Bayesian forward modelling of cosmic shear data

Natalia Porqueres, Alan Heavens, Daniel Mortlock, Guilhem Lavaux

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **502,** 3035–3044 (2021)
Advance Access publication 2021 January 25

doi:10.1093/mnras/stab204

Bayesian forward modelling of cosmic shear data

Natalia Porqueres [1]★ Alan Heavens,[1] Daniel Mortlock[1,2,3] and Guilhem Lavaux [4]

[1]*Imperial Centre for Inference and Cosmology (ICIC) & Astrophysics group, Imperial College, Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK*
[2]*Department of Mathematics, Imperial College London, London, SW7 2AZ, UK*
[3]*Department of Astronomy, Stockholm University, Albanova, SE-10691 Stockholm, Sweden*
[4]*CNRS & Sorbonne Université, UMR7095, Institut d'Astrophysique de Paris, F-75014, Paris, France*

Accepted 2021 January 21. Received 2021 January 20; in original form 2020 November 13

ABSTRACT

We present a Bayesian hierarchical modelling approach to infer the cosmic matter density field, and the lensing and the matter power spectra, from cosmic shear data. This method uses a physical model of cosmic structure formation to infer physically plausible cosmic structures, which accounts for the non-Gaussian features of the gravitationally evolved matter distribution and light-cone effects. We test and validate our framework with realistic simulated shear data, demonstrating that the method recovers the unbiased matter distribution and the correct lensing and matter power spectrum. While the cosmology is fixed in this test, and the method employs a prior power spectrum, we demonstrate that the lensing results are sensitive to the true power spectrum when this differs from the prior. In this case, the density field samples are generated with a power spectrum that deviates from the prior, and the method recovers the true lensing power spectrum. The method also recovers the matter power spectrum across the sky, but as currently implemented, it cannot determine the radial power since isotropy is not imposed. In summary, our method provides physically plausible inference of the dark matter distribution from cosmic shear data, allowing us to extract information beyond the two-point statistics and exploiting the full information content of the cosmological fields.

Key words: gravitational lensing: weak – methods: data analysis – cosmology: large-scale structure of Universe.

1 INTRODUCTION

As light from distant galaxies propagates through the Universe, it is deflected by the gravitational field induced by the large-scale structures. This deflection results in a coherent distortion of observed galaxy images, inducing small changes in the ellipticity of observed galaxies, which is known as cosmic shear. The weak gravitational lensing effect is sensitive to the geometry of the Universe and the growth of cosmic structures, making it a powerful probe to study the matter distribution and the nature of dark matter and dark energy (see e.g. Kilbinger 2015, for a review).

The next-generation surveys like *Euclid* (Euclid Collaboration 2020), *Roman Space Telescope* (Spergel et al. 2015), and the Rubin Observatory (LSST Science Collaboration 2009) will provide unprecedented precision in cosmic shear measurements, performing wide-field cosmic shear surveys and measuring large and small scales. Harvesting the information from these datasets will present a challenge. Many of the current cosmic shear analyses focus on extracting information from the correlation function or the associated power spectrum (Kitching, Heavens & Miller 2011; Heymans et al. 2013; Kitching et al. 2014; Kitching, Heavens & Das 2015; Alsing et al. 2016; Kitching et al. 2016; Hildebrandt et al. 2017; Troxel et al. 2018; Hikage et al. 2019; Taylor et al. 2019). These analyses capture the two-point statistics, but they do not fully capture the non-Gaussian information encoded in the filamentary features of the matter distribution (Bernardeau, van Waerbeke & Mellier 1997; Jain & Seljak 1997; van Waerbeke, Bernardeau & Mellier 1999; Schneider & Lombardi 2003; Takada & Jain 2003; Vafaei et al. 2010; Kayo, Takada & Jain 2013). While some approaches to access the non-Gaussian information are based on measuring high-order correlations (Bernardeau, van Waerbeke & Mellier 2003; Pen et al. 2003; Jarvis, Bernstein & Jain 2004; Semboloni et al. 2011; van Waerbeke et al. 2013; Fu et al. 2014), peak counts (Jain & van Waerbeke 2000; Dietrich & Hartlap 2010; Maturi, Fedeli & Moscardini 2011; Marian et al. 2012; Pires, Leonard & Starck 2012; Cardone et al. 2013; Lin & Kilbinger 2015; Liu et al. 2015b, a; Kacprzak et al. 2016; Petri et al. 2013; Peel et al. 2017) or using machine learning (Gupta et al. 2018), they rely on summary statistics that do not capture all the information and whose distributions are not well known.

Capturing the full information content of the large-scale structure requires a field-based approach to infer the matter distribution from observations. Böhm et al. (2017) presented a maximum likelihood estimator to reconstruct the matter density field from cosmic shear data, assuming a lognormal distribution for the density. The log-normal distribution reproduces the one- and two-point statistics but fails to reproduce higher-order statistics. Alsing et al. (2016), Alsing, Heavens & Jaffe (2017) presented a Bayesian hierarchical inference scheme to jointly infer shear maps and the corresponding power spectra, assuming Gaussian statistics of the shear field. From a Bayesian perspective, assuming a Gaussian distribution for the shear field is a well motivated prior since it constitutes the maximum

★ E-mail: n.porqueres@imperial.ac.uk

entropy prior once the mean and covariance are specified. However, more information coming from physics is available, and the Gaussian assumption is suboptimal. In this work, we address this limitation by including a gravity model in the Bayesian hierarchical model. For this, we build on the Bayesian Origin Reconstruction from Galaxies (BORG, Jasche & Kitaura 2010a; Jasche & Wandelt 2013a; Lavaux, Jasche & Leclercq 2019) framework, which employs a physical description of the dark matter dynamics and allows us to sample from the initial conditions, which are accurately described by Gaussian statistics. With this more complex data model, we get a better representation of the data, and we can extract information beyond the two-point statistics, exploiting the full information content of the shear fields.

One of the main challenges in the analysis of cosmic shear based on estimating the power spectrum is accounting for the masked regions within the survey area (see, e.g. Chon et al. 2004; Brown, Castro & Taylor 2005; Smith 2006). Our forward modelling approach circumvents these difficulties associated with the survey mask. Although the data do not provide information about the fields in the masked regions, the dynamical model still provides probabilistic information about the shear and density fields that are physically possible in those regions. In our method, the masked regions are treated as pixels with infinite noise, circumventing the need to treat unobserved areas as being cut from the analysis.

The paper is organised as follows. Section 2 describes the data model for the cosmic shear and the likelihood. Section 3 gives an overview of the Bayesian inference framework, BORG, as required for this work. In Section 4, we described the simulated data employed in testing and validating the method. The results are presented in Section 5, showing that the method provides unbiased matter density fields. In Section 7, we discuss the effect of the prior power spectrum in the results. Finally, Section 8 summarises the results.

2 THE DATA MODEL

The effect of weak gravitational lensing on a source can be described by two sky fields: the spin-2 shear, γ, which describes the distortion in the shape of the image, and the scalar convergence field, κ, which describes the change in angular size. These two fields are related to the lensing potential, ϕ, by

$$\kappa = \tfrac{1}{2}\partial\bar{\partial}\phi, \tag{1}$$

$$\gamma = \gamma_1 + i\gamma_2 = \tfrac{1}{2}\partial\partial\phi, \tag{2}$$

(see e.g. Kilbinger 2015), where γ_1 and γ_2 are the components of the shear distortion parallel and at $\pi/4$ to the coordinate axes, and $\partial = \partial_x + i\partial_y$ is the complex derivative on the sky, assuming the flat-sky approximation.

To connect the shear fields to the 3D dark matter distribution, we implemented a line-of-sight integration using the Born approximation, integrating along unperturbed paths. First, we generate convergence fields by integrating along the line-of-sight with the corresponding lensing weights as

$$\kappa(\boldsymbol{\theta}) = \frac{3H_0^2\Omega_m}{2c^2}\int_0^{r_{\rm lim}} \frac{r\,dr}{a(r)} q(r)\delta^f(r\boldsymbol{\theta}, r), \tag{3}$$

where $\boldsymbol{\theta}$ is the position on the sky, r is the comoving distance, $r_{\rm lim}$ is the limiting comoving distance, δ^f is the final density field and

$$q(r) = \int_r^{r_{\rm lim}} dr' n(r') \frac{r' - r}{r'}, \tag{4}$$

where $n(r)$ is the source galaxy distribution. In our discrete implementation, this becomes

$$\kappa^b_{mn} = \frac{3H_0^2\Omega_m}{2c^2}\sum_{j=0}^{N_2}\delta^f_{mnj}\left[\sum_{s=j}^{N_2}\frac{(r_s - r_j)}{r_s}n^b(r_s)\Delta r_s\right]\frac{r_j\Delta r_j}{a_j}. \tag{5}$$

The index b labels the tomographic bin and the subindices mn indicate the 2D pixel on the sky, whose size is chosen to include typically many sources. The sum index j indicates the slice in the radial direction, at a comoving distance r_j. N_2 is the total number of voxels along the radial axis. The voxels have a length of Δr_j. δ_f is the 3D dark matter overdensity at a scale factor a. The comoving radial distance r_s indicates the distance to the source plane. The redshift distribution of sources for each tomographic bin is given by $n^b(z_s)$. In this initial proof-of-concept work, we focus on testing the inference and investigating the extent to which the 3D density field, and the 3D matter power spectrum, can be inferred from 2D shear maps. For these tests, we used a simplified scenario, assuming flat-sky and distant observer approximations. In future work, we will drop these approximations and consider the projection effects.

