

Differentiable fuzzy cosmic web for field-level inference

P. Rossello^{1,2,*}, F.-S. Kitaura^{1,2,*}, D. Forero Sánchez^{3,4}, F. Sinigaglia^{5,6,1,2}, and G. Favole^{1,2}

¹ Instituto de Astrofísica de Canarias, s/n, E-38205 La Laguna, Tenerife, Spain

² Departamento de Astrofísica, Universidad de La Laguna, E-38206 La Laguna, Tenerife, Spain

³ Departament de Física Quàntica i Astrofísica, Universitat de Barcelona, Martí i Franquís 1, E08028 Barcelona, Spain

⁴ Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (UB), c. Martí i Franquès, 1, 08028 Barcelona, Spain

⁵ Département d'Astronomie, Université de Genève, Chemin Pegasi 51, CH-1290 Versoix, Switzerland

⁶ Institut für Astrophysik, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

Received 4 June 2025 / Accepted 26 February 2026

ABSTRACT

Context. A comprehensive analysis of the cosmological large-scale structure derived from galaxy surveys involves field-level inference, which requires a forward-modelling framework that simultaneously accounts for structure formation and tracer bias.

Aims. While structure formation models are well understood, the development of an effective field-level bias model remains challenging, particularly in the context of tracer perturbation theory within Bayesian reconstruction methods, which we address in this work.

Methods. To bridge this gap, we developed a differentiable model that integrates augmented Lagrangian perturbation theory and non-linear, non-local, and stochastic biasing. At the core of our approach is the Hierarchical Cosmic-Web Biasing Nonlocal (HICOBIAN) model, which provides a description of a field with a positive number of tracers while incorporating a long- and short-range non-local framework via cosmic-web regions and deviations from Poissonity in the likelihood. A key insight of our model is that transitions between cosmic-web regions are inherently smooth, which we implemented using sigmoid-based gradient operations. This enables a fuzzy and, hence, differentiable hierarchical cosmic-web description, making the model well-suited for machine-learning frameworks.

Results. We tested the practical implementation of this model through GPU-accelerated computations implemented in JAX, the BRIDGE code, enabling a scalable evaluation of complex biasing. Our approach accurately reproduces the primordial density field within associated error bars derived from Bayesian posterior sampling within a self-specified setting (meaning that inference is performed on data generated by the exact same forward model) as validated by two- and three-point statistics in Fourier space. Furthermore, we demonstrate that the methodology approaches the maximum encoded information consistent with Poisson noise. We also demonstrate that the bias parameters of a higher order non-local-bias model can be accurately reconstructed within the Bayesian framework for bias models with eight parameters.

Conclusions. We introduce a Bayesian field-level inference algorithm that leverages the same forward-modelling framework used in galaxy, quasar, and Lyman-alpha-forest mock-catalogue generation – including non-linear, non-local and stochastic bias with redshift space distortions – providing a unified and consistent approach to the analysis of large-scale cosmic structure.

Key words. methods: analytical – methods: data analysis – methods: numerical – methods: statistical – dark matter – large-scale structure of Universe

1. Introduction

The primordial density fluctuations provide the initial conditions for structure formation and encode statistical signatures of the early Universe. These early inhomogeneities, seeded during the initial moments after the Big Bang, evolved under gravity to form the vast cosmic-web structure we observe today. Reconstructing the initial conditions from present-day observations of luminous tracers – such as galaxies, quasars, Lyman- α forests, intensity maps, or from the 21 cm line – offers a powerful probe of fundamental physics, including inflation, dark matter, dark energy, and gravity itself. The scientific community is making significant efforts to map the matter distribution in the Universe, as exemplified by current large-scale spectroscopic surveys such as the Dark Energy Spectroscopic Instrument (*DESI*) (Levi et al. 2013; *DESI Collaboration* 2025), *Euclid* (Amendola et al. 2016; *Euclid Collaboration: Aussel et al.* 2025); and future ones such

as the Prime Focus Spectrograph (*PFS*) (Takada et al. 2014), *MULTI*plexed Survey Telescope (*MUST*) (Zhao et al. 2024), and the Nancy Grace Roman Space Telescope (*Roman*) (Wang et al. 2022).

The first attempts to recover the primordial density fluctuations from the galaxy distribution were based on inverse reconstruction methods (e.g. Bertschinger & Dekel 1989; Nusser & Dekel 1992; Monaco & Efstathiou 1999). These approaches proved to be highly effective for sharpening features such as baryon acoustic oscillations (BAOs; Eisenstein et al. 2007), and were even proposed as tools to constrain primordial non-Gaussianities (Shirasaki et al. 2021). However, the accuracy of these inverse techniques is fundamentally limited by the non-invertibility of the gravitational evolution in the non-linear regime in the absence of full phase-space information, particularly due to shell-crossing. Once multiple streams of matter overlap, information about the initial conditions is irreversibly lost, as the full phase-space dynamics – crucial for a unique reconstruction – are no longer accessible through the evolved tracer distribution alone.

* Corresponding authors: pere.rossello@iac.es, fkitaura@ull.edu.es

To overcome the limitations of inverse methods, forward modelling was introduced as a principled alternative. Instead of inverting the non-linear structure formation process, forward approaches start from initial conditions and evolve them through a physical model of gravitational dynamics to generate predictions for present-day observables (Kitaura 2013; Jasche & Wandelt 2013; Wang et al. 2013). The first application of a forward-modelling reconstruction (including peculiar motions) to observational data was presented by Kitaura et al. (2012b). All these works demonstrated the viability of sampling high-dimensional posterior distributions consistent with both data and gravitational physics, laying the foundation for later advances in bias modelling and inference methods. Subsequent developments considered improved mask handling (Jasche et al. 2015; Lavaux & Jasche 2016; Ata et al. 2021), improved gravity solvers using particle-mesh codes (Wang et al. 2014; Jasche & Lavaux 2019; Horowitz et al. 2019b), and emulators (Doeser et al. 2024). Further extensions incorporated light-cone evolution (Kitaura et al. 2021; Lavaux et al. 2019), primordial non-Gaussianities (Andrews et al. 2023), effective field theory (EFT) approaches (Stadler et al. 2023), and hydrodynamics (Horowitz & Lukic 2025). These techniques have also been applied to alternative tracers, such as the Lyman- α forest (e.g. Kitaura et al. 2012c; Horowitz et al. 2019a; Porqueres et al. 2019; Horowitz et al. 2021, 2022). A benchmark of computational efficiency in this context is presented in Bayer et al. (2023).

Bayesian statistics offers the ideal framework for field-level inference, enabling clear specification of prior knowledge and observational uncertainties (Kitaura & Enßlin 2008). In this setting, reconstruction becomes a matter of sampling from the posterior distribution of initial conditions, informed by both the underlying physics and the observed distribution of tracers.

This approach, however, faces a significant computational barrier. High-fidelity dark-matter simulations capable of resolving the halos that host galaxies require substantial computational resources – often hundreds of thousands to millions of CPU hours – rendering brute-force posterior sampling impractical (see state-of-the-art N -body simulations, e.g. Garrison et al. 2018; Chuang et al. 2019; Ishiyama et al. 2021).

To mitigate this challenge, Lagrangian perturbation theory (LPT) offers a computationally efficient alternative to full gravity solvers (Bernardeau et al. 2002), such as N -body simulations (Angulo & Hahn 2022). While being approximate, LPT accurately captures the gravitational evolution of matter on large, quasi-linear scales and serves as an effective forward model. Additional advancements have further extended and regularised these approximations, enabling reliable predictions down to megaparsec – or even sub-megaparsec – scales (Kitaura & Heß 2013; Kitaura et al. 2024).

This strategy is particularly effective when the forward model is constrained to its domain of validity and coupled to the observed tracers through an effective field-level bias prescription. The core challenge then lies in constructing a robust and physically motivated bias model. A critical first step is to ensure that large-scale bias is accurately captured, which necessitates the inclusion of non-linear and non-local contributions up to the third order (McDonald & Roy 2009).

A longstanding limitation of perturbation-theory-based bias models is that truncation at a fixed order often introduces unphysical behaviour (McDonald & Roy 2009; Schmittfull et al. 2019; Werner & Porciani 2020). Specifically, these models can yield non-positive densities and exhibit oscillatory or noisy behaviour when extended into the highly non-linear regime.

However, recognising that non-local-bias terms at different orders correspond to distinct morphological features of the cosmic web opens the door to a more stable and flexible framework (Kitaura et al. 2022). For each morphological feature, we can then assume a local non-linear-bias model tailored to the specific properties of that structure. In this way, the classification of the cosmic web into distinct patterns effectively acts as a diagonalising operation, isolating the dominant local contributions to the tracer-matter relationship within each structural environment. This enables the systematic inclusion of higher order terms while preserving physical consistency and numerical stability down to small scales, making the framework well-suited for precision field-level inference. This ansatz provides a general and flexible framework applicable to a wide range of tracer populations and has demonstrated unprecedented accuracy in field-level bias modelling of point-like tracers such as dark-matter halos (Balaguera-Antolínez et al. 2023; Coloma-Nadal et al. 2024) and the baryonic components resolved in cosmological hydrodynamical simulations – including ionised gas, neutral hydrogen, and the Lyman- α forest (Sinigaglia et al. 2021, 2022, 2024b,a).

In this paper, we introduce BRIDGE¹, a differentiable, GPU-accelerated framework for field-level inference written in JAX (Bradbury et al. 2018). In Sect. 2, we detail the technical implementation of the structure formation and bias models. Section 3 presents a series of computational tests demonstrating that the code successfully recovers both the primordial density field and the bias parameters of a long-range non-local-bias model at resolutions of 5 and 10 h^{-1} Mpc. Finally, we describe how we applied BRIDGE using a combined long- and short-range bias model, showing that the framework can robustly recover the primordial density field even in the presence of highly complex tracer bias.

2. Methods

In this section we present the technical details of the BRIDGE code. A flowchart showing the architecture of BRIDGE is presented in Fig. 1.

2.1. Field-level inference

We first consider a galaxy distribution, mapped onto a comoving volume by converting redshifts to distances, as the observed tracer field. With $\mathbf{n}^* \in \mathbb{N}^{M^3}$, we denote the observed tracer intensity on a cubic, comoving grid consisting of N^3 cells. A differentiable forward model, \mathcal{M} , maps a white-noise realisation, \mathbf{v} , which defines the initial density field through a linear power spectrum, together with tracer bias parameters \mathbf{b} and cosmological parameters $\mathbf{\Omega}$ to the expected tracer field $\bar{\mathbf{n}}$:

$$\mathcal{M}: (\mathbf{v}, \mathbf{b}, \mathbf{\Omega}) \mapsto \bar{\mathbf{n}}. \quad (1)$$

The forward model, \mathcal{M} , is constructed as a composition of three sequential maps, $\mathcal{M} = \mathcal{M}_b \circ \mathcal{M}_\Psi \circ \mathcal{M}_\delta$, where \mathcal{M}_δ is a linear transformation that maps a white-noise realisation, \mathbf{v} , to the initial overdensity field, δ_0 , at high redshift using the linear matter power spectrum (see Appendix A for details on the input field parametrisation); \mathcal{M}_Ψ is the gravity solver that evolves the initial field, δ_0 , into the non-linear dark-matter over-density field, δ , at a specified redshift using an efficient gravity solver; and \mathcal{M}_b maps the evolved matter field, δ , together with the bias parameters, \mathbf{b} ,

¹ Bayesian Reconstruction and Inference of Data-driven Generative Environments: <https://github.com/pererossello/bridge-repo>.

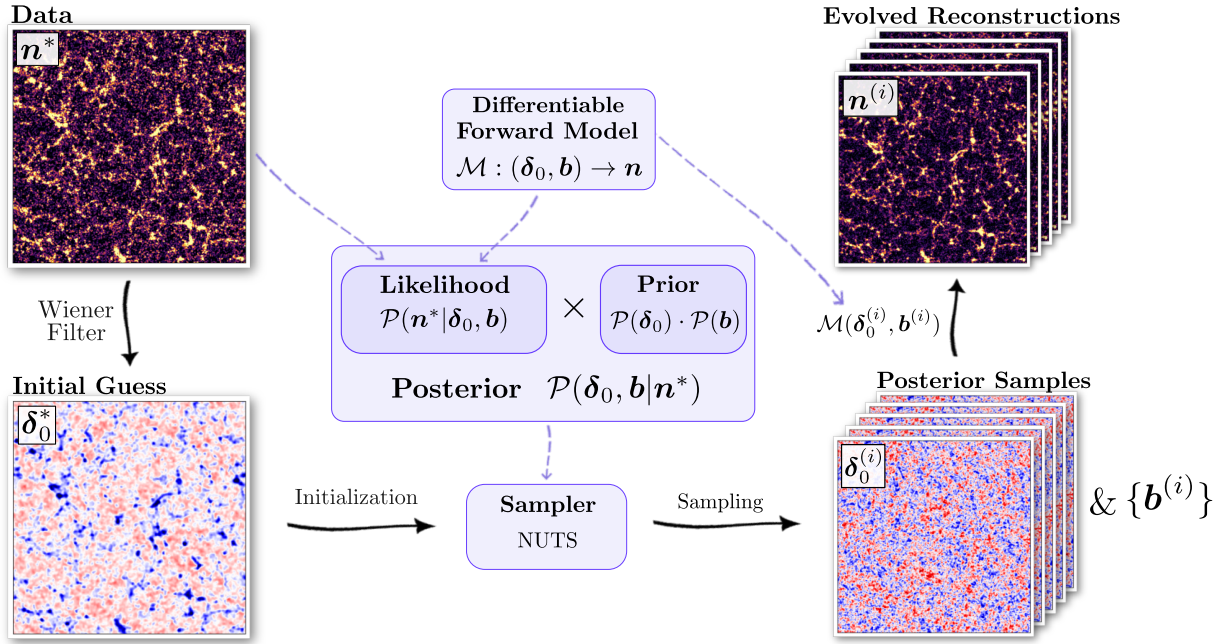


Fig. 1. Schematic overview of the BRIDGE pipeline. The framework combines a scalable, differentiable structure-formation model with a flexible field-level bias prescription, all implemented in a GPU-accelerated JAX environment. This enables efficient Bayesian inference of the primordial density field from tracer observations, with support for complex, non-local-bias models and multi-resolution analysis. The modular design allows for the seamless integration of physical models while maintaining end-to-end differentiability and high computational performance.

to the mean tracer field, \bar{n} , employing a cosmic-web-dependent, non-local-bias model.

