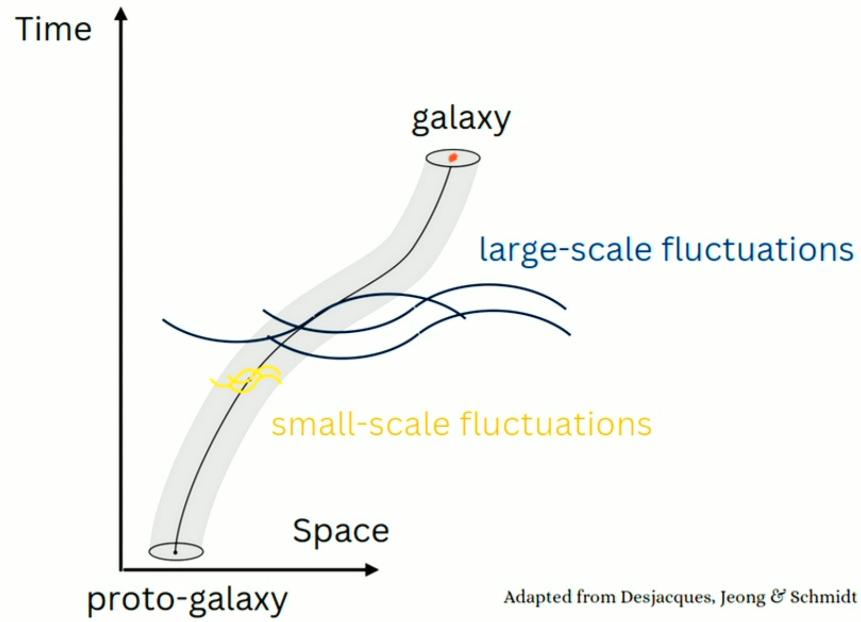
Dodelson & Schmidt
Modern Cosmology 2020

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)



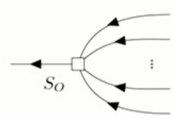
Adapted from Desjacques, Jeong & Schmidt (2016)

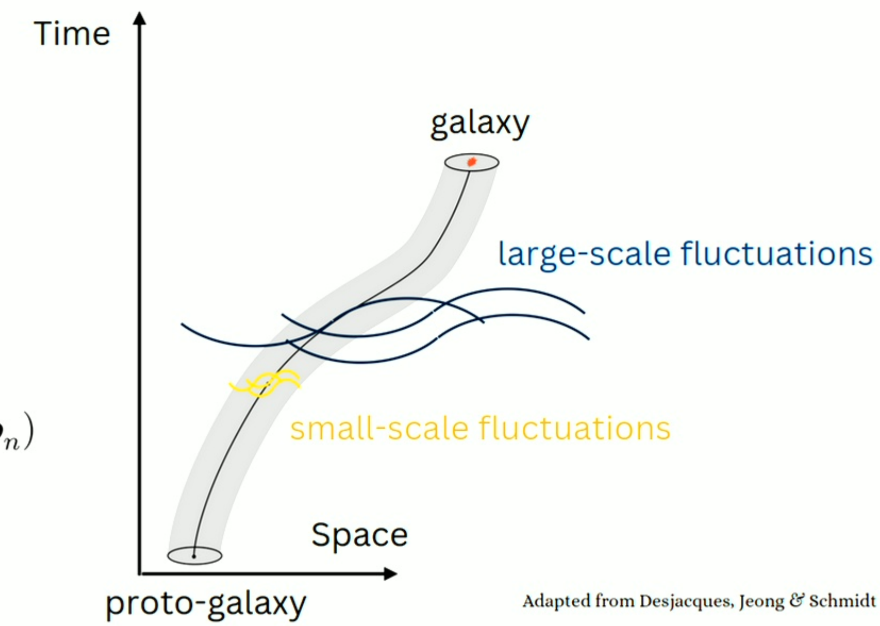
The bias expansion

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

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

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





Adapted from Desjacques, Jeong & Schmidt (2016)

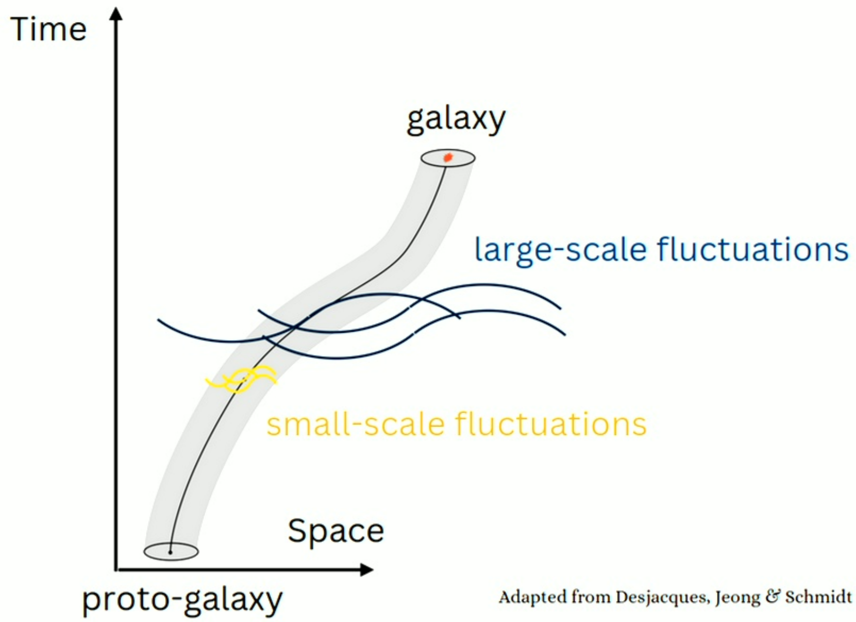
The bias expansion

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

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

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

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

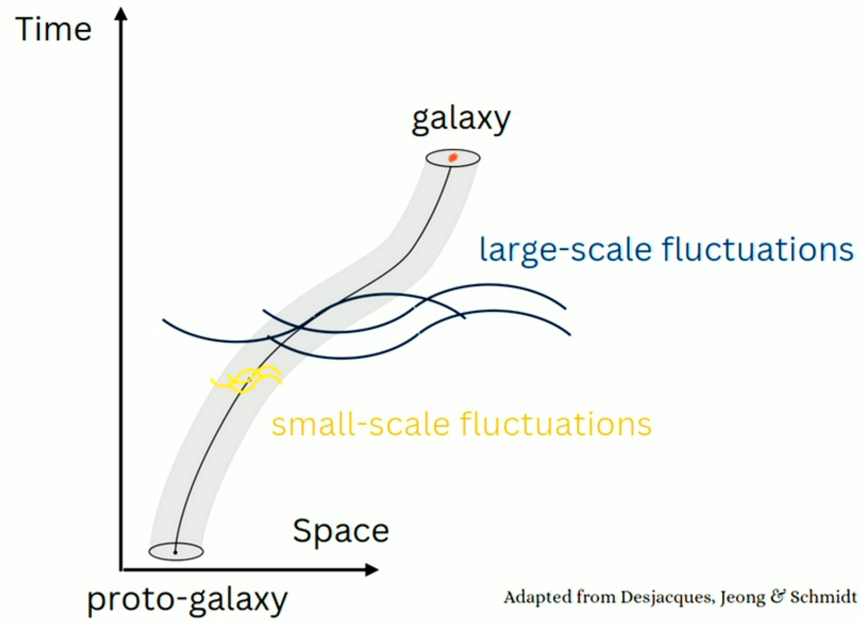


Adapted from Desjacques, Jeong & Schmidt (2016)

