

# Bayesian mass mapping with weak lensing data using KARMMA: Validation with simulations and application to Dark Energy Survey year 3 data

Supranta S. Boruah<sup>1,2,\*</sup>, Pier Fiedorowicz<sup>3,4</sup> and Eduardo Rozo<sup>1,3</sup>

<sup>1</sup>*Department of Astronomy and Steward Observatory, University of Arizona,  
933 North Cherry Avenue, Tucson, Arizona 85719, USA*

<sup>2</sup>*Department of Physics and Astronomy, University of Pennsylvania,  
Philadelphia, Pennsylvania 19104, USA*

<sup>3</sup>*Department of Physics, University of Arizona, 1118 East Fourth Street, Tucson, Arizona 85721, USA*

<sup>4</sup>*Lawrence Livermore National Laboratory, Livermore, California 94550, USA*



(Received 10 March 2024; accepted 14 June 2024; published 18 July 2024)

We update the field-level inference code KARMMA to enable tomographic forward-modeling of shear maps. Our code assumes a log-normal prior on the convergence field, and properly accounts for the cross-covariance in the lensing signal across tomographic source bins. We use mock weak lensing data from  $N$ -body simulations to validate our mass-mapping forward model by comparing our posterior maps to the input convergence fields. We find that KARMMA produces more accurate reconstructions than traditional mass-mapping algorithms. Moreover, the KARMMA posteriors reproduce all statistical properties of the input density field we tested—1- and 2-point functions, and the peak and void number counts—with  $\approx 10\%$  accuracy. Our posteriors exhibit a small bias that increases with decreasing source redshift, but these biases are small compared to the statistical uncertainties of current [Dark Energy Survey (DES)] cosmic shear surveys. Finally, we apply KARMMA to DES year 3 weak lensing data, and verify that the 2-point shear correlation function  $\xi_+$  is well fit by the correlation function of the reconstructed convergence field. This is a nontrivial test that traditional mass mapping algorithms fail.

DOI: [10.1103/PhysRevD.110.023524](https://doi.org/10.1103/PhysRevD.110.023524)

## I. INTRODUCTION

Most standard analyses of cosmological data use 2-point summary statistics such as correlation functions and power spectra for extracting information from these datasets (e.g., [1,2]). However, we lose valuable non-Gaussian information while compressing the full dataset to 2-point summary statistics. Field-based inference, where one extracts information from the full field, is emerging as a powerful alternative for optimally extracting information from cosmological data sets across a broad variety of probes, including cosmic microwave background lensing [3–5], integrated Sachs-Wolfe effect [6], galaxy clustering [7,8], peculiar velocity [9–11] and weak lensing [12–14].

In this paper, we focus on field-based inference of weak lensing data. Such an analysis requires an accurate forward model of the observed shear field. Once we have an accurate forward model for the weak lensing data, we can use it to reconstruct the unknown dark matter distribution of the Universe. Indeed, reconstructed mass maps from weak lensing surveys are crucial for extracting non-Gaussian information (e.g., [15]), for studying voids [16], and for cross-correlation studies [17,18]. As such, they are

an essential data product for stage-IV weak lensing surveys. However, the reconstruction of the unseen dark matter distribution from weak lensing shear data constitutes a nontrivial inverse problem. While standard inversion methods such as Kaiser-Squires inversion [19] are widely used, they suffer a variety of problems. For example, these methods lead to biased reconstruction in the presence of a survey mask. Furthermore, in the presence of shape noise the reconstructed mass maps do not have the expected statistical properties. For instance, the two-point functions of the recovered maps are inconsistent with direct measurements of the two point shear correlation function (see Fig. 11).

Forward modeled mass-map reconstruction naturally address these difficulties. For example, by forward-modeling the convergence field in the masked region, we get mass maps free from masking effects [20,21]. Likewise, depending on the accuracy of the forward model, we are able to reconstruct mass maps with the correct 2-point and 1-point function.

A number of different weak lensing field-based analysis methods have been proposed in the literature, though none so far have been applied to data. Running approximate numerical structure formation simulations is likely to be the most accurate forward model. This is the approach

\*Contact author: [supranta@sas.upenn.edu](mailto:supranta@sas.upenn.edu)

followed by [13,22], who used Lagrangian perturbation theory to forward model the cosmic shear data within the BORG Bayesian forward modeling framework [7,23]. However, this approach is computationally expensive and analyses are often limited to small areas with a limited resolution.

Because of these difficulties, a number of alternative field-based inference methods have tried to directly model the observed two-dimensional projected fields, thereby bypassing the need to generate a three-dimensional matter density field. For example, [24,25] developed a Bayesian hierarchical model that simultaneously inferred the mass distribution and power spectrum of the weak lensing maps in tomographic bins. However, their method assumed Gaussianity of the field and only worked on flat sky. References [26,27] proposed a similar algorithm, ALMANAC, which performs wide-field mass mapping on a sphere, but still assumes Gaussianity. Similarly, [16] sampled mass maps with DES-Y3 data using the constrained realization approach [28–30] assuming Gaussianity of the convergence field.

Since the late-time density field is highly non-Gaussian, it is important to include the non-Gaussianity of the density/convergence field in the forward model. To that end, [21,31] introduced the KARMMA algorithm. In KARMMA, the convergence field is modeled as a log-normal random field, enabling us to bypass the need to run expensive numerical simulations to reconstruct the matter density field, while simultaneously improving upon the Gaussian assumption made in other works (e.g., ALMANAC). More recently [14] introduced MIKO, a flat-sky field-level inference code that also utilizes the same log-normal prior for the convergence field. The log-normal model has been shown to be a good approximation of the late-time non-Gaussian convergence field [32,33]. Indeed, in our earlier works we demonstrated that the KARMMA posteriors from mock observations of simulated shear maps accurately reproduced a broad range of statistical properties of the corresponding input convergence fields. However, the algorithm had a severe limitation: it was only able to forward model a single tomographic source bin at a time. Here, we extend the KARMMA algorithm using the methodology of [12] to enable tomographic mass mapping on a sphere while modeling the non-Gaussianity of the convergence field. We validate the updated code by running it on mock weak lensing data based on  $N$ -body simulations. We find that the mass maps reconstructed with KARMMA have smaller errors than traditional mass mapping approaches. Further, the statistical properties of the posterior maps are consistent with those from numerical simulations. After validating our method with mock simulations, we apply KARMMA to the Dark Energy Survey (DES) year 3 (Y3) weak lensing data. To the best of our knowledge, these are the first publicly available Bayesian mass maps reconstructed with weak lensing data on the curved sky [34].

This paper is structured as follows: in Sec. II, we present the simulations and the weak lensing data used in this work. In Sec. III, we present the main features of KARMMA. Then in Sec. IV, we validate KARMMA with mock weak lensing simulations before presenting the Bayesian mass maps produced from DES-Y3 data in Sec. V. Finally, we conclude in Sec. VI. In the Appendix, we show the results of our tests with log-normal mocks.

## II. DATA

### A. Mock maps from simulations

In this section, we describe the mock weak lensing catalogs used in Sec. IV to test the performance of our code. We use the publicly available simulations of [[35], hereafter T17]. T17 produced a suite of 108 mock weak lensing catalogs by ray-tracing  $N$ -body simulations run with cosmological parameters  $\Omega_m = 0.279$ ,  $\Omega_b = 0.046$ ,  $h = 0.7$ ,  $\sigma_8 = 0.82$  and  $n_s = 0.97$ . Each mock catalog consists of full-sky shear and convergence maps on redshift shells separated by  $150h^{-1}$  Mpc. We use HEALPix maps at a resolution of  $N_{\text{side}} = 4096$  for this work.

To generate simulated shear maps, we adopt the survey properties from the DES-Y3 analysis [36,37]. Specifically, we use four tomographic redshift bins. The source redshift distribution of each bin is taken from [38], shown here in Fig. 1. The effective number density in each of the tomographic bins is  $[1.476, 1.479, 1.484, 1.461]$  arcmin $^{-2}$ . Finally, we adopt the shape noise estimate from the DES-Y3 shear catalog,  $\sigma_e = 0.261$  [39].

We produce tomographic maps of the quantity  $\mathcal{A}$  (here  $\mathcal{A}$  can be either the convergence or shear map) by taking a weighted average of the maps at various redshift shells with a given redshift distribution,  $n(z)$ , via

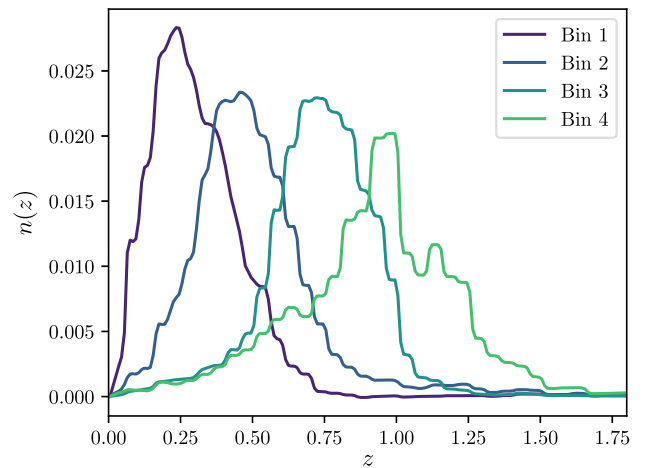


FIG. 1. The source redshift distribution,  $n(z)$ , for the DES-Y3 used in this work.

$$\mathcal{A}^i = \frac{\sum_{j=1}^{N_{\text{shells}}} n^i(z_j) \mathcal{A}_{\text{shell}}^j}{\sum_{j=1}^{N_{\text{shells}}} n^i(z_j)}, \quad (1)$$

where  $n^i$  is the redshift distribution of the  $i$ th tomographic bin and  $\mathcal{A}_{\text{shell}}^j$  is the map of the quantity  $\mathcal{A}$  on the  $j$ th redshift shell. Note that the finite redshift resolution leads to a systematic error of  $\mathcal{O}(\lesssim 5\%)$  on the power spectrum of the convergence field. Consequently, for the purposes of testing our code, we use the power spectrum measured from the T17 simulations, rather than a theory power spectrum computed using Boltzmann codes such as CAMB [40] or CLASS [41]. We change this in Sec. V while applying to DES-Y3 data.

