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

Nhat-Minh Nguyen[1, 2, *] Fabian Schmidt[3, †] Beatriz Tucci[3, ‡]
Martin Reinecke,[3] and Andrija Kostić[3]

[1]*Leinweber Center for Theoretical Physics, University of Michigan, 450 Church St, Ann Arbor, MI 48109-1040*
[2]*Department of Physics, College of Literature, Science and the Arts,*
University of Michigan, 450 Church St, Ann Arbor, MI 48109-1040
[3]*Max–Planck–Institut für Astrophysik, Karl–Schwarzschild–Straße 1, 85748 Garching, Germany*
(Dated: October 4, 2024)

We present optimal Bayesian field-level cosmological constraints from nonlinear tracers of cosmic large-scale structure, specifically the amplitude σ_8 of linear matter fluctuations inferred from rest-frame simulated dark matter halos in a comoving volume of $8\,(h^{-1}\mathrm{Gpc})^3$. Our constraint on σ_8 is entirely due to nonlinear information, and obtained by explicitly sampling the initial conditions along with tracer bias and noise parameters via a Lagrangian EFT-based forward model, `LEFTfield`. The comparison with a simulation-based inference of the power spectrum and bispectrum—likewise using the `LEFTfield` forward model—shows that, when including precisely the same modes of the same data up to $k_{\mathrm{max}} = 0.10\,h\,\mathrm{Mpc}^{-1}$ ($0.12\,h\,\mathrm{Mpc}^{-1}$), the field-level approach yields a factor of 3.5 (5.2) improvement on the σ_8 constraint, going from 20.0% to 5.7% (17.0% to 3.3%). This study provides direct insights into cosmological information encoded in galaxy clustering beyond low-order n-point functions.

Introduction.—Cosmological surveys of large-scale structure (LSS) have significantly advanced from measuring angular clustering of galaxies identified on photographic plates to mapping three-dimensional clustering with spectroscopy and optical fibers positioned by robots. Meanwhile, statistical methods to analyze galaxy clustering and other biased tracers of LSS still largely rely on modeling two- and three-point correlation functions. This leaves open the question:

How much cosmological information can be robustly extracted from galaxies and LSS tracers?

In this Letter, we address this question by comparing constraints on the amplitude of linear matter fluctuations, σ_8 [Eq. (1)] obtained with (1) the full field-level statistics, i.e. the entire three-dimensional tracer field, and (2) the combination of power spectrum plus bispectrum (P+B) summary statistics. In both setups, we analyze the same data using the same data model based on the Effective Field Theory (EFT) of LSS [1–3].

Constraints on σ_8, derived from LSS surveys and cosmic microwave background (CMB) experiments, quantify the linear growth of structure since the primordial universe, which in turn probes Dark Energy and modifications to General Relativity. Recently, σ_8, or the parameter combination $S_8 = \sigma_8\sqrt{\Omega_m/0.3}$, has been under investigation since tensions arise between CMB constraints [4–7] and galaxy weak lensing constraints [8–12], within the standard model ΛCDM. While CMB measures matter fluctuations in the early universe on the largest observable scales, galaxy weak lensing quantifies fluctuations in the late universe on smaller scales, making it hard to disentangle the physical origin of the tension: Is it a tension between physics of the early and late universe [13–15] or between that of large and small scales [16–18], or both? Separate σ_8 constraints from galaxy clustering—which alone probes growth in the late universe on large-to-intermediate scales—therefore provide an additional crucial data point to dissect this tension.

σ_8 constraint from galaxy clustering.—The initial matter fluctuations out of which LSS formed closely follow a zero-mean multivariate Gaussian distribution $\mathcal{N}(0, \mathsf{C})$ [19]. At late times, the linear growth of matter fluctuations $\delta^{(1)}(\boldsymbol{k}, z)$ in Fourier space can be decomposed into the scale-dependent transfer function $T(k)$ and the redshift-dependent growth function $D_+^{(1)}(z)$, such that $\delta^{(1)}(\boldsymbol{k}, z) = (2/5)(k^2/\Omega_m H_0^2)D_+^{(1)}(z)T(k)\mathcal{R}(\boldsymbol{k})$ for each Fourier mode $\boldsymbol{k}$, where $\mathcal{R}(\boldsymbol{k})$ is the primordial curvature perturbation, while Ω_m and H_0 are the matter density parameter and the Hubble constant, respectively. A proxy of growth up to today can be defined through the variance of linear matter fluctuations at $z = 0$:

$$\sigma_8^2 \equiv \left\langle \left(\delta_R^{(1)}(\boldsymbol{x})\right)^2 \right\rangle = \int_0^\infty \frac{dk\, k^2}{2\pi^2} P_{\mathrm{L}}(k, z=0) W_R^2(k), \tag{1}$$

where P_{L} is the linear power spectrum and W_R is a spherical window function of radius $R = 8\,h^{-1}\mathrm{Mpc}$.

Constraining σ_8 from galaxy clustering is complicated by the fact that galaxies are discrete, biased tracers of the underlying matter field. At leading order in perturbations, the observed galaxy power spectrum is related to the linear matter power spectrum through $P_{\delta_g^{\mathrm{obs.}}}(k, z) = b_\delta^2(z)P_{\mathrm{L}}(k, z) + P_\epsilon(z)$ where b_δ is the linear galaxy bias coefficient and P_ϵ is the amplitude of the galaxy stochastic contribution, both of which encapsulate the a priori unknown physics of galaxy formation. At this order, there is thus a perfect degeneracy between b_δ and σ_8, which can only be broken by considering nonlinear clustering

* nguyenmn@umich.edu
† fabians@MPA-Garching.MPG.DE
‡ tucci@MPA-Garching.MPG.DE

or redshift-space distortions (RSD). However, at leading order, RSD exhibits a further degeneracy in that it only constrains the product $f\sigma_8$, with the growth rate $f = d\ln D_+/d\ln a$.

How does nonlinear clustering information help with the $\sigma_8 - b_\delta$ degeneracy? Beyond linear order, both the matter δ and galaxy density δ_g contain advection contributions involving the Lagrangian displacement field. Einstein's equivalence principle ensures that matter and galaxies co-move on large scales [20]. The latter implies that the displacement is the same for galaxies and matter, which means that the advection contribution to δ_g can be uniquely predicted. This direct consequence of the equivalence principle therefore breaks the degeneracy between b_δ and σ_8.

This work provides the first proof that (robust) constraints on σ_8 can be extracted from nonlinear clustering of dark-matter halos in N-body simulations—directly at the field level. Field-level inference explicitly samples the full posterior of $[\sigma_8, \delta^{(1)}]$ given the data, and is hence guaranteed to be optimal within the assumed data model [21].

Our analysis extends previous works in several ways. Firstly, [22, 23] demonstrated the ability of the EFTofLSS framework to predict clustering in N-body simulations at the field level to percent-level precisions, conditioned on the true initial conditions $\hat{s}_{\text{true}}$ of the simulations. The initial conditions $\hat{s}$ of the universe, however, are not direct observables. Here, we jointly infer and marginalize over the unknown initial conditions. Thus, our σ_8 constraints are much more representative of those in a real-world scenario.

Secondly, in this work, we study N-body halos—representative of the true observable in real galaxy surveys—in contrast to previous studies that analyzed either mock data drawn from the same forward model [24–26] or idealized, unbiased tracers, namely dark-matter particles [27–29]. Since we cannot (yet) simulate realistic galaxy catalogs in a cosmological volume, model misspecification is inherent in any data model and analysis, and the robustness of the results against such misspecification must be validated. In fact, this is likely to be the key challenge for field-level inference [30, 31]. Recovering unbiased constraints on σ_8 in this work thus is a stringent test of the robustness of the EFTofLSS field-level forward model and inference framework. Given that our forward model is agnostic to details of the tracer—and based only on the equivalence principle—we expect that our results generalize to actual galaxies as well (see [32, 33] for field-level results on simulated galaxies).

Lastly, this is the first time a consistent comparison between field-level and summary statistics has been conducted on the basis of exactly the same galaxy clustering data and model.

Data.—Our primary halo sample SNG consists of main halos in the $\log_{10} M_{200m} = 12.5 - 14.0\,h^{-1}M_\odot$ mass range, identified with the ROCKSTAR halo finder [34] at redshift $z = 0.50$ in an N-body, gravity-only sim-

ulation (mean comoving number density $\bar{n} = 1.3 \cdot 10^{-3}(\,h^{-1}\text{Mpc})^{-3}$). The simulation assumes a ΛCDM cosmology with $\sigma_{8,\text{true}} = 0.850$, encompasses a comoving volume $L^3 = (2000\,h^{-1}\text{Mpc})^3$ and contains $N_{\text{particle}} = 1536^3$ particles of mass $M_{\text{particle}} = 1.8 \times 10^{11}\,h^{-1}M_\odot$ [23].

We additionally analyze the Uchuu halo sample, consisting of main halos in the $\log_{10} M_{200m} = 12.0 - 13.5\,h^{-1}M_\odot$ mass range, identified with ROCKSTAR at redshift $z = 1.03$ ($\bar{n} = 3.6 \cdot 10^{-3}(\,h^{-1}\text{Mpc})^{-3}$) in the Uchuu simulation [35], which assumes $\sigma_{8,\text{true}} = 0.816$. This simulation spans the same volume $L^3 = (2000\,h^{-1}\text{Mpc})^3$ while offering a much higher mass resolution, $N_{\text{particle}} = 12800^3$ particles of mass $M_{\text{particle}} = 3.27 \times 10^8\,h^{-1}M_\odot$.

Field-level EFT of biased tracers.—The EFTofLSS provides a perturbative framework within which the galaxy density field δ_g can be systematically expanded, order by order, in perturbations as

$$\delta_g(\boldsymbol{k}, z) = \sum_O b_O(z)O(\boldsymbol{k}, z) + \epsilon(\boldsymbol{k}, z). \qquad (2)$$

The first term on the r.h.s. of Eq. (2) includes deterministic contributions, such as $b_\delta(z)\delta(\boldsymbol{k}, z)$, and describes the biased nature of galaxies as tracers of the underlying matter field.

At each given order, there is a finite number of galaxy bias operators O, each associated with a coefficient b_O. The operators O are constructed out of local gravitational observables, obeying the equivalence principle. Further details on the galaxy bias expansion, operators and coefficients can be found in the Supplementary Material.

The second term on the r.h.s. of Eq. (2) is the stochastic contribution encapsulating the random nature of small-scale fluctuations and the discrete nature of galaxies. These are described by the noise field ϵ, which to leading order is Gaussian with RMS

$$\sigma_\epsilon(k) = \sigma_{\epsilon,0}\left[1 + \sigma_{\epsilon,k^2}k^2\right], \qquad (3)$$

where the subleading contribution $\propto k^2$ captures the nonlocality of galaxy formation.

Field-level forward model.—We forward model the bias fields O in Eq. (2) starting from Gaussian initial conditions parametrized via $\hat{s} \sim \mathcal{N}(0, 1)$, discretized on a grid of size N_{grid}, hence

$$\delta^{(1)}(\boldsymbol{k}, z) = \alpha\left[\frac{N_{\text{grid}}^3}{L^3}P_{\text{L}}(k, z)\right]^{1/2}\hat{s}(\boldsymbol{k}), \qquad (4)$$

where L is the simulation box side length, and $\alpha = \sigma_8/\sigma_{8,\text{true}}$ is the amplitude rescaling parameter. $\alpha = 1$ indicates an unbiased inference for either halo sample.

