Optimizing Simulation Parameters for Weak Lensing Analyses Involving Non-Gaussian Observables

José Manuel Zorrilla Matilla$^{1,3}$, Stefan Waterval$^{2,3}$, and Zoltán Haiman$^{1,3}$

$^1$Department of Astronomy, Columbia University, NY, USA; jzorrilla@astro.columbia.edu
$^2$Department of Physics, ETH, Zürich, Switzerland
$^3$Department of Physics, New York University Abu Dhabi, PO Box 129188 Abu Dhabi, United Arab Emirates

Received 2019 September 26; revised 2020 April 28; accepted 2020 April 30; published 2020 May 29

Abstract

We performed a series of numerical experiments to quantify the sensitivity of the predictions for weak lensing statistics obtained in ray-tracing dark matter (DM)-only simulations, to two hyper-parameters that influence the accuracy as well as the computational cost of the predictions: the thickness of the lens planes used to build past light cones and the mass resolution of the underlying DM simulation. The statistics considered are the power spectrum (PS) and a series of non-Gaussian observables, including the one-point probability density function, lensing peaks, and Minkowski functionals. Counterintuitively, we find that using thin lens planes (< 60 h$^{-1}$ Mpc on a 240 h$^{-1}$ Mpc simulation box) suppresses the PS over a broad range of scales beyond what would be acceptable for a survey comparable to the Large Synoptic Survey Telescope (LSST). A mass resolution of $7.2 \times 10^{11} h^{-1} M_{\odot}$ per DM particle (or 256$^3$ particles in a $(240 h^{-1} \text{ Mpc})^3$ box) is sufficient to extract information using the PS and non-Gaussian statistics from weak lensing data at angular scales down to 1' with LSST-like levels of shape noise.

Unified Astronomy Thesaurus concepts: Weak gravitational lensing (1797); Large-scale structure of the universe (902); Computational methods (1965)

1. Introduction

Weak gravitational lensing (WL) enables the mapping of the distribution of dark matter (DM) in the universe on large scales and as a result is a powerful probe to infer cosmological parameters such as $\sigma_8$ and $\Omega_m$ (see comprehensive reviews by, e.g., Bartelmann & Schneider 2001; Hoekstra & Jain 2008 and Kilbinger 2015). In practice, the lensing signal can be extracted from statistical measurements of the shapes of background galaxies, distorted by deflections in the path of light rays as they traverse the vicinity of matter over- and under-densities.

Upcoming surveys such as the Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration et al. 2016), the Large Synoptic Survey Telescope (LSST; LSST Science Collaboration et al. 2009), Euclid (Refregier et al. 2010), the Wide Field Infrared Survey Telescope (WFIRST; Spergel et al. 2015), and the Square Kilometer Array (SKA; Weltman et al. 2020), will provide WL data of unprecedented quality and quantity and will require correspondingly accurate and precise models to extract information from these data sets. WL observables delves into the nonlinear regime, which can be forward-modeled by ray-tracing photons through high-resolution simulated DM density fields (Schneider & Weiss 1988; Jain et al. 2000; Vale & White 2003; Hilbert et al. 2009; Heitmann et al. 2010). Simulating the large volumes encompassed by future surveys is computationally expensive, especially when a high-dimensional parameter space needs to be explored. Different ideas to reduce the cost of forward modeling WL observables have been put forward and tested. These include using approximate codes to simulate the evolution of the matter-density field, e.g., ICE-COLA (Izard et al. 2018), L-PICOLA (Howlett et al. 2015), or FastPM (Feng et al. 2016);— analytic or semi-analytic models, e.g., CAMELUS (Lin & Kilbinger 2015); and machine learning approaches, e.g., generative adversarial networks (e.g., He et al. 2019; Rodríguez et al. 2018; Mustafa et al. 2019). While analytic models can predict two-point statistics with sufficient accuracy (Barreira et al. 2018), higher-order statistics that capture non-Gaussian information from the non-linear regime require a numerical approach (Sato et al. 2009; Petri et al. 2013, 2017).

In this paper we study, within the framework of ray-tracing $N$-body simulations, the sensitivity of the power spectrum (PS) and several non-Gaussian statistics commonly used in WL studies: the one-point probability density function (PDF), lensing peak counts, and the full set of Minkowski functionals (MFs) to the two hyper-parameters with the highest impact on the computational cost of the simulations: (i) the thickness of the lens planes used to construct the past light cones in ray-tracing and (ii) the mass resolution of the $N$-body simulations used to model the underlying matter-density field. Previous studies have already tackled some of these or related aspects. For instance, Jain et al. (2000) verified the effect on the measured PS of the mass resolution of the underlying DM simulation, the grid size and interpolation method used when ray-tracing and also studied the contribution of super-sample modes to the PS variance. Sato et al. (2009) looked at the effect of the resolution of the 2D lens planes on the measured convergence PS, and evaluated the non-Gaussian contribution to its covariance matrix. Vale & White (2003) investigated the effect of mass resolution and comoving distance between lens planes on the convergence PS, skewness, and kurtosis. This work revisits the sensitivity of the measured convergence PS to the mass resolution and distance between lens planes, and extends the analysis of numerical convergence analyses to non-Gaussian statistics within the framework of a tomographic WL analysis of an LSST-like survey.

