KaRMa 2.0 - Kappa Reconstruction for Mass Mapping

Pier Fiedorowicz1*, Eduardo Rozo1, Supranta S. Boruah2
1Department of Physics, University of Arizona, Tucson, AZ 85721, USA
2Steward Observatory, University of Arizona, Tucson, AZ 85719, USA

ABSTRACT
We present KaRMa 2.0, an updated version of the mass map reconstruction code introduced in Fiedorowicz et al. (2022). KaRMa is a full-sky Bayesian algorithm for reconstructing weak lensing mass maps from shear data. It forward-models the convergence field as a realization of a lognormal field. The corresponding shear map is calculated using the standard Kaiser–Squires transformation, and compared to observations at the field level. The posterior distribution of maps given the shear data is sampled using Hamiltonian Monte Carlo chains. Our work improves on the original algorithm by making it numerically efficient, enabling full-sky reconstructions at ≈ 7 arcmin resolution with modest computational resources. These gains are made with no loss in accuracy or precision relative to KaRMa 1.0. We compare the KaRMa 2.0 posteriors against simulations across a variety of summary statistics (one-point function, two-point functions, and peak/void counts) to demonstrate our updated algorithm provides an accurate reconstruction of the convergence field at mildly non-linear scales. Unsurprisingly, the lognormal model fails as we approach non-linear scales (∆ ≥ 200), which in turn biases the map posteriors. These biases are at the 2% level in the recovered power spectrum, and at the 5% to 15% level for other statistics, depending on the resolution.

Key words: cosmology: dark matter, large-scale structure of Universe

1 INTRODUCTION

Weak gravitational lensing provides one of the most powerful methods for studying the mass distribution of the Universe and its large-scale structure. Because weak lensing is sensitive to all forms of matter, it allows us to measure the total matter density field of the Universe, including dark matter. As such, all modern photometric redshift surveys include weak lensing measurements as one of their key cosmological probes (Hikage et al. 2019; Heymans et al. 2021; Abbott et al. 2022). These current generation surveys measure the shapes of hundreds of millions of galaxies, while next-generation surveys like LSST and Euclid will observe billions of galaxies. Optimally constraining cosmology from these observations is critical, particularly as we are currently facing a ∼ 2–3σ tension in cosmological constraints between observations of the early and late Universe (Heymans et al. 2021).

Nonlinear structure growth in the late Universe causes the density field to become highly non-Gaussian. However, traditional cosmology analyses have generally relied upon constraining cosmology with two-point statistics, which fully characterize the statistical properties of field if and only if the field is Gaussian. Indeed, previous work has demonstrated that these non-Gaussianities contain significant additional cosmological information (Gatti et al. 2020). For this reason, many different methods have been developed for extracting non-Gaussian information from weak gravitational lensing data. A brief list of these methods includes peak and void count statistics (Dietrich & Hartlap 2010; Kratochvil et al. 2010; Peel et al. 2017; Shan et al. 2017; Zürcher et al. 2022), bispectra (Takada & Jain 2004; Fu et al. 2014; Jung et al. 2021), higher order moments (Peel et al. 2017; Gatti et al. 2020), Minkowski functionals (Kratochvil et al. 2012; Petri et al. 2013; Vicinanza et al. 2019), machine-learning (Gupta et al. 2018; Fluri et al. 2018; Röbi et al. 2019; Jeffrey et al. 2020b), and scattering transforms (Cheng et al. 2020).

Many methods for extracting non-Gaussian information generally assume that the summary statistics of interest is calculated on weak lensing mass maps made from the shear data. Traditionally the convergence field is recovered through the Kaiser–Squires reconstruction (Kaiser & Squires 1993). However, the quality of the Kaiser–Squires reconstruction is severely degraded by masking effects and survey noise. As a result, a number of mass map reconstruction techniques have been proposed (Wiener 1949; Jeffrey et al. 2018; Starck et al. 2021; Jeffrey et al. 2020a; Forqueres et al. 2021a; Remy et al. 2020; Alsing et al. 2015). However, the goal of reconstructing mass maps that are statistically indistinguishable from simulated maps has remained elusive. Even with improved mass map reconstructions, summary statistic-based approach are inherently sub-optimal: any summary statistics based on a mass map will necessarily lose some of the information contained in the map itself.

An optimal approach for cosmology from weak lensing surveys would rely on field-based inference. That is, instead of trying to model a lossy summary statistic of the field, one would forward-model the full observed shear field. Initial estimates of the potential gain from such methods were remarkably high (Porqueres et al. 2021a), but have since been revised downwards significantly for vanilla (i.e., ΛCDM) cosmologies (Boruah et al. 2022). However, in an upcoming work, we will demonstrate that the factor-of-two improvement in cosmological
constraints enabled by field-based inference is restored in wCDM cosmologies. Perhaps most importantly, forward modeling of the lensing field opens the door to modeling systematics at the map level (Porqueres et al. 2019; Tsapras et al. 2022), potentially enabling internal self-calibration methods for ameliorating systematics (Pyne & Joachimi 2021).

Critical to field-based inference is a forward model of the shear field from cosmological parameters. The first successful implementation was that of Alsing et al. (2017), who used a Gaussian prior to model the weak lensing fields. Porqueres et al. (2021a) used a simulation-based technique to non-linearly evolve the density field from its initial conditions to the present day, and then integrated along the line of sight to produce the convergence field. Each approach has advantages and disadvantages: the Alsing et al. (2017) model is numerically simpler, but fails to account for non-linear physics in the generative model, while the Porqueres et al. (2021a) analyses is computationally challenging. Specifically, modeling the full 3D density field is computationally expensive, making the simulation approach especially challenging to extend to wide angle surveys. In short, delivering on the promise of field-based inference requires the a fast, full-sky forward model capable of capturing the impact of non-linear structure growth.

