

Bayesian power spectrum inference with foreground and target contamination treatment

J. Jasche¹ and G. Lavaux^{2,3}

¹ Excellence Cluster Universe, Technische Universität München, Boltzmannstrasse 2, 85748 Garching, Germany
e-mail: j.jasche@tum.de

² CNRS & Sorbonne Universités, UPMC Univ. Paris 06, UMR 7095, Institut d'Astrophysique de Paris, 75014 Paris, France

³ Sorbonne Universités, Institut Lagrange de Paris (ILP), 98bis bd Arago, 75014 Paris, France

Received 31 March 2017 / Accepted 24 May 2017

ABSTRACT

This work presents a joint and self-consistent Bayesian treatment of various foreground and target contaminations when inferring cosmological power spectra and three-dimensional density fields from galaxy redshift surveys. This is achieved by introducing additional block-sampling procedures for unknown coefficients of foreground and target contamination templates to the previously presented ARES framework for Bayesian large-scale structure analyses. As a result, the method infers jointly and fully self-consistently three-dimensional density fields, cosmological power spectra, luminosity-dependent galaxy biases, noise levels of the respective galaxy distributions, and coefficients for a set of a priori specified foreground templates. In addition, this fully Bayesian approach permits detailed quantification of correlated uncertainties amongst all inferred quantities and correctly marginalizes over observational systematic effects. We demonstrate the validity and efficiency of our approach in obtaining unbiased estimates of power spectra via applications to realistic mock galaxy observations that are subject to stellar contamination and dust extinction. While simultaneously accounting for galaxy biases and unknown noise levels, our method reliably and robustly infers three-dimensional density fields and corresponding cosmological power spectra from deep galaxy surveys. Furthermore, our approach correctly accounts for joint and correlated uncertainties between unknown coefficients of foreground templates and the amplitudes of the power spectrum. This effect amounts to correlations and anti-correlations of up to 10 per cent across wide ranges in Fourier space.

Key words. large-scale structure of Universe – methods: statistical – methods: data analysis

1. Introduction

In recent years, the cosmological community has witnessed great improvements in our understanding of the Universe. This progress is particularly due to the spectacular results of the *Planck* satellite mission and deep galaxy observations such as those provided by the Baryon Oscillation Sky Survey ([Planck Collaboration I 2011](#); [Dawson et al. 2013](#); [Planck Collaboration XIII 2016](#)). These results set high standards for future analyses of cosmological data, with an ever increasing need to control uncertainties and systematic effects in observations in order not to misinterpret data when searching for cosmological signals. To address these needs, data science is challenged to provide ever more robust data models accounting for complex systematic effects and allowing for accurate marginalization over unknowns when interpreting cosmological data.

A particular challenge in existing and coming deep galaxy redshift surveys arises from the need to properly understand selection processes of galaxies from which cosmological surveys are constructed (see e.g. [Huterer et al. 2013](#)). Such an identification was conducted for mitigating star-galaxy contamination of the first SDSS photometric galaxy catalogue ([Scranton et al. 2002](#)). The problem is further exacerbated by our lack of understanding of galaxies as tracers of the underlying dark matter field when performing cosmological inference. In particular, all our indicators for completeness rely on the relatively slow homogeneous and isotropic evolution of galaxy densities relative to dark

matter densities. When the observation is further hindered by instrumental and/or terrestrial effects, this leads to a complex and challenging analysis problem.

In particular, [Ross et al. \(2011\)](#) and [Ho et al. \(2012\)](#) have identified that contamination by bright stars alter significantly the intrinsic clustering signal of the observed photometric galaxy sample at large scales. The last SDSS release based on DR12 photometry still shows this problems in the measured correlation function ([Ross et al. 2017](#)). Effects due to foreground stars, dust, seeing, and sky background intensity have the greatest potential to cause systematic deviations in the clustering signal (see e.g. [Ho et al. 2015](#); [Leistedt & Peiris 2014](#)). Selection of spectroscopic galaxies is not affected immediately by the same effect, but it is subject to other systematics, such as fibre collisions, target priority conflicts, and fibre plate fixations. All these data contaminations constitute a particular nuisance, since foreground affects are also affecting the noise properties of observed galaxy samples via varying attenuation or target contamination across the sky.

In the large-scale structure community, foreground effects are traditionally treated by weighting observed galaxies to homogenize the distribution of the traced density field across the sky (see e.g. [Ross et al. 2011, 2017](#); [Sánchez et al. 2013](#)). This approach neglects possible and sometimes counter-intuitive correlations between contamination effects and power spectrum amplitudes across wide ranges in Fourier space. It also ignores the effects of modified observational noise properties that are due to target contamination. Several methods have also been proposed

to account for additive contributions from unknown foregrounds in photometric and spectroscopic galaxy observations (see e.g. Tegmark et al. 1998; Leistedt & Peiris 2014; Ho et al. 2015). The literature on cosmic microwave background analyses also provides a plenitude of approaches to account for linear additive foreground contributions (see e.g. Tegmark & Efstathiou 1996; Hinshaw et al. 2007; Eriksen et al. 2008; Sudevan et al. 2017; Vansyngel et al. 2016; Elsner et al. 2017).

However, all these approaches do not properly account for multiplicative foreground and target contaminations that also affect the noise. In this work we expand on the idea that foreground effects are more closely related to multiplicative rather than additive contributions. This is similar to the discussion presented in Huterer et al. (2013), who used a multiplicative correction in observed galaxy densities.

In this work we seek to account for these effects when inferring cosmological power spectra from observations. Literature provides a plenitude of various statistically more or less rigorous approaches to measure power spectra. Several of these methods rely on Fourier transform based methods or exploit Karhunen-Loève or spherical harmonics decompositions (see e.g. Feldman et al. 1994; Tegmark 1995; Hamilton 1997a; Yamamoto 2003; Percival et al. 2004; Tegmark et al. 1997, 2004; Pope et al. 2004; Fisher et al. 1994; Heavens & Taylor 1995; Tadros et al. 1999; Percival 2005). Other approaches aim at inferring the real space power spectrum via likelihood methods (Ballinger et al. 1995; Hamilton 1997a,b; Tadros et al. 1999; Percival 2005).

In the Bayesian community several approaches have been proposed to jointly infer three-dimensional density fields and their corresponding cosmological power spectra (see e.g. Jasche et al. 2010; Granett et al. 2012; Alsing et al. 2016). We also note that similar approaches were explored for analyses of cosmic microwave background data (see e.g. Wandelt et al. 2004; O’Dwyer et al. 2004; Eriksen et al. 2004, 2007; Jewell et al. 2004, 2009; Larson et al. 2007).

To account for such effects of foreground and target contamination in a statistically rigorous fashion, we propose a hierarchical Bayesian approach to jointly and self-consistently infer three-dimensional density fields and corresponding power spectra and coefficients of a set of different foreground templates. This work builds upon our previously developed Bayesian inference algorithm ARES (Algorithm for REconstruction and Sampling, see e.g. Jasche et al. 2010; Jasche & Wandelt 2013; Jasche & Lavaux 2015; Lavaux & Jasche 2016).

The manuscript is structured as follows. In Sect. 2 we give a brief overview of the statistical model that we propose. First we recall in Sect. 2.1 the hierarchical Bayesian inference approach on which our code, ARES, is based. Then we describe in Sect. 2.2 the necessary modifications of the model for foreground effects and in Sect. 2.3 the modification to the original inference algorithm. In Sect. 3 we describe the generation of artificial data used to test the performance of the sampling framework in Sect. 4. Finally, we discuss our results and give an outlook on future applications in Sect. 5.

2. Bayesian inference model

This section provides a brief overview of our previously presented Bayesian inference framework ARES (see e.g. Jasche et al. 2010; Jasche & Wandelt 2013; Jasche & Lavaux 2015). We also give a detailed description of the modifications enabling us to account for foreground and target contamination in deep galaxy observations.

2.1. ARES framework

As discussed in the introduction, the current work builds upon the previously developed code ARES (see e.g. Jasche et al. 2010; Jasche & Wandelt 2013; Jasche & Lavaux 2015). This full Bayesian large-scale structure inference method aims at precision inference of cosmological power spectra from galaxy redshift surveys. Specifically, it performs joint inferences of three-dimensional density fields, cosmological power spectra, and luminosity-dependent galaxy biases and corresponding noise levels for different galaxy populations in the survey (Jasche et al. 2010; Jasche & Wandelt 2013; Lavaux & Jasche 2016).

The ARES algorithm addresses a large-scale data interpretation problem involving many millions of parameters. For the sake of clarity, we describe below the corresponding data model for a single galaxy population. The generalization of this data model to account for an arbitrary number of galaxy populations with respective stochastic and systematic uncertainties has been described in our previous works (Jasche & Wandelt 2013; Lavaux & Jasche 2016). For a single galaxy population, the data model is given as

$$N_i = \bar{N}R_i(1 + bD_i\delta_i) + \epsilon_i, \quad (1)$$

where N_i is the number of galaxies in the i th grid element, \bar{N} is the mean density of the galaxy population, R_i is the overall linear response operator of the survey, describing redshift and target completeness, b is the galaxy population bias, D_i is the cosmic growth factor at the position of i th grid element, δ_i is the density contrast at a reference redshift in this same grid element, and ϵ_i denotes random but structured instrumental noise. In addition, as described in our previous work, the observational noise is assumed to be a Gaussian approximation to Poissonian noise, neglecting the influence of the signal itself. This assumption yields the corresponding noise covariance matrix as

$$\langle \epsilon_i \epsilon_j \rangle = \bar{N}R_i \delta_{i,j}^K, \quad (2)$$

with $\delta_{i,j}^K$ being the Kronecker-Delta (i.e. equal to one if $i = j$ and zero otherwise). Finally, we assume a homogeneous and isotropic Gaussian prior for density contrast amplitudes δ_i . For further details we refer to Jasche & Wandelt (2013).