The manuscript is organized as follows. In Section 2 we describe our simulation pipeline (Section 2.1), how we quantify the impact of the hyper-parameters (Section 2.2), and the statistics used to assess that impact (Section 2.4). In Section 3 we report and discuss the results of the analysis. We summarize our conclusions in Section 4.
The galaxy densities we consider in the results in the absence of noise. Lower panel: percentage difference between the thin blue lines correspond to the convergence power spectra. The thick black lines correspond to simulations, while the thin blue lines correspond to the haloft calculations. Different dashes indicate different redshift bins for the lensed galaxies. The full lines correspond to results in the presence of shape noise (see Section 2.1 for a description of the shape-noise level considered), and partly transparent lines to the results in the absence of noise. Lower panel: percentage difference between simulations and haloft, for each redshift bin. As in the upper panel, full lines correspond to the noisy case, and partly transparent lines are for the idealized noiseless case.

2. Methods

2.1. Simulating Convergence Maps

Our analysis is based on lensing statistics measured on convergence maps obtained from ray-tracing DM-only N-body simulations. Since WL probes large-scale structures in the nonlinear regime, direct simulations offer a way to characterize the WL signal.

The N-body simulations are run using the publicly available Tree-PM code Gadget2 (Springel 2005), which evolves Gaussian initial conditions generated with NGenIC (Springel 2015). The initial conditions are defined by power spectra computed with CAMB (Lewis et al. 2000). The positions of the particles at different redshifts are used to build past light cones. The trajectory of a bundle of rays is followed along past light cones according to the multi-plane algorithm to generate synthetic convergence maps. We refer the reader to Petri (2016) for a detailed step-by-step description of our pipeline and its implementation in LensTools, and to Jain et al. (2000) for a review of the theoretical basis of the algorithms used.

All the simulations have the same comoving volume of $(240\ h^{-1}\ Mpc)^3$, phases for the initial conditions, and underlying cosmology. The cosmological parameters are consistent with Planck Collaboration et al. (2016): $(H_0,\ Omega_m,\ Omega_{\Lambda},\ Omega_b,\ w_0,\ w_a,\ n_s,\ A_s) = (67.7\ \text{km} \ s^{-1}\ \text{Mpc}^{-1},\ 0.309,\ 0.691,\ 0.0486,\ -1.0,\ 0.816,\ 0.967)$. As an illustration, in Figure 1 we compare the convergence PS measured in our highest-resolution simulation $(1024^3$ DM-only particles, or a particle mass of $1.1 \times 10^{10} M_\odot$), with flat-sky approximation predictions (see Equation (3)) computed using the haloft model (Smith et al. 2003; Takahashi et al. 2012), as implemented in the Core Cosmology Library (Chisari et al. 2019). In data smoothed at $1'$ resolution with the presence of the shape-noise levels considered here, the agreement between theory and simulations stays within 20% for $\ell < 12,000$. In noiseless data, a loss of power in the simulations is noticeable, especially for sources at low redshift. The worst case corresponds to the bin with sources at $z = 0.5$, which shows a 50% loss in power at $\ell \approx 7500$. In all cases we measure excess power at large angular scales relative to the theoretical expectation (up to $\approx 20%$ for the noiseless case and $\approx 13%$ in the presence of shape noise). The reason for this excess is the sample variance in the matter PS, as low multipoles probe scales that are comparable with the size of our simulated volume. Since the purpose of this study is the calibration of hyper-parameters, we are interested in relative changes and the sample variance at large scales should not impact our conclusions. Furthermore, we are mostly interested in relative changes on small scales, where the use of non-Gaussian observables is relevant. To perform inference from data, this sample variance would need to be tackled, for instance, by simulating a larger volume (or many independent N-body realizations of our simulation volume).

Figure 1. Comparison of convergence power spectra measured in our simulations with $1024^3$ DM particles, to those obtained in semi-analytic calculations using the flat-sky approximation and haloft. Upper panel: convergence power spectra. The thick black lines correspond to simulations, while the thin blue lines correspond to the haloft calculations. Different dashes indicate different redshift bins for the lensed galaxies. The full lines correspond to results in the presence of shape noise (see Section 2.1 for a description of the shape-noise level considered), and partly transparent lines to the results in the absence of noise. Lower panel: percentage difference between simulations and haloft, for each redshift bin. As in the upper panel, full lines correspond to the noisy case, and partly transparent lines are for the idealized noiseless case.

For each configuration, a single volume is simulated and reused through random shifts and rotations to generate as many as $\mathcal{O}(10^4)$ pseudo-independent past light cones (Petri et al. 2016). Since each convergence map covers only $3.5 \times 3.5$ deg$^2$ on the sky at this redshift, the flat-sky approximation holds. To account for their intrinsic ellipticity, Gaussian random shape noise is added independently at each of the $1024 \times 1024$ pixels in each convergence map. The standard deviation of this noise,

$$\sigma_{\text{pix}} = \sqrt{\frac{\sigma^2}{2 n_k A_{\text{pix}}}},$$

depends on the variance of the galaxies’ intrinsic ellipticity (assumed to be $\sigma_e = 0.35$), the solid area covered by a pixel ($A_{\text{pix}} = 0.04$ arcmin$^2$), and the effective galaxy number density.

