

## Accurate field-level weak lensing inference for precision cosmology

Alan Junzhe Zhou<sup>✉,\*</sup>, Xiangchong Li<sup>✉</sup>, Scott Dodelson<sup>✉</sup>, and Rachel Mandelbaum<sup>✉</sup>

*Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA;  
McWilliams Center for Cosmology, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA;  
and NSF AI Planning Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA*



(Received 12 January 2024; accepted 15 May 2024; published 22 July 2024)

We present MIKO, a catalog-to-cosmology pipeline for general flat-sky field-level inference, which provides access to cosmological information beyond the two-point statistics. In the context of weak lensing, we identify several new field-level analysis systematics (such as aliasing, Fourier mode-coupling, and density-induced shape noise), quantify their impact on cosmological constraints, and correct the biases to a percent level. Next, we find that model misspecification can lead to both absolute bias and incorrect uncertainty quantification for the inferred cosmological parameters in realistic simulations. The Gaussian map prior infers unbiased cosmological parameters, regardless of the true data distribution, but it yields overconfident uncertainties. The log-normal map prior quantifies the uncertainties accurately, although it requires careful calibration of the shift parameters for unbiased cosmological parameters. We demonstrate systematics control down to the 2% level for both models, making them suitable for ongoing weak lensing surveys.

DOI: [10.1103/PhysRevD.110.023539](https://doi.org/10.1103/PhysRevD.110.023539)

### I. INTRODUCTION

A cosmological field-level inference (FLI) compares the field-level observables (such as galaxy density [1–7], cosmic shear [8–12], or cosmic microwave background (CMB) maps [7,13,14]) directly to theoretical predictions or simulations. It treats the map pixels and cosmological parameters as random variables to be jointly modeled, with map-making and cosmological parameter inference being carried out simultaneously in a Bayesian manner.

Compared to traditional two-point analyses, FLI has several advantages:

- (i) In principle, it utilizes all the  $N$ -point statistics in the data, so there is minimal information loss.
- (ii) It is built around a conceptually simple forward model which easily extends to new physics and systematics.
- (iii) It combines the map-making and parameter inference into a single step, enabling joint inference of multiple surveys on the map level.
- (iv) It avoids the needs to compute complex high-dimensional covariance matrices.

Previous works [12,15,16] have shown that FLI can achieve superior cosmological constraints compared to two-point analyses. The goal of this work is to establish FLI methods with rigorously demonstrated systematics control at least at the level needed for current imaging surveys.

\*junzhez@andrew.cmu.edu

$$\mathcal{P}(\mathbf{c}, \mathbf{s}|\mathbf{d}) \propto G(\mathbf{d} - \mathbf{s}; 0, N) \times \mathcal{P}(\mathbf{s}|\mathbf{c}) \times \text{prior}(\mathbf{c}). \quad (1)$$

Here, the set of cosmological and astrophysical parameters are denoted  $\mathbf{c}$ ; the map pixel values (the *signal*,  $\mathbf{s}$ ; the data pixel values  $\mathbf{d}$ ; and  $G(\mathbf{x}; \mu, N)$  is a Gaussian distribution of  $\mathbf{x}$  with mean  $\mu$  and variance  $N$ . We will treat the noise as uncorrelated so that the matrix  $N$  is diagonal (in the case of lensing, related to the shape noise). The second probability on the right encodes two broad areas of complexity. First, even in the simplest case, where the observations are on the largest of scales so the fundamental fields are Gaussian, care needs to be taken to convert the true field to  $\mathbf{s}$ , the field that will be compared to the data [17]. We call this set of issues *analysis systematics*, and present our treatment of them in Sec. V below. Second, most observations are not sampling purely Gaussian fields, so understanding the distribution from which the true fields are drawn is paramount. We call this uncertainty *model misspecification* and discuss this in Sec. VI below. This paper aims to make two main points:

- (1) We have developed a pipeline to analyze weak lensing data in the flat sky case<sup>1</sup> that includes and

<sup>1</sup>The flat sky case has the advantage of requiring less computational resources and also applicable to deep surveys that cover only a fraction of the sky. Indeed, we will test our methods on realistic mocks that simulate the Hyper Suprime-Cam SSP Survey [18] year-3 shape catalog [19], and the pipeline developed here may be applicable to the Roman Space Telescope [20], current baseline wide-field imaging survey will cover about 4% of the sky.

quantifies analysis systematics such as pixelization effects, boundary conditions, and general issues that arise when moving between real and Fourier space.

- (2) In weak lensing, the choice of a Gaussian prior for  $\mathcal{P}(\mathbf{s}|\mathbf{c})$  leads to unbiased means on the extracted parameters but incorrect error bars. Meanwhile, using a log-normal prior for  $\mathcal{P}(\mathbf{s}|\mathbf{c})$  leads to correct errors bar but the means of the extracted parameters are very sensitive to the exact parametrization of the log-normal distribution.

Looking forward (Sec. IV), we will constrain the tomographic power spectrum amplitudes  $A_i$ , which is similar to a tomographically decomposed version of  $S_8$  in standard cosmological analyses. The HSC Year 3 real [21] and Fourier [22] space shear analyses both yield an approximately 4.1% constraints on  $S_8$  ( $0.769^{+0.031}_{-0.034}$  and  $0.776^{+0.032}_{-0.033}$  respectively). This translates to approximately an 8% constraint on  $A_i$ . Since we typically require the systematic uncertainty to be under 1/4 of the total error budget, we set a target of 2% for the absolute bias induced by systematic effects. Given the above numbered points, we chart out a pathway for analyses that overcomes these difficulties at this target 2% level.

This paper is organized as follows. We briefly review the basics of weak lensing in Sec. II, describe the simulated datasets in Sec. III and then give an overview of the inference pipeline in Sec. IV. The first set of results is presented in Sec. V, which deals with analysis systematics, and the second in Sec. VI, where model misspecification is discussed. We summarize and outlook in Sec. VII.

## II. WEAK LENSING BASICS

### A. Weak lensing maps

As the light from distant galaxies travels toward us, it is deflected by the intervening matter overdensity perturbation  $\delta_m$ . Consider a line-of-sight (LOS) in the  $\boldsymbol{\theta}$  direction and a light source at comoving distance  $\chi$ . The light source, although emitted at the true position  $\boldsymbol{\alpha}$ , is observed at  $\boldsymbol{\beta}$  due to gravitational lensing. In the weak lensing regime, the effect of this distortion is approximately linear

$$\left. \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\alpha}} \right|_{\boldsymbol{\alpha}, \chi} = \begin{pmatrix} 1 - \kappa(\boldsymbol{\theta}, \chi) - \gamma_1(\boldsymbol{\theta}, \chi) & -\gamma_2(\boldsymbol{\theta}, \chi) \\ -\gamma_2(\boldsymbol{\theta}, \chi) & 1 - \kappa(\boldsymbol{\theta}, \chi) + \gamma_1(\boldsymbol{\theta}, \chi) \end{pmatrix}, \quad (2)$$

where  $\kappa(\boldsymbol{\theta}, \chi)$  and  $\gamma(\boldsymbol{\theta}, \chi) = \gamma_1(\boldsymbol{\theta}, \chi) + i\gamma_2(\boldsymbol{\theta}, \chi)$  are called the convergence and the shear maps [23].

In a tomographic weak-lensing survey, we categorize the source galaxies into redshift bins with normalized galaxy density distributions  $n_{\text{gal},i}(z)$ . We are interested in the line-of-sight (LOS)-averaged convergence map, which, for a spatially flat universe, is related to the matter overdensity via

$$\kappa_i(\boldsymbol{\theta}) = \int_0^{\chi^*} d\chi W_i(\chi) \delta_m(\chi \boldsymbol{\theta}, \chi), \quad (3)$$

$$W_i(\chi) = \frac{3}{2} H_0^2 \Omega_m (1 + z(\chi)) \int_{\chi}^{\chi^*} d\chi' n_{\text{gal},i}(\chi') \frac{\chi(\chi' - \chi)}{\chi'}, \quad (4)$$

where  $H_0$  is the Hubble constant,  $\Omega_m = 8\pi G \rho_m / (3H_0^2)$ ,  $\rho_m$  is the matter density, and  $\chi^*$  is the comoving horizon.

### B. Weak lensing statistics

The  $n$ -point correlation functions of the convergence field carry significant cosmological information. The two-point cross-power spectrum between redshift bins  $i$  and  $j$  is given by (under Limber's approximation [24])

$$C_{ij}(\ell) = \int_0^{\chi^*} d\chi \frac{W_i(\chi) W_j(\chi)}{\chi} P_{\delta} \left( k = \frac{\ell}{\chi}; \chi \right), \quad (5)$$

where  $P_{\delta}(k; \chi)$  is the 3-dimensional matter power spectrum. In practice, we compute Eq. (5) using CAMB [25].

If the convergence field were completely Gaussian, then Eq. (5) would contain all the information. In reality, however, small-scale structure formation leads to non-Gaussianity characterized by extended voids and density peaks that contains important cosmological information [26,27]. For example, Fig. 1 shows three examples of convergence fields generated with models (Gaussian, log-normal, and realistic  $N$ -body ray-tracing simulation). Although these fields all share the same power spectrum, they are visibly different, each with its own one-point PDF and higher-order statistics. In this work, we model the field-level statistics to extract information not captured by the standard two-point analyses.

One of the most studied models of the convergence PDF is the log-normal distribution [9,28–30]. Let  $\kappa^G$  be a zero-mean Gaussian (transfer) field. The associated zero-mean log-normal field is defined by applying a transformation across the entire Gaussian field,

$$\kappa = \mathbb{L}(\kappa^G, a) = a(e^{\kappa^G - \sigma^G/2} - 1), \quad (6)$$

where  $a$  is the shift parameter and  $\sigma^G$  is the standard deviation of the transfer map. The log-normal field is fully characterized by its variance  $\sigma^2$  and shift parameter, giving it one more degree of freedom than the Gaussian field. The shift parameter  $a$  is particularly interesting, as it corresponds to the minimum value of the log-normal distribution. Alternatively,  $a$  can be reparametrized in terms of the skewness  $s$  [30] by

$$a = \frac{\sigma}{s} (1 + y(s) + y^{-1}(s)), \quad (7)$$

$$y(s) = \left( \frac{2 + s^2 + s\sqrt{4 + s^2}}{2} \right)^{1/3}, \quad (8)$$

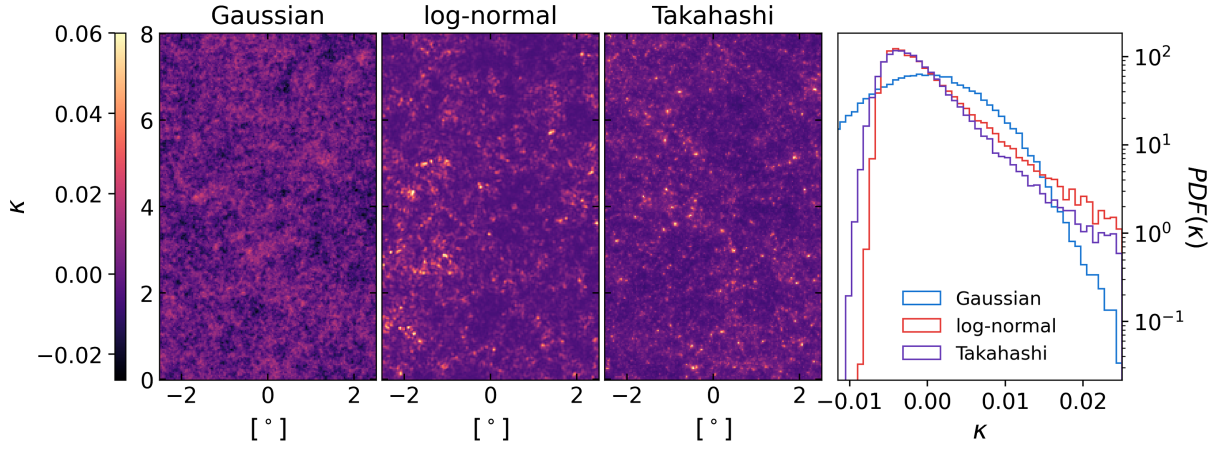


FIG. 1. Examples of Gaussian, log-normal, and Takahashi fields. All three fields have the same two-point statistics (modeled after the first redshift bin Sec. III) but a visibly different field-level appearance. Their one-point PDFs are shown in the right plot, where we see that the log-normal field has a stronger cut-off at small  $\kappa$  compared to the others. We limit the maximum color map scale to  $\kappa = 0.06$  to emphasize the visual distinction between the fields.

$$s = \langle (\kappa/\sigma)^3 \rangle, \quad (9)$$

which we find to be a more robust measurement on real simulations. The assumption about  $a$  is very important and we will discuss it in Sec. VI A 3.

