

Cosmological inference from Bayesian forward modelling of deep galaxy redshift surveys

Doogesh Kodi Ramanah^{1,2}, Guilhem Lavaux^{1,2}, Jens Jasche³, and Benjamin D. Wandelt^{1,2,4}

¹ Sorbonne Université, CNRS, UMR 7095, Institut d'Astrophysique de Paris, 98 bis bd Arago, 75014 Paris, France
e-mail: ramanah@iap.fr

² Sorbonne Universités, Institut Lagrange de Paris (ILP), 98 bis bd Arago, 75014 Paris, France

³ The Oskar Klein Centre, Department of Physics, Stockholm University, AlbaNova University Centre, 106 91 Stockholm, Sweden

⁴ Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, 10010 New York, NY, USA

Received 22 August 2018 / Accepted 7 November 2018

ABSTRACT

We present a large-scale Bayesian inference framework to constrain cosmological parameters using galaxy redshift surveys, via an application of the Alcock-Paczyński (AP) test. Our physical model of the non-linearly evolved density field, as probed by galaxy surveys, employs Lagrangian perturbation theory (LPT) to connect Gaussian initial conditions to the final density field, followed by a coordinate transformation to obtain the redshift space representation for comparison with data. We have implemented a Hamiltonian Monte Carlo sampler to generate realisations of three-dimensional (3D) primordial and present-day matter fluctuations from a non-Gaussian LPT-Poissonian density posterior given a set of observations. This hierarchical approach encodes a novel AP test, extracting several orders of magnitude more information from the cosmic expansion compared to classical approaches, to infer cosmological parameters and jointly reconstruct the underlying 3D dark matter density field. The novelty of this AP test lies in constraining the comoving-redshift transformation to infer the appropriate cosmology which yields isotropic correlations of the galaxy density field, with the underlying assumption relying purely on the geometrical symmetries of the cosmological principle. Such an AP test does not rely explicitly on modelling the full statistics of the field. We verified in depth via simulations that this renders our test robust to model misspecification. This leads to another crucial advantage, namely that the cosmological parameters exhibit extremely weak dependence on the currently unresolved phenomenon of galaxy bias, thereby circumventing a potentially key limitation. This is consequently among the first methods to extract a large fraction of information from statistics other than that of direct density contrast correlations, without being sensitive to the amplitude of density fluctuations. We perform several statistical efficiency and consistency tests on a mock galaxy catalogue, using the SDSS-III survey as template, taking into account the survey geometry and selection effects, to validate the Bayesian inference machinery implemented.

Key words. methods: data analysis – methods: statistical – cosmology: observations – large-scale structure of Universe – galaxies: statistics

1. Introduction

The past few decades have witnessed the advent of an array of galaxy redshift surveys, with the state-of-the-art catalogues mapping millions of galaxies with precision positioning and accurate redshifts. The Sloan Digital Sky Survey (SDSS; York et al. 2000; Abazajian et al. 2009; Ahn et al. 2014; Alam et al. 2015) and the Six Degree Field Galaxy Redshift Survey (6dFGRS; Jones et al. 2009) are two notable examples. Future cutting-edge surveys from the *Euclid* (Laureijs et al. 2011; Racca et al. 2016; Amendola et al. 2018) and Large Synoptic Survey Telescope (LSST; Ivezić et al. 2008) missions, currently under construction, further highlight the wealth of galaxy redshift data sets which would be available within a five to ten year time frame. Sophisticated and optimal data analysis techniques, in particular large-scale structure analysis methods, are in increasing demand to cope with the present and upcoming avalanches of cosmological and astrophysical data, and therefore optimise the scientific returns of the missions.

With the metamorphosis of cosmology into a precision (and data-driven) science, the three-dimensional (3D) large-scale structures have emerged as an essential probe of the dynamics of

structure formation and evolution to further our understanding of the Universe. The two-point statistics of the 3D matter distribution have developed into key tools to investigate various cosmological models and test different inflationary scenarios. Various techniques to measure the power spectrum and several reconstruction methods attempting to recover the underlying density field from galaxy observations are described in literature (e.g. Bertschinger & Dekel 1989; Bertschinger et al. 1991; Hoffman et al. 1994; Lahav et al. 1994; Fisher et al. 1995; Sheth 1995; Webster et al. 1997; Bistolos & Hoffman 1998; Schmoldt et al. 1999; Saunders et al. 2000; Zaroubi et al. 1999; Zaroubi 2002; Erdoğan et al. 2004, 2006), with the recent focus being on large-scale Bayesian inference methods (e.g. Kitaura & Enßlin 2008; Kitaura et al. 2009; Jasche & Kitaura 2010; Jasche et al. 2010a; Jasche & Wandelt 2012, 2013b; Jasche & Lavaux 2015, 2018). A formal and rigorous Bayesian framework provides the ideal setting to solve the ill-posed problem of inferring signals from noisy observations, while quantifying the corresponding statistical uncertainties.

The potential of such Bayesian algorithms to jointly infer cosmological constraints, nevertheless, has not yet been exploited. We present, for the first time, a non-linear Bayesian

inference framework for cosmological parameter inference from galaxy redshift surveys via an implementation of the Alcock-Paczyński (AP, [Alcock & Paczynski 1979](#)) test. We extend the hierarchical Bayesian inference machinery of Bayesian Origin Reconstruction from Galaxies (BORG; [Jasche & Wandelt 2013a](#)), originally developed for the non-linear reconstruction of large-scale structures, to constrain cosmological parameters. BORG encodes a physical model for gravitational structure formation, yielding a highly non-trivial Bayesian inverse problem. This consequently allows us to reformulate the standard problem of present 3D density field reconstruction as an inference problem for initial conditions at an earlier epoch from current galaxy observations. BORG builds upon the implementation of the Hamiltonian Monte Carlo (HMC) method ([Neal 1993](#)), initially introduced in the HAMILTONIAN DENSITY ESTIMATION and SAMPLING (HADES) algorithm ([Jasche & Kitaura 2010](#)), for efficiently sampling the high dimensional and non-linear parameter space of possible initial conditions at an earlier epoch.

In this work, the conceptual framework is to constrain the comoving-redshift coordinate transformation and therefore infer the appropriate cosmology which would result in isotropic correlations of the galaxy density field. The key aspect of this application of the AP test consequently lies in its robustness to a misspecified model and the approximations therein, yielding a near-optimal exploitation of the model predictions, without relying on its accuracy in modelling the scale dependence of the correlations of the density field. Here, we employ Lagrangian perturbation theory (LPT) as a physical description for the non-linear dynamics and perform a joint inference of initial conditions, and consequently the corresponding non-linearly evolved density fields and associated velocity fields, and cosmological parameters, from incomplete observations. This augmented framework with cosmological applications is designated as ALTAIR (ALcock-Paczyński consTrAIned Reconstruction).

The paper is organised as follows. In Sect. 2, the underlying principles of the AP test are outlined, followed by a description of the forward modelling approach and data model implemented in Sect. 3. We then test the algorithm in Sect. 5 on an artificially generated galaxy survey, with the mock generation procedure described in the preceding Sect. 4, by investigating its performance via statistical efficiency and consistency tests. In Sect. 6, we summarise the main aspects of our work and discuss further possible extensions to our algorithm in order to fully exploit its potential in deriving cosmological constraints. In Appendix A, we describe the LPT-Poissonian posterior implemented in this work, followed by the computation of the Jacobian of the comoving-redshift transformation in Appendix B. We provide a brief overview of the Hamiltonian sampling approach in Appendix C, and follow up by deriving the required equations of motion in Appendix D, with the numerical implementation outlined in Appendix E. We subsequently describe how we increase the efficiency of our cosmological parameter sampler via a rotation of the parameter space in Appendix F. Finally, we outline the derivation of the adjoint gradient for a generic 3D interpolation scheme in Appendix G.

2. The Alcock-Paczyński test

The Alcock-Paczyński (AP) test ([Alcock & Paczynski 1979](#)) is a cosmological test of the expansion of the Universe and its geometry. The main advantage of this test is that it is independent of the evolution of galaxies but depends only on the geometry of the

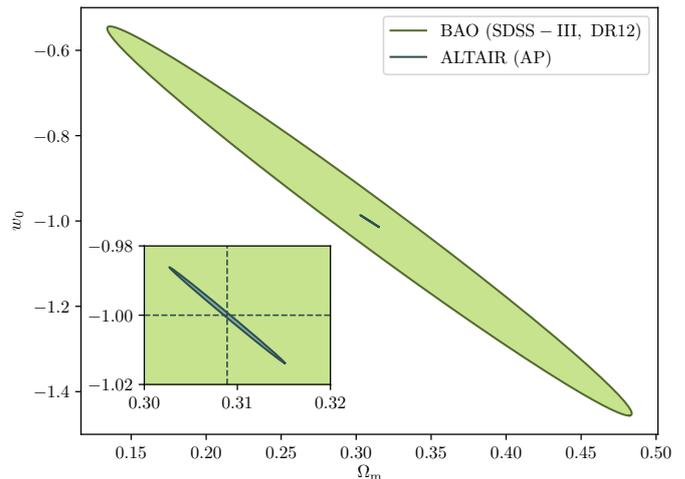


Fig. 1. Comparison of cosmological constraints from BAO measurements and our implementation of AP test in ALTAIR. The grey and green lines denote the 1σ confidence region, centred on the fiducial cosmological parameters, obtained from our AP test and BAO constraints from SDSS-III (DR 12; [Alam et al. 2017](#)), respectively. The BAO constraints have not been combined with *Planck* CMB measurements. This demonstrates the potentially unprecedented constraining power of our AP test compared to standard BAO analyses, as discussed in Sect. 2, with the inset focusing on the ALTAIR constraints where the fiducial cosmology is depicted in dashed lines. This error forecast is validated on a simulated analysis (cf. Fig. 6).

Universe. The assumption of incorrect cosmological parameters in data analysis produces distortions in the appearance of any spherical object or isotropic statistical distribution. The AP test provides a pathway to exploit this resulting spurious anisotropy to constrain the cosmological parameters. Here, we invoke the AP test to ensure that the underlying geometrical properties of isotropy of the Universe ([Friedmann 1922, 1924](#); [Lemaître 1927, 1931, 1933](#); [Robertson 1935, 1936a,b](#); [Walker 1937](#); [Saadeh et al. 2016](#)) are maintained. As such, the key underlying assumption adopted in this work relies purely on the geometrical symmetries of the cosmological principle. As a result, such a test does not employ the growth of structures to constrain cosmology, unlike cluster abundance (e.g. [Wang & Steinhardt 1998](#)).

The AP test, and various formulations thereof, have been studied extensively in the context of galaxy and quasar surveys (e.g. [Phillipps 1994](#); [Ryden 1995](#); [Ballinger et al. 1996](#); [Matsubara & Suto 1996](#); [Popowski et al. 1998](#); [de Laix & Starkman 1998](#); [López-Corredoira 2014](#); [Li et al. 2014, 2016](#)). Variants of the AP test have also been successfully applied to cosmic voids (e.g. [Sutter et al. 2012, 2014](#); [Lavaux & Wandelt 2012](#); [Hamaus et al. 2014, 2015, 2016](#)) and also to other cosmological observables like supernovae ([Blake et al. 2011](#)), the Lyman- α forest ([Hui et al. 1999](#)) and 21 centimetre emission maps ([Nusser 2005](#); [Barkana 2006](#)).

With baryon acoustic oscillations (BAOs) being a robust standard ruler, the AP test has been utilised for the simultaneous measurement of the Hubble parameter and angular diameter distance of distant galaxies (e.g. [Seo & Eisenstein 2003](#); [Blake & Glazebrook 2003](#); [Glazebrook & Blake 2005](#); [Padmanabhan & White 2008](#); [Shoji et al. 2009](#)). In Fig. 1, we depict the 1σ confidence region of the cosmological constraints inferred via our implementation of the AP test. As a comparison, we also indicate the corresponding confidence region obtained via BAO measurements from the SDSS-III (Data Release 12; [Alam et al. 2017](#)). These BAO constraints have not been combined with *Planck*

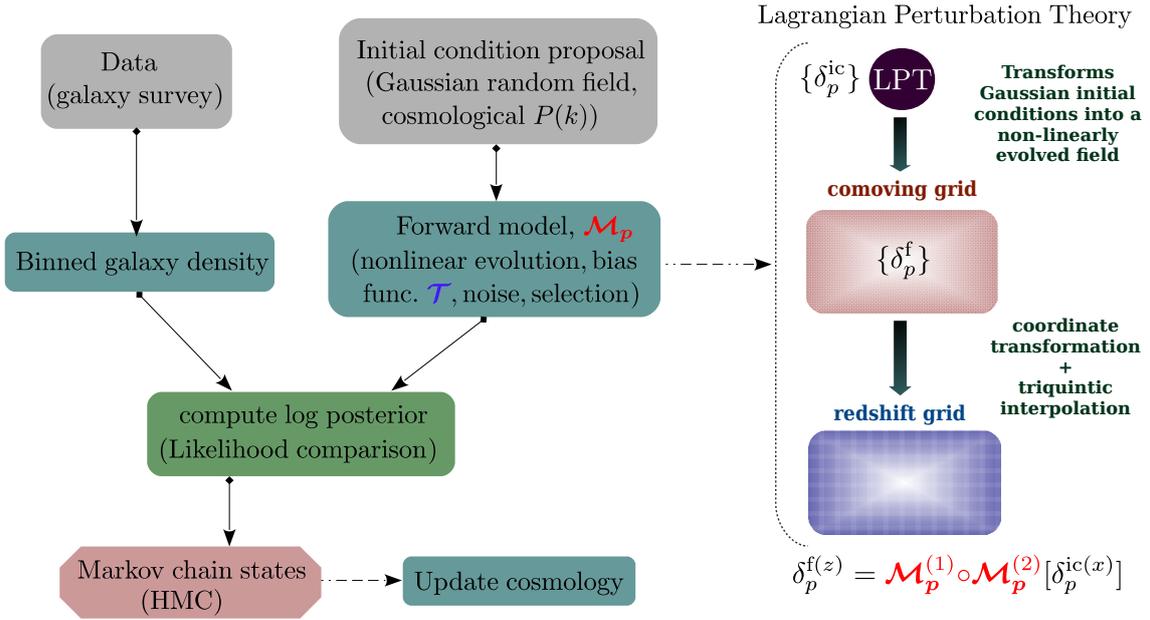


Fig. 2. Schematic representation of the reconstruction pipeline. The forward model consists of a chain of various components for the non-linear evolution from initial conditions and the subsequent transformation from comoving to redshift space for the application of the AP test. This consequently transforms the initial density field into a set of predicted observables, i.e. a galaxy distribution in redshift space, for comparison with data via a likelihood or posterior analysis.

measurements, which would significantly tighten the constraints. Nevertheless, this highlights the significant potential constraining power of our AP test compared to standard BAO analyses, while being at least as robust. While this improvement is extremely substantial for the mock SDSS-III survey considered here, we will investigate to what extent the above promise holds when applied to actual SDSS-III data in a follow-up work, as unknown systematics represent a potential caveat. We discuss Fig. 1 in more depth in Sect. 5.