We consider an LSST-like survey with a footprint of 20,000 deg$^2$ and galaxy density distribution $n(z) \propto z^2 \exp[-2z]$ (LSST Science Collaboration et al. 2009). Because a small number of redshift bins suffices to extract most of the tomographic information encoded in weak lensing data sets (Hu 1999), we use five tomographic bins. In each bin, we assume all lensed galaxies are at a fixed redshift of $[0.5, 1.0, 1.5, 2.0, 2.5]$. The galaxy densities we consider in each redshift bin are $[8.83, 13.25, 11.15, 7.36, 4.26]$ arcmin$^{-2}$.

A final smoothing is applied to both noiseless and noisy maps with a Gaussian kernel of $1'$ standard deviation. While only results from maps with shape noise are relevant for the analysis of future survey data, we show also results for smoothed, noiseless convergence maps, since the effect of different simulation choices are often more discernible in those.

2.2. Assessing the Impact of Hyper-parameters

The accuracy of the forward model as a function of different values of the hyper-parameters is assessed by comparing the statistics of observables measured over 10,048 convergence maps simulated for each configuration. As explained in Section 2.1, for each set of hyper-parameters, all 10,048 maps are generated from a single, recycled, N-body simulation.
We chose a "fiducial" configuration as a reference. The difference between an observable’s mean for all cases and the fiducial model is compared with a standard error. As standard error, we consider three standard deviations measured on the fiducial model’s maps, scaled to a survey sky area of $2 \times 10^4$ deg² (commensurate with LSST’s). This scaled standard deviation represents a lower bound on the uncertainty expected in future surveys, as it includes only the statistical error from intrinsic ellipticities. For a review of all sources of error in WL surveys, see Shirasaki & Yoshida (2014).

Ultimately, the most relevant metric is how much the lens-plane thickness affects the inference of cosmological parameters. A definite answer to this question requires the calculation of the credible contours for the parameters of interest. Within the scope of our single-cosmology numerical experiments, we instead look at the reduced $\chi^2$, which combines differences in the mean observables with their covariance matrix.

In general, a given observable is a vector $s$ with components corresponding to bins of spherical harmonic index $\ell$ for the PS, or $\kappa$ thresholds for the other statistics, such as the one-point PDF, lensing peaks, or Minkowski functionals. We compute the $\chi^2$ statistic for each configuration, considering the mean of the fiducial model as ground truth:

$$\chi^2 = (s - s_{\text{fid}})^T \widehat{C}^{-1}(s - s_{\text{fid}}),$$

where $\widehat{C}^{-1}$ is an unbiased estimator of the precision matrix. We compute (and report) results based both on the precision matrix estimated at the fiducial model, and the precision matrix estimated at each specific configuration. We used the prescription from Hartlap et al. (2007) to debias the estimator of the precision matrix. For each observable, we report the $\chi^2$ per degree of freedom. Furthermore, to account for the effect of a change in configuration on the covariance matrix, we report two values of the reduced $\chi^2$, one based on the covariance matrix for the reference configuration and another based on the configuration-specific covariance. We consider two configurations as statistically equivalent when their $\chi^2$ per degree of freedom, in the presence of noise, is less than or equal to unity.

### 2.3. Hyper-parameter Configurations

We analyzed different configurations for two hyper-parameters: the thickness of the lensing planes used in the ray-tracing algorithm, and the mass resolution (number of particles) of the N-body simulations. For each configuration, we ray-traced 10,048 convergence maps, each of these maps with an area of 12.25 deg².

The positions of the DM particles were saved at redshifts that allowed the construction of light cones with lens planes at comoving distances of $20\ h^{-1}$, $40\ h^{-1}$, $60\ h^{-1}$, $80\ h^{-1}$, and $120\ h^{-1}$ Mpc. Thinner lens planes can potentially capture more accurately the evolution of the matter-density field with redshift. However, the number of planes needed to cover a redshift range increases as the plane thickness decreases, and so do the computational and storage requirements for the simulations. In particular, the two tasks that account for the largest increase in computation time are the calculation of the gravitational potential at the planes (solving a 2D Poisson equation in Fourier space) and the computation of the Jacobian matrix that determines the light ray’s deflections at each point on the planes (Petri 2016). The fiducial case corresponds to a lens-plane thickness of $80\ h^{-1}$ Mpc, a value that has been typically used in prior work (Yang et al. 2011; Petri et al. 2013; Zorrilla Matilla et al. 2016), provides nine independent lens planes per simulation snapshot (increasing the number of pseudo-independent $\kappa$ maps that can be generated from a single N-body simulation), and is not large enough to show discreteness effects with lensed galaxies at $z = 2$ (Jain et al. 2000).

The minimum angular scale at which cosmological information can be extracted is limited by the depth of the survey, which determines the number density of galaxies whose shape can be measured, and the accuracy of the forward models used to predict the signal. Baryonic physics (Huang et al. 2019) and intrinsic alignments (Chisari et al. 2015) restrict the accuracy of current models at small scales. Matching the mass resolution of the underlying N-body simulations to the scales at which the analysis of the data is reliable saves computational resources.