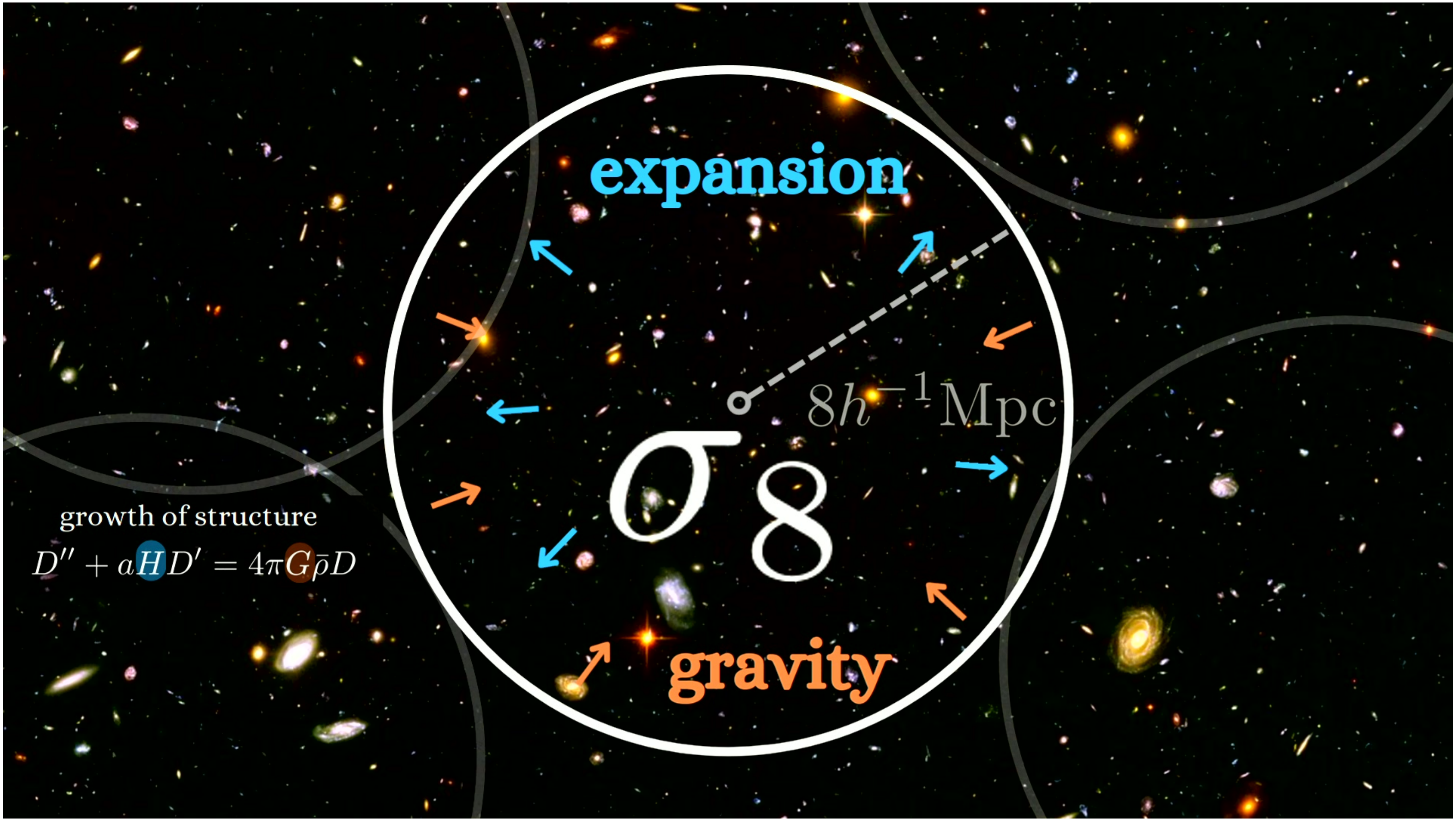
The bias expansion

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

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



Adapted from Desjacques, Jeong & Schmidt (2016)


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

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

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

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

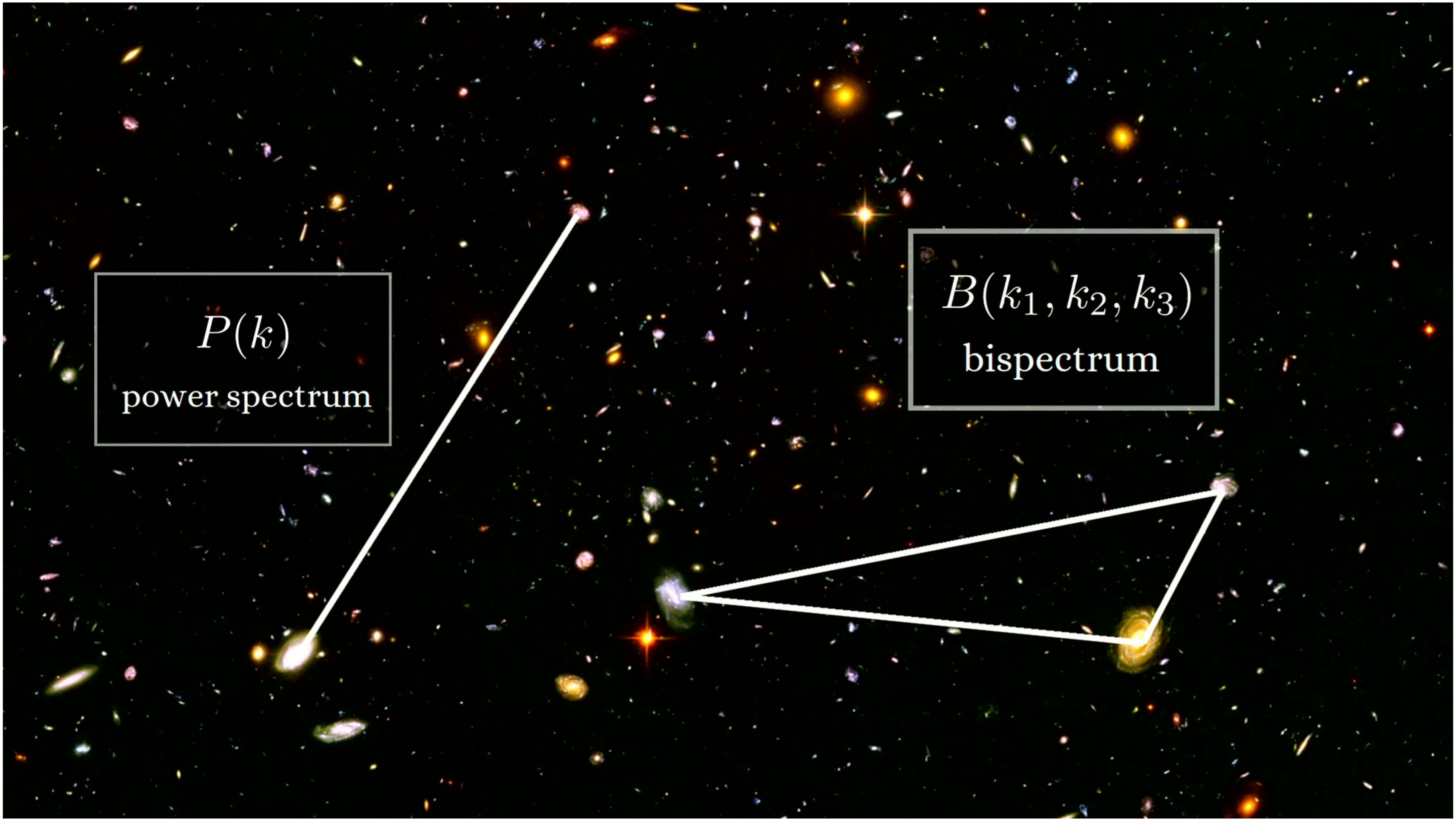
$$P_g^{\rm tree}(k) = b_1^2 P_L(k) + P_\varepsilon$$

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

$$\propto \sigma_8^2$$

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

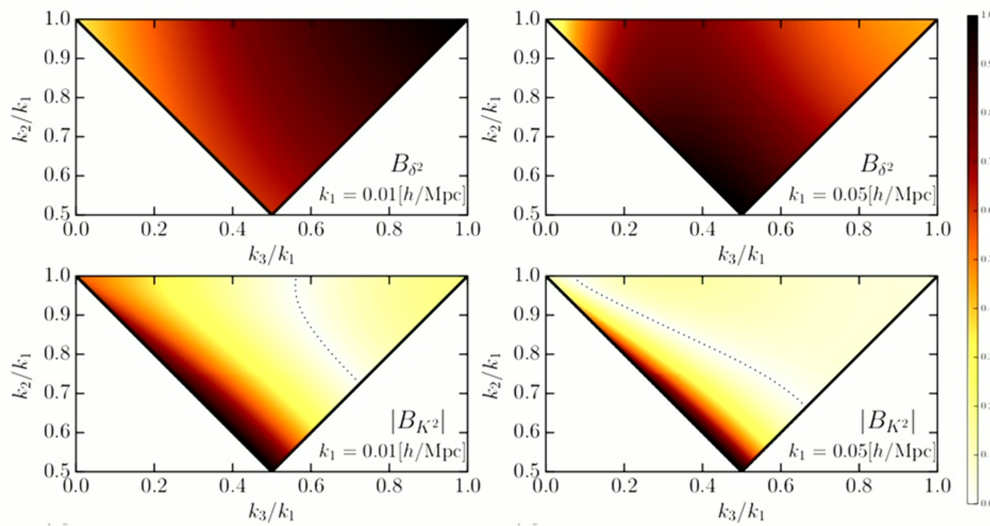
Degeneracy breaking with bispectrum

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

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

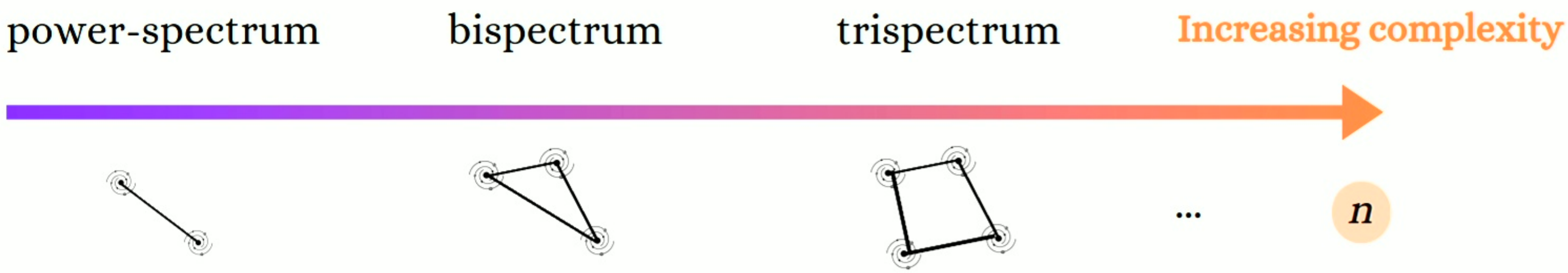
$$B_{K^2}(k_1,k_2,k_3) = \left([\hat{\boldsymbol{k}}_1\cdot\hat{\boldsymbol{k}}_2]^2 - \frac{1}{3}\right)P_L(k_1)P_L(k_2) + 2\ \text{perm.}$$

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



Adapted from Desjacques, Jeong & Schmidt (2016)