In order to provide full Bayesian uncertainty quantification, the algorithm explores the joint parameter space of density fields, power spectra, galaxy biases, and noise parameters via an efficient block sampling scheme. We show in Fig. 1 a visualization of this iterative sampling procedure including the foreground-sampling method presented in this work. Iterative execution of these respective sampling steps provides us with a valid Markov chain and a numerical representation of the full joint-target posterior distribution. We also note that we use an upgraded version of the ARES algorithm here for which we employ the messenger method discussed in Elsner & Wandelt (2013). This particular implementation of the Wiener posterior sampling method has been demonstrated to improve upon the statistical efficiency of previous implementations (Jasche & Lavaux 2015).

2.2. Foreground and target contamination model

Spectroscopic completeness is generally computed by the ratio of the number of observed spectra and the number of all photometric targets for a given area in the sky. This ratio is assumed to hold for any pointing of this area. However, in addition to galaxies, a number of unknown contamination can also contribute or

It should be remarked that foreground contributions, as modelled here, are not just mere additive contributions to the signal to infer, but they also have pronounced effect on the varying noise properties across the survey. Hence, as can be seen from Eq. (7), inferring the foreground coefficients $\{\alpha_n\}$ is a highly non-linear analysis task.

2.3. Sampling foreground coefficients

As described by the likelihood distribution given in Eq. (7), foregrounds of different catalogues as labelled by the superscript c can be sampled independently. For this reason and without loss of generality, here we provide the sampling procedure only for a single galaxy catalogue. Then the conditional posterior distribution for the coefficients $\{\alpha_n\}$ of the respective foreground templates can be written as

$$\mathcal{P}(\{\alpha_n\}|\{N_i\}, \{\delta_i\}, \{\bar{N}\}, \{b\}) \propto \mathcal{P}(\{\alpha_n\}) \mathcal{L}(\{N_i\}|\{\delta_i\}, \{\bar{N}\}, \{b\}, \{\alpha_n\}), \quad (8)$$

where $\mathcal{P}(\{\alpha_n\})$ is the prior of foreground coefficients and $\mathcal{L}(\{N_i\}|\{\delta_i\}, \{\bar{N}\}, \{b\}, \{\alpha_n\})$ is the likelihood for a single catalogue as given by Eq. (7). In the absence of any further information on the amplitudes of foreground coefficients α_n , we follow the maximal agnostic approach by using a uniform prior $\mathcal{P}(\{\alpha_n\}) = 1$. It can be seen from Eqs. (8) and (7) that the conditional posterior distribution does not factorize in the coefficients α_n for the respective foreground templates. To correctly account for the conditional dependencies between the coefficients $\{\alpha_n\}$, we propose to use a block-sampling procedure to sequentially draw random variates of α_n with respect to all other values. This is achieved by introducing the following sequence of sequential sampling steps to the full ARES framework:

$$\begin{aligned} \alpha_0 &\sim \mathcal{P}(\alpha_0|\{\alpha_n\} \setminus \alpha_0, \{N_i\}, \{\delta_i\}, \{\bar{N}\}, \{b\}) \\ \alpha_1 &\sim \mathcal{P}(\alpha_1|\{\alpha_n\} \setminus \alpha_1, \{N_i\}, \{\delta_i\}, \{\bar{N}\}, \{b\}) \\ \alpha_2 &\sim \mathcal{P}(\alpha_2|\{\alpha_n\} \setminus \alpha_2, \{N_i\}, \{\delta_i\}, \{\bar{N}\}, \{b\}) \\ &\vdots \\ \alpha_{N_{fg}-1} &\sim \mathcal{P}(\alpha_{N_{fg}-1}|\{\alpha_n\} \setminus \alpha_{N_{fg}-1}, \{N_i\}, \{\delta_i\}, \{\bar{N}\}, \{b\}) \end{aligned}$$

where the symbol “ \sim ” indicates a random draw from respective distributions. This sampling procedure integrates well into the ARES framework, as indicated in Fig. 1. Although drawing respective realizations of the foreground coefficients $\{\alpha_n\}$ is a non-linear process, a direct sampling procedure exists. The detailed derivation of the foreground coefficient sampler is presented in Appendix A. The detailed algorithm for generating the respective random variates is given in Algorithm 1.

3. Generation of Gaussian mock data

To test the validity and performance of the modified ARES sampling framework, we follow a similar approach, as discussed in our previous works (Jasche et al. 2010; Jasche & Wandelt 2013; Jasche & Lavaux 2015). These mock catalogues are generated in accordance with the data model described in Eq. (7), including various foreground effects. We generate artificial galaxy data on a cubic equidistant grid of side length $4000 h^{-1}$ Mpc consisting of 256^3 grid nodes.

First a realization of a cosmic density contrast field δ_i is drawn from a zero-mean normal distribution with covariance matrix corresponding to a cosmological power spectrum. This spectrum, including baryon acoustic oscillations, is calculated according to the prescription described in Eisenstein & Hu (1998) and Eisenstein & Hu (1999). For the numerical evaluation we assume a Λ CDM cosmology with the set of parameters given as ($\Omega_m = 0.3089$, $\Omega_\Lambda = 0.6911$, $\Omega_b = 0.0485$, $h = 0.6774$, $\sigma_8 = 0.8159$, $n_s = 0.9667$), as determined by cosmic microwave background observations of the *Planck* satellite mission (Planck Collaboration XIII 2016).

Following a similar description as discussed in Jasche & Wandelt (2013), we intend to create a realistic scenario to jointly analyse the SDSS DR7 main, the CMASS, and the LOW-Z galaxy sample while accounting for their respective systematic effects. Corresponding artificial data sets are then drawn from the distribution given in Eq. (7). These artificial data sets include the effects of noise, galaxy bias, survey geometries, selection, and foreground effects. In this work, we restrict our tests to accounting for the dominant foreground effects of dust extinction and stellar contamination. The templates for these two effects are presented in the upper panels of Fig. 2; the lower panels show the completeness masks for the respective surveys.

The map describing dust extinction has been generated straightforwardly from the SFD maps (Schlegel et al. 1998). We constructed a HEALPix map of the reddening at $N_{\text{side}} = 2048$ by linearly interpolating the values of the SFD map (Górski et al. 2005). The star map is built from different pieces of information. The first component consists of computing a MANGLE description¹ of the geometry of the spectroscopic plates. We then count the number of stars with apparent magnitudes $20.3 < i_{\text{PSF}} < 20.6$ present in each single non-overlapping polygon. We convert these MANGLE description into a HEALPix map and divide the value in each pixel by the area of the overlapping polygon. This results in a map for which each pixel has an estimated star count by steradians, thus a star density. We have chosen to average by spectroscopic plates to reduce shot-noise in the estimate. A better estimate would have been obtained from the geometrical description of the photometric tiling but at the cost of increased noise.

Furthermore, for the SDSS DR7 main sample component of the mock data, we assume a radial selection function following from a standard Schechter luminosity function with standard r -band parameters ($\alpha = -1.05$, $M_* - 5\log_{10}(h) = -20.44$), and we limit the survey to only include galaxies within an apparent Petrosian r -band magnitude range $13.5 < r < 17.6$ and within the absolute magnitude ranges $M_{\text{min}} = -17.0$ to $M_{\text{max}} = -23.0$. As usual, the radial selection function $f(z)$ is then given by the integral of the Schechter luminosity function over the range in absolute magnitude.

For the CMASS and LOW-Z component we have used numerical estimates of the selection functions by computing a histogram of the corresponding $N(d)$ in the actual data sets (e.g. for DR12 Ross et al. 2017) (d being the co-moving distance from the observer). To account for the different selection effects in the northern and southern galactic plane, we also correspondingly split our mock data sets into the CMASS and LOW-Z catalogues. The respective radial selection functions are presented in Fig. 3. The average of the product of the two-dimensional survey geometry $C^c(\hat{x}) = C_{i(\hat{x})}^c$ and the selection function $f(x)$ at each

¹ The MANGLE software is originally provided by Swanson et al. (2008).