Our “observed” shear map is a noisy realization of the true shear. Specifically, we produce our noisy shear maps at a higher resolution of  $N_{\text{side}} = 1024$  and then downgrade to a resolution of  $N_{\text{side}} = 256$  to mimic the impact of pixelization. The number of sources in each sky pixel is a Poisson random draw from the mean source density of the appropriate tomographic bin in the higher resolution maps. Note that this ignores source clustering which is an important systematic effect while considering non-Gaussian information [42]. We add to each shear component a Gaussian shape noise with zero mean and standard deviation,  $\sigma_\epsilon / \sqrt{N_p}$ , where  $N_p$  is the number of source galaxies sampled in that pixel. A resolution of  $N_{\text{side}} = 256$  corresponds to a pixel size with angular resolution of  $\sim 13$  arcmin.

### B. DES-Y3 weak lensing data

In Sec. V, we run KARMMA on the DES-Y3 weak lensing data to produce Bayesian mass maps from weak lensing data. We use the DES-Y3 METACALIBRATION [43,44] shape catalog [39]. These galaxies were divided into the 4 tomographic bins according to their estimated photometric redshifts [38]. Then we construct a HEALPIX map of estimated shear for each of the tomographic bin by taking the weighted average over the galaxy ellipticities,  $\epsilon$ , in each pixel as

$$\gamma_p^\alpha = \frac{\sum_{i \in p} w_i \epsilon_{i,\alpha}}{\sum_{i \in p} w_i}, \quad \alpha = 1, 2, \quad (2)$$

where  $w_i$  are the per-galaxy inverse variance weights,  $p$  is the pixel label, and  $\alpha$  denotes the two polarizations. The estimated variance of each shear component in the pixel  $p$  is given as [45,46]

$$\sigma_{\epsilon,p}^2 = \frac{\sum_{i \in p} w_i^2 (e_{i,1}^2 + e_{i,2}^2) / 2}{(\sum_{i \in p} w_i)^2}. \quad (3)$$

### III. MASS MAP INFERENCE WITH KARMMA

KARMMA models the convergence field as realizations of a log-normal random field. The observed shears are then forward modeled from the resulting convergence fields. We determine the posterior distribution of convergence maps by sampling the posterior using Hamiltonian Monte Carlo (HMC) methods. In this way, rather than providing a single “best-fit” convergence, we recover the full posterior distribution of convergence maps consistent with the input shear data. These posterior quantify the uncertainty in the reconstructed mass maps. We describe this algorithm in further detail below.

Our log-normal assumption states that the convergence field  $\kappa$  is a nonlinear transform of a Gaussian random field  $y$  where

$$\kappa = e^y - \lambda. \quad (4)$$

The parameter  $\lambda$  is called the shift parameter. At the resolutions of this study ( $N_{\text{side}} = 256$ , or  $\approx 13$  arcmin), a multivariate log-normal model provides a good description of the convergence field [32,33].

Here, we use the log-normal formalism developed in [12,21]. Specifically, we sample the spherical harmonics (or Fourier wave modes) of the  $y$  maps. Furthermore, following [12] we introduce new reparametrized variables,  $\mathbf{x}$ , defined via

$$y_{lm}^{(i)} = \sum_{k=1}^{N_{\text{bin}}} \mathbf{L}^{ik}(l) x_{lm}^{(k)}, \quad (5)$$

where  $y_{lm}^{(i)}$  is the spherical harmonics of the  $y$  variable defined in Eq. (4) and  $\mathbf{L}(l)$  is the Cholesky decomposition of a  $N_{\text{bin}} \times N_{\text{bin}}$  matrix constructed from the power spectrum  $C(l)$  at a wave mode  $l$  in various tomographic bins ( $N_{\text{bin}}$  being the number of tomographic bins).<sup>1</sup> The power spectrum and the shift parameters are computed at the T17 cosmological parameters using their simulations. The index  $i$  labels the tomographic bin of  $y/x$ . This reparametrization helps in two ways: (1) it decorrelates the  $\mathbf{x}$  variable in different redshift bins under the prior,<sup>2</sup> and (2) the  $\mathbf{x}$  variable now has unit variance and is described by the prior

$$\langle x_{lm}^{(i)} x_{l'm'}^{(j)} \rangle = \delta_{ij}^K \delta_{ll'}^K \delta_{mm'}^K, \quad (6)$$

where  $\delta^K$  is the Kronecker delta symbol. This redefinition makes the sampling of mass maps more efficient.

<sup>1</sup>Note that the definition of the unit variance variables differ from the definition in [12] where the  $x_{lm}$  variables were related to the  $y_{lm}$  through the eigenvalue decomposition of  $C(l)$ .

<sup>2</sup>Note that posterior samples of  $x$  can be correlated due to the likelihood.

Furthermore, even though the  $x$  maps are decorrelated under the prior, the  $y/\kappa$  maps in difference redshift bins are correlated with the theoretically expected cross-correlation. As we had shown in [12] and as we will see in Sec. V, including this cross-correlation as a part of the prior leads to massive improvement in the recovery of the small scale structures in the maps.

To better understand why adding this covariance is helpful, it is worth considering a model in which the tomographic convergence maps are uncorrelated. In this case, the posterior distribution of the maps in different tomographic bins factors out, so that the recovered convergence map in bin  $i$  depends only on the shear map in bin  $i$ . This is exactly how traditionally mass-mapping algorithms work (e.g., [47]). Thus, rather than being agnostic to the amount of covariance present in the data, analyses that “ignore” covariance in the model are enforcing a prior that the maps should not be correlated. By contrast, our correlated model results in a nonfactorizable posterior distribution which allows bin  $i$  to be informed by the data in  $j$  shear map.

Following [31], we filter our maps at  $\ell_{\max} = 2N_{\text{side}}$  to avoid aliasing. Note that the low-pass filter is applied on the  $\kappa$  field and not on the  $y$  (or  $x$ ) fields. Also note that because of the filtering, the resulting  $\kappa$  field is no longer a log-normal field.

Once we have the convergence field, the shear field is constructed using the Kaiser-Squires (KS) relation. We perform our inference on the sphere, thus requiring the curved sky version of the KS relation; the spherical harmonic coefficients of the shear field are related to the spherical harmonic coefficients of the  $\kappa$  field via [48]

$$\tilde{\gamma}_{lm} = \sqrt{\frac{(l-1)(l+2)}{l(l+1)}} \kappa_{lm}. \quad (7)$$

The shear field is then obtained as a sum over the spin-2 spherical harmonics,

$$\gamma(\boldsymbol{\theta}) = \sum_{lm} \tilde{\gamma}_{lm2} Y_{lm}(\boldsymbol{\theta}). \quad (8)$$

Since the convergence field is low-pass filtered, the shear field is also low-pass filtered. By contrast, the observed shear field is *not* filtered, which leads to some amount of

model misspecification. However, we have found that given the resolution of our maps, this slight model misspecification is easily preferable to the aliasing that occurs when we do not filter the field [31].

Finally, we use pixelized shear maps to perform the inference, so we must account for the pixel window function in our forward model. The pixelization operation is implemented in the harmonic space by multiplying each spherical harmonic mode by the spherical transform of the pixel window function,

$$\tilde{\gamma}_{lm} \rightarrow W_p^{N_{\text{side}}}(l) \tilde{\gamma}_{lm}, \quad (9)$$

where  $W_p^{N_{\text{side}}}$  is the HEALPix pixel window function at resolution of  $N_{\text{side}}$ . While in principle the pixel window function for each HEALPix pixel is different, an average pixel window function can be computed using the HEALPY function `healpy.sphtfunc.pixwin`. The need for including the pixel window function in the forward model has previously been noted in [14,21].

Given a  $\kappa$  map, the observed shear map is modeled as a noisy realization of the true shear map, where the noise is set by the shape noise of the survey. In particular, the noise is uncorrelated from pixel to pixel, so that

$$\log P(\gamma_{\text{obs}}|\mathbf{x}) = -\frac{\chi^2}{2} = -\sum_{p=1}^2 \sum_{i=1}^{N_{\text{bin}}} \sum_{\alpha=1}^{N_{\text{pix}}} \frac{[\gamma_{\alpha}^{i,p}(\mathbf{x}) - \gamma_{\alpha,\text{obs}}^{i,p}]^2}{2\sigma_{i,\alpha}^2}. \quad (10)$$

The summation is over all the pixels within the survey mask and over the two polarizations (labeled with  $p$ ).

Given an “observed” shear map, the posterior of  $\mathbf{x}$  is simply

$$P(\mathbf{x}|\gamma_{\text{obs}}) \propto P(\gamma_{\text{obs}}|\mathbf{x})P(\mathbf{x}). \quad (11)$$

Because  $\mathbf{x}$  are unit variance variables, the prior,  $P(\mathbf{x})$  is a unit variance normal distribution.

The various steps involved in our forward model are illustrated in Fig. 2. Note that all the intermediate steps are differentiable allowing us to take the gradient of the posterior with respect to the variables  $\mathbf{x}$ , as is required to sample using HMC. We implement our forward model in the Python package PyTorch [49] and use the probabilistic programming language Pyro [50] to sample the mass maps.

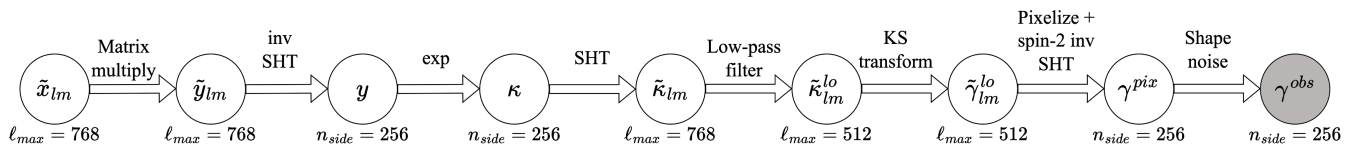


FIG. 2. Illustration of the various steps involved in the KARMMA forward model. Here, SHT refers to the spherical harmonic transform, KS refers to Kaiser-Squires. Note that each of the steps illustrated above are differentiable, thus allowing us to compute the gradient with respect to the input variables—a necessary feature for using HMC methods. The maximum  $\ell$  value of the spherical harmonic modes and the HEALPix map resolution of the maps is shown at each step. See Sec. III for details.