Inferring the cosmological parameters: **challenges**



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

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

Inferring the cosmological parameters: **challenges**

power-spectrum bispectrum trispectrum **Increasing complexity**

Analytical approximations — Estimation — Modelling

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

Measurements

Part I

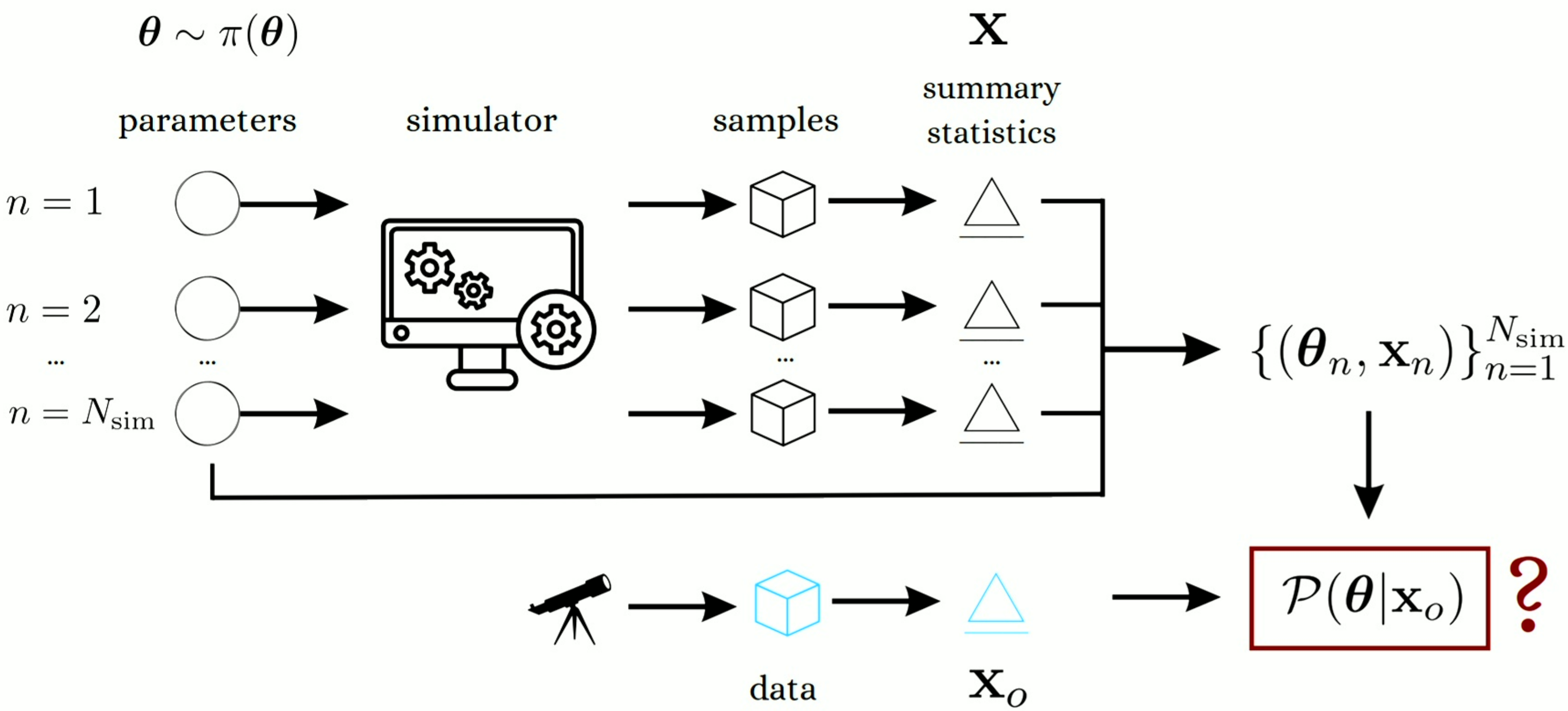
Simulation-based inference
(SBI)

SBI: the main idea

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

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

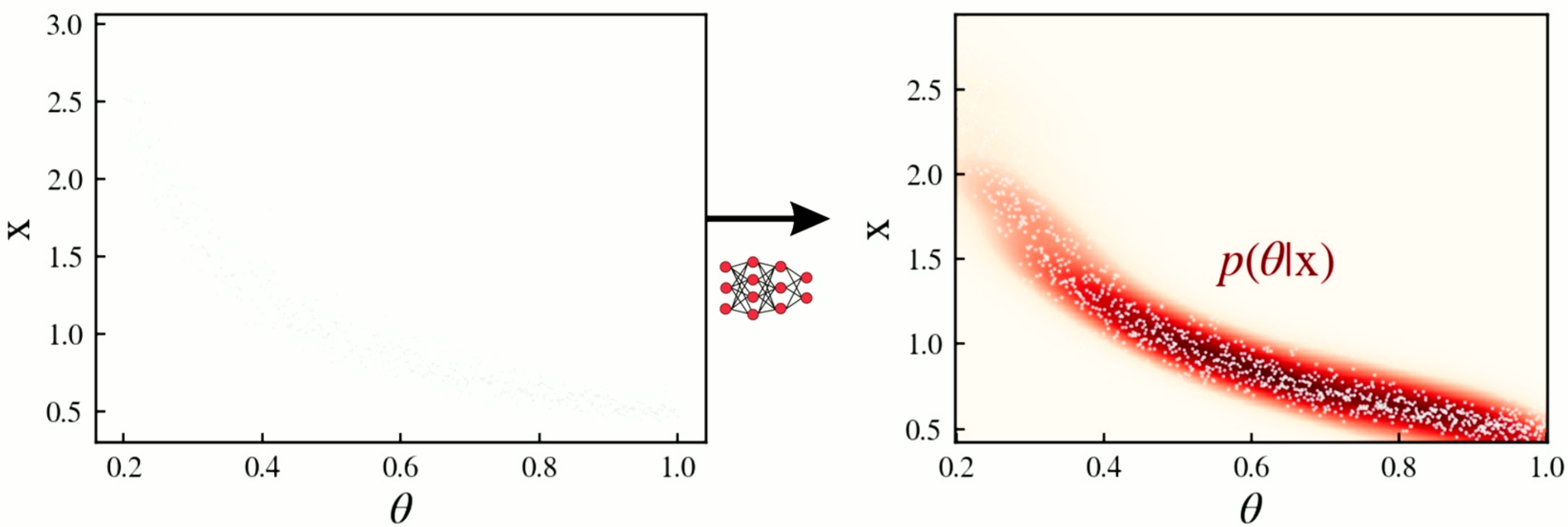
Simulation-based inference



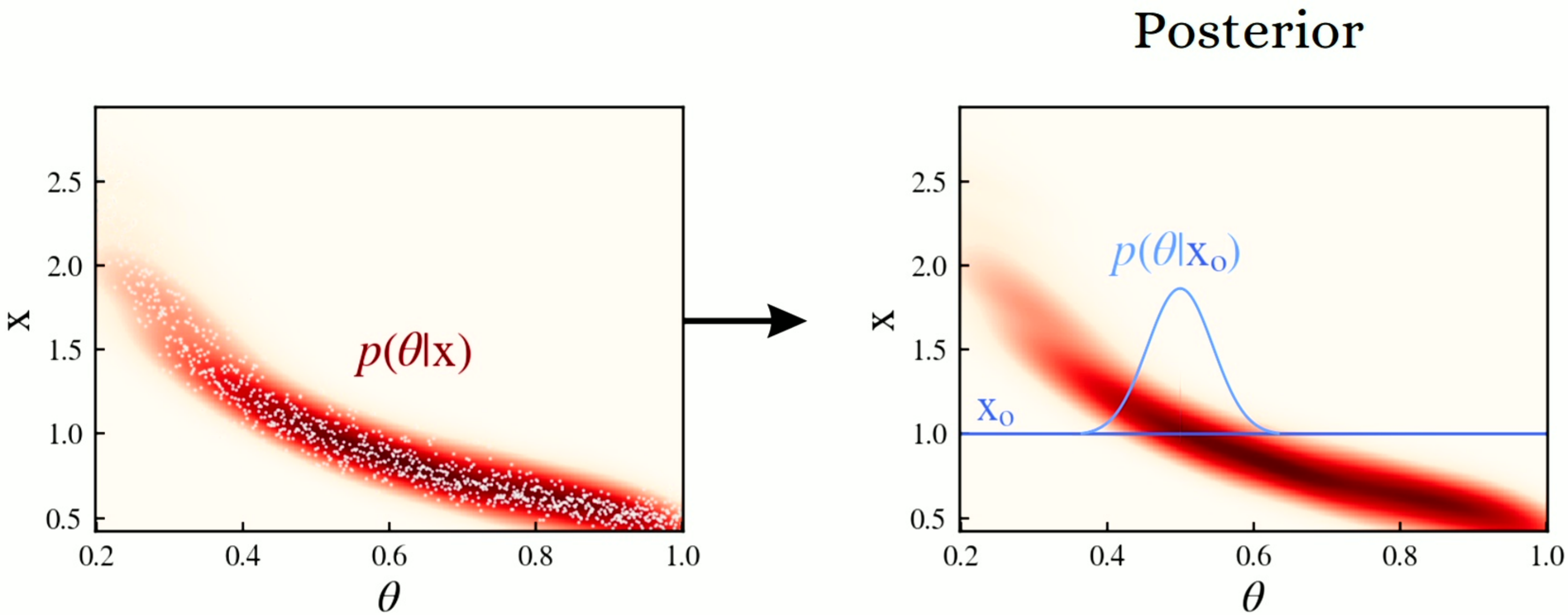
Simulation-based inference

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

Neural Posterior Estimation (NPE)