In our previous work we presented KaRMMa, a fully Bayesian framework for forward modeling the shear field and improving mass map reconstruction using a fast approximate model of the convergence field (Fiedorowicz et al. 2022). Prior work had demonstrated that the weak lensing convergence field can be accurately modeled as a log-normal random field (Taruya et al. 2002; Clerkin et al. 2016; Xavier et al. 2016) at mildly nonlinear scales. In addition, Boruah et al. (2022) demonstrated that using the lognormal model for field-level cosmology analyses results in notable gains over using a Gaussian model. For these reasons, KaRMMa models the convergence field as a lognormal random field. This first iteration of KaRMMa – KaRMMa 1.0 – successfully demonstrated the application of the lognormal model in a forward modeling-based approach for mass map reconstruction Fiedorowicz et al. (2022). In this work, we present KaRMMa 2.0, an updated version of KaRMMa. KaRMMa 2.0 improves on KaRMMa 1.0 by enabling full-sky mass map reconstructions at high resolution, a problem that was previously numerically intractable. We demonstrate that because our updated code can now operate at small, highly nonlinear scales, the failure of the log-normal model to fully capture non-linear growth now impacts mass map reconstruction at detectable levels. Nevertheless, KaRMMa 2.0 outperforms existing mass map reconstruction techniques, while being the only Bayesian reconstruction algorithm capable of running on wide-field surveys.

This paper is laid out as follows. In Section 2 we briefly review weak lensing theory and the lognormal model. We then detail our conclusions in Section 4. In Section 2 we test KaRMMa 2.0 using simulations and present our results. Subsequently, we discuss how KaRMMa’s performance compares to other reconstruction techniques. Finally, we summarize this work and present our conclusions in Section 4.

2 THE KaRMMa 2.0 ALGORITHM

2.1 Weak Lensing Formalism

Mass along the line of sight slightly distorts the observed shapes of source galaxies, resulting in magnification and shearing of the original image. For a perfectly circular source, the image becomes elliptical with ellipticity

$$\epsilon = \frac{\gamma}{1 - k}. \quad (1)$$

The convergence $k$ parametrizes the magnification of the source, and the shear is parametrized by the complex value $y$. In the weak lensing limit, $k \ll 1$ and therefore $\epsilon \approx y$. That is, by measuring the ellipticity of galaxies, we can recover a noisy estimate of the lensing shear field. In practice, we cannot measure the shear of a single source, as we do not know the shape of the undistorted source image. However, by averaging over a large number of sources (whose original ellipticities we assume are unaligned), we can measure the average value of the shear field in a given region of the sky. Finally, since convergence and shear are related, measurements of the shear enables us to infer the convergence field of the Universe.

The weak lensing shear and convergence field are both sourced by the distribution of matter along the line of sight to a source. In the Born approximation, the relation between these fields and the mass distribution of the Universe along a line of sight $\theta$ is

$$\kappa(\theta) = \int_0^\infty g(\chi)\delta(\chi, \theta)d\chi \quad (2)$$

where $\chi$ is the comoving distance, and $\delta = \rho/\bar{\rho} - 1$ is the matter density contrast. The lensing kernel $g(\chi)$ determines the sensitivity of the convergence to the density field as a function of distance, and is set by the source redshift distribution $p(\chi)$,

$$g(\chi) = \frac{3H_0^2}{2c^2} \Omega_m \int_\chi^\infty p(\chi') \frac{\chi'(\chi' - \chi)}{\chi'}d\chi'. \quad (3)$$

Note that in this equation we have assumed a flat Universe.

The shear field is related to the convergence field through the Kaiser–Squires relationship (Kaiser & Squires 1993). On the spherical sky, this relationship is most easily expressed in terms of the spherical harmonics (Castro et al. 2005),

$$\gamma_{\ell m} = -\sqrt{\frac{(\ell - 1)(\ell + 2)}{\ell(\ell + 1)}} \kappa_{\ell m}, \quad (4)$$

where $\gamma_{\ell m}$ and $\kappa_{\ell m}$ are the spherical harmonic components of the shear and convergence fields. Note that the shear field is a spin-2 field and requires spin-2 spherical harmonic transformations. This relationship allows one to predict the shear field from the convergence field. In the absence of a data mask and noise, one can reconstruct the convergence field from the shear field by simply inverting this relationship. Indeed, this approach is often taken in practice (e.g., Jeffrey et al. 2021; Oguri et al. 2017) even though the resulting mass maps can be severely biased (as demonstrated in Jeffrey et al. 2021 and Fiedorowicz et al. 2022). The KaRMMa 2.0 algorithm avoids these difficulties by forward modeling the shear field based on a lognormal model for the lensing convergence.

2.2 The Lognormal Model

The lognormal model assumes that the convergence field can be written as a simple non-linear transform of a Gaussian field. That is,

$$\kappa_i = e^{y_i - \lambda}, \quad (5)$$

where $y_i$ is a Gaussian random field and $\lambda$ is the so-called “shift” parameter. The correlation function of $y_i$ is related to the correlation function of the convergence field via

$$\xi_{yy}(\theta) = \log \left[ \frac{\xi_{kk}(\theta)}{\lambda^2 + 1} \right]. \quad (6)$$
The mean of \( y_i \) is set by the constraint that the average convergence field is zero (\( \langle \kappa \rangle = 0 \)). One finds

\[
\langle y_i \rangle = \ln \lambda - \frac{\xi_{yy}(0)}{2}.
\] (7)