Our forward model employs the Lagrangian, EFT-based forward model of cosmological density fields, LEFTfield [22]. Following the EFT principle, LEFTfield
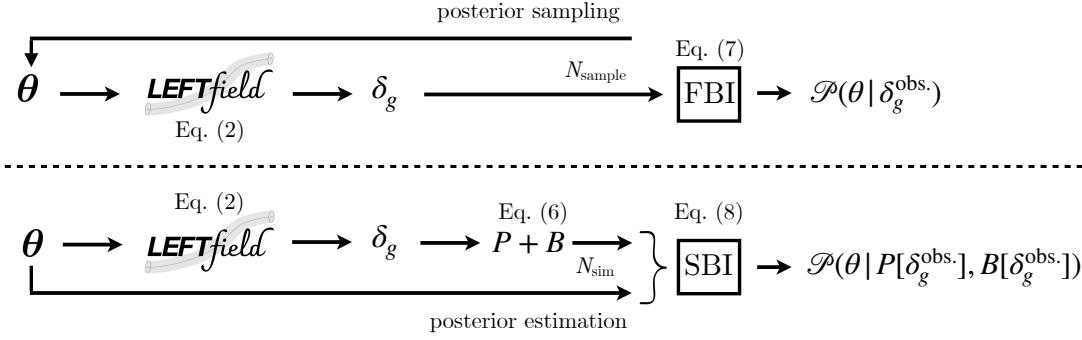
FIG. 1. Diagram of the two inference methods, FBI and SBI P+B. Both pipelines share the same forward model `LEFTfield`.

evolves all cosmological (plus auxiliary) fields up to a finite EFT cutoff Λ. We choose Λ to be a sharp-k filter that strictly filters out all Fourier modes above the cutoff [3, 36, 37]. Specifically, we implement a cubic sharp-k filter via a Fourier grid reduction [23]. Crucially, `LEFTfield` computes both $O = O(\hat{s})$ and $\partial O/\partial\hat{s}$, the latter of which proves useful for gradient-based sampling and field-level inference. We refer to Supplementary Material and [22, 26, 37, 38] for `LEFTfield` implementation and validation. Here, the new developments with respect to [22, 26, 37, 38] are: (1) a third-order model for galaxy bias, improving accuracy relative to the previous (second-order) treatment; (2) a non-uniform Fast Fourier Transform (NUFFT [39]) for grid assignment, enhancing numerical convergence and efficiency relative to previous assignment schemes; and (3) a change to $k_{\max} = \Lambda/1.2$, effectively reducing the analysis cutoff scale $k_{\max}$ relative to the initial conditions cutoff Λ, hence mitigating higher-derivative contributions.

Inference method I: FBI with explicit likelihood.—In the field-level Bayesian inference (FBI) pipeline, we evaluate and sample from an explicit field-level likelihood $\mathcal{L}_{\mathrm{FBI}}^{\mathrm{expl.}}$, depicted in the top row of Fig. 1.

Following [40], our fiducial analyses assumes Gaussianity of galaxy stochasticity and analytically marginalizes over ϵ. This leads to a Gaussian likelihood of the following form for an observed and filtered galaxy field $\delta_g^{\mathrm{obs.}}$ [36, 41]:

$$\mathcal{L}_{\mathrm{FBI}}^{\mathrm{expl.}}\left(\delta_g^{\mathrm{obs.}}\Big|\hat{s}, \alpha, \{b_O\}, \{\sigma_\epsilon\}\right) = -\frac{1}{2}\sum_{\substack{\mathbf{k}>0}}^{|\mathbf{k}|<k_{\max}}$$

$$\left[\ln 2\pi\sigma_\epsilon^2(k) + \frac{1}{\sigma_\epsilon^2(k)}\Big|\delta_g^{\mathrm{obs.}}(\mathbf{k}) - \sum_O b_O O[\alpha, \hat{s}](\mathbf{k})\Big|^2\right].$$

(5)

The $\sum_{\mathbf{k}>0}^{|\mathbf{k}|<k_{\max}}$ amounts to a spherical sharp-k filter which only includes Fourier modes $\mathbf{k}$ up to $k_{\max}$, the cutoff scale of our analyses while also excludes the $\mathbf{k} = 0$ mode. We expand $\sum_O b_O O$ to third order in the galaxy bias operators O and further analytically marginalize over the bias coefficients $\{b_O\}$ assuming weakly informative Gaussian priors (see Supplementary Material).

The final explicit FBI parameter space consists of $[\hat{s}, \alpha, \{\sigma_\epsilon\}]$. The element $\hat{s}$ is a three-dimensional grid of size $[N_{\mathrm{grid}} \times N_{\mathrm{grid}} \times N_{\mathrm{grid}}]$ containing $N_{\mathrm{mode}} = N_{\mathrm{grid}}^3$ modes of initial density fluctuations. To explore this high-dimensional posterior, following [26, 42], we employ two MCMC sampling methods: Hamiltonian Monte Carlo (HMC) [43] for $\hat{s}$—leveraging the differentiability of `LEFTfield` forward models—and slice sampling [44] for $[\alpha, \{\sigma_\epsilon\}]$.

Inference method II: SBI P+B with implicit likelihood.—Implicit-likelihood or simulation-based inference (SBI) directly learns the posterior from simulated training data without assuming any analytical form for the likelihood of the data vector [45]. Our SBI P+B pipeline is depicted in the bottom row of Fig. 1, where we closely follow the procedure detailed in [38]. We first draw the parameters $\theta \equiv [\alpha, \{b_O\}, \{\sigma_\epsilon\}]$ from their priors and simulate the galaxy fields δ_g via Eq. (2) Eq. (5) with `LEFTfield`. We then measure the power spectrum P and bispectrum B on each simulated data realization,

$$\langle\delta_g(\mathbf{k})\delta_g(\mathbf{k}')\rangle = P(k)(2\pi)^3\delta_D(\mathbf{k} + \mathbf{k}'), \tag{6a}$$

$$\langle\delta_g(\mathbf{k}_1)\delta_g(\mathbf{k}_2)\delta_g(\mathbf{k}_3)\rangle =$$
$$B(k_1, k_2, k_3)(2\pi)^3\delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3), \tag{6b}$$

following [46] (see also Eqs. (2.15–2.17) of [38]). The SBI P+B data vector contains $N_{\mathrm{bin}} + N_{\mathrm{triangle}}$ elements up to the same $k_{\max}$ used in the FBI analysis, with N_{bin} linear k-bins for the power spectrum and N_{triangle} triangle k-configurations for the bispectrum. We choose a k bin width of $\Delta k = 2k_f$, where $k_f \equiv 2\pi L^{-1}$ is the fundamental frequency. The N_{sim} samples, drawn from the joint distribution $\{\theta, P[\delta_g(\theta)], B[\delta_g(\theta)]\}$ this way, form the SBI training set. We use neural posterior estimation (NPE) [47] with masked autoregressive flows [48] from the `sbi` package [49] (see Supplementary Material).

After training, we sample the estimated posterior $\mathcal{P}_{\mathrm{P+B}}$, conditioned on the power spectrum plus bispectrum measured on the "observed" data $\left[P[\delta_g^{\mathrm{obs.}}], B[\delta_g^{\mathrm{obs.}}]\right]$, [Eq. (8)]. We employ simulation-based calibration (SBC) [50] and convergence tests to validate the SBI posteriors (see Supplementary Material). We note that the forward model employed here

assumes Gaussian noise [Eq. (5)]. Thus our bispectrum model does not contain a contribution from non-Gaussian (skewed) noise, or from a density-dependent noise variance. In the Supplementary Material, we compare our fiducial SBI P+B analysis with a variant that includes both additional stochastic contributions, but employs a restricted bias parameter set (see Supplementary Material). This variant matches current standard P+B analyses [51, 52]. We find broad consistency between both SBI P+B analyses.

Inference summary:—Explicitly, our target posteriors are

$$\mathcal{P}_{\mathrm{FBI}}\left(\alpha, \{b_O\}, \{\sigma_\epsilon\} \middle| \delta_g^{\mathrm{obs.}}\right) \propto \int \mathcal{D}\hat{s}\, \mathcal{P}\left(\hat{s}\right) \mathcal{L}_{\mathrm{FBI}}^{\mathrm{expl.}}\left(\delta_g^{\mathrm{obs.}} \middle| \hat{s}, \alpha, \{b_O\}, \{\sigma_\epsilon\}\right) \mathcal{P}\left(\alpha, \{b_O\}, \{\sigma_\epsilon\}\right), \quad (7)$$

$$\mathcal{P}_{\mathrm{P+B}}\left(\alpha, \{b_O\}, \{\sigma_\epsilon\} \middle| P[\delta_g^{\mathrm{obs.}}], B[\delta_g^{\mathrm{obs.}}]\right) \propto \mathcal{L}_{\mathrm{P+B}}^{\mathrm{impl.}}\left(P[\delta_g^{\mathrm{obs.}}], B[\delta_g^{\mathrm{obs.}}] \middle| \alpha, \{b_O\}, \{\sigma_\epsilon\}\right) \mathcal{P}\left(\alpha, \{b_O\}, \{\sigma_\epsilon\}\right), \quad (8)$$

for FBI [Eq. (7)] and SBI P+B [Eq. (8)], where $\{b_O\}$ consists of all bias parameters up to third order. We explicitly list the bias parameters $\{b_O\}$ and the priors $\mathcal{P}\left(\alpha, \{b_O\}, \{\sigma_\epsilon\}\right)$ in the Supplementary Material.

Results.—Our main results are shown in Fig. 2, where we compare α posteriors between FBI and SBI P+B. All analyses recover the ground-truth $\alpha = 1$ within 68%CL. Specifically, at $k_{\max} = 0.1\,(0.12)\,h\,\mathrm{Mpc}^{-1}$, FBI analyses constrain $\alpha = 0.976 \pm 0.056$ ($\alpha = 1.013 \pm 0.033$), a 5.9% (3.6%) constraint on α. This corresponds to a factor of 3.5 (5.2) improvement over the SBI P+B constraints, which are $\alpha = 1.014 \pm 0.200$ ($\alpha = 0.872 \pm 0.170$). An increase in the improvement of field-level constraints over low-order summary statistics with the analysis cut-off scale $k_{\max}$ is expected, since the information gain is due to the nonlinearities in the forward model, whose significance increases with wavenumber.

Both FBI and SBI P+B results show consistent posteriors between the two $k_{\max}$ values for each inference method. Their results are further consistent with each other within 0.2-σ (0.8-σ). The consistency between the two analyses (at both $k_{\max}$) stems from their common forward model, `LEFTfield`. The level of consistency further underlines the precision of `LEFTfield` on these scales.

To verify whether the above conclusions generalize, we analyze an external sample from the publicly available `Uchuu` simulation. Fig. 3 shows that the answer is affirmative: the FBI analysis yields a factor of 1.9 (2.5) improvement over that obtained with SBI P+B. Specifically, the FBI constraints are $\alpha = 0.941 \pm 0.090$ ($\alpha = 0.993 \pm 0.053$) versus the $\alpha = 1.018 \pm 0.168$ ($\alpha = 0.900 \pm 0.136$) constraints by SBI P+B, at $k_{\max} = 0.1\,(0.12)\,h\,\mathrm{Mpc}^{-1}$, in excellent agreement within 0.4-σ (0.6-σ). The improvement factors for the Uchuu sample are slightly lower than for the SNG sample. This is likely because displacement contributions to higher n-point functions are less significant at higher redshifts (due to a smaller growth factor). As these contributions provide additional information for improving σ_8 inference at the field level, their reduced importance at higher redshifts could explain the reduced improvements.
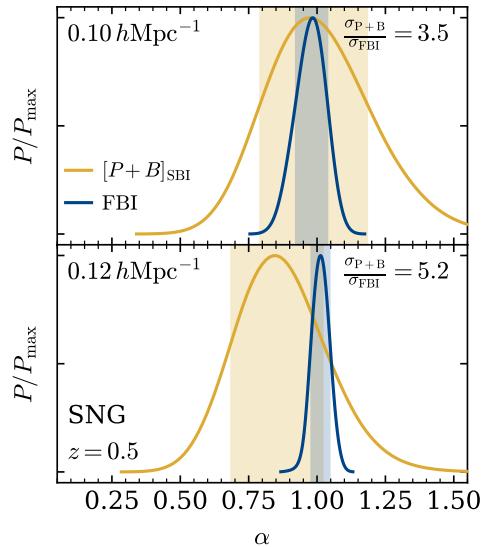

FIG. 2. Constraints on $\alpha = \sigma_8/\sigma_{8,\mathrm{true}}$, from the `SNG` sample (see text), at $k_{\max} = [0.10, 0.12]\,h\,\mathrm{Mpc}^{-1}$. Vertical bands indicate the 68% limits of the posteriors. The ratios of the 1-σ constraints between FBI (blue) and SBI P+B (yellow) are shown in the upper right corners.