In our fiducial N-body simulations we used 1024³ particles, and we ran additional simulations with 128³, 256³, and 512³ particles, which yield mass resolutions per DM particle of $5.7 \times 10^{12}\ h^{-1} M_\odot$, $7.2 \times 10^{11}\ h^{-1} M_\odot$, $9.0 \times 10^{10}\ h^{-1} M_\odot$, and $1.1 \times 10^{10}\ h^{-1} M_\odot$. We note that due to memory limitations, we increased the number of CPUs allocated to our higher-resolution simulation, resulting in a worse scaling than the one that can be achieved in an optimized setup (Springel 2015).

Tables 1 and 2 summarize the computational cost of the main steps involved in our simulation pipeline, and the disk space required for storing the different data products (in practice, not all need to be saved). Performance benchmarks are based on runs using Intel Knights Landing nodes from TACC’s Stampede2 supercomputer at the NSF XSEDE facility.

---

**Table 1: Memory and Computational Time Requirements for the Main Simulation**

| Plane thickness | Task Type | RAM (GB) | CPU Time (s) | Change CPU [%] |
|-----------------|-----------|----------|--------------|----------------|
| $20\ h^{-1}$ Mpc | Lens planes | 96 GB | 82.7 | +1450% |
| $40\ h^{-1}$ Mpc | Lens planes | 21.3 | +300% |
| $60\ h^{-1}$ Mpc | Lens planes | 9.3 | +75% |
| $80\ h^{-1}$ Mpc | Lens planes | 5.3 | | |
| $120\ h^{-1}$ Mpc | Lens planes | 2.7 | -50% |
| $20\ h^{-1}$ Mpc | Ray-tracing | 384 GB | 1621.3 | +407% |
| $40\ h^{-1}$ Mpc | Ray-tracing | 640 | +100% |
| $60\ h^{-1}$ Mpc | Ray-tracing | 426.7 | +33% |
| $80\ h^{-1}$ Mpc | Ray-tracing | 320 | | |
| $120\ h^{-1}$ Mpc | Ray-tracing | 213.3 | -33% |

**Note.** Each CPU time unit is a core hour (representing wall clock time if computed in a series). Changes in CPU time are relative to the fiducial run (in bold).
2.4. Observables

We analyze the effect of changes in the hyper-parameters used to generate the synthetic data on the convergence PS and a series of non-Gaussian observables: the one-point PDF, lensing peak counts, and the full set of MFs.

The PS is the Fourier transform of the two-point correlation function, encodes all the information available in a Gaussian random field (GRF), and can be accurately predicted from theory on large angular scales. Therefore, it is commonly used in WL analyses (The Dark Energy Survey Collaboration 2005; Hikage et al. 2019). We measured it in 20 logarithmic bins between $\ell_{\text{min}} = 200$ and $\ell_{\text{max}} = 12,000$, extending slightly above the smoothing scale of $1'$ which corresponds to $\ell \approx 10,800$. The relatively featureless power spectra can be properly characterized with 20 bins, without yielding a very large data vector when combining the 5 auto-spectra with the 15 cross-spectra for all 5 redshift bins. Since the synthetic maps cover a small field of view, we can use the flat-sky approximation to model the expected PS from measuring the ellipticities of lensed sources at a fixed radial comoving distance, $\chi$:

$$ P_\kappa(\ell) = \left( \frac{3H_0^2}{2c^2} \right)^2 \int_0^{\chi} \frac{d\chi}{a^2(\chi)} \left(1 - \frac{\chi}{\chi_s} \right)^2 P_{\kappa} \left( k = \frac{\ell}{\chi}; \chi \right). $$

(3)

Here $P_\kappa$ is the (nonlinear) matter PS, which depends on the wavenumber $k$ and evolves with redshift, and $H_0$, $\Omega_m$, and $c$ are the Hubble constant, matter-density parameter and speed of light, respectively. Shape noise contributes white noise with a total power of $P_\kappa = \sigma_\kappa^2/n_p$. Smoothing with a Gaussian kernel of width $\sigma$ multiplies the convergence PS by the beam function $b_\ell = \exp[-\ell(\ell + 1)/2\sigma^2]$.

The effect that the finite thickness of the lens planes has on the measured convergence PS can be modeled as convolving the matter-density PS with a window function. Following (Takahashi et al. 2017), it suffices to substitute $P_\kappa$ in Equation (3) by

$$ P_\kappa(k; \chi) = \int_0^{k_{\text{max}}} dk \frac{\Delta \chi}{\pi} P_\kappa(k) \text{sinc}^2 \left( \frac{k_\Delta \Delta \chi}{2} \right), $$

(4)

where $k = \sqrt{k_0^2 + k_\perp^2}$, $k_\perp = \ell/\chi$, $\text{sinc}(\chi) = \sin \chi/\chi$, and $\Delta \chi$ is the thickness of the lens planes used to build past light cones in our simulations.