In the flat-sky approximation, we can obtain the shear values from the convergence field. On a flat-sky, the shear and the convergence are related in Fourier space. We, therefore, use a discrete Fourier transform (DFT) to obtain the shear values as

$$\gamma^b_{mn} = \mathrm{DFT}^{-1}\left[\frac{(l_x + il_y)^2}{l_x^2 + l_y^2}\mathrm{DFT}\left(\kappa^b_{mn}\right)\right], \tag{6}$$

where $\vec{l} = (l_x, l_y)$ is the wave-vector written as a complex quantity. Since the convergence is also computed as part of the hierarchical model, this method has the advantage that it can analyse reduced shear,

$$g = \frac{\gamma}{1 - \kappa}, \tag{7}$$

which, rather than the shear alone, controls the shape distortion.

To analyse cosmic shear observations in our Bayesian framework, we now built a likelihood based on this data model. We assume Gaussian pixel noise for the shear, corresponding to a negative log-likelihood, $\mathcal{L} = -\log P(\hat{\gamma}|\delta^f)$, that can then be written as

$$\mathcal{L} = \sum_b\sum_{mn}\frac{\left[\hat{\gamma}^b_{1,mn} - \gamma^b_{1,mn}(\delta^f)\right]^2 + \left[\hat{\gamma}^b_{2,mn} - \gamma^b_{2,mn}(\delta^f)\right]^2}{2\sigma_b^2}, \tag{8}$$

where $\hat{\gamma} = \hat{\gamma}_1 + i\hat{\gamma}_2$ is the observed data. This is an estimate of the shear in the pixel, with a variance σ_b^2, which is determined from the shape noise and number of sources per pixel as $\sigma_\epsilon^2/N_{\rm sources}$. We note that even if the ellipticity distribution is not Gaussian, provided many sources contribute to each pixel average the noise will become Gaussian according to the central limit theorem. An alternative would be to sample from the distribution in another level of the hierarchy, but this would be expensive, so we simplify this stage by using summary statistics of estimated shear and their variance.

This likelihood is then implemented into the large-scale structure sampler of the BORG framework. The corresponding physical forward modelling approach is illustrated in Fig. 1 and proceeds as follows. Using realisations of the three-dimensional field of primordial fluctuations, the dynamical structure formation model evaluates nonlinear realisations of the dark matter distribution, accounting for the light-cone effects inherent to deep observations. Using these 3D dark matter field realisations and the data model, BORG predicts shear fields that are compared to the observed data via the likelihood in equation (8).

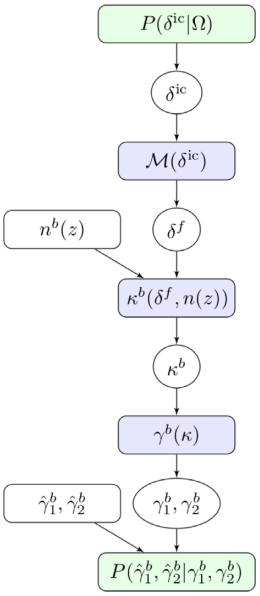

Figure 1. Hierarchical representation of the BORG inference framework for the analysis of cosmic shear data. Primordial fluctuations $\delta_{\rm ic}$ encoded in a a set of Fourier modes at $z \approx 1000$ are obtained from the prior $P(\delta_{\rm ic}|\Omega)$, where Ω represents the cosmological parameters. These initial conditions are evolved using the gravity model $\mathcal{M}(\delta^{\rm ic})$, which provides the evolved density $\delta_{\rm f}$. The evolved density and the redshift distribution of sources $n^b(z)$ are then used to compute the convergence field for each tomographic bin b, $\kappa^b(\delta^f, n^b(z))$. From the convergence, we compute the cosmic shear γ_1^b, γ_2^b in the flat-sky approximation. $\hat{\gamma}_1^b$, $\hat{\gamma}_2^b$ indicate the observational data. Purple boxes indicate deterministic transition while green boxes are probability distributions.

3 THE BORG FRAMEWORK

This work extends the previously developed BORG algorithm to analyse the spatial matter distribution underlying cosmic shear observations. In this section, we provide a summary of the algorithm. A more detailed description of the BORG framework can be found in Jasche & Wandelt (2013a), Jasche, Leclercq & Wandelt (2015), Lavaux & Jasche (2016), Jasche & Lavaux (2019), Lavaux et al. (2019).

The BORG framework is a Bayesian inference method aiming at inferring the nonlinear spatial dark matter distribution and its dynamics from cosmological datasets. The underlying idea is to fit full dynamical gravitational and structure formation models to observations. By using nonlinear structure growth models, the BORG algorithm can exploit the full statistical power of high-order statistics of the matter distribution imprinted by gravitational clustering. This dynamical model links the primordial density fluctuations to the present large-scale structures. Therefore, the forward modelling approach allows the translation of the problem of inferring nonlinear matter density fields into the inference of the spatial distribution of the primordial density fluctuations, which are well described by Gaussian statistics (Planck Collaboration 2019). The BORG algorithm, therefore, infers the initial matter fluctuations, the dark matter distribution, and its dynamical properties from observations.

Motivated by inflation theory and observational data, the BORG algorithm employs a Gaussian prior for the initial density contrast at an initial cosmic scale factor of $a \simeq 10^{-3}$, time for which density perturbations are linearly growing. Initial and evolved density fields are linked by deterministic gravitational evolution mediated by various physics models of structure growth. Specifically, BORG incorporates several physical models based on Lagrangian Perturbation Theory (LPT), fully nonlinear particle-mesh models (Jasche & Lavaux 2019), a model based on spatial COmoving Lagrangian Acceleration framework (Leclercq et al. 2020), and a semiclassical analogue to LPT (Porqueres et al. 2020). Any of these dynamical models can be straightforwardly employed within the flexible block sampling illustrated in Fig. 1. To test the inference method, in this work, we used LPT to approximately describe the gravitational clustering. However, in a future application of the method to real datasets, we will use the fully nonlinear particle-mesh (Jasche & Lavaux 2019) to have a better description of the matter density at small spatial scales, which undergo nonlinear dynamics. Though the particle mesh will be more costly, it will still be tractable. Tassev, Zaldarriaga & Eisenstein (2013) showed that the LPT begins to show significant deviations at $k >$ 0.2 h/Mpc, but using the tCOLA modification of the equation of motion we can push the precision of an LPT-like simulation close to a full N-body simulation in a few time-steps, at the field level. Typically, it can be reached in at least as little as ten time-steps to reach 90 per cent correlation at $k = 1h\text{Mpc}^{-1}$ with a full N-body simulation such as one provided by Gadget-2 (Springel 2005).

At its core, the BORG framework employs MCMC techniques. This method allows inference of the full posterior distribution from which we can quantify the uncertainties in our results. However, the inference of the density field typically involves $\mathcal{O}(10^7)$ free parameters, corresponding to the discretised volume elements of the initial conditions. To explore this high-dimensional parameter-space efficiently, the BORG framework uses a Hamiltonian Monte Carlo (HMC) method, which exploits the information in the gradients and adapts to the geometry of the problem. We need, therefore, the adjoint gradient of the data model, which transforms the error vector from the likelihood space to the initial conditions. For the case of weak lensing, we derive this gradient in Appendix A. More details about the HMC and its implementation are described in Jasche & Kitaura (2010b) and Jasche & Wandelt (2013b).

4 THE MOCK DATA

To test the inference framework, we generated mock observations of cosmic shear with 30 sources per arcmin^2 as expected for the Euclid survey, with four tomographic bins and a nontrivial survey mask. In this section, we describe the properties of the synthetic data.

Mock data are constructed by first generating Gaussian initial conditions on a Cartesian grid of size $1h^{-1}$ Gpc $\times$ $1h^{-1}$ Gpc $\times$ $4h^{-1}$ Gpc with 128 $\times$ 128 $\times$ 256 voxels. To generate primordial Gaussian density fluctuations we used a cosmological matter power spectrum including the baryonic wiggles calculated according to the prescription provided by Eisenstein & Hu (1998, 1999). We further assumed a standard ΛCDM cosmology with the following set of parameters: $\Omega_m = 0.31$, $\Omega_\Lambda = 0.69$, $\Omega_b = 0.049$, $h = 0.6711$, $\sigma_8 = 0.8$, $n_s = 0.9624$. Here $\text{H}_0 = 100h$ km s^{-1} Mpc^{-1}.

To generate realisations of the nonlinear density field, we evolve the Gaussian primordial fluctuations via LPT. This involves simulating displacements for $256^2 \times 512$ particles in the LPT simulation, accounting for light-cone effects inherent to deep observations. Final density fields are constructed by estimating densities via the cloud-in-cell scheme from simulated particles on the Cartesian grid. A cosmic shear field is generated by applying the data model described in Section 2, assuming the redshift distributions for tomographic bins shown in Fig. 2. Finally, we added Gaussian pixel-noise to the shear with variance corresponding to 30 sources per arcmin^2, as expected to be obtained from the Euclid survey (Laureijs et al. 2011), and with an error on intrinsic ellipticity given by $\sigma_\epsilon = 0.3$ (Kilbinger 2015). The total of 30 sources per arcmin^2 is then equally distributed

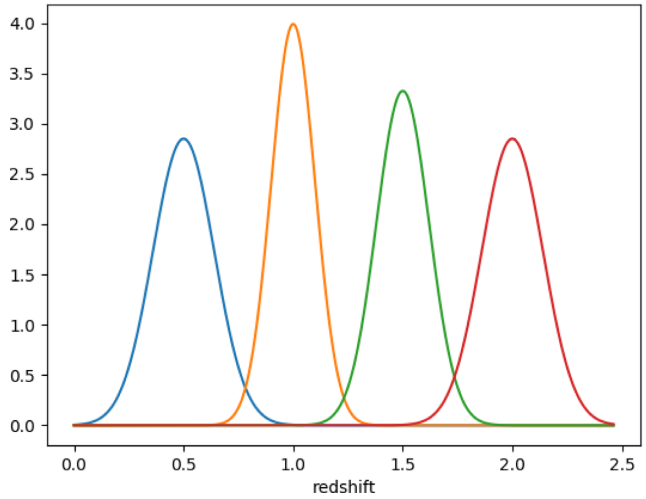

Figure 2. Redshift distributions of sources for the four tomographic bins considered in this analysis.