Finally, we will need to relate the two-point correlation function of the transfer field to that of the log-normal field. This is given by

$$\xi_{ij}^G(\theta) = \log \left( \frac{\xi_{ij}(\theta)}{a_i a_j} + 1 \right). \quad (10)$$

The two-point correlation function is related to the power spectrum by the classic Hankel transform

$$\xi_{ij}(\theta) = \mathcal{H}^{-1}[C_{ij}(\ell)](\theta) \quad (11)$$

$$= 2\pi \int d\theta \theta C_{ij}(\ell) J_0(\ell\theta), \quad (12)$$

where  $J_0$  is the Bessel function of the first kind.

### C. Weak lensing observables

A weak lensing survey measures the shapes of galaxies to infer the convergence and/or the shear [31]. For a galaxy with observed complex shear  $\gamma_o$ ,

$$\gamma_o = \gamma + \gamma_I + \gamma_n. \quad (13)$$

The first term on the right hand side is the lensing shear,  $\gamma_I$  is the coherent distortion of galaxy shape due to intrinsic alignment (IA) [32,33], and  $\gamma_n$  is the galaxy shape noise caused by random intrinsic galaxy shape and image noise. In this study, we leave  $\gamma_I = 0$  and model  $\gamma_n$  as a zero-mean Gaussian random variable with standard deviation  $\gamma_{\text{rms}}$ .

Additionally, higher-order differences between reduced shear and shear [34] are not considered in our analysis.

## III. SIMULATIONS AND DATASETS

We model our analysis after the HSC Year 3 survey specifications [19,35]. We use three sets of independent simulations to test our analysis pipeline across the entire galaxy catalog to cosmological constraints process. In this section, we discuss the process of converting galaxy catalog to the observed pixelized shear maps  $\gamma_i^{\text{obs}}$ . We will show in Sec. V that this procedure introduces significant statistical artefacts in  $\gamma_i^{\text{obs}}$  that must be reflected in our forward model for unbiased cosmological constraints.

The first simulation comes from Takahashi *et al.* [36], where the authors perform multiple-lens plane ray-tracing on high-resolution N-body simulation. This simulation uses the WMAP 9 years result [37] as its fiducial cosmology, which we also adopt through out our work. The Takahashi simulation has been extensively tested and used in HSC to derive covariance of 2-point summary statics [38]. It provides independent full sky realizations of  $\kappa(\boldsymbol{\theta}, \chi)$  and  $\gamma(\boldsymbol{\theta}, \chi)$  at the 38 comoving radial shells. These maps are provided in HEALPIX, format [39,40] with a resolution of NSIDE = 8192. We use the radial shells to construct the 4 tomographic  $\kappa_i(\boldsymbol{\theta})$  and  $\gamma_i(\boldsymbol{\theta})$  maps using the HSC Year 3 redshift distribution (Fig. 2). Next, we locate rectangular patches of side lengths  $(L_x, L_y) = (8^\circ, 5^\circ)$  and the patches are separated by at least  $8^\circ$  to avoid spatial correlations. We generate galaxies with tomographic effective number density  $\rho_i$  with true  $\kappa$  and  $\gamma$  values according to their positions.

We set  $\rho_i$  to be 36.1, 52.0, 39.6, and 20.6  $\text{arcmin}^{-2}$  for the four redshift bins, corresponding to 10 times the HSC Year 3 W12H field specification (explained below). Then, we add Gaussian shape noise to individual galaxies

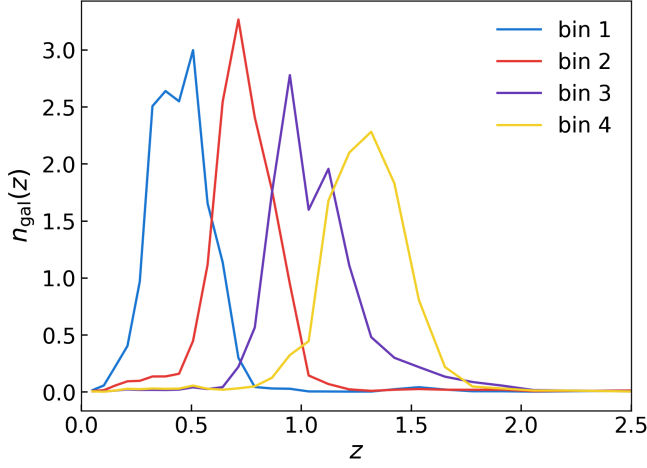


FIG. 2. The normalized PDF of the tomographic redshift distributions of the HSC Year 3 galaxy shape catalog [35].

according to Eq. (13) with  $\gamma_{\text{rms}} = 0.26$  [41]. We construct the pixelized maps  $\kappa_i^{\text{true}}$  and  $\gamma_i^{\text{true}}$  using the noiseless catalog, and construct  $\gamma_i^{\text{obs}}$  using the noisy catalog. In both cases, the pixels have size  $\Delta = 3'$ .

The choice of  $\rho_i$  merits a few comments. For a pixel  $p$ , the effective shape noise has variance

$$N_i(p) = \frac{\gamma_{\text{rms}}^2}{\rho_i(p)\Delta^2}. \quad (14)$$

We are mostly interested in the bias of modeling choices and the robustness of our posterior estimation in the presence of noise and cosmic variance. Thus, given a finite computational resource, we want to reduce the statistical uncertainty and perform as many independent HSC-like experiments as possible. We achieve this goal by setting  $\rho_i$  to 10 times the fiducial HSC number density [19] to reduce the effective shape noise.

Due to simulation artefacts [36], the Takahashi convergence maps does not recover the theoretical power spectrum [Eq. (5)] with the precision required for this work. Therefore, we need calibrate the theory when performing inference on the Takahashi mocks. The exact procedure is detailed in the Appendix. The calibrated theoretical power spectrum agrees with the simulation within 1% for  $\ell < 24000$  for all redshift bins.

The second and third sets of simulations are conceptually similar. Using the same WMAP9 cosmology, we make full sky log-normal and Gaussian realizations of convergence and shear at resolution of NSIDE = 8192, and then construct shear catalogs and flat sky maps as before. We generate the log-normal maps using FLASK [30]. For these maps, we observe a larger-than-1% surplus of power above  $\ell = 5000$ . Therefore, we only use  $\ell \leq 5000$  for both the log-normal maps and the log-normal theory power in the analyses.

For all three simulations, we use 4 independent full-sky simulations to generate 80 nonoverlapping flat-sky patches.

#### IV. INFERENCE FRAMEWORK

Our inference framework MIKO, graphically illustrated in Fig. 3, is a hierarchical Bayesian network that forward models from the cosmological parameter to the observed noisy fields. We focus on the lensing convergence and shear fields in this work, but the framework is set up to include galaxy density and CMB fields as well [7]. The weak lensing pipeline includes two models; the first generates the observable fields using a Gaussian map prior, and the second uses a log-normal prior.

MIKO has four main components. Starting from the top of the flow chart, we first draw cosmological parameters  $\mathbf{c}$  and

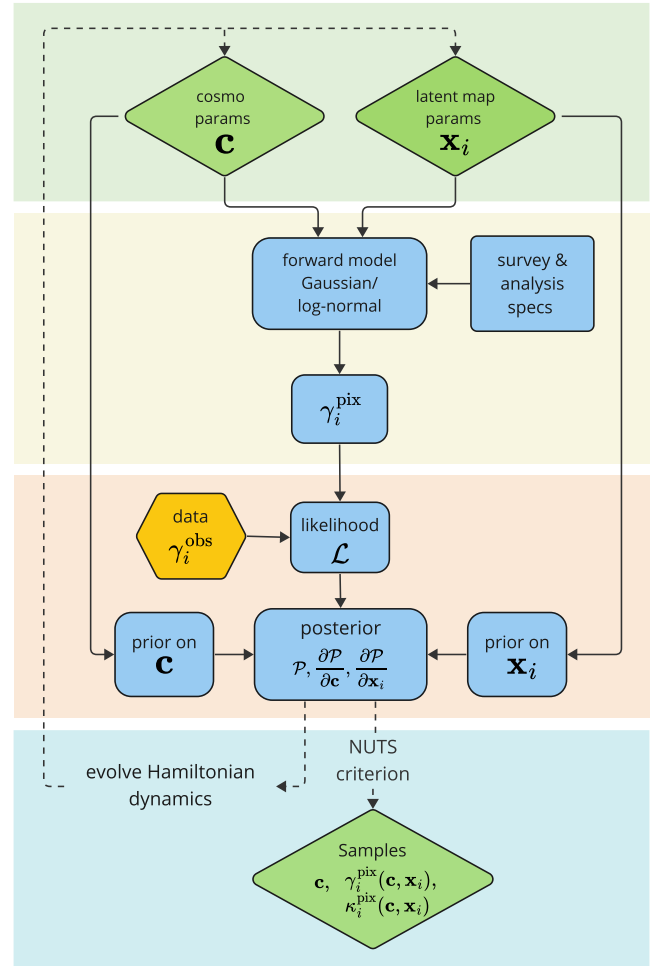


FIG. 3. This flow chart describes MIKO's weak lensing field level forward model and inference pipeline (Sec. IV). The pipeline has four steps (from top to bottom): generating latent parameters, forward modeling the noiseless fields, computing the likelihood given the data, and sampling. In the chart, diamonds represent sampled parameters, squares the model-relevant functions, and the hexagon the observed data vector.

the latent tomographic map parameters  $\mathbf{x}_i$  from their prior distributions. For the Gaussian model, we set the free cosmological parameters to be the amplitudes of the tomographic power spectra,

$$C_{ij} \longrightarrow A_i A_j C_{ij}. \quad (15)$$

where  $C_{ij}$  is the power spectrum at the fiducial cosmology calibrated to the simulations<sup>2</sup> (Sec. III). Inference of  $A_i$  is a powerful test of the  $\Lambda$ CDM model. Consider the redshift-modulated growth function  $A(z)D(z)$ . In this context,  $A_i$  can be thought of as  $A(z)$  integrated over the tomographic window function  $W_i(z)$ . Thus, a detection of  $A_i \neq 1$  implies that the growth history near the redshift interval of bin  $i$  deviates from the  $\Lambda$ CDM model. Through out this study, we use a broad and uniform prior on  $A_i$  over the interval [0.5, 1.5]. For the log-normal model, the free parameters are both  $A_i$  and the shift parameters  $a_i$ . The prior on  $a_i$  is more complicated and we will delay its discussion until Sec. VI. The latent map parameters  $\mathbf{x}_i$  represent the field-level degrees of freedom and are drawn from the unit normal distribution.

Next, we use the cosmological parameters and survey and analysis specifications to define a forward model, which transforms  $\mathbf{x}_i$  to noiseless and pixelized tomographic convergence and shear maps  $\kappa_i^{\text{pix}}$  and  $\gamma_i^{\text{pix}}$  on one (or more) flat sky patch(es). It is important to note that the process of generating  $\kappa_i^{\text{pix}}$  and  $\gamma_i^{\text{pix}}$  includes not only the cosmological model but also the survey and analysis-specific systematics model. The process of going from  $\mathbf{x}_i$  to  $\gamma_i^{\text{pix}}$  is explained in detail in Sec. V). We then compare the noiseless  $\gamma_i^{\text{pix}}$  to the data  $\gamma_i^{\text{obs}}$  and build the Bayesian posterior function given by

$$\begin{aligned} & \log \mathcal{P}(\{\gamma_i^{\text{pix}}(\mathbf{c}, \mathbf{x}_i)\} | \{\gamma_i^{\text{obs}}\}) \\ &= -\sum_i \frac{(\gamma_i^{\text{obs}} - \gamma_i^{\text{pix}}(\mathbf{c}, \mathbf{x}_i))^2}{2N} + \log \text{prior}(\{\gamma_i^{\text{pix}}(\mathbf{c}, \mathbf{x}_i)\}), \end{aligned} \quad (16)$$

where  $N$  is the shape noise variance map given by Eq. (14). We implement all computations using differentiable programming in JAX.

In the last step, we connect this differentiable posterior function to the NUMPYRO [42] implementation of the Hamiltonian Monte Carlo No-U-Turn sampler (HMC NUTS) [43–45] to efficiently sample from the high-dimensional joint posterior space of both the cosmological and the map parameters. The result of the inference is comprised of the posterior samples of the cosmological parameters  $\mathbf{c}$  and the maps  $\kappa^{\text{pix}}$  and  $\gamma^{\text{pix}}$ . We could also

<sup>2</sup>In real observations, we can simply set  $C_{ij}$  to the prediction of the fiducial cosmology.

obtain samples of the continuous fields (without the pixelization effect) if desired.