The crucial aspect of our AP test is that it does not assume that the correlation function is correctly modelled. This robustness to a misspecified model is illustrated explicitly in Sect. 5, where we demonstrate that the shape of the prior power spectrum adopted in the inference framework does not impact on the inferred cosmological constraints (cf. Fig. 10). As a result of this robustness, our AP test has a definite edge over standard approaches. Moreover, it has been pointed out that other cosmological tests, such as the luminosity distance – redshift relation, can be considered as generalised formulations of the AP test (Mukherjee & Wandelt 2018), further underlining the strength of the approach presented in this work.

3. The forward modelling approach

The large-scale structure (LSS) posterior implemented in this work, based on the BORG framework (Jasche & Wandelt 2013a), is described in depth in Appendix A. A key component of the inference framework is the forward model \mathcal{M}_p which links the initial conditions $\delta_p^{\text{ic}(r)}$ to the redshift space representation of the evolved density field $\delta_p^{\text{f}(z)}$ as follows:

$$\begin{aligned} \delta_p^{\text{f}(z)} &= \mathcal{M}_p(1 + \delta_p^{\text{ic}(r)}) = \mathcal{M}_p^{(1)} \circ \mathcal{M}_p^{(2)}(1 + \delta_p^{\text{ic}(r)}) = \mathcal{M}_p^{(1)}(\rho_p^{\text{f}(r)}) \\ &= \mathcal{J}_p \left(\sum_{i,j} \mathcal{E}_{ij}^{-1} \rho_j^{\text{f}(r)} x_p^{\alpha(i)} y_p^{\beta(i)} z_p^{\gamma(i)} \right) - 1, \end{aligned} \quad (1)$$

where $\rho_p^{\text{f}(r)} \equiv 1 + \delta_p^{\text{f}(r)}$ is the final density field in comoving space. The forward model consists of two components, $\mathcal{M}_p = \mathcal{M}_p^{(1)} \circ \mathcal{M}_p^{(2)}$. The first component, $\mathcal{M}_p^{(1)} \equiv \mathcal{G}_p(a, \{\delta_p^{\text{ic}}\})$, contains a physical description of the non-linear dynamics, and consequently propagates the initial conditions forward in time using LPT, yielding a non-linearly evolved final density field in comoving space, $\delta_p^{\text{f}(r)}$.

To encode the AP test, we incorporate another component in the forward model that takes care of the coordinate transformation from comoving (\mathbf{r}) to redshift (z) space, encoded in $\mathcal{M}_p^{(1)}$ (cf. Fig. 2). Schematically, we construct a second grid in redshift space, which involves a triquintic interpolation (fifth order interpolation scheme in three dimensions) on the comoving grid. This interpolation scheme is described in Appendix G, with the notation employed in Eq. (1) clearly laid out. The corresponding Jacobian factor of this transformation, $|\mathcal{J}_p^z|$ (cf. Appendix B), entails cosmological dependence and is consequently included in the AP test as well as through the direct coordinate dependence \mathcal{E}_{ij} .

The redshift space representation then allows for comparison with data via the likelihood or posterior. The essence of this AP test to constrain cosmological parameters can be summarised as follows: The Bayesian inference machinery explores the various cosmological expansion histories and selects the cosmology-dependent evolution pathways which result in isotropic correlations of the galaxy density field.

Figure 2 illustrates the reconstruction scheme implemented in ALTAIR. First, galaxies are projected from the survey onto a 3D grid, such that we have a distribution of galaxies in redshift space and this constitutes our observable. We then generate a 3D density field according to Gaussian initial conditions (homogeneous prior) with a reference power spectrum, typically Λ CDM cosmology. The forward model subsequently transforms the initial density field into a set of predicted observables which are then compared to data via a likelihood or posterior analysis.

And conversely, given the position of galaxies, we can infer this density field.

While we implement LPT to approximately describe gravitational non-linear structure formation in this work, other more adequate physical descriptions such as 2LPT or the non-perturbative particle mesh (see recent upgrade of BORG in Jasche & Lavaux 2018) can be straightforwardly employed, within the flexible block sampling approach described in Appendix A.3, by upgrading the first component, $\mathcal{M}_p^{(2)} \equiv \mathcal{G}_p(a, \{\delta_p^{\text{ic}}\})$, of our forward model (cf. Fig. 2). Nevertheless, our implementation of the AP test exploits essentially the isotropy of the correlation function, such that there is no explicit dependence on the accuracy of modelling the scale dependence of the correlations, rendering this method robust to a misspecified model and the approximations therein.

3.1. The galaxy data model

Galaxies can be considered as tracers of the matter fluctuations since they follow the gravitational potential of the underlying matter distribution, with the statistical uncertainty due to the discrete nature of the galaxy distribution usually modelled by a Poissonian distribution (e.g. Layzer 1956; Peebles 1980). Poissonian likelihoods have emerged as the standard for non-linear LSS inference (e.g. Jasche et al. 2010b; Jasche & Kitaura 2010; Kitaura et al. 2010). The Poissonian likelihood distribution implemented in ALTAIR, for multiple subcatalogues or galaxy observations labelled by the index g , can be expressed as follows:

$$\mathcal{L}(\{N_p^g\}|\lambda_p^g) = \prod_p \frac{(\lambda_p^g)^{N_p^g} e^{-\lambda_p^g}}{N_p^g!}, \quad (2)$$

where N_p^g is the observed galaxy number counts in redshift space in the given voxel p . λ_p^g is the expected number of galaxies at this given position and is related to the final density field δ_p^f , in redshift space, via

$$\lambda_p^g(\{\delta_p^f\}, \{\theta_i\}) = R_p^g(\theta_i) \mathcal{T} \left[1 + \delta_p^f \right], \quad (3)$$

where R_p^g is the overall linear response operator of the survey that incorporates the survey geometry and selection effects, and θ_i corresponds to a set of cosmological parameters. \mathcal{T} is the galaxy biasing model which accounts for the fact that galaxies do not trace exactly the underlying matter distribution, and are therefore biased tracers with clustering properties that do not exactly mirror those of dark matter (Kaiser 1984). This is currently one of the most challenging and unresolved issues hindering the analysis of galaxy distributions in non-linear regimes (e.g. see the review by Desjacques et al. 2018; Schmidt et al. 2018).

In this work, we adopt the standard approach of a local, but non-linear bias function, in particular, the phenomenological model proposed by Neyrinck et al. (2014), such that the above Poisson intensity field can be expressed as

$$\lambda_p^g(\{\delta_p^f\}, \{\theta_i\}, \{\bar{N}^g\}, \{b_i^g\}) = R_p^g \bar{N}^g \left[1 + \delta_p^f \right]^\beta e^{-\rho^g [1 + \delta_p^f]^{-\epsilon^g}}, \quad (4)$$

where the bias function, described by four parameters, \bar{N}^g , the mean density of tracers, and $\{b_i^g\} = \{\beta, \rho^g, \epsilon^g\}$, is a truncated power law bias model with the additional exponential function suppressing galaxy clustering in under dense regions. This bias model, with a power law and an exponential at low densities, were found to be in good agreement with standard excursion set

and local-growth-factor models (for more details, see Neyrinck et al. 2014). The main limitation of this bias model is that it is purely local. Nevertheless, it is more adequate than a simplistic linear bias model and mitigates in practice the deficiencies of our physical model (LPT) at the considered resolution. The expected number of galaxies can subsequently be related to the initial conditions δ_p^{ic} via the forward model \mathcal{M}_p , as described above, due to the deterministic nature of structure formation, as implied by the Dirac delta function in Eq. (A.2).

The logarithm of the likelihood from Eq. (2) can therefore be expressed, in terms of the initial conditions, as

$$\begin{aligned} \ln \mathcal{L} \left[\{N_p^g\} | \mathcal{M}_p(\{\delta_p^{\text{ic}}\}), \{\theta_i\}, \{\bar{N}^g\}, \{b_i^g\} \right] \\ = - \sum_p \left\{ \lambda_p^g(\{\delta_p^f\}, \{\theta_i\}, \{\bar{N}^g\}, \{b_i^g\}) \right. \\ \left. - N_p^g \ln \left[\lambda_p^g(\{\delta_p^f\}, \{\theta_i\}, \{\bar{N}^g\}, \{b_i^g\}) \right] + \ln(N_p^g!) \right\}. \end{aligned} \quad (5)$$

We therefore have a likelihood distribution that encodes the statistical process describing the generation of galaxy observations given a specific realisation of 3D initial conditions. This data model is inherently non-linear as a result of the galaxy biasing model employed and also due to the signal dependence of Poissonian noise, which does not behave as an additive nuisance.

3.2. The augmented joint posterior distribution

The augmented joint posterior distribution corresponds to the following:

$$\begin{aligned} \mathcal{P}(\{\delta_p^{\text{ic}}\}, \{\bar{N}^g\}, \{b_i^g\}, \{\theta_i\} | \{N_p^g\}, \mathbf{S}) \\ \propto \mathcal{L} \left[\{N_p^g\} | \mathcal{M}_p(\{\delta_p^{\text{ic}}\}), \{\theta_i\}, \{\bar{N}^g\}, \{b_i^g\} \right] \\ \times \Pi(\{\delta_p^{\text{ic}}\} | \mathbf{S}) \Pi(\{\bar{N}^g\}, \{b_i^g\}) \Pi(\{\theta_i\}), \end{aligned} \quad (6)$$

where the Π 's correspond to the respective priors for each parameter. Hence, given our forward model \mathcal{M} , which incorporates sequential components of structure formation and coordinate transformation, as described above, the complex task of modelling accurate priors for the statistical behaviour of present-day matter fluctuations can be recast into a Bayesian inference problem for the initial conditions. From this joint posterior distribution, we can construct the various conditional posterior distributions for each parameter of interest. The modular statistical programming approach adopted in outlined in Appendix A.3.

Since the non-linear LSS analysis has been reformulated as an initial conditions statistical inference problem, as described by the joint posterior distribution Eq. (6), this method depends solely on forward evaluations, and consequently has a definite edge over traditional approaches of initial conditions inference that require backward integration of the equations of motion or the inversion of the flow of time (e.g. Nusser & Dekel 1992). The latter methods are prone to erroneous fluctuations in the initial density and velocity fields, resulting from spurious growth of decaying modes. Moreover, such schemes are hindered by survey incompleteness which requires the knowledge of the complex and, as yet, unknown multi-variate probability distribution for the matter fluctuations, to render the backward integration of non-linear models physically meaningful via constrained realisations. In comparison, the forward modelling approach here conveniently accounts for survey masks and statistical uncertainties in the initial conditions, which amounts to modelling straightforward uncorrelated Gaussian processes, to generate data constrained realisations of the initial and evolved density fields.

3.3. The cosmological parameter posterior distribution

In this work, we sample the present-day values of matter density and dark energy equation of state parameters, $\{\theta_i\} = \{\Omega_m, w_0\}$, via the following conditional posterior distribution,

$$\begin{aligned} \mathcal{P}(\{\theta_i\}|\{N_p^g\}, \{\delta_p^{\text{ic}}\}, \{\bar{N}^g\}, \{b_i^g\}, \mathbf{S}) \\ = \mathcal{P}(\{\theta_i\}|\{N_p^g\}, \{\delta_p^{\text{ic}}\}, \{\bar{N}^g\}, \{b_i^g\}), \end{aligned} \quad (7)$$

assuming conditional independence of the cosmological power spectrum, that is, Fourier transform of the covariance matrix \mathbf{S} , once the density field is known. This assumption holds as we are only probing the cosmological expansion in this work, with the power spectrum anchored with a fiducial cosmology. We quantitatively demonstrate the validity of this assumption in Sect. 5 by comparing the entropy of prior information against that of posterior information (cf. Fig. 9). We defer power spectrum sampling to a future work. Applying Bayes' identity, and using the joint posterior distribution from Eq. (6), we obtain

$$\begin{aligned} \mathcal{P}(\{\theta_i\}|\{N_p^g\}, \{\delta_p^{\text{ic}}\}, \{\bar{N}^g\}, \{b_i^g\}) \\ = \frac{\mathcal{P}(\{\theta_i\}, \{\delta_p^{\text{ic}}\}, \{\bar{N}^g\}, \{b_i^g\}|\{N_p^g\})}{\mathcal{P}(\{\delta_p^{\text{ic}}\}, \{\bar{N}^g\}, \{b_i^g\}|\{N_p^g\})} \\ = \frac{\mathcal{L}[\{N_p^g\}|\mathcal{M}_p(\{\delta_p^{\text{ic}}\}), \{\theta_i\}, \{\bar{N}^g\}, \{b_i^g\}]}{\mathcal{P}(\{\delta_p^{\text{ic}}\}, \{\bar{N}^g\}, \{b_i^g\}|\{N_p^g\})} \\ \times \Pi(\{\delta_p^{\text{ic}}\}|\mathbf{S}) \Pi(\{\bar{N}^g\}, \{b_i^g\}) \Pi(\{\theta_i\}) \\ \propto \mathcal{L}[\{N_p^g\}|\mathcal{M}_p(\{\delta_p^{\text{ic}}\}), \{\theta_i\}, \{\bar{N}^g\}, \{b_i^g\}] \times \Pi(\{\theta_i\}), \end{aligned} \quad (8)$$

after omitting the terms without any cosmological parameter dependence.

Since this work attempts to demonstrate the capabilities of ALTAIR to constrain the cosmological parameters, as proof of concept, we set uniform prior distributions on θ_i . To sample from the above marginal posterior distribution, we make use of a slice sampling procedure (e.g. Neal 2000, 2003). After obtaining a realisation of θ_i , we need to update the comoving-redshift coordinate transformation in the forward model.

We adopt a dynamical dark energy model, in particular, the standard Chevallier-Polarski-Linder (CPL) parameterisation (Chevallier & Polarski 2001; Linder 2003), where the evolution of the dark energy equation of state parameter w is a linear function of the scale factor a , as follows:

$$w = w_0 + (1 - a)w_a. \quad (9)$$

In this work, we set $w_a = 0$ and infer the present-day value w_0 . Moreover, we impose the assumption of flatness, that is, $\Omega_k = 0$, such that the dark energy density is $\Omega_{\text{de}} = 1 - \Omega_m$.

Due to the correlation between Ω_m and w_0 , we perform a rotation of the (Ω_m, w_0) parameter space, using orthonormal basis transformations derived from the covariance matrix, to improve the efficiency of the slice sampler. This procedure is outlined in Appendix F, with the significant gain in efficiency illustrated in Fig. F.1. The corresponding sampler for the bias parameters is outlined in Appendix A.4.

4. Generation of a mock galaxy catalogue

We describe the generation of an artificial galaxy survey using as template the Sloan Digital Sky Survey (SDSS-III), consisting of four galaxy subcatalogues, to validate the methodology

described in the previous sections. The procedure implemented here for the mock generation is essentially based on the descriptions provided in Jasche & Kitaura (2010) and Jasche & Wandelt (2013a).

At first, we generate a realisation for the initial density contrast δ_p^{ic} from a normal distribution with zero mean and covariance corresponding to an underlying cosmological power spectrum for the matter distribution. This power spectrum, including baryonic wiggles, is computed following the prescription provided in Eisenstein & Hu (1998, 1999), assuming a standard Λ cold dark matter (Λ CDM) cosmology with the set of cosmological parameters ($\Omega_m = 0.3089$, $\Omega_\Lambda = 0.6911$, $\Omega_b = 0.0486$, $h = 0.6774$, $\sigma_8 = 0.8159$, $n_s = 0.9667$) from Planck (Planck Collaboration XIII 2016). This yields a 3D Gaussian initial density field in an equidistant Cartesian grid with $N_{\text{side}} = 128$, consisting of 128^3 voxels, where each voxel corresponds to a discretised volume element with a physical voxel size of $15.6 h^{-1}$ Mpc, and comoving box length of $4000 h^{-1}$ Mpc. As a result, this implies a total of $\sim 2.1 \times 10^6$ inference parameters, corresponding to the amplitude of the primordial density fluctuations at the respective grid nodes.