Simulation-based inference



Neural Density Estimation

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

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 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

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

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

Neural Density Estimation

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

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

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

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

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

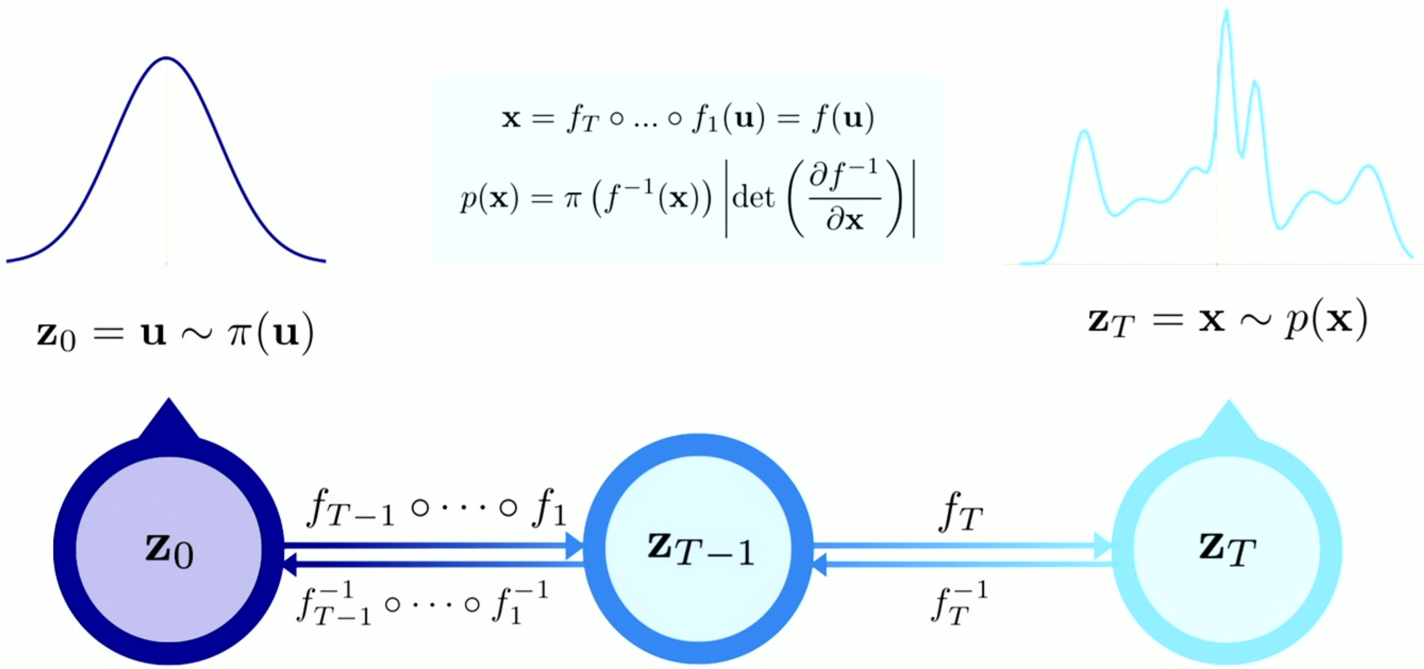
$$= \int d\boldsymbol{\theta}\, d\mathbf{x}\, p(\boldsymbol{\theta}, \mathbf{x}) \log\left(\frac{p(\mathbf{x}|\boldsymbol{\theta})}{q_{\phi}(\mathbf{x}|\boldsymbol{\theta})}\right)$$

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

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

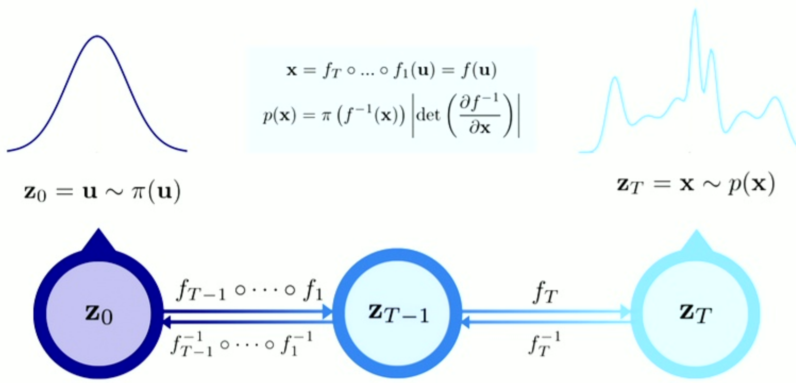
$$\{(\boldsymbol{\theta}_n, \mathbf{x}_n)\}_{n=1}^{N_{\text{sim}}}$$

Normalizing Flows



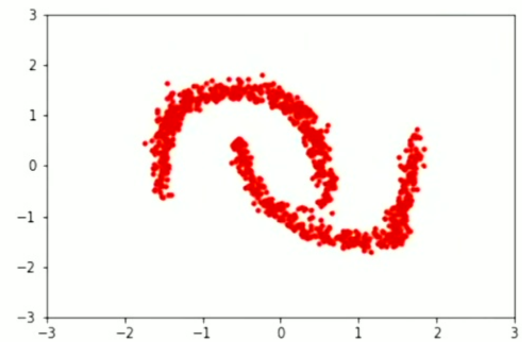
Tucci, Schmidt (2023)

Normalizing Flows



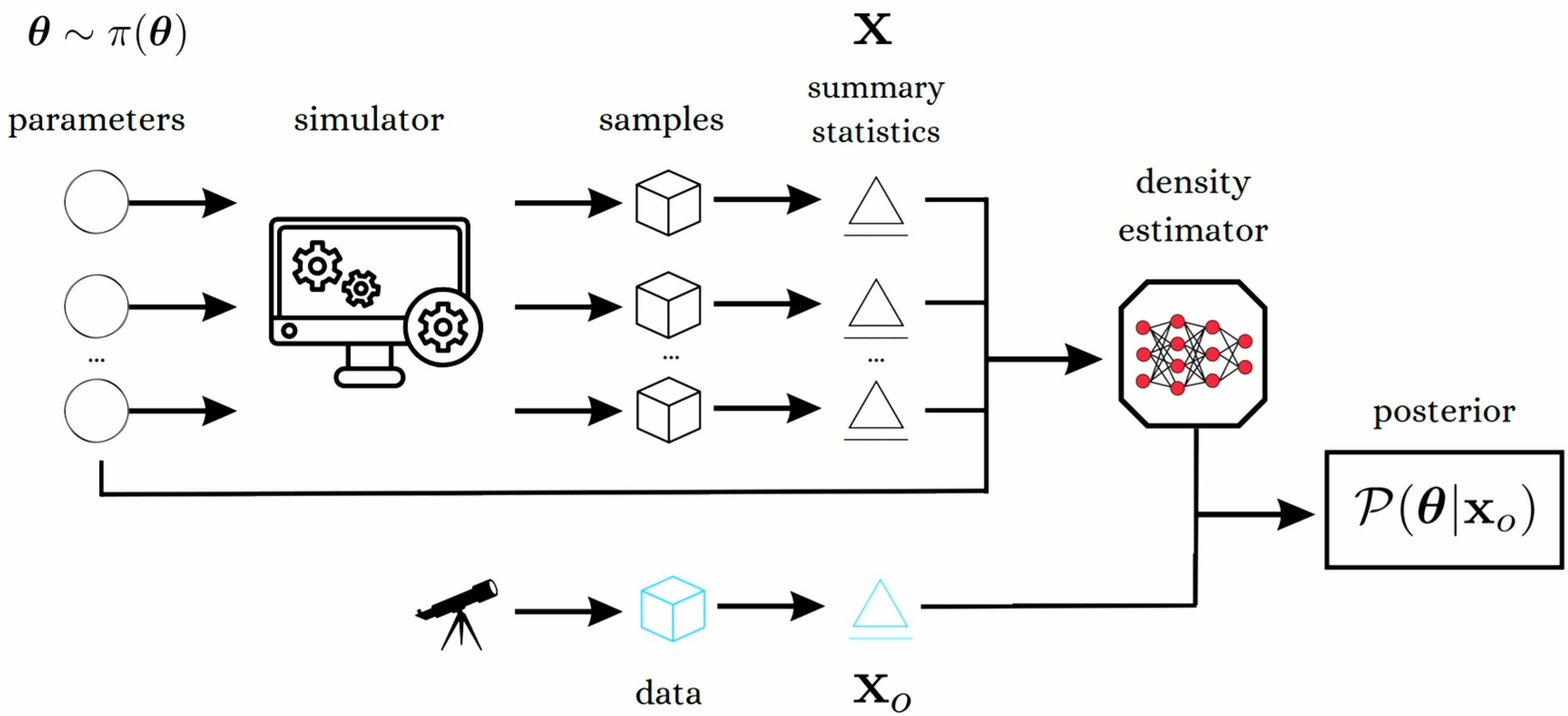
Tucci, Schmidt (2023)