We perform consistency checks on both models. This is done by first using the model to generate noisy shear maps and using the same model to infer  $A_i$  from it. We use 10 chains for each experiment for all the main analyses in this paper. Each chain is initialized at the maximum a posterior estimate calculated with a stochastic variational inference procedure parametrized with the  $\delta$ -distributions. We then run approximately 500 warm-up NUTS steps followed by 100 sample collection NUTS steps. We use a target probability of 0.7 and a maximum tree depth of 9 for our NUTS sampler. We also check the chain convergence for the posterior samples.

## V. FORWARD MODELING AND ANALYSIS SYSTEMATICS

The first main result of this study is to present a catalog-to-cosmology pipeline and demonstrate that we can control the systematic uncertainty on  $A_i$  to be within 2% (a target motivated by our arguments in Sec. I). We begin this section by laying out the algorithm that generates the maps. We then identify and catalog the systematics that significantly bias the cosmological results and provide the appropriate remedies. This section is a detailed view of the “forward model” box in Fig. 3.

### A. Map making

Let us first consider generating a set of correlated tomographic convergence maps with Gaussian priors. To do this, we first generate independent standard normal variables (which we shall call the latent map parameters)  $\tilde{\mathbf{x}}_i(\boldsymbol{\ell})$  on the Fourier grid. Next, we transform these latent parameters to the Fourier space convergence maps  $\tilde{\kappa}_i(\boldsymbol{\ell})$  by (we will denote all Fourier space quantities using  $\tilde{\cdot}$ )

$$\tilde{\kappa}_i(\boldsymbol{\ell}) = L_{ij}(\boldsymbol{\ell}) \tilde{\mathbf{x}}_j(\boldsymbol{\ell}), \quad (17)$$

where  $L_{ij}(\boldsymbol{\ell})$  is the Cholesky decomposition of  $A_i A_j C_{ij}(\boldsymbol{\ell})$ . In the trivial example with one tomographic bin and the amplitude fixed to its fiducial value ( $A = 1$ ), this would simply mean multiplying  $\mathbf{x}(\boldsymbol{\ell})$  by  $\sqrt{C(\boldsymbol{\ell})}$ , the RMS of the kappa map.

Now let us discuss the log-normal forward model. For the log-normal maps, we first need to compute the power spectrum of the Gaussian transfer map, this is given by

$$C_{ij}^G = \mathcal{H} \left[ \log \left( \frac{\mathcal{H}^{-1}[A_i A_j C_{ij}]}{a_i a_j} + 1 \right) \right] \quad (18)$$

where  $\mathcal{H}[\cdot]$  is the Hankel transform.

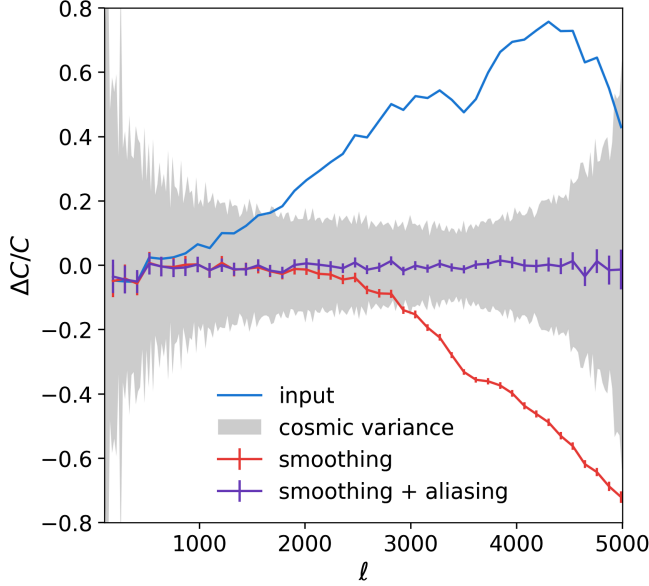


FIG. 4. The impact of smoothing and aliasing on  $C$  is presented in terms of the fractional difference relative to  $\langle C(\kappa^{\text{true}}) \rangle$ , with the gray region representing the  $1\sigma$  spread of  $C(\kappa^{\text{true}})$ . The full-sky power spectrum is shown in blue.  $\langle C(\kappa^{\text{pix}}) \rangle$ , generated by a forward model that accounts for smoothing but not aliasing, is shown in red. When both effects are included in the forward model (with aliasing parameter  $r = 1.6$ ),  $\langle C(\kappa^{\text{pix}}) \rangle$  is depicted in purple. Error bars indicate the standard error of the mean. Neglecting either the smoothing or aliasing effects, or both, in the forward model results in an order-of-unity bias on  $C$ .

We then generate a Gaussian transfer map just as we did in the Gaussian model, Fourier transform it into the real space, and apply the log-normal transformation in Eq. (6) to obtain the log-normal map.

For both the Gaussian and the log-normal model, once we obtain the  $\kappa_i$ , we compute the shear maps using the Kaiser-Squire relationship [46], which is most easily expressed in the Fourier space by

$$\tilde{\gamma}_1(\ell) = \mathbb{K}_1(\tilde{\kappa}) = \frac{\ell_x^2 - \ell_y^2}{|\ell|^2} \tilde{\kappa}(\ell) \quad (19)$$

$$\tilde{\gamma}_2(\ell) = \mathbb{K}_2(\tilde{\kappa}) = \frac{2\ell_x\ell_y}{|\ell|^2} \tilde{\kappa}(\ell) \quad (20)$$

## B. Smoothing and aliasing

The process of pixelizing galaxy catalogs into shear maps (Sec. III) introduces field-level statistical artifacts—smoothing and aliasing—in  $\gamma^{\text{obs}}$ . Both effects introduce corrections to  $C$  that are on the same order as  $C$  itself on small scales. For example, Fig. 4 shows that the full-sky theory  $C$  (blue) and the distribution of  $\kappa^{\text{true}}$  for Gaussian mocks (gray) differ by as much as 80%. An unbiased forward model should generate  $\kappa^{\text{pix}}$  such that  $\langle C(\kappa^{\text{pix}}) \rangle = \langle C(\kappa^{\text{true}}) \rangle$ . Neglecting the pixelization effect significantly biases the cosmological results. To the best of our knowledge, this is the first precise quantification and correction of the aliasing effect in the context of field-level cosmology.

Let us start by drawing a large number of galaxies from a continuous and noiseless convergence field  $\kappa$ . Mathematically, the process of map-making (Sec. III) is equivalent to first smooth  $\kappa$  with a pixel kernel and then sampling it at the center of each pixel (Fig. 5). Both effects are most easily described in the Fourier space. Smoothing is described by

$$\tilde{\kappa}^{\text{smooth}} = \text{sinc}\left(\frac{\ell_x\Delta}{2\pi}\right)\text{sinc}\left(\frac{\ell_y\Delta}{2\pi}\right)\tilde{\kappa}, \quad (21)$$

Meanwhile, the innocent-looking sampling procedure not only discretizes the Fourier space but also aliases  $\kappa$ . The net effect increases the small-scale (near the Nyquist frequency of the observed map,  $\ell_N = \pi/\Delta$ ) power of the pixelized map by an order of unity and breaks the statistical isotropy of  $\kappa$ . More precisely, pixelizing the continuous field  $\kappa$  means [47,48]

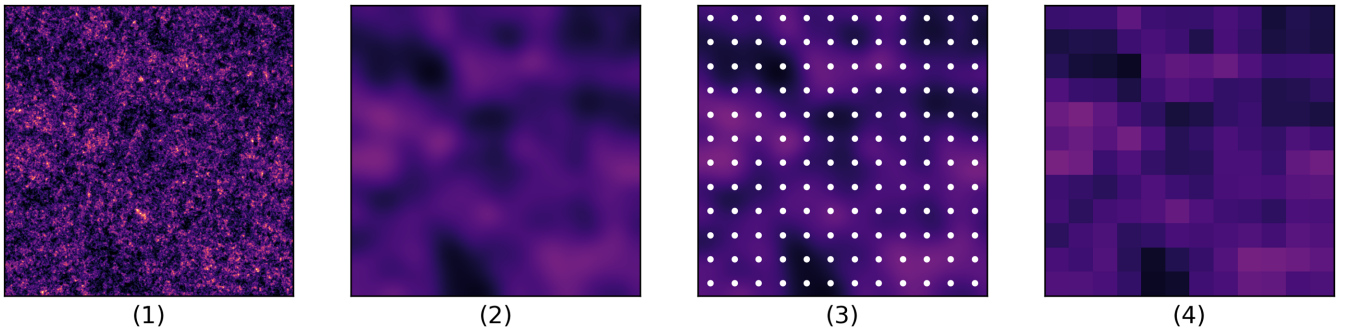


FIG. 5. Illustration of the mathematical operations implied by the map-making process (Sec. III). Plot (1) shows a continuous field. During map-making, the field is first smoothed with a pixel kernel (2). The smoothed field is then sampled at the pixel centers, indicated by white dots in (3). The values at these pixel centers constitute the pixelized map seen in (4). Both the smoothing and sampling processes introduce significant statistical artifacts into the pixelized maps, which must be accounted for in the forward model.

$$\tilde{\kappa}^{\text{alias}}(\ell) = (\tilde{\text{III}} * \tilde{\kappa})(\ell) + \underbrace{\sum_{(k_1, k_2) \neq (0,0)} (\tilde{\text{III}} * \tilde{\kappa})(\ell_x - k_1 \ell_N, \ell_y - k_2 \ell_N)}_{\text{aliases}}, \quad (22)$$

$$C^{\text{alias}}(\ell) = C(\ell) + \underbrace{\sum_{(k_1, k_2) \neq (0,0)} C_{ij}(\ell_x - k_1 \ell_N, \ell_y - k_2 \ell_N)}_{\text{aliases}}. \quad (23)$$

We define  $*$  as the (Fourier space) convolution and the Kronecker comb function as

$$\text{III}(\theta) = \sum_{n_x, n_y \in \mathbb{Z}} \delta_K(\theta_x - n_x \Delta) \delta_K(\theta_y - n_y \Delta), \quad (24)$$

where  $\delta_K(x) = 1$  when  $x = 0$  and 0 otherwise.

In short, Eq. (22) tells us that each mode of  $\tilde{\kappa}^{\text{alias}}(\ell)$  is a superposition of  $\tilde{\kappa}(\ell)$  and its ‘‘aliases’’ at higher frequencies. Equation (23) tells us that the power spectrum of the pixelized map acquires a two-dimensional  $\ell$  dependence and is no longer isotropic. Furthermore,  $C^{\text{alias}}(\ell)$  is also the superposition of  $C(\ell)$  and its aliases. One immediate consequence is that, around the Nyquist frequency,  $C^{\text{alias}}$  is at least 100% higher than  $C$ . Although Eq. (22) suggests that accounting for infinitely many aliases at increasingly larger  $\ell$ 's is necessary to fully capture aliasing, in practice, aliases above  $\sqrt{2}r\ell_N$  with  $r = 1.6$  are sufficiently suppressed by the smoothing kernel.

We implement Eqs. (21) and (22) in MIKO to model the pixelization effect. The results are demonstrated in Fig. 4 for the Gaussian model. If we only include smoothing (red) in the forward model, the  $\langle C(\kappa^{\text{pix}}) \rangle$  is two times lower near  $\ell_N$ , as expected. When both smoothing and aliasing are included (purple), the model generates  $\kappa^{\text{pix}}$ 's that align with  $\kappa^{\text{true}}$  on the two-point level across all scales. This behavior is consistently observed across all the cross-power spectra and is also confirmed for the log-normal model.

We now test the bias introduced by the pixelization effect on the cosmological parameters  $A_i$ . We setup three different Gaussian models. Each model includes the smoothing correction but differs in the aliasing parameter  $r$ , set at 1 (no aliasing correction), 1.3, and 1.6 respectively. We run each model on 80 HSC-like experiments, using Gaussian mocks. The results are presented in Fig. 6. Neglecting aliasing in the forward model biases  $A_i$  higher. The direction of the bias is expected since aliasing amplifies the small-scale power in  $\gamma^{\text{obs}}$  which can only be compensated by increasing  $A_i$ . Neglecting aliasing leads to a 2.5% absolute bias on the cosmological parameters exceeding our error budget. However, setting  $r = 1.6$  is generally

sufficient to model the aliasing effect. We will use  $r \geq 1.6$  for all the following analyses.