Changes in the mass resolution of the $N$-body simulations will also affect the measured $\kappa$ PS. On one hand, the matter power will be suppressed on scales below the simulation resolution. On the other hand, an additional shot-noise component will result from the discreteness of the simulated matter-density field. Both effects can be accounted for as in Vale & White (2003), by smoothing the matter PS with a Gaussian kernel of with $\sigma_N$ and adding a Poisson noise contribution corresponding to the finite number density $n_p$ of simulated particles,

$$ P_\kappa''(k; \chi) = P_\kappa(k; \chi) \exp \left(-\frac{\sigma_N^2 k^2}{2} + \frac{1}{n_p} \right). $$

(5)

Most scales in our simulated maps probe nonlinear structures, and as a result there is non-Gaussian information not captured by the convergence PS. To assess the impact of the different simulation parameters on the non-Gaussian information content of convergence maps, we analyze the effect on the one-point PDF, the distribution of lensing peaks, and the Minkowski functionals.

The convergence PDF has been shown to help break the PS degeneracy in the $\Omega_{\text{m}}-\sigma_8$ plane, improving cosmological constraints from the PS alone (e.g., Wang et al. 2009; Patton et al. 2017). It has also been used in order to assess the ability of WL measurements to infer the neutrino masses (Liu & Madhavacheril 2019). We use the binned PDF measured on our (smoothed) simulated maps within the range $[-3\sigma, 5\sigma]$, where $\sigma_\kappa$ is the standard deviation of the maps corresponding to the highest mass resolution (1024$^3$ particles), and the fiducial, 80 $h^{-1}$ Mpc, lens planes. For each redshift bin in the noiseless case, we use $\sigma_\kappa = [0.007, 0.014, 0.019, 0.023, 0.027]$, and for the maps that include shape noise $\sigma_\kappa = [0.034, 0.031, 0.035, 0.043, 0.055]$. We compare the measurements over simulations with the analytic expectation for GRFs with variance

$$ \sigma_\kappa^2 = \int_0^\infty \frac{\ell d\ell}{2\pi} P_\kappa(\ell). $$

(6)
Lensing peaks are local maxima of the convergence field; their distribution as a function of their height, $\kappa$, was proposed to constrain cosmological parameters (Dietrich & Hartlap 2010; Kratochvil et al. 2010), and has been used extensively and successfully ever since (Liu et al. 2015a, 2015b; Kacprzak et al. 2016; Martinet et al. 2018). We measured peak histograms in 50 linear bins covering the range $k_{-0.5, 5.0}$ in units of the root mean square (rms) shape noise in each redshift bin. For the maps smoothed on 1’ scales, the rms of the shape noise for the redshift bins $0.5, 1.0, 1.5, 2.0, 2.5$ is $\sigma_n = [0.033, 0.027, 0.030, 0.036, 0.048]$. As we did for the PDF, we compare measurements with the predictions for GRFs with the same convergence PS as the simulated $\kappa$ maps. The analytic expressions for the number of peaks in GRFs were derived in Bond & Efstathiou (1987). We reproduce here the relevant equations for convenience. The number of lensing peaks in a GRF for a given height $\nu = \kappa/\sigma_0$ is given by

$$n_{\text{peaks}}(\nu) = \frac{1}{(2\pi)^{3/2}\theta^2} \exp \left(-\frac{\nu^2}{2}\right) G(\gamma, \gamma\nu),$$

where $\theta = \sqrt{2}\sigma_1/\sigma_2$ and $\gamma = \sigma_1^2/(\sigma_0\sigma_2)$ depend on the moments of the convergence PS

$$\sigma_p^2 \int_0^\infty \frac{\ell d\ell}{2\ell} \ell^2 P_{\ell}(\ell)$$

and

$$G(\gamma, x) = \left(x^2 - \gamma^2\right) \left[1 - \frac{1}{2} \text{erfc}\left(\frac{x}{\sqrt{2(1 - \gamma^2)}}\right)\right]$$

$$+ x(1 - \gamma^2) \frac{\exp\left(-\frac{x^2}{2(1 - \gamma^2)}\right)}{\sqrt{2\pi(1 - \gamma^2)}}$$

$$+ \frac{\exp\left(-\frac{x^2}{3 - 2\gamma^2}\right)}{\sqrt{3 - 2\gamma^2}} \left[1 - \frac{1}{2} \text{erfc}\left(\frac{x}{\sqrt{2(1 - \gamma^2)(3 - 2\gamma^2)}}\right)\right].$$

Finally, the three Minkowski functionals that can be defined in a two-dimensional random field, $\{V_0, V_1, V_2\}$, measure the area, length, and genus of the subset of points in the random field whose value exceeds a given threshold (Mecke et al. 1994; Schmalzing et al. 1996). While $V_0$ conveys the same information as the PDF, we include it in our measurements for completeness. They have been extensively used to measure or extract non-Gaussian information from cosmological data sets, including weak lensing (Munshi et al. 2012; Petri et al. 2013; Shirasaki & Yoshida 2014; Ling et al. 2015; Marques et al. 2019).

We measured the three MFs in 28 linear bins covering the range $\kappa \in [-2, 5]$ in units of the rms shape noise for each redshift bin, and compared it with analytical predictions for GRFs. The MFs of two-dimensional GRFs are determined by the moments of their PS, and their expression as a function of $\nu$.