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

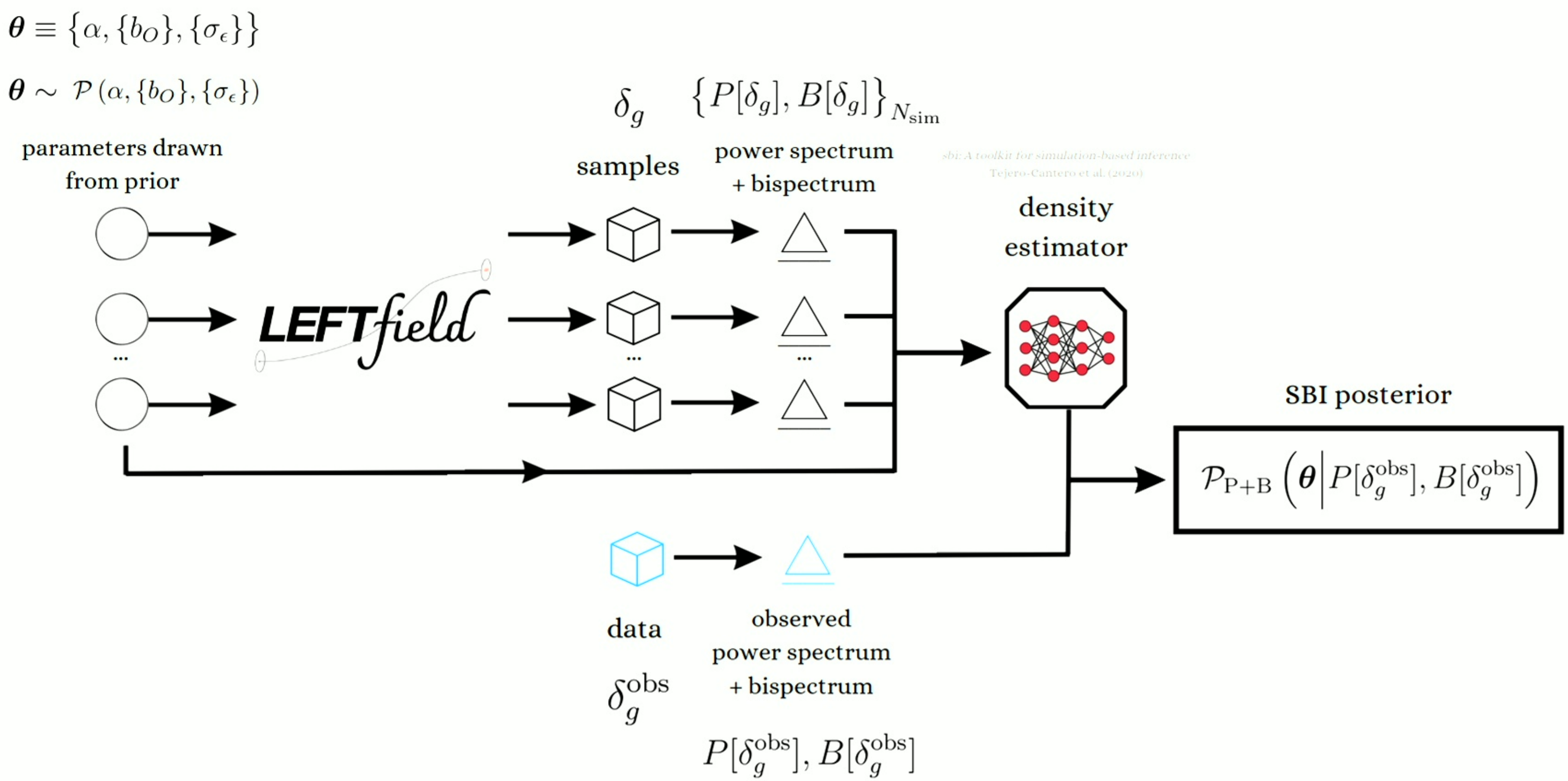


Credits: Miles Cranmer

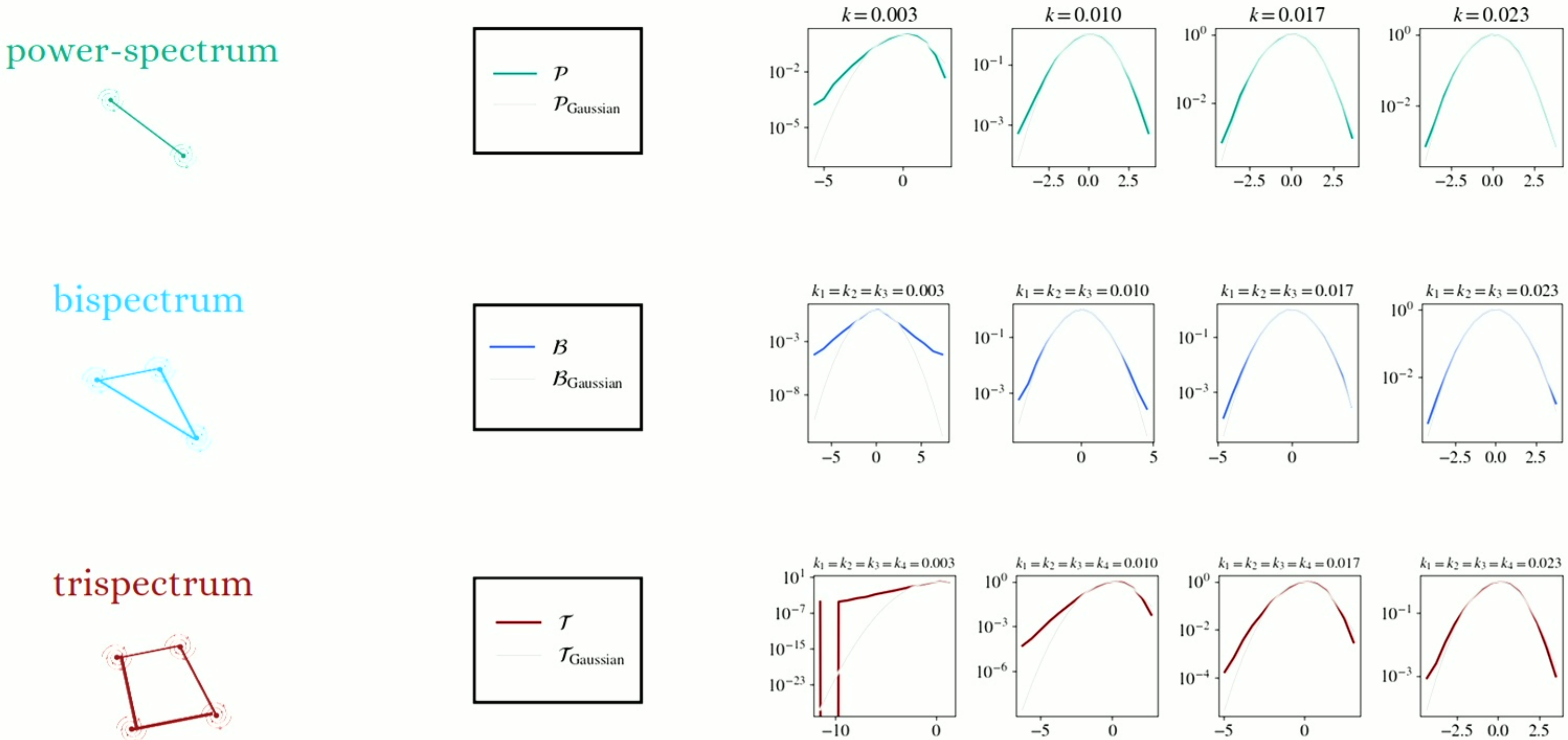
Simulation-based inference



Simulation-based inference for galaxy clustering



On the Gaussianity assumption of the n-point functions

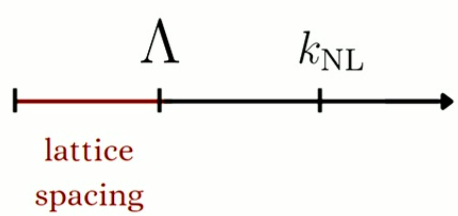


The forward model

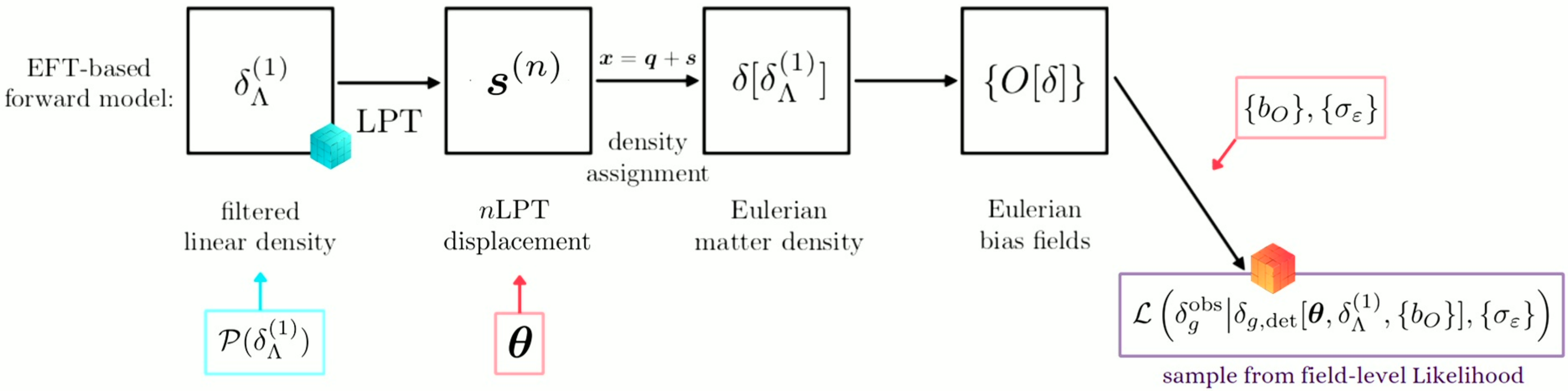
based on the EFTofLSS & the bias expansion

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



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)