### C. Boundary conditions

We use the Kaiser-Squire relationship Eqs. (19) and (20) to transform  $\kappa$  into  $\gamma$  in the forward model. The Kaiser-Squire relationship is exact on an infinite plane. However, our forward model is defined on a finite grid which has a periodic boundary condition inherited from the fast Fourier transform. This mismatch introduces field-level errors around the edges of the maps as shown in Fig. 7. We can alleviate this bias by generating maps larger than the survey footprint and defining the likelihood on the inner regions. More specifically, we test how different amounts of margins, measured by the fractional length of the short side of the map, affect the final cosmological parameter constraints. The result is shown in Fig. 8. We find that the boundary condition mismatch drives  $A_i$  higher. Furthermore, a 80% margin is sufficient to reduce the bias

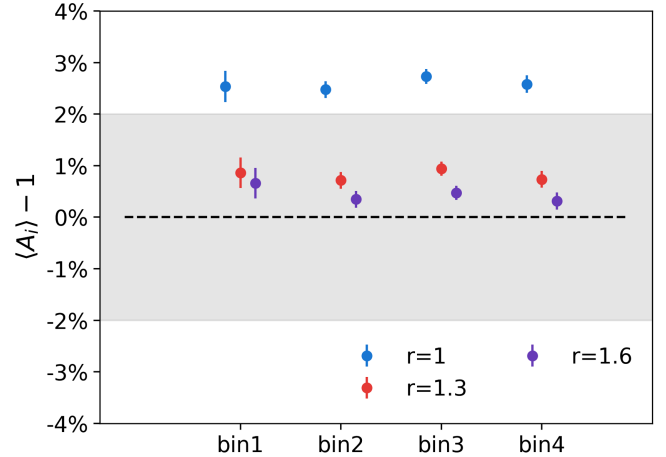


FIG. 6. The absolute bias of  $A_i$  measured by applying the Gaussian model to Gaussian mocks. Each color corresponds to a Gaussian model with a different aliasing parameter  $r$ . Each data point represents the average inferred  $A_i$  over 80 independent HSC-like experiments, with the error bars denoting the standard error of the mean. The experiment is set with an 80% margin (see Sec. VC and Fig. 8). The gray region indicates the 2% systematics control target.

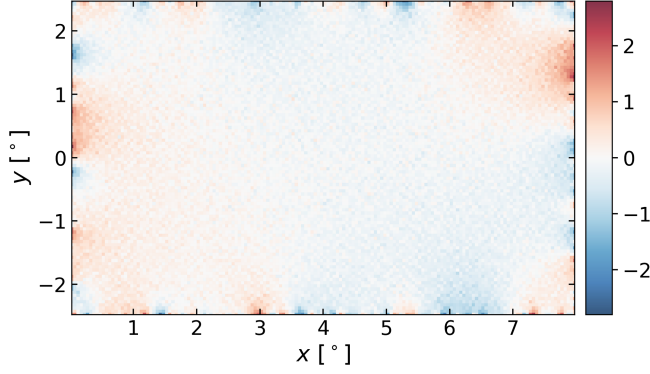


FIG. 7. The Kaiser-Squire reconstruction error of a (Gaussian) shear field assuming periodic boundary condition. We define the reconstruction error as  $\epsilon = (\mathbb{K}_1\kappa - \gamma_1)/\Delta_\gamma$ , where  $\mathbb{K}_1$  is the Kaiser-Squire operator that transforms  $\kappa$  to  $\gamma_1$  assuming a nonperiodic boundary at spatial infinity.  $\Delta_\gamma$  is the standard deviation of the shear field. We observe the most error along the boundaries.

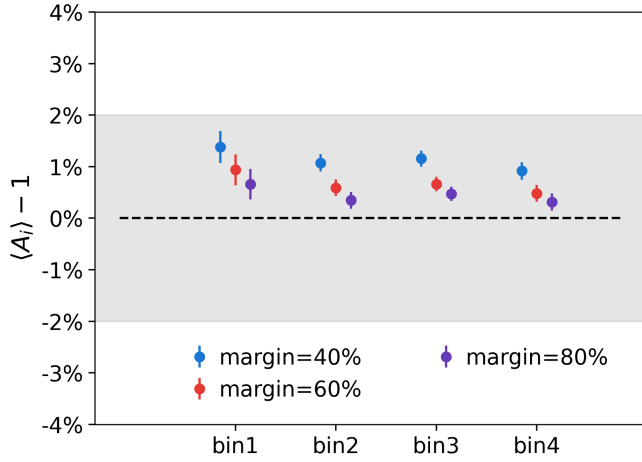


FIG. 8. The absolute bias of  $A_i$  measured by applying the Gaussian model to Gaussian mocks. Each color corresponds to a Gaussian model with a different margin size (as a fraction of the length of the short side of the patch). Each data point represents the average inferred  $A_i$  over 80 independent HSC-like experiments, with the error bars denoting the standard error of the mean. The experiment is set with  $r = 1.6$  (see Sec. V B and Fig. 6). The purple points corresponds to the same purple points in Fig. 6. The gray region indicates the 2% systematics control target.

below 0.5%. We will use this choice of margin for all the following analyses.

#### D. Fourier mode coupling

As pointed out by Xavier *et al.* [30] and Tessore *et al.* [49], the log-normal model transformation is not local in Fourier space and strongly couples the small-scale modes. A finite Fourier grid implies a strict band-limit and therefore changes the expected behavior of the transformation. To faithfully simulate a log-normal map up to  $\ell_{\max}$ , we

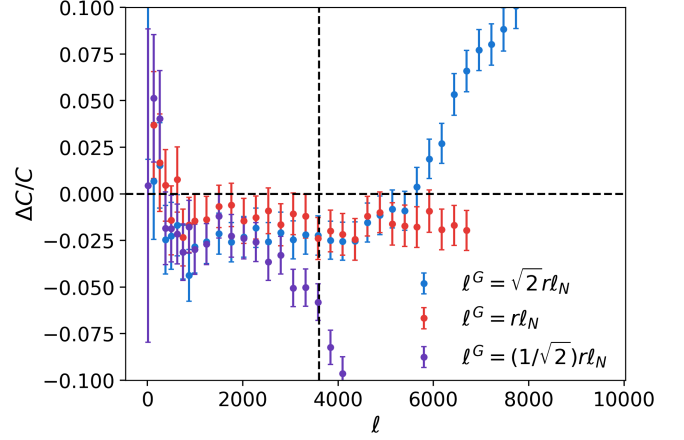


FIG. 9. The impact of the Gaussian transfer map scale-cut,  $\ell^G$ , on the recovered  $C$  is presented in terms of the fractional difference relative to the input  $C$ . The error bars represent the standard error of the mean. The vertical line denotes the Nyquist frequency,  $\ell_N$ , of  $\gamma^{\text{obs}}$ . Setting  $\ell^G = r\ell_N$  (red) recovers the input  $C$  with the highest accuracy. The observed residual additive bias is related to the limited sky coverage of the patch.

need to generate a Gaussian transfer map up to  $\ell^G > \ell_{\max}$ . For example,  $\ell^G > 4\ell_{\max}$  is required for percent-level accuracy when simulating full-sky log-normal maps using FLASK (as is done for our log-normal mocks Sec. III). The iterative correction method proposed by Tessore *et al.* [49] can, in theory, produce logarithmic normal maps with accurate power spectra with  $\ell^G \approx \ell_{\max}$ . We, however, did not achieve the same success when applying the same idea to the flat sky case, potentially due to the boundary condition and the nonisotropy of the Fourier grid.

Nevertheless, we found an empirical solution to reduce this error. In Fig. 9, we plot the scale-dependent error of the log-normal  $C(\kappa)$  relative to the input  $C$  as a function of  $\ell^G$ . If we set  $\ell^G$  to the maximum aliasing scale  $\sqrt{2}r\ell_N$  (blue),  $C(\kappa)$  starts to deviate from the input beyond  $\ell_N$  (vertical line). To obtain the correct  $C(\kappa)$  beyond  $\ell_N$ , we find it useful to force  $\kappa^G$  to be isotropic by removing powers beyond  $r\ell_N$  (red). If we impose an even stricter scale-cut, the log-normal maps exhibit lower band power below  $\ell_N$ . Regardless of  $\ell^G$ , the log-normal maps still have a deficit of large-scale power. This appears to be a consequence of the small survey footprint or the periodic boundary effect, and this error is reduced when generating larger maps. For the log-normal model, since the Fourier mode coupling issues that we have discussed so far interact together, it is difficult to quantify each error separately. However, looking ahead, we will show that using  $\ell^G = r\ell_N$  with  $r = 1.9$  enables us to control the systematic bias for the log-normal model to 2%.

#### E. Number density-induced shape noise

Finally, we observe another novel source of error that impacts FLI—the finite galaxy density can induce a shape



noise-like error on the pixelized maps even when no shape noise is introduced. When we average the galaxy shapes within a pixel, the distribution of random galaxy positions samples the inhomogeneous shear field, leading to statistical fluctuation of the pixel value. Formally, for a pixel  $p$  centered at  $\theta_p$ , we expect the galaxy's true shape to differ from  $\gamma(\theta_p)$  by

$$\Delta\gamma \approx \frac{\partial\gamma}{\partial\theta}(\theta_p)\Delta\theta = \ell\gamma\Delta\theta. \quad (25)$$

This term behaves like a white noise that only appears in the autospectra of  $\gamma^{\text{pix}}$ . Therefore, we call this the number density-induced shape noise, or density shot noise in short. The impact of density shot noise is characterized in Fig. 10, where we see that it impacts the small scale autopower spectrum by about 20% near the pixel resolution scales. We did not observe this noise in the cross-spectra. To correct for this systematics, we first measure the power spectrum of this density-induced shape noise term ( $C_{ii}^{\text{dsn}}$ ) directly from mock catalogs for each redshift bin  $i$  (as shown in Fig. 10). Then, when we generate convergence maps following Eq. (17) (for the Gaussian model) and (18) (for the log-normal model), we can simply change the cosmological power spectrum  $C_{ii}$  via

$$C_{ii} \rightarrow C_{ii} + C_{ii}^{\text{dsn}}. \quad (26)$$

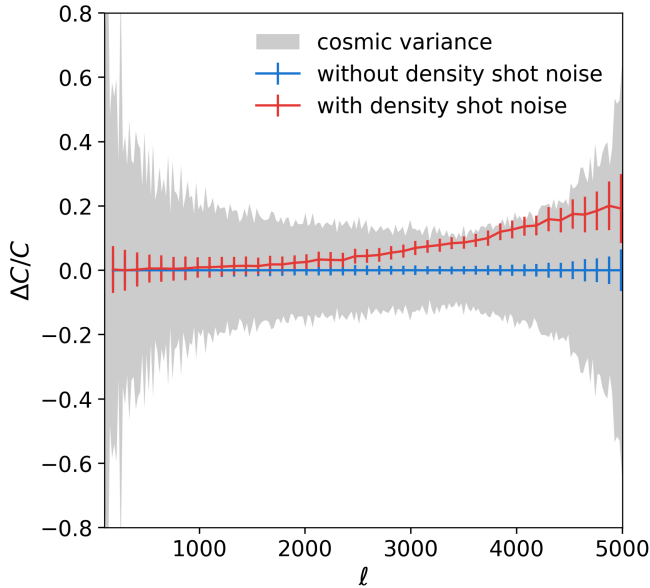


FIG. 10. The impact of number density-induced shape noise on  $C$  is presented in terms of the fractional difference relative to  $\langle C(\kappa^{\text{true}}) \rangle$  where the number density-induced shape noise is removed (achieved by creating maps using  $\rho_i$  and  $\gamma_{\text{rms}}^2$  at ten times the fiducial value). The high density case is represented in blue, and its  $1\sigma$  spread is shown in gray (as in Fig. 4).  $\langle C(\kappa^{\text{true}}) \rangle$  calculated with fiducial values for  $\rho_i$  and  $\gamma_{\text{rms}}$  is shown in red. The error bars denote the standard error of the mean.

In this way, the shear maps ( $\gamma_i^{\text{pix}}$ ) generated by the forward model automatically include the density shot noise. The expression  $(\gamma_i^{\text{obs}} - \gamma_i^{\text{pix}}(\mathbf{c}, \mathbf{x}_i))^2/2N$  in Eq. (16) still accurately describes the regular shape noise model and does not require additional modification.