The shift parameter, \( \lambda \), and correlation function, \( \xi_{kk}(\theta) \), are specified by the cosmological parameters. The former can be fit from simulations\(^1\) and the latter using tools like CAMB (Lewis et al. 2000) or CLASS (Blas et al. 2011). We use PyCCL (Chisari et al. 2019b) which calls CAMB to compute the non-linear matter power spectrum using halofit (Smith et al. 2003; Takahashi et al. 2012). When applying our code to the mocks described in Section 3 we then use the prescription from Takahashi et al. (2017a) to correct the power spectrum for simulation artifacts arising from the ray-tracing scheme and the finite angular resolution of the mocks. After applying the provided correction, we find residual artifacts remain at the few percent level, which we remove by fitting a small multiplicative correction to the power spectrum measured. The necessary bias is fit using a regularized polynomial regression (polynomial order 5, regularization \( \alpha = 0.5 \)). For more information on the lognormal model, we refer the reader to Xavier et al. (2016).

### 2.3 Improvements Over KaRMMa 1.0

In Fiedorowicz et al. (2022), we characterized the convergence field in configuration space: the value of the field \( y \) at each pixel was a model parameter, and a Gaussian prior was imposed on the resulting map. This approach suffers from a critical deficiency: the covariance matrix of the resulting mass map is both extremely large and dense (i.e. not diagonal). To invert this matrix, KaRMMa 1.0 relied on Singular Value Decomposition (SVD), an \( O(n^3) \) operation. Our algorithm quickly became compute and memory limited at a resolution of \( \approx 14 \) arcmin in a DES Y1-footprint (\( \approx 28 \) arcmin for a DES Y5-footprint).

In this work, we switch from using the real space values of the pixelized field \( y \) to using the corresponding spherical harmonic coefficients. Thanks to the homogeneity and isotropy of the Universe, these harmonics are uncorrelated. The corresponding covariance matrix is therefore diagonal, with the entries corresponding to the power-spectrum of the field: \( \langle \delta_{lm} \delta_{lm'} \rangle = C_l \delta_{l,l'} \delta_{m,m'} \). In this approach, the most computationally expensive operations performed as part of the modeling pipeline are the spherical harmonic transforms, which scale as \( O(n^{1.5}) \). Consequently, KaRMMa 2.0 easily outperforms KaRMMa 1.0 in run time while benefiting from a drastically smaller memory footprint. Moreover, by modeling the spherical harmonic coefficients we reconstruct the convergence map across the full spherical sky, thereby eliminating masking artifacts in our reconstruction.

In practice, our model parameters are not the \( x_{lm} \) themselves, but rather the quantities

\[
x_{lm} = \frac{y_{lm}}{\sqrt{C_l}}.
\] (8)

Consequently, the set of spherical harmonics \( \{x_{lm}\} \) is normally distributed with a unit covariance matrix,

\[
P(x) \propto \exp \left( -\frac{1}{2} x^T \Sigma x \right).
\] (9)

This parametrization helps decorrelate the cosmological parameters from the relative phases and amplitudes of \( \kappa \) (Boruah et al. 2022). While KaRMMa 2.0 does not yet allow for marginalization over cosmological parameters when inferring mass maps, we intend to include this feature in future releases of the code.

### 2.4 Likelihood Model and Parameter Sampling

We model the observed shear as a noisy version of the true shear field. The latter is related to the convergence field via equation 4. We assume the source density is large enough that the shape noise in each spatial pixel can modeled using a Gaussian distribution. The full posterior distribution in our model is therefore

\[
P(\tilde{\kappa} \mid \tilde{y}_{\text{obs}}) \propto P(\tilde{\kappa}) \times P(\tilde{y}_{\text{obs}} \mid \tilde{\kappa})
\]

\[
\propto P(\tilde{\kappa}) \times \exp \left[ -\frac{1}{2} \sum_i (y_{i,\text{obs}} - y_{i}(\tilde{\kappa}))^2 \sigma^2_{y,i} \right].
\] (10)

To compute \( y(\tilde{\kappa}) \), we proceed as follows:

(i) For each spherical harmonic, set \( y_{lm} = \sqrt{C_l} x_{lm} \).
(ii) Perform a spherical harmonic transform to get \( y \) in pixel-space.
(iii) Perform the lognormal transformation to arrive at \( \kappa \).
(iv) Use the Kaiser-Squires relation to obtain \( \gamma \) via spherical harmonic transforms.

We sample the \( x_{lm} \) parameters using the Pyro (Bingham et al. 2019) probabilistic programming language. Pyro takes advantage of PyTorch for autograd calculations and GPU acceleration, enabling efficient Hamiltonian Monte Carlo (Neal 2012) chain sampling. To use Pyro with the HEALPix spherical harmonic transforms, we implemented custom PyTorch autograd functions\(^2\), which call the underlying HEALPix routines for the forward and backward (gradient) calculations. Unfortunately, we cannot take advantage of GPU acceleration because the HEALPix spherical harmonic routines do not have a GPU implementation. Should such routines become available in the future, we anticipate that KaRMMa 2.0 run time would dramatically decrease. For sampling, we use the NUTS (Hoffman & Gelman 2011) algorithm which automatically tunes the step size and trajectory length for HMC, thereby reducing the autocorrelation between samples. We perform 100 steps of burn-in followed by 200 samples for each MCMC chain generated in this work. In Appendix A, we demonstrate that despite the small number of samples — particularly given the high dimensionality of our parameter space — the posterior distribution of the summary statistics considered is already well sampled by this point. We attribute this remarkable convergence rate to the fact that wide survey footprints already contain many independent sky patches. Consequently, one realization of the survey area contains within it many nearly-independent samples. This fact boosts the number of “independent realizations” used in characterizing the distribution of weak lensing summary statistics, enabling the posterior distributions of our summary statistics to be well converged. Of course, the full posterior of our mass maps is not adequately sampled using only 200 mass maps.
Figure 1. Visual comparison of map reconstructions from KaRMMa performed at several resolutions ($N_{\text{side}} = 128, 256,$ and 512). Each reconstruction is performed for the same mock realization but at a different resolution. For each resolution, we show a randomly chosen convergence map as sampled from the posterior distribution. In the bottom right corner, we show the true convergence map at high resolution ($N_{\text{side}} = 2048$). As we increase the reconstruction resolution, we can see that we begin to recover smaller-scale features of the underlying convergence field. This figure highlights how our updated method can now efficiently forward model the weak lensing field at high resolution.