For each cell, we assume the data arise from an independent discrete distribution with mean \bar{n}_i , which may include additional parameters, \mathbf{p} , modelling survey completeness or super-Poissonian dispersion. The likelihood, which incorporates the forward model, reads

$$\mathcal{L}(\mathbf{v}, \mathbf{b}, \mathbf{p}, \boldsymbol{\Omega}) \equiv \mathcal{P}(\mathbf{n}^* | \mathbf{v}, \mathbf{b}, \mathbf{p}, \boldsymbol{\Omega}), \quad (2)$$

where \mathcal{P} denotes the probability of observing \mathbf{n}^* given the model parameters.

Performing Bayesian inference at the field level amounts to sampling from the posterior distribution of the latent initial conditions and model parameters, given the observed tracer field \mathbf{n}^* . This posterior is given by

$$\mathcal{P}(\mathbf{v}, \mathbf{b}, \mathbf{p}, \boldsymbol{\Omega} | \mathbf{n}^*) \propto \mathcal{P}(\mathbf{v}) \mathcal{P}(\mathbf{b}, \mathbf{p}, \boldsymbol{\Omega}) \mathcal{L}(\mathbf{v}, \mathbf{b}, \mathbf{p}, \boldsymbol{\Omega}), \quad (3)$$

where $\mathcal{P}(\mathbf{v})$ encodes prior information on the initial conditions, and $\mathcal{P}(\mathbf{b}, \mathbf{p}, \boldsymbol{\Omega})$ represents prior knowledge on the bias and cosmological parameters. The forward model was constructed such that the latent field, \mathbf{v} , is whitened, and therefore $\mathcal{P}(\mathbf{v})$ is an N^3 -dimensional multi-variate Gaussian with zero mean and identity covariance. The bias and additional likelihood model parameters, \mathbf{b} and \mathbf{p} , which in our case were strictly positive, were assigned independent log-normal priors. The probability density function of the log-normal distribution for the bias parameters implemented in the BRIDGE code is given by

$$f(b; \mu, \sigma) = \frac{1}{b\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln b - \mu)^2}{2\sigma^2}\right], \quad b > 0, \quad (4)$$

where μ and σ are fixed hyperparameters that define the mean and standard deviation of $\ln b$. Throughout the remainder of this work, we fixed the cosmological parameters, $\boldsymbol{\Omega}$, to their fiducial values and omitted their explicit dependence for notational simplicity.

2.2. Posterior sampling

We exploited the differentiability of the forward model to efficiently sample the high-dimensional posterior using gradient-based Hamiltonian Monte Carlo (HMC) sampling (Duane et al. 1987). HMC was first applied to Bayesian large-scale-structure inference by Jasche & Kitaura (2010) and subsequently to observational data by Jasche et al. (2010).

We implemented within the BRIDGE code various posterior sampling methods, such as second- and fourth-order HMC sampling and the Blanes et al. (2014) HMC variant. The numerical tests shown in this work were finally performed with the publicly available No-U-Turn Sampler (NUTS) in NumPyro (Hoffman & Gelman 2011; Phan et al. 2019). During the adaptation (burn-in) phase, the algorithm tunes the step size, trajectory length, and the mass matrix, which is kept diagonal due to memory constraints. During sampling, we fixed the adapted mass matrix and step size and integrated the trajectories with a NUTS tree depth equal to 10.

To accelerate convergence, each chain was initialised from a deterministic state obtained by Wiener-filtering the observed tracer density contrast (Zaroubi et al. 1995), $\delta_{\text{tr}}^* \equiv \mathbf{n}^* N^3 / N_{\text{tr}} - 1$, back to an approximate initial density field, $\delta_0^{(0)}$, which was transformed into white noise, $\mathbf{v}^{(0)}$; N_{tr} is the total number of tracers. In this first step, we assumed the linear power spectrum to be known and ignored peculiar motions.

2.3. Gravity solver

We modelled large-scale-structure (LSS) formation with Augmented Lagrangian perturbation theory (ALPT; Kitaura & Heß 2013), which merges second-order LPT (2LPT; Zel'dovich 1970; Buchert 1994; Bouchet et al. 1995; Catelan 1995) for long-range tidal displacements and a spherical-collapse (SC) solution (Bernardeau 1994; Neyrinck 2016) for short-range

dynamics. This was achieved by decomposing the particle displacement field as implemented in the BRIDGE code:

$$\Psi = \Psi_L + \Psi_S, \quad (5)$$

with $\Psi_L = \mathcal{K}(q, r_s) \circ \Psi_{2\text{LPT}}$, $\Psi_S = (1 - \mathcal{K}) \circ \Psi_{\text{SC}}$, and \mathcal{K} a smoothing kernel of fixed radius, $r_s = 4 h^{-1} \text{Mpc}$, following Kitaura & Heß (2013). In future work, we plan to sample this parameter within the Bayesian framework and consider nLPT for the long-range interaction.

This hybrid formulation retains analytic dependence on cosmological parameters while suppressing shell crossing. Coupled with empirical bias schemes such as the Bias Assignment Method (BAM). (Balaguera-Antolínez et al. 2018, 2020), ALPT reproduces N -body halo statistics down to $k \lesssim 0.4 h \text{Mpc}^{-1}$, the scale at which most cosmological information resides (Hahn & Villaescusa-Navarro 2021). An even higher accuracy was achieved with the HICOBAN bias model considered in this work (see Sect. 2.7).

We note that ALPT is substantially more efficient in terms of both computational speed and memory usage compared to fast particle mesh (PM) methods (e.g. Tassev et al. 2013; Feng et al. 2016; Klypin & Prada 2018), as shown in the comparison work by Blot et al. (2019), making it especially advantageous for scalable Bayesian inference frameworks. Moreover, ALPT can be extended via iteration (eALPT) to reach sub-megaparsec accuracy (Kitaura et al. 2024) by indirectly emulating the effect of a viscous stress tensor in the equations of motion, thereby capturing aspects of non-linear dynamics such as vorticity. Recent GPU-based PM methods have shown noteworthy advancements (e.g. Modi et al. 2021; Li et al. 2024). Alternative promising approaches to approximate structure formation are offered by emulators trained on large ensembles of simulations (Kodi Ramanah et al. 2020; Conceição et al. 2024). However, for the time being, we opt to rely on analytic solutions with explicit cosmological dependence.

2.4. Bias model framework

When employing an approximate gravity solver, the evolved density field should not be directly trusted at the particle level. Instead, it is only reliable on a coarse mesh with a resolution of a few megaparsecs, where the approximation remains accurate. As a result, the bias relation between matter and tracers is no longer deterministic – as in approaches that apply a halo finder directly to particle distributions – it must be modelled as an effective field-level bias (Schmittfull et al. 2019). In this framework, the luminous tracers of LSS exhibit a complex, non-linear, non-local, and stochastic relationship with the underlying matter field (McDonald & Roy 2009; Desjacques et al. 2018). Accurately capturing this relationship is essential for unbiased inference and requires flexible and physically motivated modelling approaches that operate at the field level.

We took advantage of a positive, non-local-bias modelling framework that employs local-bias expansions within distinct morphological regions of the cosmic web (Kitaura et al. 2022; Coloma-Nadal et al. 2024; Sinigaglia et al. 2024b).

2.5. Deterministic local-bias model

The bias relation between tracers and the underlying matter field has long been known to be non-linear. One of the earliest approaches involved a local perturbative expansion of the tracer

overdensity, introduced by Fry & Gaztanaga (1993):

$$\delta_{\text{tr}} = \sum_{m=0}^{\infty} \frac{b_m}{m!} \delta^m, \quad (6)$$

where δ_{tr} denotes the tracer overdensity, δ the matter overdensity, and b_m the bias coefficients. While conceptually straightforward, this expansion suffers from the drawback that it may yield unphysical negative tracer densities when truncated at any order.

To address this issue, an alternative logarithmic formulation was proposed around the same time by Cen & Ostriker (1993):

$$\log(1 + \delta_{\text{tr}}) = \sum_{m=0}^{\infty} \frac{c_m}{m!} [\log(1 + \delta)]^m, \quad (7)$$

which ensures positivity of the tracer field by construction. To a linear order, this expansion corresponds to a power-law bias model, which has been particularly effective in modelling the Lyman- α forest. In this context, it leads to the fluctuating Gunn–Peterson approximation, where the optical depth, τ , is related to the matter density field by $\tau \propto (1 + \delta)^\alpha$ (Bi & Davidsen 1997). An analogous relationship was found for the functional dependency between the molecular gas temperature and dark-matter density (Hui & Gnedin 1997).

Interestingly, a power-law bias model in Eulerian space can be related to a linear bias in Lagrangian space – i.e. in the coordinate system of the initial conditions – according to the continuity equation, which leads to a log-normal evolution of the density field (Coles & Jones 1991). In this picture, a linear Lagrangian bias naturally evolves into a power-law bias at later times.

In the context of galaxy bias, the polynomial perturbative expansion has historically been preferred. A foundational physical interpretation of bias was introduced even earlier by Kaiser (1984) through the excursion set model, where galaxies preferentially form in high-density peaks of the matter distribution. Building on this idea, Kitaura et al. (2014) proposed a hybrid model that combines the power-law bias of Cen & Ostriker (1993) with the threshold bias of Kaiser (1984) to model galaxy number counts as a function of the underlying dark-matter field. This framework was later refined by Neyrinck et al. (2014), the authors of which replaced the hard threshold with an exponential cutoff. The resulting model captures the average bias behaviour of halos with remarkable accuracy, particularly in low-density environments.

Although the power-law and threshold bias models are degenerate with respect to two-point statistics – both amplifying power on large scales – this degeneracy can be broken using three-point statistics, which provide a means to distinguish between different biasing mechanisms (see also Kitaura et al. 2015 for the definitions of three-point statistics calculations used in this work). The hybrid model has proven especially useful in this regard and was employed to generate thousands of mock catalogues of luminous red galaxies for the final BOSS data release (see the PATCHY mocks in Kitaura et al. 2016b).

To date, the suppression of tracer densities has primarily focused on low-density regions, following the picture introduced by Kaiser. However, another important effect in the context of halo and galaxy formation is halo exclusion (Baldauf et al. 2013), which accounts for the saturation of tracer objects in highly overdense environments. This mechanism introduces a natural suppression of tracer abundance in high-density regions, complementing the low-density suppression from the original threshold bias picture (Coloma-Nadal et al. 2024). A similar concept has also been applied in modelling the Lyman- α forest,

where saturation effects in flux absorption become significant in dense regions (Sinigaglia et al. 2024b).

Motivated by these considerations, we modelled and implemented the expected tracer number density in the BRIDGE code using a power-law component modulated by both a low-pass and a high-pass filter:

$$\bar{n}_i = C (1 + \delta_i)^\alpha \exp \left[- \left(\frac{1 + \delta_i}{\rho} \right)^\epsilon \right] \exp \left[- \left(\frac{1 + \delta_i}{\rho'} \right)^{-\epsilon'} \right], \quad (8)$$

where C is a normalisation constant defined such that $\sum_i \bar{n}_i = \sum_i n_i^* \equiv N_{\text{tr}}$, where N_{tr} is the total number of tracers in the data. All parameters are positive. We assume that the observed tracer count in each cell is a stochastic realisation drawn from a discrete probability distribution, $\mathcal{P}_i^{\text{tr}}$, characterised by a mean expected value, \bar{n}_i , provided by the forward model.

2.6. Stochastic bias model

The coarse resolution inherent to the field-level framework necessitates the introduction of a stochastic component to the bias model (Dekel & Lahav 1999; Sheth & Lemson 1999), which is not explicitly present in deterministic high-resolution N -body simulations. Nonetheless, the lack of full phase-space information after shell-crossing also introduces an uncertainty that requires a probabilistic treatment in the likelihood.

The most basic discretisation assumes a Poisson likelihood, where the variance equals the mean tracer count. The complete integration of this likelihood beyond the second moment within the Bayesian inference framework has been studied for a long time (Kitaura & Enßlin 2008; Kitaura et al. 2010). However, this assumption holds only in very limited regimes – typically for low tracer densities and in relatively homogeneous environments. As originally predicted by Peebles (1980), (anti-)correlations between tracers below the mesh resolution introduce deviations from pure Poisson statistics. These unresolved sub-grid correlations give rise to either sub- or super-Poissonian noise characteristics. To account for this, various probability distribution functions (PDFs) have been proposed in the literature (Saslaw & Hamilton 1984; Sheth 1995), most of which introduce additional degrees of freedom to model the over-dispersion commonly observed in galaxy and halo distributions relative to the Poisson expectation (Somerville et al. 2001; Casas-Miranda et al. 2002). Pellejero-Ibañez et al. (2020) demonstrated that the negative binomial distribution can be parametrised to reproduce a wide range of functional dependencies in the over-dispersed deviations from Poisson statistics.