## VI. MODEL MISSPECIFICATION

We have demonstrated that by accounting for various analysis systematics—including aliasing, boundary effects, mode coupling, and density-induced shape noise—we can obtain unbiased cosmological constraints from observed pixelized shear maps. A remaining modeling uncertainty is the field-level prior on the convergence field. It is anticipated that this uncertainty could potentially induce an absolute bias on the inferred parameters as well as affect the uncertainty quantification. In this section, we test both the Gaussian and the log-normal models using Gaussian, log-normal, and Takahashi mocks.

As we have shown above, a significant difference exists between the realistic observed shear maps and the maps generated by the forward model (which we shall call pipeline mocks). Previous works [8,9] have demonstrated that it is possible to use the forward model to analyze the pipeline mocks and correctly recover the cosmological parameters. Here, we follow Fiedorowicz *et al.* [10] and go one step further toward realism by using MIKO to analyze the raw mocks themselves. A further feature of this analysis is that, to the best of our knowledge, this is the first study of its kind to extract cosmological parameters (here the  $A_i$ ) using realistic mocks and taking into account the range of analysis systematics that must be considered for real data analysis.

### A. Impact on cosmological parameters

#### 1. The Gaussian model

First, we run the Gaussian model on all three sets of mocks. For each model-data pair, we perform 80 HSC-like experiments on independent patches, utilizing analysis systematics models recommended by Sec. V. For each experiment, we obtain the joint posterior distribution of  $A_i$  along with the map parameters. For instance, the constraint on  $A_i$  for a single Takahashi patch is depicted as the red contour in Fig. 11. The  $A_i$ 's are uncorrelated in the Gaussian model, consistent with the findings in Zhou and Dodelson [7]. There is considerable cosmic variance within a  $40 \text{ deg}^2$  patch. To determine whether the pipeline has an absolute bias on  $A_i$ , we compute  $\langle A_i \rangle - 1$  over independent experiments and present the results in the top panel of Fig. 12. The top part of the top panel shows that the Gaussian model does not exhibit absolute bias for all simulated datasets, regardless of whether the data is simulated with a Gaussian model or not.

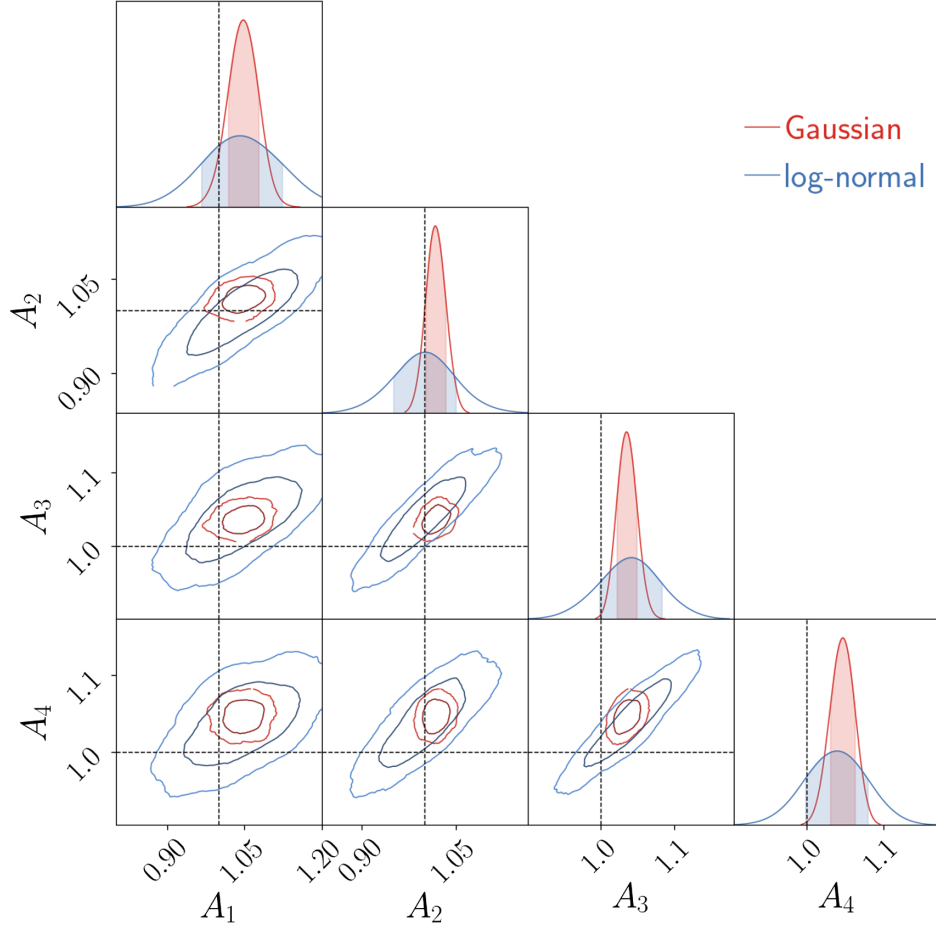


FIG. 11. The joint posterior distribution of  $A_i$  inferred from a 40 deg<sup>2</sup> Takahashi patch using the Gaussian (red) and the log-normal (blue) models. The log-normal model employs an informative prior on  $a_i$  that is jointly sampled (not shown here) with  $A_i$ . Compared to the Gaussian model, the log-normal model exhibits larger uncertainties and a stronger correlation between the  $A_i$ 's. The contours show the  $1\sigma$  and  $2\sigma$  level sets.

To assess the impact of model misspecification on our uncertainty estimates, let us consider the one-dimensional marginal posterior distribution of a particular  $A_i$ . Ideally,  $x\%$  (the observed probability) of all the experiments should find  $A_i = 1$  contained within the  $x\%$  (the expected probability) credible interval. The observed probability as a function of the expected probability is called the P-P plot. It is shown in the lower part of the top panel of Fig. 12, after adjusting for mean absolute bias to concentrate on the error bars. Perfect uncertainty prediction would align the P-P curves with the line  $y = x$ . This alignment occurs only without model misspecification. For instance, applying the Gaussian model to log-normal or Takahashi mocks results in overly optimistic uncertainty estimates for all  $A_i$  values, particularly for log-normal mocks where the truth is included within the one standard deviation interval less than 20% of the time. This is consistent with the observation of Boruah *et al.* [9], although the extent to which the Gaussian model underestimates the uncertainty is much more severe in our case. The bottom line is that assuming a Gaussian prior leads to an unbiased estimate of the

cosmological parameters but a *biased* estimate of the errors on those parameters.

## 2. The “unreasonable” effectiveness of the Gaussian priors

Why does the Gaussian map prior not induce any absolute bias on  $A_i$  when it is applied to clearly non-Gaussian fields? The answer lies in the PDF of  $\tilde{\kappa}$  in Fourier space. In Fig. 13, we plot the PDFs of  $\tilde{\kappa}^{\text{true}}$  for Gaussian, log-normal, and Takahashi mocks at three different  $\ell$  values. Interestingly, for all mocks and across all ranges of  $\ell$ , the PDF of  $\tilde{\kappa}^{\text{true}}$  is consistent with a Gaussian distribution. Similar phenomena have previously been observed in three-dimensional simulations, for example, by Matsubara [50] and Qin *et al.* [51]. Motivated by these observations, we can express the PDFs of non-Gaussian (prior<sup>NG</sup>) and Gaussian (prior<sup>G</sup>) convergence fields as follows (ignoring aliasing and other analysis systematics for simplicity)

$$\text{prior}^{\text{NG}}(\tilde{\kappa}) = \text{prior}^{\text{G}}(\tilde{\kappa})[1 + \delta(\tilde{\kappa})]. \quad (27)$$

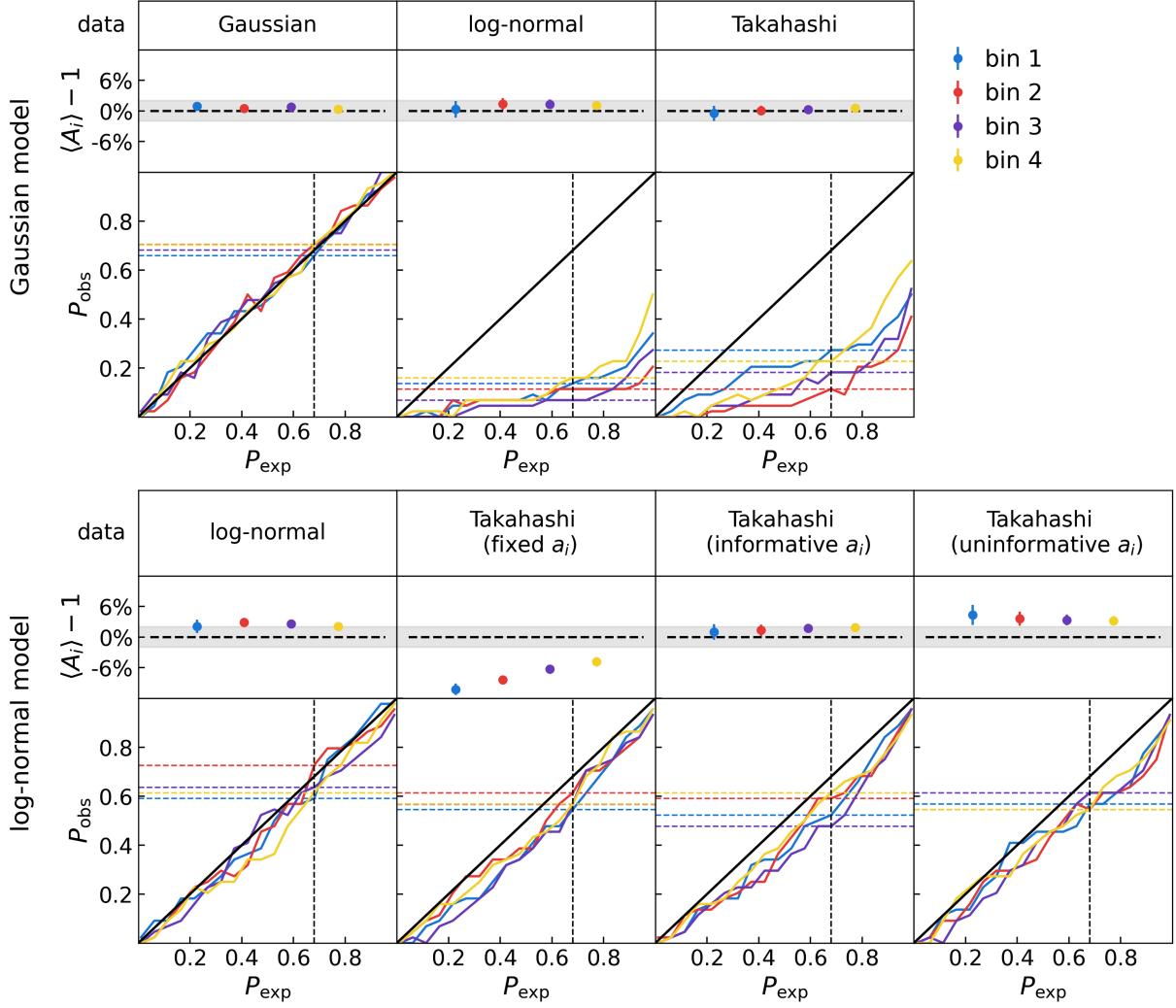


FIG. 12. The top panel displays the results of the Gaussian model applied to Gaussian, log-normal, and Takahashi mocks. Within this panel, the first row indicates the absolute bias of  $A_i$ , averaged over 80 independent HSC-like experiments. The gray area denotes the systematic error control target of 2%. The second row presents the P-P plot for  $A_i$  (with the mean bias subtracted), where the  $x$ -axis denotes the expected probability ( $P_{\text{exp}}$ ) that a credible interval contains the true  $A_i$ , and the  $y$ -axis denotes the observed probability ( $P_{\text{obs}}$ ). The vertical dashed line indicates  $P_{\text{obs}} = 68\%$ , and the colored horizontal lines show where the P-P curves intersect with this line. The Gaussian model demonstrates no absolute bias irrespective of model misspecification but exhibits overconfidence in the error bars when applied to realistic data. The bottom panel illustrates the absolute bias and the P-P plot for  $A_i$  using the log-normal model, applied to log-normal and Takahashi mocks. For the Takahashi mocks, three scenarios are tested with the log-normal model concerning the shift parameters: (1) fixed at values measured from noiseless true mocks (fixed  $a_i$ ), (2) allowed to vary within a narrow and informative range, and (3) allowed to vary within a broad and uninformative range. The log-normal model generally yields accurate uncertainties, yet the absolute bias is highly sensitive to the choice of  $a_i$ .