3 SIMULATION TESTS

3.1 Method

We use the suite of 108 full-sky weak lensing simulations generated for the HSC survey in Takahashi et al. (2017a)\textsuperscript{3} to test the KaRMMa 2.0 algorithm. The simulations were generated with a fixed flat-$\Lambda$CDM cosmology with cosmological parameters set by the WMAP Y9 (Hinshaw et al. 2013) results: $\Omega_m = 0.279$, $h = 0.7$, $\sigma_8 = 0.82$, and $n_s = 0.97$. We use the simulations to construct DES Y3-like mocks using the Y3 survey mask, shape noise, and non-tomographic source redshift distribution ($dn/dz$) (Gatti et al. 2021).

To add shape noise, we first randomly draw the number of galaxies in each pixel as a Poisson realization given the expected number of galaxies. We then add Gaussian shape noise to the shears with variance $\sigma_\gamma^2 = \sigma_e^2/N_i$ where $\sigma_e$ is the shape noise and $N_i$ is the number of source galaxies in pixel $i$. We construct our mock shear maps using the HEALPix (Gorski et al. 2005) pixelization scheme of the sphere at HEALPix resolutions of $N_{\text{side}} = 128, 256,$ and 512.

For each of the 108 simulations, and each of the three resolutions we considered, we run the KaRMMa 2.0 algorithm. HEALPix maps suffer aliasing such that only modes up to $\ell_{\text{max}} = 2N_{\text{side}}$ can be accurately recovered from a map at resolution $N_{\text{side}}$. For this reason, we reconstruct only the low-pass filtered convergence field up to $\ell_{\text{max}} = 2N_{\text{side}}$. Because of mode-mixing introduced by the exponentiation in equation 5, if $y$ is band-limited to $\ell_{\text{max}} = 2N_{\text{side}}$, the resulting $\kappa$ field will not be band-limited; instead, it has slightly ($< 1\%$) suppressed power for $\ell$ approaching $\ell_{\text{max}}$, and a small amount of power at $\ell > \ell_{\text{max}}$. Removing the excess high $\ell$ power with a low-pass filter is not a good solution: the bias at $\ell > \ell_{\text{max}}$ remains, while the removed power degrades the resulting 1-point distributions. Instead, we model the $y$ field up to $\ell = 3N_{\text{side}} - 1$ (i.e., $y_{\ell m} = 0$ for $\ell >= 3N_{\text{side}} - 1$), and low-pass filter the resulting $\kappa$ map to $\ell_{\text{max}} = 2N_{\text{side}}$. Doing so ensures that the low-pass filtered convergence field has both the correct power spectrum and one-point distribution.

Figure 1 shows a visual comparison between random samples from the KaRMMa 2.0 posteriors along with the true convergence field. The true map is shown at a very high resolution ($N_{\text{side}} = 2048$, or just under 1 arcmin.) We see that: 1) all the maps have similar large scale structure; and 2) the resolution increases steadily as we move from $N_{\text{side}} = 128$ to $N_{\text{side}} = 512$ without degrading the qualitative aspect of the maps. This can be better seen in Figures 2 and 3, which zoom into 10 deg $\times$ 10 deg zoom patches centered on: A) the largest

\textsuperscript{2} https://pytorch.org/docs/stable/notes/extending.html#extending-torch-autograd

\textsuperscript{3} http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allsky_raytracing/
density peak in the sky (Fig. 2); and B) the least dense void in the sky (Fig. 3) for one randomly chosen simulation.

We characterize the performance of KaRMMa 2.0 using a range of summary statistics, namely: 1) the binned power spectrum; 2) the correlation function; 3) the one-point function; 4) peak counts; and 5) void counts. We stress that in these tests we are comparing the summary statistics of our maps marginalized over the full posterior distribution of mass maps, as opposed to calculating the summary statistics of the single maximum-posterior mass map. This is an important point: the maximum a posteriori maps suppress unresolved density fluctuations, giving rise to biased summary statistics. For further discussion, we refer the reader to Friedorowicz et al. (2022).

The comparison is performed as follows. For a given cosmology $\theta$ and summary-statistic $\bar{s}$, we calculate the mean and covariance matrix of the distribution

$$P(\bar{s}|\theta) = \int P(\bar{s}|\tilde{x})P(\tilde{x}|\theta)d\tilde{x}. \quad (11)$$

To do so, we calculate the summary statistics $\bar{s}$ in each of the maps in the posterior. We then evaluate the mean and covariance matrix of $\bar{s}$ across all posterior samples. The mean summary statistic $\langle \bar{s} \rangle$ is compared to the mean statistic evaluated from the input simulations. Because KaRMMa 2.0 produces low-pass filtered maps, we low-pass filter the simulations at a high resolution and then downgrade to the resolution of the KaRMMa maps before computing the summary statistics in the simulation. The low pass filtering at high resolution is necessary to avoid aliasing in our “truth” maps.