We implemented the negative binomial distribution to model $\mathcal{P}_i^{\text{tr}}$, thereby accounting for potential over-dispersion in the tracer counts. This distribution was first considered in the context of field-level bias in Kitaura et al. (2014) and later implemented in the Bayesian framework in Ata et al. (2015). Our implementation of the negative binomial distribution in this work relies on the gamma–Poisson mixture modelling, in which the Poisson rate, λ_i , is treated as a random variable drawn from a gamma distribution. The corresponding hierarchical formulation implemented in the BRIDGE code is given by

$$\lambda_i \sim \text{Gamma}(\beta, \bar{n}_i/\beta) \quad (9)$$

$$n_i^* \sim \text{Poisson}(\lambda_i) \quad (10)$$

where \bar{n}_i denotes the expected tracer count in cell i , and β is the over-dispersion parameter. In the limit $\beta \rightarrow \infty$, the gamma distribution becomes sharply peaked, recovering the standard Poisson

likelihood. In the gamma distribution, the first argument is the shape parameter and the second argument is the scale parameter. With this model, the (negative log) likelihood reads

$$- \log \mathcal{L}(\mathbf{v}, \mathbf{b}, \beta) \equiv \sum_{i=1}^{N^3} - \log \mathcal{P}_i^{\text{tr}}(n_i^* | \mathcal{M}(\mathbf{v}, \mathbf{b}), \beta), \quad (11)$$

with $\mathbf{b} = (\alpha, \rho, \epsilon, \rho', \epsilon')$ from Eq. (8). While sub-Poissonian behaviour is indirectly captured by the high-density damping exponential factor, an explicit treatment can be implemented using the Conway–Maxwell–Poisson distribution (Daly & Gaunt 2016) or extensions of the binomial distribution (Vos-Ginés et al. 2024).

2.7. Non-local bias and cosmic-web dependence

In the previous sections, we introduced a local non-linear-bias model to describe the average bias relation and subsequently incorporated scatter to account for stochasticity. However, this stochastic component should be understood as a feature of the coarse-grained nature of the effective bias approach, without superseding deterministic contributions from non-local-bias effects.

Galaxy bias is shaped not only by the local matter density, but also by the surrounding large-scale tidal environment and the small-scale curvature of the density field. These non-local dependencies stem from the intrinsic nature of gravitational collapse and galaxy formation. A general framework for incorporating such effects into bias expansions was introduced by McDonald & Roy (2009). In another particularly relevant study, Chan et al. (2012) connected long-range tidal tensor invariants to higher order bias terms. Kitaura et al. (2022) later related the invariants of the Hessian of a field to the morphological classifications of the cosmic web.

These developments motivated the extension of bias models to explicitly include cosmic-web morphology (see Appendix B). Galaxies inhabiting different large-scale structures, including voids, sheets, filaments, and knots, exhibit systematically different biasing behaviours, as shown in simulations and observations (e.g. Nuza et al. 2014; Filho et al. 2015).

Several approaches exist for incorporating cosmic-web morphology. One method involves binning tidal field tensor invariants, which can result in a high-dimensional classification if multiple invariants and their dependence on local density are included (Kitaura et al. 2022). Alternatively, one can compress this information into four characteristic environments, sacrificing some modelling granularity for interpretability – this leads to the so-called Φ -web classification, based on the large-scale tidal field (Hahn et al. 2007).

To recover finer biasing distinctions, we introduced a hierarchical structure by embedding a second classification based on the Hessian of the density field (the δ -web) within each Φ -web region. This short-range non-local bias captures small-scale tidal effects rooted in tidal torque theory (Heavens & Peacock 1988). The result is the HICOBAN model, which partitions the space into 16 regions (Coloma-Nadal et al. 2024). For each region, it applies a different local-bias relation, which is determined by the specific bias parameters of model Eq. (8). This approach gains physical insight with computational efficiency and outperforms fine binning of the tidal field in terms of both precision and interpretability, as both short- and long-range non-local biases are incorporated.

Given that gravitational evolution is inherently non-local, the shortcomings of ALPT relative to full gravity solvers

can be effectively absorbed into non-local-bias terms. This explains the excellent performance of more recent studies employing ALPT (e.g. Balaguera-Antolínez et al. 2023; Coloma-Nadal et al. 2024). This insight provides a foundation for extending the bias framework to capture deviations arising from alternative gravity models (García-Farieta et al. 2024), where characteristic bias parameters and higher order statistics including Legendre multipole expansions would serve as indicators of such modified dynamics.

In the context of this work, it is important to note that most cosmic-web classifiers rely on hard, non-differentiable thresholds (e.g. eigenvalue sign changes), which are incompatible with gradient-based inference techniques such as HMC sampling. To address this, we adopted a fuzzy, differentiable classification scheme using sigmoid-based transitions. This enables smooth, physically motivated boundaries between web types and integrates seamlessly with our differentiable, field-level inference framework.

2.8. Differentiable fuzzy cosmic-web classification

In the top-down structure-formation scenario proposed by Zel'dovich (1970), cosmic structures form via anisotropic collapse governed by the tidal field tensor. The process unfolds along the tensor's eigenvectors, with collapse first occurring along the direction associated with the largest eigenvalue – a phenomenon known as pancake formation. The term cosmic web was coined by Bond et al. (1996) to describe the large-scale structure of the Universe, characterised by the emergence of distinct morphological components such as filaments, walls, and nodes arising from anisotropic gravitational collapse. This interpretation was supported by shape parameters such as ellipticity and prolateness, derived from the eigenvalues of the deformation tensor and linked to the Zel'dovich approximation. A more systematic and operational classification of the cosmic web was later proposed by Hahn et al. (2007), the authors of which used the eigenvalues of the tidal (equal to the velocity shear within the Zel'dovich approximation) tensor to categorise regions of space into knots, filaments, sheets, and voids, depending on the number of collapsing directions. This method provided a quantitative tool to identify web environments in cosmological simulations.

While widely used, cosmic-web classification schemes remained largely phenomenological, lacking a firm quantitative foundation. A more rigorous interpretation began to emerge with the connection to the framework of non-local bias, which provided a physically motivated context for understanding how large-scale structure influences halo and galaxy formation – an approach we adopted in this work (see Appendix B).

We now consider a cosmic-web classification scheme based on the tidal field tensor, $\partial_j \partial_k \Phi$, where $\Phi = \nabla^{-2} \delta$ is the gravitational potential. Let $\lambda_i^{(1)} < \lambda_i^{(2)} < \lambda_i^{(3)}$ be the eigenvalues of the tidal field tensor at the i th cell. The cosmic-web classification at the i th cell can be compactly expressed as

$$W_i = \sum_{n=1}^3 h(\lambda_i^{(n)} - \lambda_{\text{th}}), \quad (12)$$

where h is the Heaviside step function and λ_{th} is a free parameter. In this work, we set $\lambda_{\text{th}} = 0.05$ (see Forero-Romero et al. 2009 for the introduction of a threshold in the cosmic-web classification). For recent studies regarding the choice of λ_{th} , see Coloma-Nadal et al. (2024) and Olex et al. (2025). W_i counts the number of eigenvalues above the threshold; the four resulting values of 0, 1, 2, and 3 correspond to voids, sheets, filaments, and

knots, respectively. This cosmic-web definition is problematic if introduced in a forward model that aims to be differentiable with respect to ν . This is because the cosmic-web classification mesh, \mathbf{W} , depends on ν through the potential of $\delta = \mathcal{M}_\Psi \circ \mathcal{M}_\delta(\nu)$, but it does so in a manifestly non-differentiable way due to the presence of the Heaviside step function. As a consequence, the gradients $\partial W_i / \partial q_j$ are ill-defined, which precludes the use of gradient-based sampling methods. To address this, we relied within the BRIDGE code on a fuzzy cosmic-web classification defined through the sigmoid weights, with

$$w_i^{(n)} = \sigma[s(\lambda_i^{(n)} - \lambda_{\text{th}})], \quad (13)$$

where $\sigma(x) = (1 + e^{-x})^{-1}$, and s is the steepness parameter controlling how sharply $w_i^{(n)}$ transitions around λ_{th} . Because σ is strictly increasing and the eigenvalues are sorted, the weights inherit the order $w^{(1)} > w^{(2)} > w^{(3)}$. We converted these cumulative exceedance probabilities into mutually exclusive ‘fuzzy’ memberships for knots (K), filaments (F), sheets (S), and voids (V):

$$p_i^{(\text{K})} = w_i^{(3)}, \quad (14)$$

$$p_i^{(\text{F})} = w_i^{(2)} - w_i^{(3)}, \quad (15)$$

$$p_i^{(\text{S})} = w_i^{(1)} - w_i^{(2)}, \quad (16)$$

$$p_i^{(\text{V})} = 1 - w_i^{(1)}, \quad (17)$$

so that they represent the probabilities of having exactly three, two, one, or zero eigenvalues above the threshold, and, by construction, $p_i^{(\text{V})} + p_i^{(\text{S})} + p_i^{(\text{F})} + p_i^{(\text{K})} = 1$. After introducing this smooth classification scheme, the forward model remained differentiable. Regarding the steepness parameter s , a larger value makes the fuzzy classification closer to the original classification, but it also risks causing numerical instabilities in the gradients. We therefore chose the largest s that preserves stable gradient computations. In Fig. 2, we show a comparison of the ‘sharp’ and smooth cosmic-web classification.

We now consider a cosmic-web-dependent bias model where each of the bias parameters from the local-bias model depend on the cosmic-web classification of the voxel. That is, given a bias parameter of $b \in \{\alpha, \rho, \epsilon, \rho', \epsilon'\}$, it now takes different values in each of the cosmic-web regions: $b^{(\text{K})}, b^{(\text{F})}, b^{(\text{S})}, b^{(\text{V})}$. To ensure differentiability, we used the fuzzy cosmic-web classification, and the bias parameter at voxel i takes the value

$$b_i = b^{(\text{K})} p_i^{(\text{K})} + b^{(\text{F})} p_i^{(\text{F})} + b^{(\text{S})} p_i^{(\text{S})} + b^{(\text{V})} p_i^{(\text{V})}. \quad (18)$$

2.9. Redshift-space-distortion modelling

Redshift-space distortions (RSDs) arise when the observed position of a tracer deviates from its Hubble-flow distance due to peculiar velocities, making them sensitive probes of the growth rate of structure formation (see Kaiser 1987; Hamilton 1998). Redshift-space distortions are governed by two distinct types of peculiar motions: coherent large-scale flows, which arise from gravitational infall towards overdense regions, and small-scale random motions within quasi-virialised structures. For a variety of ways of modelling coherent flows, we refer the reader to Kitaura et al. (2012a) and references therein (see also Wang et al. 2012). There are several strategies for incorporating RSDs within a Bayesian framework (we chose the third), which we list below.

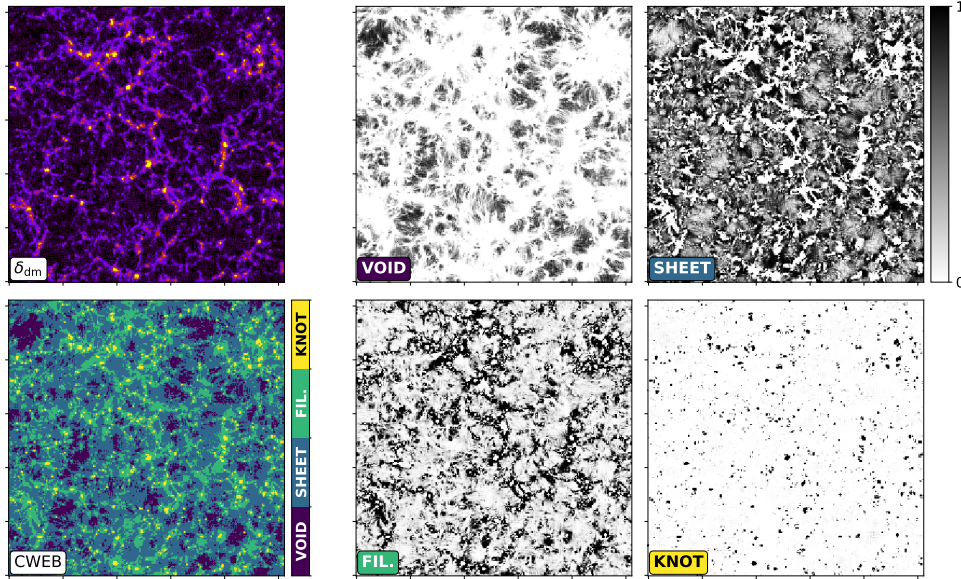


Fig. 2. Example of fuzzy cosmic-web classification with $N = 256$ and $\Delta L = 1.7 h^{-1} \text{ Mpc}$. Top left: Evolved dark-matter density contrast field. Bottom left: ‘Hard’ Φ -web classification. Right panels: Fuzzy membership weights, $p_i^{(V)}$, $p_i^{(S)}$, $p_i^{(F)}$, and $p_i^{(K)}$, for voids, sheets, filaments, and knots, respectively, as defined in Eq. (17).

1. Iterative sampling of the redshift-to-real-space mapping. This approach iteratively reconstructs the real-space distribution from redshift-space observations using Gibbs sampling (Kitaura & Enßlin 2008; Kitaura et al. 2012c, 2016a). It is inspired by earlier iterative RSD correction methods (see Yahil et al. 1991; Monaco & Efstathiou 1999).
2. Forward modelling of RSDs at the tracer level. Originally introduced by Kitaura et al. (2012b), this method evolves tracers from sampled initial conditions to redshift space, where their final positions are compared with observed galaxy distributions to constrain the primordial density field. Small-scale virial motions can be addressed in two ways: either by collapsing elliptical groups of galaxies – as proposed by Tegmark et al. (2004) – and modelling only coherent flows, as in the original study, or by retaining fingers-of-God features and modelling them with a stochastic velocity dispersion component, as done by Heß et al. (2013).
3. Forward modelling of RSDs at the dark-matter-field level. In this work, we followed the approach of mapping the matter field directly to redshift space in a fully differentiable Bayesian inference framework. We set the observer at the center of the cubical volume and computed the RSD effects accordingly along lines of sight. See Bos et al. (2019) for a first implementation of such a method and calculation details.