To ensure a sufficient resolution of inferred final density fields, we oversample the initial density field by a factor of eight, thereby evaluating the LPT model with 256^3 particles in every sampling step. The following step is to scale this 3D distribution of initial conditions to a cosmological scale factor of $a_{\text{init}} = 0.001$ via multiplication with a cosmological growth factor $D^+(a_{\text{init}})$. These initial conditions are subsequently evolved forward in time, in non-linear fashion, with LPT providing an approximate description of gravitational LSS formation. We then construct a final 3D non-linearly evolved density field δ_p^f from the resulting particle distribution via the cloud-in-cell (CIC) method (e.g. Hockney & Eastwood 1988).

To generate the mock galaxy survey, we essentially need to simulate the inhomogeneous Poissonian process described by Eq. (2), by drawing random samples from the distribution, on top of the final density field δ_p^f . In this work, we generate a mock data set with realistic features emulating the highly structured survey geometry and selection effects of the SDSS-III survey, with the observed sky completeness depicted in the left and right panels of Fig. 3, respectively, for the CMASS and LOW-Z components of the SDSS-III survey. To account for the different selection effects in the northern and southern galactic planes, each component is further divided into two subcatalogues, corresponding to north/south galactic caps (NGC/SGC). The respective radial selection functions for these four subcatalogues are illustrated in Fig. 4. Here, the selection functions are numerical estimates obtained by binning the corresponding distribution of tracers $N(d)$ in the CMASS and LOW-Z components (e.g. Ross et al. 2017), where d is the comoving distance from the observer.

The projection of the completeness functions into the 3D volume produces the 3D observation mask C_p^g . The survey properties are described by the galaxy selection function and the 3D completeness function, with the product of these two functions yielding the survey response operator R_p^g . More specifically, it is the average of the product of the 2D survey geometry $C_p^g(\hat{x}) = C_{p(\hat{x})}^g$ and the selection function $f(x)$ at each volume element of the 3D grid:

$$R_p^g = \frac{1}{|\mathcal{V}_p|} \int_{\mathcal{V}_p} d^3x C_p^g(\hat{x}) f^g(x), \quad (10)$$

where the volume occupied by the p th voxel is indicated by \mathcal{V}_p , with $|\mathcal{V}|$ the volume of the set \mathcal{V} .

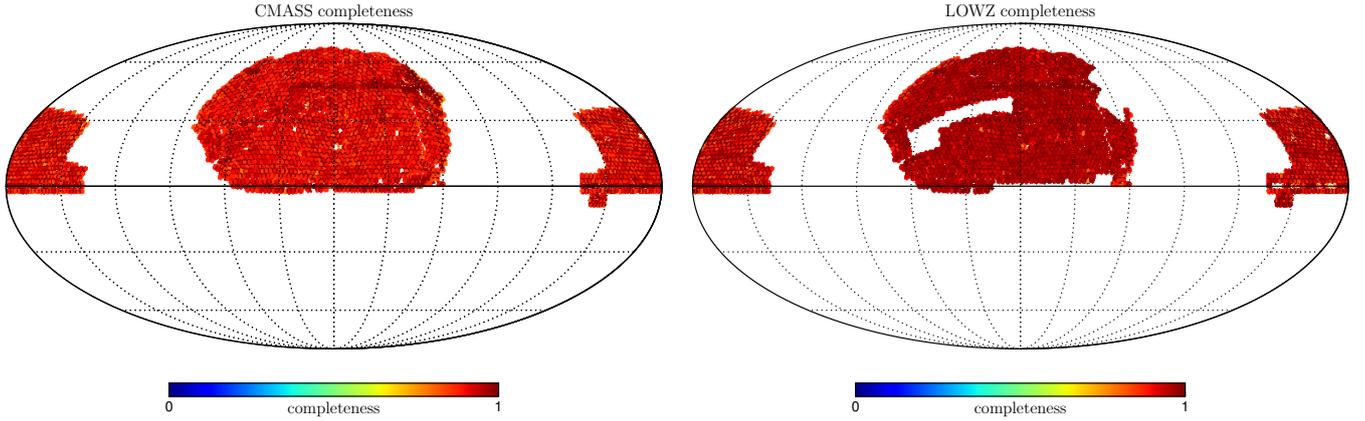


Fig. 3. Observed sky completeness, used to generate and analyse the mock catalogue in this work, are illustrated in the *left* and *right* panels, corresponding to the CMASS and LOW-Z components of the SDSS-III survey, respectively.

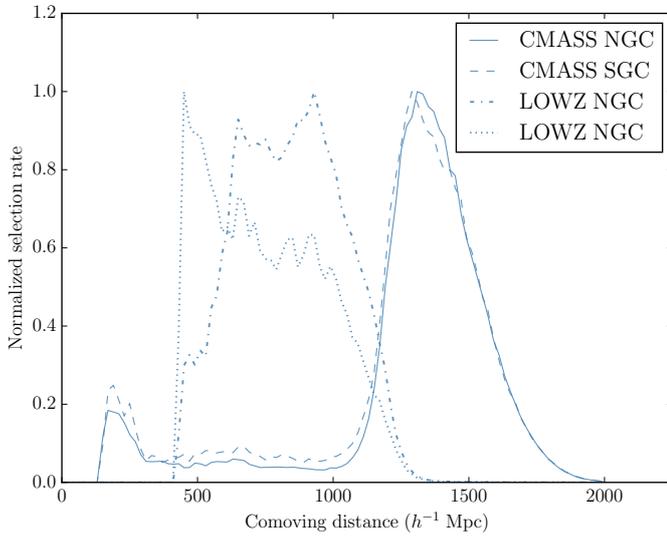


Fig. 4. Radial selection functions for the CMASS sample, in solid and dashed lines for the north galactic cap (NGC) and south galactic cap (SGC), respectively. The corresponding radial selection functions for the LOW-Z sample are depicted in dash-dotted (NGC) and dotted lines (SGC). These selection functions are used to generate the mock data to emulate features of the actual SDSS-III BOSS data.

Finally, we generate the four artificial galaxy subcatalogues, labelled by g , by Poisson sampling on the grid with $\bar{N}_p^g = \{110.42, 122.94, 71.43, 205.48\}$, resulting in a total number of 997828 galaxies. The values of \bar{N}_p^g are chosen such that the mock catalogue reflects the characteristics of the actual SDSS-III data which contains around one million tracers.

5. Results

We perform a series of tests to evaluate the performance of our algorithm in a realistic context by applying ALTAIR on the simulated SDSS-III-like galaxy catalogue described in the previous section. In particular, the focus is on the burn-in and convergence behaviour of our method, which are key indicators of the overall numerical feasibility and statistical efficiency for real data applications. To validate the conceptual framework for cosmological parameter inference and the robustness of our implementation of the AP test, the reproducibility of the input

cosmology and the correlations with the other inferred quantities such as galaxy bias are also of interest.

The Markov chains for the cosmological parameters, displayed in Fig. 5, were initialised with an over-dispersed state. This figure consequently illustrates an initial burn-in phase, lasting ~ 250 MCMC steps, where the Markov chains follow a persistent drift towards the high probability region of the parameter space. The rotation of the (Ω_m, w_0) parameter space before slice sampling, as described in Appendix F, reduces the burn-in period significantly, by roughly a factor of five, as shown in the right panel of Fig. F.1, resulting in improved sampling efficiency.

The corresponding marginal and joint posterior distributions for the cosmological parameters are displayed in Fig. 6, demonstrating the capability of ALTAIR to infer tight constraints from galaxy redshift surveys. This robust AP test fully exploits the high information content from the cosmic expansion as a result of probing a deep redshift range, where the distortion is more pronounced, yielding the following cosmological constraints: $\Omega_m = 0.3080 \pm 0.0036$ and $w_0 = -0.998 \pm 0.008$. As a comparison, the SDSS-III (DR12, BAO + *Planck*) constraints are as follows: $\Omega_m = 0.310 \pm 0.005$ and $w_0 = -1.01 \pm 0.06$ (Alam et al. 2017), further highlighting the significant constraining power of our AP test. We acknowledge the significant difference in the size of uncertainties. A back of the envelope computation of the information gain is as follows: Considering a sphere of $100 h^{-1}$ Mpc for BAO against all voxels in $4000 h^{-1}$ Mpc, $N_{\text{BAO}} = 4000^3 / (4\pi \times 100^3 / 3) = 15278$, compared to $N_{\text{vox}} = 128^3$, yields an improvement of $\sqrt{N_{\text{vox}}/N_{\text{BAO}}} = 11.7$, which provides an order of magnitude of our uncertainties on the cosmological parameters. This is an approximate attempt to quantify the information gain from including smaller scales (in our work, $\sim 0.17 h^{-1}$ Mpc) than the BAO scale by essentially counting the number of modes. However, the above argument does not imply that employing finer resolutions will result in an infinite gain of information. There is a saturation of information at a certain resolution due to the slow variation in the density fields across neighbouring voxels, such that further refinement beyond this limit will not yield any additional information.

To verify that the cosmological information stems purely from the geometric distortion due to the cosmic expansion, that is, the AP test, we perform the following experiment: We deactivate the cosmic expansion component in our forward model and sample the cosmological parameters using only the LPT evolved density field. We consequently recover the

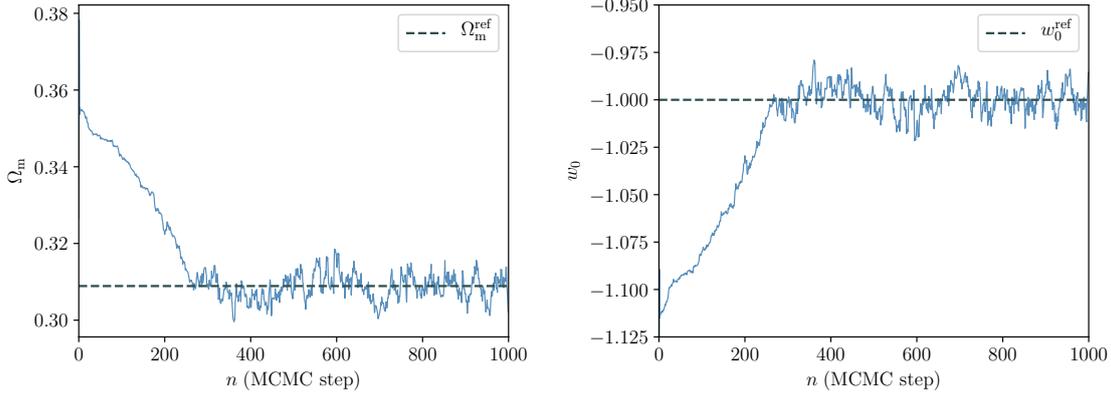


Fig. 5. MCMC chains for the cosmological parameters, for the first 1000 samples, with the reference cosmology employed in the mock generation indicated by the horizontal dashed lines. An initial burn-in phase lasting ~ 250 Markov transitions is illustrated by the coherent drift of the Markov chain towards the preferred region in parameter space.

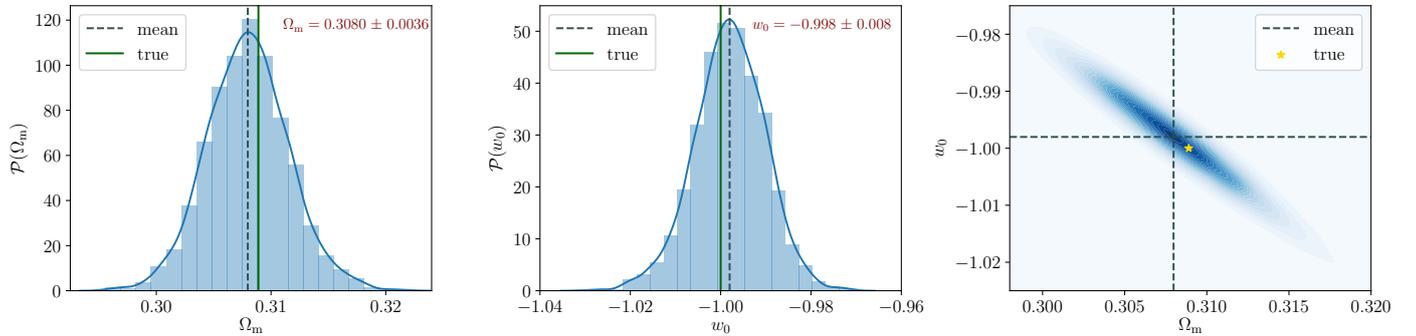


Fig. 6. Marginal posteriors for Ω_m (left panel) and the dark energy equation of state, w_0 (middle panel), for ~ 3000 MCMC realisations, ignoring the burn-in phase of ~ 250 Markov steps. The corresponding mean and standard deviation for each parameter are indicated in the top right corner of each plot. The joint posterior (right panel) for Ω_m and w_0 , depicting the high level of correlation between these two parameters. The highly informative distortion due to the cosmic expansion, as a result of probing a deep redshift range, yields extremely tight constraints on the above cosmological parameters. As a consistency test, this validates our implementation of the AP test to correctly recover the input cosmology.

corresponding prior distributions for the marginal posteriors of the cosmological parameters. As a result, this implies that the information derives purely from the geometry and not from the clustering of the non-linearly evolved density field, at least for the test case with the physical voxel size considered in this work.

The mean initial and final density fields computed from ~ 4000 realisations (ignoring the burn-in phase) are illustrated in the left panel of Fig. 7 for a particular slice of the 3D field. The maps reveal that on average we can recover highly detailed and well-defined structures in the observed regions. In particular, the filamentary nature of the non-linearly evolved density field can be clearly seen, while the Gaussian nature of the initial conditions can also be deduced visually. However, in poorly or not observed regions, the ensemble mean density field approaches the cosmic mean density, as expected, since regions lacking any observational information should on average reflect the cosmic mean. The uncertainty in these regions is accurately accounted for in the inference process, as demonstrated by the right panel, since each sampled density field is a constrained realisation, which means that these regions are augmented with statistically correct information.

In the top panel of Fig. 8, we illustrate the power spectra reconstructed from all the realisations of 3D initial density field, obtained from the posterior via the HMC sampler. This is a self-consistency test to confirm that the sampled density fields are

in agreement with the reference (prior) power spectrum adopted for the mock generation, as substantiated quantitatively in the bottom panel, where the ratio of the a posteriori power spectra to the prior distribution is shown. The measured power spectra therefore demonstrate that the individual realisations possess the correct power throughout the entire domain of Fourier modes considered in this work. Moreover, they do not exhibit any spurious power artefacts typically resulting from erroneous treatment of survey characteristics, such as survey geometry and selection effects, and galaxy biases, implying that such effects have been properly accounted for.

In order to verify the impact of the prior power spectrum on the actual inference of cosmological parameters, we illustrate, in Fig. 9, the distributions of power spectra computed using the inferred cosmology and the prior analytic prescription (Eisenstein & Hu 1998, 1999) and the reconstructed power spectra from Fig. 8, via their respective summary statistics, normalised with respect to the fiducial power spectrum. The scatter in the latter a posteriori power spectra reconstructed from the sampled density field realisations is significantly higher than distribution of the former prior spectra. This implies that the entropy of prior information is much lower than that of posterior information, thereby demonstrating that the prior power spectrum does not influence the cosmological parameter inference via the AP test and justifying the assumption made in Sect. 3.3.