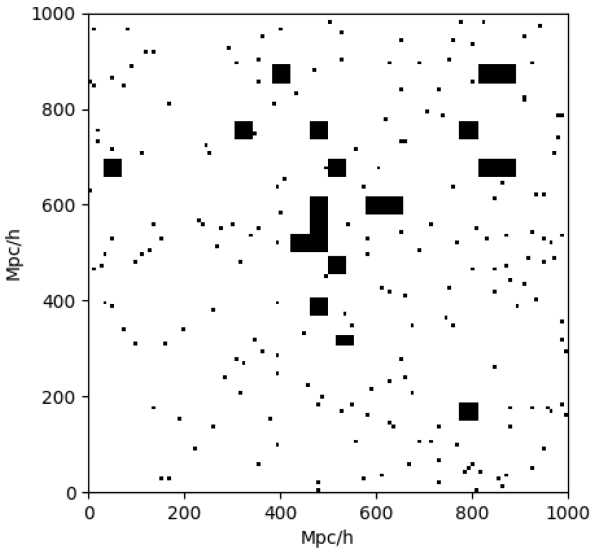

Figure 3. Mask used to generate the mock data. The masked regions, with no contributing sources, are indicated in black. There are two different scales, corresponding to unobserved regions and bright stars.

between the bins, corresponding to 7.5 galaxies per arcmin^2 for each tomographic bin. This corresponds to a signal-to-noise ratio of $S/N = 0.5$. We added a nontrivial survey mask, shown in Fig. 3. Since the data provides no direct information in the masked regions, these are treated as pixels with infinite noise.

5 RESULTS

Here, we present the results of applying our algorithm to simulated cosmic shear data. We show that our method infers unbiased density fields and corresponding power spectra at all scales considered in this work. We also perform a posterior predictive test for the shear, showing that the inferred densities can explain the data within the noise uncertainty.

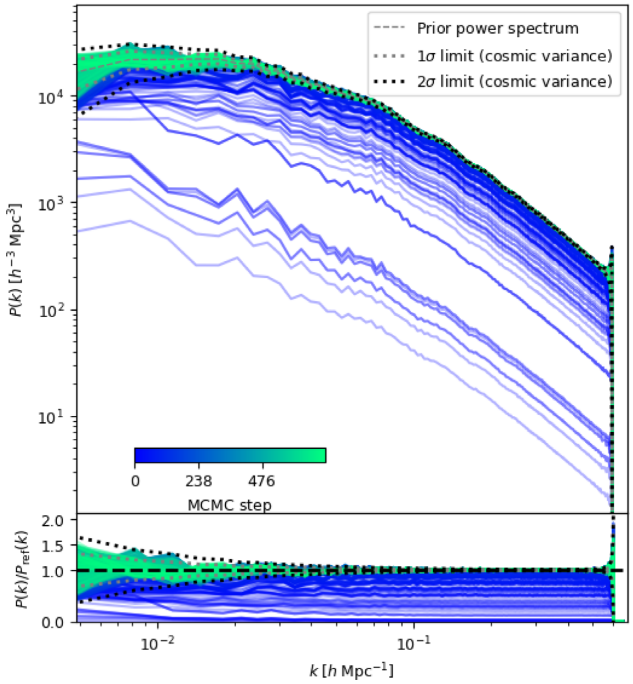

Figure 4. Burn-in of the posterior initial matter power spectra. The colour scale shows the evolution of the matter power spectrum with the number of samples. The dashed lines indicate the underlying power spectrum and the 1- and 2-σ cosmic variance limits. The Markov chain is initialised with a Gaussian initial density field scaled by a factor 10^{-3} and the amplitudes of the power spectrum systematically drift towards the fiducial values, recovering the true matter power spectrum at the end of the warm-up phase.

5.1 The warm-up phase of the sampler

In this first Bayesian approach, we keep the cosmology fixed, and specify a prior on the initial power spectrum. However, the power spectrum of the inferred matter distribution is conditioned by the data, and we can use the posterior $P(k)$ as a diagnostic for the effectiveness of the inference since the power spectrum of the posterior samples may differ from the prior. To monitor the initial warm-up phase of the Markov sampler, we follow a similar approach to our previous works (Jasche & Wandelt 2013a; Jasche & Lavaux 2017; Ramanah et al. 2019; Jasche & Lavaux 2019; Porqueres et al. 2019a, b): we initialised the Markov chain with an overdispersed state and traced the systematic drift of inferred quantities towards their preferred regions in the parameter space. Specifically, we initialised the Markov chain with a random Gaussian initial density field scaled by a factor 10^{-3} and monitored the drift of corresponding posterior power spectra during the warm-up phase. Fig. 4 presents the results of this exercise, showing successive measurements of the posterior power spectrum during the initial warm-up phase. The amplitudes of the posterior power spectrum show a systematic drift towards their fiducial values. By the end of the warm-up phase, the sampler has found an unbiased representation of the initial power spectrum at all Fourier modes considered in this work. Starting the sampler from an over-dispersed state, therefore, provides us with an important diagnostics to test the validity of the sampling algorithm.

5.2 Inferred density fields

As discussed above, our method uses a forward modelling approach, fitting a physical dynamical model to shear data and employs an

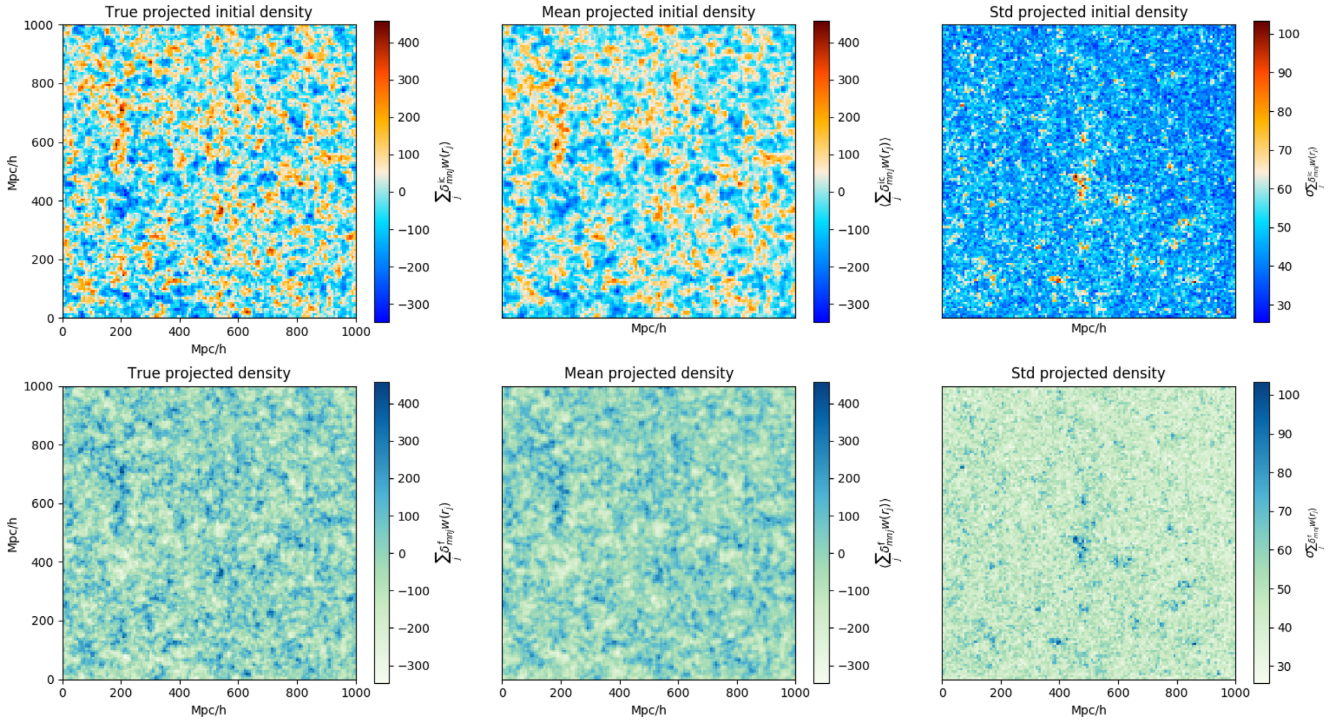

Figure 5. Projections of the ground truth initial (left-hand upper panel), final density field (left-hand lower panel), inferred ensemble mean initial (middle upper panel) and ensemble mean final (middle-lower panel) density field computed from 500 MCMC samples. Since the information on the radial direction is not very informative, the density fields are projected on the sky, and the different slices of the 3D density field are weighted with the distribution of sources. Comparison between these panels shows that the method recovers the structure of the true projected density field with high accuracy. Right-hand panels show standard deviations of inferred amplitudes of the initial (upper right-hand panel) and final density fields (lower right-hand panel). The regions of high uncertainty correspond to the masked regions, where there are no contributing sources.

MCMC sampler to explore the parameter space. This provides the full posterior distribution, from which we draw samples of the initial matter fluctuations.