We note that the modelling of RSDs for biased tracers involving non-linear transformations – such as cosmic voids and Lyman- α forests – requires the inclusion of velocity bias. For the Lyman- α forest, RSD affects the optical depth, τ , but the observable is the transmitted flux $F = \exp(-\tau)$, introducing a non-trivial mapping. Accurate modelling in these cases must account for velocity-bias effects (McDonald et al. 2000; Seljak 2012; Sinigaglia et al. 2022, 2024b). Similarly, cosmic voids are identified from the galaxy distribution in redshift space, which effectively applies a non-linear and non-local transformation (Chuang et al. 2017). For the time being, we restrict the implementation tests within the BRIDGE code to unbiased peculiar motions ($b_v = 1$); i.e. the dark-matter-density contrast in redshift space is given

by

$$\delta(s) = \delta\left(\mathbf{r} + b_v \frac{\mathbf{v} \cdot \mathbf{r}}{aHr^2} \mathbf{r}\right), \quad (19)$$

with s and \mathbf{r} being the redshift- and real-space coordinates, respectively; \mathbf{v} the peculiar velocity; a the scale factor; and H the Hubble constant for a given Hubble distance.

3. Validation of field-level inference with fuzzy cosmic-web bias

To validate the BRIDGE code, we considered three numerical tests with different resolutions and bias models: TEST1, TEST2, and TEST3 (see Table 1). Results of TEST2 and TEST3 are presented in Appendix C. The numerical tests consist of the following steps.

1. Generate a ground-truth number-count catalogue of biased tracers from an initial Gaussian density field with the same forward model employed during inference. All runs are performed to redshift $z = 0$ with a particular random seed and specific bias parameters as listed in Tables 2 and 3.
2. Perform field-level inference of the posterior describing the white-noise and bias parameters for TEST1 and TEST2, applying the BRIDGE code.
3. Perform quality assessment of the reconstructions by evaluating the recovery of both the initial and final conditions using two-point and three-point statistical analyses.

Regarding the hyperparameters of the bias parameter prior of Eq. (4), the power-law (α) and negative-binomial (β) hyperparameters are $\mu_\alpha = 0.27$, $\sigma_\alpha = 0.30$, $\mu_\beta = 2.12$, and $\sigma_\beta = 0.20$.

3.1. Motivation of the numerical tests

The numerical tests were chosen to demonstrate a number of scientific relevant cases. We list these below.

- TEST1 and TEST2: The non-local-bias description used here is based on the Φ -web classification. Balaguera-Antolínez et al. (2018) demonstrated that this

Table 1. Setting the different numerical cases.

	TEST1	TEST2	TEST3
ΔL [h^{-1} Mpc]	10	5	8
L [h^{-1} Mpc]	1280	640	1024
k_{nyq} [$h \text{Mpc}^{-1}$]	0.31	0.63	0.39
$k_{\text{nyq}}^{\text{non-iso}}$ [$h \text{Mpc}^{-1}$]	0.54	1.10	0.68
Non-local bias	$\Phi\text{-web}_j$	$\Phi\text{-web}_j$	$\Phi\delta\text{-web}_{jk}$
Local bias	$\{\alpha_j, \beta_j\}$	$\{\alpha_j, \beta_j\}$	$\{\alpha_{jk}, \beta_{jk}, \epsilon_{jk}, \rho_{jk}\}$

Notes. In all cases, we used a mesh of 128^3 voxels with cubical volumes of distinct side lengths, L , and resolutions, $\Delta L = L/N$. We consider the Nyquist frequency $k_{\text{nyq}} = \pi/\Delta L$ and the non-isotropic Nyquist frequency $k_{\text{nyq}}^{\text{non-iso}} = \sqrt{3} k_{\text{nyq}}$ of a cubical Fourier-space mesh. The output redshift is $z = 0$. There are 16 combinations of Φ - and δ -web regions: $j, k \in [1, 2, 3, 4]$.

Table 2. Values of bias parameters by cosmic-web region of the reference for the TEST1 and TEST2 cases (8 bias parameters).

Test		Voids	Sheets	Filaments	Knots
TEST1	α	1.10	1.03	1.15	1.23
	β	10.10	13.70	12.10	14.40
TEST2	α	1.05	1.11	1.23	1.30
	β	7.10	8.70	8.10	9.40

Table 3. Values of bias parameters by combined cosmic-web regions for the TEST3 case (64 bias parameters).

Φ -web	δ -web	α	β	ϵ	ρ
Voids	Voids	1.29	5.80	1.02	0.22
	Sheets	1.05	14.30	1.06	0.26
	Filaments	1.14	11.87	0.97	0.16
	Knots	1.34	10.60	1.00	0.20
Sheets	Voids	1.04	10.74	1.00	0.20
	Sheets	1.06	11.57	0.97	0.17
	Filaments	1.11	12.59	0.98	0.18
	Knots	1.05	13.67	1.06	0.26
Filaments	Voids	0.94	5.71	1.02	0.22
	Sheets	1.13	13.42	0.93	0.13
	Filaments	1.27	7.75	0.96	0.16
	Knots	1.20	11.21	1.01	0.21
Knots	Voids	1.18	8.36	0.99	0.19
	Sheets	1.01	11.24	0.96	0.16
	Filaments	1.07	10.99	0.94	0.14
	Knots	1.17	9.71	1.08	0.28

Notes. We show sub-divisions of δ -web types within each Φ -web classification (see Appendix B and Coloma-Nadal et al. 2024).

model can accurately reproduce key statistical properties of the halo distribution, including one-point PDFs, two-point power spectra at percent-level accuracy, and bispectra with reasonable agreement, generally within 15% (see also Balaguera-Antolínez et al. 2020, Pellejero-Ibañez et al. 2020, and Balaguera-Antolínez et al. 2023, who confirm these results for different halo resolutions). These studies confirm the Φ web as a physically meaningful and effective framework for modelling non-local bias within a reasonable degree of accuracy through non-parametric approaches.

However, for field-level inference applications – such as the one developed here – a parametric bias model is required to enable efficient sampling and marginalisation within a Bayesian framework. The same non-local-bias treatment incorporating an equivalent non-linear-bias prescription to the one considered here (see Eq. 8) was employed by Sinigaglia et al. (2024b,a) to model the Lyman- α forest. Their work demonstrated excellent agreement with high-resolution hydrodynamical simulations, achieving power-spectrum accuracy within 5% up to $k \sim 1 h \text{Mpc}^{-1}$, along with consistent bispectrum predictions. In both cases, i.e. haloes and Lyman- α forests, the underlying structure-formation model was ALPT. For both the TEST1 and TEST2 cases, we jointly sampled the bias parameters starting with an initial guess randomly sampled from the prior distribution (see Eq. 4).

- TEST1 focuses on a resolution of $10 h^{-1}$ Mpc, corresponding to a Nyquist frequency of $k \simeq 0.3 h \text{Mpc}^{-1}$. This resolution surpasses the typical scale used in most cosmological analyses, which are commonly limited to $k < 0.2 h \text{Mpc}^{-1}$ (see e.g. Ivanov et al. 2025). Also, the ideal resolution at which BAO reconstruction techniques are applied is slightly lower, i.e. $15 h^{-1}$ Mpc (see e.g. Paillas et al. 2025).
- TEST2 focuses on a resolution of $5 h^{-1}$ Mpc with a Nyquist frequency of $k \simeq 0.6 h \text{Mpc}^{-1}$. This resolution has been reported to be high enough to produce accurate galaxy catalogues combined with a sub-grid model based on ALPT (see Forero Sánchez et al. 2024).
- TEST3: This case study incorporates the full HICOBAN non-local bias model, which has been shown to significantly enhance accuracy in both power spectra and bispectra, achieving full compatibility with N -body-based catalogues across a wide range of scales. While Coloma-Nadal et al. (2024) demonstrated that this model remains valid at mesh resolutions below $4 h^{-1}$ Mpc, in this study we restricted our analysis to a coarser resolution of $8 h^{-1}$ Mpc. This choice is motivated by the fact that the corresponding Nyquist frequency, $k \simeq 0.4 h \text{Mpc}^{-1}$, marks the scale beyond which cosmological information is effectively saturated (Hahn & Villaescusa-Navarro 2021). For computational reasons, we kept the bias parameters fixed in this case, assuming they are known. This assumption is reasonable if the parameters can be extracted from high-fidelity reference catalogues, as was done for the PATCHY BOSS mocks, where the MULTIDARK reference catalogue was designed to reproduce the observed galaxy distribution with halo abundance matching (Rodríguez-Torres et al. 2016). Even more sophisticated reference catalogues are now being developed to reproduce the clustering statistics of the observed DESI catalogues (DESI Collaboration 2025), from which the parameters of the HICOBAN model can be accurately extracted (Favole et al. 2025).

We considered $\rho' = 0$ and $\epsilon' = 1$ for simplicity, interpreting this by continuity for $\delta > -1$, so the high-density damping factor, $\exp\left[-\left(\frac{1+\delta_i}{\rho'}\right)^{-\epsilon'}\right]$, was set to one in all three test cases.

3.2. Numerical results of TEST1

The first numerical test was performed on a grid with 128^3 voxels within a comoving volume of $1280 h^{-1}$ Mpc per side, corresponding to a spatial resolution of $10 h^{-1}$ Mpc. We used a Φ -web-dependent bias modelled with a power-law and negative-binomial likelihood for each region, for a total of eight

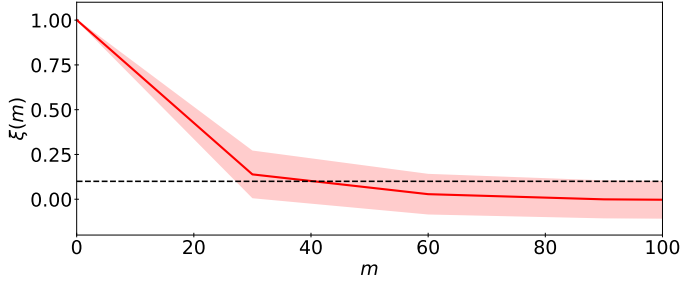


Fig. 3. TEST1: Mean auto-correlation $\xi(n)$ as a function of lag, n , for a chain of length of $M \approx 1000$, evaluated over 10^4 randomly selected voxels. The red curve shows the ensemble average, $\langle \xi(n) \rangle$, across parameters, with the shaded band indicating the $\pm 1\sigma$ dispersion. The horizontal dashed line marks the threshold, $\xi = 0.1$, used to assess mixing efficiency.

bias parameters. Using this model and a fixed random seed, we constructed a mock catalogue of object number counts, \mathbf{n}^* , in redshift space, which we define as the ground truth. The priors of the power-law (α) and negative-binomial (β) bias parameters were set to be log-normally distributed (see Eq 4) with fixed hyperparameters: $\mu_\alpha = 0.27$, $\sigma_\alpha = 0.30$, $\mu_\beta = 2.12$, $\sigma_\beta = 0.20$.

The HMC chain was initialised with a state $(\mathbf{v}^{(0)}, \mathbf{b}^{(0)})$, with $\mathbf{v}^{(0)}$ being the white-noise representation of the density contrast coming from Wiener-filtering, \mathbf{n}^* , and $\mathbf{b}^{(0)}$ being bias parameters randomly sampled from their priors. During burn-in, chain convergence was achieved after approximately 250 samples. The adapted step size was $\sim 10^{-3}$, and the NUTS trajectory length adaptation saturated at 1024 leapfrog steps (corresponding to a maximum tree depth of 10).

After convergence, we assessed the correlation length in the white-noise samples, based on their autocorrelation at lag m :

$$\xi_i(m) = \frac{1}{\sigma_i^2(M-m)} \sum_{j=0}^{M-m} (v_i^{(j)} - \bar{v}_i)(v_i^{(j+m)} - \bar{v}_i), \quad (20)$$

where $v_i^{(j)}$ is the i th voxel of the j th sample, M the total number of samples of a given chain run, \bar{v}_i the within-chain mean, and σ_i^2 the within-chain variance. For a chain of $M = 1000$ samples, we computed $\xi_i(m)$ for lags $m = 0$ to 100 across 10^4 randomly selected voxels. Figure 3 shows the mean auto-correlation over this subset.

We defined the correlation length as the smallest lag, n , such that $\langle \xi(n) \rangle \leq 0.1$, yielding an estimate of ~ 40 samples. Given this auto-correlation, we drew 15 000 post-burn-in states, providing about 500 effectively independent posterior samples for the subsequent statistical analysis.

We assessed the computing times of single-gradient calculations across different mesh sizes and bias configurations finding that they are below one second (see Fig. 4).

In addition, we performed the Gelman–Rubin convergence diagnostic, which confirms that the drawn samples have converged to the same target distribution (see Fig. 5).