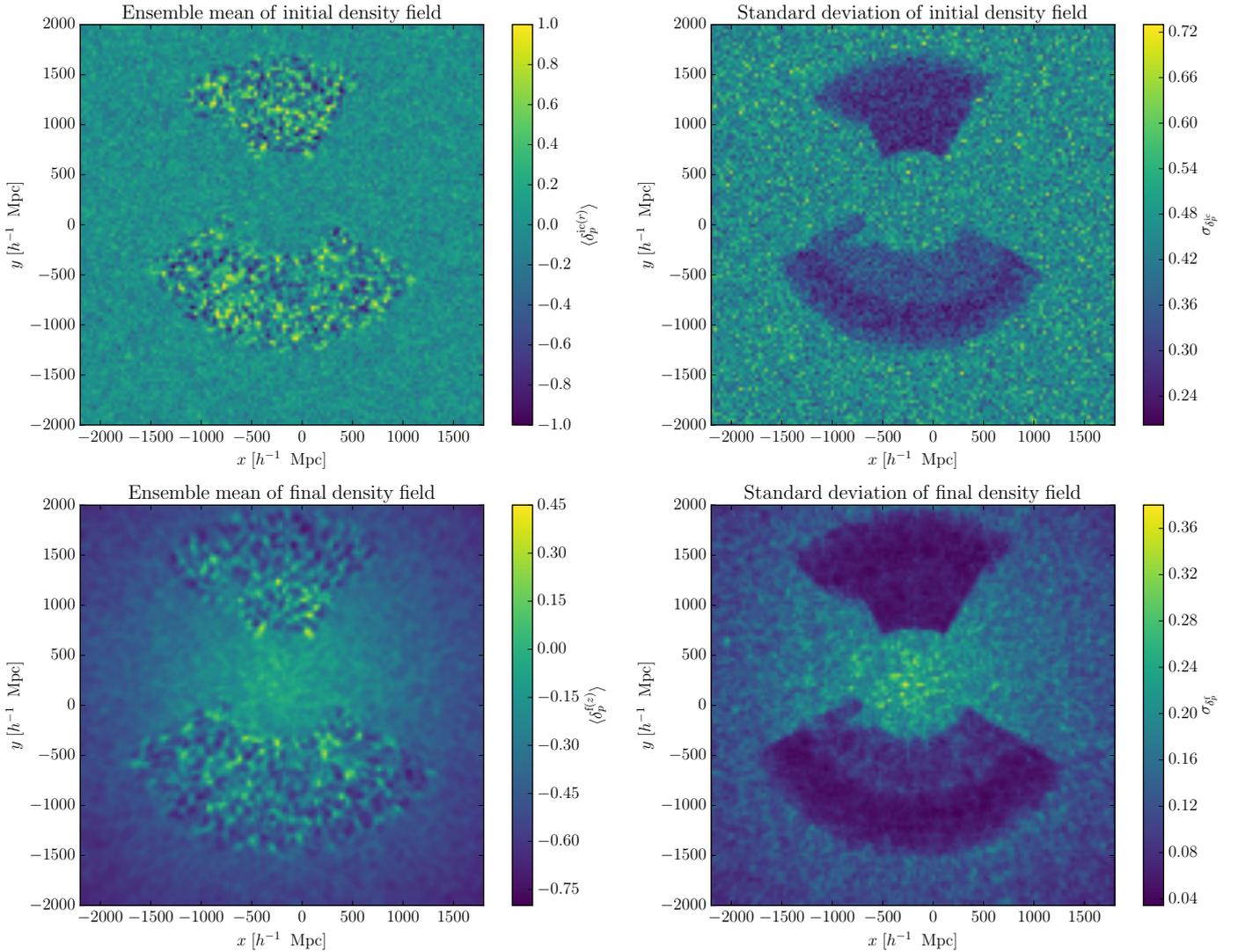


Fig. 7. Mean and standard deviation maps for the initial (*top panels*) and final density fields (*bottom panels*), computed from the MCMC realisations, with the same slice through the 3D density fields being illustrated above. In unobserved or masked regions, the density fields are not constrained by data, and they average out to the cosmic mean density. However, in observed regions, the Gaussian nature of the initial conditions and the filamentary nature of the non-linearly evolved density field is manifest. The corresponding variance is therefore higher in regions devoid of data.

We further demonstrate this robustness of our AP test by employing a modified prior power spectrum in the inference procedure. By adopting a different cosmology ($\Omega_m = 0.40$, $w_0 = -0.85$), we modify the shape of the power spectrum, and subsequently apply `ALTAIR` on the same mock catalogue from Sect. 4. As shown in Fig. 10, we recover the fiducial cosmological parameters employed in the mock generation, although with slightly larger uncertainties than for the original run by roughly 15%. This test case therefore explicitly highlights the robustness of our implementation of the AP test to a misspecified model since it does not optimise the information from the scale dependence of the correlations of the density field, but rather from the isotropy of the field.

From the correlation matrix of Ω_m , w_0 and the galaxy bias parameters for each of the four subcatalogues, illustrated in Fig. 11, we deduce the extremely weak correlation between the cosmological constraints and the bias. This is a key positive aspect of our method, as galaxy bias remains nevertheless a highly active and challenging field of research (see, for e.g.,

[Desjacques et al. 2018](#)), due to its complex non-linear behaviour on intermediate and small scales, which may potentially limit the effectiveness of traditional methods of cosmological parameter inference ([Pollina et al. 2018](#)).

Moreover, this insensitivity to the galaxy bias implies that our method does not rely on absolute density fluctuation amplitudes, but on the actual location of matter. This entails that our AP test exploits the geometrical structure of the density field and not its absolute amplitude since the power spectrum does not influence the inferred cosmological constraints. To the best of our knowledge, this is a novel aspect of cosmological inference, with most of our current understanding of cosmology based on measurements of density contrast amplitudes. We present, therefore, one of the first methods which extracts a large fraction of information from statistics other than that of direct density contrast correlations, without relying on the power spectrum or bispectrum. Our method consequently yields complementary information to state-of-the-art methods.

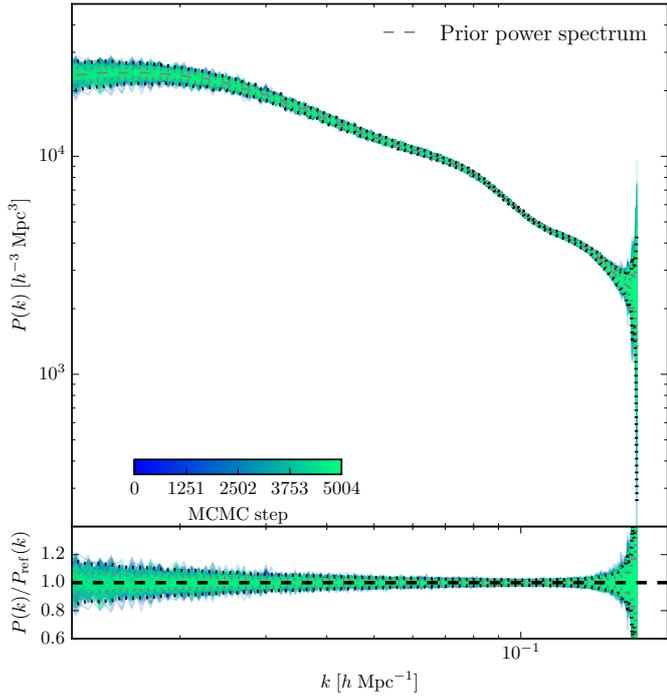


Fig. 8. *Top panel:* reconstructed power spectra from the inferred posterior initial density field realisations for all sampling steps of the Markov chain. The MCMC samples are colour-coded according to their sample number, as indicated by the colour bar. This is a self-consistency test to verify whether the sampled density fields are in accordance with the reference power spectrum, depicted in dashed lines, employed in the mock generation. *Bottom panel:* deviation from the true underlying solution, with the thick dashed lines representing the prior power spectrum. The thin grey and black dotted lines correspond to the Gaussian 1σ and 2σ limits, respectively.

6. Summary and conclusions

We presented the implementation of a robust AP test that performs a detailed fit of the cosmological expansion via a non-linear and hierarchical Bayesian LSS inference framework. This forward modelling approach employs LPT as a physical description for the non-linear dynamics and sequentially encodes the cosmic expansion effect for joint inference of the cosmological parameters and underlying 3D density fields, while also fitting the mean density of tracers and bias parameters. In essence, this inference machinery explores the various cosmological expansion histories and selects the cosmology-dependent evolution pathways which yield isotropic correlations of the galaxy density field, thereby constraining cosmology.

We demonstrated the application of our algorithm *ALTAIR* on an artificially generated galaxy catalogue, consisting of four sub-catalogues, that emulates the highly structured survey geometry and selection effects of SDSS-III. We performed a series of statistical efficiency and consistency tests to validate the methodology adopted and showcased its potential to yield tight constraints on cosmological parameters from current and future galaxy redshift surveys. The main strength of our implementation of the AP test lies in its robustness to a misspecified model and its inherent approximations, thereby near-optimally exploiting the model predictions, without relying on its accuracy in modelling the scale dependence of the correlations of the field.

Moreover, another key aspect of our approach, resulting from the robustness to a misspecified model, is that the

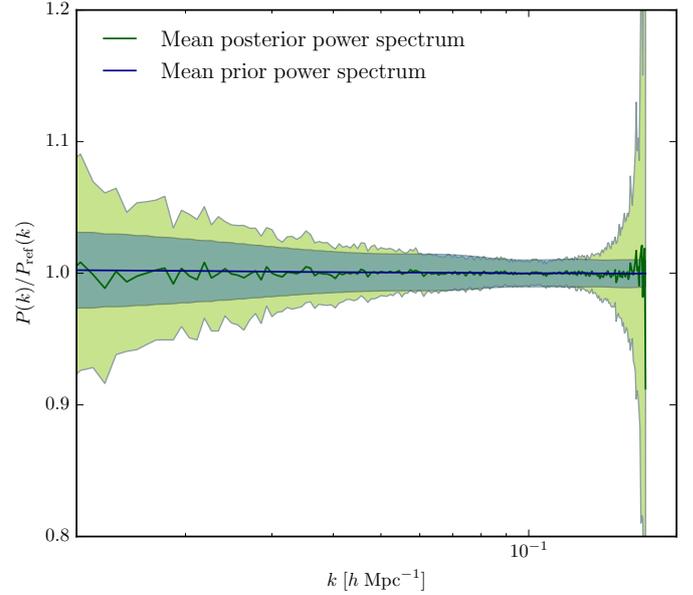


Fig. 9. Summary statistics of the reconstructed power spectra from the inferred posterior initial density field realisations depicted in Fig. 8, with the ensemble mean indicated by a solid green line. The solid blue line corresponds to the ensemble mean of the power spectra realisations generated using the inferred cosmological parameters. The shaded regions indicate their respective 1σ confidence region, i.e. 68% probability volume. This plot shows that the prior information entropy is inferior to the posterior information entropy, due to the narrower distribution of the former. The prior power spectrum adopted, as a result, does not impact significantly on the cosmological parameter inference via the AP test.

cosmological constraints show extremely weak dependence on galaxy bias. This yields two crucial advantages. First, this is especially interesting as the lack of a sufficient physical description of this bias remains a potential limiting factor for standard approaches of cosmological parameter inference from such redshift surveys. Furthermore, this lack of sensitivity to the bias also implies that our method does not depend on the absolute density fluctuation amplitudes. This is therefore among the first methods to extract a large amount of information from statistics other than that of direct density contrast correlations, without relying on the power spectrum or bispectrum, thereby providing complementary information to state-of-the-art techniques.

There is scope for further development of the *ALTAIR* framework, such as incorporating power spectrum inference, which is a highly non-trivial undertaking. We also intend to augment the current formalism to include the treatment of redshift space distortions, which is key for unbiased constraints on the cosmological parameters, and apply *ALTAIR* on state-of-the-art galaxy redshift catalogues for cosmological inference¹.

Acknowledgements. We express our appreciation to the anonymous reviewer for their constructive comments that helped us to improve the manuscript. We are grateful to Franz Elsner, Florent Leclercq and Natalia Porqueres for interesting discussions and/or for their comments on a draft version of the paper. We thank Tom Charnock for his help with coining the name *ALTAIR* in a coherent context. We also thank Jeremy Tinker for his assistance in interpreting the SDSS-III (DR12) science products. We acknowledge financial support from the ILP LABEX (under reference ANR-10-LABX-63) which is financed by French state funds managed by the ANR within the Investissements d’Avenir programme under reference ANR-11-IDEX-0004-02. This work was supported by the ANR BIG4 project, grant ANR-16-CE23-0002 of the French Agence Nationale de la

¹ <https://aquila-consortium.org>

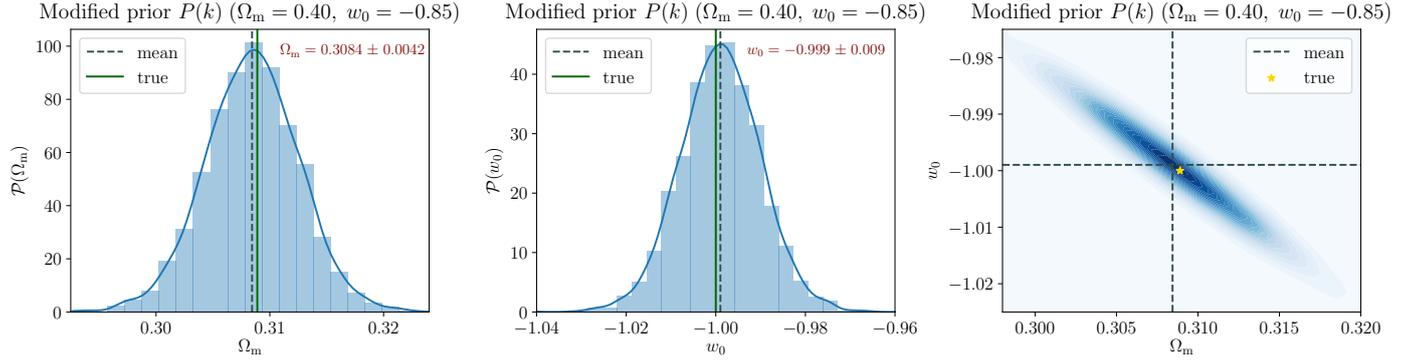


Fig. 10. Same as Fig. 6, but employing a different prior power spectrum ($\Omega_m = 0.40$, $w_0 = -0.85$), for ~ 3000 MCMC realisations. By recovering the fiducial cosmological parameters employed in the mock generation, this test case explicitly highlights the robustness of our approach to the shape of the prior power spectrum adopted. The corresponding uncertainties are slightly larger than for the original run by around 15%.