In the variant SBI case, which resembles current standard P+B analyses, the improvement factors between FBI and SBI P+B constraints on α at $k_{\max} = 0.1\,(0.12)\,h\,\mathrm{Mpc}^{-1}$ are 3.5 (5.2) for the SNG halo sample and 2.3 (3.5) for the Uchuu halo sample.

Summary and discussion.—In this Letter, we have presented the first σ_8 constraints from field-level inference on fully nonlinear biased tracers, specifically N-body halos in their comoving rest frame. Our constraints are based on the validity of the EFTofLSS on quasilinear scales, and rigorously marginalize over fully nonlinear scales. While we have not incorporated RSD due to tracer peculiar velocities, which are usually employed to constrain the combination $f\sigma_8$, our analysis remains a useful demonstration: First, a separate σ_8 constraint, combined with
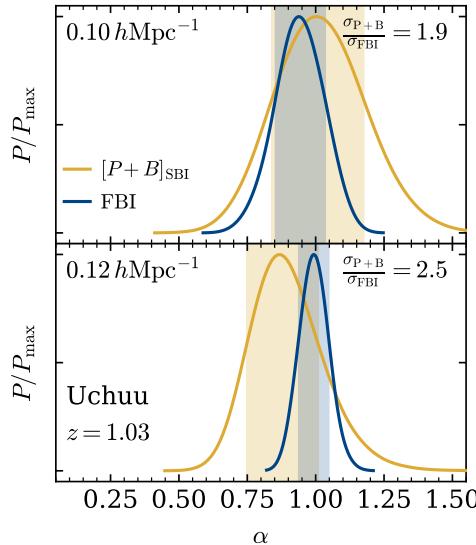
FIG. 3. Similar to Fig. 2, but obtained from the `Uchuu` sample (see text), also at $k_{\mathrm{max}} = [0.10, 0.12]\, h\,\mathrm{Mpc}^{-1}$.

$f\sigma_8$ as inferred from *linear* RSD, allows for a direct inference of the growth rate f, which is a sensitive probe of dark energy and gravity [53–56], much more so than the combination $f\sigma_8$. Second, if a given galaxy sample is affected by line-of-sight-dependent selection bias, then the leading RSD contribution is perfectly degenerate with the leading selection bias term [57, 58], i.e. in the same way as b_1 is degenerate with σ_8 in the analysis presented here; however, there are analogous higher-order protected RSD contributions which break the selection bias-growth rate degeneracy [59]. Hence, even when moving to a redshift-space analysis, being able to break the $b_1 - \sigma_8$ degeneracy remains extremely helpful for σ_8 and f constraints.

We compare our field-level results with a simulation-based inference based on summary statistics, namely the power spectrum and bispectrum. Using the same field-level forward model in both analyses, we demonstrate that the field-level approach significantly outperforms the summary statistics [Figs. 2 and 3]. While previous works, e.g. [36, 60], have shown a correspondence of field-level inference with power spectrum and bispectrum, this only holds when expanding the field-level likelihood at second order in perturbation theory. Here, our forward model includes third-order bias, and is thus expected to incorporate information from higher n-point functions. Indeed, our results show that, even on quasilinear scales, there is significant cosmological information beyond the power spectrum and bispectrum.

In future work, we will explore whether additional low-order summaries could extract this information, such as the trispectrum (4-point function).

While we have focused on dark-matter halos here, we demonstrate in [33] that this conclusion holds for simulated galaxies as well, as expected from EFT principles. Looking forward to FBI on observed data, [61, 62] provided the theoretical framework for and demonstrated a successful implementation of RSD into the `LEFTfield` forward model, respectively. This should enable a field-level analysis in redshift space and open up the access to additional cosmological information, as argued above.

We stress that we have not attempted to push our analysis to even smaller scales, instead aiming for converged posteriors at conservative scale cuts of $k_{\mathrm{max}} \leq 0.12\, h\,\mathrm{Mpc}^{-1}$ [63]. Already in this case, our results indicate that field-level inference enables robust constraints on the growth of structure, independent of the growth rate f, at the few-percent level even within a modest volume of $8\,(h^{-1}\mathrm{Gpc})^3$. This should allow for correspondingly improved constraints on cosmological parameters, in the standard ΛCDM as well as extended models, using the upcoming DESI [64, 65], Euclid [66, 67] and PFS [68] data.

[1] D. Baumann, A. Nicolis, L. Senatore, and M. Zaldarriaga, JCAP **07**, 051 (2012), arXiv:1004.2488 [astro-ph.CO].

[2] J. J. M. Carrasco, M. P. Hertzberg, and L. Senatore, JHEP **09**, 082 (2012), arXiv:1206.2926 [astro-ph.CO].

[3] S. M. Carroll, S. Leichenauer, and J. Pollack, Phys. Rev. D **90**, 023518 (2014), arXiv:1310.2920 [hep-th].

[4] N. Aghanim *et al.* (Planck), Astron. Astrophys. **641**, A6 (2020), [Erratum: Astron.Astrophys. 652, C4 (2021)], arXiv:1807.06209 [astro-ph.CO].

[5] S. Aiola *et al.* (ACT), JCAP **12**, 047 (2020), arXiv:2007.07288 [astro-ph.CO].

[6] L. Balkenhol *et al.* (SPT-3G), Phys. Rev. D **108**, 023510 (2023), arXiv:2212.05642 [astro-ph.CO].

[7] M. S. Madhavacheril *et al.* (ACT), Astrophys. J. **962**, 113 (2024), arXiv:2304.05203 [astro-ph.CO].

[8] C. Heymans *et al.*, Astron. Astrophys. **646**, A140 (2021), arXiv:2007.15632 [astro-ph.CO].

[9] T. M. C. Abbott *et al.* (DES), Phys. Rev. D **105**, 023520 (2022), arXiv:2105.13549 [astro-ph.CO].

[10] T. M. C. Abbott *et al.* (Kilo-Degree Survey, Dark Energy Survey), Open J. Astrophys. **6**, 2305.17173 (2023), arXiv:2305.17173 [astro-ph.CO].

[11] S. Sugiyama *et al.*, Phys. Rev. D **108**, 123521 (2023), arXiv:2304.00705 [astro-ph.CO].

[12] J. U. Lange, A. P. Hearin, A. Leauthaud, F. C. van den Bosch, E. Xhakaj, H. Guo, R. H. Wechsler, and J. DeRose, Mon. Not. Roy. Astron. Soc. (2023), 10.1093/mnras/stad473, arXiv:2301.08692 [astro-ph.CO].

[13] V. Poulin, J. L. Bernal, E. D. Kovetz, and M. Kamionkowski, Phys. Rev. D **107**, 123538 (2023), arXiv:2209.06217 [astro-ph.CO].

[14] N.-M. Nguyen, D. Huterer, and Y. Wen, Phys. Rev. Lett. **131**, 111001 (2023), arXiv:2302.01331 [astro-ph.CO].

[15] M.-X. Lin, B. Jain, M. Raveri, E. J. Baxter, C. Chang, M. Gatti, S. Lee, and J. Muir, (2023), arXiv:2308.16183 [astro-ph.CO].

[16] A. Amon and G. Efstathiou, Mon. Not. Roy. Astron. Soc. (2022), 10.1093/mnras/stac2429, arXiv:2206.11794 [astro-ph.CO].

[17] K. K. Rogers, R. Hložek, A. Laguë, M. M. Ivanov, O. H. E. Philcox, G. Cabass, K. Akitsu, and D. J. E. Marsh, JCAP **06**, 023 (2023), arXiv:2301.08361 [astro-ph.CO].

[18] C. Preston, A. Amon, and G. Efstathiou, Mon. Not. Roy. Astron. Soc. **525**, 5554 (2023), arXiv:2305.09827 [astro-ph.CO].

[19] Y. Akrami *et al.* (Planck), Astron. Astrophys. **641**, A7 (2020), arXiv:1906.02552 [astro-ph.CO].

[20] V. Desjacques, D. Jeong, and F. Schmidt, Phys. Rept. **733**, 1 (2018), arXiv:1611.09787 [astro-ph.CO].

[21] We note that some previous studies in the context of cosmological parameter inference, e.g. [69, 70], used the term "field-level inference" for a different analysis. As their inference pipeline relied on neural network data compression, there is in fact no guarantee that this neural compression is lossless and the encoded summary statistics capture all relevant cosmological information.

[22] F. Schmidt, JCAP **04**, 033 (2021), arXiv:2012.09837 [astro-ph.CO].

[23] F. Schmidt, JCAP **04**, 032 (2021), arXiv:2009.14176 [astro-ph.CO].

[24] D. K. Ramanah, G. Lavaux, J. Jasche, and B. D. Wandelt, Astron. Astrophys. **621**, A69 (2019), arXiv:1808.07496 [astro-ph.CO].

[25] A. Andrews, J. Jasche, G. Lavaux, and F. Schmidt, Mon. Not. Roy. Astron. Soc. **520**, 5746 (2023), arXiv:2203.08838 [astro-ph.CO].

[26] A. Kostić, N.-M. Nguyen, F. Schmidt, and M. Reinecke, JCAP **07**, 063 (2023), arXiv:2212.07875 [astro-ph.CO].

[27] U. Seljak, G. Aslanyan, Y. Feng, and C. Modi, JCAP **12**, 009 (2017), arXiv:1706.06645 [astro-ph.CO].

[28] A. E. Bayer, U. Seljak, and C. Modi, in *40th International Conference on Machine Learning* (2023) arXiv:2307.09504 [astro-ph.CO].

[29] N. Mudur, C. Cuesta-Lazaro, and D. P. Finkbeiner (2023) arXiv:2312.07534 [astro-ph.CO].

[30] N.-M. Nguyen, F. Schmidt, G. Lavaux, and J. Jasche, JCAP **03**, 058 (2021), arXiv:2011.06587 [astro-ph.CO].

[31] P. Villanueva-Domingo and F. Villaescusa-Navarro, Astrophys. J. **937**, 115 (2022), arXiv:2204.13713 [astro-ph.CO].

[32] A. Barreira, T. Lazeyras, and F. Schmidt, J. Cosmology Astropart. Phys. **2021**, 029 (2021), arXiv:2105.02876 [astro-ph.CO].

[33] Beyond-2pt Collaboration, E. Krause, Y. Kobayashi, A. N. Salcedo, M. M. Ivanov, T. Abel, K. Akitsu, R. E. Angulo, G. Cabass, S. Contarini, C. Cuesta-Lazaro, C. Hahn, N. Hamaus, D. Jeong, C. Modi, N.-M. Nguyen, T. Nishimichi, E. Paillas, M. Pellejero Ibañez, O. H. E. Philcox, A. Pisani, F. Schmidt, S. Tanaka, G. Verza, S. Yuan, and M. Zennaro, arXiv e-prints , arXiv:2405.02252 (2024), arXiv:2405.02252 [astro-ph.CO].

[34] P. S. Behroozi, R. H. Wechsler, and H.-Y. Wu, Astrophys. J. **762**, 109 (2013), arXiv:1110.4372 [astro-ph.CO].

[35] T. Ishiyama *et al.*, Mon. Not. Roy. Astron. Soc. **506**, 4210 (2021), arXiv:2007.14720 [astro-ph.CO].

[36] F. Schmidt, F. Elsner, J. Jasche, N. M. Nguyen, and G. Lavaux, JCAP **01**, 042 (2019), arXiv:1808.02002 [astro-ph.CO].