As a first step, we conducted a visual assessment of the true and reconstructed maps of the initial and final density fields (see Fig. 6). The reconstructed samples exhibit a high degree of visual similarity to the ground truth for both the initial and final conditions. On a more quantitative level, we find that the standard deviations of the reconstructed fields display structured patterns at amplitudes approximately an order of magnitude lower than the true density field at the initial conditions. In contrast, the

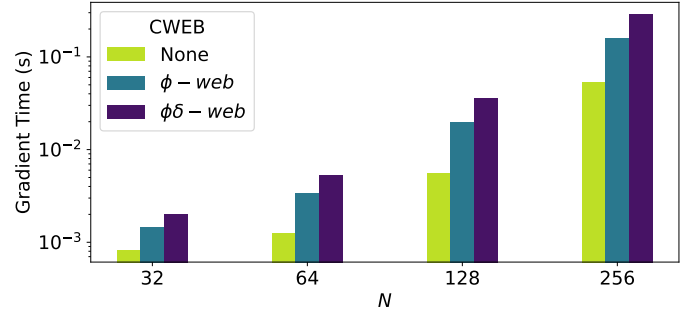


Fig. 4. Computing times for single-gradient evaluations of the forward model employing ALPT evolution and the HICOBIAN bias model for different mesh sizes, run on a single NVIDIA A100-SXM4 GPU equipped with 40 GB of on-board HBM2 memory. For reference, the full burn-in and sampling procedure for the 128^3 case in this study required approximately 100 GPU hours.

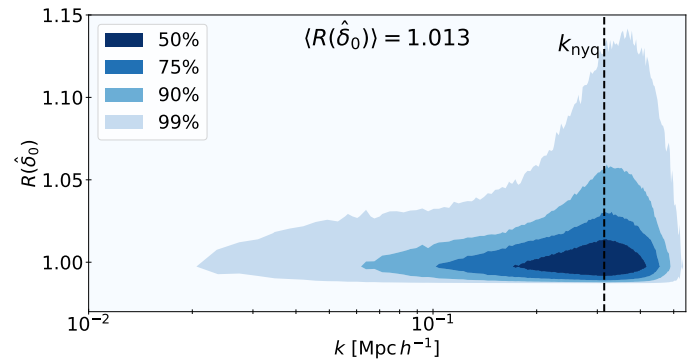


Fig. 5. Gelman–Rubin convergence diagnostic computed for three independently initialised chains of 250 samples after convergence. This statistic is calculated for all cells in Fourier space, and we show their distribution with scale, k . Values close to one indicate that the drawn samples have converged to the same distribution.

final conditions exhibit higher variance, reflecting the increased non-linearity and complexity of the evolved structures.

To quantitatively assess the quality of the reconstruction, we computed both two- and three-point statistics in Fourier space for the initial (Fig. 7) and final conditions (Fig. 8). Specifically, we evaluated the monopole and quadrupole moments of the power spectrum, along with the reduced bispectrum for a configuration particularly sensitive to non-linear effects ($k_1 = 0.1$ and $k_2 = 0.2 \text{ h Mpc}^{-1}$). In addition, we computed the propagator, defined as the cross-correlation between the mock tracer density contrast and the initial density reference. The motivation for investigating the quadrupole and bispectrum of a the reconstructed Gaussian field was to assess how well the samples recover deviations introduced by cosmic variance in these statistics.

In Figs. 7 and 8, it is remarkable to observe that the monopoles, quadrupoles, and bispectra are accurately recovered not only in the final conditions, but also in the initial conditions, even reproducing the specific features inherent to the particular cosmic variance realisation. Although local-bias models fail to reproduce three-point statistics without short- and long-range non-local terms (Coloma-Nadal et al. 2024), we show that only neglecting long-range non-local bias also leads to significant differences in the bispectrum (see the third panel in Fig. 8).

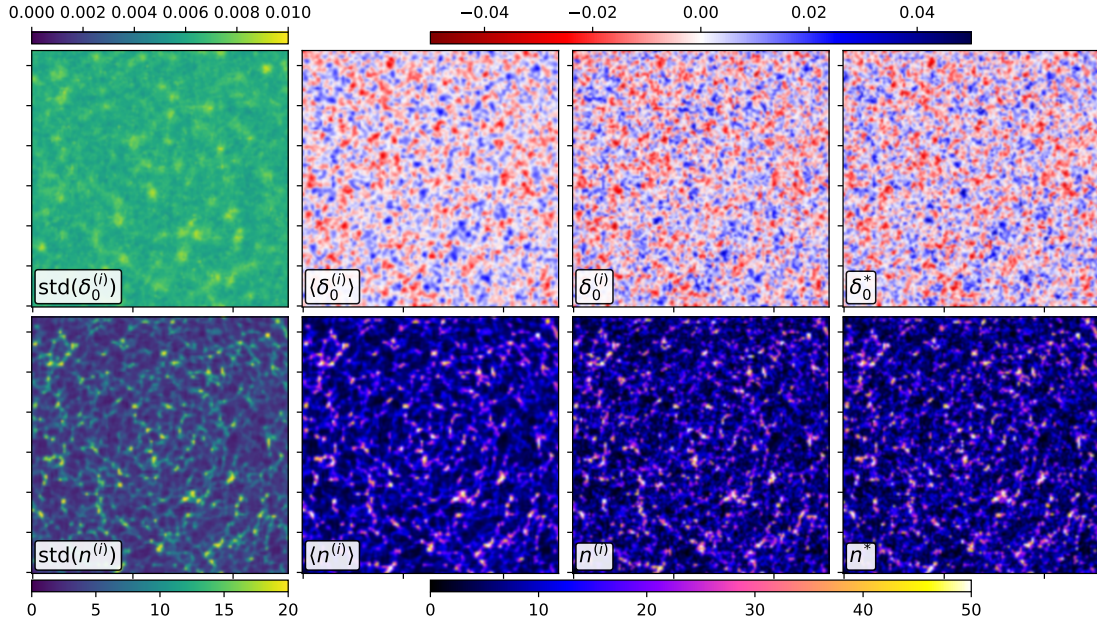


Fig. 6. TEST1: Spatial summary of 500 independent reconstructions ($N = 128$). Slices of $\Delta L = 10 h^{-1}$ Mpc (one voxel width). Top row: Reconstructed initial density contrast. Bottom row: Reconstructed tracer field. Columns from left to right: (1) Pixel-wise standard deviation across the 500 independent reconstructions; (2) mean reconstructed field; (3) one representative reconstruction sample; and (4) true reference field.

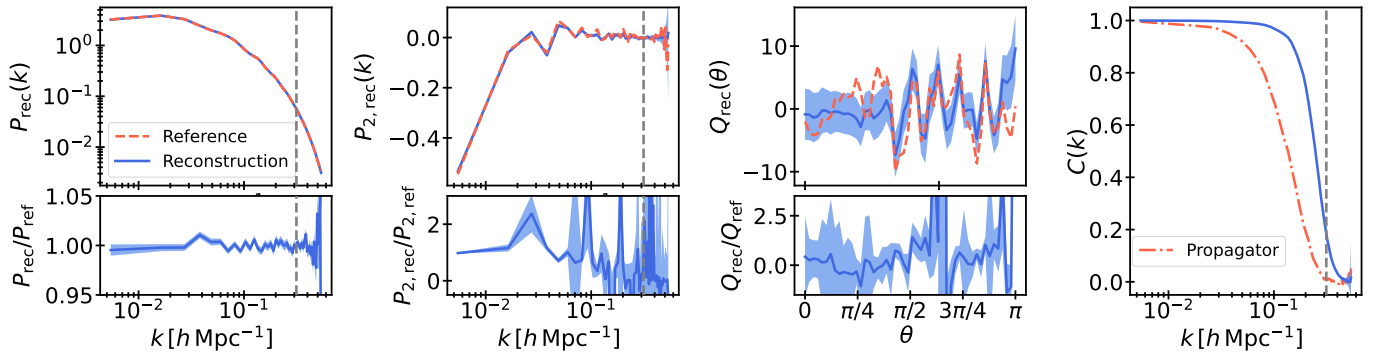


Fig. 7. TEST1: Comparison of summary statistics of the initial density contrast field for the mock reference (dashed red line) and 500 independent reconstructions (solid blue line). Columns from left to right: (1) Monopole power spectrum; (2) quadrupole; (3) reduced bispectrum as a function of triangle opening angle, θ ; and (4) cross-correlation coefficient $C(k)$ between the reference and reconstruction (solid blue line) and the propagator (dashed red line), defined as the cross-correlation between the mock tracer density contrast and the initial density reference. In columns (1)–(3), the lower sub-panels display the ratios with respect to the reference. Shaded bands denote the 1σ intervals from the posterior samples. Vertical dashed lines indicate the isotropic Nyquist frequency, $k_{\text{nyq}}/\sqrt{3}$.

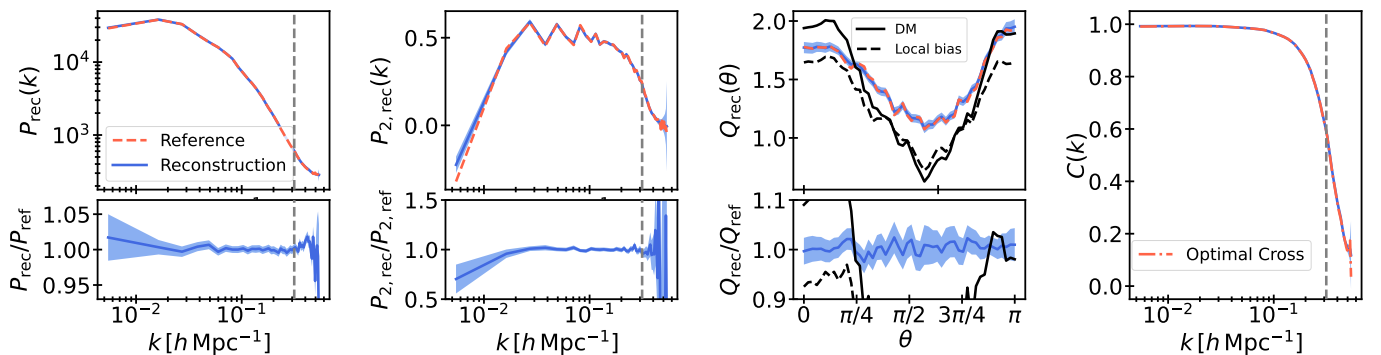


Fig. 8. TEST1: Comparison of summary statistics of the mock reference and 500 independent reconstructions evolved through the forward model. Panel definitions are identical to those in Fig. 7, except that in the rightmost panel the propagator is replaced by the maximal tracer field cross-correlation and that we show the bispectrum of the dark-matter field and a tracer field using a local bias (without cosmic web) with bias parameters set to those of knots.

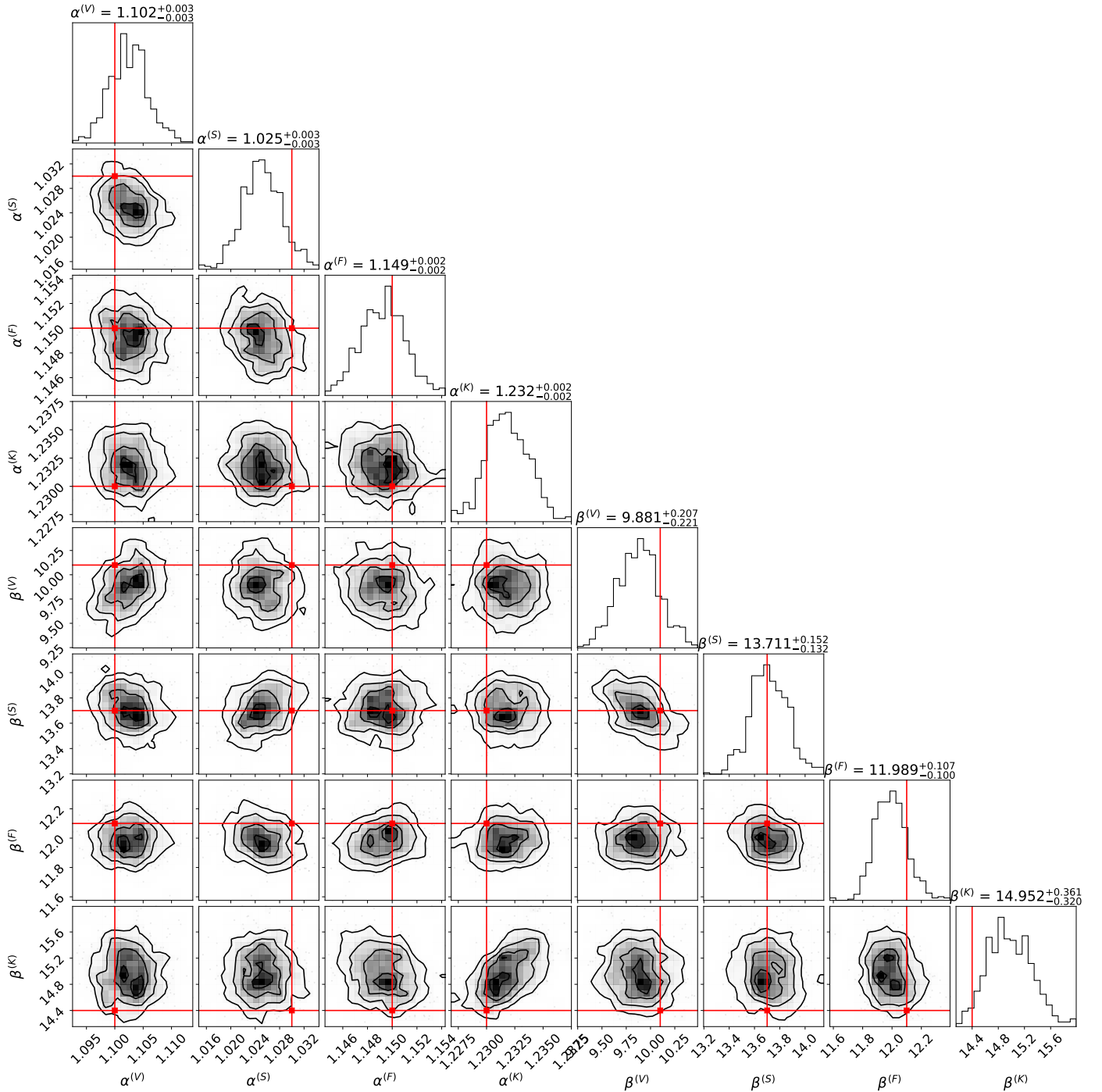


Fig. 9. TEST1: Posterior distributions for the 500 independent samples of the α and β bias parameters in each cosmic-web region of the Φ web. Red dots and lines show the values of the parameters for the ground-truth catalogue. The posteriors are smoothed with a 2D Gaussian filter with standard deviation equal to 0.75 bin widths. TEST2 yields equivalent results.