![Figure 2. Effect of lens-plane thickness on convergence power spectra. Each column shows, for a different redshift bin, the percentage difference in the mean auto-power spectrum measured over 10,048 convergence maps for a given thickness of the lens planes, relative to the fiducial value of 80 $h^{-1}$ Mpc. The top row corresponds to noiseless data, and the bottom row corresponds to data in the presence of shape noise (in all cases, maps were smoothed at 1’ resolution). Thick lines represent measurements over simulated maps, while thin lines represent theoretical predictions following Takahashi et al. (2017). For comparison, a standard error is shown in shaded gray, corresponding to three standard deviations of the measurements in the fiducial case, scaled to a $2 \times 10^4$ deg$^2$ survey.](image-url)
Goodness of Fit for Different Lens-plane Thickness Con

| Thickness  | Model-dependent | Noisy | Power-spectrum-reduced $\chi^2$ |
|------------|----------------|-------|-------------------------------|
| 20 $h^{-1}$ Mpc | 152.98 | 98.78 | 8.23 (11.72) 5.59 (7.95) |
| 40 $h^{-1}$ Mpc | 17.33 (16.72) | 15.73 | 0.83 (1.17) 0.74 (1.04) |
| 60 $h^{-1}$ Mpc | 6.96 (4.43) | 6.91 (4.34) | 0.17 (0.23) 0.17 (0.24) |
| 120 $h^{-1}$ Mpc | 9.64 (5.25) | 8.84 (5.09) | 0.16 (0.21) 0.16 (0.22) |

PDF-reduced $\chi^2$

| Thickness  | Model-dependent | Noisy | Peak-histogram-reduced $\chi^2$ |
|------------|----------------|-------|--------------------------------|
| 20 $h^{-1}$ Mpc | 30.67 (40.27) | 18.02 | 10.25 (13.31) 7.50 (9.78) |
| 40 $h^{-1}$ Mpc | 3.09 (3.96) | 2.51 (3.20) | 1.33 (1.66) 1.23 (1.53) |
| 60 $h^{-1}$ Mpc | 0.65 (0.72) | 0.62 (0.69) | 0.41 (0.46) 0.41 (0.44) |
| 120 $h^{-1}$ Mpc | 0.48 (0.52) | 0.51 (0.54) | 0.38 (0.39) 0.39 (0.39) |

V$^2$-reduced $\chi^2$

| Thickness  | Model-dependent | Noisy | V$^2$-reduced $\chi^2$ |
|------------|----------------|-------|-----------------------|
| 20 $h^{-1}$ Mpc | 58.55 (82.06) | 32.13 | 18.16 (25.01) 13.17 (18.27) |
| 40 $h^{-1}$ Mpc | 5.38 (7.39) | 4.37 (6.01) | 2.16 (2.79) 1.95 (2.54) |
| 60 $h^{-1}$ Mpc | 0.99 (1.19) | 0.97 (1.16) | 0.54 (0.58) 0.54 (0.57) |
| 120 $h^{-1}$ Mpc | 0.70 (0.81) | 0.81 (0.94) | 0.39 (0.41) 0.40 (0.41) |

V$^1$-reduced $\chi^2$

| Thickness  | Model-dependent | Noisy | V$^1$-reduced $\chi^2$ |
|------------|----------------|-------|-----------------------|
| 20 $h^{-1}$ Mpc | 56.34 (78.72) | 41.79 | 19.70 (26.83) 13.89 (19.10) |
| 40 $h^{-1}$ Mpc | 6.61 (9.04) | 5.78 (7.93) | 2.25 (2.93) 2.05 (2.68) |
| 60 $h^{-1}$ Mpc | 1.20 (1.47) | 1.16 (1.42) | 0.60 (0.56) 0.62 (0.57) |
| 120 $h^{-1}$ Mpc | 1.87 (2.40) | 2.78 (3.66) | 0.41 (0.38) 0.41 (0.37) |

V$^0$-reduced $\chi^2$

| Thickness  | Model-dependent | Noisy | V$^0$-reduced $\chi^2$ |
|------------|----------------|-------|-----------------------|
| 20 $h^{-1}$ Mpc | 46.94 (64.68) | 33.04 | 18.26 (24.98) 13.32 (18.30) |
| 40 $h^{-1}$ Mpc | 5.26 (7.00) | 4.53 (6.05) | 2.13 (2.79) 1.96 (2.57) |
| 60 $h^{-1}$ Mpc | 1.11 (1.25) | 1.11 (1.23) | 0.61 (0.45) 0.62 (0.44) |
| 120 $h^{-1}$ Mpc | 1.43 (1.61) | 1.67 (1.94) | 0.40 (0.32) 0.40 (0.31) |

Note. Configurations that yield good fits ($\chi^2 \leq 1$), implying that they are indistinguishable from the fiducial case, are highlighted in bold. Power spectrum: values for a range of $\ell \in [200, 12000]$ and in parentheses $\ell \in [200, 3532]$. PDF: values for a range of $\sigma \in [-3.0, 5.0]$ in units of the shape-noise rms, and in parentheses $\sigma \in [-3.0, 3.1]$. Peak counts: values for a range of $\sigma \in [-0.5, 5.0]$ in units of the shape-noise rms, and in parentheses $\sigma \in [-0.5, 4.0]$. MFs: values for a range of $\sigma \in [-2.0, 5.0]$ in units of the shape-noise rms, and in parentheses $\sigma \in [-2.0, 3.0]$.