Fig. 5 shows projections of the true fields, and the ensemble mean and variances of inferred three-dimensional fields. The mean and variance are estimated from 500 samples of the posterior distribution (the correlation length is ≈80 samples). To compare the ground truth to the inferred mean density field, we computed the projection of the density fields on the sky since the radial information is not very constraining. In this projection, the different slices of the 3D density field are weighted with the distribution of the lensing sources. A first visual comparison between ground truth and the inferred ensemble mean initial and final density fields shows that the algorithm correctly recovered the large-scale structures from cosmic shear data. As expected, the mean of the initial density samples exhibits a small degree of smoothing, a feature that is known from the Wiener filtering solution for Gaussian fields and Gaussian prior.

The right-hand panels of Fig. 5 show the corresponding standard deviations of the projected densities, which are estimated from the posterior samples. The high uncertainty regions correspond to the masked areas, where there are no contributing sources. While the data do not provide direct information about the density field in these masked regions, the dynamical model still provides probabilistic information about the density fields that are physically plausible in those regions. These results indicate that the method can deal with nontrivial survey masks, and account for the uncertainty in the unobserved areas. The standard deviation of the initial conditions is homogeneous, apart from mask effect, indicating that the dynamical model correctly propagates the information between the primordial matter fluctuations and the final density field.

5.3 Posterior predictive tests

Posterior predictions allow testing whether the inferred density fields provide accurate explanations of the data (see, e.g. Gelman et al. 2004). Generally, posterior predictive tests provide good diagnostics about the adequacy of data models in explaining observations and identifying possible systematic problems with the inference. In this section, we predicted the shear and convergence fields as the average computed from 500 posterior samples.

Fig. 6 presents the result of this test for one tomographic bin, showing that the posterior predicted shear and convergence recover the features of the true fields. The masked regions show higher standard deviation, indicating that the method can account for the uncertainties in the unobserved areas and provide probabilistic information of the physically plausible shear in those regions. While Fig. 6 shows a visual comparison, Fig. 7 shows the residuals between the true and the mean posterior-predicted shear fields. The green line in the plot indicates the noise distribution, showing that the distribution of the residuals is narrower than the noise distribution, and, therefore, the inferred quantities can explain the data at subnoise level, with the additional constraints coming from the cosmological prior.

Fig. 8 shows the power spectrum of the posterior-predicted shear, measured as the average of predicted shear fields from 500 randomly-drawn samples. The predicted power spectra match the true shear power spectrum at all scales. We note that we computed the true power spectrum from the noiseless data as the posterior-predicted shear does not contain noise. These shear power spectra are pure *E*-mode since they are obtained from the lensing equations under the Born approximation.

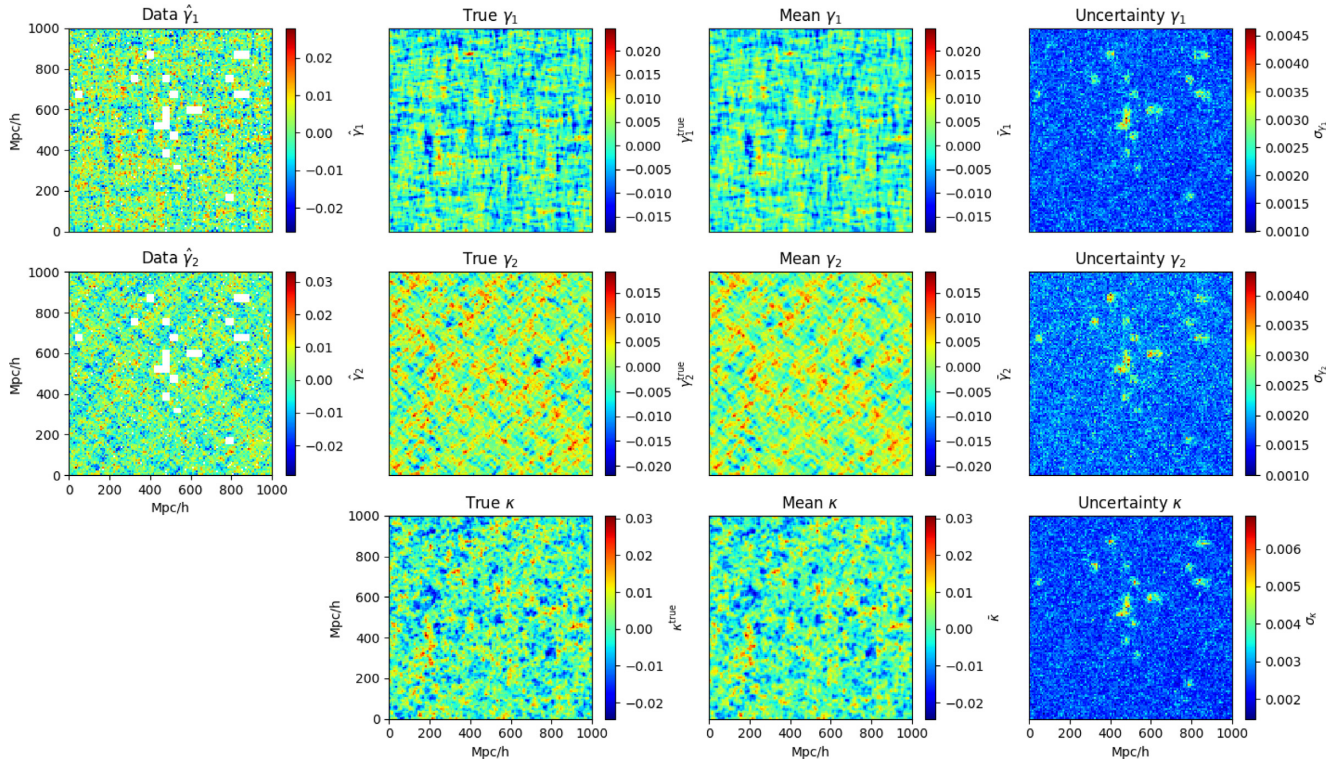

Figure 6. Posterior predicted shear and convergence for one tomographic bin. The left-hand column shows the shear data, including noise and masked regions; the second column shows the true shear and convergence fields, and third and fourth columns show the mean and standard deviation of the posterior-predicted shear and convergence, computed from 500 posterior samples. The method recovers the true cosmic shear correctly. The regions with higher standard deviation correspond to the masked regions.

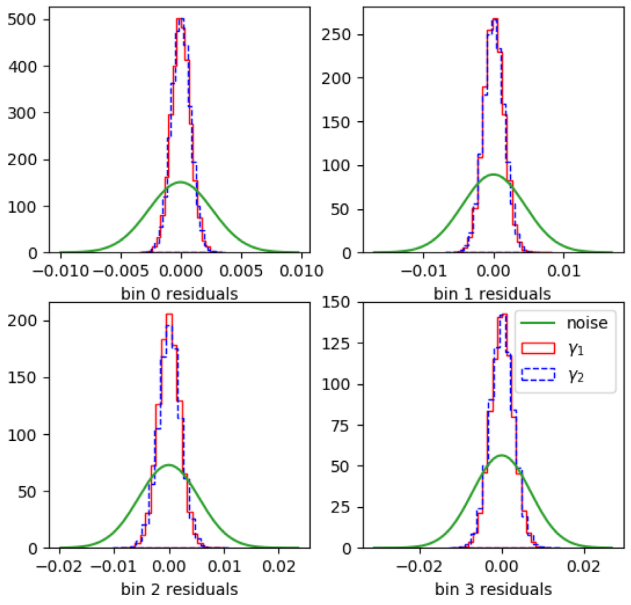

Figure 7. Histogram of the residuals computed as the difference between the posterior predicted shear and the true shear. We note that the true shear does not include the noise. The residuals distribution is narrower than the distribution of pixel noise in the data, indicated in green, showing that the method recovers the true shear at subnoise level with additional constraints from the cosmological prior.

Fig. 9 shows the correlation coefficient between the posterior-predicted shear maps and the true shear. This is readily understood, as it is similar to the Wiener solution for the posterior of a statistically homogeneous Gaussian field with signal power S and Gaussian noise power N, (see e.g Jeffrey, Heavens & Fortio 2018). In this case the mean posterior has power suppressed by $(S^{-1} + N^{-1})N^{-1}$ and this is the correlation coefficient. As a result, we expect the tomographic bin centred at the lowest redshift (bin 0) to have a lower correlation coefficient because the lensing power at low redshift is smaller.

6 DISTRIBUTION OF THE CONVERGENCE FIELD

Previous Bayesian approaches (Alsing et al. 2016) rely on a Gaussian prior for the shear data. The Gaussian prior is well justified when only the mean and variance are known since it is the least informative prior. It is important to understand that the samples are not Gaussian fields, since they are conditioned on the data, so non-gaussianity in the field may be imposed by the data. However, we can make use of the fact that we have more information available from knowledge of gravitational physics and, for this reason, we include a gravity model in our Bayesian hierarchical model. The advantage here is that we sample from the initial field, which we know to be Gaussian, so rather than relying on an uninformative prior for the final shear fields, we use the correct Gaussian distribution for the initial conditions. What we do not yet do in this model is to vary the prior parameters of the power spectrum (as Alsing et al. (2016) do), and this will be the subject of future work. As described in Section 2, we obtained the