In the absence of model misspecification, the likelihood analyses should yield unbiased  $A_i^{\text{NG}}$ , where

$$\langle A_i^{\text{NG}} \rangle = \int d\gamma^{\text{obs}} \int d\tilde{\kappa} A_i \mathcal{P}(\gamma^{\text{obs}} | \tilde{\kappa}, A_i) \text{prior}^{\text{NG}}(\tilde{\kappa}) \quad (28)$$

$$= \int d\gamma^{\text{obs}} \int d\tilde{\kappa} A_i \mathcal{P}(\gamma^{\text{obs}} | \tilde{\kappa}, A_i) \text{prior}^{\text{G}}(\tilde{\kappa}) [1 + \delta(\tilde{\kappa})] \quad (29)$$

$$= \langle A_i^{\text{G}} \rangle + \text{correction}[\delta] \quad (30)$$

Therefore, as long as the perturbation  $\delta$  is small,  $\langle A_i^{\text{G}} \rangle = \langle A_i^{\text{NG}} \rangle$  is unbiased.

The second question pertains to why the Gaussian model fails to predict the correct uncertainty of  $A_i$ . To address this, we first examine the correlation between  $C(\ell_1)$  and  $C(\ell_2)$  across the Gaussian model, the log-normal model, and the Takahashi mocks, as illustrated in Fig. 14. As expected, the Gaussian model assumes there is no correlation between the power of two different  $\ell$  modes. Nonetheless, correlations are evident in both the log-normal simulations and the realistic Takahashi mocks. Consequently, the Gaussian

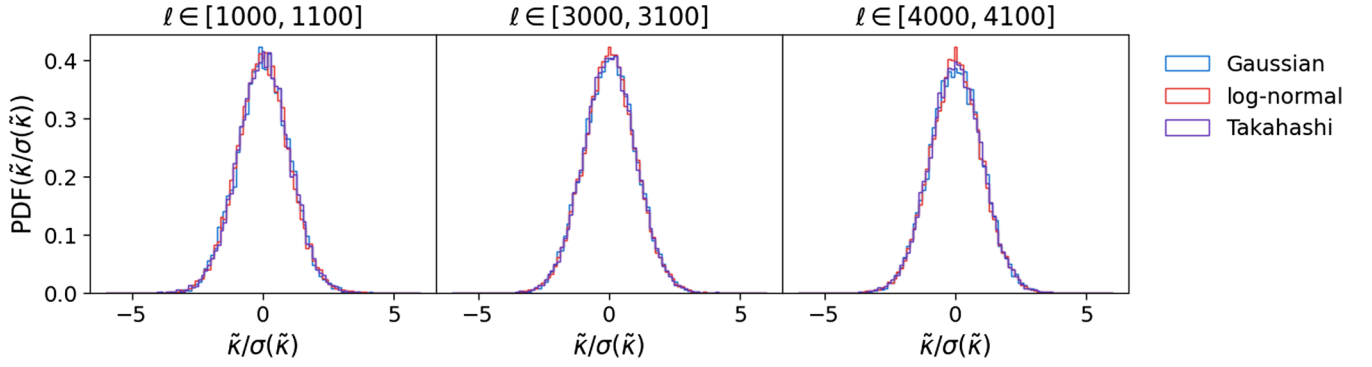


FIG. 13. Standardized one-point PDF of  $\tilde{\kappa}(\ell)$  for Gaussian, log-normal, and Takahashi mocks at three distinct  $\ell$  values. Across all mocks and  $\ell$  values, the distribution of  $\tilde{\kappa}(\ell)$  remains consistently Gaussian, which elucidates the Gaussian model’s ability to infer cosmological parameters without absolute bias, notwithstanding model misspecification.

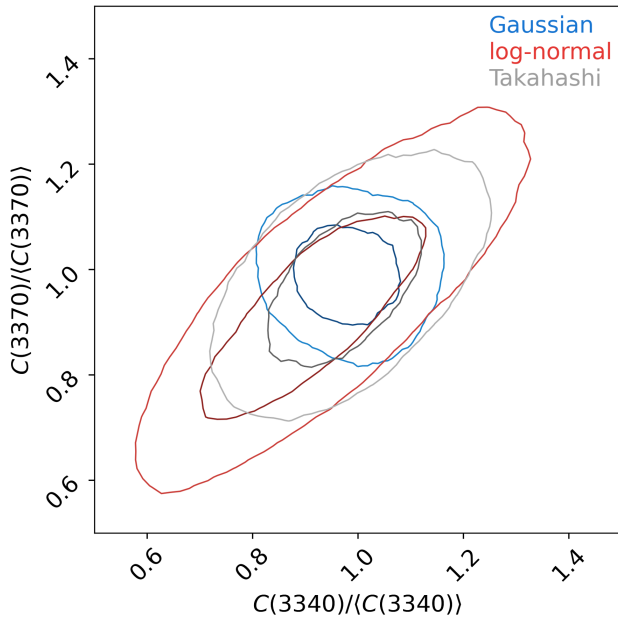


FIG. 14. The correlation between  $C(\ell_1)/\langle C(\ell_1) \rangle$  and  $C(\ell_2)/\langle C(\ell_2) \rangle$  for Gaussian (blue), log-normal (red), and Takahashi (gray) mocks. In this example, we have chosen  $\ell_1 = 3340$  and  $\ell_2 = 3370$ , and  $C$  to be the autopower spectrum of the first redshift bin. The contours show the  $0.3\sigma$  and  $0.8\sigma$  level sets. The Gaussian model assumes no mode correlations. The log-normal model assumes more mode correlation than the realistic mocks.

model presupposes that each  $\ell$  mode contributes independently, while in actuality, the different  $\ell$  modes are correlated and thus provide less information than assumed. As a result, the Gaussian model reports a smaller error bar on the inferred power spectrum amplitude  $A_i$ , an error analogous to assuming a larger  $f_{\text{sky}}$  than what is true in reality. Another perspective on this issue is provided by examining the cosmic variance predicted by the Gaussian, log-normal, and Takahashi maps, as shown in Fig. 15. The Gaussian model predicts smaller cosmic variance than the

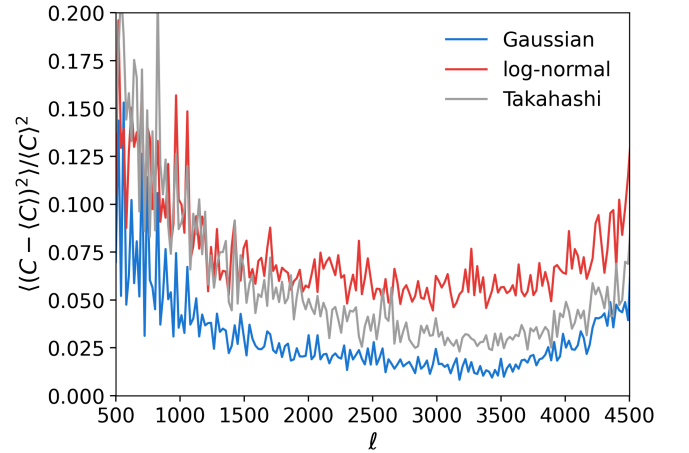


FIG. 15. Normalized variance of the power spectrum for Gaussian (blue), log-normal (red), and Takahashi (gray) mocks, with  $C$  representing the autopower spectrum of the first redshift bin. Notably, the Gaussian mocks exhibit smaller cosmic variance relative to the Takahashi mocks, whereas the log-normal mocks demonstrate greater cosmic variance in comparison.

other two models across all scales, which then translates to the narrower uncertainty bounds on the estimated power spectrum amplitude  $A_i$ .

### 3. The log-normal model

The log-normal model depends on both  $A_i$  and the shift parameter  $a_i$ . We first run the log-normal model on the 44 log-normal mocks fixing  $a_i$  at the truth. The pipeline recovers the correct  $A_i$  with an absolute bias within  $2\%^3$ . We also report excellent uncertainty quantification. This result is shown in the bottom panel of Fig. 12 (column 1).

<sup>3</sup>The slight positive bias is due to the deficit of large-scale powers discussed in Sec. VD. This error can be alleviated by using larger patches.

Model misspecification also arises when applying the log-normal model to the Takahashi mocks. Specifically, since the Takahashi mocks do not adhere to a strict log-normal distribution, as shown in Fig. 1, a definitive ground truth for  $a_i$  is absent. Previous studies have approached  $a_i$  in various ways: for instance, Alsing *et al.* [8] treated  $a_i$  as fixed parameters, whereas Boruah *et al.* [9] and Fiedorowicz *et al.* [10] employed perturbation theory for approximation. As we will show, the modeling of  $a_i$  is arguably the primary source of bias with the log-normal model and warrants careful consideration.

In our first approach, we compute  $a_i^{\text{fit}}$  by fitting the one-point PDF to the noiseless Takahashi mocks using Eq. (7). We then perform inference on the Takahashi mocks assuming  $a_i = a_i^{\text{fit}}$ . This method approximates a perfect perturbation theory prediction for  $a_i$ . The result is presented in the bottom panel of Fig. 12 (column 2). Surprisingly, even when using the optimally fitted  $a_i^{\text{fit}}$ , the log-normal

model infers  $A_i$  with a significant negative bias. This bias is most pronounced, reaching up to 10%, in the lowest redshift bins where the convergence field exhibits the highest non-Gaussianity. Consequently, we deduce that merely employing the log-normal shift parameter that most closely fits the PDF of the actual convergence field is not only insufficient but also incorrect for acquiring unbiased constraints on cosmological parameters.

In our second approach, we adopt an agnostic stance on  $a_i$ , treating it as a variable during sampling. Observing that the likelihood analysis typically favors  $a_i > a_i^{\text{fit}}$ , we implement a flat prior on  $a_i$  in the range  $[0.9a_i^{\text{fit}}, 1.3a_i^{\text{fit}}]$ . The joint posterior of  $A_i$ , as depicted in Fig. 11, reveals that the uncertainties reported by the log-normal model are larger—yet appear more realistic—than those by the Gaussian model, with  $A_i$  exhibiting positive correlation. Furthermore, the joint posterior of  $A_i$  and  $a_i$  (Fig. 16) indicates a positive correlation between them, with the

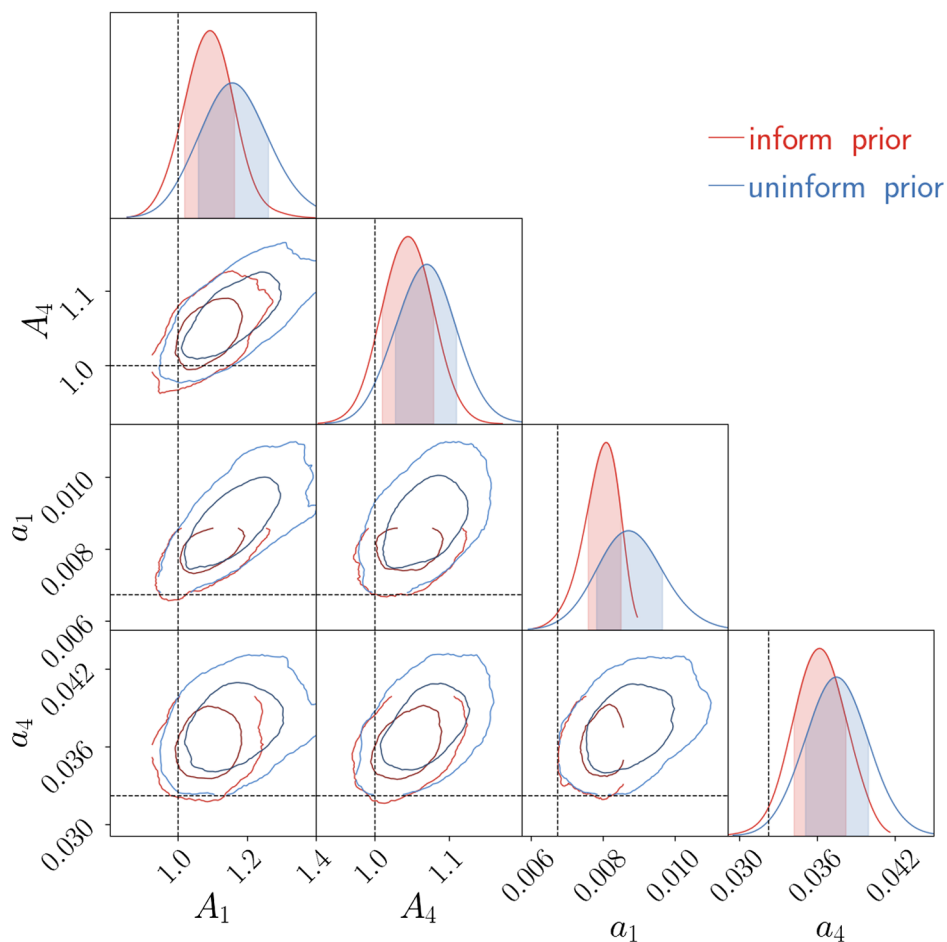


FIG. 16. The joint posterior distribution of  $A_i$  and  $a_i$ , as inferred from a Takahashi patch using the log-normal model. We compare an informative prior (red) and an uninformative prior (blue) on  $a_i$ . For simplicity, only the first and the last redshift bins are shown. The dashed lines across  $A_i$  represent  $A_i = 1$ , and the dashed lines across  $a_i$  represent  $a_i^{\text{fit}}$ —the best-fit values from noiseless data. Three important conclusions are (1)  $A_i$  and  $a_i$  are positively correlated, (2) the log-normal model prefers  $a_i$  much larger than  $a_i^{\text{fit}}$ , and (3) the informative prior truncates the distribution of  $a_i$ . Therefore, when the prior is relaxed,  $A_i$  shifts to higher values. The contours show the 1σ and 2σ level sets.