While a perturbative expansion has been found for MFs for non-Gaussian random fields (Munshi et al. 2012), they have been shown to converge too slowly at the relevant ~arcminute smoothing scales (Petri et al. 2013), so we do not investigate them here.

3. Results and Discussion

We discuss the effect on lensing statistics of the comoving distance between lens planes (plane thickness) in Section 3.1 and mass resolution in Section 3.2.

3.1. Lens Plane Thickness

We show the mean percentage difference in the PS for noiseless and noisy convergence maps for different lens-plane thicknesses (or comoving distance between lens planes) relative to the fiducial case of 80 $h^{-1}$ Mpc in Figure 2. For each plane thickness, the PS is measured and averaged over 10,048 convergence maps. For clarity, we do not display the results for the 10 cross-spectra between redshift bins, but the conclusions are not altered. The gray band in each panel represents a standard error corresponding to three times the standard deviation measured over the fiducial maps, scaled to a $2 \times 10^5$ deg$^2$ LSST-like survey. This does not incorporate the effect of off-diagonal terms in the covariance matrix, which are fully utilized (including auto and cross-spectra) in the $\chi^2$ statistic described in Section 2.2 to assess the significance of the differences between each configuration and the fiducial case.

The effect of the finite width of lens planes, measured on the simulations, matches well the expectation from the model described in Takahashi et al. (2017), in which the matter PS is convolved with the window function corresponding to a spherical lens shell. Since the differences between the fiducial and the 60 $h^{-1}$ Mpc case are in general sub-percent level, they fall in the linear range of the $\gamma$-scale and appear to be larger than those for other cases, but they are not. As the lens planes cut structures, there is an appreciable loss of power for thin planes throughout the multipole range considered in this study. The presence of shape noise renders those differences statistically insignificant for the range of thicknesses 60 $h^{-1}$–120 $h^{-1}$ Mpc (see Table 3).

The effect on the PDF is shown in Figure 3, where the PDF change measured from maps built using different lens-plane thicknesses is plotted (thick lines) together with predictions for Gaussian random fields (thin lines). The GRFs have the same PS as the simulated $\kappa$ maps. The overall impact on the GRFs is straightforward: the loss in power translates into a smaller variance for the random field (see Equation (6)). As a result we would expect, for planes thinner than the fiducial, a suppression in the tails of the PDF compensated by an enhancement near the peak of the PDF (and the opposite for a the case with thicker lens planes). Qualitatively, the measurements in the presence of shape noise follow that trend, which is to be expected since the shape-noise model considered is Gaussian.
In the limit of noiseless data, while there is still some qualitative agreement, the quantitative differences with the GRF model are evident. From a practical perspective, lens planes whose thickness lies in the range $h^{-1}Mpc$ yield convergence maps with statistically indistinguishable PDFs, even in the absence of shape noise (see Section 3.1).

The sensitivity of the distribution of lensing peak counts to the width of the lens planes used during ray-tracing is shown in Figure 4. For noiseless data, lens planes thinner than $h^{-1}Mpc$ overproduce low-significance peaks while underproducing higher-significance ones. Table 3 shows how the results from using lens planes within the range $60–120 h^{-1}Mpc$ are statistically indistinguishable. The addition of shape noise...
reduces the differences further, to the point where using thinner planes, at $40 \, h^{-1} \, $Mpc, could be safe. As seen from the PS, there is not a strong case to move from the fiducial thickness of $80 \, h^{-1} \, $Mpc.

As for the PDF, the changes in the lensing peak counts in $\kappa$ maps are qualitatively similar to those in GRFs with the same PS, but there are quantitative differences. The effect is also similar to the one seen for the PDF, and can be understood
The axis labels indicate the redshift bins used. For example, the slice $z_s = 1.5$ planes generated from our $512^3$ particle simulation with the variance of their distribution intuitively: thicker lens planes provide more power on larger scales that modulates the heights of local maxima, increasing the variance of their distribution (and conversely for thinner planes).

Past studies have found that the use of the Born approximation, in which the convergence is directly estimated by weighting the projected matter density, can induce a significant bias to predictions of non-Gaussian observables such as the skewness and kurtosis (Petri et al. 2017). Since in the limit of very thick planes, the multi-plane ray-tracing algorithm is similar to the Born approximation (except for the lack of redshift evolution in the matter-density field within each plane), any attempt to predict peak histograms beyond the range of thickness explored in this study would need to be validated.

The effect of lens-plane thickness on the MFS is different for each functional (see Figure 5). As expected, shape noise “Gaussianizes” the convergence field and as a result measured differences follow qualitatively the predictions for GRFs. This is not the case in the limit of noiseless data. The $\kappa$ field is less sensitive to changes in the plane thickness than what would be expected for GRFs, especially $V_1$. For $V_2$, the region of $\kappa \approx 0$, where the functional changes sign, yields noisy measurements.