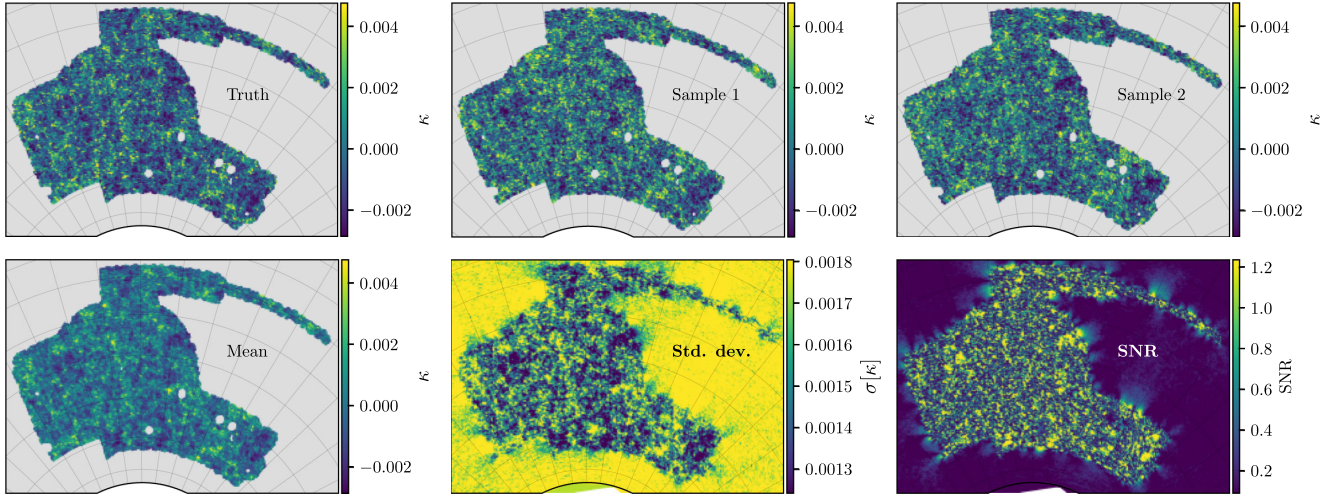


FIG. 3. An illustration of the properties of the KARMMA posteriors for the first tomographic redshift bin in our simulations. The input convergence map used to generate our mock observations is shown in the top left panel. The center and right panels on the top row shows two randomly chosen KARMMA map samples. The mean map and the corresponding pixel-by-pixel rms uncertainty are shown in the bottom left and bottom center panels, respectively. The bottom right panel shows the signal to noise ratio computed from the sample mean and standard deviation.

We use the No U-Turn Sampler [51] to sample  $\mathbf{x}$  (or equivalently  $\kappa$ ).

We emphasize that in the current implementation, the transformation from the  $x$  variables to convergence are conditioned on the power spectrum and shift parameter characterizing the log-normal prior. Both of these are cosmology dependent, and therefore one must specify a cosmological model to sample from these maps. One can, in principle, simultaneously sample the cosmological parameters characterizing the prior as part of the inference process as described in [12]. In this way, KARMMA can produce joint cosmology and mass map posteriors. However, we have not yet characterized the shift parameter  $\lambda$  as a function of cosmology, nor have we validated the resulting cosmological inference pipeline. For this reason, we postpone this simultaneous inference process to future work.

#### IV. VALIDATING KARMMA WITH MOCK SIMULATIONS

We test the performance of KARMMA using the simulations presented in Sec. II A. The multivariate log-normal prior requires the power spectrum and the shift parameters as its input. We measure these quantities in the simulations, averaging over all 108 simulations, and use these as the inputs to the log-normal prior.<sup>3</sup> We then run KARMMA on all 108 mock data sets created from the T17 data. This allows us to check how well the log-normal forward model

<sup>3</sup>We remind the reader that the use of the simulated power spectrum is necessary due to the small differences between the power spectrum in the simulations and theory power spectra. When applied to data, we replace the power spectrum by the appropriate theory power spectrum for our fiducial cosmology.

performs on mock weak lensing data from  $N$ -body simulations. Each KARMMA run requires  $\sim \mathcal{O}(50\text{--}100)$  CPU core hours to produce 100 independent samples on Advanced Micro Devices Inc., Zen2 processors.

Figures 3 and 4 illustrate the properties of the KARMMA posteriors when run on our simulated dataset. The figures correspond to the mass maps from tomographic redshift bins 1 and 4, respectively. The input mass map used to generate the simulation is shown on the top left panel. The center and right panels on the top row show two randomly chosen samples for our posterior distribution of maps. The mean and standard deviation of the posteriors are shown in the bottom left and bottom center panels, respectively. The bottom right panels shows the signal-to-noise ratio defined as

$$\text{SNR} = \frac{\text{Mean}(\kappa)}{\sigma[\kappa]}. \quad (12)$$

By comparing the input convergence map to the KARMMA posterior and mean map, we can see that KARMMA recovers the dark matter distribution of the input map. We also see that the higher redshift bins have a higher signal-to-noise ratio, as expected.

The fidelity of our reconstructed mass maps can be quantified using the correlation coefficient between the reconstructed map and the input  $\kappa$  maps defined as

$$\rho_c(\ell) = \frac{C_{\text{rec}\times\text{input}}(\ell)}{\sqrt{C_{\text{input}}(\ell)C_{\text{rec}}(\ell)}}, \quad (13)$$

where  $C_{\text{rec}\times\text{input}}$  is the cross power spectrum between the reconstructed and the input map, while  $C_{\text{rec}/\text{input}}$  is the auto power spectrum of the reconstructed/input map. We

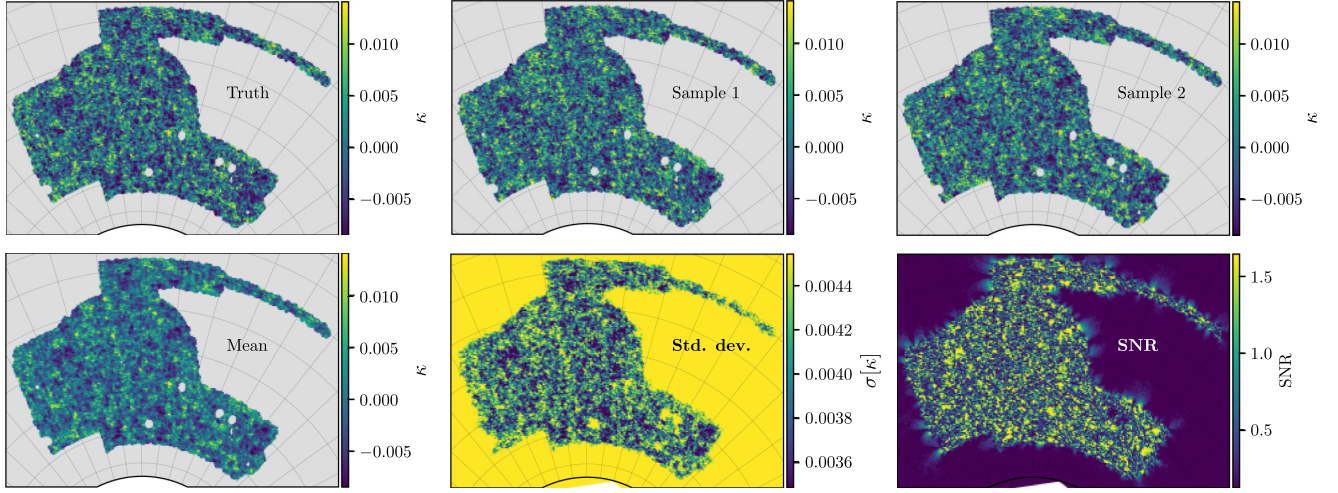


FIG. 4. Same as Fig. 3, but for the fourth tomographic bin. Compared to Fig. 3, the structures are produced with a higher SNR in the higher redshift bins.

compute the cross-correlation as a function of scale in harmonic space. The cross-correlation between the mean of the posteriors maps and the corresponding input map is shown in Fig. 5. As we can see from the figure, the cross-correlation on large scales (i.e., low  $\ell$ ) is higher than on smaller scales. This is because the signal-to-noise ratio on small scales is reduced by the shape noise. We also see that the cross-correlation of the lowest redshift bin is the lowest among the various redshift bins due to the lower signal-to-noise ratio of the data in this bin.

Next we assess the quality of the reconstructed mass maps by computing the pixel-by-pixel errors in the output mass maps. The error between the convergence value in the  $i$ th pixel of a reconstructed KARMMA sample (labeled  $s$ ) and the simulation is

$$\Delta\kappa_i = \kappa_{\text{KARMMA},s}^i - \kappa_{\text{sim}}^i. \quad (14)$$

In Fig. 6, we compare the histogram of the convergence residuals for two different maps: (1) a randomly chosen map from our posterior sample; and (2) a traditional Kaiser-Squires mass map. Both maps are generated using the same input shear map. We find the errors in the KARMMA mass maps are always much lower than those in Kaiser-Squires maps. The standard deviation of the KARMMA residuals are 85%, 77%, 66% and 60% lower than the standard

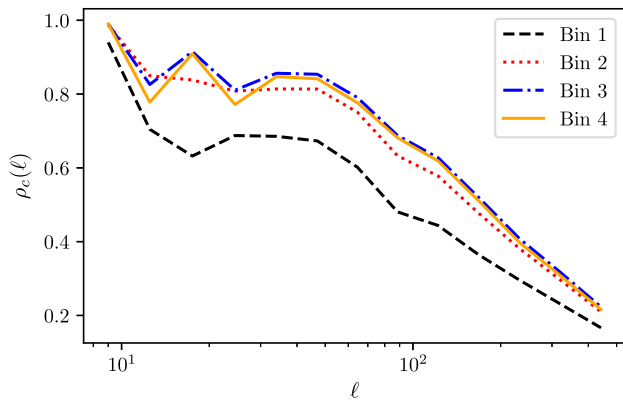


FIG. 5. Cross-correlation of the mean KARMMA map with the true simulation map. The different lines indicate the cross-correlation in the different tomographic bins. As we can see the cross-correlation is the highest for the high-redshift bins where the signal-to-noise ratio is the highest. Furthermore, we see a higher cross-correlation at large scales (low  $\ell$ s) than at small scales that are dominated by the shape noise.