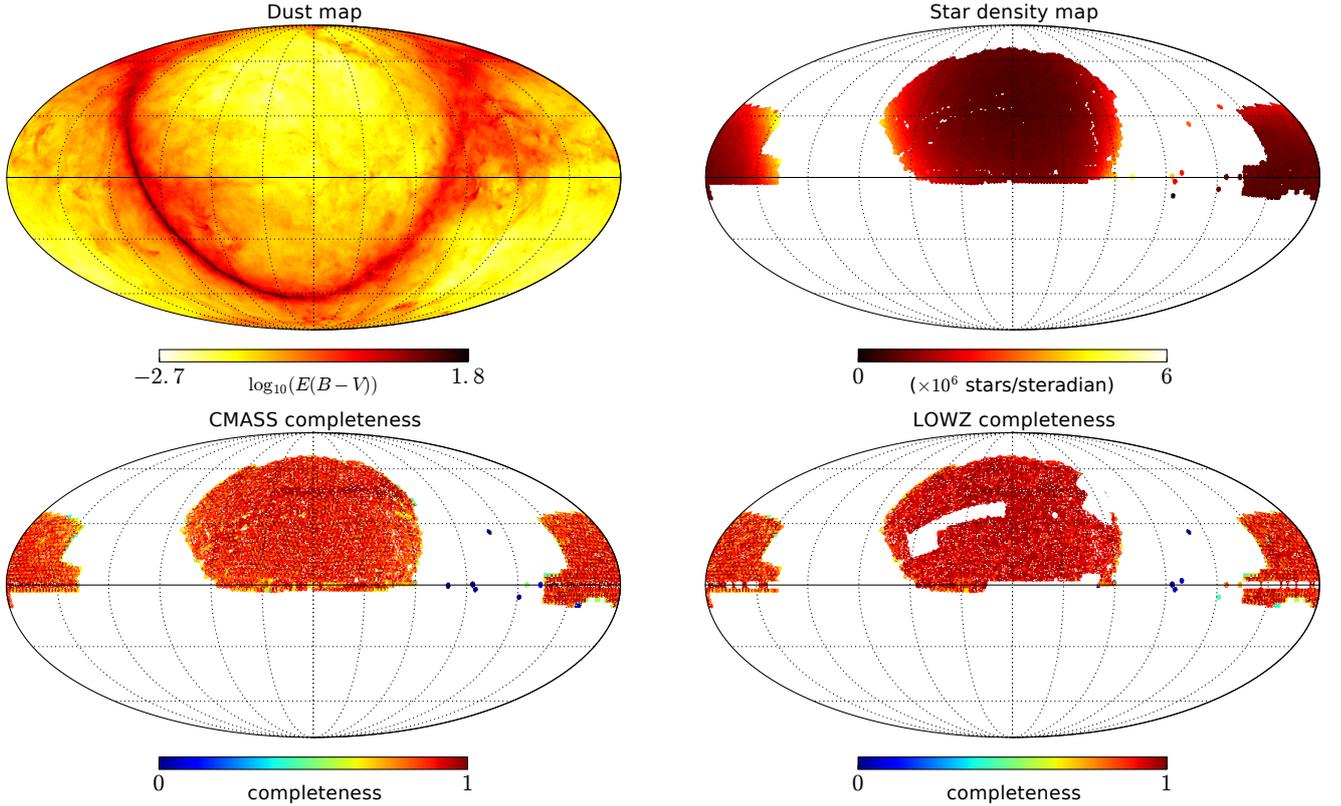


Fig. 2. Foreground templates (*top row*) and the observed sky completenesses (*bottom row*) used to generate and analyse the mock catalogue in this work. The *upper left panel* shows the reddening map derived from the data of [Schlegel et al. \(1998\)](#). The *upper right panel* is a star map count obtained as detailed in Sect. 3. The *lower left panel* gives the observed completeness for the mock CMASS survey and the *lower right panel* for the mock LOW-Z survey. These maps have been generated from SDSS-DR12 data ([Eisenstein et al. 2011](#)).

grid element in the three-dimensional volume yields the survey response operator:

$$R_i^c = \frac{1}{|\mathcal{V}_i|} \int_{\mathcal{V}_i} d^3x C^c(\hat{x}) f^c(x), \quad (9)$$

with $|\mathcal{V}|$ the volume of the set \mathcal{V} , \mathcal{V}_i indicating the volume represented by the i th grid element.

Given these definitions and a realization of the three-dimensional density field δ_i , realizations of artificial galaxy observations for respective catalogues labelled with c can be obtained by evaluating

$$N_i^c = \bar{N}^c R_i^c M_i^c (1 + b^c \delta_i) + \sqrt{\bar{N}^c R_i^c M_i^c} \epsilon_i, \quad (10)$$

where ϵ_i is a white-noise field drawn from zero-mean and unit variance normal distribution.

4. Results

In this section we discuss results obtained by applying the modified ARES algorithm to artificial mock data. In particular, we focus on the validity and statistical efficiency of the algorithm.

4.1. Statistical efficiency of the sampler

To test the statistical efficiency of our sampler, we follow a standard test procedure as described in previous works (see e.g. [Jasche et al. 2010](#); [Jasche & Wandelt 2013](#); [Jasche & Lavaux 2015](#)). We test the initial burn-in phase by starting the sampler

from an over-dispersed state and monitoring transitions in parameter space as a sequence of steps in the Markov chain. Typically, this test reveals a coordinated drift of inference parameters towards their target values. After the chain moved to a preferred region in parameter space, it started to correctly explore the target posterior distribution via the random-walk Gibbs sampling approach. At this stage it is assumed that the sampler has passed the initial burn-in phase, and we start recording samples of the Markov chain. In Fig. 4 we show the sequence of sampled posterior power spectra during the burn-in phase. For this test we started from an over-dispersed state by multiplying the initial guess of a Gaussian random density contrast field by a factor of 0.1. Figure 4 nicely demonstrates the initial drift towards the preferred region in parameter space. The initial burn-in phase consists of about ~ 2000 Markov steps. As a side remark, we note that generation of individual samples requires an investment of ~ 1.87 CPU hours per sample for the present scenario of dealing with five different galaxy sub-catalogues and two foregrounds.

To further test the statistical efficiency of the Gibbs sampling procedure, we estimated its efficiency by generating independent Markov samples. Generally, subsequent samples in a Markov chain are correlated and do not qualify for independent samples of the target posterior distribution. To estimate how many independent samples can be drawn from a Markov chain with a given number of transition steps, one has to determine the length over which sequential samples are correlated. This correlation length characterizes the statistical efficiency of generating independent realizations of a parameter θ as follows:

$$C(\theta)_n = \frac{1}{N-n} \sum_{i=0}^{N-n} \frac{\theta^i - \langle \theta \rangle}{\sqrt{\text{Var}(\theta)}} \frac{\theta^{i+n} - \langle \theta \rangle}{\sqrt{\text{Var}(\theta)}}, \quad (11)$$

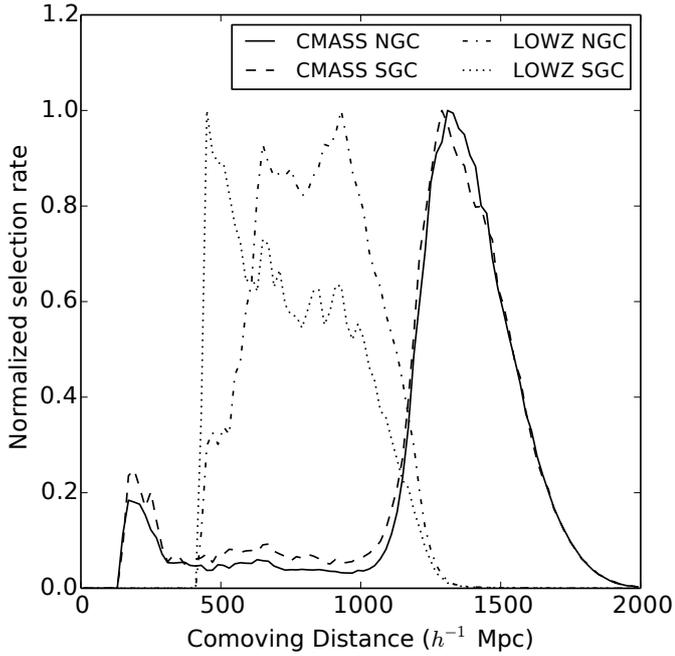


Fig. 3. Radial selection functions for the CMASS and LOW-Z sample that we have used to generate the mock data to closely resemble the actual SDSS3 BOSS data. The CMASS selection is shown as the thick solid (north galactic cap) and dashed line (south galactic cap). The LOW-Z selection is shown as thin dash-dotted (north galactic cap) and dotted lines (south galactic cap).

where n is the distance in the chain measured in iterations, $\langle \theta \rangle = 1/N \sum_i \theta^i$ and $\text{Var}(\theta) = 1/N \sum_i (\theta^i - \langle \theta \rangle)^2$, and N is the total number of samples in the chain.

As an illustration, we show in Fig. 5 the correlation length for the power spectrum amplitudes of different modes in Fourier space, as indicated in the plot. The typical correlation length for a BOSS like survey analysis is of the order of ~ 100 Markov transitions. These results demonstrate the numerical feasibility of complex full Bayesian analyses of present and next-generation surveys.

4.2. Inference of template coefficients

As discussed above, the proposed hierarchical Bayesian inference machine aims to account for the uncertainties of systematic effects arising from foreground effects. The ARES framework correctly accounts for all joint and correlated uncertainties of different inference parameters and even across the five different mock galaxy surveys, as used in this work. To illustrate this, we present in Fig. 6 two-dimensional marginal posterior distributions for the corresponding foreground template coefficients. The different panels indicate various degrees of correlation between different foreground coefficients across the five mock catalogues. We would also like to point out that generally, distributions for foreground template coefficients are highly non-Gaussian and can have sharp transitions that are due to the requirement that effective contamination templates are required to have a positive sign.

As an interesting by-product of our sampling procedure, we are able to provide effective survey response maps that account for the a priori unknown systematics that are due to foreground and target contamination. In particular, mean and standard deviations of these maps can be estimated by evaluating Eq. (4) for

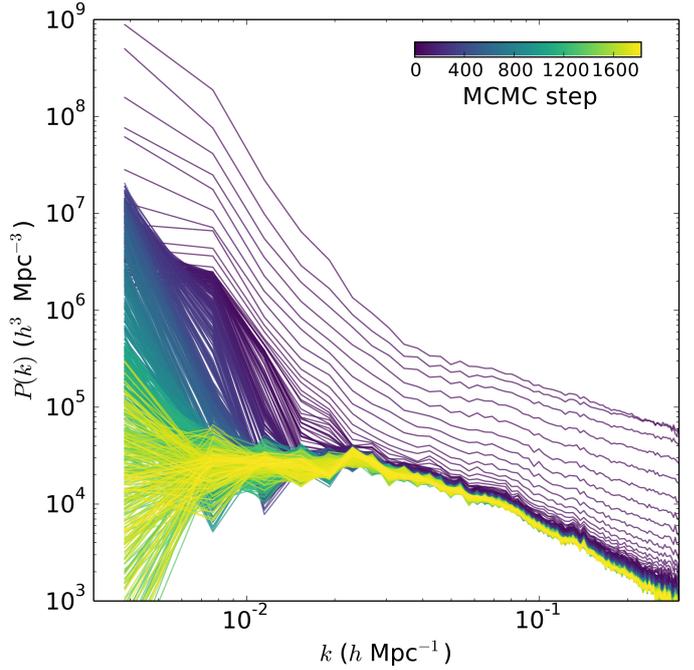


Fig. 4. Sequence of posterior power spectrum realizations during the initial sampling steps of the Markov chain as indicated by the colour bar in the plot. The Markov chain performs a coherent drift from an over-disperse initial state towards the preferred region in parameter space. This initial burn-in phase lasts for ~ 2000 Markov transitions.

every foreground template parameter coefficient in the Markov chain and multiplying it with the estimated completeness map $C_{i,\text{obs}}$. The result is demonstrated in Fig. 7. Most corrections appear at the boundary of the mask. These are the regions most affected by foreground stars in our test scenario. We note that the corresponding standard deviations, as shown in the right panel of Fig. 7, also show increased uncertainty in these regions. This demonstrates that the algorithm accounts for a larger uncertainty for more unreliable portions of the data and optimally extracts cosmological information from observations, as discussed in the following.