[37] F. Schmidt, G. Cabass, J. Jasche, and G. Lavaux, JCAP **11**, 008 (2020), arXiv:2004.06707 [astro-ph.CO].

[38] B. Tucci and F. Schmidt, JCAP **05**, 063 (2024), arXiv:2310.03741 [astro-ph.CO].

[39] A. H. Barnett, J. Magland, and L. af Klinteberg, SIAM Journal on Scientific Computing **41**, C479 (2019), arXiv:1808.06736 [math.NA].

[40] G. Cabass and F. Schmidt, JCAP **07**, 051 (2020), arXiv:2004.00617 [astro-ph.CO].

[41] G. Cabass and F. Schmidt, JCAP **04**, 042 (2020), arXiv:1909.04022 [astro-ph.CO].

[42] J. Jasche and G. Lavaux, Astron. Astrophys. **625**, A64 (2019), arXiv:1806.11117 [astro-ph.CO].

[43] R. Neal, in *Handbook of Markov Chain Monte Carlo* (2011) pp. 113–162.

[44] R. M. Neal, arXiv e-prints , physics/0009028 (2000), arXiv:physics/0009028 [physics.data-an].

[45] K. Cranmer, J. Brehmer, and G. Louppe, Proceedings of the National Academy of Science **117**, 30055 (2020), arXiv:1911.01429 [stat.ML].

[46] R. Scoccimarro, S. Colombi, J. N. Fry, J. A. Frieman, E. Hivon, and A. Melott, Astrophys. J. **496**, 586 (1998), arXiv:astro-ph/9704075.

[47] D. S. Greenberg, M. Nonnenmacher, and J. H. Macke, arXiv e-prints , arXiv:1905.07488 (2019), arXiv:1905.07488 [cs.LG].

[48] G. Papamakarios, T. Pavlakou, and I. Murray, arXiv e-prints , arXiv:1705.07057 (2017), arXiv:1705.07057 [stat.ML].

[49] A. Tejero-Cantero, J. Boelts, M. Deistler, J.-M. Lueckmann, C. Durkan, P. Gonçalves, D. Greenberg, and J. Macke, The Journal of Open Source Software **5**, 2505 (2020), arXiv:2007.09114 [cs.LG].

[50] S. Talts, M. Betancourt, D. Simpson, A. Vehtari, and A. Gelman, arXiv e-prints , arXiv:1804.06788 (2018), arXiv:1804.06788 [stat.ME].

[51] O. H. E. Philcox and M. M. Ivanov, Phys. Rev. D **105**, 043517 (2022), arXiv:2112.04515 [astro-ph.CO].

[52] G. D'Amico, Y. Donath, M. Lewandowski, L. Senatore, and P. Zhang, (2022), arXiv:2206.08327 [astro-ph.CO].

[53] E. V. Linder, Phys. Rev. **D72**, 043529 (2005), arXiv:astro-ph/0507263 [astro-ph].

[54] E. V. Linder and R. N. Cahn, Astropart. Phys. **28**, 481 (2007), arXiv:astro-ph/0701317.

[55] E. V. Linder, Phys. Rev. D **79**, 063519 (2009), arXiv:0901.0918 [astro-ph.CO].

[56] Y. Wen, N.-M. Nguyen, and D. Huterer, JCAP **09**, 028 (2023), arXiv:2304.07281 [astro-ph.CO].

[57] C. M. Hirata, MNRAS **399**, 1074 (2009), arXiv:0903.4929 [astro-ph.CO].

[58] Z. Zheng, R. Cen, H. Trac, and J. Miralda-Escudé, ApJ **726**, 38 (2011), arXiv:1003.4990 [astro-ph.CO].

[59] N. Agarwal, V. Desjacques, D. Jeong, and F. Schmidt, JCAP **03**, 021 (2021), arXiv:2007.04340 [astro-ph.CO].

[60] G. Cabass, M. Simonović, and M. Zaldarriaga, Phys. Rev. D **109**, 043526 (2024), arXiv:2307.04706 [astro-ph.CO].

[61] G. Cabass, JCAP **01**, 067 (2021), arXiv:2007.14988 [astro-ph.CO].

[62] J. Stadler, F. Schmidt, and M. Reinecke, JCAP **10**, 069 (2023), arXiv:2303.09876 [astro-ph.CO].

[63] The cost of our forward model scales directly as $\sim \Lambda^3 \ln \Lambda$.

[64] A. Aghamousa *et al.* (DESI), (2016), arXiv:1611.00036 [astro-ph.IM].

[65] S. Goldstein, M. Park, M. Raveri, B. Jain, and L. Samushia, Phys. Rev. D **107**, 063530 (2023), arXiv:2207.01612 [astro-ph.CO].

[66] A. Blanchard *et al.* (Euclid), Astron. Astrophys. **642**, A191 (2020), arXiv:1910.09273 [astro-ph.CO].

[67] B. Bose *et al.* (Euclid), (2023), arXiv:2311.13529 [astro-ph.CO].

[68] M. Takada *et al.* (PFS), PASJ **66**, R1 (2014), arXiv:1206.0737 [astro-ph.CO].

[69] P. Lemos *et al.*, in *40th International Conference on Machine Learning* (2023) arXiv:2310.15256 [astro-ph.CO].

[70] C. Hahn *et al.*, (2023), arXiv:2310.15246 [astro-ph.CO].

[71] H. Rubira and F. Schmidt, JCAP **01**, 031 (2024), arXiv:2307.15031 [astro-ph.CO].

[72] M. Mirbabayi, F. Schmidt, and M. Zaldarriaga, JCAP **07**, 030 (2015), arXiv:1412.5169 [astro-ph.CO].

[73] P. McDonald and A. Roy, JCAP **08**, 020 (2009), arXiv:0902.0991 [astro-ph.CO].

[74] V. Assassi, D. Baumann, D. Green, and M. Zaldarriaga, JCAP **08**, 056 (2014), arXiv:1402.5916 [astro-ph.CO].

[75] F. Elsner, F. Schmidt, J. Jasche, G. Lavaux, and N.-M. Nguyen, JCAP **01**, 029 (2020), arXiv:1906.07143 [astro-ph.CO].

[76] N. Hamaus, U. Seljak, V. Desjacques, R. E. Smith, and T. Baldauf, Phys. Rev. D **82**, 043515 (2010), arXiv:1004.5377 [astro-ph.CO].

[77] M. Schmittfull, M. Simonović, V. Assassi, and M. Zaldarriaga, Phys. Rev. D **100**, 043514 (2019), arXiv:1811.10640 [astro-ph.CO].

[78] M. M. Ivanov, O. H. E. Philcox, T. Nishimichi, M. Simonović, M. Takada, and M. Zaldarriaga, Phys. Rev. D **105**, 063512 (2022), arXiv:2110.10161 [astro-ph.CO].

[79] See Sec. 2.5.2-2.5.3 in [20].

[80] D. P. Kingma and J. Ba, arXiv e-prints , arXiv:1412.6980 (2014), arXiv:1412.6980 [cs.LG].

Supplementary Material

Appendix A: Forward model

The `LEFTfield` forward models were detailed in [22], specifically their Sec. 2. Here, we describe the specific procedure we follow to construct the Eulerian density field δ and bias fields O out of a particular realization of the linear density field $\delta^{(1)}$:

1. Apply a cubic sharp-k filter with cutoff Λ to the linear density field $\delta^{(1)} \to \delta_\Lambda^{(1)}$ in Fourier space. Nonlinear gravitational evolution induces mode coupling, hence contributions from arbitrarily small-scale modes. This cutoff is therefore required for either the Gaussian initial conditions or linear density field [37, 71].

2. Construct the Lagrangian displacement field $\mathbf{s}(\boldsymbol{k}) = \sum_n \mathbf{s}^{(n)}(\boldsymbol{k})$, where n is the order of Lagrangian Perturbation Theory (LPT); in this work, we adopt $n = 2$, i.e. 2LPT.

3. Evaluate the displacement field and displace mass particles from Lagrangian to Eulerian positions accordingly in real space.

4. Assign the displaced mass particles to the Eulerian grid of size $N_{\text{grid}}^{\text{Eul}}$ via a non-uniform-to-uniform fast Fourier transform (NUFFT, see below).

5. Apply another cubic sharp-k filter with cutoff $\Lambda' = 1.2 k_{\text{max}}$.

6. Construct the bias fields O out of the filtered Eulerian density grid $\delta_{\Lambda'}$. We choose $\Lambda = 1.2\Lambda' = 1.44 k_{\text{max}}$ throughout.

A new feature of `LEFTfield` in this work is the introduction of the NUFFT algorithm for density assignment in step 4 above. NUFFT implements the nonuniform-to-uniform discrete Fourier transform $f(\vec{x}) \to \tilde{f}(\vec{k})$ by assigning pseudoparticles at positions $\vec{x}_i$ with weights $f(\vec{x}_i)$ to a supersampled grid $(N_{\text{grid}}^{\text{NUFFT}})^3$ (typically $N_{\text{grid}}^{\text{NUFFT}} = 1.2 - 2 N_{\text{grid}}^{\text{Eul}}$) using a specific assignment kernel with compact support (roughly 4-16 grid cells). The algorithm proceeds to perform an FFT on the supersampled grid, deconvolves the assignment kernel, and resizes the grid in Fourier space to the target uniform discrete Fourier transform $\tilde{f}(\vec{k})$. This method is approximate, but accuracy close to machine precision can be obtained for very reasonable computational effort [39].

1. Galaxy bias expansion

The EFT galaxy bias expansion expands the local galaxy density field δ_g in a set of galaxy bias operators, i.e. fields, constructed recursively out of the gravitational potential Φ. The latter relates to the local matter density field δ through the Poisson equation

$$\nabla^2 \Phi(\boldsymbol{x}, \tau) = (aH)^2 \Omega_m \delta(\boldsymbol{x}, \tau), \tag{A1}$$

in which the scale factor a is a function of the conformal time τ.

At each n-th order, the set includes local gravitational operators that can be classified into two categories: (1) operators that involve two spatial derivatives for each instance of Φ and combinations thereof, plus (2) higher-derivative operator(s) that involve more than two spatial derivatives acting on a single instance of Φ [20].

Let us first consider category (1). Following [20, 72] and Einstein notation, we define

$$\Pi_{ij}^{[1]}(\boldsymbol{x}, \tau) = \frac{2}{3\Omega_m a H^2} \partial_{x_i} \partial_{x_j} \Phi(\boldsymbol{x}, \tau) = K_{ij}(\boldsymbol{x}, \tau) + \frac{1}{3} \delta_{ij}^K \delta(\boldsymbol{x}, \tau), \tag{A2}$$

where $\mathbf{K} \equiv \mathcal{D}\delta$ denotes the local tidal field tensor and δ^K denotes the Kronecker delta. The superscript [1] indicates that the lowest-order contribution to this operator is at $n = 1$ order in perturbation theory.

It follows that $\mathbf{\Pi}^{[n]}$ is given by the recursive relation:

$$\Pi_{ij}^{[n]} = \frac{1}{(n-1)!} \left[(aHf)^{-1} \frac{\mathrm{D}}{\mathrm{D}\tau} \Pi_{ij}^{[n-1]} - (n-1) \Pi_{ij}^{[n-1]} \right], \tag{A3}$$

where $\frac{\mathrm{D}}{\mathrm{D}\tau}$ is the convective derivative with respect to τ.