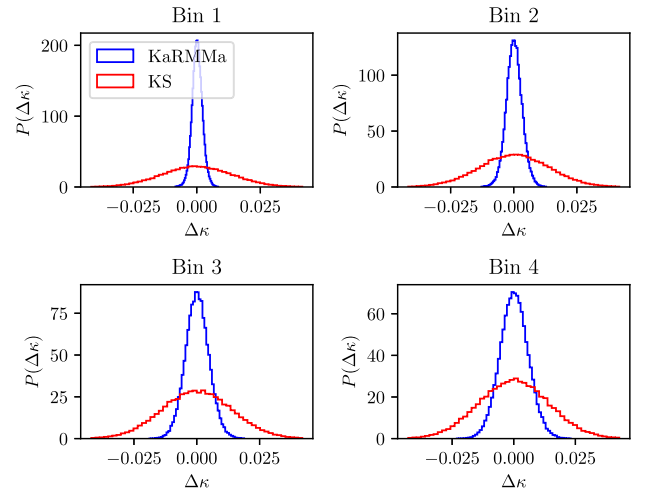


FIG. 6. Comparison of the residual errors in a randomly chose KARMMA mass map sample (blue) and Kaiser-Squires mass map (red) created with the same data. The different panels show the results in different tomographic bins. As we can see from the figure, the residuals of the KARMMA mass maps are much lower than that of the Kaiser-Squires mass maps, showing the better quality of the KARMMA mass maps. The improvement is especially noteworthy for the lowest redshift bin, where the signal-to-noise ratio is the lowest.

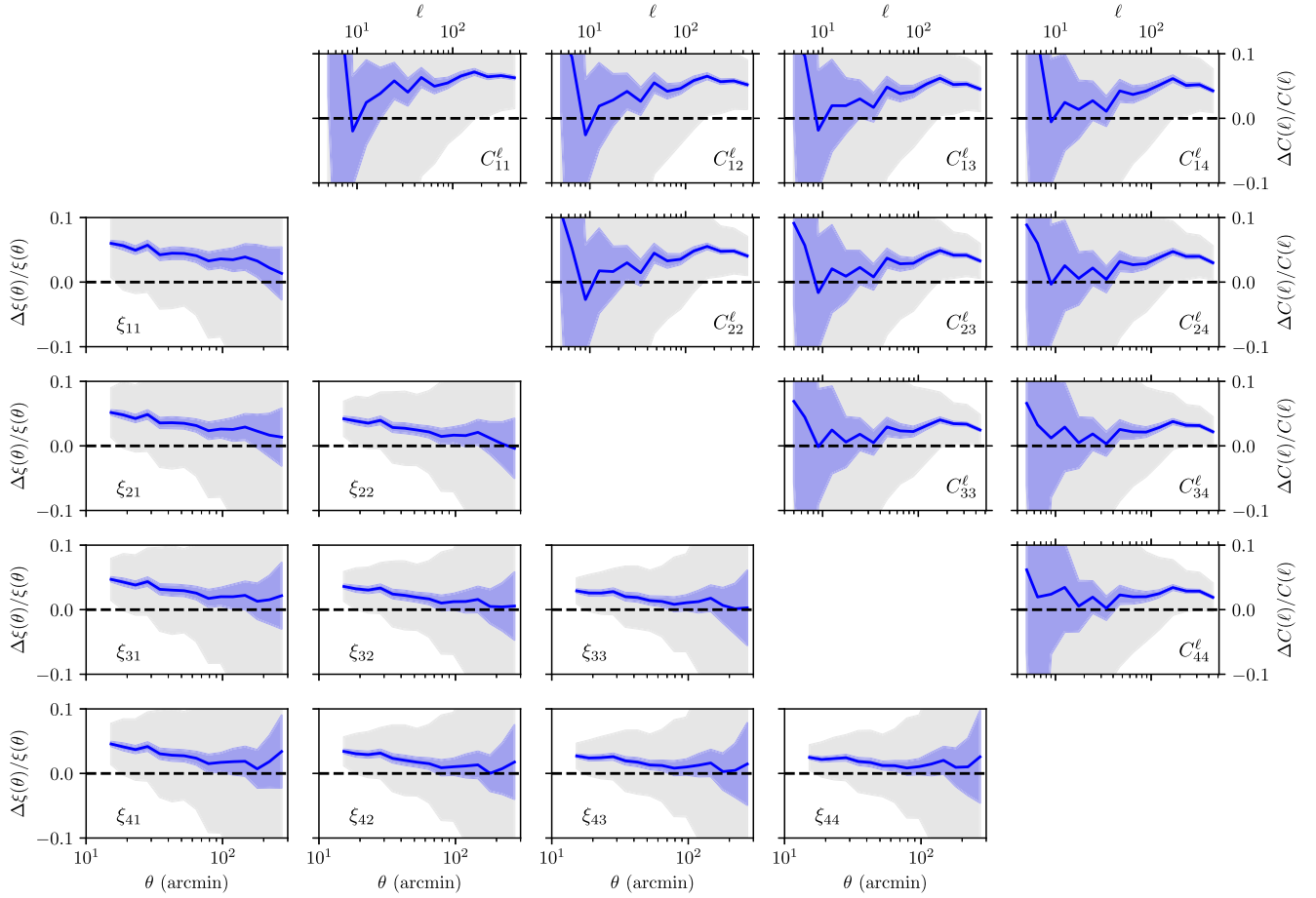


FIG. 7. Fractional error between the 2-point functions of the input simulation maps and those estimated from the KARMMA posteriors. The bottom-left triangles shows the correlation functions between the different tomographic bins, while the upper-right triangles shows the corresponding power spectra computed using pseudo- $C(l)$ s. Each panel corresponds to a different cross-bin combination, as labeled. The blue lines show the difference in the map statistics averaged across the full 108 simulated maps. The gray bands shows the mean error in one map as estimated using the KARMMA posteriors, while the blue band shows the error on the mean estimated using all 108 simulations. The difference between the simulated maps and our posteriors increases with decreasing redshift and smaller scales, and is typically  $\lesssim 5\%$ .

deviations of the KS residuals in the 4 tomographic bins, respectively. The improvement is especially noticeable for the lowest redshift bin, where the signal-to-noise ratio is the lowest. As shown in [12], adequately accounting for the covariance between tomographic bins dramatically improves the signal-to-noise ratio of the  $\kappa$  reconstruction at low redshift by allowing us to extract information about low redshift structures contained in the high redshift tomographic bins.

As we now demonstrate, one of the principal advantages of KARMMA is that the posterior maps faithfully reproduce the statistical properties of the input convergence field. Figure 7 shows the difference between the mean 2-point function in our posterior maps and that of the input convergence map in both real and harmonic space. For the latter, we compute pseudo- $C(l)$ s using the publicly available code NaMaster [52]. The blue line shows the mean difference, averaged across all 108 simulations. The gray

bands show the mean error in the posterior for one simulation, while the blue bands show the error on the mean. The difference in the two point functions between the input simulated maps and our reconstructed maps is  $\lesssim 5\%$ , with the difference increasing with decreasing redshift. These biases arise because of the failure of the log-normal model to accurately model nonlinear structure formation. In particular, the Appendix demonstrates that we do not observe any such biases when we test KARMMA using log-normal convergence maps as input.

Apart from the 2-point function, the log-normal model also captures many non-Gaussian features of the convergence field. Figure 8 compares three such statistics, namely the probability density function (PDF) of  $\kappa$  values in each pixel (i.e., the 1-point function), and the peak and void counts of the  $\kappa$  maps. The peaks (voids) of the map are defined as the number of pixels which have the highest (lowest)  $\kappa$  value among all its neighbors. These statistics are