4.3. Inferred three-dimensional density fields

Although this work focusses primarily on the inference of cosmological power spectra, it is instructive to also study inferred three-dimensional density fields. In particular, we would like to highlight the effect of foreground contaminations on the inference of density fields from galaxy redshift surveys. To do so, we compare two ARES analyses of the generated mock galaxy catalogue, with and without foreground treatment. From these two Markov chains, we can calculate the ensemble mean density and corresponding variance fields. Results are presented in Fig. 8. The analysis without a detailed treatment of foreground contaminations shows residual large-scale features and erroneous power particularly close to the survey boundaries. These regions are affected the most by stellar contamination, as indicated by the corresponding foreground map shown in Fig. 2.

In contrast, our ARES run with detailed Bayesian foreground treatment shows a homogeneous density distribution throughout the entire observed domain. We also note that variance maps of the corrected and uncorrected ARES run look very similar. Apparently, the erroneous features in the uncorrected ARES analysis do not affect the corresponding variance map. Since the data

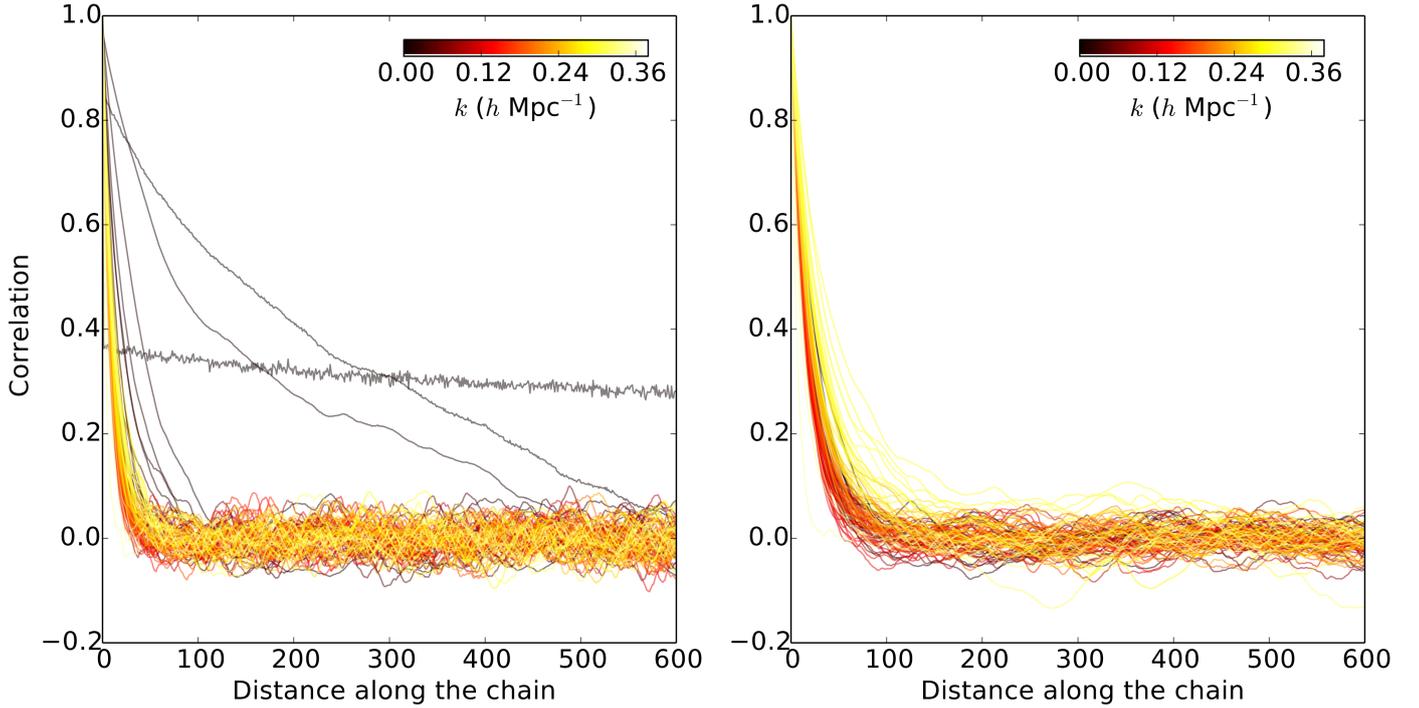


Fig. 5. Correlation length of the sampler for different modes as indicated by the colour bar on the left. Most of the modes have typical correlation lengths of the order of 100 samples. Some large-scale modes exhibit a longer correlation length, as explained in the text.

model does not account for the systematic uncertainties associated with foreground contaminations for the uncorrected run, the reconstructed erroneous large-scale power in the field will be fully attributed to the inferred large-scale structure. From Fig. 8 it is visually evident that our data model including the treatment of foreground contaminations is much more robust against such misinterpretations. In the following we consider the inferred cosmological power spectra in more detail.

4.4. Inferred cosmological power spectra

One of the most important features of the ARES algorithm is its ability to jointly infer three-dimensional density fields, corresponding cosmological power spectra, galaxy biases, noise levels, and coefficients of several foreground templates, including a detailed treatment of all joint and correlated uncertainties. Since the ARES framework yields proper Markov chains, we are able to correctly marginalize over all joint uncertainties when focussing on the analysis of specific target quantities such as the cosmological power spectrum. Specifically, [Jasche & Wandelt \(2013\)](#) have demonstrated that the ARES algorithm reveals and correctly treats the anti-correlation between bias amplitudes and power spectrum, a 20% effect across wide ranges in Fourier space. Here we also take the unknown coefficients of several foreground templates into account.

To study the impact of foreground contamination on the analyses of cosmological power spectra in deep galaxy surveys, we compare inference results obtained from our two ARES runs with and without corresponding corrections. The results for inferred power spectra are presented in Fig. 9, where we show the univariate marginal posterior distribution for power spectrum amplitudes at different modes in Fourier space. For the Markov chain without foreground treatment, we can clearly observe excessive power at the largest scales. This observation corresponds

to the excessive large-scale power observed in corresponding inferred three-dimensional density fields, as discussed above. In addition, we can observe a slight bias with respect to the true underlying power spectrum from which mock observations were generated. This can easily be understood by inspecting the data model described in Eq. (5), which shows a certain potential for a degeneracy between foreground coefficients and galaxy biases. If foreground effects are not treated correctly, then some of the foreground contributions will erroneously be compensated for by sampled galaxy bias amplitudes, introducing the offset between true and recovered power spectra shown in Fig. 9. In contrast, inferred power spectra for the run with foreground treatment are unbiased with respect to the true underlying power spectrum over the full domain of Fourier modes considered in this work. The shape of the recovered power spectrum at the largest scales is in excellent agreement with the true fiducial model.

We further studied the effect of different foreground coefficients on power spectrum amplitudes at different scales in Fourier space. In particular, we calculated the cross-correlation matrix between the foreground coefficients of the five mock catalogues and power spectrum amplitudes from the posterior samples of the corresponding Markov chain. Results are presented in Fig. 10. Correlations and anti-correlations can amount to up to 10 per cent across all modes in Fourier space.

Additionally, we tested whether the sampler correctly accounts for the combined effects of foreground contaminations, galaxy biases, and unknown noise amplitudes by estimating the co-variance matrix of inferred power spectra from the ARES runs. Figure 11 shows that the co-variance matrix for both runs exhibit a strongly diagonal shape, indicating that the algorithm correctly accounted for the otherwise erroneous mode coupling introduced by survey geometry and foreground effects. Residual off-diagonal contributions amount to less than 10 per cent.