### 3.2 Results

Figures 4 - 8 compare the KaRMMa 2.0 posteriors for each of the summary statistics to “truth” as determined from the simulations. All plots share a common layout where the left panel overplots the KaRMMa and simulation summary statistics at each of the resolutions tested. Note the simulation summary statistics change with resolution since the simulation map is always low-pass filtered at $\ell_{\text{max}} = 2N_{\text{side}}$. The bands in the left panel are the 1$\sigma$ scatter on the summary statistics as measured using the KaRMMa posterior distributions. The three panels on the right show the residuals for the summary statistics at each of three resolutions we consider. Unlike the left panel, the colored bands here are the error on the mean residual as estimated from the 108 independent simulations, i.e. they are smaller than the bands in the left-hand plot by a factor of $\sqrt{108}$. The black bands centered on zero are the error on the mean due to cosmic variance in the simulations (i.e., they do not involve KaRMMa in any way).

Figure 4 compares the power spectrum in the simulations and at each resolution. We avoid the prior-dominated parts of the sky by setting the convergence field to zero outside the survey boundary for both simulations and the KaRMMa posteriors. Since the mask effects are identical, we have not corrected the power spectrum measurements for mask effects. We find that the samples generated by KaRMMa correctly recover the power spectrum at large scales. However, at small non-linear scales the reconstructed power is larger than that in simulations by $\approx 2\%$. Because we are able to make this comparison using 108 simulations, this small excess is detected with high significance.

The vertical red line in the residual plots is the non-linear scale $\ell_{NL}$ defined as follows. First, given a source at redshift $z$, we set $\ell_{NL}(z) = d_A(z)k_{NL}$ with $k_{NL}(z)$ predicted by P2CCL as the inverse variance of the displacement field from linear theory. The non-linear scale for a tomographic bin is set by averaging $\ell_{NL}(z)$ over the source redshift distribution. The fact that the KaRMMa posterior for the two-point function starts to deviate from truth at $\ell \gtrsim \ell_{NL}$ demonstrates the lognormal model fails on strongly non-linear scales. Indeed, figure 5 compares the corresponding correlation function posteriors. There, we see the bias of the KaRMMa posteriors at zero angular separation increases with increasing resolution. The vertical red line in figure 5 corresponds to the non-linear angular scale $\ell_{NL} = \pi/\ell_{NL}$.

In principle, the recovered power is subject to a second source of bias: aliasing. In real data, the observed shear field is generated by an underlying density field that has power at all scales. When averaging over sky pixels, power in unresolved modes can “leak” into the observed angular ranges, leading to an artificial boost in power (aliasing). This leakage would in turn bias our posteriors. To test the impact of aliasing, we reran our analyses but modified the input shear field so that it contained no high $\ell$ power. That is, our true underlying density field (and the corresponding shear field) only contains well resolved modes that aren’t aliased. Shape noise is then added to the low-pass filtered shear field in each pixel independently (as would be normally) to produce the observed shear field for this test. The posteriors recovered in this case are marginally less biased than in the run subject to aliasing. This test demonstrates that the bias in Figures 4 and 5 is dominated by the failure of the log-normal model on non-linear scales. Aliasing from high $\ell$-power in the shear maps is only a minor contaminant in our reconstructions.

Figure 6 compares the posteriors of the 1-point function to simulations. At the lowest resolution of $N_{\text{side}} = 128$ (27 arcmin), we find that the KaRMMa samples recover the one-point function in simulations without any bias at all. However, the agreement between the KaRMMa posteriors and simulations becomes increasingly worse with increasing resolution. At $N_{\text{side}} = 512$ (7 arcmin), the difference in the 1-point function can be as large as 15%. Figures 7 and 8 comparing the posteriors of the peak and void counts are qualitatively similar to Figure 6: good agreement at low ($N_{\text{side}} = 128$), but increasing differences as resolution improves.

### 3.3 Comparison to Other Work

There are many proposals in the literature for how to improve on the traditional Kaiser–Squires algorithm for mass map reconstruction. These broadly fall into two categories: “best guess” estimators and fully probabilistic reconstructions.

“Best guess” estimators produce a single map which maximizes a specified metric or likelihood function. This class of methods include:

(i) Wiener filtering (Wiener 1949): Wiener filtering reconstructs the maximum posterior map under the assumption the convergence field is a Gaussian random field.

(ii) GLIMPSE (Lanusse et al. 2016; Jeffrey et al. 2018): GLIMPSE returns the maximum posterior map after adopting a prior of sparsity on the wavelet coefficients of the mass map.

(iii) MCALens (Starck et al. 2021): MCALens combines both the Gaussian and sparsity priors to arrive at a single “best guess” map.

(iv) DeepMass (Jeffrey et al. 2020a): DeepMass uses machine learning to predict the convergence map by minimizing the mean squared error reconstruction loss.

Jeffrey et al. (2021) studied a variety of mass map reconstruction methods, including Wiener filtering and GLIMPSE. They found these two methods clearly outperform the Kaiser–Squires algorithm when comparing the pixel-level reconstruction quality (Pearson correlation coefficient and mean squared error). However, they also fail to recover the correct statistical properties of the true mass maps. This failure is generic to all “best guess” mass map reconstruction efforts: “best guess” maps suppress the amplitude of noise-dominated
Fiedorowicz et al.

Figure 2. Sample KaRMMa maps zoomed in on a 10 deg. × 10 deg. patch centered around the densest peak in the sky for a randomly chosen simulation. Each panel corresponds to a different resolution, as labelled. The bottom-right panel is the true density field sampled at high ($N_{side} = 2048$) resolution.