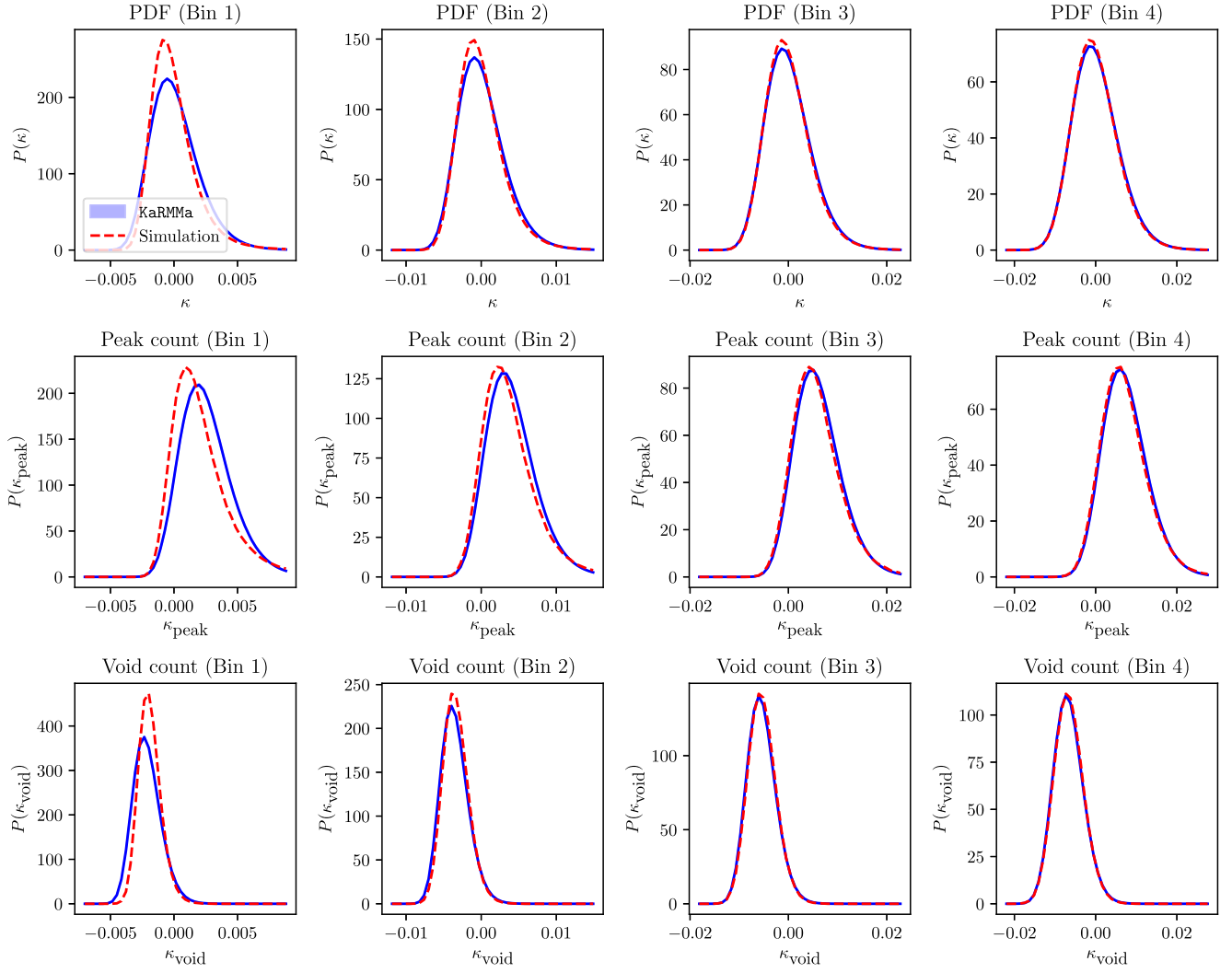


FIG. 8. Comparison of non-Gaussian statistics of the KARMMA maps (blue) and of the mock simulations (red dashed). The different rows corresponds to the three non-Gaussian statistics computed here—the PDF of  $\kappa$  values (top panels), the peak counts (middle panels) and the void counts (bottom panels). The different columns show these statistics in the 4 tomographic bins. As we can see, all these non-Gaussian summary statistics are well approximated by KARMMA at high redshift, whereas at low-redshift, these summary statistics are significantly biased due to the breakdown of the log-normal model.

well known to be sensitive to the non-Gaussian information of the convergence field [53,54]. We see that the non-Gaussian statistics of the KARMMA maps are close to those in the simulations, particularly for the high redshift bins. However, the differences between the simulated maps and our posteriors increases with decreasing redshift, being particularly large for the lowest redshift bin. This is not surprising, as it is at this redshift that nonlinear structure formation is most important. We note that because the signal-to-noise ratio of the convergence in an individual pixel is low ( $S/N \lesssim 1$ ), the one point of the posteriors is nearly identical to that of the prior.

## V. BAYESIAN MASS MAPS WITH DES-Y3 DATA

We run KARMMA on pixelized shear maps produced from the DES-Y3 shape catalogs as detailed in Sec. II B. The input  $C(l)$  for this run is computed for the T17 cosmological parameters using PYCCL [55]. We use the shift parameters computed from the T17 simulations for these KARMMA runs. We produce a chain with 500 independent samples. The maps showing the mean and standard deviation of the posteriors in each of the four tomographic bins is shown in Fig. 9. The figure highlights the fact that

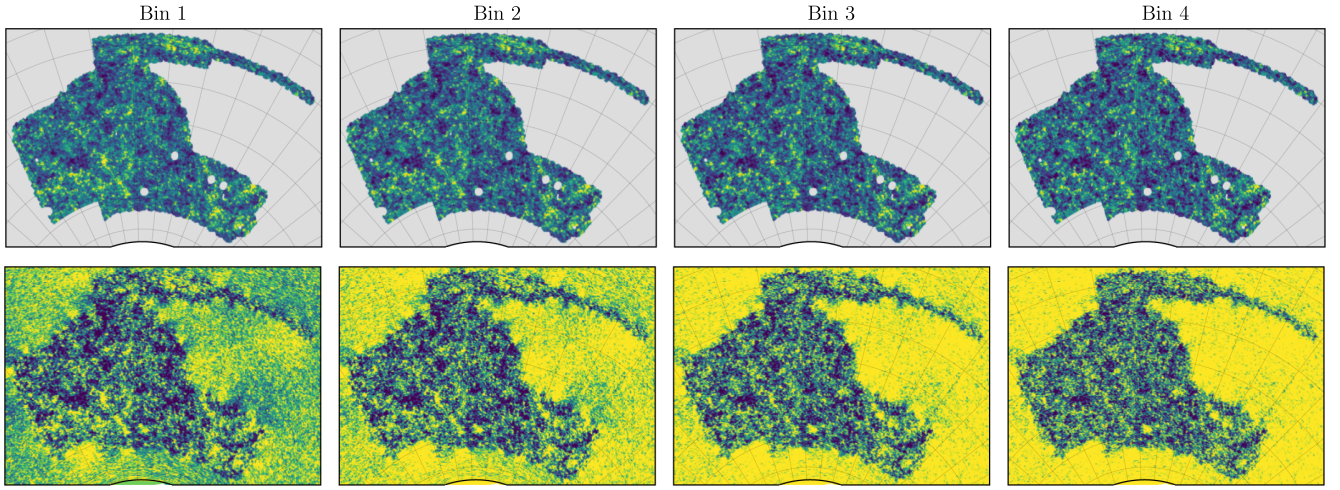


FIG. 9. The mean (top) and standard deviation (bottom) of the KARMMA posterior maps for DES-Y3 data in each of the 4 tomographic bins, as labeled.

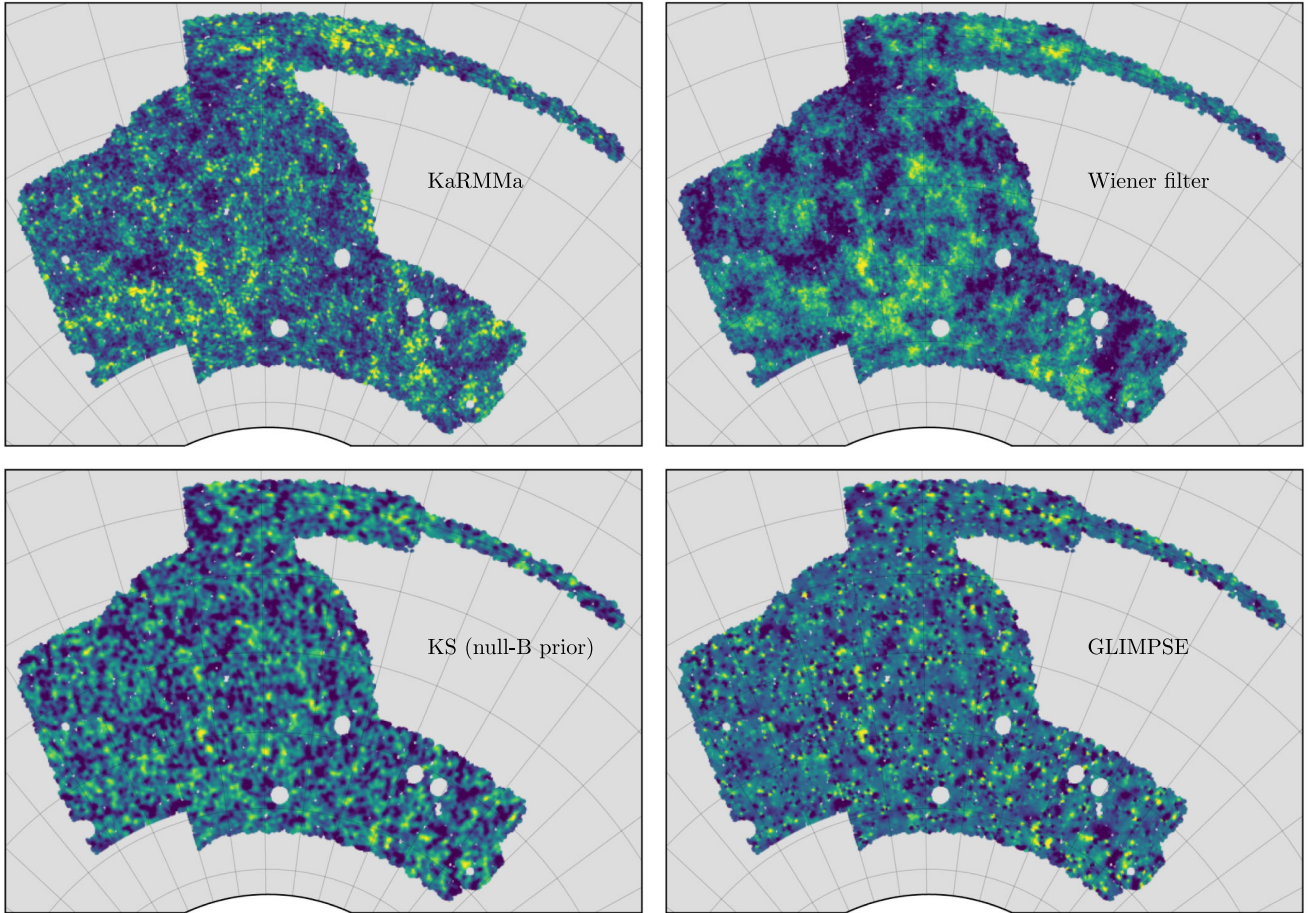


FIG. 10. Comparison of the mass map for the lowest tomographic redshift bin in the DES-Y3 data. Each panel corresponds to a different algorithm: KARMMA mean map (top left), Wiener filtering (top right), Kaiser-Squires with null B-mode prior (bottom left), and GLIMPSE (bottom right). The fact that KARMMA adequately models the covariance between tomographic redshift bins enable us to robustly recover structure on scales significantly smaller than those resolved in other algorithms.