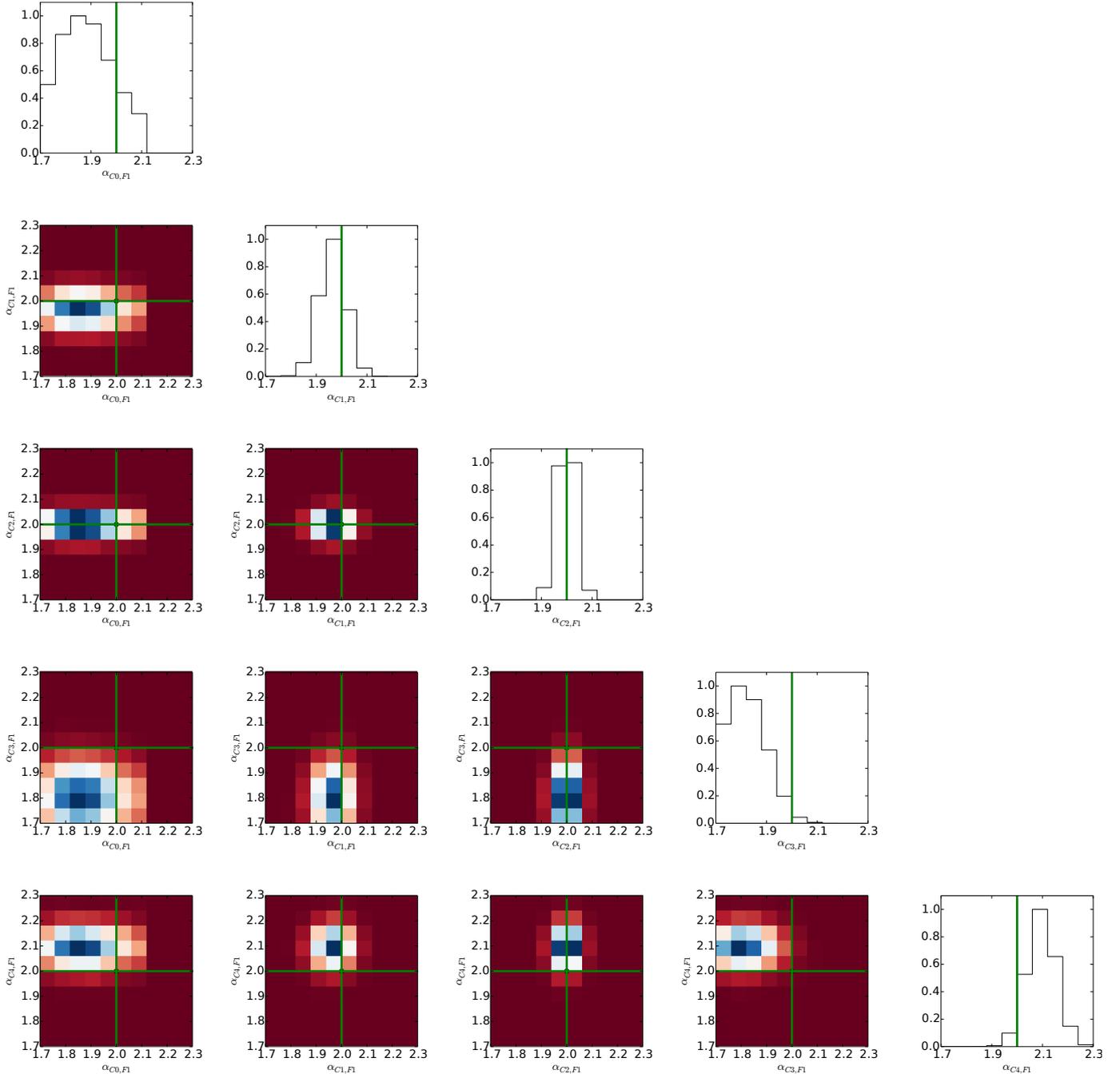


Fig. 6. Marginalized two-dimensional probability distributions for foreground contamination coefficients of the five mock catalogues used in this analysis. Green lines indicate the true values of foreground coefficients as used for the generation of the artificial mock galaxy survey. The different panels show various degrees of correlation between inferred coefficients that are correctly accounted for by our Markov chain.

These results clearly demonstrate the feasibility of dealing with strong but unknown foreground contaminations when inferring cosmological power spectra from deep galaxy observations.

5. Summary and conclusion

Major challenges for the analysis of next-generation deep galaxy redshift surveys arise from the requirement to account for an increasing amount of systematic and stochastic uncertainties. Foreground effects and target contaminations due to stars and dust, for instance, can greatly affect the observation of galaxies. If not accounted for properly, these effects can yield erroneous modulations of galaxy number counts across the sky,

which hinders the immediate inference of power spectra and three-dimensional density fields from such galaxy samples.

To address this issue, we have described a fully Bayesian treatment of unknown foreground contamination in the process of inferring cosmological power spectra from deep galaxy surveys. In particular, we built upon the previously presented Bayesian inference framework ARES (for details see [Jasche et al. 2010](#); [Jasche & Wandelt 2013](#); [Jasche & Lavaux 2015](#)). The ARES algorithm aims to jointly infer three-dimensional density fields, the corresponding cosmological power spectrum, luminosity-dependent galaxy biases, and unknown noise levels. Because it is a full Bayesian inference engine, ARES furthermore correctly provides joint and correlated

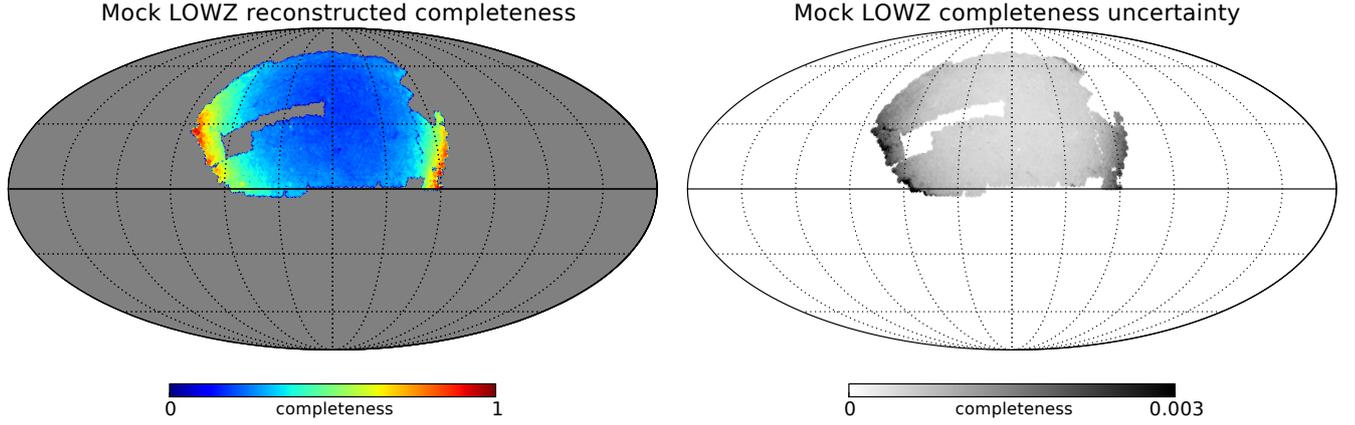


Fig. 7. Ensemble mean of the effective survey response operator (*left panel*) and corresponding standard deviation map (*right panel*). The ensemble mean is renormalized by the highest pixel value, as the absolute value does not have a meaning independent of the mean density \bar{N} and the radial selection function. The two above maps should be compared to the north galactic cap of the map in *the lower right panel* of Fig. 2. The ensemble mean is quite different owing to the introduced star contamination, which could introduce contamination in targets. This manifests itself by an over-completeness on the edge of the map. The *right* map shows a similar trend, but touching the uncertainty on the selection this time.

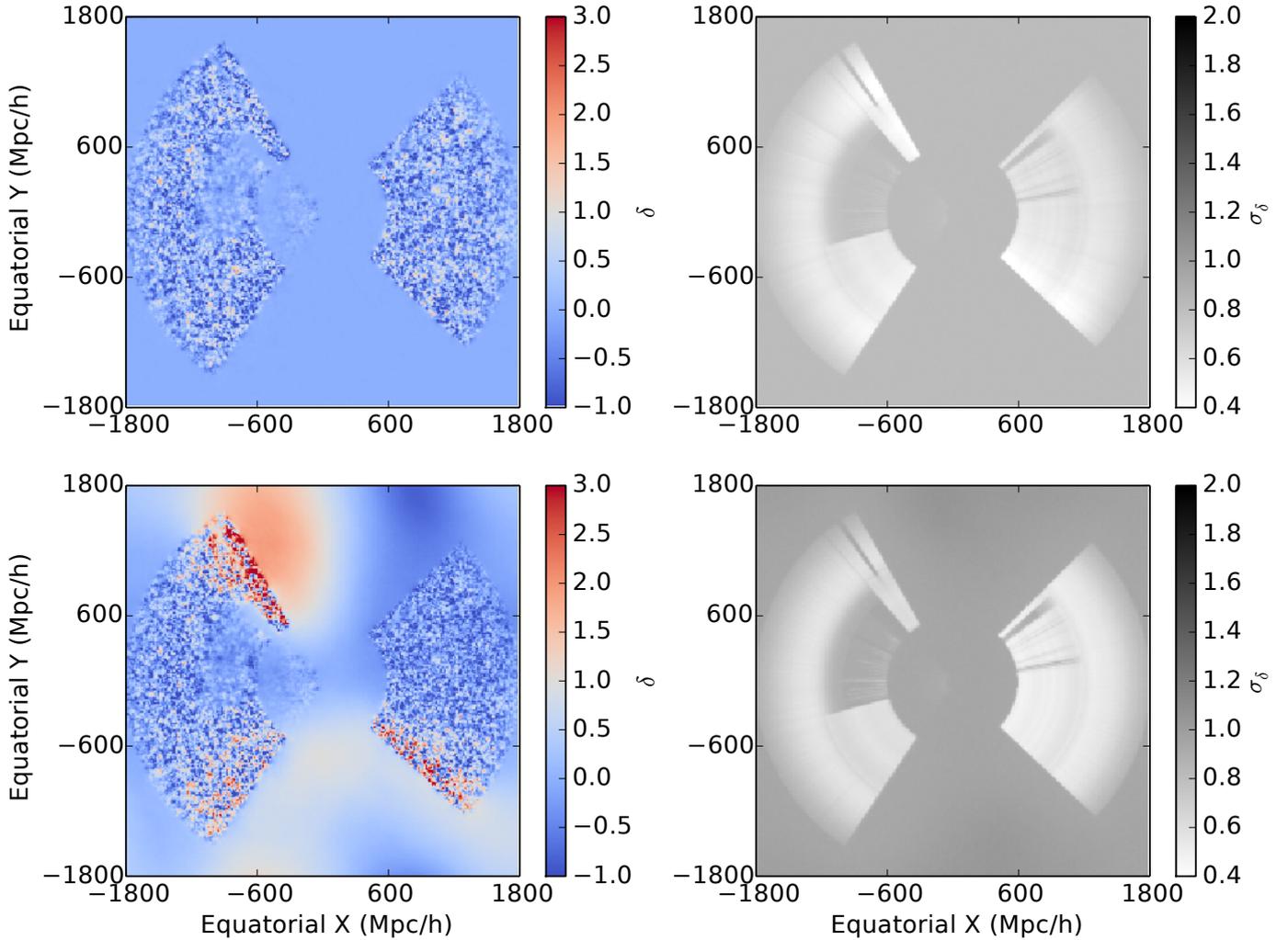


Fig. 8. Slices through three-dimensional ensemble mean (*left panels*) and variance fields (*right panels*). The *top panels* show results obtained with foreground correction, while the *bottom panels* show results without any foreground correction. As for the power spectrum, we find an excessive large-scale power when foreground corrections are not applied. When the foreground is computed self-consistently, the result is a non-contaminated reconstruction. The variance fields are also affected, as is shown by the notably darker bottom on average compared to the top slice, which indicates higher variance.