The propagator ensures a significant gain in information, demonstrating the effectiveness of the reconstruction. To evaluate how close the result is to an ideal scenario, we computed the optimal cross at final conditions by generating a synthetic-tracer number-count field from the evolved ground-truth dark-matter field using the exact same bias parameters and stochastic prescription as in the model, but a different seed. The cross-correlation between this ideal tracer field and the ground truth serves as a benchmark, represented by the dashed-dotted red line in the figure. This allowed us to assess whether the recon-

struction approaches the theoretical limit given the assumed bias model. As we see in Fig. 8, the red and blue lines of the right panel are on top of each other, indicating that the optimal cross-correlation limited by shot noise has been reached.

In Fig. 9, we show a corner plot of the posterior distributions of the eight parameters of the bias model compared to their values in the reference catalogue. The power-law parameters, α , were recovered with sub-percent accuracy, while the negative binomial β -parameter was recovered with precision of a few percent. The absence of significant correlations between α bias

parameters in distinct cosmic-web regions supports the validity and stability of our inference framework.

Overall, the consistency of all diagnostics – maps, summary statistics, and bias-parameter posteriors – demonstrates that the complete field-level inference pipeline reliably recovers both the underlying density field and the cosmic-web-dependent bias parameters at the targeted resolution. The numerical studies TEST2 and TEST3 further confirm the validation of the BRIDGE code (details can be found in Appendix C). However, it is remarkable that we did not recover the original bias parameters when we allowed them to vary freely in case TEST3.

4. Conclusions

In this work we developed and validated a novel Bayesian field-level inference framework that incorporates a physically motivated, differentiable, and non-local bias model – HICOBIAN – based on the hierarchical cosmic-web. This model allows for smooth transitions between different cosmic environments via a fuzzy classification approach, making it well suited for machine-learning integration and GPU acceleration.

We demonstrated that, within a self-specified setting where inference is performed on data generated by the exact same forward model, our approach accurately reconstructs the initial density field, achieving high fidelity even in the presence of redshift-space distortions and complex, non-linear, non-local, and stochastic bias. The reconstruction quality was evaluated using two- and three-point statistics, including the power-spectrum monopole and quadrupole, the reduced bispectrum, and the propagator, showing excellent agreement with the ground truth and reaching the optimal limit set by shot noise.

We further tested the robustness of the model under various conditions, including varying spatial resolutions and complex bias scenarios involving four to 16 cosmic-web regions, each characterised by distinct bias parameters. These included power-law bias, exponential threshold bias, and deviations from Poisson statistics, totalling eight to 64 bias parameters. Across all settings, the method maintained its high accuracy, confirming the generality and adaptability of our approach. However, for the complex bias models, the original bias parameters were not recovered when left to vary freely. This hints at the complexity of bias modelling and the potential limitations of field-level inference approaches that ignore intra-halo dynamics. As demonstrated in recent work (e.g. Favole et al. 2025), a realistic halo-to-galaxy connection based on high-resolution N -body simulations requires a large number of parameters to capture processes such as halo exclusion, satellite dynamics, quenching, and intra-halo flows.

In future work, we plan to investigate the performance of the BRIDGE code with high-fidelity synthetic reference catalogues and directly apply it to redshift survey data across various tracers, including DESI-tracers, such as bright galaxies, luminous red galaxies, emission-line galaxies, quasars, and Lyman- α forests. This will require the incorporation of light-cone evolution and selection effects and a more detailed analysis of redshift-space distortions in the highly non-linear regime. Moreover, our results suggest that achieving full consistency may require a joint Bayesian inference framework that incorporates intra-halo dynamics informed by detailed N -body simulations to break the degeneracy in the bias parameters.

By implementing this framework in the GPU-accelerated BRIDGE code using JAX, we enable efficient and scalable inference on large datasets. This work thus marks a significant step towards a unified, interpretable, and fully differentiable frame-

work for cosmological data analysis, with direct applicability to tracers of the large-scale structure.

Acknowledgements. The authors thank the Spanish Ministry of Science and Innovation for funding the Big Data of the Cosmic-Web project (PID2020-120612GB-I00/AEI/10.13039/501100011033, PI:FSK), under which this work was conceived and carried out. We are also grateful to the Instituto de Astrofísica de Canarias (IAC) for its continued support through the Cosmology with LSS Probes project (PI:FSK). The authors thank Jesús Sanz Serna and José María Coloma-Nadal for discussions. PR acknowledges support from the IAC Research Summer Grant 2024 for the project “Bayesian Inference of the LSS with ALPT and HICOBIAN bias using JAX”. This grant enabled his contribution to this work prior to the start of his master’s thesis on the same topic, as reflected in this manuscript. He also thanks Gustavo Yepes, the Universidad Autónoma de Madrid (UAM), and the Department of Theoretical Physics for their hospitality and for providing access to computing facilities during his visit – resources that were instrumental to the development of this work.

References

- Amendola, L., Appleby, S., Avgoustidis, A., et al. 2016, arXiv e-prints [arXiv:1606.00180]
- Andrews, A., Jasche, J., Lavaux, G., & Schmidt, F. 2023, *MNRAS*, 520, 5746
- Angulo, R. E., & Hahn, O. 2022, *LRCAs*, 8, 1
- Ata, M., Kitaura, F.-S., & Müller, V. 2015, *MNRAS*, 446, 4250
- Ata, M., Kitaura, F.-S., Lee, K.-G., et al. 2021, *MNRAS*, 500, 3194
- Balaguera-Antolínez, A., Kitaura, F.-S., Pellejero-Ibáñez, M., Zhao, C., & Abel, T. 2018, *MNRAS*, 483, L58
- Balaguera-Antolínez, A., Kitaura, F.-S., Pellejero-Ibáñez, M., et al. 2020, *MNRAS*, 491, 2565
- Balaguera-Antolínez, A., Kitaura, F.-S., Alam, S., et al. 2023, *A&A*, A130
- Baldauf, T., Seljak, U., Smith, R. E., Hamaus, N., & Desjacques, V. 2013, *Phys. Rev. D*, 88, 083507
- Bayer, A. E., Seljak, U., & Modi, C. 2023, arXiv e-prints [arXiv:2307.09504]
- Bernardeau, F. 1994, *ApJ*, 427, 51
- Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, *Phys. Rep.*, 367, 1
- Bertschinger, E., & Dekel, A. 1989, *ApJ*, 336, L5
- Bi, H., & Davidsen, A. F. 1997, *ApJ*, 479, 523
- Blanes, S., Casas, F., & Sanz-Serna, J. M. 2014, *SIAM J. Sci. Comput.*, 36, A1556
- Blot, L., Crocce, M., Sefusatti, E., et al. 2019, *MNRAS*, 485, 2806
- Bond, J. R., Kofman, L., & Pogosyan, D. 1996, *Nature*, 380, 603
- Bos, E. G. P., Kitaura, F.-S., & van de Weygaert, R. 2019, *MNRAS*, 488, 2573
- Bouchet, F. R., Colombi, S., Hivon, E., & Juszkiewicz, R. 1995, *A&A*, 296, 575
- Bradbury, J., Frostig, R., Hawkins, P., et al. 2018, *JAX: composable transformations of Python+NumPy programs*, <http://github.com/google/jax>
- Buchert, T. 1994, *MNRAS*, 267, 811
- Casas-Miranda, R., Mo, H. J., Sheth, R. K., & Boerner, G. 2002, *MNRAS*, 333, 730
- Catelan, P. 1995, *MNRAS*, 276, 115
- Cen, R., & Ostriker, J. P. 1993, *ApJ*, 417, 415
- Chan, K. C., Scoccimarro, R., & Sheth, R. K. 2012, *Phys. Rev. D*, 85, 083509
- Chuang, C.-H., Kitaura, F.-S., Liang, Y., et al. 2017, *Phys. Rev. D*, 95, 063528
- Chuang, C.-H., Yepes, G., Kitaura, F.-S., et al. 2019, *MNRAS*, 487, 48
- Coles, P., & Jones, B. 1991, *MNRAS*, 248, 1
- Coloma-Nadal, J. M., Kitaura, F.-S., García-Farieta, J. E., et al. 2024, *JCAP*, 2024, 083
- Conceição, M., Krone-Martins, A., da Silva, A., & Moliné, Á. 2024, *A&A*, 681, A123
- Daly, F., & Gaunt, R. E. 2016, *ALEA: Latin Am. J. Probability Math. Stat.*, 13, 635
- Dekel, A., & Lahav, O. 1999, *ApJ*, 520, 24
- DESI Collaboration (Abdul-Karim, M., et al.) 2025, arXiv e-prints [arXiv:2503.14745]
- Desjacques, V., Jeong, D., & Schmidt, F. 2018, *Phys. Rep.*, 733, 1
- Doerer, L., Jamieson, D., Stopyra, S., et al. 2024, *MNRAS*, 535, 1258
- Duane, S., Kennedy, A., Pendleton, B. J., & Roweth, D. 1987, *Phys. Lett. B*, 195, 216
- Eisenstein, D. J., Seo, H.-J., Sirko, E., & Spergel, D. N. 2007, *ApJ*, 664, 675
- Euclid Collaboration (Aussel, H., et al.) 2026, *A&A*, in press, <https://doi.org/10.1051/0004-6361/202554610>
- Favole, G., Kitaura, F. S., Hadzhiyska, B., et al. 2025, *ApJ*, submitted [arXiv:2512.04362]
- Feng, Y., Chu, M.-Y., Seljak, U., & McDonald, P. 2016, *MNRAS*, 463, 2273
- Filho, M. E., Sánchez Almeida, J., Muñoz-Tuñón, C., et al. 2015, *ApJ*, 802, 82