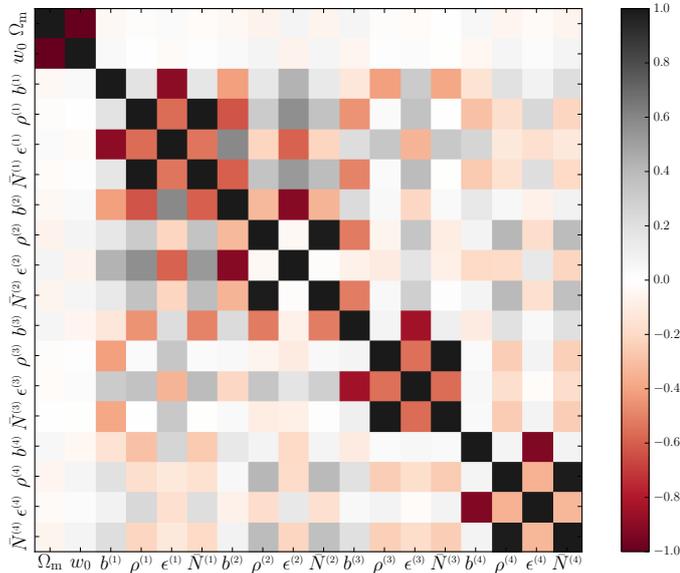


Fig. 11. Correlation matrix of the galaxy bias and cosmological parameters, normalised using their respective variance. This illustrates the weak correlation between the inferred cosmological constraints and the galaxy bias. The lack of dependence on the currently unknown phenomenon of galaxy biasing is therefore a key highlight of our implementation of the AP test for cosmological parameter inference.

Recherche. BDW is supported by the Simons Foundation. This work has made use of the *Horizon/Beyond* cluster at the Institut d’Astrophysique de Paris. This work is supported, through the maintenance of *Horizon/Beyond* cluster, under the name ORIGIN by the Domaine d’Intérêt Majeur (DIM) Astrophysique et Conditions d’Apparition de la Vie (ACAV), and received financial support from Région Île-de-France. We thank Stéphane Rouberol for running smoothly this cluster for us. This work is done within the Aquila Consortium.

References

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, *ApJS*, **182**, 543
Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2014, *ApJS*, **211**, 17
Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *ApJS*, **219**, 12
Alam, S., Ata, M., Bailey, S., et al. 2017, *MNRAS*, **470**, 2617
Alcock, C., & Paczynski, B. 1979, *Nature*, **281**, 358
Amendola, L., Appleby, S., Avgoustidis, A., et al. 2018, *Liv. Rev. Rel.*, **21**, 2
Ballinger, W. E., Peacock, J. A., & Heavens, A. F. 1996, *MNRAS*, **282**, 877
Barkana, R. 2006, *MNRAS*, **372**, 259
Bertschinger, E., & Dekel, A. 1989, *ApJ*, **336**, L5
Bertschinger, E., & Dekel, A. 1991, in *Large-scale Structures and Peculiar Motions in the Universe*, eds. D. W. Latham, & L. A. N. da Costa, *ASP Conf. Ser.*, **15**, 67

Bistolas, V., & Hoffman, Y. 1998, *ApJ*, **492**, 439
Blake, C., & Glazebrook, K. 2003, *ApJ*, **594**, 665
Blake, C., Glazebrook, K., Davis, T. M., et al. 2011, *MNRAS*, **418**, 1725
Bouchet, F. R., Colombi, S., Hivon, E., & Juszkiewicz, R. 1995, *A&A*, **296**, 575
Buchert, T., Melott, A. L., & Weiss, A. G. 1994, *A&A*, **288**, 349
Chevallier, M., & Polarski, D. 2001, *Int. J. Mod. Phys. D*, **10**, 213
Coles, P., & Jones, B. 1991, *MNRAS*, **248**, 1
de Laix, A. A., & Starkman, G. 1998, *ApJ*, **501**, 427
Desjacques, V., Jeong, D., & Schmidt, F. 2018, *Phys. Rep.*, **733**, 1
Duane, S., Kennedy, A. D., Pendleton, B. J., & Roweth, D. 1987, *Phys. Lett. B*, **195**, 216
Eisenstein, D. J., & Hu, W. 1998, *ApJ*, **496**, 605
Eisenstein, D. J., & Hu, W. 1999, *ApJ*, **511**, 5
Erdođdu, P., Lahav, O., Zaroubi, S., et al. 2004, *MNRAS*, **352**, 939
Erdođdu, P., Lahav, O., Huchra, J. P., et al. 2006, *MNRAS*, **373**, 45
Fisher, K. B., Lahav, O., Hoffman, Y., Lynden-Bell, D., & Zaroubi, S. 1995, *MNRAS*, **272**, 885
Friedmann, A. 1922, *Z. Angew. Phys.*, **10**, 377
Friedmann, A. 1924, *Z. Angew. Phys.*, **21**, 326
Frigo, M., & Johnson, S. G. 2005, *Proc. IEEE*, **93**, 216
Galassi, M., Davies, J., Theiler, J., et al. 2009, *GNU Scientific Library reference manual Network Theory* (UK: Bristol)
Gaztanaga, E., & Yokoyama, J. 1993, *ApJ*, **403**, 450
Glazebrook, K., & Blake, C. 2005, *ApJ*, **631**, 1
Hamaus, N., Sutter, P. M., Lavaux, G., & Wandelt, B. D. 2014, *J. Cosmol. Astropart. Phys.*, **12**, 013
Hamaus, N., Sutter, P. M., Lavaux, G., & Wandelt, B. D. 2015, *J. Cosmol. Astropart. Phys.*, **11**, 036
Hamaus, N., Pisani, A., Sutter, P. M., et al. 2016, *Phys. Rev. Lett.*, **117**, 091302
Hastings, W. K. 1970, *Biometrika*, **57**, 97
Hockney, R. W., & Eastwood, J. W. 1988, *Computer simulation using particles* (New York, London: Taylor & Francis)
Hoffman, Y. 1994, in *Unveiling Large-Scale Structures Behind the Milky Way*, eds. C. Balkowski, & R. C. Kraan-Korteweg, *ASP Conf. Ser.*, **67**, 185
Hubble, E. 1934, *ApJ*, **79**, 8
Hui, L., Stebbins, A., & Burles, S. 1999, *ApJ*, **511**, L5
Ivezic, Z., Tyson, J. A., & Abel, B. 2008, ArXiv e-prints [arXiv:0805.2366]
Jasche, J., & Kitaura, F. S. 2010, *MNRAS*, **407**, 29
Jasche, J., & Lavaux, G. 2015, *MNRAS*, **447**, 1204
Jasche, J., & Lavaux, G. 2018, ArXiv e-prints [arXiv:1806.11117]
Jasche, J., & Wandelt, B. D. 2012, *MNRAS*, **425**, 1042
Jasche, J., & Wandelt, B. D. 2013a, *MNRAS*, **432**, 894
Jasche, J., & Wandelt, B. D. 2013b, *ApJ*, **779**, 15
Jasche, J., Kitaura, F. S., Wandelt, B. D., & Enßlin, T. A. 2010a, *MNRAS*, **406**, 60
Jasche, J., Kitaura, F. S., Li, C., & Enßlin, T. A. 2010b, *MNRAS*, **409**, 355
Jasche, J., Leclercq, F., & Wandelt, B. D. 2015, *J. Cosmol. Astropart. Phys.*, **1**, 036
Jones, D. H., Read, M. A., Saunders, W., et al. 2009, *MNRAS*, **399**, 683
Kaiser, N. 1984, *ApJ*, **284**, L9
Kayo, I., Taruya, A., & Suto, Y. 2001, *ApJ*, **561**, 22
Kitaura, F.-S. 2012, *MNRAS*, **420**, 2737
Kitaura, F. S., & Enßlin, T. A. 2008, *MNRAS*, **389**, 497
Kitaura, F. S., Jasche, J., Li, C., et al. 2009, *MNRAS*, **400**, 183
Kitaura, F.-S., Jasche, J., & Metcalf, R. B. 2010, *MNRAS*, **403**, 589

- Kitaura, F.-S., Gallerani, S., & Ferrara, A. 2012, *MNRAS*, **420**, 61
- Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, *ApJS*, **192**, 18
- Lahav, O., Fisher, K. B., Hoffman, Y., Scharf, C. A., & Zaroubi, S. 1994, *ApJ*, **423**, L93
- Laureijs, R., Amiaux, J., & Arduini, S. 2011, ArXiv e-prints [arXiv:1110.3193]
- Lavaux, G., & Jasche, J. 2016, *MNRAS*, **455**, 3169
- Lavaux, G., & Wandelt, B. D. 2012, *ApJ*, **754**, 109
- Layzer, D. 1956, *AJ*, **61**, 383
- Lekien, F., & Marsden, J. 2005, *J. Numer. Methods Eng.*, **63**, 455
- Lemaître, G. 1927, *Ann Soc. Sci. Bruxelles*, **47**, 49
- Lemaître, G. 1931, *MNRAS*, **91**, 483
- Lemaître, G. 1933, *Ann Soc. Sci. Bruxelles*, **53**
- Li, X.-D., Park, C., Forero-Romero, J. E., & Kim, J. 2014, *ApJ*, **796**, 137
- Li, X.-D., Park, C., Sabiu, C. G., et al. 2016, *ApJ*, **832**, 103
- Linde, A. 2008, *Inflationary cosmology* (Berlin, Heidelberg: Springer), 1
- Linder, E. V. 2003, *Phys. Rev. Lett.*, **90**, 091301
- López-Corredoira, M. 2014, *ApJ*, **781**, 96
- Matsubara, T., & Suto, Y. 1996, *ApJ*, **470**, L1
- Matsumoto, M., & Nishimura, T. 1998, *ACM Trans. Modeling Comput. Simulation*, **8**, 3
- Moutarde, F., Alimi, J.-M., Bouchet, F. R., Pellat, R., & Ramani, A. 1991, *ApJ*, **382**, 377
- Mukherjee, S., & Wandelt, B. D. 2018, *Phys. Rev. Lett.* submitted [arXiv:1808.06615]
- Neal, R. 2000, *Ann. Stat.*, **31**, 705
- Neal, R. M. 1993, *Technical Report CRG-TR-93-1*
- Neal, R. M. 2003, *Ann. Stat.*, **705**
- Neyrinck, M. C., Aragón-Calvo, M. A., Jeong, D., & Wang, X. 2014, *MNRAS*, **441**, 646
- Nusser, A. 2005, *MNRAS*, **364**, 743
- Nusser, A., & Dekel, A. 1992, *ApJ*, **391**, 443
- Padmanabhan, N., & White, M. 2008, *Phys. Rev. D*, **77**, 123540
- Peebles, P. J. E. 1980, *The large-scale structure of the universe* (Princeton: Princeton University Press)
- Phillipps, S. 1994, *MNRAS*, **269**, 1077
- Planck Collaboration XIII. 2016, *A&A*, **594**, A13
- Planck Collaboration XVII. 2016, *A&A*, **594**, A17
- Pollina, G., Hamaus, N., & Paech, K. 2018, *MNRAS*, submitted [arXiv:1806.06860]
- Popowski, P. A., Weinberg, D. H., Ryden, B. S., & Osmer, P. S. 1998, *ApJ*, **498**, 11
- Racca, G. D., Laureijs, R., & Stagnaro, L. 2016, in *SPIE Conf. Ser., Proc. SPIE*, **9904**, 990400
- Robertson, H. P. 1935, *ApJ*, **82**, 284
- Robertson, H. P. 1936a, *ApJ*, **83**, 187
- Robertson, H. P. 1936b, *ApJ*, **83**, 257
- Ross, A. J., Beutler, F., Chuang, C.-H., et al. 2017, *MNRAS*, **464**, 1168
- Ryden, B. S. 1995, *ApJ*, **452**, 25
- Saadeh, D., Feeney, S. M., Pontzen, A., Peiris, H. V., & McEwen, J. D. 2016, *Phys. Rev. Lett.*, **117**, 131302
- Saunders, W., & Ballinger, W. E. 2000, in *Mapping the Hidden Universe: The Universe behind the Milky Way - The Universe in HI*, eds. R. C. Kraan-Korteweg, P. A. Henning, & H. Andernach, *ASP Conf. Ser.*, **218**, 181
- Schmidt, F., Elsner, F., Jasche, J., Nguyen, N. M., & Lavaux, G. 2018, ArXiv e-prints [arXiv:1808.02002]
- Schmoldt, I. M., Saar, V., Saha, P., et al. 1999, *AJ*, **118**, 1146
- Scoccimarro, R. 2000, *ApJ*, **544**, 597
- Scoccimarro, R., & Sheth, R. K. 2002, *MNRAS*, **329**, 629
- Seo, H.-J., & Eisenstein, D. J. 2003, *ApJ*, **598**, 720
- Sheth, R. K. 1995, *MNRAS*, **277**, 933
- Shoji, M., Jeong, D., & Komatsu, E. 2009, *ApJ*, **693**, 1404
- Sutter, P. M., Lavaux, G., Wandelt, B. D., & Weinberg, D. H. 2012, *ApJ*, **761**, 187
- Sutter, P. M., Pisani, A., Wandelt, B. D., & Weinberg, D. H. 2014, *MNRAS*, **443**, 2983
- Taylor, J. F., Ashdown, M. A. J., & Hobson, M. P. 2008, *MNRAS*, **389**, 1284
- Tegmark, M., Blanton, M. R., Strauss, M. A., et al. 2004, *ApJ*, **606**, 702
- Walker, A. G. 1937, *Proc. London Math. Soc.*, **42**, 90
- Wang, L., & Steinhardt, P. J. 1998, *ApJ*, **508**, 483
- Webster, M., Lahav, O., & Fisher, K. 1997, *MNRAS*, **287**, 425
- York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *AJ*, **120**, 1579
- Zaroubi, S. 2002, *MNRAS*, **331**, 901
- Zaroubi, S., Hoffman, Y., & Dekel, A. 1999, *ApJ*, **520**, 413

Appendix A: The LPT-Poissonian posterior

In this section, we describe the large-scale structure (LSS) posterior distribution implemented in this work. We demonstrate how the complex problem of exploring the high dimensional joint posterior distribution can be reduced to a set of distinct steps via a multiple block sampling scheme.

A.1. The density posterior distribution

The primary objective is to fully characterise the 3D cosmic LSS in observations via a numerical representation of the corresponding LSS posterior using sophisticated Markov chain Monte Carlo (MCMC) techniques, in particular to provide data constrained realisations of a set of plausible 3D density contrast amplitudes underlying a given set of galaxy observations. The posterior distribution for the evolved density field fluctuation δ_p^f can be formulated, in a general context, via Bayes' identity as:

$$\mathcal{P}(\{\delta_p^f\}|\{N_p^g\}) = \frac{\mathcal{L}(\{N_p^g\}|\{\delta_p^f\})\Pi(\{\delta_p^f\})}{\Pi(\{N_p^g\})}, \quad (\text{A.1})$$

where N_p^g is the observed number of galaxies in voxel p , at position \mathbf{x}_p in the sky, in redshift space, for the g th galaxy sample, with the prior $\Pi(\{\delta_p^f\})$ incorporating our a priori knowledge of the present-day matter fluctuations in the Universe, the likelihood $\mathcal{L}(\{N_p^g\}|\{\delta_p^f\})$ and the normalizing factor given by the evidence $\Pi(\{N_p^g\})$.

A major stumbling block consequently arises, as discussed extensively in [Jasche & Wandelt \(2013a\)](#), since the inference framework requires a suitable prior $\Pi(\{\delta_p^f\})$ which adequately describes the physical behaviour of the gravitationally evolved density field. Nevertheless, as elaborated in the following section, most attempts made in this direction so far have been based on heuristic approximations and the absence of a closed form description of the present day matter fluctuations encoded in a multi-variate probability density distribution still persists.