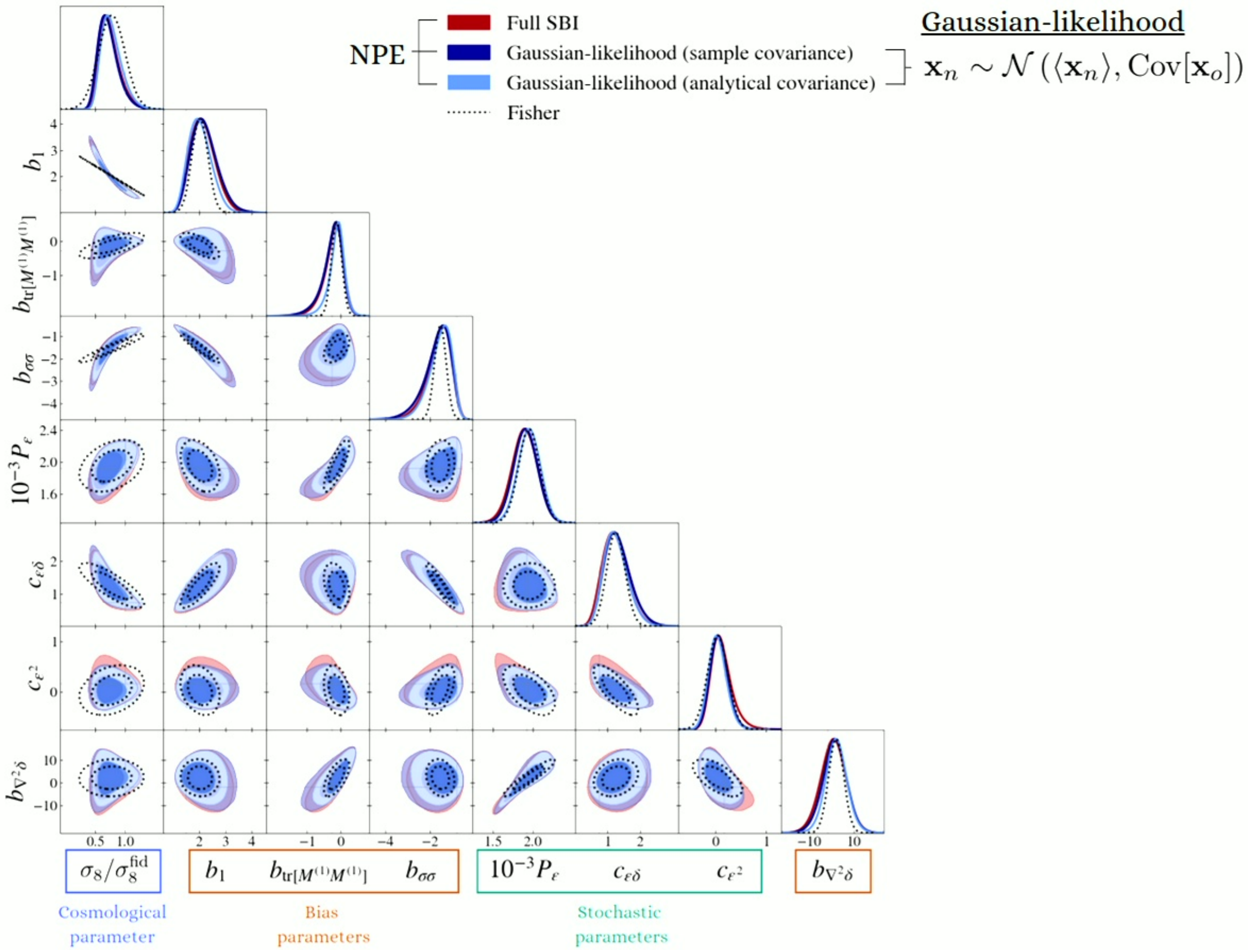
JCAP

Cosmological constraints

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

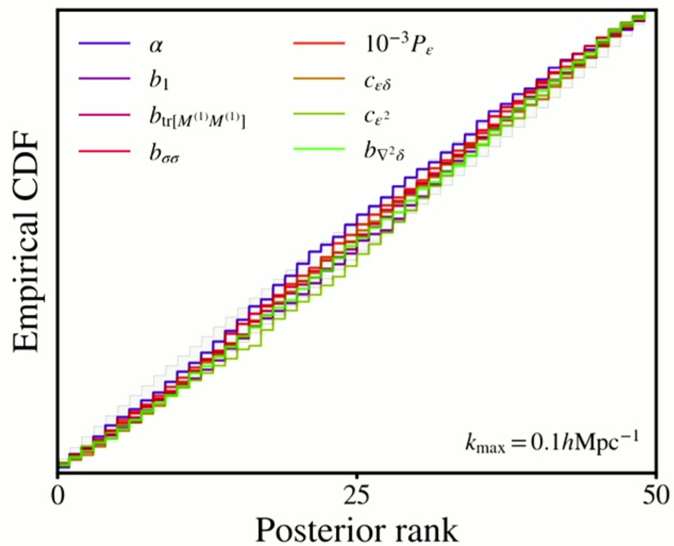
$k_{\rm max} = \Lambda = 0.1 h{\rm Mpc}^{-1}$

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

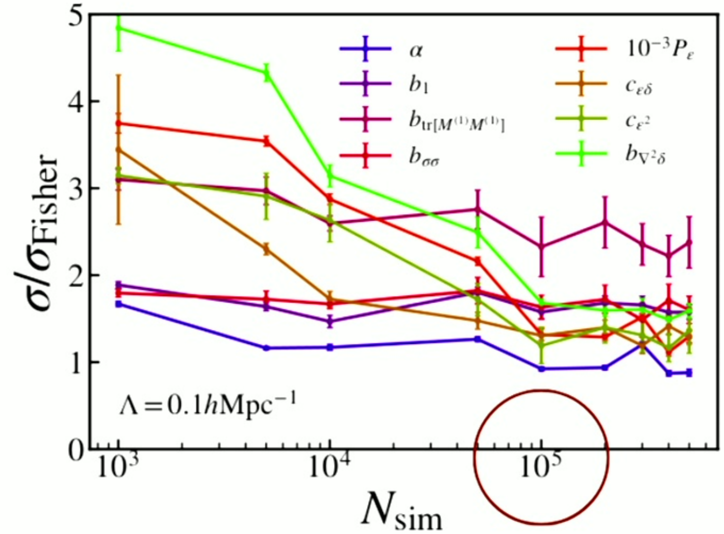


Tests of inference

Simulation-based calibration



Convergence

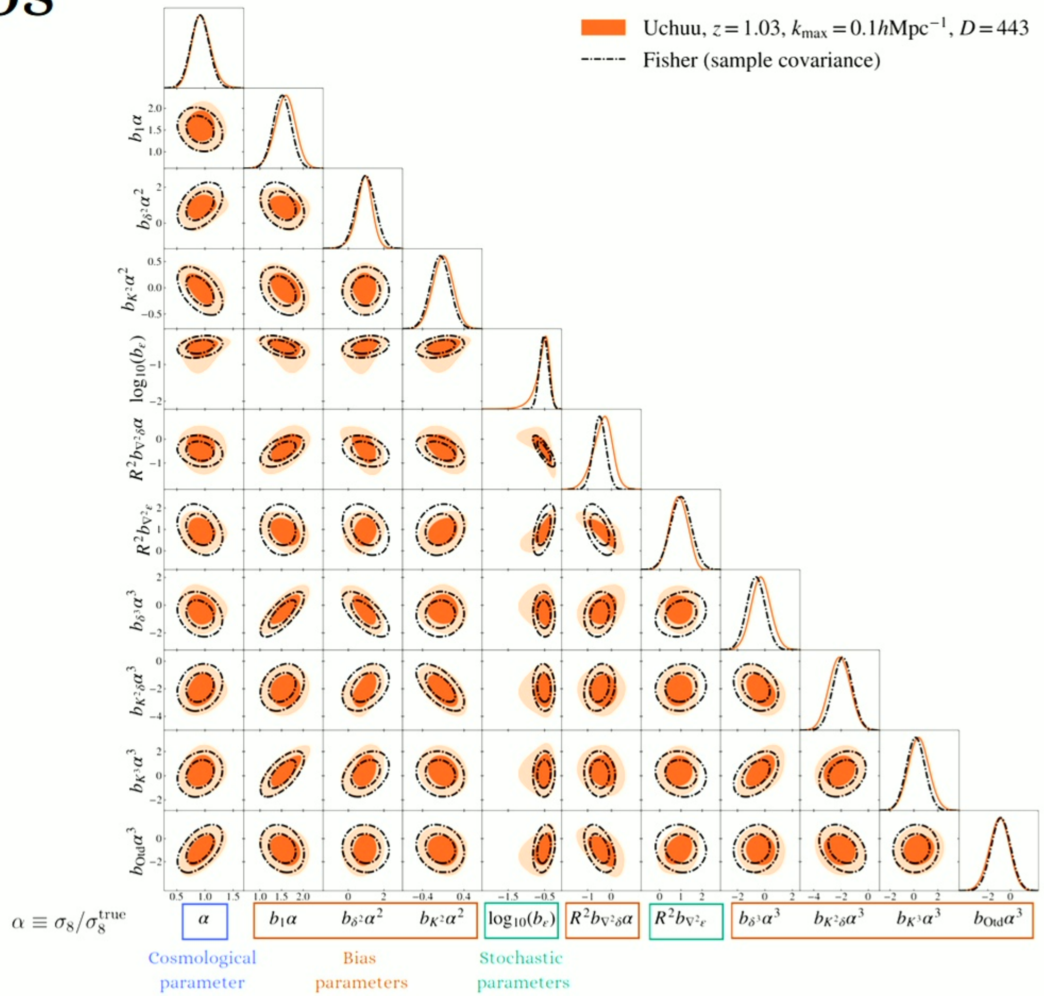


SBI on dark-matter halos

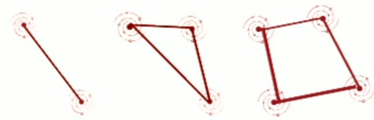
Breaking degeneracy between σ_8 and bias parameters with the galaxy power-spectrum and bispectrum