- Forero Sánchez, D., Kitaura, F. S., Sinigaglia, F., Coloma-Nadal, J. M., & Kneib, J. P. 2024, *JCAP*, 2024, 001
- Forero-Romero, J. E., Hoffman, Y., Gottlöber, S., Klypin, A., & Yepes, G. 2009, *MNRAS*, 396, 1815
- Fry, J. N., & Gaztanaga, E. 1993, *ApJ*, 413, 447
- García-Farieta, J. E., Balaguera-Antolínez, A., & Kitaura, F.-S. 2024, *A&A*, 690, A27
- Garrison, L. H., Eisenstein, D. J., Ferrer, D., et al. 2018, *ApJS*, 236, 43
- Hahn, C., & Villaescusa-Navarro, F. 2021, *JCAP*, 2021, 029
- Hahn, O., Porciani, C., Carollo, C. M., & Dekel, A. 2007, *MNRAS*, 375, 489
- Hamilton, A. J. S. 1998, *Astrophys. Space Sci. Libr.*, 231, 185
- Heavens, A., & Peacock, J. 1988, *MNRAS*, 232, 339
- Heß, S., Kitaura, F.-S., & Gottlöber, S. 2013, *MNRAS*, 435, 2065
- Hoffman, M. D., & Gelman, A. 2011, arXiv e-prints [arXiv:1111.4246]
- Horowitz, B., & Lukic, Z. 2025, arXiv e-prints [arXiv:2502.02294]
- Horowitz, B., Lee, K.-G., White, M., Krolewski, A., & Ata, M. 2019a, *ApJ*, 887, 61
- Horowitz, B., Seljak, U., & Aslanyan, G. 2019b, *JCAP*, 2019, 035
- Horowitz, B., Zhang, B., Lee, K.-G., & Kooistra, R. 2021, *ApJ*, 906, 110
- Horowitz, B., Dornfest, M., Lukic, Z., & Harrington, P. 2022, *ApJ*, 941, 42
- Hui, L., & Gnedin, N. Y. 1997, *MNRAS*, 292, 27
- Ishiyama, T., Prada, F., Klypin, A. A., et al. 2021, *MNRAS*, 506, 4210
- Ivanov, M. M., Obuljen, A., Cuesta-Lazaro, C., & Toomey, M. W. 2025, *Phys. Rev. D*, 111, 063548
- Jasche, J., & Kitaura, F.-S. 2010, *MNRAS*, 407, 29
- Jasche, J., & Lavaux, G. 2019, *A&A*, 625, A64
- Jasche, J., & Wandelt, B. D. 2013, *MNRAS*, 432, 894
- Jasche, J., Kitaura, F. S., Li, C., & Enßlin, T. A. 2010, *MNRAS*, 409, 355
- Jasche, J., Leclercq, F., & Wandelt, B. 2015, *JCAP*, 2015, 036
- Kaiser, N. 1984, *ApJ*, 284, L9
- Kaiser, N. 1987, *MNRAS*, 227, 1
- Kitaura, F.-S. 2013, *MNRAS*, 429, L84
- Kitaura, F.-S., & Enßlin, T. A. 2008, *MNRAS*, 389, 497
- Kitaura, F.-S., & Heß, S. 2013, *MNRAS*, 435, L78
- Kitaura, F.-S., Jasche, J., & Metcalf, R. B. 2010, *MNRAS*, 403, 589
- Kitaura, F.-S., Angulo, R. E., Hoffman, Y., & Gottlöber, S. 2012a, *MNRAS*, 425, 2422
- Kitaura, F.-S., Erdoğdu, P., Nuza, S. E., et al. 2012b, *MNRAS*, 427, L35
- Kitaura, F.-S., Gallerani, S., & Ferrara, A. 2012c, *MNRAS*, 420, 61
- Kitaura, F.-S., Yepes, G., & Prada, F. 2014, *MNRAS*, 439, L21
- Kitaura, F.-S., Gil-Marín, H., Scóccola, C. G., et al. 2015, *MNRAS*, 450, 1836
- Kitaura, F.-S., Ata, M., Angulo, R. E., et al. 2016a, *MNRAS*, 457, L113
- Kitaura, F.-S., Rodríguez-Torres, S., Chuang, C.-H., et al. 2016b, *MNRAS*, 456, 4156
- Kitaura, F.-S., Ata, M., Rodríguez-Torres, S. A., et al. 2021, *MNRAS*, 502, 3456
- Kitaura, F.-S., Balaguera-Antolínez, A., Sinigaglia, F., & Pellejero-Ibañez, M. 2022, *MNRAS*, 512, 2245
- Kitaura, F.-S., Sinigaglia, F., Balaguera-Antolínez, A., & Favole, G. 2024, *A&A*, 683, A215
- Klypin, A., & Prada, F. 2018, *MNRAS*, 478, 4602
- Kodi Ramanah, D., Charnock, T., Villaescusa-Navarro, F., & Wandelt, B. D. 2020, *MNRAS*, 495, 4227
- Lavaux, G., & Jasche, J. 2016, *MNRAS*, 455, 3169
- Lavaux, G., Jasche, J., & Leclercq, F. 2019, *Systematic-Free Inference of the Cosmic Matter Density Field from SDSS3-BOSS Data*
- Levi, M., Bebek, C., Beers, T., et al. 2013, arXiv e-prints [arXiv:1308.0847]
- Li, Y., Modi, C., Jamieson, D., et al. 2024, *ApJS*, 270, 36
- McDonald, P., & Roy, A. 2009, *JCAP*, 2009, 020
- McDonald, P., Miralda-Escudé, J., Rauch, M., et al. 2000, *ApJ*, 543, 1
- Modi, C., Lanusse, F., & Seljak, U. 2021, *Astron. Comput.*, 37, 100505
- Monaco, P., & Efstathiou, G. 1999, *MNRAS*, 308, 763
- Neyrinck, M. C. 2016, *MNRAS*, 455, L11
- Neyrinck, M. C., Aragón-Calvo, M. A., Jeong, D., & Wang, X. 2014, *MNRAS*, 441, 646
- Nusser, A., & Dekel, A. 1992, *ApJ*, 391, 443
- Nuza, S. E., Kitaura, F.-S., Heß, S., Libeskind, N. I., & Müller, V. 2014, *MNRAS*, 445, 988
- Olex, E., Hellwing, W. A., & Knebe, A. 2025, *A&A*, 696, A142
- Paillass, E., Ding, Z., Chen, X., et al. 2025, *JCAP*, 2025, 142
- Peebles, P. J. E. 1980, *The large-scale structure of the universe* (Princeton University Press)
- Pellejero-Ibañez, M., Balaguera-Antolínez, A., Kitaura, F.-S., et al. 2020, *MNRAS*, 493, 586
- Phan, D., Pradhan, N., & Jankowiak, M. 2019, arXiv e-prints [arXiv:1912.11554]
- Porqueres, N., Jasche, J., Lavaux, G., & Enßlin, T. 2019, *A&A*, 630, A151
- Rodríguez-Torres, S. A., Chuang, C.-H., Prada, F., et al. 2016, *MNRAS*, 460, 1173
- Saslaw, W. C., & Hamilton, A. J. S. 1984, *ApJ*, 276, 13
- Schmittfull, M., Simonović, M., Assassi, V., & Zaldarriaga, M. 2019, *Phys. Rev. D*, 100, 043514
- Seljak, U. 2012, *JCAP*, 2012, 004
- Sheth, R. K. 1995, *MNRAS*, 274, 213
- Sheth, R. K., & Lemson, G. 1999, *MNRAS*, 304, 767
- Shirasaki, M., Sugiyama, N. S., Takahashi, R., & Kitaura, F.-S. 2021, *Phys. Rev. D*, 103, 023506
- Sinigaglia, F., Kitaura, F.-S., Balaguera-Antolínez, A., et al. 2021, *ApJ*, 921, 66
- Sinigaglia, F., Kitaura, F.-S., Balaguera-Antolínez, A., et al. 2022, *ApJ*, 927, 230
- Sinigaglia, F., Kitaura, F.-S., Nagamine, K., & Oku, Y. 2024a, *ApJ*, 971, L22
- Sinigaglia, F., Kitaura, F.-S., Nagamine, K., Oku, Y., & Balaguera-Antolínez, A. 2024b, *A&A*, 682, A21
- Somerville, R. S., Lemson, G., Sigad, Y., et al. 2001, *MNRAS*, 320, 289
- Stadler, J., Schmidt, F., & Reinecke, M. 2023, *JCAP*, 2023, 069
- Takada, M., Ellis, R. S., Chiba, M., et al. 2014, *PASJ*, 66, R1
- Tassev, S., Zaldarriaga, M., & Eisenstein, D. J. 2013, *JCAP*, 6, 036
- Tegmark, M., Blanton, M. R., Strauss, M. A., et al. 2004, *ApJ*, 606, 702
- Vos-Ginés, B., Avila, S., Gonzalez-Perez, V., & Yepes, G. 2024, *MNRAS*, 530, 3458
- Wang, H., Mo, H. J., Yang, X., & van den Bosch, F. C. 2012, *MNRAS*, 420, 1809
- Wang, H., Mo, H. J., Yang, X., & Van den Bosch, F. C. 2013, *ApJ*, 772, 63
- Wang, H., Mo, H. J., Yang, X., Jing, Y. P., & Lin, W. P. 2014, *ApJ*, 794, 94
- Wang, Y., Zhai, Z., Alavi, A., et al. 2022, *ApJ*, 928, 1
- Werner, K. F., & Porciani, C. 2020, *MNRAS*, 492, 1614
- Yahil, A., Strauss, M. A., Davis, M., & Huchra, J. P. 1991, *ApJ*, 372, 380
- Zaroubi, S., Hoffman, Y., Fisher, K. B., & Lahav, O. 1995, *ApJ*, 449, 446
- Zel'dovich, Y. B. 1970, *A&A*, 5, 84
- Zhao, C., Huang, S., He, M., et al. 2024, arXiv e-prints [arXiv:2411.07970]

Appendix A: White-noise parametrization

In the forward model we have adopted a white-noise parametrization, \mathbf{v} , of i.i.d. standard normal variates, which we transform into the linear density field through the linear operator

$$\mathcal{M}_\delta(\mathbf{v}) \equiv \delta_0 = (\Delta x)^{-3} \mathbf{F}^{-1} \sqrt{\mathbf{P}} \mathbf{F} \mathbf{v}, \quad (\text{A.1})$$

where \mathbf{F} is the discrete Fourier-transform (DFT) linear operator on an N^3 mesh of side length L , given by

$$\mathbf{F}_{nm} = (\Delta x)^3 \exp(-i\mathbf{k}_n \cdot \mathbf{x}_m), \quad (\text{A.2})$$

$$\mathbf{F}_{nm}^{-1} = L^{-3} \exp(i\mathbf{k}_m \cdot \mathbf{x}_n). \quad (\text{A.3})$$

Here \mathbf{x}_n represents the Cartesian direction vector in comoving space at a voxel indexed by n . The same holds for the wave-vector \mathbf{k}_n in the Fourier-transformed space. We have defined the diagonal linear matter power spectrum matrix $\mathbf{P}_{nm} \equiv \delta_{nm}^K P_{\text{lin}}(|\mathbf{k}_n|)$. The density contrast, δ_0 , is then Gaussian-distributed with zero mean and covariance:

$$\mathbf{C} = (\Delta x)^{-3} \mathbf{F}^{-1} \mathbf{P} \mathbf{F}. \quad (\text{A.4})$$

In principle, one can choose to parametrize the field either in terms of δ_0 directly or through the underlying white noise \mathbf{v} . Sampling δ_0 directly removes the linear transformation from the forward model but the fast Fourier transforms must still be performed at every Markov chain Monte Carlo step to evaluate the negative log prior and to draw new field realisations. We have tested both parametrizations and did not observe significant differences in computational cost or sampling behaviour. Nonetheless, we favour the white-noise parametrization \mathbf{v} for its simpler implementation.

Another option is to parametrize the field directly in Fourier space. Defining

$$\hat{\mathbf{u}} = \sqrt{\frac{2N^3}{L^6}} \mathbf{F} \mathbf{v}, \quad (\text{A.5})$$

the components of $\hat{\mathbf{u}}$ are i.i.d standard normal variates for both real and imaginary parts. The density contrast is then given by

$$\delta_0 = \sqrt{\frac{L^3}{2}} \mathbf{F}^{-1} \sqrt{\mathbf{P}} \hat{\mathbf{u}}. \quad (\text{A.6})$$

With this parametrization every parameter corresponds to a definite scale/wavenumber. We expected a different chain sampling behaviour, especially when using diagonal mass matrix adaptation during burn-in, but we saw no significant differences with respect to the white noise parametrization. We note that this parametrization requires the careful construction of a Hermitian-packed Fourier ($N, N, N/2 + 1$)-shaped array produced by a 3D real-to-complex fast Fourier transform out of N^3 random real numbers, with explicit handling of the array structure and Nyquist planes.

A third option we considered is to separate each Fourier mode into amplitude and phase, $\hat{\delta}_0 = |\delta_0| \exp(i\varphi)$. The parameters to be sampled are then split into:

- $N^3/2 - 4$ phases uniformly distributed in $[0, 2\pi)$
- $N^3/2 - 4$ amplitudes following a Rayleigh distribution.
- 8 Nyquist modes following a Gaussian distribution.

This parametrization may be advantageous if we want to fix the amplitudes to the theoretical linear matter power spectrum, effectively reducing the parameter space in half. But again, for our work, we did not come across any obvious advantages of using this approach over the more straightforward white-noise parametrization.

Appendix B: Nonlocal bias and cosmic-web relation

Let us consider a scalar field η , which could be, for example, the matter overdensity field $\eta = \delta$ or the gravitational potential $\eta = \Phi$. We begin by constructing the Hessian matrix of this field, defined as

$$\mathcal{H}_{jk} = \frac{\partial^2 \eta}{\partial x_j \partial x_k} = \eta_{,jk}, \quad (\text{B.1})$$

which encodes the second derivatives of η with respect to spatial coordinates.

We then compute the eigenvalues of the Hessian, denoted by $\lambda^{(1)} \geq \lambda^{(2)} \geq \lambda^{(3)}$. These eigenvalues characterize the local curvature of the field in each principal direction. From them, we define the three rotationally invariant scalar quantities, or invariants, of the Hessian:

- $I_1 = \lambda^{(1)} + \lambda^{(2)} + \lambda^{(3)}$, the trace of the Hessian,
- $I_2 = \lambda^{(1)}\lambda^{(2)} + \lambda^{(1)}\lambda^{(3)} + \lambda^{(2)}\lambda^{(3)}$, the sum of principal minors,
- $I_3 = \lambda^{(1)}\lambda^{(2)}\lambda^{(3)}$, the determinant of the Hessian.

According to [Kitaura et al. \(2022\)](#), the invariants of the Hessian matrix can be directly linked to different regions of the cosmic-web. For simplicity, and without loss of generality, this classification can be expressed for a threshold $\lambda_{\text{th}} = 0$ as follows:

- Knots: $I_3 > 0, I_2 > 0, I_1 > \lambda^{(1)}$
- Filaments: $I_3 < 0, I_2 < 0,$ or $I_3 < 0, I_2 > 0, \lambda^{(3)} < I_1 < \lambda^{(3)} - \frac{\lambda^{(2)}\lambda^{(3)}}{\lambda^{(1)}}$
- Sheets: $I_3 > 0, I_2 < 0,$ or $I_3 < 0, I_2 > 0, \lambda^{(1)} - \frac{\lambda^{(2)}\lambda^{(3)}}{\lambda^{(1)}} < I_1 < \lambda^{(1)}$
- Voids: $I_3 < 0, I_2 > 0, I_1 < \lambda^{(1)}$.

Moreover, these invariants can be directly related to the nonlocal bias operators commonly used in perturbation theory. Specifically for $\eta = \Phi$, i.e., long-range nonlocal bias:

- $\delta = I_1$, the local density contrast,
- $s^2 = \frac{2}{3}I_1^2 - 2I_2$, the tidal shear squared,
- $s^3 = -I_1I_2 + 3I_3 + \frac{2}{9}I_1^3$, the cubic tidal bias term.

This formulation thus bridges cosmic-web morphology with perturbative bias theory, showing that both short-range ($\eta = \delta$) and long-range ($\eta = \Phi$) nonlocal bias terms can be systematically described in terms of the same invariant-based framework. We refer to the cosmic-web classification based on $\eta = \delta$ as the δ -web, and when based on $\eta = \Phi$, as the Φ -web.

Appendix C: Numerical results of TEST2 and TEST3

Here we present the summary statistics resulting from TEST2 and TEST3. The case of TEST2 – when the resolution is doubled to $5 h^{-1}$ Mpc with respect to TEST1 – yields similarly accurate results across all summary statistics (see Figs. [C.1](#), [C.2](#), and [C.3](#)).

The additional numerical case TEST3 in which the bias parameters are assumed to be known, allows us to implement the most complex bias model as in [Coloma-Nadal et al. \(2024\)](#). Specifically, we define distinct bias prescriptions across the hierarchical cosmic-web, using a classification based on both the tidal field tensor and the Hessian of the density field. This results in 16 regions, each characterised by its own power-law, exponential threshold, and stochastic bias parameters. As shown in Figs. [C.4](#) [C.5](#), and [C.6](#), the reconstruction maintains the same level of precision under this more sophisticated setting, as in TEST1 and TEST2, assuming, however, the bias parameters to be known.