Up to third order in the Eulerian basis, the complete set of operators in category (1) is given by

1st-order $\qquad$ $\mathrm{tr}\left[\mathbf{\Pi}^{[1]}\right]$

2nd-order $\qquad$ $\mathrm{tr}\left[\mathbf{\Pi}^{[1]}\,\mathbf{\Pi}^{[1]}\right], \mathrm{tr}\left[\mathbf{\Pi}^{[1]}\right]\mathrm{tr}\left[\mathbf{\Pi}^{[1]}\right],$

3rd-order $\qquad$ $\mathrm{tr}\left[\mathbf{\Pi}^{[1]}\,\mathbf{\Pi}^{[1]}\,\mathbf{\Pi}^{[1]}\right], \mathrm{tr}\left[\mathbf{\Pi}^{[1]}\,\mathbf{\Pi}^{[1]}\right]\mathrm{tr}\left[\mathbf{\Pi}^{[1]}\right], \mathrm{tr}\left[\mathbf{\Pi}^{[1]}\right]\mathrm{tr}\left[\mathbf{\Pi}^{[1]}\right]\mathrm{tr}\left[\mathbf{\Pi}^{[1]}\right], \mathrm{tr}\left[\mathbf{\Pi}^{[1]}\mathbf{\Pi}^{[2]}\right].$

The above list, albeit complete (at third order), is only one possible linear combination of Eulerian cubic bias operators O in category (1). In fact, there exists different equivalent linear combinations and conventions in the literature, e.g. [73, 74]. We refer readers to App. C.2 of [20] for further details and discussions.

The full set of EFT bias operators O up to third order, following the convention in [20, 72], therefore is

$$O \in \left[\delta, \delta^2, K^2, \delta^3, K^3, \delta K^2, O_{\mathrm{td}}, \nabla^2\delta\right], \tag{A4}$$

where we include one operator from category (2), namely the leading-order higher-derivative bias operator $\nabla^2\delta$. At third order, a nonlocal operator from category (1) appears in Eq. (A4), namely

$$O_{\mathrm{td}} \equiv \frac{8}{21}K_{ij}^{(1)}\mathcal{D}^{ij}\left[\left(\delta^{(1)}\right)^2 - \frac{3}{2}\left(K_{ij}^{(1)}\right)^2\right]. \tag{A5}$$

Eq. (A4) explicitly provides the EFT bias operators considered in the modeling of the FBI analysis and SBI P+B Fisher forecast.

For both inference methods (FBI and SBI P+B), we use `LEFTfield` to construct the Eulerian galaxy bias fields O in Eq. (A4) out of the filtered Eulerian matter density field $\delta_{\Lambda'}$ at the tracer redshift z, following the procedure outlined in the previous section of Supplementary Material.

2. Marginalization over galaxy bias

Eq. (5) can be rewritten as

$$\mathcal{L}_{\mathrm{FBI}}^{\mathrm{expl.}}\left(\delta_g^{\mathrm{obs.}}\Big|\hat{s}, \alpha, \{b_O\}, \{\sigma_\epsilon\}\right) = \mathbb{N}\exp\left[-\frac{1}{2}\sum_{\boldsymbol{k}>0}^{k_{\mathrm{max}}}\ln\sigma_\epsilon^2(k)\right] \times \exp\left\{-\frac{1}{2}C + \sum_O b_O B_O - \frac{1}{2}\sum_{O,O'}b_O b_{O'} A_{OO'}\right\}, \tag{A6}$$

where $\mathbb{N}$ is a parameter-independent normalization constant and following [75], we introduce a scalar C, a vector B_O and a matrix $A_{OO'}$:

$$C = \sum_{\boldsymbol{k}>0}^{k_{\mathrm{max}}}\frac{1}{\sigma_\epsilon^2(k)}\left|\delta_g^{\mathrm{obs.}} - \sum_{O_{\mathrm{unmarg.}}}b_{O_{\mathrm{unmarg.}}}O_{\mathrm{unmarg.}}(\boldsymbol{k})\right|^2$$

$$B_O = \sum_{\boldsymbol{k}>0}^{k_{\mathrm{max}}}\frac{\left[\delta_g^{\mathrm{obs.}}(\boldsymbol{k}) - \sum_{O_{\mathrm{unmarg.}}}b_{O_{\mathrm{unmarg.}}}O_{\mathrm{unmarg.}}(\boldsymbol{k})\right]O^*(\boldsymbol{k})}{\sigma_\epsilon^2(k)}$$

$$A_{OO'} = \sum_{\boldsymbol{k}>0}^{k_{\mathrm{max}}}\frac{O(\boldsymbol{k})O'^*(\boldsymbol{k})}{\sigma_\epsilon^2(k)}. \tag{A7}$$

Eq. (A6) thereby separates the n galaxy bias coefficients $\{b_O\}$ to be marginalized over and the rest $\{b_O\}_{\mathrm{unmarg.}}$ to be left unmarginalized.

For Gaussian priors assumed in Eq. (A9), integrating out $\{b_O\}$ in Eq. (A6) (using Gaussian integrals) yields

$$\mathcal{L}_{\mathrm{FBI}}^{\mathrm{expl.}}\left(\delta_g^{\mathrm{obs.}}\Big|\hat{s}, \alpha, \{b_O\}_{\mathrm{unmarg.}}, \{\sigma_\epsilon\}\right) = \left(\prod_O\int db_O\right)P\left(\delta_g^{\mathrm{obs.}}\Big|\hat{s}, \{b_O\}\right)$$

$$= \mathbb{N}(2\pi)^{n/2}\left|A_{OO'}\right|^{-1/2}\exp\left[-\frac{1}{2}\sum_{\boldsymbol{k}>0}^{k_{\mathrm{max}}}\ln\sigma_\epsilon^2(k)\right]$$

$$\times \exp\left\{-\frac{1}{2}C(\{b_O\}) + \frac{1}{2}\sum_{O,O'}B_O(\{b_O\})(A^{-1})_{OO'}B_{O'}(\{b_O\})\right\}. \tag{A8}$$
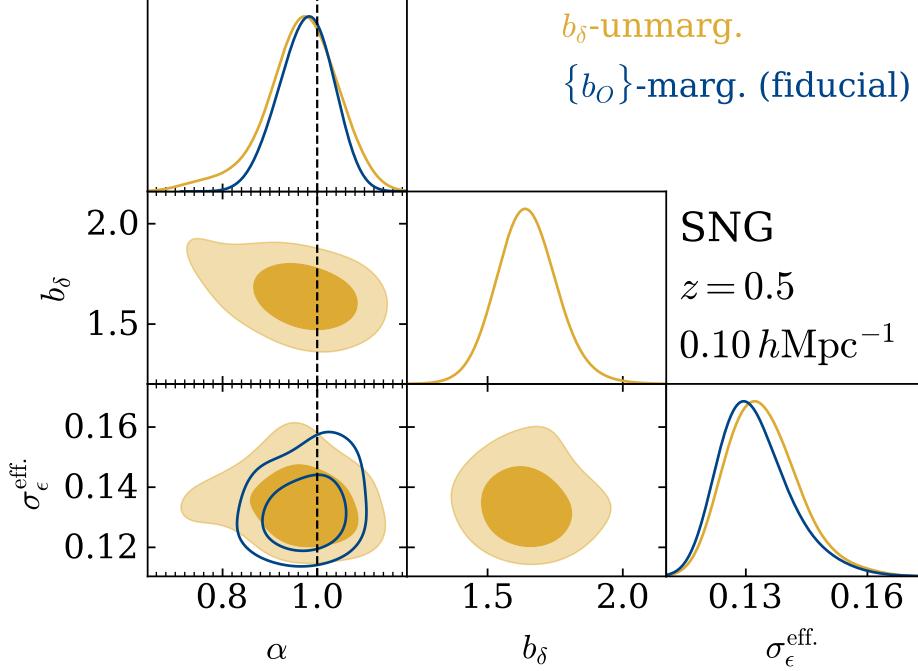
FIG. 4. Posterior consistency between the (fiducial) $\{b_O\}$-marginalized likelihood (blue, open) and the b_δ-unmarginalized likelihood (yellow, filled). Panels show the 1D and 2D constraints on the amplitude rescaling parameters $\alpha = \sigma_8/\sigma_{8,\text{fid.}}$, the linear galaxy bias coefficient b_δ, and the galaxy stochasticity parameters $[\sigma_{\epsilon,0}, \sigma_{\epsilon,k^2}]$. Contours indicate the 68% and 95% confidence intervals.

In the FBI analysis, we analytically marginalize the entire set of bias coefficients $\{b_O\}$. Analytical marginalization improves the MCMC sampling efficiency by reducing the dimensions of parameter space to explore. A validation of Eq. (A8) on mock data can be found in App. F of [26].

In this work, we validate the analytical marginalization by re-analyzing the SNG sample at $k_{\max} = 0.10\,h\,\text{Mpc}^{-1}$ using a likelihood that marginalizes over all galaxy bias coefficients in Eq. (A4) except the linear bias coefficient b_δ. In Fig. 4, we compare the posteriors obtained with this "b_δ-unmarg." likelihood with those obtained in our fiducial FBI analysis using the "$\{b_O\}$-marg." likelihood [Eq. (A8)]. All 1D and 2D marginal posteriors are consistent with their corresponding counterparts. The b_δ-unmarginalized posterior, as projected down to the $b_\delta - \alpha$ plane, showcases the ability of FBI to break the $b_\delta - \sigma_8$ degeneracy by extracting information from nonlinear clustering: both b_δ and α are well constrained.

We note however that the "b_δ-unmarg." analysis assumes a prior consisting of the product of the fiducial Gaussian prior (see next section) and a uniform prior, i.e. $\mathcal{P}(b_\delta|\text{unmarg}) = \mathcal{N}(1.0, 5.0)\,\mathcal{U}(0., 7.)$, and is thus slightly mismatched with the marginalized chains. However, given the tight constraint obtained on b_δ, we do not expect this small mismatch to be quantitatively relevant. However, the "b_δ-unmarg." chains have only reached roughly 21 effective samples, as compared to the substantially larger sample size of the fiducial chains [Table I]. We therefore only focus on the overall consistency between the "b_δ-unmarg." and the fiducial "$\{b_O\}$-marg." posteriors.

3. Priors

a. *FBI priors.* In the FBI analysis, we assume the following priors on the amplitude rescaling parameter α, bias coefficients $\{b_O\}$ in Eq. (A4), and the noise parameters $\{\sigma_\epsilon\}$:

$$\mathcal{P}(\alpha) = \mathcal{U}(0.5, 1.5),$$
$$\mathcal{P}(b_\delta) = \mathcal{N}(1.0, 5.0),$$
$$\mathcal{P}(b_{\delta^2}) = \mathcal{P}(b_{K^2}) = \mathcal{P}(b_{\delta^3}) = \mathcal{P}(b_{\delta K^2}) = \mathcal{P}(b_{K^3}) = \mathcal{P}(b_{O_{\text{td}}}) = \mathcal{N}(0.0, 1.0),$$
$$\mathcal{P}(b_{\nabla^2 \delta}) = \mathcal{N}(0.0, 5.0),$$
$$\mathcal{P}(\sigma_{\epsilon,0}) = \mathcal{U}(0.8\sigma_{\epsilon,\text{Poisson}}, 100.), \quad \mathcal{P}(\sigma_{\epsilon,k^2}) = \mathcal{U}(-10.0, 100.0), \tag{A9}$$

where $\mathcal{U}(a, b)$ represents a uniform distribution between the lower bound a and upper bound b, while $\mathcal{N}(\mu, \sigma)$ represents a normal distribution with mean μ and standard deviation σ.

$\sigma_{\epsilon,\mathrm{Poisson}}$ in Eq. (A9) is the expected (Gaussian) RMS noise amplitude based on the Poisson shot noise expectation given the comoving number density of the tracers considered. We find it necessary to impose an informative lower limit on the noise level, as the FBI MCMC chains tend to get stuck in a region of unphysically low σ_ϵ. The causes of this drift are still under investigation. The lower limit $0.8\sigma_{\epsilon,\mathrm{Poisson}}$ we impose in this work is significantly *below* what previous studies of halo stochasticity have found [76, 77].