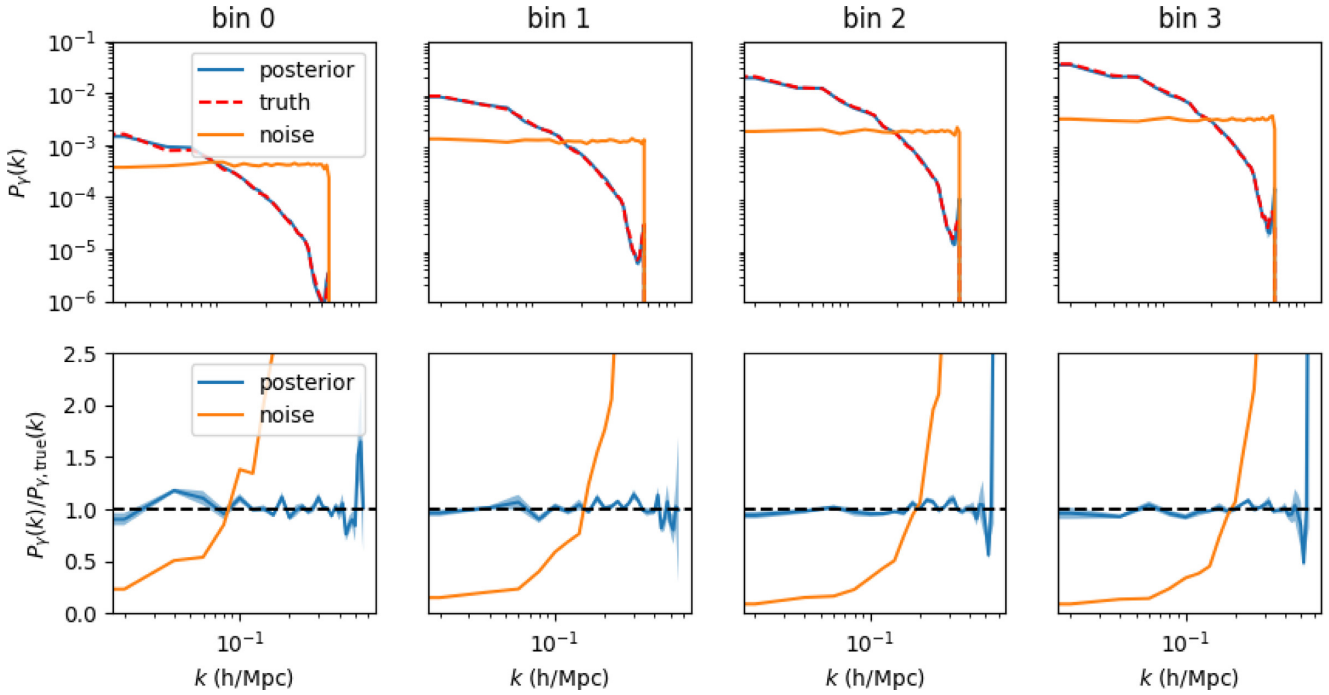

Figure 8. Posterior power spectrum of the shear field compared to the power spectrum of the true shear for each tomographic bin. The posterior power spectrum is the averaged of the power spectrum measured in 500 posterior samples. The orange line shows the noise power spectrum. The bottom plots show the ratio between the posterior and the true power spectrum, showing that the method recovers the true shear power spectrum at all scales.

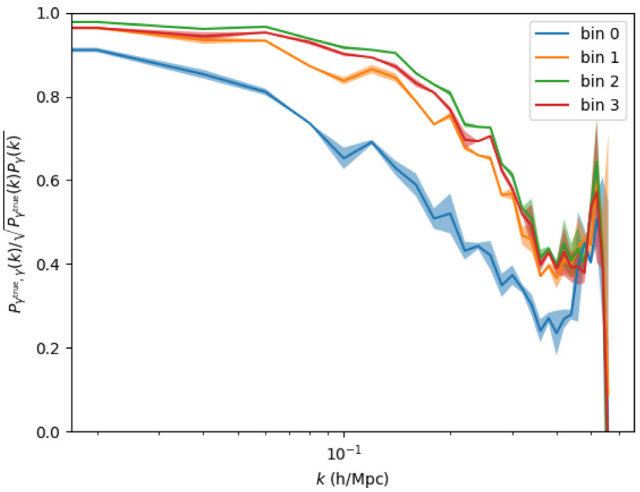

Figure 9. Correlation coefficient between the posterior-predicted and true shear fields, from 500 posterior samples. This shows the expected Wiener-filter-like suppression of power where the noise is high.

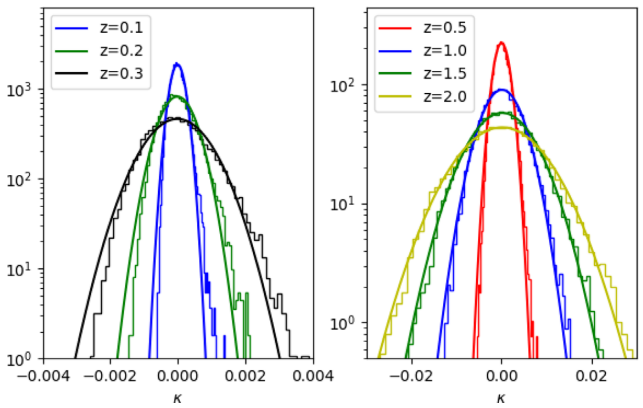

Figure 10. Distribution of the convergence field for tomographic bins centred at different redshifts. For comparison, a Gaussian with the same mean and variance is plotted on top. The convergence distribution is skewed for tomographic bins centered at $z > 0.5$.

shear on a flat-sky from the convergence, which is computed from the nonlinear density field. Using our forward model, we have computed the convergence field in a wider range of tomographic bins, centred at different redshifts but with the same bin width ($\sigma = 0.1$), to illustrate more clearly how non-Gaussianity in the 1-point distribution evolves. Fig. 10 shows the distribution of these convergence fields. While the convergence shows a Gaussian distribution for tomographic bins centred at $z > 0.5$, it is skewed for tomographic bins at lower redshifts. This indicates that the Gaussian approximation is accurate at large redshift, but it is suboptimal for low-redshift bins.

7 PRIOR TEST

As discussed in Section 5.1, in this approach, we keep the cosmology fixed, and specify a prior on the initial power spectrum. Although our method does not sample the power spectrum, the power spectrum of the posterior samples is conditioned by the data. This means that, if the data require it, the posterior power spectrum deviates from the prior, and the density samples have a power spectrum that differs from the prior. To demonstrate that, we tested our method with mock data generated with $h = 0.6$, $\sigma_8 = 0.55$, $\Omega_m = 0.7$ and analyse these data with $h = 0.677$, $\sigma_8 = 0.8$, $\Omega_m = 0.3$, such that $\Omega_m h^2$ changes by a factor 2. These changes in the cosmological parameters give a different shape and amplitude of the power spectrum, as can be seen in Fig. 11. The resolution, total number of sources per

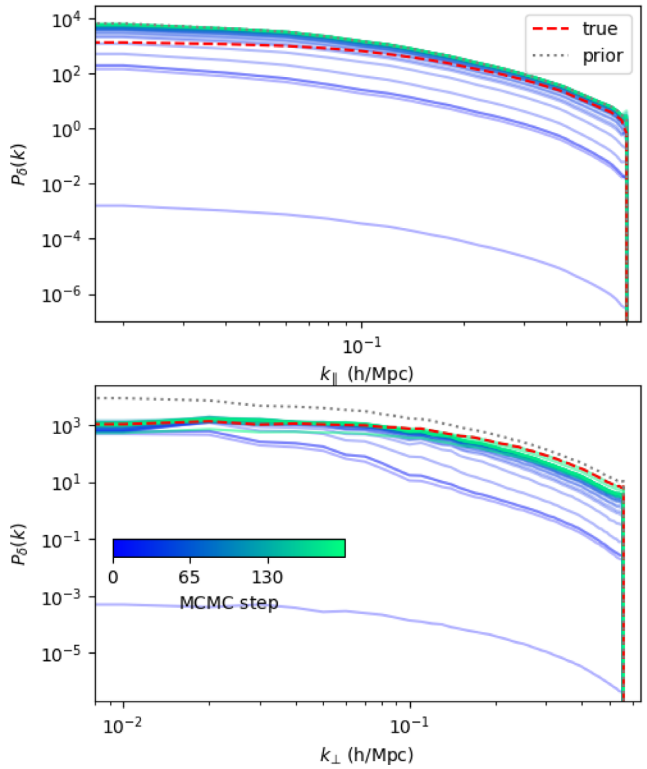

Figure 11. Burn-in phase of the dark matter power spectra for the prior test. We show the power spectrum for the modes parallel to the line-of-sight, $k_{||}$ (upper panel), and the perpendicular modes, $k_\perp$ (lower panel), i.e. in the plane of the sky. The dotted line indicates the prior matter power spectrum, and the truth is indicated by the red dashed line. The colour scale indicates the sample number in the Markov chain. After the burn-in phase, the method recovers the correct matter power on the plane of the sky. However, the data is not constraining in the radial direction and the prior dominates. We note that as currently implemented, isotropy is not a requirement.

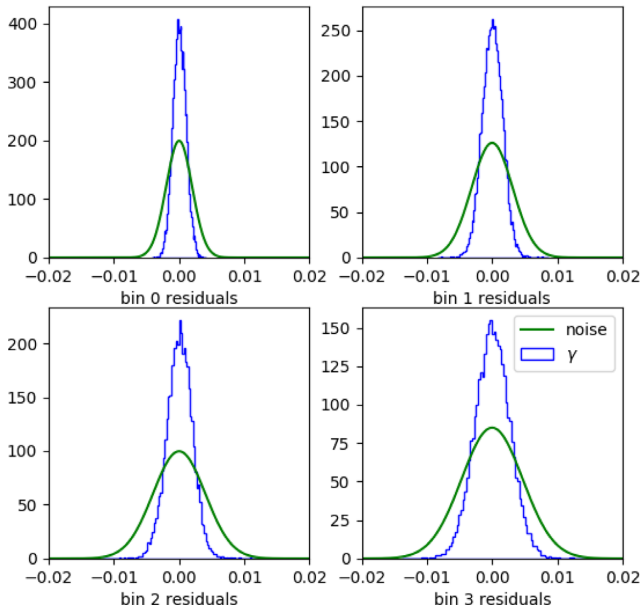

Figure 12. Histogram of the shear residuals for the prior test. The residuals are computed as the difference between the mean posterior-predicted shear and true shear. The distribution of shear residuals is narrower than the noise distribution, indicated in green.

square arcmin, the mask and tomographic bins remain as described in Section 4.