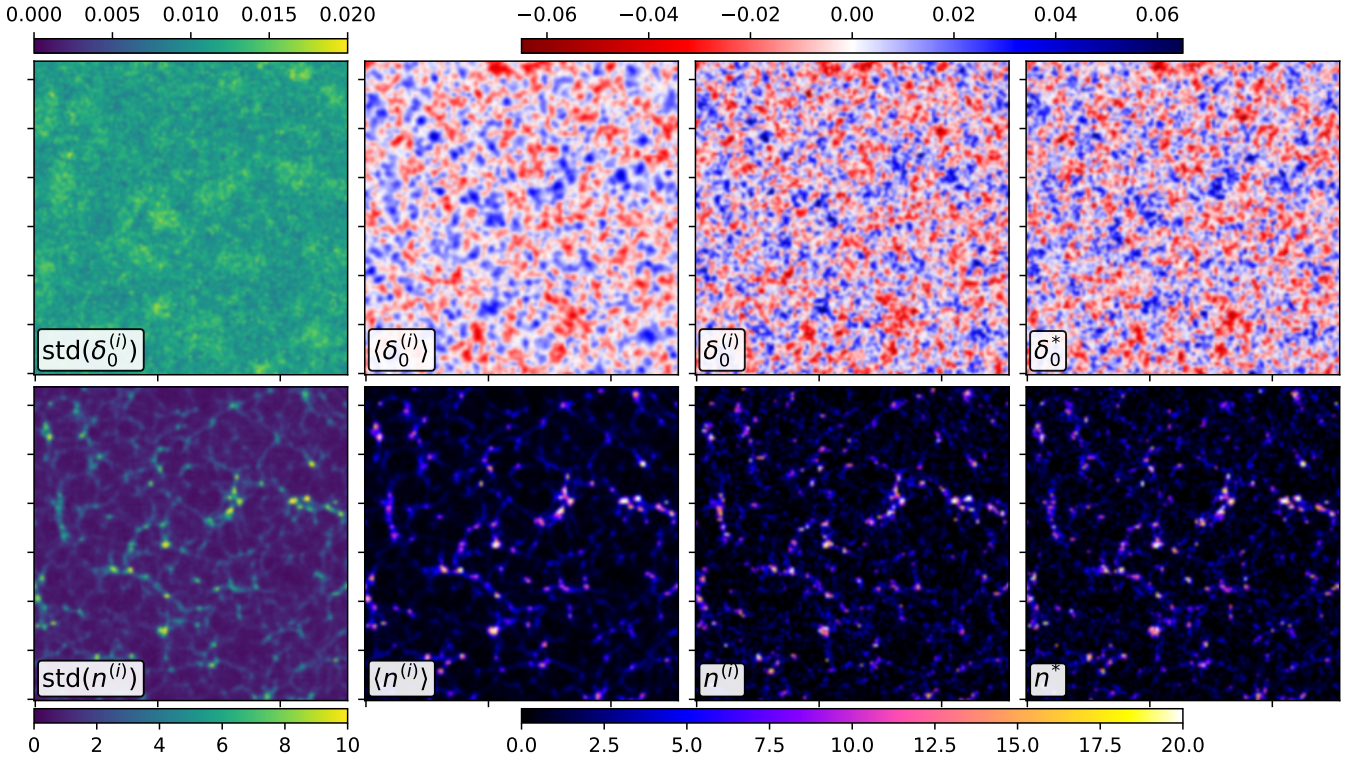


Fig. C.1. TEST2: Analogous to Fig. 6.

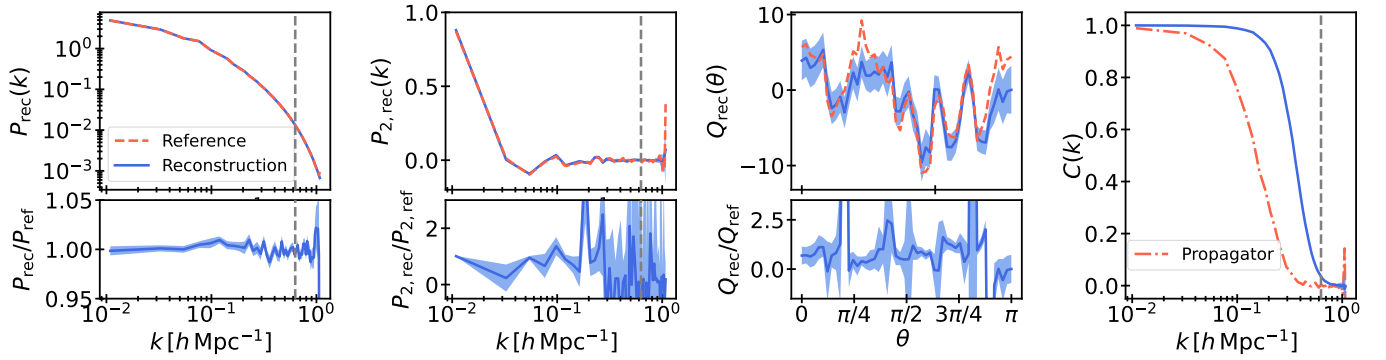


Fig. C.2. TEST2: Initial conditions. Analogous to Fig. 7.

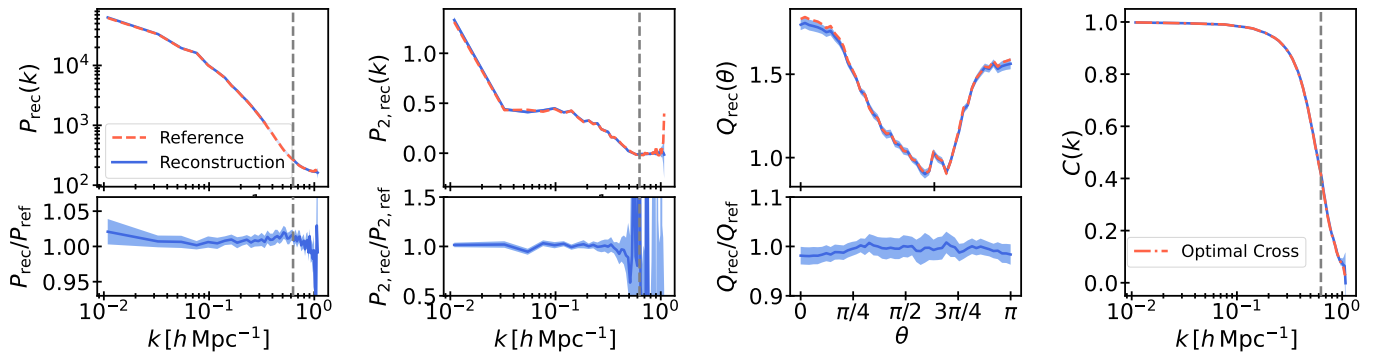


Fig. C.3. TEST2: Final conditions. Analogous to Fig. 8.

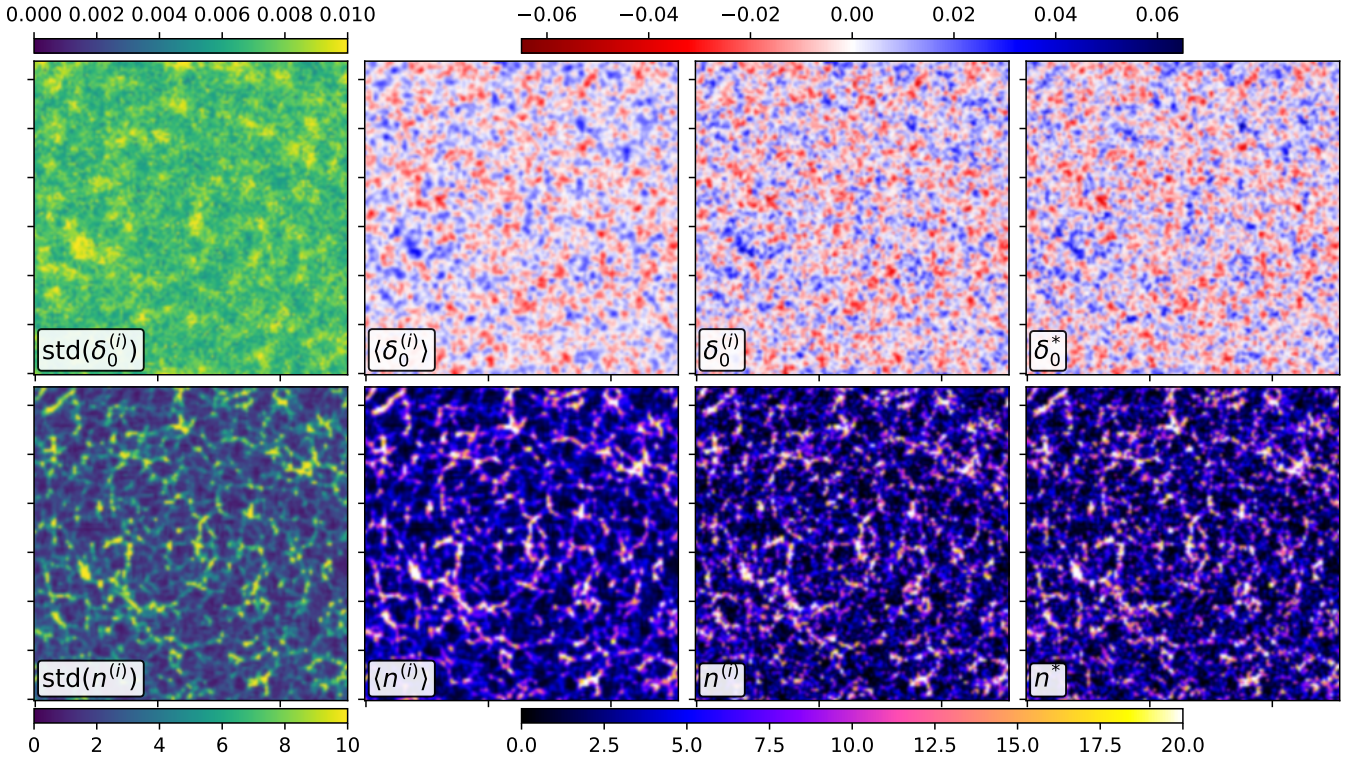


Fig. C.4. TEST3: Analogous to Fig. 6.

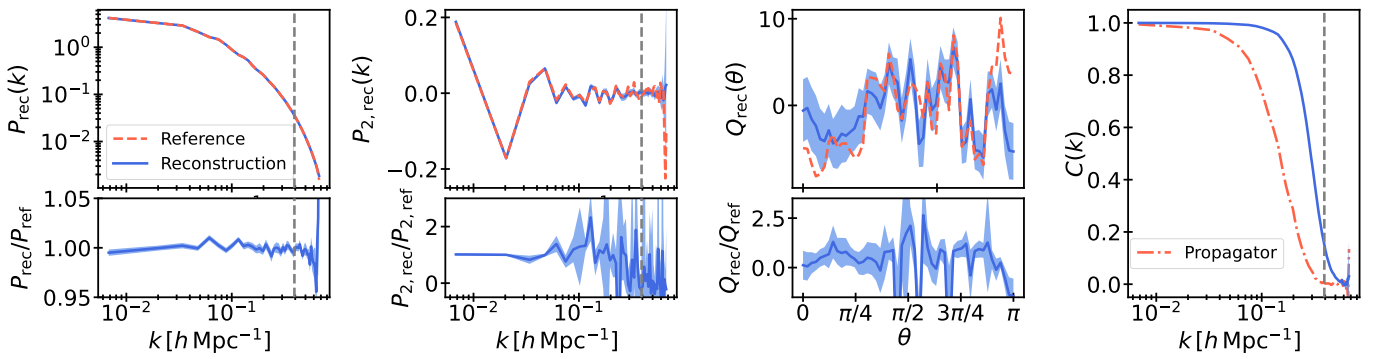


Fig. C.5. TEST3: Initial conditions. Analogous to Fig. 7.

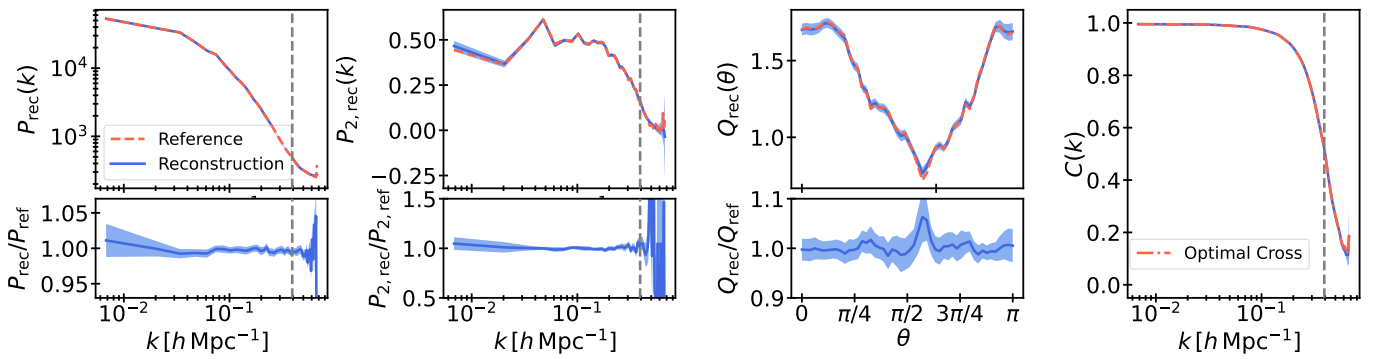


Fig. C.6. TEST3: Final conditions. Analogous to Fig. 8.