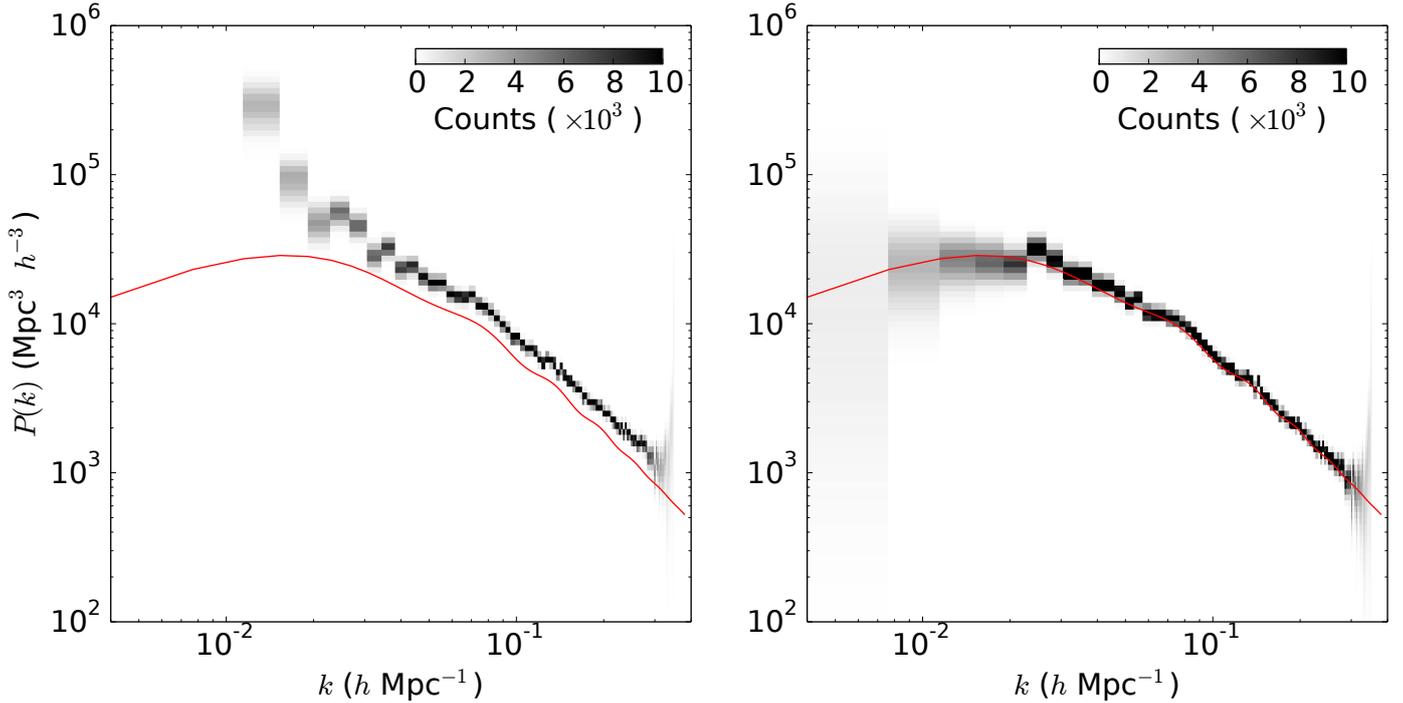


Fig. 9. Univariate posterior distributions of power spectrum amplitudes for a test without (*left panel*) and with (*right panel*) foreground corrections over the full range of Fourier modes considered in this work. Red lines correspond to the true underlying cosmological power spectrum from which mock data sets were generated. The *left panel* clearly shows that uncorrected foreground effects yield excessive power for large-scale modes and also introduce an overall biased result. In contrast, *the right panel* shows results obtained from our test with foreground corrections. Clearly, a detailed treatment of all foreground effects permits us to obtain an unbiased measurement of power spectrum amplitudes over the full range of Fourier modes.

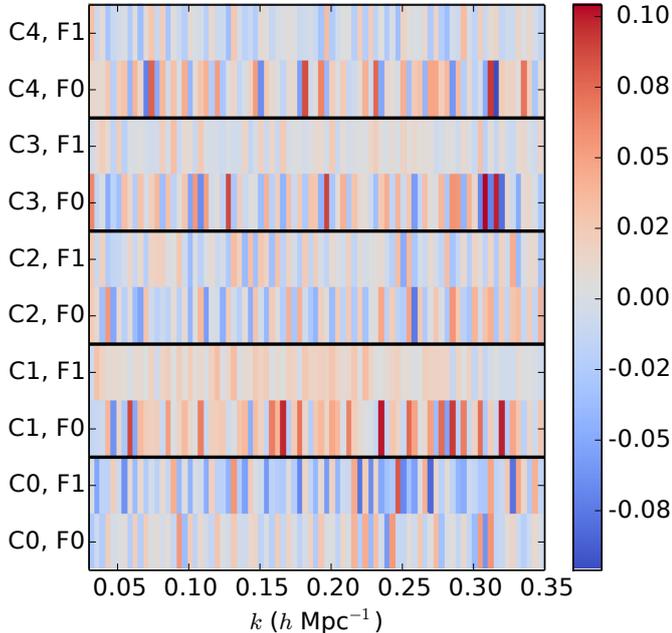


Fig. 10. Cross-correlation between power spectrum amplitudes at different Fourier modes and negative (labelled F0) as well as positive (labelled F1) foreground coefficients for the five respective catalogues (labelled C0 to C4).

uncertainties for all target quantities by performing an efficient Markov chain Monte Carlo sampling within a block sampling scheme, as indicated in Fig. 1.

We have extended this hierarchical Bayesian framework by also including an additional sampling procedure to account for

foreground and target contamination effects. As discussed in Sect. 2.2, such contaminating effects particularly affect the estimation of the spectroscopic completeness of a given galaxy survey. A naive estimation of the probability of acquisition of galaxy spectra at given positions in the sky, by calculating ratios of observed galaxy spectra and all observed photometric targets, ignores the possibility that photometric targets are contaminated by foreground stars or dust extinction. Such effects are likely to artificially increase or decrease the estimated number of galaxy targets in observations. In consequence, these effects introduce an artificial modulation of observed galaxy number densities across the sky, which in turn yields erroneous large-scale power in inferred cosmological power spectra.

As demonstrated in Sect. 2.2, foreground and target contaminations can be accounted for by describing the mismatch between actual galaxies and observed photometric targets as multiplicative correction factors to estimated spectroscopic completeness maps. These correction factors can be described as combinations of various templates for foreground effects, such as introduced by stars or dust, and corresponding unknown template coefficients. The aim of this work is to jointly infer these template coefficients together with the three-dimensional density field, corresponding cosmological power spectra, galaxy biases, and unknown noise levels. This goal can be achieved by adding an additional sampling scheme for foreground coefficients to the block-sampling framework of the ARES algorithm. Interestingly, as discussed in Sect. 2.3 and Appendix A, we were able to derive a direct sampling scheme by introducing auxiliary random fields and marginalizing over them. The corresponding algorithm to generate random realizations of foreground template coefficients is given in Algorithm 1.

We have tested the performance of improved ARES algorithms through applications to mock galaxy observations. To

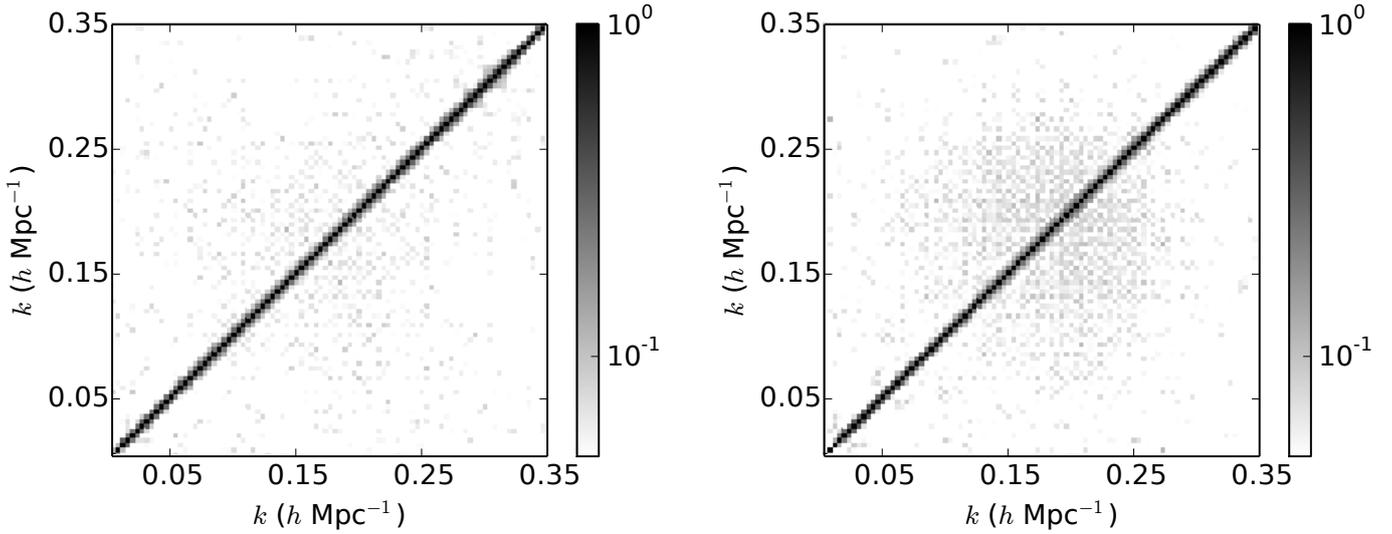


Fig. 11. Correlation matrix of power spectrum amplitudes with respect to their mean value in the case of unaccounted foregrounds (*left panel*) and modelled foregrounds (*right panel*). The correlation matrix is normalized using the variance of the power spectrum amplitudes. We note that the colour map is truncated at a correlation level of 5×10^{-2} . We estimate that all values below this threshold are too noisy to be cleanly represented and discard them.

further evaluate the foreground and target contamination effects on the inference of cosmological power spectra, we have compared two test scenarios, with and without treatment of foreground effects. Corresponding artificial galaxy observations were self-consistently generated according to the data model described in Sect. 3. These artificial observations sought to emulate realistic features of the SDSS DR7 main, the LOW-Z, and the CMASS galaxy sample. This effectively resulted in five artificial galaxy surveys that are jointly handled by the ARES framework, while self-consistently accounting for their respective systematic and stochastic uncertainties, including survey geometries, selection effects, galaxy biasing, and foregrounds.