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

SBI on halos



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

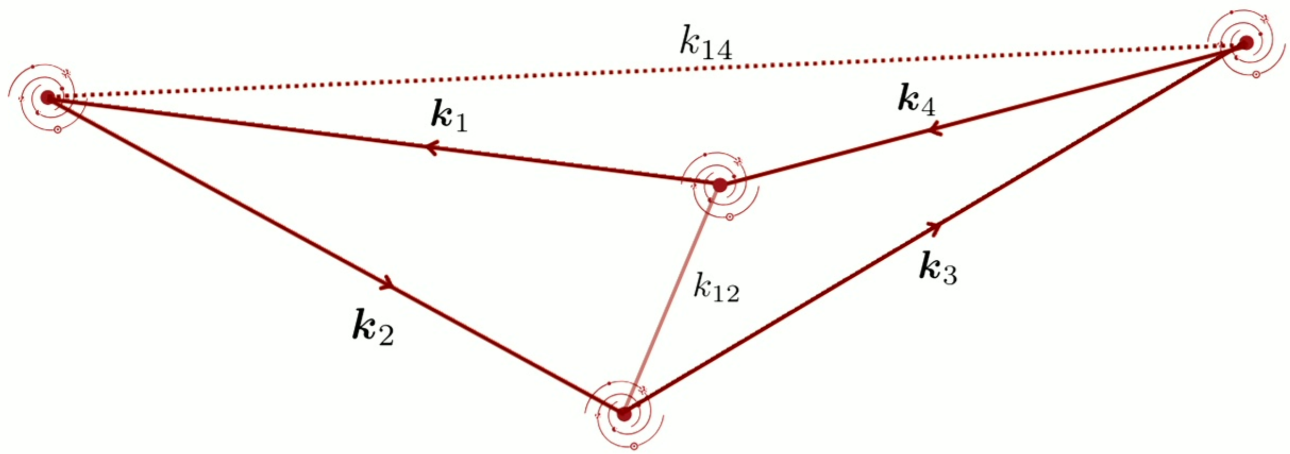


What if we add the galaxy **trispectrum**?

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

Tucci & Schmidt (in prep.)

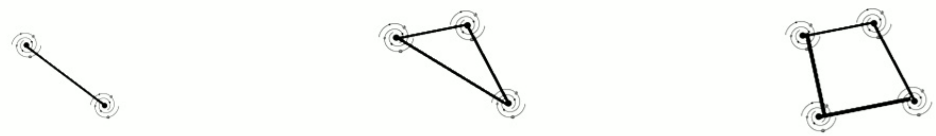
Trispectrum: the estimator



Jung+23, Coulton+23, Goldstein+24

Inferring the cosmological parameters: **challenges**

power-spectrum bispectrum trispectrum **Increasing complexity**



... n

where to stop?
is there a better way?

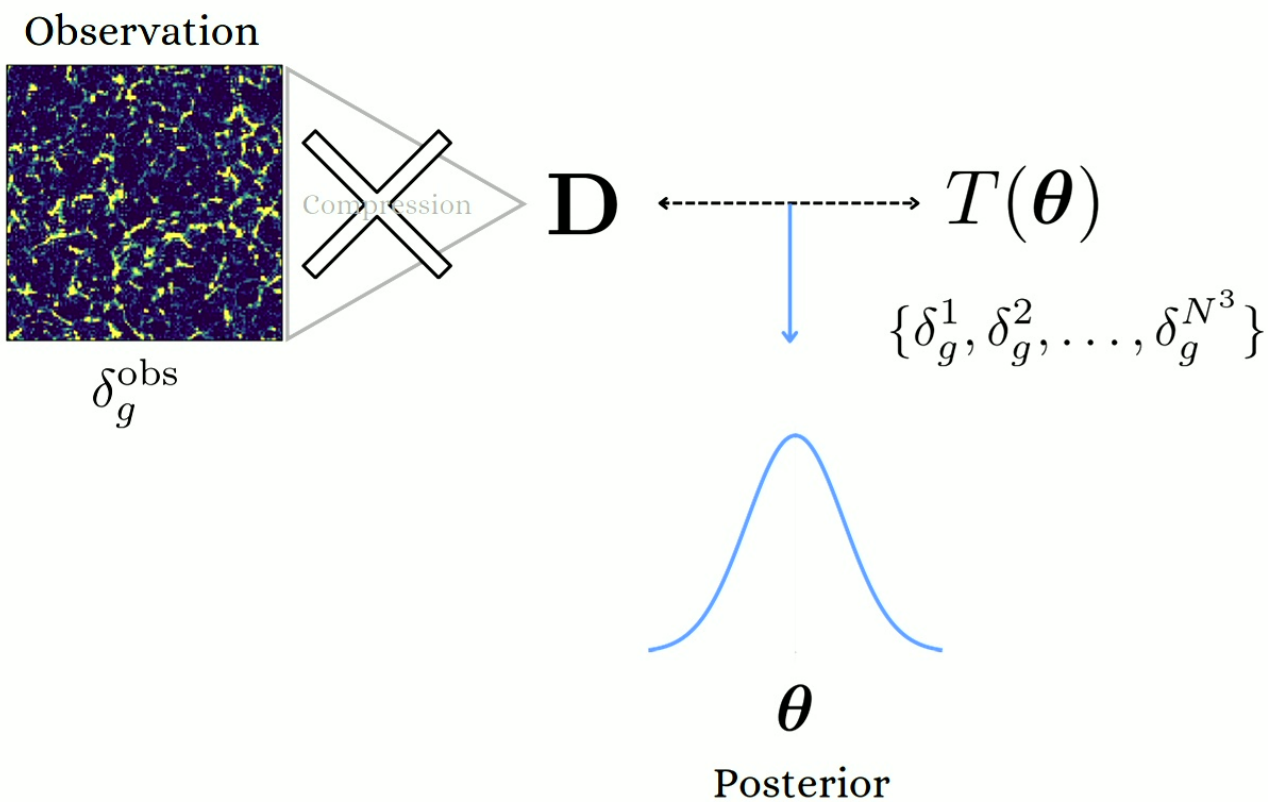
Analytical approximations — Estimation — Modelling

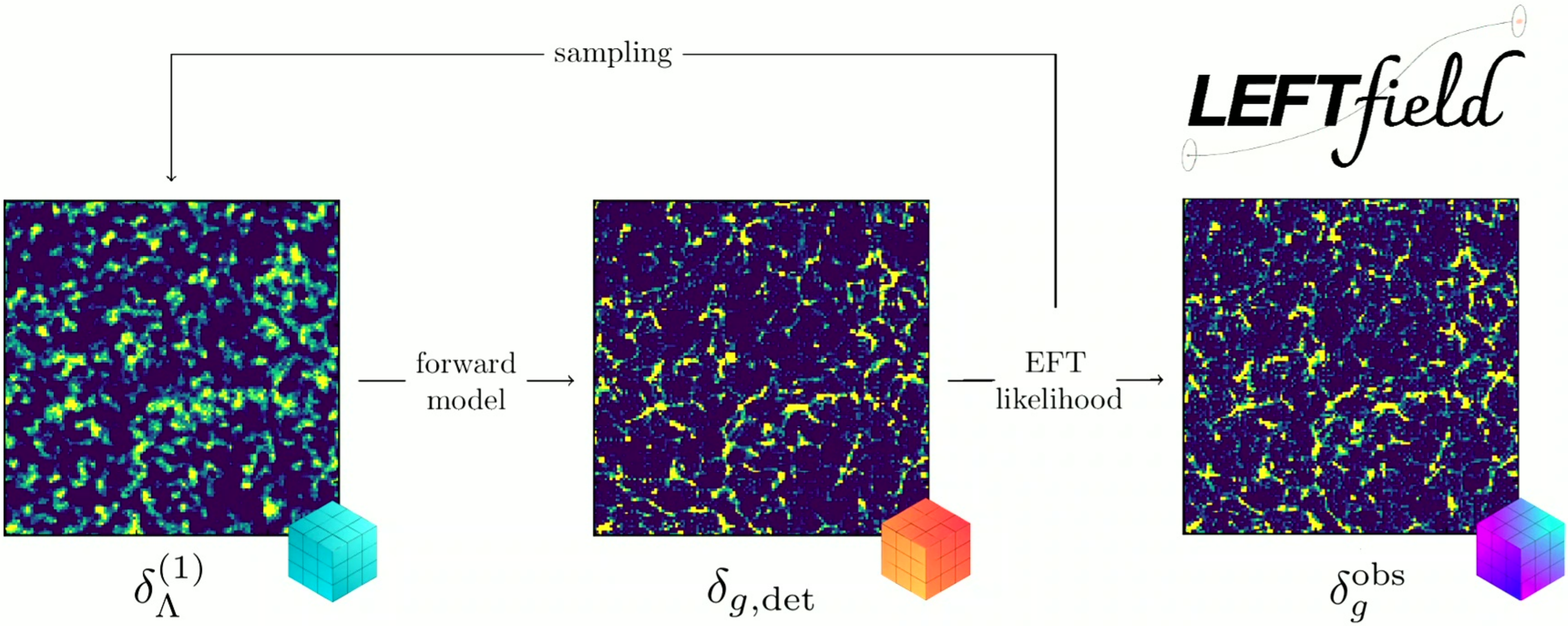
$$-2\ln \mathcal{L}(\mathbf{D}|\boldsymbol{\theta}) = \left(\mathbf{D}-\mathbf{T}(\boldsymbol{\theta})\right)\cdot \mathbf{C}^{-1}\cdot\left(\mathbf{D}-\mathbf{T}(\boldsymbol{\theta})\right)$$

Measurements

SBI

FBI: the main idea



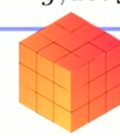


Credits: Julia Stadler

Field level Likelihood

Mode by mode
data and theory
comparison!