However, [Jasche & Wandelt \(2013a\)](#) proposed an elegant approach to circumvent this key impediment based on the following assertions: There is substantial evidence that primordial seed fluctuations at redshifts $z \sim 1000$ can be modelled as a Gaussian random field to great accuracy (e.g. [Linde 2008](#); [Komatsu et al. 2011](#); [Planck Collaboration XVII 2016](#)), consistent with inflationary theories and CMB observations. Moreover, the evolution of the initial conditions relies solely on deterministic gravitational structure formation processes. Therefore, a conceptually reasonable alternative to modelling the complex statistical behaviour of the non-linear matter distribution is to formulate the inference problem at the level of the initial conditions adequately described by Gaussian statistics. This constitutes the conceptual foundation of the BORG framework ([Jasche & Wandelt 2013a](#)).

Given a forward model \mathcal{M}_p that connects the initial conditions $\delta_p^{\text{ic},(r)}$, in comoving (r) space, to the redshift (z) space representation of the final density field $\delta_p^{\text{f},(z)}$, we can therefore express the conditional posterior for the evolved density field as

$$\mathcal{P}(\{\delta_p^f\}|\{\delta_p^{\text{ic}}\}) = \prod_p \delta^{\text{D}}[\delta_p^f - \mathcal{M}_p(\{\delta_p^{\text{ic}}\})], \quad (\text{A.2})$$

where $\delta^{\text{D}}(x)$ denoting the Dirac delta distribution encapsulates the assumption that the structure formation process is deterministic. Within this generic framework, the forward model may

be generalised to a chain of arbitrary components linking the primordial density fluctuations to the present-day density contrast. Nevertheless, at its crux lies a cosmic structure formation model $\mathcal{G}_p(a, \{\delta_p^{\text{ic}}\})$ for the non-linear evolution of initial conditions into a final density field at a given scale factor a , that is, $\delta_p^{\text{f},(r)} = \mathcal{G}_p(a, \{\delta_p^{\text{ic},(r)}\})$. The forward model implemented in this work to encode the AP test is described in Sect. 3 above.

We can then obtain a prior distribution for δ_p^f via a two-step sampling procedure: First, a realisation of δ_p^{ic} is generated from the prior distribution $\Pi(\{\delta_p^{\text{ic}}\})$ and subsequently evolved with a given forward model $\mathcal{M}_p(\{\delta_p^{\text{ic}}\})$. This essentially implies generating samples from the joint prior distribution of δ_p^{ic} and δ_p^f :

$$\begin{aligned} \mathcal{P}(\{\delta_p^f, \{\delta_p^{\text{ic}}\}) &= \Pi(\{\delta_p^{\text{ic}}\})\mathcal{P}(\{\delta_p^f|\{\delta_p^{\text{ic}}\}) \\ &= \Pi(\{\delta_p^{\text{ic}}\}) \prod_p \delta^{\text{D}}[\delta_p^f - \mathcal{M}_p(\{\delta_p^{\text{ic}}\})], \end{aligned} \quad (\text{A.3})$$

after plugging in the conditional posterior distribution from Eq. (A.2).

Assuming a normal distribution with zero mean and covariance \mathbf{S} corresponding to an underlying cosmological power spectrum for the initial conditions δ_p^{ic} , the joint prior distribution can be expressed as

$$\begin{aligned} \mathcal{P}(\{\delta_p^f, \{\delta_p^{\text{ic}}\}|\mathbf{S}) &= \Pi(\{\delta_p^{\text{ic}}\}|\mathbf{S}) \prod_p \delta^{\text{D}}[\delta_p^f - \mathcal{M}_p(\{\delta_p^{\text{ic}}\})] \\ &= \frac{e^{-\frac{1}{2} \sum_{p,q} \delta_p^{\text{ic}} \mathbf{S}^{-1} \delta_q^{\text{ic}}}}{\det(2\pi\mathbf{S})} \prod_p \delta^{\text{D}}[\delta_p^f - \mathcal{M}_p(\{\delta_p^{\text{ic}}\})]. \end{aligned} \quad (\text{A.4})$$

Reformulating the statistical inference problem in terms of the initial conditions δ_p^{ic} results in the following joint posterior distribution of δ_p^{ic} and δ_p^f , conditional on the observed galaxy number counts N_p^g ,

$$\mathcal{P}(\{\delta_p^f, \{\delta_p^{\text{ic}}\}|\{N_p^g\}, \mathbf{S}) = \frac{\mathcal{L}(\{N_p^g\}|\{\delta_p^f, \{\delta_p^{\text{ic}}\})\mathcal{P}(\{\delta_p^f, \{\delta_p^{\text{ic}}\}|\mathbf{S})}{\Pi(\{N_p^g\}|\mathbf{S})}, \quad (\text{A.5})$$

after making the dependence on the underlying power spectrum explicit. Assuming that the galaxy observations are conditionally independent of the initial conditions once the final density field is known, that is, $\mathcal{L}(\{N_p^g\}|\{\delta_p^f, \{\delta_p^{\text{ic}}\}) = \mathcal{L}(\{N_p^g\}|\{\delta_p^f\})$ and using the joint prior distribution from Eq. (A.4) leads to the LSS posterior distribution:

$$\begin{aligned} \mathcal{P}(\{\delta_p^f, \{\delta_p^{\text{ic}}\}|\{N_p^g\}, \mathbf{S}) &= \mathcal{L}(\{N_p^g\}|\{\delta_p^f\}) \frac{\Pi(\{\delta_p^{\text{ic}}\}|\mathbf{S})}{\Pi(\{N_p^g\}|\mathbf{S})} \\ &\quad \times \prod_p \delta^{\text{D}}[\delta_p^f - \mathcal{M}_p(\{\delta_p^{\text{ic}}\})]. \end{aligned} \quad (\text{A.6})$$

By marginalizing over the final density field δ_p^f , we finally obtain our posterior distribution as follows:

$$\mathcal{P}(\{\delta_p^{\text{ic}}\}|\{N_p^g\}, \mathbf{S}) = \mathcal{L}[\{N_p^g\}|\mathcal{M}_p(\{\delta_p^{\text{ic}}\})] \frac{\Pi(\{\delta_p^{\text{ic}}\}|\mathbf{S})}{\Pi(\{N_p^g\}|\mathbf{S})}, \quad (\text{A.7})$$

thereby connecting present galaxy observations N_p^g to their corresponding primordial density fields δ_p^{ic} via a given forward

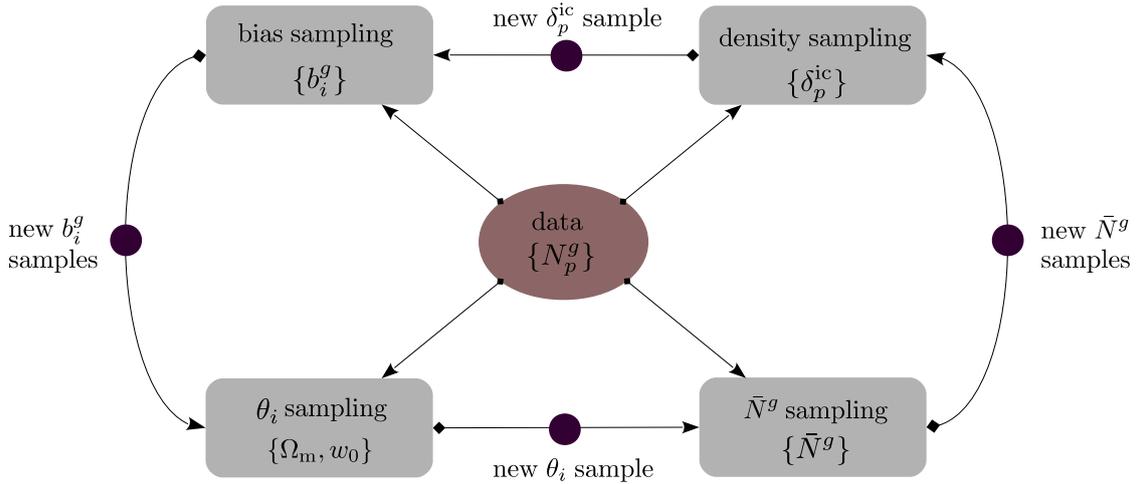


Fig. A.1. Schematic representation of the multi-step iterative block sampling procedure implemented in `ALTAIR`. Initially, a realisation of the 3D density contrast is obtained conditional on galaxy observations, followed by the other parameters being sampled conditional on the respective previous samples. Subsequent iterations of this procedure yield samples from the full joint posterior distribution within a modular statistical programming framework.

model, $\mathcal{M}_p(\{\delta_p^{ic}\})$. We must therefore sample from the highly non-linear and non-Gaussian posterior above to obtain realisations of the 3D initial density fields conditioned on the galaxy observations via the sophisticated HMC method described below in Appendix C.

A.2. Choice of density prior

Standard Wiener filtering approaches employ isotropic Gaussian priors on the present-day density contrast which is justified for inference on the largest scales (e.g. Zaroubi et al. 1999; Zaroubi 2002; Erdoğan et al. 2004, 2006; Kitaura & Enßlin 2008; Kitaura et al. 2009; Jasche et al. 2010a; Jasche & Lavaux 2015), where the density field can be reasonably approximated as a Gaussian random field. Although the exact probability distribution for the density field in the mildly and strongly non-linear regimes is not known, the deviation from Gaussianity is well-established and we therefore require a non-Gaussian prior to effectively capture the details of the highly complex filamentary nature of the present day cosmic web.

Lognormal density priors were subsequently proposed (Coles & Jones 1991) and this proved to be an adequate phenomenological choice (e.g. Hubble 1934; Peebles 1980; Gaztanaga & Yokoyama 1993; Kayo et al. 2001), albeit with some limitations, for modelling the evolved matter field in the mildly non-linear regime (Kitaura et al. 2009; Jasche & Kitaura 2010; Jasche et al. 2010b). The use of Edgeworth expansions to construct prior distributions involving third order moments has also been proposed in the literature (Kitaura 2012).

In this work, we rely on Lagrangian perturbation theory (LPT) to model cosmic structure formation, which accounts for non-local effects of gravitational mass transport from initial to final positions. It has been extensively shown that LPT provides a sufficient description of the cosmic LSS on large scales, where it is capable of reproducing the exact one-, two- and three-point statistics of the evolved field, while still being a reasonable approximation to the higher order statistics (e.g. Moutarde et al. 1991; Buchert et al. 1994; Bouchet et al. 1995; Scoccimarro 2000; Scoccimarro & Sheth 2002).

The essence of the above approach is that it provides a pathway to recover the high-order statistics of the matter

distribution using only the 2-point statistics of the initial conditions, obviating the need for additional parameters to describe higher order statistics for the matter inference problem. Moreover, since our model encodes the dynamics, the reconstruction of the large-scale velocity field is automatically generated without incurring the expense of an increased parameter space.

A.3. Modular statistical programming

In this work, we exploit the modular statistical programming framework inherent in the `BORG` algorithm to encode the AP test as an additional component to the original block sampling machinery. This approach allows us to model any Bayesian hierarchical problem to take into account any observational systematics via data models of higher complexity. Here, we account for the unknown parameters $\{b_i^g\}$ of the galaxy bias model described in Sect. 3.1 and unknown normalisations $\{\bar{N}^g\}$ for distinct galaxy samples, as illustrated in Fig. A.1. These last normalisations, in practice, scale as the noise amplitudes as we are using a per-voxel Poissonian likelihood.

Conceptually, within such a framework, the overall aim is to characterise fully the augmented joint posterior $\mathcal{P}(\{\delta_p^{ic}\}, \{\bar{N}^g\}, \{b_i^g\}, \{\theta_i\} | \{N_p^g\}, \mathbf{S})$ of different tracer populations labelled by the index g . Since it is not computationally feasible and advisable to sample directly from the high dimensional joint posterior distribution, we make use of an important theorem on Metropolis-Hastings block sampling which allows us to break the high dimensional sampling problem into a series of lower dimensional ones (Hastings 1970). We therefore dissect the exploration of the full joint parameter space into a sequence of conditional sampling procedures. The block sampling approach consists of drawing samples from the following conditional probability distributions:

$$\begin{aligned}
 (1) \quad & \{\delta_p^{ic}\}^{s+1} \sim \mathcal{P}(\{\delta_p^{ic}\} | \{\theta_i\}^s, \{N_p^g\}^s, \{\bar{N}^g\}^s, \{b_i^g\}^s, \mathbf{S}) \\
 (2) \quad & \{\bar{N}^g\}^{s+1} \sim \mathcal{P}(\{\bar{N}^g\} | \{\theta_i\}^s, \{N_p^g\}^s, \{\delta_p^{ic}\}^s, \{b_i^g\}^s, \mathbf{S}) \\
 (3) \quad & \{b_i^g\}^{s+1} \sim \mathcal{P}(\{b_i^g\} | \{\theta_i\}^s, \{\bar{N}^g\}^s, \{N_p^g\}^s, \{\delta_p^{ic}\}^s, \mathbf{S}) \\
 (4) \quad & \{\theta_i\}^{s+1} \sim \mathcal{P}(\{\theta_i\} | \{\bar{N}^g\}^s, \{N_p^g\}^s, \{\delta_p^{ic}\}^s, \{b_i^g\}^s, \mathbf{S}). \tag{A.8}
 \end{aligned}$$

In the above expressions, s denotes the sampling step and the symbol \smile indicates sampling from the distribution on the right hand side. A series of iterations of the individual sampling steps above will converge to the target distribution which corresponds the full joint posterior distribution (Hastings 1970). Hence, by simultaneously exploring the 3D initial density field $\{\delta_p^{\text{ic}}\}$, the galaxy bias parameters $\{b_i^g\}$, the galaxy density contrast normalisations $\{\bar{N}^g\}$ and the cosmological parameters $\{\theta_i\}$ via an implementation of high dimensional MCMC methods in a multiple block sampling procedure, we can obtain samples from the desired joint posterior distribution.

A.4. The bias posterior distribution

The formalism for the data model is presented in a generic context in Sect. 3.1, such that it can be implemented for two or more different galaxy surveys, distinct subsamples of a given catalogue or even different cosmological probes of the LSS. The advantage of splitting a galaxy sample into various subsamples is that we can treat the respective systematic and statistical uncertainties of each subsample separately, thereby allowing us to account for the distinct clustering behaviour of galaxy populations in the LSS sample via their respective bias parameters.

These additional parameters can be trivially incorporated in the flexible block sampling approach adopted here, as described in Sect. A.3. The mean numbers of galaxies, \bar{N}^g , for the various subsamples are essential for defining the density contrast of the galaxy distribution, with an erroneous value of \bar{N}^g resulting in a non-zero value of the mean, yielding spurious large-scale power in the inferred density fields (Jasche & Wandelt 2013b). Due to a lack of a priori knowledge, we must perform a joint inference of the set of four \bar{N}^g and b_i^g bias parameters, together with initial and final density fields, to take into account possible cross-correlations and interdependencies. Unlike traditional approaches (e.g. Tegmark et al. 2004), here we infer the bias factors directly from the relation between the data and the density field.