Figure 3. Same as Figure 2, but now zoomed in on a 10 deg. × 10 deg. patch centered around the least dense void in the sky.

modes, resulting in a reduction in power at small scales. For example, at $\ell \sim 500$, the reconstructed power spectrum using either Wiener filtering or Glimpse was biased low by more than 80%. By contrast, at the same scale, the samples generated by KaRMMa recover the true power spectrum at the level of 2% error.

Probabilistic reconstruction algorithms — which includes KaRMMa — sample the probability distribution $P(\psi|\gamma_{obs})$. At present, we are aware of three other such algorithms. Porqueres et al. (2021a) approaches this problem by modeling the initial 3D density field of the Universe. The initial conditions are non-linearly evolved using Lagrangian perturbation theory, and integrated along the line of sight to generate 3D density maps of the Universe. Their angular resolution is comparable to ours, but are restricted to relatively small patches of sky for which the flat sky approximation is valid. Based on their Fig. 8, we estimate they recover the true shear power spectrum with an accuracy of $\approx 5\%$, comparable to but slightly worse than KaRMMa. They do not compare their mass maps to simulations using non-Gaussian summary statistics.

The second algorithm we are aware of is that of Remy et al. (2020), who uses machine learning in the form of neural score estimation to sample the posterior distribution of mass maps. Comparing the accuracy of their approach to ours is impossible. While the $\approx 10\%$ accuracy on the power spectrum reported in that work appears to be worse than that of KaRMMa, the two algorithms are operating in completely opposite regimes. Specifically, their simulations are designed to match COSMOS data, which has an area of just 1.64 deg$^2$. Consequently, the largest scales they constrain are roughly the smallest scales we consider. Their success in recovering the power spectrum at such small scales is a testament to the power of machine learning techniques (see also Dai & Seljak 2022; Mustafa et al. 2019; Yiu et al. 2021), though we caution that the large theoretical uncertainty at such small scales (e.g., due to baryonic processes, Chisari et al. 2019a) represent a significant challenge for modeling these scales.

Alsing et al. (2015) developed a probabilistic approach for jointly sampling the posterior distribution of shear maps and the shear power spectrum, with the shear modeled as a Gaussian random field. The authors tested their reconstruction on a 100 deg$^2$ mock survey at an angular resolution of roughly 5 arcmin using the flat sky approximation. The power spectrum of their reconstructed shear maps recovers the power spectrum of the true shear field to within the uncertainty of their posterior distribution. However, as they do not show power spectrum residuals, it is difficult to compare the accuracy of their reconstruction to ours. We note that in Fiedorowicz et al. (2022) we found that adopting a Gaussian prior for the convergence field enables one to recover maps that are of comparable quality to KaRMMa in terms of the accuracy of the resulting power spectra. However, the recovered non-Gaussian summary statistics exhibit a larger bias than those obtained using KaRMMa.

In summary, the statistical properties of mass maps are much better reproduced by probabilistic sampling algorithms than by “best guess” maps. Compared to the other three probabilistic algorithms currently available in the literature, KaRMMa achieves either comparable or somewhat better accuracy in terms of the reconstructed power spectrum. We have also, for the first time, quantitatively characterized the performance of a probabilistic algorithm using non-Gaussian statistics. At the time of writing, KaRMMa is the only such algorithm capable of operating on large survey areas for which the flat sky approximation is invalid.
Figure 4. Left: Comparison of the reconstructed power spectrum from KaRMMa samples to the input simulations at multiple resolutions. The bands represent the 1σ scatter due to cosmic variance. The right panels show the KaRMMa residuals at a resolution of $N_{\text{side}} = 128$ (27 arcmin, top), $N_{\text{side}} = 256$ (14 arcmin, middle), and $N_{\text{side}} = 512$ (7 arcmin, bottom). The bands are the 1σ error on the mean. The horizontal dashed lines are at ± 2%, and the red vertical line marks the non-linear scale. As ℓ approaches the non-linear scale, the recovered power spectrum becomes biased by ± 2% for all resolutions.

4 SUMMARY AND CONCLUSIONS

We have presented an updated version of KaRMMa, an algorithm for mass map reconstruction. KaRMMa forward models the observed shear field subject to a lognormal prior. By reparameterizing the problem in terms of the spherical harmonic coefficients of the transformed convergence field (i.e., $y = \ln(k + \lambda)$), we are able to decorrelate our model parameters in the prior. This feature, in combination with the existence of efficient spherical harmonic transformations enables us to improve the scaling of our algorithm to efficiently perform reconstruction at resolutions that were formerly computationally intractable. As with our first iteration of the algorithm (Fiedorowicz et al. 2022), KaRMMa 2.0 is fully Bayesian and generates a library of maps that sample the posterior distribution of maps consistent with the observed shear field.

Figures 4 through 8 summarize our results. The figures compare the KaRMMa 2.0 posterior samples to simulations for a variety of summary statistics: power spectrum, correlation function, 1-point function, and peak and void counts. In all cases, KaRMMa 2.0 returns accurate (5% or better) posteriors for all statistics at $N_{\text{side}} = 128$ resolution. As resolution increases, the accuracy of the posteriors decreases, with systematic differences in non-Gaussian summary statistics reaching as high as ± 15% at $N_{\text{side}} = 512$.

Given that KaRMMa 2.0 can now return high resolution posteriors, the main difficulty moving forward is our ability to resolve strongly non-linear scales. At these scales, the lognormal approximation fails to be adequate, driving the biases observed in Figures 4 through 8. To overcome this limitation, we need to improve upon the lognormal approximation. Ideally, one would replace the lognormal model with full gravity simulations. Doing so is nontrivial, and remains an active area of research (Jasche & Wandelt 2013; Porqueres et al. 2021b,a). Moving forward, we intend instead to augment the lognormal model with machine-learning techniques that enable us to accurately reconstruct the true non-linear density field (Fiedorowicz et al. in preparation). Future releases of KaRMMa will also enable: 1) tomographic mass map reconstruction (Boruah et al. 2022); and 2) marginalization over cosmological parameters (Boruah et al., in preparation).