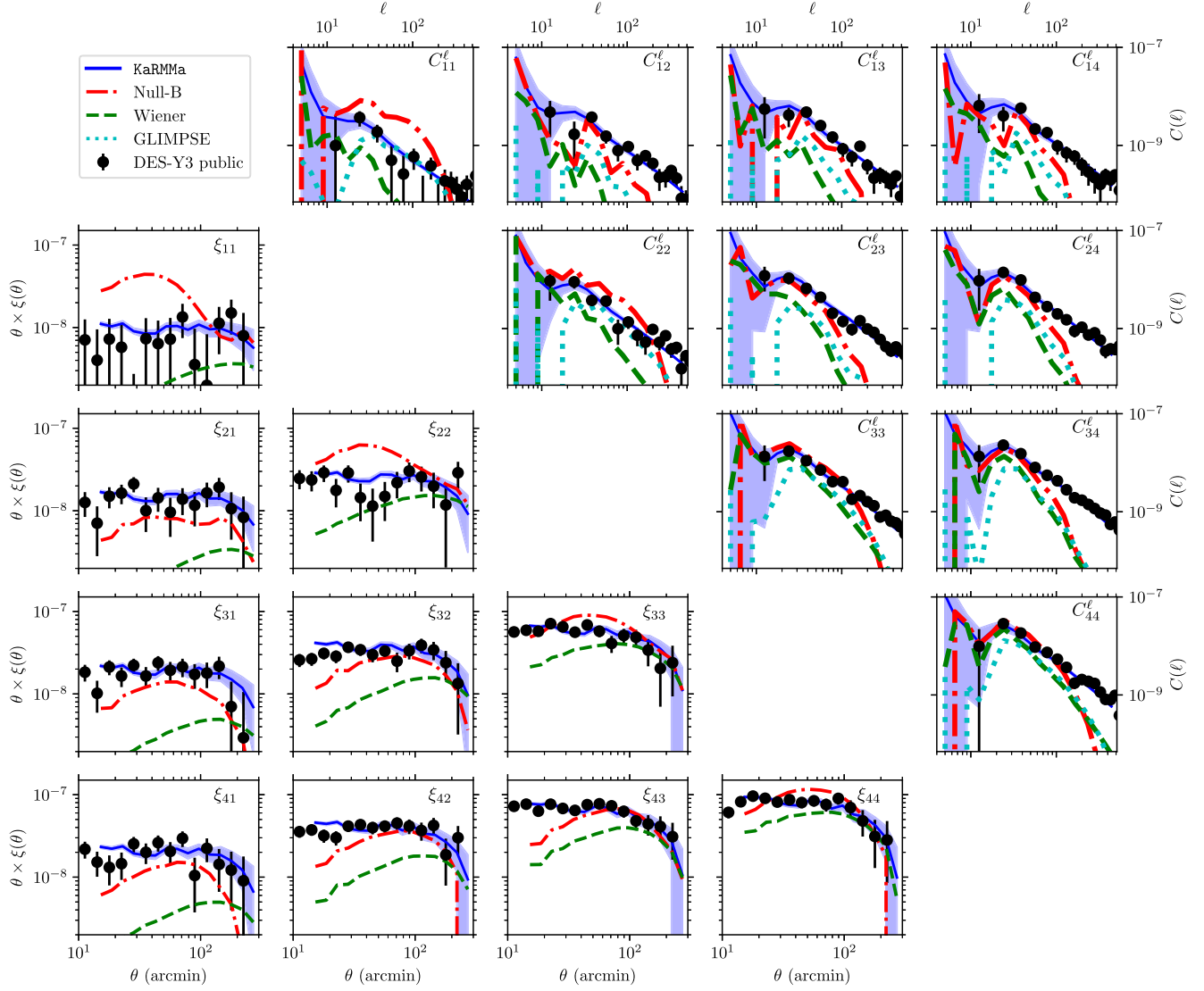


FIG. 11. Comparison of the correlation function (bottom left) and the pseudo- $C(\ell)$  (top right) of mass maps created from DES-Y3 data with the same data vectors computed directly from the DES-Y3 shape catalog (black). The blue shaded region shows the 95% confidence interval of the summary statistics of KARMMA posteriors. The correlation function and pseudo- $C(\ell)$ s of DES-Y3 mass maps created with different methods, namely Null-B prior KS (red dash dotted), Wiener filter (green dashed), GLIMPSE (cyan dotted), are also shown in the figure. As we can see from the figure, none of the mass mapping methods except KARMMA reproduce the correlation functions of the DES-Y3 weak lensing data.

KARMMA produces full posterior distributions of the mass maps, which is particularly important if these maps are to be used for cross-correlation studies (e.g., [56]).

Figure 10 compares the mean of our posterior maps for the lowest tomographic redshift bin to the publicly available mass maps released by the DES collaboration [47]. Specifically, we compare our mean mass map to their null B-mode prior Kaiser-Squires method, Wiener filter method and the wavelet transform based GLIMPSE method [57]. We find that KARMMA can resolve smaller structures than any of the other methods. This improvement is driven by the fact that KARMMA self-consistently accounts for the covariance across all tomographic redshift bins, which, as noted

earlier, allows us to extract information on the low-redshift density field from the high-redshift shear data [12].

In Fig. 11, we compare the correlation function and the pseudo  $C(\ell)$  of the DES-Y3 mass maps reconstructed using KARMMA and the same publicly available mass maps used in Fig. 10. In that figure, we also show the correlation functions and the pseudo- $C(\ell)$ s measured by the DES collaboration with black dots.<sup>4</sup> As we can see from the figure, barring KARMMA, none of the other methods recover

<sup>4</sup>DES-Y3 data vectors are available at <https://des.ncsa.illinois.edu/releases/y3a2/Y3key-products> and the pseudo- $C(\ell)$  measurements are taken from [46].

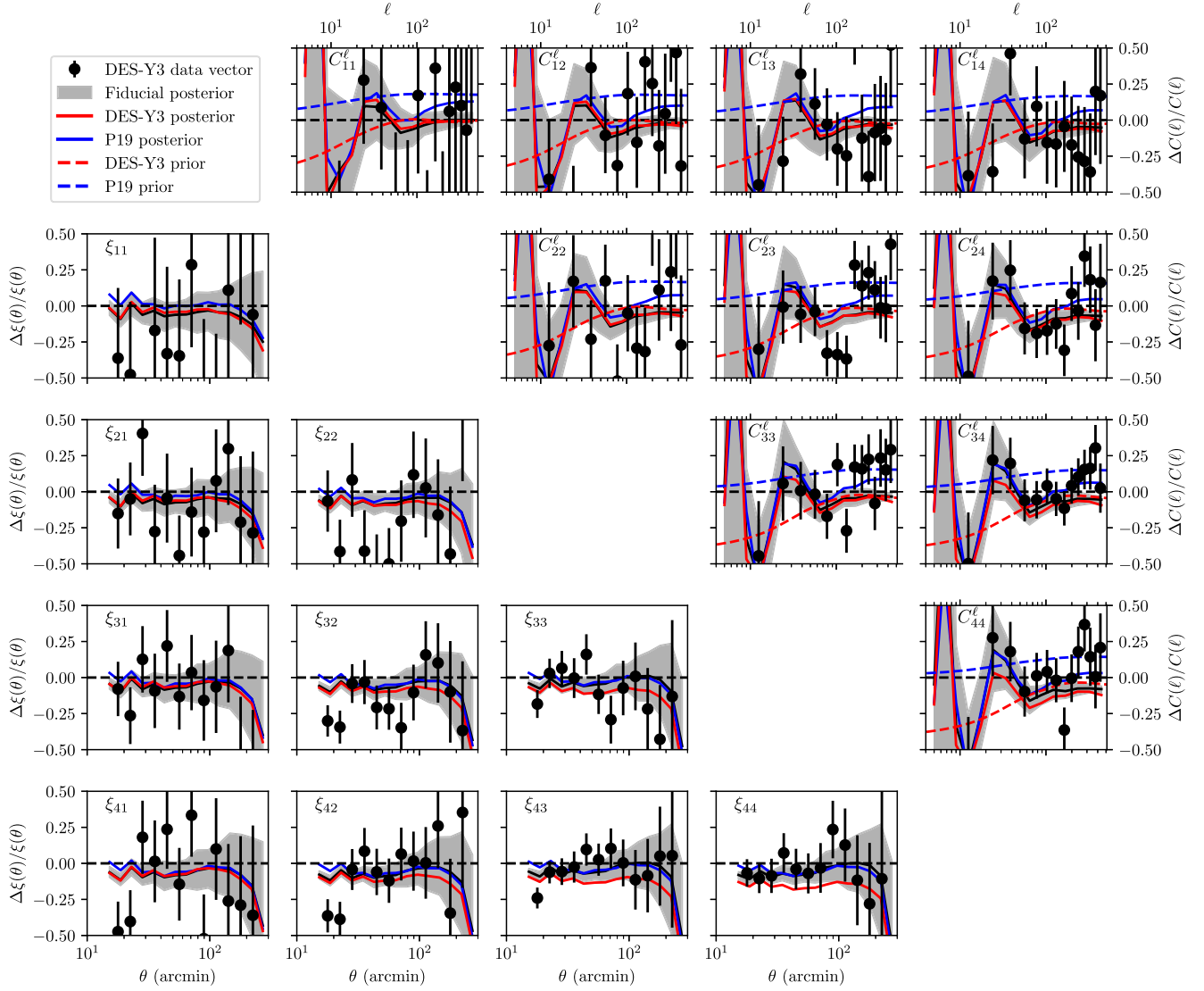


FIG. 12. The fractional difference in the recovered correlation function and pseudo  $C(\ell)$ s for KARMMA runs with different cosmological parameters with respect to the theory predictions from T17 cosmology. The gray shaded region shows the 95% confidence interval for the KARMMA run with the fiducial cosmological parameters of T17. The blue (red) solid line shows the mean of the KARMMA posterior with Planck (DES-Y3) cosmological parameters. The blue (red) dashed lines in the  $C(\ell)$  plot shows the input  $C(\ell)$  prior for the corresponding cosmological parameters. As we can see from the figure, the  $C(\ell)$ s on large scales are not substantially impacted by the input cosmological parameters. However, the small-scale power depends on the input cosmological parameters.

the expected 2-point functions in the mass maps. This is unsurprising. Indeed, the fact that best point-estimates of the convergence map are expected to be biased further highlights the importance of being able to properly sample the posterior distribution of the maps.