These artificial data were then used to test the statistical performance of our sampling framework. We tested the burn-in behaviour of the algorithm by starting the Markov chain from an over-dispersed state. As described in Sect. 4.1, we started the chain with an initial power spectrum scaled by a factor 0.1. The following burn-in behaviour then manifested itself by a coherent drift of sequential power spectrum samples towards preferred regions in parameter spaces. We estimated the burn-in phase to be completed after ~ 2000 sampling transitions. The statistical efficiency of the sampler was estimated by measuring the correlation length between subsequent posterior power spectrum samples. As demonstrated in Sect. 4.1, the sampler exhibited a correlation length of a few hundred samples. This left us with a numerically efficient sampling framework to explore cosmological power spectra in deep galaxy surveys. It should be remarked that various possibilities exist to further improve statistical efficiencies, but details are left to future publications.

The results for inferring foreground and target contamination template coefficients are described in Sect. 4.2. We have used two realistic foreground templates describing foreground stars in the galaxy and dust extinction. Furthermore, our implementation is general enough to account for an arbitrary amount of foreground templates. To demonstrate the feasibility of inferring these parameters jointly with the cosmological power spectrum, three-dimensional density fields, galaxy biases, and noise levels, we presented two-dimensional marginalized distributions for template coefficients. These results showed that different contamination contributions can be recovered within ≤ 2.3 sigma of their input values. Given the inferred foreground coefficients,

it is furthermore possible to reconstruct an effective completeness mask and corresponding uncertainties. As expected, in the example tested in this work, the uncertainties in the recovered completeness masks are largest in regions where stellar contaminations of the target distribution are also large.

In Sect. 4.3 we studied the effect of foreground and target contaminations on the inference of three-dimensional density fields from galaxy surveys. To do this, we have compared a run with and without foreground treatment. When foreground effects are ignored when density fields are inferred, it yields excessive large-scale power particularly in regions that are most affected by these contaminations. In contrast, a detailed Bayesian treatment of foreground systematics yields inferred density fields showing a homogeneous distribution of power across the inference domain. It must be remarked that if foreground contaminations are not explicitly modelled within the data model, then their effects will be attributed to a real signal.

This result is also in agreement with inferred cosmological posterior power spectra, as presented in Sect. 4.4. The inferred power spectrum of the run without foreground treatment agrees with the visual impression obtained from corresponding three-dimensional density fields. In particular, it reflects the observed excess in large-scale power. Ignoring foreground effects may furthermore lead to an overall bias across the entire Fourier domain with respect to the true underlying power spectrum. In contrast, a detailed Bayesian foreground treatment yields inferred power spectra in agreement with the underlying fiducial truth over the entire range of Fourier modes, as considered in this work.

We also tested different foreground effects on the inference of power spectrum amplitudes by estimating their correlations with foreground template coefficients throughout the full Fourier range. These results showed that correlations and anti-correlations can amount to up to 10 percent throughout wide ranges in Fourier space.

The ARES algorithm also accounts for artificial mode coupling between power spectrum amplitudes as introduced by survey geometries and completeness masks. To demonstrate this, we estimated the correlation matrix of power spectrum amplitudes from our Markov runs. These tests showed that residual artificial mode coupling is typically much less than 10 percent. These results indicate the validity of the algorithm in scenarios

with heavily masked data. We are currently using our method to process the actual BOSS data, which will be presented in our companion paper (Lavaux & Jasche, in prep.).

As demonstrated in this work, a detailed treatment of foreground and target contamination is essential to recover unbiased estimates of three-dimensional density fields and corresponding cosmological power spectra from present and next-generation surveys. The proposed ARES algorithm provides a joint and statistically rigorous Bayesian inference framework to achieve this goal and prevent misinterpretation of observations.

Acknowledgements. This research was supported by the DFG cluster of excellence “Origin and Structure of the Universe” (www.universe-cluster.de). This work made in the ILP LABEX (under reference ANR-10-LABX-63) was supported by French state funds managed by the ANR within the Investissements d’Avenir programme under reference ANR-11-IDEX-0004-02. The Parkes telescope is part of the Australia Telescope, which is funded by the Commonwealth of Australia for operation as a National Facility managed by CSIRO. We acknowledge financial support from “Programme National de Cosmologie and Galaxies” (PNCG) of CNRS/INSU, France. This work was supported in part by the ANR grant ANR-16-CE23-0002.

References

- Alsing, J., Heavens, A., Jaffe, A. H., et al. 2016, *MNRAS*, **455**, 4452
- Anderson, L., Aubourg, E., Bailey, S., et al. 2012, *MNRAS*, **427**, 3435
- Ballinger, W. E., Heavens, A. F., & Taylor, A. N. 1995, *MNRAS*, **276**, L59
- Dagpunar, J. 1988, *Principles of Random Variate Generation* (Oxford University Press)
- Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, *AJ*, **145**, 10
- Eisenstein, D. J., & Hu, W. 1998, *ApJ*, **496**, 605
- Eisenstein, D. J., & Hu, W. 1999, *ApJ*, **511**, 5
- Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, *AJ*, **142**, 72
- Elsner, F., & Wandelt, B. D. 2013, *A&A*, **549**, A111
- Elsner, F., Leistedt, B., & Peiris, H. V. 2017, *MNRAS*, **465**, 1847
- Eriksen, H. K., O’Dwyer, I. J., Jewell, J. B., et al. 2004, *ApJS*, **155**, 227
- Eriksen, H. K., Huey, G., Banday, A. J., et al. 2007, *ApJ*, **665**, L1
- Eriksen, H. K., Jewell, J. B., Dickinson, C., et al. 2008, *ApJ*, **676**, 10
- Feldman, H. A., Kaiser, N., & Peacock, J. A. 1994, *ApJ*, **426**, 23
- Fisher, K. B., Scharf, C. A., & Lahav, O. 1994, *MNRAS*, **266**, 219
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, **622**, 759
- Granett, B. R., Guzzo, L., Coupon, J., et al. 2012, *MNRAS*, **421**, 251
- Hamilton, A. J. S. 1997a, *MNRAS*, **289**, 285
- Hamilton, A. J. S. 1997b, *MNRAS*, **289**, 295
- Heavens, A. F., & Taylor, A. N. 1995, *MNRAS*, **275**, 483
- Hinshaw, G., Nolta, M. R., Bennett, C. L., et al. 2007, *ApJS*, **170**, 288
- Ho, S., Cuesta, A., Seo, H.-J., et al. 2012, *ApJ*, **761**, 14
- Ho, S., Agarwal, N., Myers, A. D., et al. 2015, *J. Cosmol. Astropart. Phys.*, **5**, 040
- Huterer, D., Cunha, C. E., & Fang, W. 2013, *MNRAS*, **432**, 2945
- Jasche, J., & Lavaux, G. 2015, *MNRAS*, **447**, 1204
- Jasche, J., & Wandelt, B. D. 2013, *ApJ*, **779**, 15
- Jasche, J., Kitaura, F. S., Wandelt, B. D., & Enßlin, T. A. 2010, *MNRAS*, **406**, 60
- Jewell, J., Levin, S., & Anderson, C. H. 2004, *ApJ*, **609**, 1
- Jewell, J. B., Eriksen, H. K., Wandelt, B. D., et al. 2009, *ApJ*, **697**, 258
- Larson, D. L., Eriksen, H. K., Wandelt, B. D., et al. 2007, *ApJ*, **656**, 653
- Lavaux, G., & Jasche, J. 2016, *MNRAS*, **455**, 3169
- Leistedt, B., & Peiris, H. V. 2014, *MNRAS*, **444**, 2
- O’Dwyer, I. J., Eriksen, H. K., Wandelt, B. D., et al. 2004, *ApJ*, **617**, L99
- Percival, W. J. 2005, *MNRAS*, **356**, 1168
- Percival, W. J., Burkey, D., Heavens, A., et al. 2004, *MNRAS*, **353**, 1201
- Planck Collaboration I. 2011, *A&A*, **536**, A1
- Planck Collaboration XIII. 2016, *A&A*, **594**, A13
- Pope, A. C., Matsubara, T., Szalay, A. S., et al. 2004, *ApJ*, **607**, 655
- Ross, A. J., Beutler, F., Chuang, C.-H., et al. 2017, *MNRAS*, **464**, 1168
- Ross, A. J., Ho, S., Cuesta, A. J., et al. 2011, *MNRAS*, **417**, 1350
- Sánchez, A. G., Kazin, E. A., Beutler, F., et al. 2013, *MNRAS*, **433**, 1202
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, **500**, 525
- Scranton, R., Johnston, D., Dodelson, S., et al. 2002, *ApJ*, **579**, 48
- Swanson, M. E. C., Tegmark, M., Hamilton, A. J. S., & Hill, J. C. 2008, *MNRAS*, **387**, 1391
- Sudevan, V., Aluri, P. K., Yadav, S. K., Saha, R., & Souradeep, T. 2017, *ApJ*, **842**, 62
- Tadros, H., Ballinger, W. E., Taylor, A. N., et al. 1999, *MNRAS*, **305**, 527
- Tegmark, M. 1995, *ApJ*, **455**, 429
- Tegmark, M., & Efstathiou, G. 1996, *MNRAS*, **281**, 1297
- Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, *ApJ*, **480**, 22
- Tegmark, M., Hamilton, A. J. S., Strauss, M. A., Vogeley, M. S., & Szalay, A. S. 1998, *ApJ*, **499**, 555
- Tegmark, M., Blanton, M. R., Strauss, M. A., et al. 2004, *ApJ*, **606**, 702
- Vansyngel, F., Wandelt, B. D., Cardoso, J.-F., & Benabed, K. 2016, *A&A*, **588**, A113
- Wandelt, B. D., Larson, D. L., & Lakshminarayanan, A. 2004, *Phys. Rev. D*, **70**, 083511
- Yamamoto, K. 2003, *ApJ*, **595**, 577