Despite its current limitations, we consider KaRMMa 2.0 an unquestionable success. First, even at $N_{\text{side}} = 512$, the ± 15% biases in our posteriors are a dramatic improvement relative to other mass mapping techniques available in the literature (see Section 3.3). Second, our algorithm is one of the few approaches that works in a spherical sky. Third, our algorithm is fully Bayesian, returning proper posteriors as opposed to a single “best guess” mass map. This last point is particularly important, since “best guess” maps are necessarily strongly biased (see Fiedorowicz et al. 2022). For all these reasons, KaRMMa 2.0 represents a significant step forward in our ability to reconstruct the matter density field of the Universe from weak lensing data.

Upon publication, the updated version of the KaRMMa code will
Figure 5. Comparison of the reconstructed correlation function from KaRMMa to the true correlation function. The layout of this plot is the same as in Figure 4. Similarly to the power spectrum, we find that the correlation function from KaRMMa deviates from truth at small scales and worsens as we increase the reconstruction resolution. The vertical line in the right panels corresponds to the non-linear angular scale $\theta_{NL} = \pi/\ell_{NL}$.

be made publicly available on GitHub at https://github.com/pierfied/karmma.

ACKNOWLEDGEMENTS

ER and PF are supported by NSF grant 2009401. ER also receives support from DOE grant DE-SC0009913.

We would like to thank Takahashi et al. (2017b) and collaborators for making their full-sky weak lensing simulations publicly available. Without them, this work would not have been possible.

DATA AVAILABILITY

The data underlying this article was derived from Takahashi et al. (2017b) which can be accessed at: http://cosmo.phys.hirosaki-u.ac.jp/takahasi/allskyraytracing/. The derived data generated in this research will be shared on reasonable request to the corresponding author.

REFERENCES

Abbott T., et al., 2022, Physical Review D, 105
Alising J., Heavens A., Jaffe A. H., Kiessling A., Wandelt B., Hoffmann T., 2015, Monthly Notices of the Royal Astronomical Society, 455, 4452
Alising J., Heavens A., Jaffe A. H., 2017, MNRAS, 466, 3272
Bingham E., et al., 2019, J. Mach. Learn. Res., 20, 973–978
Blas D., Lesgourgues J., Tram T., 2011, Journal of Cosmology and Astroparticle Physics, 2011, 034–034
Boruah S. S., Rozo E., Fiedorowicz P., 2022, Map-based cosmology inference with lognormal cosmic shear maps, doi:10.48550/ARXIV.2204.13216, https://arxiv.org/abs/2204.13216
Castro P. G., Heavens A. F., Kitching T. D., 2005, Phys. Rev. D, 72, 023516
Cheng S., Ting Y.-S., Ménard B., Bruna J., 2020, Monthly Notices of the Royal Astronomical Society, 499, 5902
Chisari N. E., et al., 2019a, The Open Journal of Astrophysics, 2
Chisari N. E., et al., 2019b, The Astrophysical Journal Supplement Series, 242, 2
Clerkin L., et al., 2016, Monthly Notices of the Royal Astronomical Society, 466, 1444
Dai B., Seljak U., 2022, Translation and Rotation Equivariant Normalizing Flow (TRENF) for Optimal Cosmological Analysis, doi:10.48550/ARXIV.2202.05282, https://arxiv.org/abs/2202.05282
Dietrich J. P., Hartlap J., 2010, Monthly Notices of the Royal Astronomical Society, 402, 1049–1058
Fiedorowicz P., Rozo E., Boruah S. S., Chang C., Gatti M., 2022, Monthly Notices of the Royal Astronomical Society, 512, 73
Fluri J., Kacprzak T., Refregier A., Amara A., Lucchi A., Hofmann T., 2018, Physical Review D, 98
Friedrich O., Uhlemann C., Villaescusa-Navarro F., Baldauf T., Manera M., Nishimichi T., 2020, Monthly Notices of the Royal Astronomical Society, 498, 464–483
Fu L., et al., 2014, Monthly Notices of the Royal Astronomical Society, 441, 2725
Figure 6. Comparison of the reconstructed one-point function from KaRMMa to the true one-point function. The layout of this plot is the same as in Figure 4. At low resolution, we find that KaRMMa does a good job of recovering the true one-point function at about the 5% level. At higher resolutions, the one-point function from KaRMMa becomes biased at around the 15% level, indicating a failure of the lognormal model at small scales.