We run two additional KARMMA analyses with different priors to test the sensitivity of the recovered mass maps to the input cosmological parameters. We use the best fit cosmological parameters from Planck collaboration cosmological analysis [58] and DES-Y3  $3 \times 2$  pt analysis [1]. For each of these cosmologies, we recompute the power spectrum and shift parameters. The shift parameter at different cosmologies is computed by rescaling the shift

parameter in the T17 simulations with the ratio of the shift parameter predicted using CosMomentum [59]. That is, we only rely on analytic predictions for the shift parameter for the purposes of rescaling the shift parameter from our fiducial cosmology to a new cosmology. The results of these runs are shown in Fig. 12, where we plot the fractional difference of the recovered correlation function and pseudo- $C(\ell)$ s for the KARMMA runs with different cosmological parameters with respect to the T17 2-point functions. As we can see from the figure, the power spectrum on large scales is largely insensitive to the input prior. On large scales, the signal-to-noise ratio is high and therefore the large-scale modes are determined by the

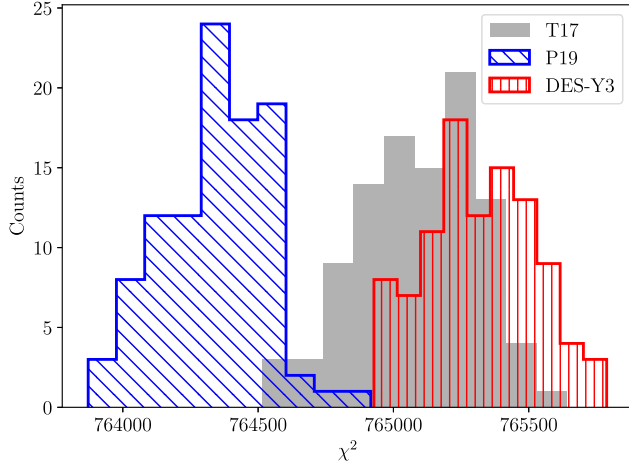


FIG. 13.  $\chi^2$  distribution from the KARMMA posteriors with different cosmological parameters. The posteriors with DES-Y3 cosmological parameters (red vertical lines) have the highest  $\chi^2$ , followed by the T17 (gray histograms) and P19 (blue diagonal lines). The fact that these distributions are clearly differentiated suggests a field-level cosmological analysis of the DES shear maps may be able to distinguish between the Planck and DES-Y3 cosmologies.

likelihood. By contrast, the signal-to-noise ratio on small scales is low, which renders the posterior at small scales sensitive to the adopted prior.

In Fig. 13 we compare the distribution of  $\chi^2$  values computed using Eq. (10) for each map in the KARMMA posterior. Interestingly, we see that the KARMMA maps produced using a Planck cosmology provide a better fit for the DES-Y3 data than the DES-Y3 or T17 cosmological parameters. However, we caution against interpreting this result as evidence for a Planck cosmology since we did not account for observational systematics in our analysis. Rather, we want to highlight the difference in the likelihood between the Planck and DES-Y3 posteriors suggests that a field-level analysis of the DES-Y3 data may have the potential to distinguish between these different cosmologies.

## VI. CONCLUSION

The KARMMA algorithm [21,31] models the convergence field as a lognormal random field. By doing so, KARMMA can perform a forward-modeled Bayesian reconstruction of the density field of the Universe as constrained by cosmic shear data. In this paper, we have extended the KARMMA framework so that it can simultaneously forward model multiple tomographic redshift bins while fully accounting for their expected covariance. This enhancement significantly improves the signal-to-noise ratio of the reconstructed maps, particularly in the low signal-to-noise regime (see Sec. IV and [12]). We validated our method on simulated cosmic shear data generated using  $N$ -body simulations, and demonstrated that the KARMMA posteriors accurately

reproduce a variety of Gaussian and non-Gaussian statistics of the input simulated maps (see Figs. 7 and 8). However, we do find evidence of small  $\lesssim 5\%$ – $10\%$  biases which increase with decreasing redshifts. These biases arise due to the failure of the log-normal model to correctly capture non-linear structure formation. These results suggest that further improvements require improving the field-level prior assumed for the convergence field. Some options for such an improvement include analytic extensions such as the double log model of [8] or using generative artificial intelligence models [60–63].

Following our simulation tests, we applied KARMMA to DES-Y3 weak lensing data to produce the first Bayesian forward-modeled tomographic mass maps from stage-III weak lensing data on a sphere. These maps are also the first to self-consistently account for the covariance between tomographic redshift bins as part of the mass map reconstruction algorithm, which in turn improved the signal-to-noise of the resulting maps, particularly at low redshifts. We show that these maps have the correct theoretically expected statistical properties such as the 2-point function [34].

An important limitation in our work is the relatively coarse angular resolution of our lensing maps ( $\approx 13$  arcmin). Since the information on small scales is highly non-Gaussian, one might expect that field-level inference will become even more useful at higher resolution. However, this leads to increased computational demands due to (1) spherical harmonic transforms at a higher resolution and (2) a larger parameter space. Significantly improving resolution over that achieved in this work will require improved computational spherical harmonics code and faster sampling methods. Improved spherical harmonic transform codes have been produced using GPU acceleration (e.g., [64]) or spherical Fourier neural operators [65]. Improvement in sampling can be achieved using improved sampling methods such as microcanonical Hamiltonian Monte Carlo methods [66,67].

Despite this limitation, the inclusion of tomographic mass mapping on a sphere represents an important step towards enabling a full field-level cosmological analysis of cosmic shear data. In addition to varying cosmology, our future work will focus on incorporating weak lensing systematics as part of the inference process. The field-level approach is particularly well suited for incorporating some of these systematics. For example, it has been shown that the non-Gaussian clustering of galaxies can improve the inference of photometric redshifts [68,69]. Likewise, Bayesian photo- $z$  inference algorithms [70–73] can naturally be extended to incorporate density field inference to further improve photometric redshift estimation. Similarly, we will need to model intrinsic alignments at the field level in order to properly extract information using our forward modeling framework [74]. Fortunately, we see no reason why any of these challenges should be prohibitive,

suggesting that field-based inference of cosmological parameters from cosmic shear data could soon be realized.

### ACKNOWLEDGMENTS

We thank Marco Gatti for providing the DES-Y3 shape catalog, and for useful comment on the manuscript that helped improved the clarity of our presentation. The computation presented here was performed on the High Performance Computing (HPC) resources supported by the University of Arizona Technology and Research Initiative Fund, University Information Technology Services, and Research, Innovation, and Impact (RII) and maintained by the UArizona Research Technologies Department. S. S. B. was supported by the Department of Energy Cosmic Frontier program, Grant No. DE-SC0020215, and NSF Grant No. 2009401. E. R.'s work is supported is supported by NSF Grant No. 2009401.

E. R. also receives funding from DOE Grant No. DE-SC0009913 and NSF Grant No. 2206688.

### APPENDIX: TESTS WITH LOG-NORMAL MOCKS

In Fig. 14, we show the fractional error in the correlation function, pseudo- $C(\ell)$ s and the 1-pt PDF in the KARMMA posterior maps on runs with log-normal mocks. For the ease of visualization, we only show the results for the first redshift bins, where we saw in Sec. IV that the uncertainty was the largest. As we can see from the figure, we do not see the  $\sim \mathcal{O}(5\%)$  bias in the 2-pt functions or the even larger bias in the 1-pt PDF. This demonstrates that the biases we see in our runs with  $N$ -body simulations arise because of model misspecification due to the assumed log-normal prior.

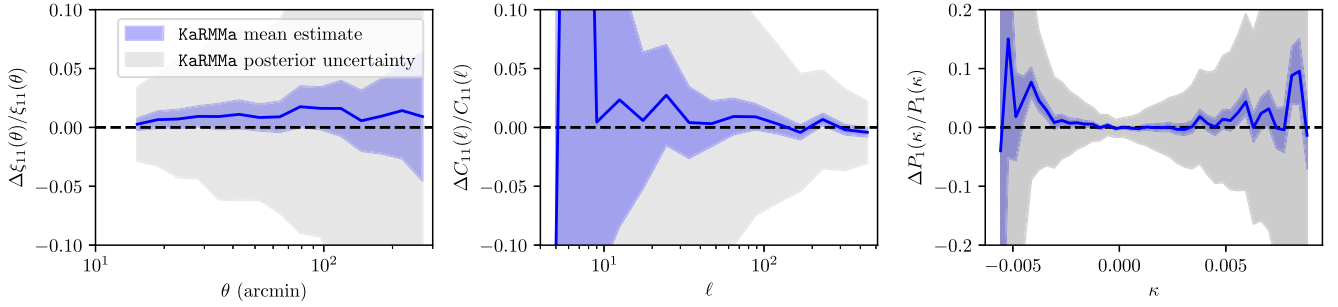


FIG. 14. Fractional error in the correlation function (left), pseudo- $C(\ell)$  (center) and the 1-pt PDF (right) of the first redshift bin of the KARMMA posteriors for runs on log-normal mocks. As we can see, we do not see the large biases in these summary statistics that we saw for the KARMMA runs on  $N$ -body simulations in Sec. IV. Thus, it shows that the biases seen in Sec. IV and Figs. 7 and 8 arises due to the model misspecification due to the log-normal prior.