We note that our priors on $\{b_O\}$ are consistent with other EFT cosmology analyses, e.g. [33, 51, 78]. Further, in [33], we have explicitly verified that uniform flat priors (see their Eqs. (20-21)), Gaussian priors (our Eq. (A9)), or anything in between, on $\{b_O\}$ do not affect the σ_8 constraint.

b.   *SBI* P+B *priors.*   The SBC tests for asserting the uncertainties of the estimated SBI posterior are only feasible when the posterior is amortized, i.e., not restricted to a single observation. However, performing an amortized NPE by drawing samples from the considerably wide FBI prior would require a huge amount of simulations for convergence.

We therefore follow the same strategy described in [38]: (1) We initiate the training sequence using the same priors as in the FBI analysis and a sequential NPE (SNPE) is performed over multiple rounds, where the posterior of each round is used as proposal for the next one. This allows us to exclude the prior regions where the posterior has negligible support. (2) The training data for the final NPE posterior is then sampled from a wider Gaussian than the posterior obtained in the last SNPE round. This allows for faster SBC analysis and for convergence with fewer simulations, but still guarantees that both SBI and FBI analysis are consistent and that none of the parameters are prior dominated.

Another subtlety employed in SBI is that the bias parameters $b_{O^{(n)}}$, where n stands for the order of the bias operators as listed in Eq. (A4), are correspondingly scaled to $\alpha^n b_{O^{(n)}}$ when doing inference. This avoids prior volume effects by reducing degeneracies from the SBI data vector. Additionally, the higher-derivative biases $b_{\nabla^2\delta}$ and $b_{\nabla^2\varepsilon}$ are made dimensionless after a scaling by R_*^{-2}, where R_* is the characteristic scale of halo formation (in this work, we assume $R_* = 5h^{-1}\mathrm{Mpc}$).

4.   Eulerian and Lagrangian bias expansions

In both FBI and SBI P+B fiducial analyses, we adopt the Eulerian basis for the EFT galaxy bias expansion, as described in Sec. II of the Supplementary Material. Another valid choice of basis for the EFT bias fields O and bias expansion is the Lagrangian space [79]. In principle, we can construct O at the Lagrangian coordinates $\boldsymbol{q} = \boldsymbol{x}(\tau = 0)$ and advect them to the final Eulerian coordinates $\boldsymbol{x}(\boldsymbol{q}, \tau)$, following the procedure in Sec. 3 of [22]. As the displacement field contains only perturbative modes and is protected by the equivalence principle of corrections from bias terms, UV effects in displacement operations are not of concern. Therefore, a cut in the evolved density field, as listed by step 5 in App. A1, is unnecessary in the Lagrangian-bias forward model.

Both bias expansions are equivalent at fixed order in perturbation theory but differ in higher-order terms. Thus, a comparison of both cases provides an estimate of the importance of higher-order terms in the bias expansion.

In Fig. 5, we show the constraints obtained with each bias model, for the fiducial halo sample at $z = 0.5$ and the analysis cutoff scale $k_{\mathrm{max}} = 0.10\,h\,\mathrm{Mpc}^{-1}$. Posteriors from the two analyses are consistent with one another, taking into account the fact that the Lagrangian chains only consist of roughly 13 effective samples in total (and hence moments of its estimated posteriors are not expected to be accurate).

Appendix B: Posterior sampling and validation

1.   FBI—MCMC chain initialization

Regardless of the initialization, once converged, MCMC chains should sample the same underlying distribution. MCMC convergence is particularly relevant in the case of high-dimensional input data and sampling space, such as that of FBI.

In this work, we explore two strategies to initialize a MCMC chain. Specifically, for each FBI analysis, we initialize one MCMC chain from the true initial conditions (true-phase initialization, TPI) $\hat{s}_{\mathrm{true}}$, and multiple chains at different random initial conditions (random-phase initialization, RPI) $\hat{s}$. Monitoring and comparing behaviors between TPI and RPI chains help us verify that the chains have converged.

In Fig. 6, we show a trace plot of the parameter α in the FBI chains, both TPI (light blue) and RPI. After the warm-up phase, they are virtually indistinguishable, all sampling the underlying posterior.
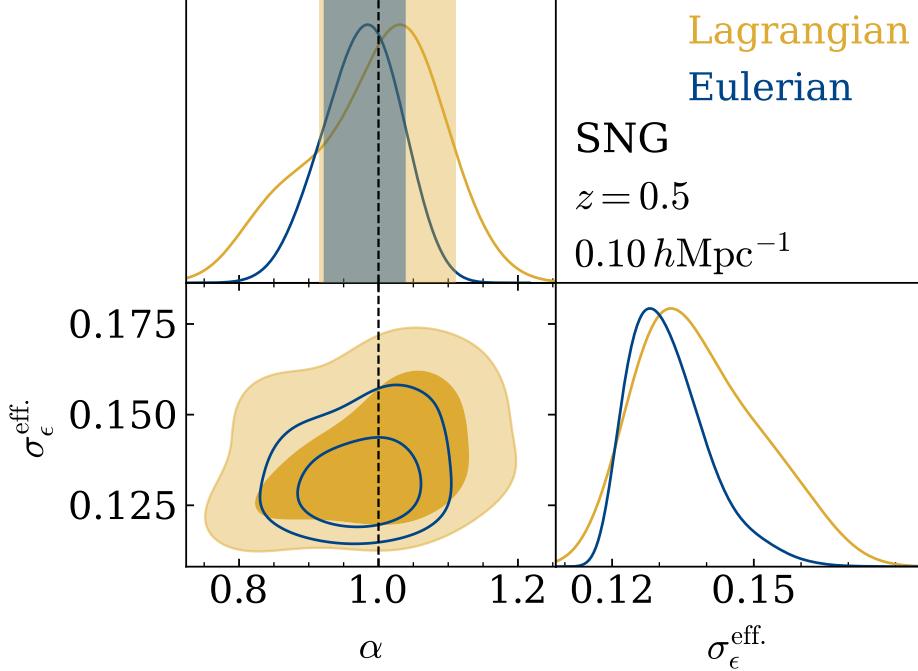
FIG. 5. Posterior consistency between the (fiducial) Eulerian bias and the Lagrangian bias. Constraints on the amplitude rescaling parameter $\alpha = \sigma_8/\sigma_{8,\mathrm{fid.}}$ and the effective noise amplitude $\sigma_\epsilon^{\mathrm{eff.}} = \sigma_{\epsilon,0}[1 + \sigma_{\epsilon,k^2} k_{\mathrm{max}}^2]$, obtained with the fiducial Eulerian bias model (blue, open) and with the alternative Lagrangian bias model (yellow, filled). Contours indicate the 68% and 95% confidence intervals of the corresponding 2D marginal posteriors. Vertical bands indicate the 68% limits of the corresponding 1D marginal posteriors.

2. FBI—Posterior of initial conditions

Our field-level constraints include the initial conditions $\hat{s}$ of the observed universe volume. Fig. 7 displays the coverage of the initial condition posterior, in different summary statistics.

Firstly, the histograms in the top panel represent the probability density of $\hat{s} - \hat{s}_{\mathrm{true}}$ where $\hat{s}$ are samples from the posterior, while $\hat{s}_{\mathrm{true}}$ is the true initial conditions. The distribution perfectly follows a normal distribution $\mathcal{N}(0,1)$, indicating that our $\hat{s}$ posterior in FBI correctly recovers the first two moments of the true underlying $\mathcal{P}(\hat{s}|\delta_g^{\mathrm{obs.}})$.

Secondly, the middle panel shows square roots of the ratios between posterior $\hat{s}$ power spectra over the $\hat{s}_{\mathrm{true}}$ power spectrum. The quantity—representing a transfer function—is consistent with 1 across all Fourier modes, as expected.

Lastly, in the bottom panel, we show the variance of $\hat{s} - \hat{s}_{\mathrm{true}}$, which goes to zero for zero noise and a perfect forward model. The inferred variance is broadly consistent with the Poisson noise expectation of the data; see also the extended discussion in [26].

3. FBI—Parameter posterior convergence

Given the (extremely) high-dimensional space in FBI, posterior convergence is a particular concern. In Fig. 8, we show the marginal posteriors of α and $\sigma_\epsilon^{\mathrm{eff.}} = \sigma_{\epsilon,0}[1 + \sigma_{\epsilon,k^2} k_{\mathrm{max}}^2]$ from the same set of MCMC chains shown in Fig. 6. These chains come from the FBI analysis of the main sample at the analysis cutoff scale $k_{\mathrm{max}} = 0.1\,h\,\mathrm{Mpc}^{-1}$. The posterior contours are consistent with each other, indicating that these chains —-individually and jointly—sample the same underlying posterior distribution.

Markov chain transitions and samples are correlated. Therefore, the MCMC sampling error on the mean of α can be estimated by $\sigma_\alpha/\mathrm{ESS}$ where ESS is the effective number of independent samples, i.e. effective sample size of α. Table I reports the ESS for each FBI analysis in Fig. 2 and Appendix B 5.
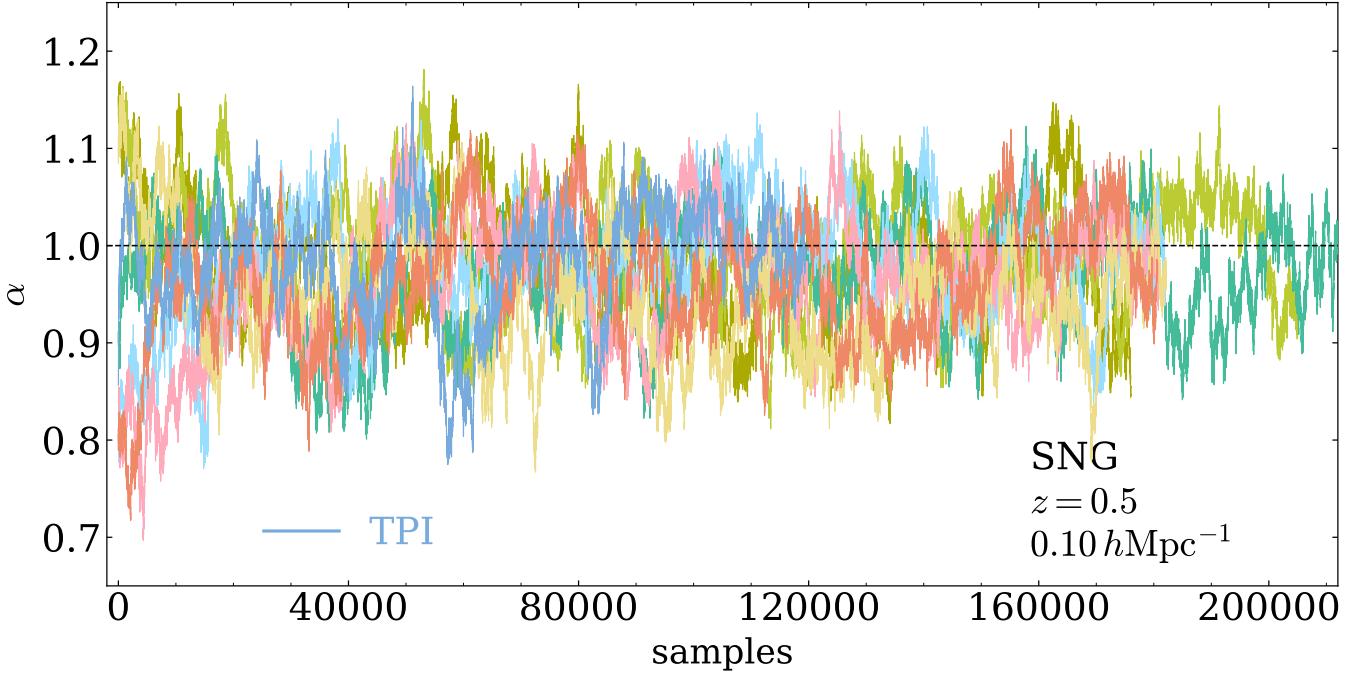
FIG. 6. Trace plot of the amplitude rescaling parameter $\alpha = \sigma_8/\sigma_{8,\mathrm{fid.}}$ for TPI and RPI chains in the FBI analysis. Different colors indicate chains starting at distinct initial points in parameter space.