Fig. 11 presents the evolution of the power spectrum for this test, showing the burn-in phase of the Markov chain. In the plane of the sky, the posterior power spectrum agrees with the truth, demonstrating that the method is sensitive to the true power spectrum even when this differs from the prior. However, in the line-of-sight direction, the data is not constraining, and the prior dominates. These results are consistent with Simon, Taylor & Hartlap (2009). As currently implemented, isotropy is not required, and the method cannot determine the radial power accurately. The recovered matter power spectrum, however, suffices to explain the data, obtaining shear residuals below the noise level, as can be seen in Fig. 12. Fig. 13 shows the lensing power spectrum, indicating that the method recovers the correct lensing power at large scales where the noise is low. However, the small scales have lower S/N, and the posterior power spectrum drifts towards the prior. To avoid this prior sensitivity, we need to sample from the power spectrum as well as the field. This will be the subject of a future paper.

In future work, we will extend our Bayesian hierarchical model to jointly sample the cosmological parameters and the density field, using an approach compatible with the one presented in Ramanah et al. (2019). We expect that this future extension of the method will constrain the matter power spectrum, also in the radial direction, through the assumption of isotropy, but we do not anticipate being able to recover the small-scale radial distribution at the field level, because of the width of the lensing kernel and the distance uncertainties. Meanwhile, the test presented here shows that the posterior power spectrum is conditioned by the data despite using a fixed cosmology.

8 SUMMARY AND DISCUSSION

We have developed a Bayesian physical forward model to infer the matter density field and primordial fluctuations from cosmic shear data. This framework consists of a Gaussian prior for the primordial fluctuations, a dynamical structure formation model that links the initial conditions and the evolved density field, and a likelihood based on a data model of the cosmic shear in the flat-sky approximation.

This forward modelling approach allows us to go beyond the common analyses of cosmic shear based on two-point statistics. While many studies of the cosmic shear focus on the power spectrum or the correlation function, the nonlinear dynamics of the large-scale structure encode significant information in higher-order statistics associated with the filamentary structure of the cosmic web. Our dynamical forward model reproduces the filamentary matter distribution and, in this way, allows using every data point, rather than relying on summary statistics that do not capture all the information and whose distributions are not well known. By employing a more accurate gravity model, our method also improves over previous Bayesian hierarchical approaches that assumed a Gaussian distribution of the shear field (Alsing et al. 2016).

We have tested our inference method with simulated data with four tomographic bins, a survey mask, and 30 sources per arcmin^2 as expected for the Euclid survey. These tests demonstrate that our method recovers the unbiased matter distribution and initial matter power spectrum from cosmic shear data. Posterior predictive tests showed that the inferred quantities are known to the sub-noise

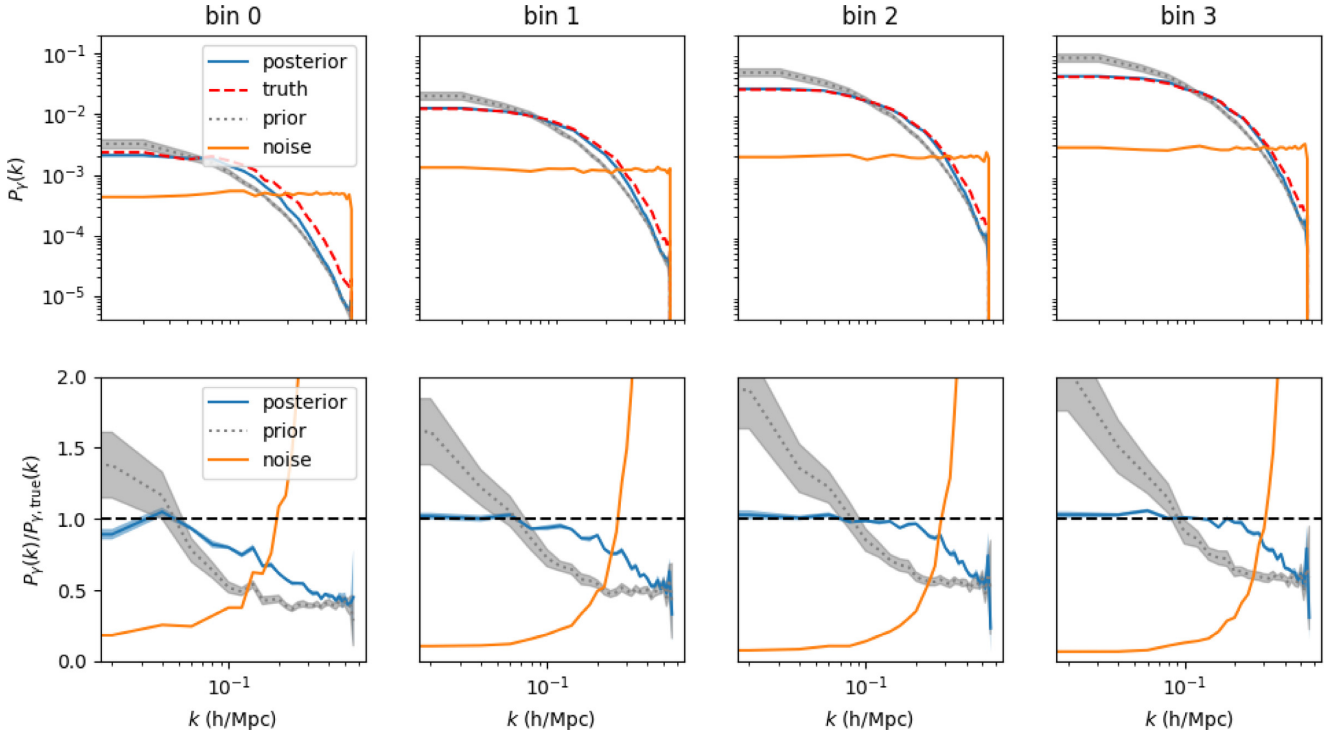

Figure 13. Posterior power spectrum of the shear field compared to the power spectrum of the true shear for the prior test. The grey dotted line indicates the shear power spectrum computed using the prior cosmology, which differs from the truth. The orange line indicates the power spectrum of the noise. The bottom plots show the ratio between the posterior, prior, and noise and the true power spectrum, showing that the method recovers the true lensing power spectrum where the signal-to-noise is high, but is suppressed in the low S/N regions where the prior is low.

level, with additional constraints coming from the cosmological prior.

Although our framework currently uses a fixed cosmology, we have shown that the method recovers the true power spectrum when this differs from the prior where the signal-to-noise is high. While we do not sample the power spectrum, the posterior power spectrum deviates from the prior if the data require it. To illustrate this, we performed a test using different values of H_0, σ_8 and Ω_m to generate the mock data and to analyse them. In this case, the prior power spectrum, therefore, differs from the true power spectrum. This test demonstrated that our method is sensitive to the underlying cosmology, and the power spectrum of the density samples is conditioned by the data, recovering the true matter power spectrum across the sky and the lensing power spectrum. However, in the radial direction, the data is not informative, and the prior dominates since we have not imposed isotropy. In future work, we will extend our Bayesian hierarchical approach to sample the cosmological parameters, through both the geometry and the power spectrum and its growth. We expect that this extension will also constrain the matter power spectrum in the radial direction through the imposition of isotropy, and remove the prior power spectrum sensitivity.

To summarise, this work demonstrates the feasibility of detailed and physically plausible inference of the large-scale structure from cosmic shear data. The proposed approach, therefore, improves the shear data model from previous methods by including a physical description of gravity, providing a better representation of the data and allowing us to extract information beyond the two-point statistics. In future work, we will explore the constraints on cosmology that this approach provides by jointly sampling the initial conditions and the cosmological parameters.

ACKNOWLEDGEMENTS

This work was supported by STFC through Imperial College Astrophysics Consolidated Grant ST/5000372/1. GL acknowledges financial support from the ANR BIG4, under reference ANR-16-CE23-0002. This work was carried out within the Aquila Consortium.[1]

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Alsing J., Heavens A., Jaffe A. H., Kiessling A., Wandelt B., Hoffmann T., 2016, MNRAS, 455, 4452
Alsing J., Heavens A., Jaffe A. H., 2017, MNRAS, 466, 3272
Bernardeau F., van Waerbeke L., Mellier Y., 1997, A&A, 322, 1
Bernardeau F., van Waerbeke L., Mellier Y., 2003, A&A, 397, 405
Böhm V., Hilbert S., Greiner M., Enßlin T. A., 2017, Phys. Rev. D, 96, 123510
Brown M. L., Castro P. G., Taylor A. N., 2005, MNRAS, 360, 1262
Cardone V. F., Camera S., Mainini R., Romano A., Diaferio A., Maoli R., Scaramella R., 2013, MNRAS, 430, 2896
Chon G., Challinor A., Prunet S., Hivon E., Szapudi I., 2004, MNRAS, 350, 914
Dietrich J. P., Hartlap J., 2010, MNRAS, 402, 1049
Eisenstein D. J., Hu W., 1998, ApJ, 496, 605
Eisenstein D. J., Hu W., 1999, ApJ, 511, 5
Euclid Collaboration, 2020, A&A, 635, A139