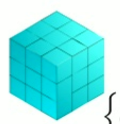
$$\ln \mathcal{L}\left(\delta_g^{\rm obs}\middle|\delta_{g,\rm det}[\boldsymbol{\theta},\delta_\Lambda^{(1)},\{b_O\}],\{\sigma_\varepsilon\}\right) = -\frac{1}{2}\sum_{k<k_{\rm max}}\left[\frac{1}{\sigma_\varepsilon^2(k)}\left|\delta_g^{\rm obs}(\boldsymbol{k})-\delta_{g,\rm det}[\boldsymbol{\theta},\delta_\Lambda^{(1)},\{b_O\}](\boldsymbol{k})\right|^2+\ln[2\pi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



↓ HMC

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

Full posterior
including initial
conditions!



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

3rd order bias expansion

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

Nhat-Minh Nguyen
(IPMU)

Fabian Schmidt
(MPA)

Field-level Bayesian Inference

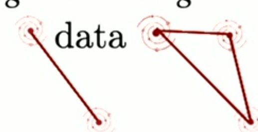
$\boldsymbol{\theta}$ → LEFTfield → δ_g → ($N_{\rm sample}$) → FBI → $\mathscr{P}(\theta \,|\, \delta_g^{\rm obs.})$

cosmology and initial conditions — forward model — model prediction — data

$\boldsymbol{\theta}$ → LEFTfield → δ_g → $P + B$ → ($N_{\rm sim}$) → SBI → $\mathscr{P}(\theta \,|\, P[\delta_g^{\rm obs.}], B[\delta_g^{\rm obs.}])$

summary statistics

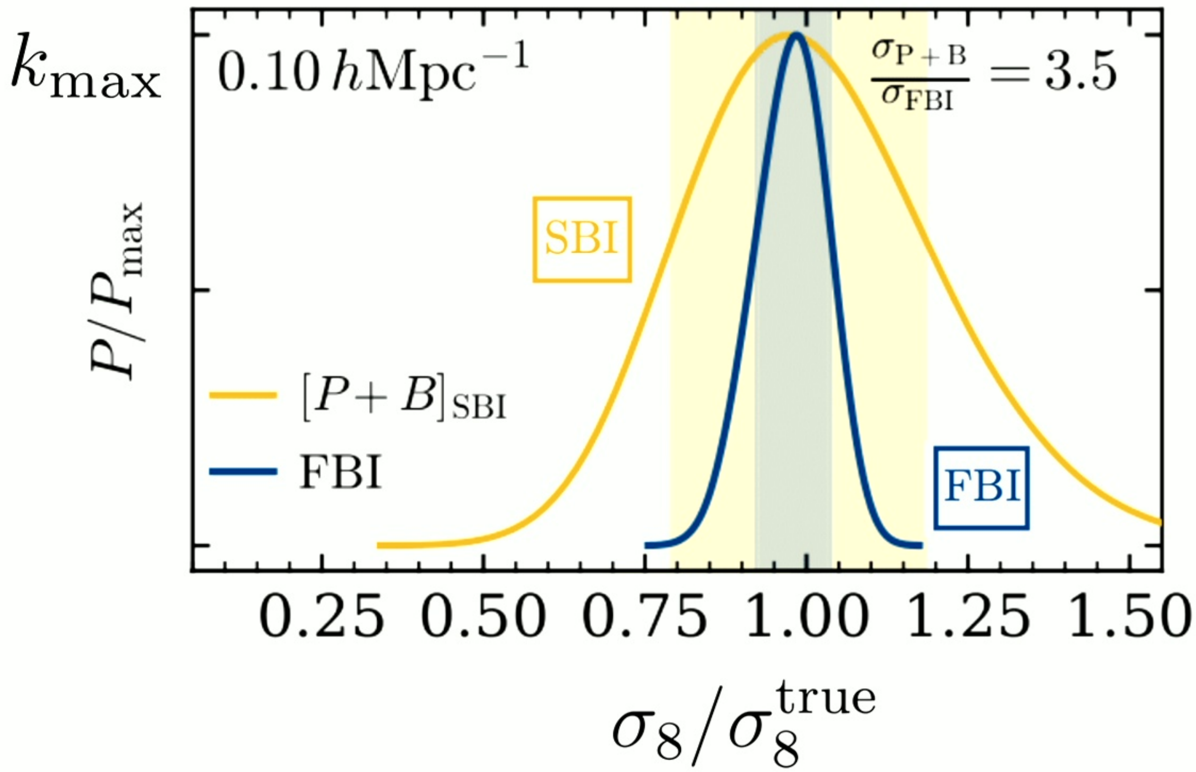
data



Simulation-Based Inference

A lot of reliable information at the field-level!

SNG halos



3.5 improvement factor!

On the Bispectrum stochasticity

Perturbation Theory

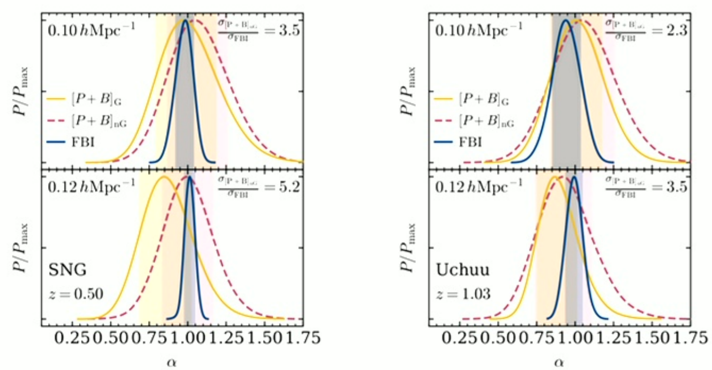
$$\langle \delta_g(k_1)\delta_g(k_2)\delta_g(k_3)\rangle'^{\rm LO}_{\rm stoch} = B_\varepsilon + 2b_1 P_{\varepsilon\varepsilon_\delta}(P_{\rm m}(k_1) + 2\ \text{perm.})$$

Forward Model with Non-Gaussian Noise

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

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

$$\varepsilon \sim \mathcal{N}(0,\sigma_\varepsilon^2)$$

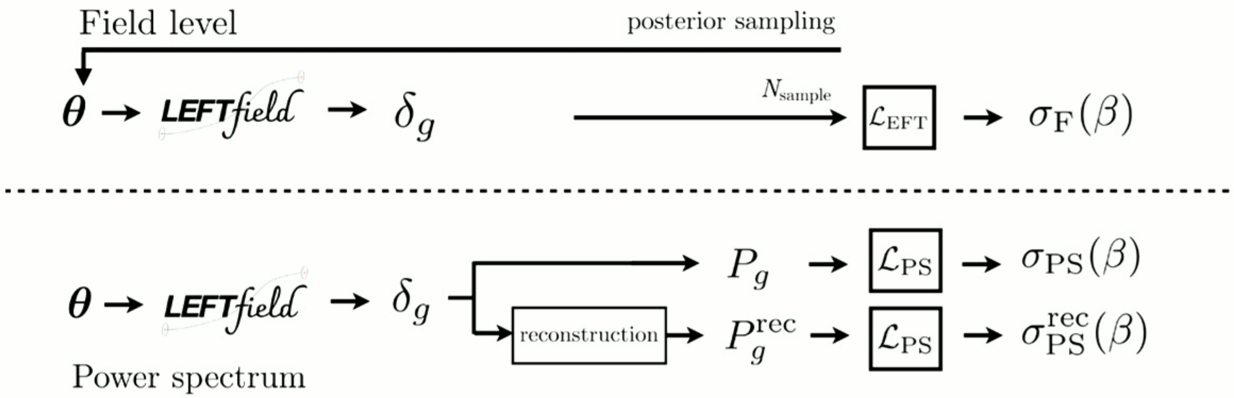


$$O_{\rm det} \in [\delta, \delta^2, K^2, O_{\rm td}, \nabla^2\delta]$$

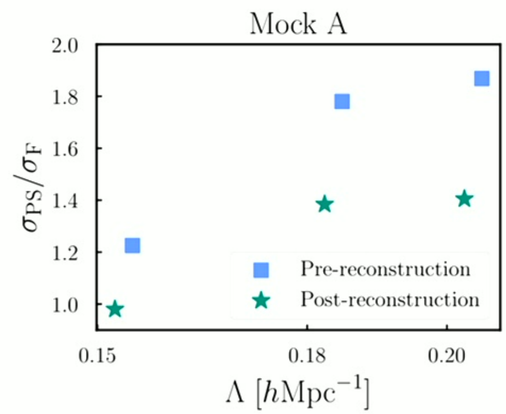
$$O_{\rm stoch} \in [\varepsilon, \varepsilon\delta, \varepsilon^2, \nabla^2\varepsilon]$$

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