Analysis	ESS [samples]
SNG, $k_{\max} = 0.10\,h\,\mathrm{Mpc}^{-1}$	200
SNG, $k_{\max} = 0.12\,h\,\mathrm{Mpc}^{-1}$	100
Uchuu, $k_{\max} = 0.10\,h\,\mathrm{Mpc}^{-1}$	180
Uchuu, $k_{\max} = 0.12\,h\,\mathrm{Mpc}^{-1}$	60

TABLE I. Estimates for the effective sample size of α (middle column) and the classical G-R statistics (right column) in the FBI analyses.

4. SBI P+B—Posterior diagnostics

a. Analysis settings. For the scale cuts considered in this work, namely $k_{\max} = 0.1\,(0.12)\,h\,\mathrm{Mpc}^{-1}$, the power spectrum bins have dimension $N_{\mathrm{bins}} = 15\,(18)$, corresponding to a total of $D = 443\,(714)$ dimensions for the SBI data vector. We use a total simulation budget of $N_{\mathrm{sim}} = 5 \times 10^5$ for all posteriors shown with the except of the SNG halo sample at higher cutoff, which uses $N_{\mathrm{sim}} = 10^6$, as motivated by the convergence tests below. We use 10^5 samples from the posterior for plotting.

b. Hyperparameters and training. We use the SNPE method of [47] with 10 atoms for atomic proposals and masked autoregressive flows [48] with 10 autoregressive layers, each constructed using two fully-connected tanh layers with 100 hidden units. We train the models by stochastically minimizing the loss using the Adam optimizer [80] with learning rate of 5×10^{-4} and batch size of 50. 10% of the samples are used for validation and we stop training if the validation set loss did not improve after 20 consecutive epochs.

c. Simulation-based calibration (SBC). We use SBC [50] to analyze the uncertainties of the final estimated posterior. A healthy posterior should lead to uniformly distributed rank statistics for all parameters. If the SBI underestimates (overestimates) the true posterior variance for some parameter, one expects a "U-shaped" ("∩-shaped") rank distribution. As shown in Fig. 9, the ranks are uniformly distributed, indicating that the estimated posteriors passed the calibration test.

d. Convergence tests. For SBI, we check for posterior convergence (and uncertainty quantification) by monitoring the estimated uncertainties on α as a function of the simulation budget, N_{sim}. We show the SBI convergence tests in Fig. 10, where the standard deviations of α posteriors are normalized by the same constraints from the corresponding
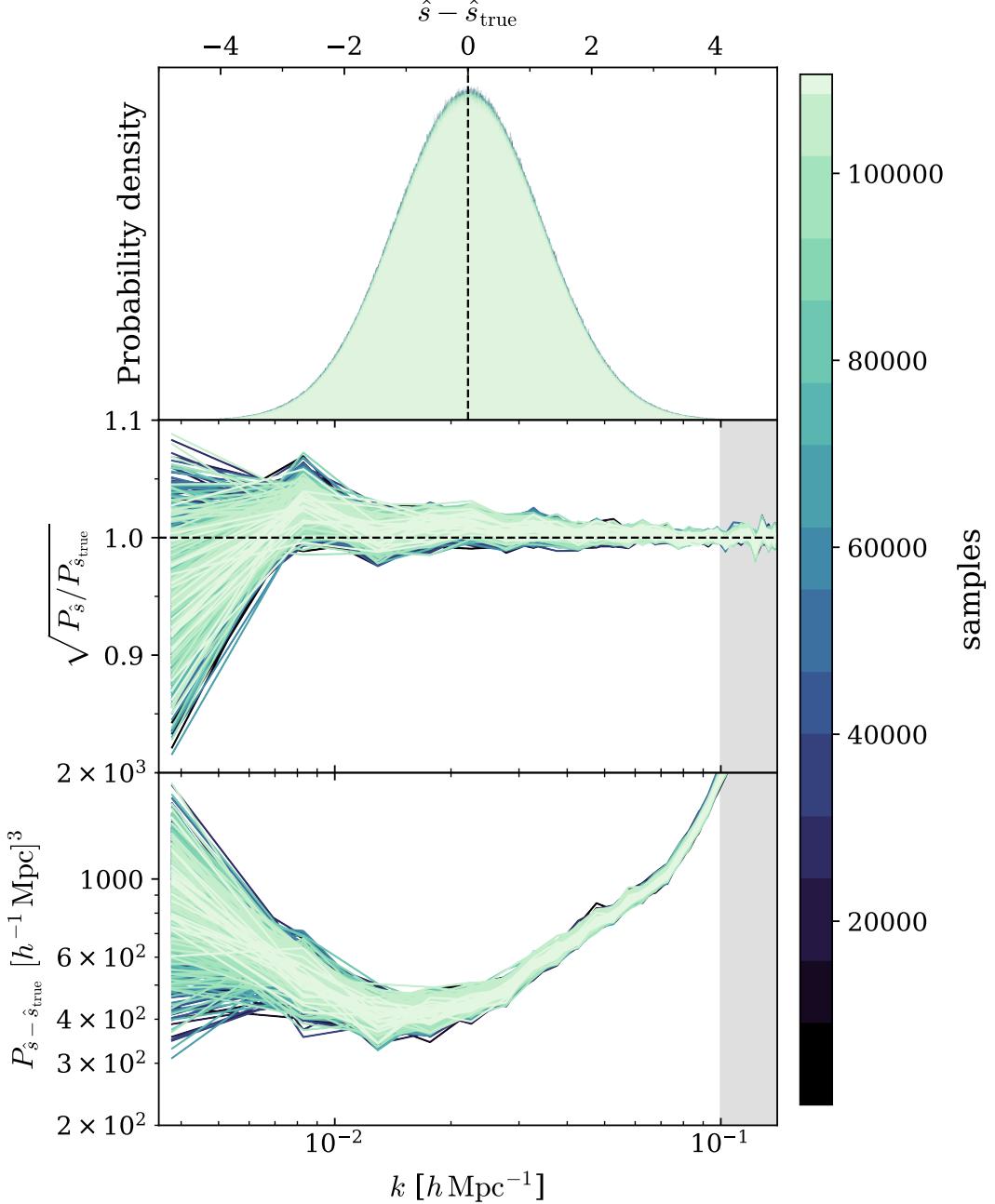
FIG. 7. Trace plots of different statistics of the posterior initial conditions $\hat{s}$ from our TPI chain in Fig. 6. For visibility, in this plot we remove the warm-up phase and further thin the samples by a factor of 100. The grey vertical bands in the lower two panels indicate scales above the cutoff $k_{\mathrm{max}} = 0.1\,h\,\mathrm{Mpc}^{-1}$. Top: Distribution of the difference between posterior and true initial conditions, $\hat{s} - \hat{s}_{\mathrm{true}}$. Middle: Square root of the ratios between posterior and true power spectrum of initial conditions $P_{\hat{s}}/P_{\hat{s}_{\mathrm{true}}}$. Bottom: Variance of the residuals between posterior and true initial conditions, $P_{\hat{s}-\hat{s}_{\mathrm{true}}}$.

Fisher analyses. We perform this test for all parameters in the SBI P+B analyses, but show only the case of α to avoid clutter.

  e. *Fisher analysis.*   The 1σ error on the parameter θ_α is estimated by $\sqrt{(F^{-1})_{\alpha\alpha}}$, where F is the Fisher information matrix, evaluated at the SBI posterior mode, θ_{MAP}. Following [38], we compute F by taking numerical derivatives with respect to the model parameters of the mean of the data vector obtained by averaging over 1000 LEFTfield data realizations. We estimate the sample covariance from 10^5 simulations evaluated at θ_{MAP}. For the Fisher analysis when including all third-order bias parameters, we use the θ_{MAP} obtained by SBI from the restricted case (i.e., with
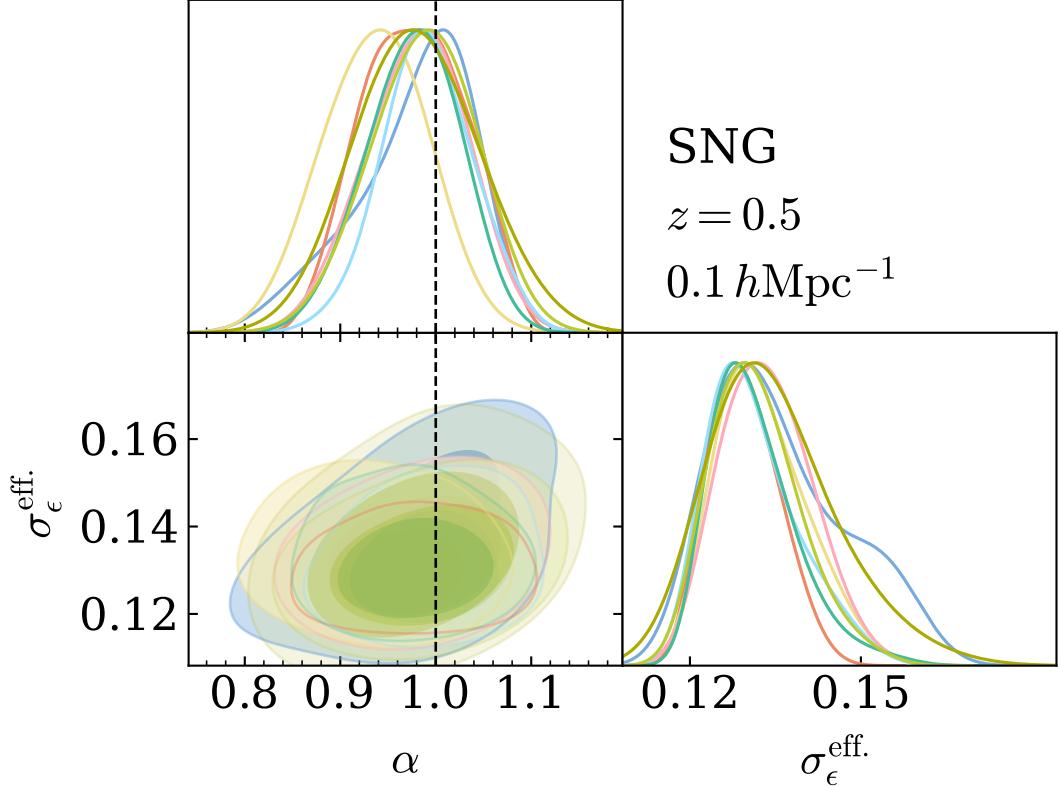
FIG. 8. Posterior consistency between MCMC chains with different initializations. Each color maps to the chain with the same color in Fig. 7. Panels show 1D and 2D marginal posteriors of $[\alpha, \{\sigma_\epsilon^{\text{eff.}}\}]$. Contours indicate the 68% and 95% credible intervals.
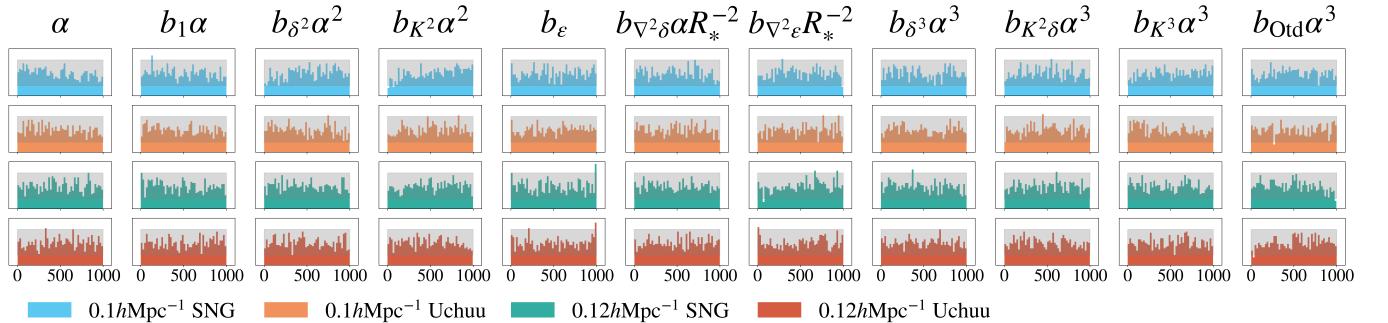