The conditional posterior distribution for the bias parameters, given a realisation of the density field and galaxy number counts for the respective subcatalogues, can be expressed as:

$$\mathcal{P}\left(\{\bar{N}^g\}, \{b_i^g\} \mid \{\theta_i\}, \{N_p^g\}, \{\delta_p^{\text{ic}}\}\right) \propto \mathcal{L}\left[\{N_p^g\} \mid \mathcal{M}_p\left(\{\delta_p^{\text{ic}}\}\right), \{\theta_i\}, \{\bar{N}^g\}, \{b_i^g\}\right] \times \Pi\left(\{\bar{N}^g\}, \{b_i^g\}\right), \quad (\text{A.9})$$

following an analogous reasoning to that described in Sect. 3.3. Adopting a standard maximally agnostic philosophy, we set uniform prior distributions for the bias parameters:

$$\Pi\left(\{\bar{N}^g\}, \{b_i^g\}\right) = \Theta\left(\{\bar{N}^g\}\right) \Theta\left(\{b_i^g\}\right), \quad (\text{A.10})$$

where the Heaviside function $\Theta(x)$ ensures that the parameters are positive, as required by the bias model. The non-linear shape of the galaxy biasing function, as given in Eq. (4), poses a particular challenge as no straightforward sampling procedure can be derived. We therefore sample from the above conditional bias posterior distribution via sequential slice sampling steps (e.g. Neal 2000, 2003) to ensure unit acceptance rates.

Appendix B: Jacobian of comoving-redshift transformation

The derivation of the Jacobian, \mathcal{J}_r^z , of the coordinate transformation between comoving (\mathbf{r}) and redshift (z) transformation is described below.

Since we desire the redshift space representation of the density field,

$$\rho_z[\mathbf{z}(\mathbf{r})] = \rho_r(\mathbf{r}) \left| \frac{\partial \mathbf{z}}{\partial \mathbf{r}} \right|^{-1} = \rho_r(\mathbf{r}) |\mathcal{J}_r^z|^{-1} = \rho_r(\mathbf{r}) |\mathcal{J}_z^r|, \quad (\text{B.1})$$

or in terms of the density contrast,

$$\delta_z[\mathbf{z}(\mathbf{r})] = [1 + \delta_r(\mathbf{r})] |\mathcal{J}_r^z|^{-1} - 1, \quad (\text{B.2})$$

where \mathcal{J}_z^r is the corresponding Jacobian for the converse (redshift to comoving) transformation.

The set of three indices $\{i, j, k\}$ below are spatial coordinates $\{1, 2, 3\}$. With the functional dependence of $\mathbf{z} = f(\mathbf{r})$, we have

$$z_i = \frac{r_i}{r} f(\mathbf{r}) = \bar{z}(r) \frac{r_i}{r}, \quad (\text{B.3})$$

where we defined $r = \sqrt{\sum_j r_j^2}$, and \bar{z} is the cosmological redshift. Taking the derivative yields

$$\begin{aligned} dz_i &= d\bar{z}(r) \frac{r_i}{r} + \bar{z}(r) d\left(\frac{r_i}{r}\right) \\ &= \sum_j \frac{\partial \bar{z}}{\partial r} \frac{\partial r}{\partial r_j} dr_j \frac{r_i}{r} + \bar{z} \frac{dr_i}{r} - \bar{z} r_i d\left(\frac{1}{r}\right), \end{aligned} \quad (\text{B.4})$$

after applying the chain rule. Differentiating with respect to the comoving coordinate, and using $d(1/r) = -\sum_j (1/r^3) r_j dr_j$,

$$\begin{aligned} \frac{\partial z_i}{\partial r_k} &= \frac{\partial \bar{z}}{\partial r} \frac{\partial r}{\partial r_k} \frac{r_i}{r} + \frac{\bar{z}}{r} \delta_{ik}^k - \frac{z_i r_k}{r^3} \\ &= \left(\frac{\partial \bar{z}}{\partial r} - \frac{\bar{z}}{r} \right) \frac{r_i r_k}{r^2} + \frac{\bar{z}}{r} \delta_{ik}^k, \end{aligned} \quad (\text{B.5})$$

where we made use of $r dr = \sum_k r_k dr_k$ which follows from the definition $r = \sqrt{\sum_k r_k^2}$. To compute $\partial \bar{z} / \partial r$, we start from the definition of comoving distance,

$$r = \frac{c}{H_0} \int_0^{\bar{z}} \frac{dz}{\mathcal{E}(z)}, \quad (\text{B.6})$$

with the conventional definition for $\mathcal{E}(z)$,

$$\mathcal{E}(z) = \sqrt{\Omega_{\text{de}}(1+z)^{3(1+w)} + \Omega_k(1+z)^2 + \Omega_m(1+z)^3 + \Omega_r(1+z)^4}, \quad (\text{B.7})$$

where c is the speed of light and H_0 is the Hubble constant today, such that

$$\frac{\partial \bar{z}}{\partial r} = \left(\frac{H_0}{c} \right) \mathcal{E}(\bar{z}). \quad (\text{B.8})$$

Finally, the Jacobian matrix can be expressed as follows:

$$\mathcal{J}_{z\{ik\}}^r = \frac{\partial z_i}{\partial r_k} = \left(\frac{H_0}{c} \mathcal{E}(\bar{z}) - \frac{\bar{z}}{r} \right) \frac{r_i r_k}{r^2} + \frac{\bar{z}}{r} \delta_{ik}^k, \quad (\text{B.9})$$

and we obtain the desired $|\mathcal{J}_{r\{ik\}}^z|$ by taking the reciprocal of $|\mathcal{J}_{z\{ik\}}^r|$.

Appendix C: Hamiltonian sampling

The exploration of the LPT-Poissonian posterior, described by Eq. (6), requires highly non-linear reconstruction methods and is therefore numerically intensive. From the Bayesian viewpoint considered in this work, a single estimate of the density field is not of particular interest, with the desired scientific output being a sampled representation of the multi-dimensional LSS posterior. Extracting any relevant statistical summary, such as mean, mode or variance, given this posterior representation, is then straightforward. Furthermore, the propagation of uncertainties to the final inferred quantities is smooth and coherent.

In the absence of a known procedure to directly sample from a LPT-Poissonian distribution, a Metropolis-Hastings sampling scheme with an accept-reject step must be implemented. But such a mechanism suffers from the well-known drawback of possible high rejection rate, where a low acceptance rate of proposed samples will render the method numerically inefficient. This is especially significant since the inference of 3D density fields typically involves extremely high number of inference parameters. For instance, in this work, there are over 2×10^6 free parameters, which correspond to primordial density fluctuation amplitudes δ_p^{ic} at respective grid nodes. Exploring this high dimensional parameter space via a random walk, as in conventional MCMC methods, consequently results in a high rejection rate. This is an understatement, as in practice, the acceptance rate is virtually zero, with one chance over 10^6 of going in the right direction, without taking into account the step size.

To limit the number of samples and alleviate the numerical scaling of the method, we implement a Hamiltonian Monte Carlo (HMC) scheme. This method guarantees an acceptance rate of unity in the absence of numerical errors. The numerical efficiency of the Hamiltonian sampling lies in deterministically proposing new samples to the Metropolis-Hastings algorithm, based on techniques to follow dynamical particle motion in potentials. The HMC algorithm therefore maintains a reasonable sampling efficiency in high dimensional spaces by suppressing the dominant random walk behaviour.

The HMC method has been successfully implemented for the inference of density fields in the non-linear regime, and has been found to be very efficient (e.g. Jasche & Kitaura 2010; Jasche et al. 2010b; Kitaura et al. 2012; Jasche & Wandelt 2013a,b). Moreover, this approach is less prone to spurious effects induced by numerical inaccuracies due to the final accept/reject step of the Metropolis-Hastings sampler ensuring correctness of the target density. Numerical inaccuracies would only be detrimental to the efficiency, without compromising the correctness.

We briefly review the underlying framework and rationale of the HMC method. An excellent in-depth review of this approach is provided in Duane et al. (1987) and Neal (1993), with its application to LSS inference illustrated in Jasche & Kitaura (2010), Jasche et al. (2010b), Jasche & Wandelt (2013a,b) and Jasche & Lavaux (2018).

The essence of the HMC method is as follows: If we want to generate random variates according to a desired probability distribution $\mathcal{P}(\{x_i\})$, with $\{x_i\}$ being a set of N elements x_i , then we may interpret the negative logarithm of the distribution as a potential $\psi(x)$,

$$\psi(x) = -\ln(\mathcal{P}(x)). \quad (\text{C.1})$$

We can subsequently formulate a corresponding Hamiltonian that describes the dynamics in the multi-dimensional phase space by adding a kinetic term to the above potential. To this end, we introduce a ‘‘momentum’’ variable p_i and a ‘‘mass matrix’’ \mathbf{M}

as nuisance parameters, as follows:

$$H = \frac{1}{2} \sum_{i,j} p_i M_{ij}^{-1} p_j + \psi(x). \quad (\text{C.2})$$

This Hamiltonian, as in classical mechanics, describes the dynamics in a high dimensional parameter space. We then obtain the new target probability distribution by exponentiating the above Hamiltonian:

$$e^{-H} = \mathcal{P}(\{x_i\}) \exp\left(\frac{1}{2} \sum_{i,j} p_i M_{ij}^{-1} p_j\right). \quad (\text{C.3})$$

The form of the Hamiltonian ensures that the new joint distribution can be separated into a Gaussian distribution for the momenta $\{p_i\}$ and the target distribution $\mathcal{P}(\{x_i\})$. This implies that the two sets of variables $\{x_i\}$ and $\{p_i\}$ are independent, and hence, we can obtain samples from the target distribution $\mathcal{P}(\{x_i\})$ simply by marginalizing over the auxiliary momenta.

The next step is to generate a random variate from the joint distribution, which is proportional to $\exp(-H)$. We must first draw a set of momenta from the distribution defined by the kinetic energy term, which is an N -dimensional Gaussian with covariance \mathbf{M} . To obtain a new sample from the joint distribution, we allow the system to evolve deterministically, from a given initial point in the high dimensional parameter space $(\{x_i\}, \{p_i\})$, according to Hamilton’s equations:

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial p_i} \quad (\text{C.4})$$

$$\frac{dp_i}{dt} = \frac{\partial H}{\partial x_i} = -\frac{\partial \psi(x)}{\partial x_i}. \quad (\text{C.5})$$

Integrating the above equations of motion for the time variable t up to a fixed pseudo-time τ results in the new position $(\{x'_i\}, \{p'_i\})$ in phase space. This new point in phase space is accepted according to the standard Metropolis-Hastings acceptance rule:

$$\mathcal{P}_{\mathcal{A}} = \min\{1, \exp[-(H(\{x'_i\}, \{p'_i\}) - H(\{x_i\}, \{p_i\}))]\}. \quad (\text{C.6})$$

Since the equations of motion provide a solution to a Hamiltonian system, the energy or the Hamiltonian described by Eq. (C.2) must be conserved. A direct consequence of this conservation of the Hamiltonian is that the HMC approach yields an acceptance rate of unity. In practice, however, the acceptance rate may be lower due to numerical inaccuracies in the integration scheme. After accepting a new sample, the mechanism proceeds by discarding the momentum counterpart and restarts the sampling procedure by randomly drawing a new set of momenta. The individual momenta $\{p_i\}$ do not have to be stored in memory which is numerically convenient, and marginalisation simply implies discarding them.

The HMC sampling scheme can therefore be summarised in two steps. The first step is essentially a Gibbs sampling step which results in a new set of Gaussian distributed momenta. The second step involves computing the deterministic dynamical trajectory on the posterior surface. We obtain samples of the desired target distribution by marginalizing over the nuisance parameters, that is, by discarding the auxiliary momenta $\{p_i\}$.

Appendix D: Equations of motion for LSS inference

As described in Appendix C, the Hamiltonian Monte Carlo (HMC) approach allows the exploration of the non-linear LSS

posterior via the study of the Hamiltonian dynamics in the high dimensional parameter space. To ensure a high acceptance rate, the HMC exploits the gradient of the logarithmic posterior distribution to optimally explore this parameter space, such that the algorithm also requires the derivatives of the posterior distribution, as outlined below.

From the LSS posterior illustrated in Eq. (6), we can derive the corresponding forces required to evaluate the Hamiltonian trajectories. The Hamiltonian potential $\psi(\delta^{\text{ic}})$ can be written as

$$\begin{aligned} \psi(\delta^{\text{ic}}) &= -\ln \left[\mathcal{P} \left(\{\delta_p^{\text{ic}}\}, \{\bar{N}^g\}, \{b^g\}, \{\theta_i\} \mid \{N_p^g\}, \mathbf{S} \right) \right] \\ &= \psi_{\text{likelihood}}(\delta^{\text{ic}}) + \psi_{\text{prior}}(\delta^{\text{ic}}), \end{aligned} \quad (\text{D.1})$$

where the potential $\psi_{\text{prior}}(\delta^{\text{ic}})$ corresponds to

$$\psi_{\text{prior}}(\delta^{\text{ic}}) = \frac{1}{2} \delta^{\text{ic},\dagger} \mathbf{S}^{-1} \delta^{\text{ic}} = \frac{1}{2} \sum_{ij} \delta_i^{\text{ic},*} S_{ij}^{-1} \delta_j^{\text{ic}}, \quad (\text{D.2})$$

while $\psi_{\text{likelihood}}(\delta^{\text{ic}})$ is given by

$$\begin{aligned} \psi_{\text{likelihood}}(\delta^{\text{ic}}) &= \sum_p \left\{ \lambda_p^g \left(\{\delta_p^f\}, \{\theta_i\}, \{\bar{N}^g\}, \{b_i^g\} \right) \right. \\ &\quad \left. - N_p^g \ln \left[\lambda_p^g \left(\{\delta_p^f\}, \{\theta_i\}, \{\bar{N}^g\}, \{b_i^g\} \right) \right] \right\}. \end{aligned} \quad (\text{D.3})$$

Using the definition of the Hamiltonian potential $\psi(\delta^{\text{ic}})$ from Eq. (D.1), we obtain the corresponding equations describing the Hamiltonian forces by differentiating with respect to δ^{ic} :

$$\frac{\partial \psi(\delta^{\text{ic}})}{\partial \delta_i^{\text{ic}}} = \frac{\partial \psi_{\text{prior}}(\delta^{\text{ic}})}{\partial \delta_i^{\text{ic}}} + \frac{\partial \psi_{\text{likelihood}}(\delta^{\text{ic}})}{\partial \delta_i^{\text{ic}}}. \quad (\text{D.4})$$

The prior term is trivially obtained as

$$\frac{\partial \psi_{\text{prior}}(\delta^{\text{ic}})}{\partial \delta_i^{\text{ic}}} = \sum_j \delta_j^{\text{ic},*} S_{ij}^{-1}. \quad (\text{D.5})$$

The corresponding likelihood term, however, cannot be derived in straightforward fashion. We compute this adjoint gradient as a sequential application of linear operations, as follows:

$$\frac{\partial \psi_{\text{likelihood}}(\delta^{\text{ic}})}{\partial \delta_s^{\text{ic}}} = \sum_q \frac{\partial \rho_q^{(r)}}{\partial \delta_s^{\text{ic}}} \sum_{\bar{p}} \frac{\partial \rho_{\bar{p}}^{(z)}}{\partial \rho_q^{(r)}} \sum_{p,g} \frac{\partial \lambda_p^g}{\partial \rho_{\bar{p}}^{(z)}} \frac{\partial \psi_{\text{likelihood}}}{\partial \lambda_p^g}, \quad (\text{D.6})$$

where we explicitly expressed the coordinate transformation from comoving (\mathbf{r}) to redshift (\mathbf{z}) space, and $\rho_p = 1 + \delta_p^f$ is the density field. Equation (D.6) constitutes a sequence of matrix vector applications, as follows:

$$\tilde{\mathcal{G}}_{qs} \equiv \frac{\partial \rho_q^{(r)}}{\partial \delta_s^{\text{ic}}} \quad (\text{D.7})$$

$$\mathcal{Q}_{\bar{p}q} \equiv \frac{\partial \rho_{\bar{p}}^{(z)}}{\partial \rho_q^{(r)}} = \sum_i \mathcal{E}_{i\bar{p}}^{-1} \sum_{\bar{p}} \mathcal{J}_{\bar{p}} x_{\bar{p}}^{\alpha(i)} y_{\bar{p}}^{\beta(i)} z_{\bar{p}}^{\gamma(i)} \quad (\text{D.8})$$

$$\mathcal{K}_{p\bar{p}} \equiv \sum_g \frac{\partial \lambda_p^g}{\partial \rho_{\bar{p}}^{(z)}} = \sum_g \delta_{p,\bar{p}}^k \frac{\lambda_p^g}{\rho_{\bar{p}}^{(z)}} \left[\rho^g \epsilon^g \rho_{\bar{p}}^{-\epsilon_g(r)} + \beta \right] \quad (\text{D.9})$$

$$v_p \equiv \frac{\partial \psi_{\text{likelihood}}}{\partial \lambda_p^g} = 1 - \frac{N_p^g}{\lambda_p^g \left(\{\delta_p^f\}, \{\theta_i\}, \{\bar{N}^g\}, \{b_i^g\} \right)}, \quad (\text{D.10})$$

where \mathcal{J} is the Jacobian of the comoving-redshift transformation, $\tilde{\mathcal{G}}$ indicates the derivative of LPT (cf. Appendix D in Jasche

& Wandelt 2013a) and $\delta_{p,\bar{p}}^k$ is the Kronecker delta. The derivation of the adjoint gradient for the triquintic interpolation is illustrated in Appendix G.