maximum likelihood estimate for  $a_i$  substantially exceeding  $a_i^{\text{fit}}$ . This is consistent with the previous observation that fixing  $a_i$  to  $a_i^{\text{fit}}$  induces a negative bias on  $A_i$ . Nonetheless, Fig. 16 also illustrates that the prior range for  $a_i$  truncates the posterior distribution, hence we refer to this as the informative prior model. Through 44 independent experiments, we find that imposing an informative prior on  $a_i$  can significantly reduce the absolute bias on  $A_i$  to within 2% while maintaining accurate uncertainty estimation for  $A_i$ . The P-P curves in the bottom panel of Fig. 12 quantify the extent to which the uncertainty estimation is accurate when the log-normal prior is used.

Lastly, we impose a uniform and uninformative prior on  $a_i$  over the positive reals. We explicitly check that the prior does not truncate the posterior distribution. An example is shown in Fig. 16. As the positive correlation between  $a_i$  and  $A_i$  suggest, this model creates a positive absolute bias on  $A_i$  (bottom panel of Fig. 12, column 4). The bias is most severe in the lowest redshift bin at the 4% level.

To summarize, the three choices for  $a_i$ 's prior suggest that, in order to obtain unbiased cosmological constraints on real data using the log-normal model, we can neither use  $a_i$  directly measured from the noiseless mocks nor treat it as a variable that is completely free. Otherwise, although we obtain good uncertainty quantification, we risk significant bias in cosmological parameters. The existence of the absolute bias is not too surprising given that the log-normal also misspecifies the field-level priors.

## B. Map reconstructions

MIKO also provides the joint posterior distribution of all the map pixels— $(\kappa_i^{\text{pix}})_s$  (the bottom diamond in the

flowchart Fig. 3), where  $s$  denotes the sample index. For simplicity, let us focus specifically on one experiment on the Takahashi mock, and on the reconstruction of the first redshift bin convergence field. The results for both the Gaussian and log-normal (informative prior) models are summarized in Fig. 17.

Let us start with the first row. The first and second columns depict the observed noisy shear field,  $\gamma^{\text{obs}}$ , and the true convergence,  $\kappa^{\text{true}}$ , respectively. The final three columns pertain exclusively to the log-normal model, illustrating the mean posterior map  $\langle \kappa^{\text{pix}} \rangle$ , the standard deviation map  $\sigma$ , and the signal-to-noise ratio map (S/N). For these last three maps, we zoom in on a  $1.5 \text{ deg} \times 1.5 \text{ deg}$  patch in the inset plots below. In these insets, we also compare the log-normal results with those of the Gaussian model.

Theoretically, we do not expect  $\langle \kappa^{\text{pix}} \rangle$  to match  $\kappa^{\text{pix}}$  exactly. Instead,  $\langle \kappa^{\text{pix}} \rangle$  is *roughly* analogous to a Wiener-filtered version of  $\kappa^{\text{pix}}$ . This filtering smooths the map on small scales where shot noise dominates (see Zhou and Dodelson [7] for a detailed discussion on the properties of the posterior maps and the maps' power spectra). Both the Gaussian and the log-normal model recover the large-scale features of the true convergence maps. The log-normal model, however, resolves the density peaks much better (insets in the second row). The adjacent histogram plot further supports this claim—the log-normal model's  $\text{PDF}(\langle \kappa^{\text{pix}} \rangle)$  agrees with  $\text{PDF}(\kappa^{\text{true}})$  up to the highest  $\kappa$  values. In contrast, the Gaussian model's  $\text{PDF}(\langle \kappa^{\text{pix}} \rangle)$  decays much faster. On the other hand, the log-normal model's  $\text{PDF}(\langle \kappa^{\text{pix}} \rangle)$  shows a hard cutoff in the low  $\kappa$  region. The Gaussian model captures the voids much better

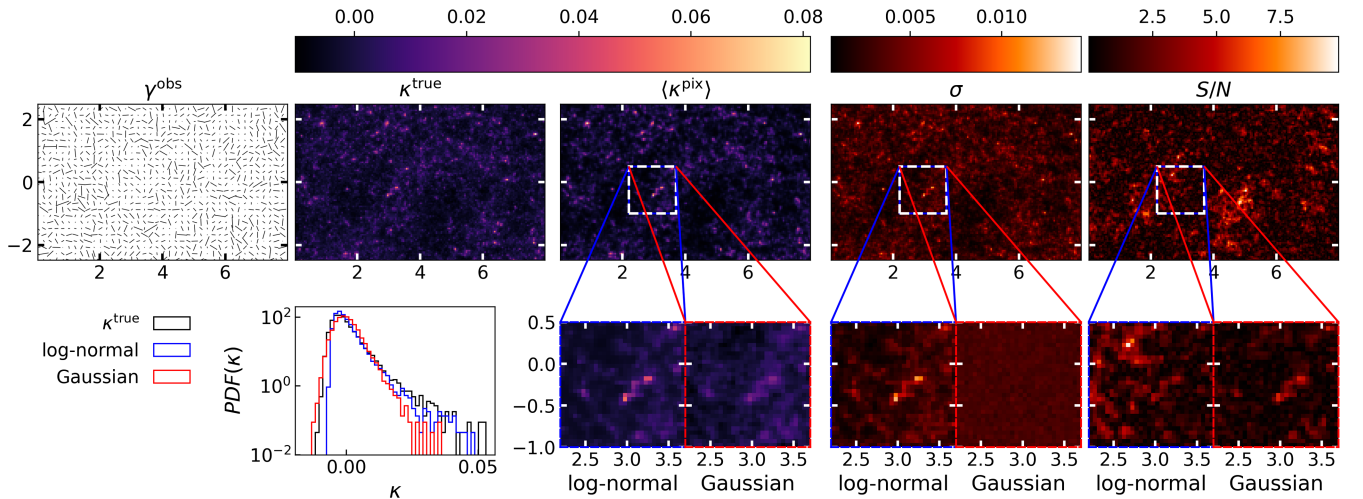


FIG. 17. Field-level reconstruction for the first redshift bin of a single Takahashi patch using the log-normal (with informative prior on  $a_i$ ) and the Gaussian models. The first row, from left to right, are the observed noisy shear map  $\gamma^{\text{obs}}$ , the true convergence map  $\kappa^{\text{true}}$ , the mean posterior map  $\langle \kappa^{\text{pix}} \rangle$  (log-normal model), the standard deviation map  $\sigma$  (log-normal model), and the S/N map (log-normal model). In the second row, the leftmost plot compares the histograms of  $\kappa^{\text{true}}$  with those of  $\langle \kappa^{\text{pix}} \rangle$  for both Gaussian and log-normal models. The three inset plots zoom in on  $1.5 \text{ deg} \times 1.5 \text{ deg}$  region on the patch, contrasting the log-normal results (left) with the Gaussian results (right).

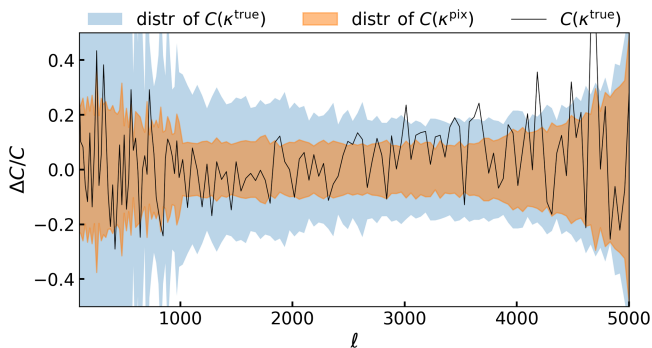


FIG. 18. Power spectra is presented in terms of the fractional difference relative to  $\langle C(\kappa^{\text{pix}}) \rangle$ , where  $\kappa^{\text{pix}}$  are map samples inferred from a single Takahashi patch using the log-normal model with an informative prior on  $a_i$ . The power spectrum distribution of the sampled maps,  $C(\kappa^{\text{pix}})$ , is depicted in orange, while the distribution of the mock fields,  $C(\kappa^{\text{true}})$ , is shown in blue to indicate cosmic variance. The power spectrum of the true noiseless field being inferred is shown in black. In this context,  $C$  specifies the autopower spectrum of the first redshift bin. The posterior samples are in agreement with  $\kappa^{\text{true}}$  at the two-point level within the bounds of uncertainty.

in comparison. This aligns with the findings of Fiedorowicz *et al.* [10], who reported that the log-normal model is less successful in reconstructing the correct void count distribution at higher resolutions.

The log-normal and the Gaussian model also differ in their prediction of the pixel-level uncertainty. (We show the  $1\sigma$  uncertainty for each pixel in the fourth column and the comparison between the log-normal model and the Gaussian model in the inset plots below.) The variance predicted by the log-normal model shows a positive correlation with the absolute pixel value, whereas the Gaussian model predicts uniform uncertainty across all pixels. In the rightmost column, we display the signal-to-noise ratio (S/N). Interestingly, both models predict identical S/N for the peaks, yet the log-normal model yields a higher S/N for the voids due to its nonuniform uncertainty predictions.

For each  $\kappa^{\text{pix}}$  sample, we can compute its power spectrum  $C(\kappa^{\text{pix}})$ . As proved in Zhou and Dodelson [7], in the absence of model misspecification, the distribution of  $C(\kappa^{\text{pix}})$  is expected to follow that of  $C(\kappa^{\text{true}})$ . The distribution of  $C(\kappa^{\text{pix}})$  for the log-normal model with informative prior on  $a_i$  is depicted in Fig. 18, where we indeed confirm that  $C(\kappa^{\text{pix}})$  recovers  $C(\kappa^{\text{true}})$  within uncertainty.

## VII. SUMMARY

We have presented a percent-level accurate field-level cosmological analysis pipeline, MIKO, that infers the joint posterior distribution of the cosmological parameters and the tomographic convergence fields. We identified several analysis systematics such as aliasing, boundary effect, mode-coupling, and density-induced shape noise, many

of which have not been captured or modeled in previous studies. We measured their impact on cosmological parameter constraints and proposed effective mitigation methods that control the systematic uncertainty down to the 2% level. We then explore the problem of model misspecification (using wrong field-level priors) by testing the Gaussian and log-normal forward models on three sets of mock data (Gaussian, log-normal, and ray-tracing simulations). The Gaussian model shows no absolute bias in the  $A_i$  (the amplitude of the power spectrum in the  $i^{\text{th}}$  redshift bin) even with model misspecification, but yields overconfident uncertainty on cosmological parameters. The log-normal provides accurate uncertainty quantification, but is sensitive to the choice of the shift parameters,  $a_i$ . In particular, the shift parameters fitted to the simulations can incur 10% absolute bias on power spectrum amplitudes. After accounting for analysis systematics (and calibrating  $a_i$ 's prior for the log-normal model), both models recover  $A_i$  with an absolute bias less than 2%. Our result demonstrates the accuracy and unbiased nature of this field-level cosmological analysis framework, a prerequisite for it to provide meaningful, tight cosmological parameter constraints.