FIG. 9. Rank distributions from the SBC tests for the SBI posteriors shown in Fig. 2 and Fig. 3. The grey shaded area indicates the 99% confidence interval of an uniform distribution. Upper, middle and lower panels correspond to the SNG ($k_{\max} = 0.1\,h\,\mathrm{Mpc}^{-1}$), Uchuu ($k_{\max} = 0.1\,h\,\mathrm{Mpc}^{-1}$), and SNG ($k_{\max} = 0.12\,h\,\mathrm{Mpc}^{-1}$) analyses, respectively.

only b_{Otd}), and set the fiducial values of the other third-order bias parameters to zero. We use one standard deviation of the SBI posterior as step size for the derivatives of the parameters corresponding to the restricted case, and unity as step size for the other third-order bias parameters, motivated by the SBI priors. We also sum the Gaussian FBI priors to the Fisher matrix. We test convergence of the Fisher errors with respect to step size, fiducial point, number of simulations for covariance estimation and for the mean in the derivatives (not shown). We show in Fig. 11 the posterior obtained by SBI for one of the cases considered in this work and a comparison to a Fisher analysis using sample covariance. $\log(b_\epsilon)$ (with basis 10) is shown instead of b_ϵ since its prior is bounded at zero, while Fisher assumes Gaussianity of the parameter posteriors.
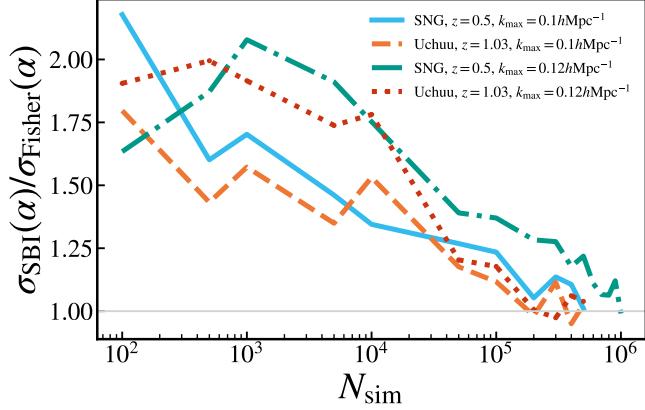
FIG. 10. Convergence of the standard deviations of α posteriors in SBI P+B analyses with increasing simulation budget N_{sim} used for posterior estimation. Each value is normalized by the corresponding α constraint from a Fisher analysis, for each of the cases considered.

5.   SBI P+B—Non-Gaussian stochasticity

The analysis described in the main text only includes the Gaussian contribution to galaxy stochasticity. In standard P+B analysis, one usually includes a density-dependent stochastic contribution and a non-Gaussian white noise for the bispectrum, i.e. $\delta_g^{\mathrm{stoch.}}(\boldsymbol{x}) = [b_\epsilon + b_{\epsilon\delta}\delta(\boldsymbol{x}) + b_{\nabla^2\epsilon}\nabla^2]\epsilon_0(\boldsymbol{x}) + b_{\epsilon^2}[\epsilon_0^2(\boldsymbol{x}) - \langle\epsilon_0^2\rangle]$, where ϵ_0 is a unit Gaussian random field (see App. A of [38]). Although the density-dependent noise is more important than the contribution $\propto k^2$ [41], it is technically more challenging to implement in FBI [23, 40]. In addition, it is usually the case that only the third-order bias operator that enters the power spectrum at one loop order [74] is constrained, which corresponds to the operator O_{td} in our bias parameterization.

We therefore consider an additional analysis where we fix all third-order bias parameters to zero with the exception of $b_{O\mathrm{td}}$, which is inferred together with the two additional stochastic contributions b_{ϵ^2} and $b_{\epsilon\delta}$. We show in Fig. 12 the results of such analysis, which leads to similar constraints on α than the case with only Gaussian noise. We apply convergence and calibration tests also to these posteriors (not shown).
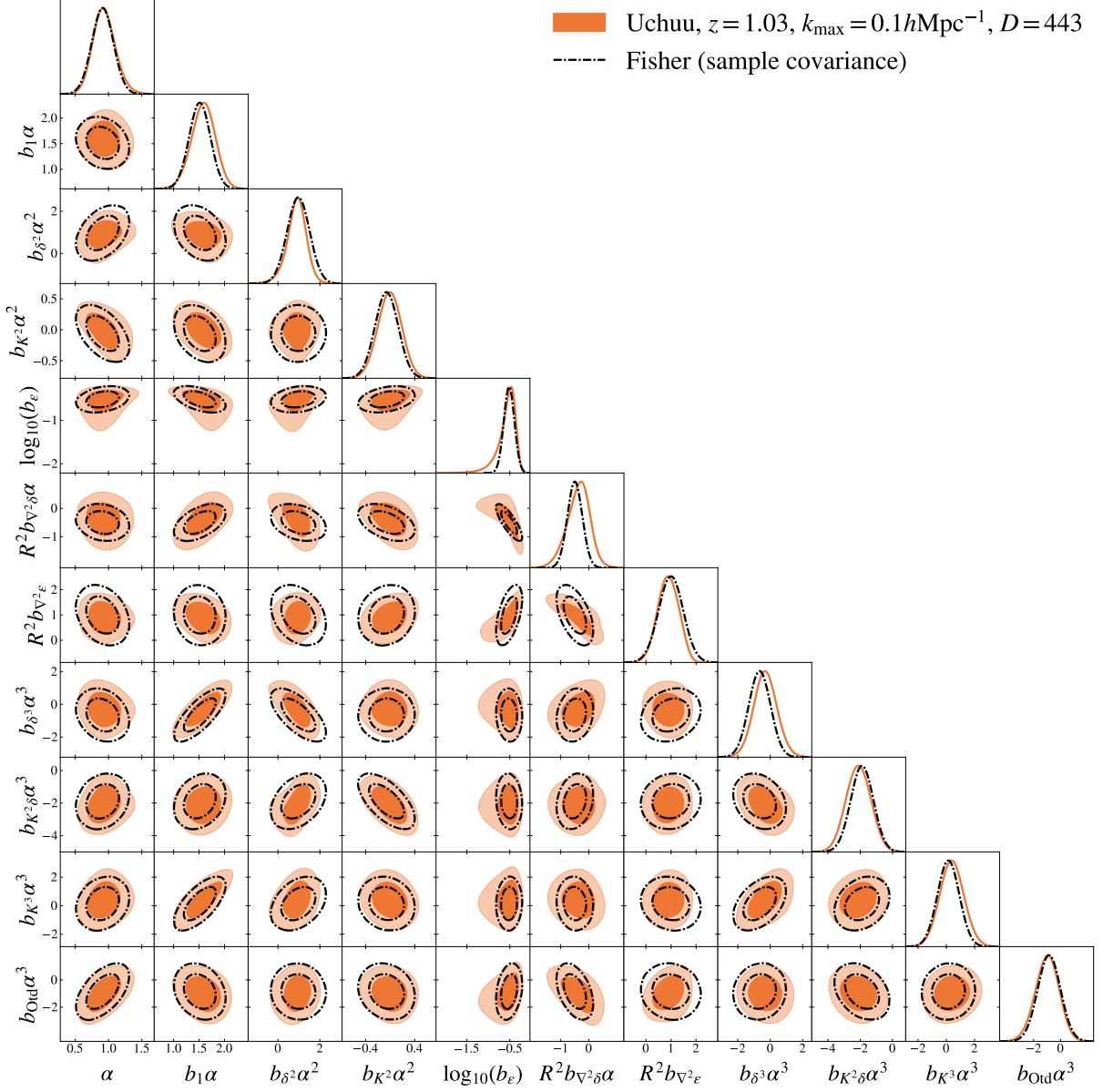
FIG. 11. Parameter posteriors obtained in the SBI P+B analysis of the Uchuu halo sample at redshift $z = 0.5$ with $k_{\max} = 0.1\,h\,\mathrm{Mpc}^{-1}$. Contours indicate 68% and 95% credible intervals. Dashed lines correspond to the Fisher analysis for this case, using sample covariance for the data vector.
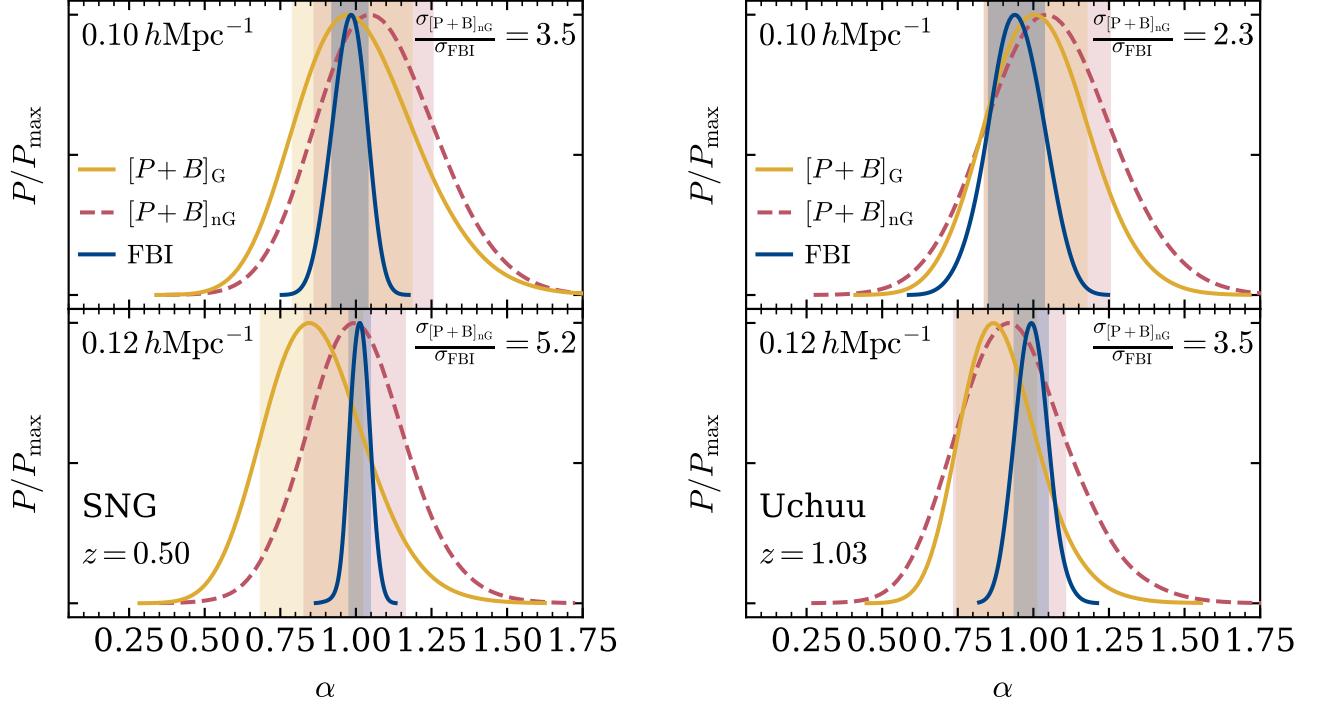
FIG. 12. Similar to Fig. 2 (left panels) and Fig. 3 (right panels) in the main text, but including the α posteriors for SBI P+B analyses that include non-Gaussian (nG) contributions to galaxy stochasticity, as described in Appendix B 5. Vertical bands indicate the 68% limits of the posteriors. The ratios of the 1-σ constraints between FBI (blue) and SBI $[P+B]_{\rm nG}$ (pink) are shown in the upper right corners.