Using the above results, the two equations of motion (C.4) and (C.5) for the Hamiltonian system can be expressed as follows:

$$\frac{d\hat{\delta}_i^{\text{ic}}}{dt} = \sum_j M_{ij}^{-1} p_j, \quad (\text{D.11})$$

and

$$\frac{d\hat{p}_i}{dt} = - \sum_j \delta_j^{\text{ic},*} S_{ij}^{-1} - \sum_{q,\bar{p},p} \tilde{\mathcal{G}}_{qi} \mathcal{Q}_{\bar{p}q} \mathcal{K}_{p\bar{p}} v_p, \quad (\text{D.12})$$

where the hats denote Fourier space representation. We can now obtain new samples from the LSS posterior by following the dynamical evolution of the Hamiltonian system, governed by Eqs. (D.11) and (D.12), in phase space.

Appendix E: Numerical implementation

Our numerical implementation of the augmented version of BORG, that incorporates cosmological parameter inference, while assuming a more realistic non-linear bias model, is designated as ALTAIR. It employs the FFTW3 library for fast Fourier transforms (Frigo & Johnson 2005), whose feature of parallel transforms allows for straightforward parallelisation of our code. For random number generation, we make use of the GNU scientific library (gsl; Galassi et al. 2009), in particular, the Mersenne Twister MT19937, with 32-bit word length, from the `gsl_rng_mt19937` routine. The use of the Mersenne Twister algorithm as a pseudo-random number generator for Monte Carlo simulations has been proven to be efficient and dependable (Matsumoto & Nishimura 1998).

E.1. The leapfrog scheme

In order to obtain samples from the LSS posterior, we must integrate the equations of motion (D.11) and (D.12) numerically over a pseudo-time interval τ by means of a leapfrog scheme. Ideally, we want to solve the equations exactly for an optimal acceptance rate. As such, the choice of this integrator is motivated by several reasons. Primarily, the leapfrog scheme is convenient as it is a symplectic integrator and is therefore exactly reversible, thereby maintaining detailed balance (Duane et al. 1987) of the Markov chain. The high accuracy of such an integration scheme is conducive to the conservation of the total Hamiltonian along a given trajectory, within the limit of numerical errors, resulting in high acceptance rates. It is also numerically stable and allows for simple propagation of errors. To avoid resonant trajectories, τ is drawn from a uniform distribution. The numerical implementation of the ‘‘kick-drift-kick’’ leapfrog scheme here follows closely the descriptions provided in Jasche & Kitaura (2010) and Jasche & Wandelt (2013a), which we refer the reader to for more elaborate and complementary explanations.

E.2. Hamiltonian mass

The mass matrix \mathbf{M} contains a large number of parameters that must be tuned to optimise the performance of the HMC sampler. Conceptually, this Hamiltonian mass characterises the inertia of individual parameters as they move through the parameter

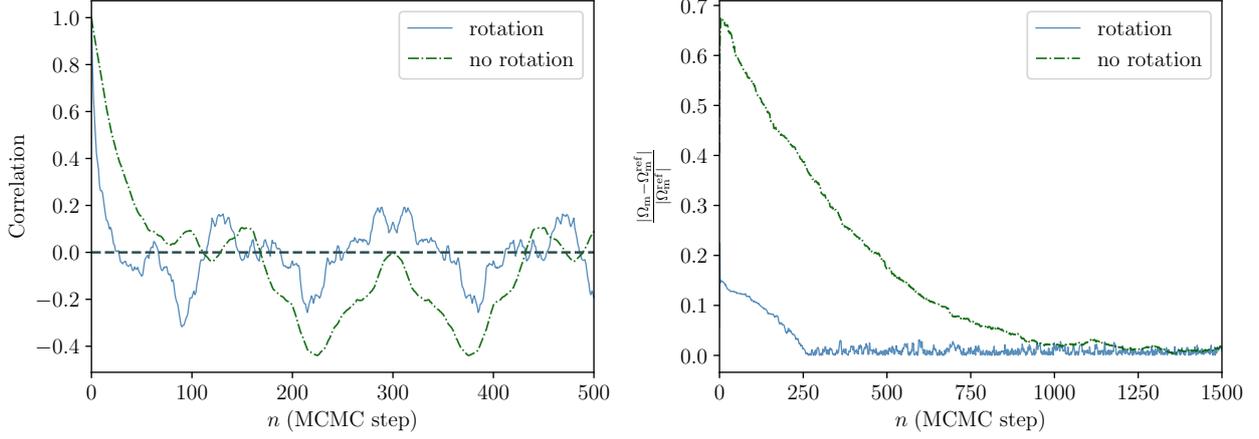


Fig. F.1. *Left panel:* correlation length of the Ω_m Markov chain. The chain has a correlation length of the order of 25 samples. The dashed-dotted lines depict the longer correlation length without the rotation of the cosmological parameter space, by nearly a factor of five. *Right panel:* burn-in phase for both scenarios, illustrated via the deviation from the reference cosmology. The dashed-dotted lines depict the MCMC chains obtained without rotating the parameter space, where the burn-in phase consequently requires around five times more iterations. The longer correlation length and burn-in phase highlight the loss in efficiency compared to the more sophisticated sampler.

space. As a result, the choice of \mathbf{M} is a compromise to limit slow exploration efficiency due to extremely large masses and avoid large numerical errors of the integration scheme due to extremely light masses, and is therefore a trade-off between efficiency and acceptance rate.

A general prescription for the Hamiltonian masses, when the density fluctuations δ_p^f can be characterised as a Gaussian random field, is to set them inversely proportional to the variance of that specific δ_p^f (Taylor et al. 2008). However, this prescription has been found to be effective even for the case of non-Gaussian distributions, such as the LPT-Poissonian posterior encountered in this work (Jasche & Wandelt 2013a; Jasche et al. 2015; Lavaux & Jasche 2016). Since the leapfrog scheme requires the inversion of \mathbf{M} , a diagonal representation of \mathbf{M} is numerically convenient, given the extremely high dimensionality of the problem. We therefore choose \mathbf{M} to be inversely proportional to the cosmological power spectrum, $\mathbf{M} = \mathbf{S}^{-1}$, which is diagonal in its Fourier basis. This choice, moreover, improves the efficiency of the HMC sampler and results in faster convergence since it accounts for the correlation structure of the underlying density field (Jasche & Wandelt 2013a).

Appendix F: Rotation of cosmological parameter space

To increase the efficiency of the cosmological parameter sampler, we rotate the (Ω_m, w_0) parameter space, using orthogonal basis transformations, before slice sampling. The corresponding covariance matrix can be orthonormally decomposed as $\mathbf{C} = \mathbf{Q}^\dagger \mathbf{\Lambda} \mathbf{Q}$, where $\mathbf{Q}^\dagger \mathbf{Q} = \mathbb{1}$.

We perform a rotation of the θ vector of cosmological parameters, about the mean $\bar{\theta}$, using the θ transformation:

$$\tilde{\theta} = \mathbf{\Lambda}^{-\frac{1}{2}} \mathbf{Q} (\theta - \bar{\theta}), \quad (\text{F.1})$$

where the tilde denotes the transformed variable. A key point is that we must rotate back to the original frame when computing the logarithm of the posterior via

$$\theta = \mathbf{Q}^\dagger \mathbf{\Lambda}^{\frac{1}{2}} \tilde{\theta} + \bar{\theta}. \quad (\text{F.2})$$

This is implemented at each step of the Markov chain, such that the covariance used to compute the orthogonal basis operators

must be learnt progressively as follows:

$$\mathbf{C}_{n+1} = \frac{n}{n+1} \mathbf{C}_n + \frac{(\theta_n - \bar{\theta})^\dagger (\theta_n - \bar{\theta})}{n}, \quad (\text{F.3})$$

at the given n th step. The above transformation helps to decorrelate the variables involved when sampling, thereby reducing the burn-in period significantly. This procedure is especially effective with highly correlated parameters, as is the case for (Ω_m, w_0) (cf. Fig. 6).

We investigate the statistical efficiency of our cosmological parameter sampler by computing the correlation length of the Markov chain. Subsequent samples in the chain are generally correlated and hence do not qualify as independent samples of the target posterior distribution. The correlation length characterises the statistical efficiency of generating independent samples of a given parameter θ , as follows:

$$C(\theta)_n = \frac{1}{N-n} \sum_{i=0}^{N-n} \left(\frac{\theta^i - \langle \theta \rangle}{\sqrt{\text{Var}(\theta)}} \frac{\theta^{i+n} - \langle \theta \rangle}{\sqrt{\text{Var}(\theta)}} \right), \quad (\text{F.4})$$

where n is the number of transition steps in the chain, $\langle \theta \rangle$ and $\text{Var}(\theta)$ correspond to the mean and variance, respectively, $\langle \theta \rangle = 1/N \sum_i \theta^i$ and $\text{Var}(\theta) = 1/N \sum_i (\theta^i - \langle \theta \rangle)^2$, while N is the total number of realisations in the chain. This correlation length therefore indicates the number of independent realisations that can be drawn from the MCMC chain with a given number of transition steps. For an unbiased comparison, we computed the correlation length using the same chain length for both scenarios, while ignoring their respective burn-in phases. From the left panel of Fig. F.1, we deduce that the correlation length for the Ω_m chain is of the order of 25 samples. Without the rotation of (Ω_m, w_0) parameter space, the correlation length is longer by nearly a factor of five, further demonstrating the gain in efficiency obtained with the more sophisticated sampler.

Appendix G: Adjoint gradient for generic 3D interpolation

The 3D interpolation can be expressed generally, for any arbitrary n th order of interpolation, as:

$$P_{\alpha\beta\gamma} = \sum_{i=0}^n \sum_{j=0}^n \sum_{k=0}^n a_{ijk} x_\alpha^i y_\beta^j z_\gamma^k = \sum_{i,j,k} a_{ijk} \mathcal{E}_{\alpha\beta\gamma}^{ijk}, \quad (\text{G.1})$$

where \mathbf{P} is the interpolated surface and \mathcal{E} is a $(n+1)^3 \times (n+1)^3$ matrix required for computing the vector of interpolation coefficients \mathbf{a} . In this work, we implement a triquintic interpolation scheme, which corresponds to $n = 5$, with the indices $\{i, j, k\}, \{\alpha, \beta, \gamma\} = \{0, 1, 2, 3, 4, 5\}$ constituting the particular indexing scheme employed to denote the voxels involved. \mathcal{E} is therefore a 216×216 matrix for the triquintic scheme.

The above system of equations can be reformulated as a matrix for the linear system described by $\mathbf{P} = \mathcal{E}\mathbf{a}$, such that the vector of interpolation coefficients can be computed in straightforward fashion through matrix inversion, $\mathbf{a} = \mathcal{E}^{-1}\mathbf{P}$. The advantage of this approach is that \mathcal{E}^{-1} can be computed only once and stored, and then used for interpolation at any location inside the cube (e.g. [Lekien & Marsden 2005](#)).

The derivative of the generic 3D interpolation is required as a component of the adjoint gradient at the core of the HMC sampler described in [Appendix C](#). The 3D interpolated density field, in redshift space (cf. [Eq. \(B.1\)](#)), can be written explicitly as:

$$\tilde{\rho}_k = \mathcal{J}_k \left(\sum_{i,m} \mathcal{E}_{im}^{-1} \rho_m x_k^{\alpha(i)} y_k^{\beta(i)} z_k^{\gamma(i)} \right), \quad (\text{G.2})$$

where \mathcal{J} is the Jacobian of the comoving-redshift transformation from [Appendix B](#), evaluated at the location of the mesh element k . Here, we employed a compact notation, where $\{x, y, z\}$

correspond to the fractional steps with respect to a reference node in the grid, that is, interpolation weights, and $\{\alpha, \beta, \gamma\}$ label the powers of the elements of \mathcal{E} via:

$$i = \alpha(n+1)^2 + \beta(n+1) + \gamma, \quad (\text{G.3})$$

with the index $m \in \{i, j, k\}$ from [Eq. \(G.1\)](#). Using the chain rule, the resulting gradient \mathbf{w} , after application of the incoming gradient \mathbf{v} from the previous components in the forward model, is obtained as follows:

$$\begin{aligned} w_a &= \sum_k v_k \frac{\partial \tilde{\rho}_k}{\partial \rho_a} \\ &= \sum_k \mathcal{J}_k v_k \sum_{i,m} \mathcal{E}_{im}^{-1} \delta_{ma}^k x_k^{\alpha(i)} y_k^{\beta(i)} z_k^{\gamma(i)} \\ &= \sum_i \mathcal{E}_{ia}^{-1} \sum_k \mathcal{J}_k v_k x_k^{\alpha(i)} y_k^{\beta(i)} z_k^{\gamma(i)} \\ &= \sum_{i,k} \mathcal{E}_{ia}^{-1} \xi_{ik}, \end{aligned} \quad (\text{G.4})$$

where $\xi_{ik} \equiv \mathcal{J}_k v_k x_k^{\alpha(i)} y_k^{\beta(i)} z_k^{\gamma(i)}$ represents the contribution to the adjoint gradient from a specific voxel labelled by the indexing scheme adopted. This is convenient as \mathcal{E}^{-1} is already available from the execution of the 3D interpolation in the forward model.