The model misspecification bias identified in this work carries important practical implications for future research. Given that the log-normal distribution results in prior-dependent cosmological constraints, it is crucial to calibrate prior( $a_i$ ) before analyzing real data. We have demonstrated the feasibility of constructing prior( $a_i$ ) to infer unbiased cosmology from mocks. Nonetheless, there is no assurance that the same prior( $a_i$ ) will work on real data distributions. Future studies should examine the robustness of prior( $a_i$ ) across different simulations or investigate alternative one-point distributions that both reflect the non-Gaussianity of the field and remain unaffected by priors on nuisance parameters. Given the similar cosmological constraining powers derived from both the Gaussian and the log-normal models, it currently seems more practical to utilize the Gaussian model and calibrate the uncertainty estimates based on independent mock datasets.

Many analysis systematics identified in this work are applicable to general field-level inference problems involving pixelized maps. For instance, we can employ the same aliasing correction when modeling the observed temperature and polarization maps CMB at the field level. Otherwise, we risk a positive absolute bias in the inferred amplitude of the power spectrum. The error induced by model misspecification is also not weak lensing-specific.

In the next step, we will implement galaxy intrinsic alignment [32], redshift error [35], and point spread function [52] systematics into MIKO and apply to the HSC Year 3 data. Implementing efficient algorithms to incorporate correlated pixel noise is also important [14]. Another direction is to leverage emulator-based models to constrain more general cosmological parameters [53–55].

## ACKNOWLEDGMENTS

We thank Masahiro Takada, Fabian Schmidt, and Bhuvnesh Jain on detailed and insightful comments on the draft of this paper. A.Z. thanks Chirag Modi for discussion on sampling, Benjamin Horowitz for discussion on the log-normal map prior, and Yin Li for discussion on the FFTLog Hankel transform in early stages of the project. A. Z. also expresses his gratitude to the Yukawa Institute for Theoretical Physics at Kyoto University for hosting the YITP workshop YITP-W-23-02 on “Future Science with CMB x LSS.” The discussions during the workshop were instrumental in completing this work. X.L. and R.M. were supported by a grant from the Simons Foundation (Simons Investigator in Astrophysics, Award ID No. 620789). This work was supported by NSF Award No. 2020295.

## APPENDIX: CALIBRATION OF THE TAKAHASHI SIMULATIONS’ POWER SPECTRA

The convergence and shear maps calculated from the Takahashi simulations have a scale-dependent and redshift-dependent bias on the two-point level when compared to the theoretical prediction [Eq. (5)]. When we analyze the Takahashi simulations, we need to correct our theoretical model to match the simulation. At low redshift, the dominating systematics is the discretization of the lens plane. During ray tracing, the simulation assumes the light rays are deflected by lens planes of finite thickness  $\Delta^{\text{ST}} = 150h^{-1}$  Mpc. We therefore introduce a correction in Eq. (5) by convolving the  $P_\delta(k; \chi)$  with a window function [36]

$$P_\delta(k; \chi) \longrightarrow \Delta^{\text{ST}} \int \frac{dk_r}{2\pi} P_\delta(k; \chi) \text{sinc}^2\left(\frac{k_r \Delta^{\text{ST}}}{2}\right) \quad (\text{A1})$$

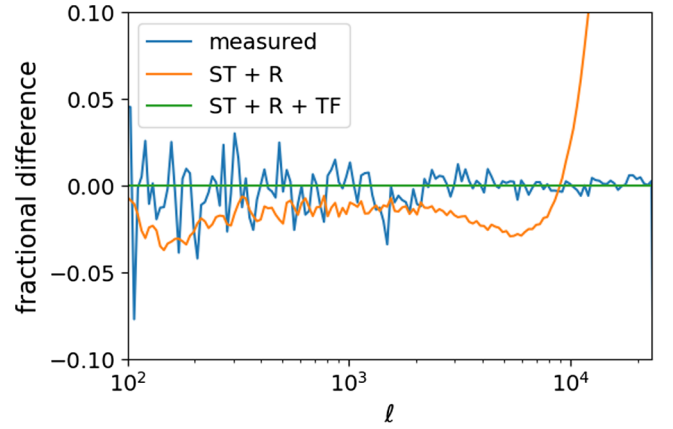


FIG. 19. Power spectra of the first redshift bin in terms of the fractional difference relative to the adopted theoretical  $C$  for the Takahashi mocks (green). The mean of  $C$ , as measured on Takahashi mocks, is depicted in blue. The theoretical  $C$ , corrected only for shell thickness and resolution, is shown in orange. This correction is found to be insufficient for meeting our accuracy requirements. Abbreviations used: ST for shell-thickness correction, R for resolution correction, and TF for transfer function correction.

where  $k_r$  is the radial wave vector. At high- $\ell$ , the bias is mostly due to the pixelization effect. This can be corrected by [36]

$$C_{ij}(\ell) \longrightarrow \frac{C_{ij}(\ell)}{1 + (\ell/(1.6 \text{ NSIDE}))^2} \quad (\text{A2})$$

These two corrections reduce the bias below 5% for  $\ell < 10000$  as shown in orange in Fig. 19. Finally, we fit a transfer function so that the bias falls below 1% for  $\ell < 24000$  for all redshift bins (shown in green).

- 
- [1] A. Andrews, J. Jasche, G. Lavaux, and F. Schmidt, *Mon. Not. R. Astron. Soc.* **520**, 5746 (2023).
  - [2] J. Jasche and B. D. Wandelt, *Mon. Not. R. Astron. Soc.* **432**, 894 (2013).
  - [3] J. Jasche, F. Leclercq, and B. D. Wandelt, *J. Cosmol. Astropart. Phys.* 01 (2015) 036.
  - [4] G. Lavaux and J. Jasche, *Mon. Not. R. Astron. Soc.* **455**, 3169 (2016).
  - [5] F. Schmidt, *J. Cosmol. Astropart. Phys.* 04 (2021) 033.
  - [6] J. Stadler, F. Schmidt, and M. Reinecke, *J. Cosmol. Astropart. Phys.* 10 (2023) 069.
  - [7] A. J. Zhou and S. Dodelson, *Phys. Rev. D* **108**, 083506 (2023).
  - [8] J. Alsing, A. F. Heavens, and A. H. Jaffe, *Mon. Not. R. Astron. Soc.* **466**, 3272 (2017).
  - [9] S. S. Boruah, E. Rozo, and P. Fiedorowicz, *Mon. Not. R. Astron. Soc.* **516**, 4111 (2022).
  - [10] P. Fiedorowicz, E. Rozo, and S. S. Boruah, *arXiv:2210.12280*.
  - [11] A. Loureiro, L. Whiteway, E. Sellentin, J. S. Lefaurie, A. H. Jaffe, and A. F. Heavens, *Open J. Astrophys.* **6** (2023), 10.21105/astro.2210.13260.
  - [12] N. Porqueres, A. Heavens, D. Mortlock, G. Lavaux, and T. L. Makinen, *arXiv:2304.04785*.
  - [13] M. Millea, E. Anderes, and B. D. Wandelt, *Phys. Rev. D* **102**, 123542 (2020).
  - [14] M. Millea *et al.*, *Astrophys. J.* **922**, 259 (2021).
  - [15] S. S. Boruah and E. Rozo, *Mon. Not. R. Astron. Soc.* **527**, L162 (2023).



- [16] F. Leclercq and A. Heavens, *Mon. Not. R. Astron. Soc.* **506**, L85 (2021).
- [17] N.-M. Nguyen, F. Schmidt, G. Lavaux, and J. Jasche, *J. Cosmol. Astropart. Phys.* **03** (2021) 058.
- [18] H. Aihara *et al.*, *Publ. Astron. Soc. Jpn.* **70**, S4 (2018).
- [19] X. Li *et al.*, *Publ. Astron. Soc. Jpn.* **74**, 421 (2022).
- [20] D. Spergel *et al.*, [arXiv:1503.03757](https://arxiv.org/abs/1503.03757).
- [21] X. Li *et al.*, *Phys. Rev. D* **108**, 123518 (2023).
- [22] R. Dalal *et al.*, *Phys. Rev. D* **108**, 123519 (2023).
- [23] M. Bartelmann and P. Schneider, *Phys. Rep.* **340**, 291 (2001).
- [24] D. N. Limber, *Astrophys. J.* **117**, 134 (1953).
- [25] A. Lewis, A. Challinor, and A. Lasenby, *Astrophys. J.* **538**, 473 (2000).
- [26] Z. Li, J. Liu, J. M. Z. Matilla, and W. R. Coulton, *Phys. Rev. D* **99**, 063527 (2019).
- [27] J. Liu, A. Petri, Z. Haiman, L. Hui, J. M. Kratochvil, and M. May, *Phys. Rev. D* **91**, 063507 (2015).
- [28] V. Böhm, S. Hilbert, M. Greiner, and T. A. Enßlin, *Phys. Rev. D* **96**, 123510 (2017).
- [29] S. Hilbert, J. Hartlap, and P. Schneider, *Astron. Astrophys.* **536**, A85 (2011).
- [30] H. S. Xavier, F. B. Abdalla, and B. Joachimi, *Mon. Not. R. Astron. Soc.* **459**, 3693 (2016).
- [31] R. Mandelbaum *et al.*, *Astrophys. J. Suppl. Ser.* **212**, 5 (2014).
- [32] M. A. Troxel and M. Ishak, *Phys. Rep.* **558**, 1 (2015).
- [33] E. Tsaprazi, N.-M. Nguyen, J. Jasche, F. Schmidt, and G. Lavaux, *J. Cosmol. Astropart. Phys.* **08** (2022) 003.
- [34] E. Krause *et al.*, [arXiv:2105.13548](https://arxiv.org/abs/2105.13548).
- [35] M. M. Rau, R. Dalal, T. Zhang, X. Li, A. J. Nishizawa, S. More, R. Mandelbaum, H. Miyatake, M. A. Strauss, and M. Takada, *Mon. Not. R. Astron. Soc.* **524**, 5109 (2023).
- [36] R. Takahashi, T. Hamana, M. Shirasaki, T. Namikawa, T. Nishimichi, K. Osato, and K. Shiroyama, *Astrophys. J.* **850**, 24 (2017).
- [37] G. Hinshaw *et al.*, *Astrophys. J. Suppl. Ser.* **208**, 19 (2013).
- [38] M. Shirasaki, T. Hamana, M. Takada, R. Takahashi, and H. Miyatake, *Mon. Not. R. Astron. Soc.* **486**, 52 (2019).
- [39] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelman, *Astrophys. J.* **622**, 759 (2005).
- [40] A. Zonca, L. Singer, D. Lenz, M. Reinecke, C. Rosset, E. Hivon, and K. Gorski, *J. Open Source Software* **4**, 1298 (2019).
- [41] R. Mandelbaum *et al.* (The LSST Dark Energy Science Collaboration), [arXiv:1809.01669](https://arxiv.org/abs/1809.01669).
- [42] D. Phan, N. Pradhan, and M. Jankowiak, [arXiv:1912.11554](https://arxiv.org/abs/1912.11554).
- [43] E. Bingham *et al.*, [arXiv:1810.09538](https://arxiv.org/abs/1810.09538).
- [44] M. D. Hoffman and A. Gelman, *J. Mach. Learn. Res.* **15**, 1593 (2014).
- [45] R. M. Neal, MCMC using Hamiltonian dynamics, in *Handbook of Markov Chain Monte Carlo*, 1st ed., edited By Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng (Chapman and Hall/CRC, New York, 2011), p. 619.
- [46] N. Kaiser and G. Squires, *Astrophys. J.* **404**, 441 (1993).
- [47] J. W. Kirchner, *Phys. Rev. E* **71**, 066110 (2005).
- [48] E. Sefusatti, M. Crocce, R. Scoccimarro, and H. Couchman, *Mon. Not. R. Astron. Soc.* **460**, 3624 (2016).
- [49] N. Tessore, A. Loureiro, B. Joachimi, M. von Wietersheim-Kramsta, and N. Jeffrey, *Open J. Astrophys.* **6** (2023), [10.21105/astro.2302.01942](https://doi.org/10.21105/astro.2302.01942).
- [50] T. Matsubara, *Astrophys. J. Suppl. Ser.* **170**, 1 (2007).
- [51] J. Qin, J. Pan, Y. Yu, and P. Zhang, *Mon. Not. R. Astron. Soc.* **514**, 1548 (2022).
- [52] T. Zhang *et al.*, *Mon. Not. R. Astron. Soc.* **525**, 2441 (2023).
- [53] Y. Kobayashi, T. Nishimichi, M. Takada, and H. Miyatake, *Phys. Rev. D* **105**, 083517 (2022).
- [54] H. Miyatake *et al.*, *Phys. Rev. D* **108**, 123517 (2023).
- [55] T. Nishimichi *et al.*, *Astrophys. J.* **884**, 29 (2019).