Gatti M., et al., 2020, Monthly Notices of the Royal Astronomical Society, 498, 4060–4087
Gatti M., et al., 2021, Monthly Notices of the Royal Astronomical Society, 504, 4312
Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, The Astrophysical Journal, 622, 759
Gupta A., Matilla J. M. Z., Hsu D., Haiman Z., 2018, Physical Review D, 97
Heymans C., et al., 2021, Astronomy & Astrophysics, 646, A140
Hikage C., et al., 2019, PASJ, 71, 43
Hinshaw G., et al., 2013, The Astrophysical Journal Supplement Series, 208, 19
Hoffman M. D., Gelman A., 2011, The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo, doi:10.48550/ARXIV.1111.4246, https://arxiv.org/abs/1111.4246
Jasche J., Wandelt B. D., 2013, Monthly Notices of the Royal Astronomical Society, 432, 894–913
Jeffrey N., et al., 2018, Monthly Notices of the Royal Astronomical Society, 479, 2871–2888
Jeffrey N., Lanusse F., Lahav O., Starck J.-L., 2020a, Monthly Notices of the Royal Astronomical Society, 492, 5023–5029
Jeffrey N., Alsing J., Lanusse F., 2020b, Monthly Notices of the Royal Astronomical Society, 501, 954–969
Jeffrey N., et al., 2021, Monthly Notices of the Royal Astronomical Society
Jung G., Namikawa T., Liguori M., Munshi D., Heavens A., 2021, The integrated angular bispectrum of weak lensing (arXiv:2102.05521)
Kaiser N., Squires G., 1993, ApJ, 404, 441
Kratochvil J. M., Haiman Z., May M., 2010, Physical Review D, 81
Kratochvil J. M., Lim E. A., Wang S., Haiman Z., May M., Huffenberger K., 2012, Physical Review D, 85
Lanusse F., Starck J. L., Leonard A., Pires S., 2016, A&A, 591, A2
Lewis A., Challinor A., Lasenby A., 2000, The Astrophysical Journal, 538, 473–476
Mustafa M., Bard D., Bhimji W., Lukić Z., Al-Rfou R., Kratochvil J. M., 2019, Computational Astrophysics and Cosmology, 6
Neal R. M., 2012, MCMC using Hamiltonian dynamics (arXiv:1206.1901)
Oguri M., et al., 2017, Publications of the Astronomical Society of Japan, 70
Peel A., Liu C.-A., Lanusse F., Leonard A., Starck J.-L., Kilbinger M., 2017, Astronomy & Astrophysics, 599, A79
Petri A., Haiman Z., Hui L., May M., Kratochvil J. M., 2013, Physical Review D, 88
Porqueres N., Kodama Matilla J. M., Hsu D., Haiman Z., Csabai I., 2019, MNRAS, 490, 1843
Pyne S., Joachimi B., 2021, MNRAS, 503, 2300
Remy B., Lanusse F., Ramzi Z., Liu J., Jeffrey N., Starck J.-L., 2020, Probabilistic Mapping of Dark Matter by Neural Score Matching (arXiv:2011.08271)
Ribli D., Pataki B. Á., Zorilla Matilla J. M., Hsu D., Haiman Z., Csabai I., 2019, MNRAS, 490, 1843
Shan H., et al., 2017, Monthly Notices of the Royal Astronomical Society, 474, 1116–1134
Smith R. E., et al., 2003, Monthly Notices of the Royal Astronomical Society, 341, 1311
Starck J. L., Themelis K. E., Jeffrey N., Peel A., Lanusse F., 2021, Weak

MNRAS 000, 1–10 (2015)
lensing mass reconstruction using sparsity and a Gaussian random field (arXiv:2102.04127)

Takada M., Jain B., 2004, Monthly Notices of the Royal Astronomical Society, 348, 897

Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, The Astrophysical Journal, 761, 152

Takahashi R., Hamana T., Shirasaki M., Namikawa T., Nishimichi T., Osato K., Shiroyama K., 2017a, The Astrophysical Journal, 850, 24

Takahashi R., Hamana T., Shirasaki M., Namikawa T., Nishimichi T., Osato K., Shiroyama K., 2017b, The Astrophysical Journal, 850, 24

Takada M., Hamana T., Kayo I., Futamase T., 2002, The Astrophysical Journal, 571, 638

Tsaprazi E., Nguyen N.-M., Jasche J., Schmidt F., Lavaux G., 2022, J. Cosmology Astropart. Phys., 2022, 003

Vicinanza M., Cardone V.F., Maoli R., Scaramella R., Er X., Tereno I., 2019, Physical Review D, 99

Wiener N., 1949, Extrapolation, interpolation, and smoothing of Stationary time series: With engineering applications. MIT Press

Xavier H. S., Abdalla F. B., Joachimi B., 2016, Monthly Notices of the Royal Astronomical Society, 459, 3693

Yiu T. W. H., Fluri J., Kacprzak T., 2021, A tomographic spherical mass map emulator of the KiDS-1000 survey using conditional generative adversarial networks. doi:10.48550/ARXIV.2112.12741, https://arxiv.org/abs/2112.12741

Zürcher D., et al., 2022, MNRAS, 511, 2075

**APPENDIX A: CHAIN LENGTH VALIDATION**

Throughout this work we use a default chain length of 200 samples for each mock realization. It is therefore unlikely that our maps fully sample the posterior space of mass maps. In fact, due to the extreme dimensionality of the problem (e.g., ~ 2.5 million parameters for $N_{\text{side}} = 512$ maps), it is virtually impossible to sample the entire parameter space adequately. However, we do find that the statistical properties of the sampled mass maps are well converged. To verify this claim, for one of our mock realizations at $N_{\text{side}} = 128$ we run a chain where we generate 20,000 samples. This chain is 100 times longer than our default chains. We then compare the recovered statistical properties from the default and extended chain lengths. Figure A1 shows the one and two-point statistics of the sample maps for each chain length. From this plot, it is clear that our default chain length of 200 samples is adequate and accurately recovers the statistical properties from the extended runs. As stated previously, we attribute these results to a large survey footprint that effectively contains many independent sky patches.

This paper has been typeset from a TeX/LaTeX file prepared by the author.
Figure 8. Comparison of the reconstructed void counts distribution from KaRMMa to the true void counts distribution. The layout of this plot is the same as in Figure 4. Identically to peak counts, we find that KaRMMa correctly recovers the void counts distribution at low resolution but fails at high resolution.
Figure A1. Comparison between our default chain length of 200 samples and a longer chain with 20,000 samples over two-point (left) and one-point statistics (right) of KaRMMa samples for a single mock realization. The error bars represent the $1\sigma$ uncertainty from the posterior. It is clear that, despite the small number of samples, the statistical properties of the sample maps are well converged.