[1] https://aquila-consortium.org

Fu L. et al., 2014, MNRAS, 441, 2725
Gelman A., Carlin J. B., Stern H. S., Rubin D. B., 2004, Bayesian Data Analysis, 2nd edn, Chapman and Hall/CRC
Gupta A., Matilla J. M. Z., Hsu D., Haiman Z., 2018, Phys. Rev. D, 97, 103515
Heymans C. et al., 2013, MNRAS, 432, 2433
Hikage C. et al., 2019, PASJ, 71, 43
Hildebrandt H. et al., 2017, MNRAS, 465, 1454
Jain B., Seljak U., 1997, ApJ, 484, 560
Jain B., van Waerbeke L., 2000, ApJ, 530, L1
Jarvis M., Bernstein G., Jain B., 2004, MNRAS, 352, 338
Jasche J., Kitaura F. S., 2010a, MNRAS, 407, 29
Jasche J., Kitaura F. S., 2010b, MNRAS, 407, 29
Jasche J., Lavaux G., 2017, A&A, 606, A37
Jasche J., Lavaux G., 2019, A&A, 625, A64
Jasche J., Wandelt B. D., 2013a, MNRAS, 432, 894
Jasche J., Wandelt B. D., 2013b, MNRAS, 432, 894
Jasche J., Leclercq F., Wandelt B. D., 2015, J. Cosmol. Astropart. Phys., 1, 036
Jeffrey N., Heavens A. F., Fortio P. D., 2018, Astron. Comput., 25, 230
Kacprzak T. et al., 2016, MNRAS, 463, 3653
Kayo I., Takada M., Jain B., 2013, MNRAS, 429, 344
Kilbinger M., 2015, Reports on Progress in Physics, 78, 086901
Kitching T. D. et al., 2014, MNRAS, 442, 1326
Kitching T. D., Heavens A. F., Miller L., 2011, MNRAS, 413, 2923
Kitching T. D., Heavens A. F., Das S., 2015, MNRAS, 449, 2205
Kitching T. D., Verde L., Heavens A. F., Jimenez R., 2016, MNRAS, 459, 971
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Lavaux G., Jasche J., 2016, MNRAS, 455, 3169
Lavaux G., Jasche J., Leclercq F., 2019, preprint (arXiv:1909.06396)
Leclercq F., Faure B., Lavaux G., Wandelt B. D., Jaffe A. H., Heavens A. F., Percival W. J., 2020, A&A, 639, A91
Lin C.-A., Kilbinger M., 2015, A&A, 583, A70
Liu X. et al., 2015b, MNRAS, 450, 2888
Liu J., Petri A., Haiman Z., Hui L., Kratochvil J. M., May M., 2015a, Phys. Rev. D, 91, 063507
LSST Science Collaboration, 2009, preprint (arXiv:0912.0201)
Marian L., Smith R. E., Hilbert S., Schneider P., 2012, MNRAS, 423, 1711
Maturi M., Fedeli C., Moscardini L., 2011, MNRAS, 416, 2527
Peel A., Lin C.-A., Lanusse F., Leonard A., Starck J.-L., Kilbinger M., 2017, A&A, 599, A79
Pen U.-L., Zhang T., van Waerbeke L., Mellier Y., Zhang P., Dubinski J., 2003, ApJ, 592, 664
Petri A., Haiman Z., Hui L., May M., Kratochvil J. M., 2013, Phys. Rev. D, 88, 123002
Pires S., Leonard A., Starck J.-L., 2012, MNRAS, 423, 983
Planck Collaboration, 2020, A&A, 641, A9
Porqueres N., Kodi Ramanah D., Jasche J., Lavaux G., 2019a, A&A, 624, A115
Porqueres N., Jasche J., Lavaux G., Enßlin T., 2019b, A&A, 630, A151
Porqueres N., Hahn O., Jasche J., Lavaux G., 2020, A&A, 642, A139
Ramanah D. K., Lavaux G., Jasche J., Wand elt B. D., 2019, A&A, 621, A69
Schneider P., Lombardi M., 2003, A&A, 397, 809
Semboloni E., Schrabback T., van Waerbeke L., Vafaei S., Hartlap J., Hilbert S., 2011, MNRAS, 410, 143
Simon P., Taylor A. N., Hartlap J., 2009, MNRAS, 399, 48
Smith K. M., 2006, Phys. Rev. D, 74, 083002
Spergel D. et al., 2015, preprint (arXiv:1503.03757)
Springel V., 2005, MNRAS, 364, 1105
Takada M., Jain B., 2003, MNRAS, 344, 857
Tassev S., Zaldarriaga M., Eisenstein D. J., 2013, J. Cosmol. Astropart. Phys., 2013, 036
Taylor P. L., Kitching T. D., Alsing J., Wandelt B. D., Feeney S. M., McEwen J. D., 2019, Phys. Rev. D, 100, 023519
Troxel M. A. et al., 2018, Phys. Rev. D, 98, 043528
Vafaei S., Lu T., van Waerbeke L., Semboloni E., Heymans C., Pen U.-L., 2010, Astropart. Phys., 32, 340
van Waerbeke L. et al., 2013, MNRAS, 433, 3373
van Waerbeke L., Bernardeau F., Mellier Y., 1999, A&A, 342, 15

APPENDIX A: ADJOINT GRADIENT OF THE DATA MODEL

The inference of the density field requires inferring the amplitudes of the primordial density at different volume elements of a regular grid, commonly between 128^3 and 256^3 volume elements. This implies 10^6 to 10^7 free parameters. To explore this high-dimensional parameter space efficiently, the BORG framework employs a Hamiltonian Monte Carlo (HMC) method, which adapts to the geometry of the problem by using the information in the gradients. Therefore, this algorithm requires the derivatives of the forward model. In this appendix, we derive the adjoint gradient of the shear model, which linearly transforms the error vector from the likelihood space to the parameter space of initial conditions.

More specifically, the HMC relies on the availability of a gradient of the posterior distribution. Therefore, we need to compute the gradient of the log-likelihood with respect to the initial density contrast, $\delta^{\rm ic}$.

$$\frac{\partial \log \mathcal{L}}{\partial \delta_p^{\rm ic}} = \sum_b \sum_{mn} \frac{\partial \log \mathcal{L}}{\partial \gamma_{mn}^b} \frac{\partial \gamma_{mn}^b}{\partial \delta_p^{\rm ic}} \tag{A1}$$

$$\begin{aligned}\frac{\partial \log \mathcal{L}}{\partial \delta_p^{\rm ic}} &= \sum_b \sum_{mn} \frac{\hat{\gamma}_{1,mn}^b - \gamma_{1,mn}^b}{\sigma_{mn}^2} \frac{\partial \gamma_{1,mn}^b}{\partial \delta_p^{\rm ic}} \\ &\quad + \frac{\hat{\gamma}_{2,mn}^b - \gamma_{2,mn}^b}{\sigma_{mn}^2} \frac{\partial \gamma_{2,mn}^b}{\partial \delta_p^{\rm ic}}\end{aligned} \tag{A2}$$

$$\begin{aligned}\frac{\partial \log \mathcal{L}}{\partial \delta_p^{\rm ic}} &= \sum_b \sum_{mn} \frac{\hat{\gamma}_{1,mn}^b - \gamma_{1,mn}^b}{\sigma_{mn}^2} \mathrm{DFT}^{-1} \left[\frac{l_x^2 - l_y^2}{l_x^2 + l_y^2} \mathrm{DFT} \left(\frac{\partial \kappa_{mn}}{\partial \delta_p^{\rm ic}} \right) \right] \\ &\quad + \sum_b \sum_{mn} \frac{\hat{\gamma}_{2,mn}^b - \gamma_{2,mn}^b}{\sigma_{mn}^2} \mathrm{DFT}^{-1} \left[\frac{-2 l_x l_y}{l_x^2 + l_y^2} \mathrm{DFT} \left(\frac{\partial \kappa_{mn}}{\partial \delta_p^{\rm ic}} \right) \right]\end{aligned}$$

with

$$\frac{\partial \kappa_{mn}}{\partial \delta_p^{\rm ic}} = \frac{3 H_0^2 \Omega_m}{2c^2} \left[\sum_{s=p}^{N_2} \frac{(r_s - r_p)}{r_s} n^b(r_s) \Delta r_s \right] \frac{r_p \Delta r_p}{a} \frac{\partial \mathcal{M}(a, \delta^{\rm ic})}{\partial \delta_p^{\rm ic}} \tag{A3}$$

where $\mathcal{M}(a, \delta^{\rm ic})$ is the dynamical forward model.

This paper has been typeset from a TEX/LATEX file prepared by the author.

List of astronomical key words (Updated on 2020 January)

This list is common to *Monthly Notices of the Royal Astronomical Society, Astronomy and Astrophysics*, and *The Astrophysical Journal*. In order to ease the search, the key words are subdivided into broad categories. No more than *six* subcategories altogether should be listed for a paper.

The subcategories in boldface containing the word 'individual' are intended for use with specific astronomical objects; these should never be used alone, but always in combination with the most common names for the astronomical objects in question. Note that each object counts as one subcategory within the allowed limit of six.

The parts of the key words in italics are for reference only and should be omitted when the keywords are entered on the manuscript.