Appendix A: Sampling foreground coefficients

In this appendix, we derive the sampling procedure for a single foreground template coefficient α_k . As described in Sect. 2.3, the full joint distribution of all template coefficients may be sampled by performing a sequential iterative block sampling procedure. In the absence of any a priori information on the foreground coefficient α_k , we follow a maximally conservative approach by assuming a uniform prior distribution. Using Eq. (8), the logarithm of the conditional posterior distribution for a foreground coefficient in a single galaxy catalogue can then be written as

$$\begin{aligned} \log \mathcal{P}(\alpha_k | \{\alpha_n\} \setminus \alpha_k, \{N_i\}, \{\delta_i\}, \{\bar{N}\}) = \\ - \frac{1}{2} \sum_{i=0}^v \frac{[N_i - \bar{N} M_i(\{\alpha_n\}) R_i (1 + D_i b \delta_i)]^2}{\bar{N} M_i(\{\alpha_n\}) R_i} \\ - \frac{1}{2} \log(\bar{N} M_i(\{\alpha_n\}) R_i). \end{aligned} \quad (\text{A.1})$$

We can simplify the notation by noting that different foreground templates contribute multiplicatively an effective survey response operator. We can therefore collapse all multiplicative foreground contributions, except the one currently under consideration, into the effective survey response operator given as

$$\tilde{R}_i = \bar{N} R_i \prod_{n \neq k} (1 - \alpha_n F_{n,i}), \quad (\text{A.2})$$

where k labels the currently considered foreground template, and we have used Eq. (4) to factorize foreground contributions. We furthermore introduce the vector

$$A_i = \tilde{R}_i (1 + D_i b \delta_i). \quad (\text{A.3})$$

With this notation, the conditional posterior distribution for α_k simplifies to

$$\begin{aligned} \log \mathcal{P}(\alpha_k | \{\alpha_n\} \setminus \alpha_k, \{N_i\}, \{\delta_i\}, \{\bar{N}\}) = \\ - \frac{1}{2} \sum_{i=0}^v \frac{[N_i - (1 - \alpha_k F_{k,i}) A_i]^2}{(1 - \alpha_k F_{k,i}) \tilde{R}_i} - \frac{1}{2} \log[(1 - \alpha_k F_{k,i}) \tilde{R}_i]. \end{aligned} \quad (\text{A.4})$$

The expression can be further compressed by introducing the following indexed quantities:

$$B_i = N_i - A_i, \quad (\text{A.5})$$

$$C_{k,i} = A_i F_{k,i}, \text{ and} \quad (\text{A.6})$$

$$\gamma_{k,i} = F_{k,i} \tilde{R}_i. \quad (\text{A.7})$$

Consequently, the conditional posterior distribution can be expressed as

$$\begin{aligned} \log(\mathcal{P}(\alpha_k | \{\alpha_n\} \setminus \alpha_k, \{N_i\}, \{\delta_i\}, \{\bar{N}\})) = \\ - \frac{1}{2} \sum_{i=0}^v \frac{(B_i + \alpha_k C_{k,i})^2}{(\tilde{R}_i - \alpha_k \gamma_{k,i})} - \frac{1}{2} \log(\tilde{R}_i - \alpha_k \gamma_{k,i}). \end{aligned} \quad (\text{A.8})$$

In order for Eq. (A.8) to represent a proper probability distribution, the following positivity requirement for the variances needs to hold:

$$\forall i, \tilde{R}_i - \alpha_k \gamma_{k,i} > 0. \quad (\text{A.9})$$

Similarly, it is required that

$$\forall i, \tilde{R}_i > 0, \quad (\text{A.10})$$

which states that the survey response operator should be positive definite. To ensure these requirements, we split the effective survey response operator \tilde{R}_i as follows:

$$\tilde{R}_i = \tilde{R}'_{i,k} + \omega \gamma_{k,i}. \quad (\text{A.11})$$

By also requiring $\tilde{R}'_{i,k} > 0$, we obtain the following requirement for the scalar quantity ω :

$$\tilde{R}'_{i,k} = \tilde{R}_i - \omega \gamma_{k,i} > 0, \text{ which brings us to} \quad (\text{A.12})$$

$$\omega < \frac{\tilde{R}_i}{\gamma_{k,i}} \forall i. \quad (\text{A.13})$$

We therefore choose ω to be

$$\omega = \min_i \left(\frac{\tilde{R}_i}{\gamma_{k,i}} \right). \quad (\text{A.14})$$

With these definitions, we can express the conditional posterior distribution as

$$\begin{aligned} \log(\mathcal{P}(\alpha_k | \{\alpha_n\} \setminus \alpha_k, \{N_i\}, \{\delta_i\}, \{\bar{N}\})) = \\ - \frac{1}{2} \sum_{i=0}^v \frac{(B_i + \alpha_k C_{k,i})^2}{\tilde{R}'_{i,k} + (\omega - \alpha_k) \gamma_{k,i}} \\ - \frac{1}{2} \log(\tilde{R}'_{i,k} + (\omega - \alpha_k) \gamma_{k,i}). \end{aligned} \quad (\text{A.15})$$

We note that this distribution can be conveniently described as the marginalization over a set of auxiliary fields $\{t_i\}$:

$$\begin{aligned} \mathcal{P}(\alpha_k | \{\alpha_n\} \setminus \alpha_k, \{N_i\}, \{\delta_i\}, \{\bar{N}\}) \propto \\ \prod_i \int_{-\infty}^{\infty} dt_i \frac{e^{-\frac{1}{2} \sum_{i=0}^v \frac{(B_i + \alpha_k C_{k,i} - t_i)^2}{(\omega - \alpha_k) \gamma_{k,i}}} e^{-\frac{1}{2} \frac{t_i^2}{\tilde{R}'_{i,k}}}}{\sqrt{2\pi (\omega - \alpha_k) \gamma_{k,i}} \sqrt{2\pi \tilde{R}'_{i,k}}}. \end{aligned} \quad (\text{A.16})$$

This approach therefore follows a similar line of reasoning as discussed in our previous work when presenting a messenger field Gibbs sampler (Jasche & Lavaux 2015). Here we propose to jointly sample the template coefficient parameter α_k with the auxiliary messenger field t_i through a two-step block-sampling procedure. First we generate realizations of the messenger field t_i conditional on the current value of α_k , and then we draw a new value of α_k conditional on the given realization of t_i . We loop 10 times over this small block to ensure a minimal mixing of the variables. To simplify the notation, we introduce the following change of variable:

$$\xi = \omega - \alpha_k. \quad (\text{A.17})$$

This yields the joint distribution:

$$\begin{aligned} \mathcal{P}(\xi, \{t_i\} | \{\alpha_n\} \setminus \alpha_k, \{N_i\}, \{\delta_i\}, \{\bar{N}\}) \propto \\ \prod_i \frac{e^{-\frac{1}{2} \sum_{i=0}^v \frac{(B_i + (\xi - \omega) C_{k,i} - t_i)^2}{\xi \gamma_{k,i}}} e^{-\frac{1}{2} \frac{t_i^2}{\tilde{R}'_{i,k}}}}{\sqrt{2\pi \xi \gamma_{k,i}} \sqrt{2\pi \tilde{R}'_{i,k}}}. \end{aligned} \quad (\text{A.18})$$

As can be easily confirmed, sampling the messenger field t_i amounts to generating normal random variates with the following means and variances:

$$\mu_{i,k} = \frac{\tilde{R}'_{i,k}}{\xi\gamma_{k,i} + \tilde{R}'_{i,k}} (B_i + (\xi - \omega) C_{k,i}) \quad (\text{A.19})$$

and

$$\sigma_{i,k} = \frac{\tilde{R}'_{i,k} \xi \gamma_{k,i}}{\xi\gamma_{k,i} + \tilde{R}'_{i,k}}. \quad (\text{A.20})$$

To generate realizations of the ξ values, we introduce the following quantities:

$$w_k = \sum_i \frac{(B_i - C_{k,i} \omega - t_i)^2}{\gamma_{k,i}} = \sum_i w_{k,i}(t_i) \quad (\text{A.21})$$

and

$$z_k = \sum_i \frac{(C_{k,i})^2}{\gamma_{k,i}} = \sum_i z_{k,i}(t_i). \quad (\text{A.22})$$

With this definition the conditional distribution to sample ξ turns into a generalized inverse Gaussian (GIG) distribution given as

$$\mathcal{P}(\xi | \{t_i\}, \{\alpha_n\} \setminus \alpha_k, \{N_i\}, \{\delta_i\}, \{\bar{N}\}) \propto \frac{1}{(\xi)^{\frac{N_v}{2}}} e^{-\frac{1}{2} \left(\frac{w_k}{\xi} + \xi z_k \right)}, \quad (\text{A.23})$$

where N_v is the number of observed grid elements. The GIG distribution can be conveniently sampled with standard approaches as described in the literature (see e.g. [Dagpunar 1988](#)). Finally, to obtain a sample for the foreground coefficient α_k , we invert the transformation in Eq. (A.17):

$$\alpha_k = \omega - \xi. \quad (\text{A.24})$$

The respective realizations of the t_i field are no longer required and are immediately discarded, which amounts to a marginalization over the t_i values. An efficient sampling algorithm that avoids storing the full t_i vector is proposed in Algorithm 1.

Algorithm 1 Algorithm derived in Appendix A to sample the values of α_k , the multiplicative coefficient attached to the k th foreground.

```

1: function SAMPLE_ONE_α (αk,  $\tilde{R}'$ , B, C, γ, ω, k)
2:   ξ ← ω - αk
3:   wk ← 0
4:   zk ← 0
5:   for i = 0 → (LENGTH(B) - 1) do
6:     ti ~ G(μi,k, σi,k2)
7:     wk ← wk + (Bi - Ck,iω - ti)2/γk,i
8:     zk ← zk + (Ck,i)2/γk,i
9:   end for
10:  ξ ~ GIG(wk, zk)
11:  αk ← ω - ξ
12:  return αk
13: end function

```

▶ we loop over all grid elements of B
 ▶ Here we generate a new value for the messenger field t_i . It is a simple normally distributed value centred on $\mu_{i,k}$ with a variance $\sigma_{i,k}^2$.
 ▶ Accumulate w_k and z_k .
 ▶ We sample ξ from a GIG distribution.
 ▶ at that point we have generated a new α_k from the conditional probability distribution in Eq. (A.4).