- 
- [1] DES Collaboration, *Phys. Rev. D* **105**, 023520 (2022).
  - [2] eBOSS Collaboration, *Phys. Rev. D* **103**, 083533 (2021).
  - [3] M. Millea, E. Anderes, and B. D. Wandelt, *Phys. Rev. D* **100**, 023509 (2019).
  - [4] M. Millea, E. Anderes, and B. D. Wandelt, *Phys. Rev. D* **102**, 123542 (2020).
  - [5] M. Millea *et al.*, *Astrophys. J.* **922**, 259 (2021).
  - [6] A. J. Zhou and S. Dodelson, *Phys. Rev. D* **108**, 083506 (2023).
  - [7] J. Jasche and G. Lavaux, *Astron. Astrophys.* **625**, A64 (2019).
  - [8] L. Porth, G. M. Bernstein, R. E. Smith, and A. J. Lee, *Mon. Not. R. Astron. Soc.* **518**, 3344 (2023).
  - [9] S. S. Boruah, G. Lavaux, and M. J. Hudson, *Mon. Not. R. Astron. Soc.* **517**, 4529 (2022).
  - [10] J. Prideaux-Ghee, F. Leclercq, G. Lavaux, A. Heavens, and J. Jasche, *Mon. Not. R. Astron. Soc.* **518**, 4191 (2023).
  - [11] A. E. Bayer, C. Modi, and S. Ferraro, *J. Cosmol. Astropart. Phys.* **06** (2023) 046.
  - [12] S. S. Boruah, E. Roza, and P. Fiedorowicz, *Mon. Not. R. Astron. Soc.* **516**, 4111 (2022).
  - [13] N. Porqueres, A. Heavens, D. Mortlock, and G. Lavaux, *Mon. Not. R. Astron. Soc.* **509**, 3194 (2022).
  - [14] A. Junzhe Zhou, X. Li, S. Dodelson, and R. Mandelbaum, *arXiv:2312.08934*.
  - [15] Euclid Collaboration, *Astron. Astrophys.* **675**, A120 (2023).
  - [16] A. Kovács *et al.*, *Mon. Not. R. Astron. Soc.* **510**, 216 (2022).
  - [17] J. Liu and J. C. Hill, *Phys. Rev. D* **92**, 063517 (2015).
  - [18] A. Hojjati *et al.*, *Mon. Not. R. Astron. Soc.* **471**, 1565 (2017).
  - [19] N. Kaiser and G. Squires, *Astrophys. J.* **404**, 441 (1993).
  - [20] B. Mawdsley *et al.* (DES Collaboration), *Mon. Not. R. Astron. Soc.* **493**, 5662 (2020).

- [21] P. Fiedorowicz, E. Rozo, S. S. Boruah, C. Chang, and M. Gatti, *Mon. Not. R. Astron. Soc.* **512**, 73 (2022).
- [22] N. Porqueres, A. Heavens, D. Mortlock, and G. Lavaux, *Mon. Not. R. Astron. Soc.* **502**, 3035 (2021).
- [23] J. Jasche and B. D. Wandelt, *Mon. Not. R. Astron. Soc.* **432**, 894 (2013).
- [24] J. Alsing, A. Heavens, A. H. Jaffe, A. Kiessling, B. Wandelt, and T. Hoffmann, *Mon. Not. R. Astron. Soc.* **455**, 4452 (2016).
- [25] J. Alsing, A. Heavens, and A. H. Jaffe, *Mon. Not. R. Astron. Soc.* **466**, 3272 (2017).
- [26] A. Loureiro, L. Whiteaway, E. Sellentin, J. S. Lefaurie, A. H. Jaffe, and A. F. Heavens, *Open J. Astrophys.* **6** (2023), 10.21105/astro.2210.13260.
- [27] E. Sellentin, A. Loureiro, L. Whiteaway, J. S. Lefaurie, S. T. Balan, M. Olamaie, A. H. Jaffe, and A. F. Heavens, *arXiv:2305.16134*.
- [28] Y. Hoffman and E. Ribak, *Astrophys. J. Lett.* **380**, L5 (1991).
- [29] S. Zaroubi, Y. Hoffman, K. B. Fisher, and O. Lahav, *Astrophys. J.* **449**, 446 (1995).
- [30] D. Kodi Ramanah, G. Lavaux, and B. D. Wandelt, *Mon. Not. R. Astron. Soc.* **490**, 947 (2019).
- [31] P. Fiedorowicz, E. Rozo, and S. S. Boruah, *arXiv:2210.12280*.
- [32] H. S. Xavier, F. B. Abdalla, and B. Joachimi, *Mon. Not. R. Astron. Soc.* **459**, 3693 (2016).
- [33] L. Clerkin *et al.*, *Mon. Not. R. Astron. Soc.* **466**, 1444 (2017).
- [34] The code is publicly available at <https://github.com/Supranta/KaRMMA.git>. KARMMA DES year 3 mass maps are publicly available at <https://zenodo.org/records/10672062>.
- [35] R. Takahashi, T. Hamana, M. Shirasaki, T. Namikawa, T. Nishimichi, K. Osato, and K. Shiroyama, *Astrophys. J.* **850**, 24 (2017).
- [36] A. Amon (DES Collaboration), *Phys. Rev. D* **105**, 023514 (2022).
- [37] L. F. Secco (DES Collaboration), *Phys. Rev. D* **105**, 023515 (2022).
- [38] J. Myles *et al.*, *Mon. Not. R. Astron. Soc.* **505**, 4249 (2021).
- [39] M. Gatti *et al.*, *Mon. Not. R. Astron. Soc.* **504**, 4312 (2021).
- [40] A. Lewis, A. Challinor, and A. Lasenby, *Astrophys. J.* **538**, 473 (2000).
- [41] J. Lesgourgues, *arXiv:1104.2934*.
- [42] M. Gatti (DES Collaboration), *Mon. Not. R. Astron. Soc.* **527**, L115 (2024).
- [43] E. Huff and R. Mandelbaum, *arXiv:1702.02600*.
- [44] E. S. Sheldon and E. M. Huff, *Astrophys. J.* **841**, 24 (2017).
- [45] A. Nicola, C. García-García, D. Alonso, J. Dunkley, P. G. Ferreira, A. Slosar, and D. N. Spergel, *J. Cosmol. Astropart. Phys.* **03** (2021) 067.
- [46] C. Doux (DES Collaboration), *Mon. Not. R. Astron. Soc.* **515**, 1942 (2022).
- [47] N. Jeffrey, M. Gatti (DES Collaboration), *Mon. Not. R. Astron. Soc.* **505**, 4626 (2021).
- [48] C. G. R. Wallis, M. A. Price, J. D. McEwen, T. D. Kitching, B. Leistedt, and A. Plouviez, *Mon. Not. R. Astron. Soc.* **509**, 4480 (2022).
- [49] A. Paszke *et al.*, in *Advances in Neural Information Processing Systems 32, Vancouver, BC* (Curran Associates, Inc., 2019), pp. 8024–8035.
- [50] E. Bingham, J. P. Chen, M. Jankowiak, F. Obermeyer, N. Pradhan, T. Karaletsos, R. Singh, P. A. Szerlip, P. Horsfall, and N. D. Goodman, *J. Mach. Learn. Res.* **20**, 1 (2019), <https://jmlr.org/papers/v20/18-403.html>.
- [51] M. D. Hoffman and A. Gelman, *arXiv:1111.4246*.
- [52] D. Alonso, J. Sanchez, A. Slosar (LSST Dark Energy Science Collaboration), *Mon. Not. R. Astron. Soc.* **484**, 4127 (2019).
- [53] J. Liu, A. Petri, Z. Haiman, L. Hui, J. M. Kratochvil, and M. May, *Phys. Rev. D* **91**, 063507 (2015).
- [54] C. T. Davies, M. Cautun, B. Giblin, B. Li, J. Harnois-Déraps, and Y.-C. Cai, *Mon. Not. R. Astron. Soc.* **507**, 2267 (2021).
- [55] N. E. Chisari (LSST Dark Energy Science Collaboration), *Astrophys. J. Suppl. Ser.* **242**, 2 (2019).
- [56] N.-M. Nguyen, J. Jasche, G. Lavaux, and F. Schmidt, *J. Cosmol. Astropart. Phys.* **12** (2020) 011.
- [57] F. Lanusse, J. L. Starck, A. Leonard, and S. Pires, *Astron. Astrophys.* **591**, A2 (2016).
- [58] Planck Collaboration, *Astron. Astrophys.* **641**, A6 (2020).
- [59] O. Friedrich, C. Uhlemann, F. Villaescusa-Navarro, T. Baldauf, M. Manera, and T. Nishimichi, *Mon. Not. R. Astron. Soc.* **498**, 464 (2020).
- [60] P. Fiedorowicz, E. Rozo, S. Boruah, W. Coulton, S. Ho, and G. Fabbian, in *Proceedings of the 39th International Conference on Machine Learning (ICML 2022), Baltimore, Maryland* (2022), p. 8.
- [61] T. W. H. Yiu, J. Fluri, and T. Kacprzak, *J. Cosmol. Astropart. Phys.* **12** (2022) 013.
- [62] B. Dai and U. Seljak, *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2309624121 (2024).
- [63] M. Shirasaki and S. Ikeda, *arXiv:2310.17141*.
- [64] C. Tian, S. Li, and H. Liu, *Astrophys. J. Suppl. Ser.* **263**, 21 (2022).
- [65] B. Bonev, T. Kurth, C. Hundt, J. Pathak, M. Baust, K. Kashinath, and A. Anandkumar, *arXiv:2306.03838*.
- [66] J. Robnik, G. B. De Luca, E. Silverstein, and U. Seljak, *arXiv:2212.08549*.
- [67] A. E. Bayer, U. Seljak, and C. Modi, *arXiv:2307.09504*.
- [68] J. Jasche and B. D. Wandelt, *Mon. Not. R. Astron. Soc.* **425**, 1042 (2012).
- [69] E. Tsaprazi, J. Jasche, G. Lavaux, and F. Leclercq, *arXiv:2301.03581*.
- [70] B. Leistedt, D. J. Mortlock, and H. V. Peiris, *Mon. Not. R. Astron. Soc.* **460**, 4258 (2016).
- [71] B. Leistedt, J. Alsing, H. Peiris, D. Mortlock, and J. Leja, *Astrophys. J. Suppl. Ser.* **264**, 23 (2023).
- [72] C. Sánchez and G. M. Bernstein, *Mon. Not. R. Astron. Soc.* **483**, 2801 (2019).
- [73] A. Alarcon, C. Sánchez, G. M. Bernstein, and E. Gaztañaga, *Mon. Not. R. Astron. Soc.* **498**, 2614 (2020).
- [74] E. Tsaprazi, N.-M. Nguyen, J. Jasche, F. Schmidt, and G. Lavaux, *J. Cosmol. Astropart. Phys.* **08** (2022) 003.