General
editorials, notices
errata, addenda
extraterrestrial intelligence
history and philosophy of astronomy
miscellaneous
obituaries, biographies
publications, bibliography
sociology of astronomy
standards

Physical data and processes
acceleration of particles
accretion, accretion discs
asteroseismology
astrobiology
astrochemistry
astroparticle physics
atomic data
atomic processes
black hole physics
chaos
conduction
convection
dense matter
diffusion
dynamo
elementary particles
equation of state
gravitation
gravitational lensing: micro
gravitational lensing: strong
gravitational lensing: weak
gravitational waves
hydrodynamics
instabilities
line: formation
line: identification
line: profiles
magnetic fields
magnetic reconnection
(magnetohydrodynamics) MHD
masers
molecular data
molecular processes
neutrinos
nuclear reactions, nucleosynthesis, abundances
opacity
plasmas
polarization
radiation: dynamics
radiation mechanisms:general
radiation mechanisms: non-thermal
radiation mechanisms: thermal
radiative transfer
relativistic processes
scattering
shock waves
solid state: refractory
solid state: volatile
turbulence
waves

Astronomical instrumentation, methods and techniques
atmospheric effects
balloons
instrumentation: adaptive optics
instrumentation: detectors
instrumentation: high angular resolution
instrumentation: interferometers
instrumentation: miscellaneous
instrumentation: photometers
instrumentation: polarimeters
instrumentation: spectrographs
light pollution
methods: analytical
methods: data analysis
methods: laboratory: atomic
methods: laboratory: molecular
methods: laboratory: solid state
methods: miscellaneous
methods: numerical
methods: observational
methods: statistical
site testing
space vehicles
space vehicles: instruments
techniques: high angular resolution
techniques: image processing
techniques: imaging spectroscopy
techniques: interferometric
techniques: miscellaneous
techniques: photometric
techniques: polarimetric
techniques: radar astronomy
techniques: radial velocities
techniques: spectroscopic
telescopes

Astronomical data bases
astronomical data bases: miscellaneous
atlases
catalogues
surveys
virtual observatory tools

Software
software: data analysis
software: development
software: documentation
software: public release
software: simulations

Astrometry and celestial mechanics
astrometry
celestial mechanics
eclipses
ephemerides
occultations
parallaxes
proper motions
reference systems
time

The Sun
Sun: abundances
Sun: activity
Sun: atmosphere
Sun: chromosphere
Sun: corona
Sun: coronal mass ejections (CMEs)
Sun: evolution
Sun: faculae, plages
Sun: filaments, prominences
Sun: flares
Sun: fundamental parameters
Sun: general
Sun: granulation
Sun: helioseismology
Sun: heliosphere
Sun: infrared
Sun: interior
Sun: magnetic fields
Sun: oscillations
Sun: particle emission
Sun: photosphere
Sun: radio radiation
Sun: rotation
(Sun:) solar–terrestrial relations
(Sun:) solar wind
(Sun:) sunspots
Sun: transition region
Sun: UV radiation
Sun: X-rays, gamma-rays

Planetary systems
comets: general

comets: individual: . . .
Earth
interplanetary medium
Kuiper belt: general

Kuiper belt objects: individual: . . .
meteorites, meteors, meteoroids
minor planets, asteroids: general

minor planets, asteroids: individual: . . .
Moon
Oort Cloud
planets and satellites: atmospheres
planets and satellites: aurorae
planets and satellites: composition
planets and satellites: detection
planets and satellites: dynamical evolution and stability
planets and satellites: formation
planets and satellites: fundamental parameters
planets and satellites: gaseous planets
planets and satellites: general

planets and satellites: individual: . . .
planets and satellites: interiors
planets and satellites: magnetic fields
planets and satellites: oceans
planets and satellites: physical evolution
planets and satellites: rings
planets and satellites: surfaces
planets and satellites: tectonics
planets and satellites: terrestrial planets
planet–disc interactions
planet–star interactions
protoplanetary discs
zodiacal dust

Stars
stars: abundances
stars: activity
stars: AGB and post-AGB
stars: atmospheres
(stars:) binaries *(including multiple)*: close
(stars:) binaries: eclipsing
(stars:) binaries: general
(stars:) binaries: spectroscopic
(stars:) binaries: symbiotic
(stars:) binaries: visual
stars: black holes
(stars:) blue stragglers
(stars:) brown dwarfs
stars: carbon
stars: chemically peculiar
stars: chromospheres
(stars:) circumstellar matter
stars: coronae
stars: distances
stars: dwarf novae
stars: early-type
stars: emission-line, Be
stars: evolution
stars: flare
stars: formation
stars: fundamental parameters
(stars:) gamma-ray burst: general
(stars:) **gamma-ray burst: individual: . . .**
stars: general
(stars:) Hertzsprung–Russell and colour–magnitude diagrams
stars: horizontal branch
stars: imaging
stars: individual: . . .
stars: interiors

stars: jets
stars: kinematics and dynamics
stars: late-type
stars: low-mass
stars: luminosity function, mass function
stars: magnetars
stars: magnetic field
stars: massive
stars: mass-loss
stars: neutron
(stars:) novae, cataclysmic variables
stars: oscillations *(including pulsations)*
stars: peculiar *(except chemically peculiar)*
(stars:) planetary systems
stars: Population II
stars: Population III
stars: pre-main-sequence
stars: protostars
(stars:) pulsars: general
(stars:) **pulsars: individual: . . .**
stars: rotation
stars: solar-type
(stars:) starspots
stars: statistics
(stars:) subdwarfs
(stars:) supergiants
(stars:) supernovae: general
(stars:) **supernovae: individual: . . .**
stars: variables: Cepheids
stars: variables: Scuti
stars: variables: general
stars: variables: RR Lyrae
stars: variables: S Doradus
stars: variables: T Tauri, Herbig Ae/Be
(stars:) white dwarfs
stars: winds, outflows
stars: Wolf–Rayet

Interstellar medium (ISM), nebulae
ISM: abundances
ISM: atoms
ISM: bubbles
ISM: clouds
(ISM:) cosmic rays
(ISM:) dust, extinction
ISM: evolution
ISM: general
(ISM:) HII regions
(ISM:) Herbig–Haro objects

ISM: individual objects: . . .
(except planetary nebulae)
ISM: jets and outflows
ISM: kinematics and dynamics
ISM: lines and bands
ISM: magnetic fields
ISM: molecules
(ISM:) photodissociation region (PDR)
(ISM:) planetary nebulae: general
(ISM:) **planetary nebulae: individual: . . .**
ISM: structure
ISM: supernova remnants

The Galaxy
Galaxy: abundances
Galaxy: bulge
Galaxy: centre
Galaxy: disc
Galaxy: evolution
Galaxy: formation
Galaxy: fundamental parameters
Galaxy: general
(Galaxy:) globular clusters: general
(Galaxy:) **globular clusters: individual: . . .**
Galaxy: halo
Galaxy: kinematics and dynamics
(Galaxy:) local interstellar matter
Galaxy: nucleus
(Galaxy:) open clusters and associations: general
(Galaxy:) **open clusters and associations: individual: . . .**
(Galaxy:) solar neighbourhood
Galaxy: stellar content
Galaxy: structure

Galaxies
galaxies: abundances
galaxies: active
galaxies: bar
(galaxies:) BL Lacertae objects: general
(galaxies:) **BL Lacertae objects: individual: . . .**
galaxies: bulges
galaxies: clusters: general

galaxies: clusters: individual: . . .
galaxies: clusters: intracluster medium
galaxies: disc
galaxies: distances and redshifts
galaxies: dwarf
galaxies: elliptical and lenticular, cD
galaxies: evolution
galaxies: formation
galaxies: fundamental parameters
galaxies: general
galaxies: groups: general

galaxies: groups: individual: . . .
galaxies: haloes
galaxies: high-redshift

galaxies: individual: . . .
galaxies: interactions
(galaxies:) intergalactic medium
galaxies: irregular
galaxies: ISM
galaxies: jets
galaxies: kinematics and dynamics
(galaxies:) Local Group
galaxies: luminosity function, mass function
(galaxies:) Magellanic Clouds
galaxies: magnetic fields
galaxies: nuclei
galaxies: peculiar
galaxies: photometry
(galaxies:) quasars: absorption lines
(galaxies:) quasars: emission lines
(galaxies:) quasars: general

(galaxies:) **quasars: individual: . . .**
(galaxies:) quasars: supermassive black holes
galaxies: Seyfert
galaxies: spiral
galaxies: starburst
galaxies: star clusters: general

galaxies: star clusters: individual: . . .
galaxies: star formation
galaxies: statistics
galaxies: stellar content
galaxies: structure

Cosmology
(cosmology:) cosmic background radiation
(cosmology:) cosmological parameters
(cosmology:) dark ages, reionization, first stars
(cosmology:) dark energy
(cosmology:) dark matter
(cosmology:) diffuse radiation
(cosmology:) distance scale
(cosmology:) early Universe
(cosmology:) inflation
(cosmology:) large-scale structure of Universe
cosmology: miscellaneous
cosmology: observations
(cosmology:) primordial nucleosynthesis
cosmology: theory

Resolved and unresolved sources as a function of wavelength
gamma-rays: diffuse background
gamma-rays: galaxies
gamma-rays: galaxies: clusters
gamma-rays: general
gamma-rays: ISM
gamma-rays: stars
infrared: diffuse background
infrared: galaxies
infrared: general
infrared: ISM
infrared: planetary systems
infrared: stars
radio continuum: galaxies
radio continuum: general
radio continuum: ISM
radio continuum: planetary systems
radio continuum: stars
radio continuum: transients
radio lines: galaxies
radio lines: general
radio lines: ISM
radio lines: planetary systems
radio lines: stars
submillimetre: diffuse background
submillimetre: galaxies
submillimetre: general
submillimetre: ISM
submillimetre: planetary systems
submillimetre: stars
ultraviolet: galaxies
ultraviolet: general
ultraviolet: ISM
ultraviolet: planetary systems
ultraviolet: stars
X-rays: binaries
X-rays: bursts
X-rays: diffuse background
X-rays: galaxies
X-rays: galaxies: clusters
X-rays: general
X-rays: individual: . . .
X-rays: ISM
X-rays: stars

Transients
(transients:) black hole mergers
(transients:) black hole - neutron star mergers
(transients:) fast radio bursts
(transients:) gamma-ray bursts
(transients:) neutron star mergers
transients: novae
transients: supernovae
transients: tidal disruption events