In terms of $\chi^2$ (Table 3), the MFS are more demanding than other non-Gaussian observables. While in the presence of shape noise, lens planes in the range of $60-120 \ h^{-1} \ Mpc$ yield measurements that are statistically equivalent, that is not the case anymore in the limit of noiseless data for $V_1$ and $V_2$.

As explained in Section 2.2, the reduced $\chi^2$ allows us to assess the impact on inference, of the differences in the measured statistics from synthetic maps generated with different hyper-parameters. Changes in the covariance matrix are taken into consideration when computing $\chi^2$. We use either the covariance matrix from maps in the reference configuration, or an alternative covariance from maps recomputed after changing the thickness of the lens planes. In both cases, we use the full covariance to compute $\chi^2$, including, for instance, the 10 cross-power spectra not shown in Figure 2. As an illustration, we display the full covariance matrix for the PS measured on noisy maps in Figure 6.

The effect of lens planes in the covariance is small (for example, less than 20% for the dominant, diagonal elements for the PS). As a result, the conclusions are the same regardless of which covariance is used. Also, Barreira et al. (2018) have analyzed the impact of non-Gaussian contributions to the lensing PS covariance on parameter inference. They find that mean changes in the covariance matrix of $\approx 20\%$ translate to $\leq 5\%$ changes in the parameter uncertainties.

Table 3 collects the results for the observables under study. With the shape-noise levels considered, lens planes whose thickness is in the range of $60-120 \ h^{-1} \ Mpc$ yield statistically indistinguishable measurements for all the statistics we have studied, and are therefore safe to use. For a scheme, like ours, that recycles $N$-body snapshots to build different past light cones, there is a tradeoff between a loss in power induced by the lens plane’s window function and the number of lens planes that can be built from a single $N$-body simulation. The fiducial choice of $80 \ h^{-1} \ Mpc$ seems reasonable, since it allows us to generate as many as a few $\times 10^4$ pseudo-independent convergence maps from a single simulation (Petri et al. 2016). In the limit of noiseless data this is not the case anymore for the PS and MFS. For the PS, the effect of the lens-plane thickness can be incorporated analytically.

These results do not take into consideration the uncertainty in the measured covariance matrices: if they are overestimated, our results could change. Nevertheless, the required errors must be substantial: the true covariance would need to be overestimated by more than 150% (its true elements would be smaller than $0.4 \times$ the elements of the matrix measured on our simulations) for the Minkowski functionals (the most sensitive statistic to this hyper-parameter) measured on maps generated with $80 \ h^{-1} \ Mpc$ lensing planes to be statistically distinguishable from those measured on maps generated with $120 \ h^{-1} \ Mpc$ lensing planes.

### 3.2. Mass Resolution

Figure 7 displays the mean percentage difference in the PS for noiseless and noisy maps for different mass resolutions, compared to the highest-resolution case. The number, of particles in the four configurations we tested are 128$^3$, 256$^3$, 512$^3$, and 1024$^3$. As done in Figure 2 with the sensitivity to the lens-plane thickness, a standard error corresponding to three standard deviations for the fiducial case is shaded for reference in Figure 7.

The main difference between the power spectra is an increase in power on small scales, with a relative loss of power on intermediate scales, as the mass resolution decreases. Qualitatively, this behavior matches what is expected from the matter PS: a loss of power at intermediate scales as the mass resolution decreases (captured by the first term in Equation (5)) followed by a sharp increase at small scales due to shot noise. Similar effects have been described in past studies (Jain et al. 2000; Vale & White 2003). We cannot fit all the data with the simple model in Equation (5), though. For example, a smoothing scale of $\sigma_N = 5\%$ of the average inter-particle separation gives results that are qualitatively in line with the relative power change when going from 256$^3$ to 1024$^3$ particles, but not for the other resolutions tested. An effect not taken into consideration in that simple model is that of the softening length on the relaxation time of the simulations. All
of our simulations kept that length constant, and the softening for the lower resolution runs may have been insufficient. Another effect that can induce a loss in power is the reduction in the linear growth factor due to the discreteness of the initial conditions, as described in Heitmann et al. (2010). The presence of shape noise, which dominates at small scales, mitigates both effects, rendering the measured statistic more forgiving to the resolution of the underlying N-body simulation.

The effect of mass resolution on the one-point PDF is not as straightforward as that for the lens-plane thickness. The GRF model does not work particularly well for noiseless data, as can be seen in Figure 8. Except for the lowest-resolution simulation, shape noise still keeps any differences within the standard error of the simulations themselves. In the absence of noise, 512$^3$ and 1024$^3$ particles seem to yield equivalent results.

Lensing peaks are a more robust statistic to the mass resolution (see Figure 9), and the configurations with 256$^3$ and 512$^3$ particles are statistically indistinguishable from the 1024$^3$ fiducial for both noisy and noiseless data. The model with the lowest mass resolution yields histograms whose differences from the fiducial case clearly exceed the statistical uncertainty. The difference is more important for low-significance peaks, which can be induced by the Poisson shot noise present in the underlying N-body simulation. Shape noise reduces the differences at the low-significance tail, where peaks are noise-dominated. The GRF model for peak counts seems to work better than that for the PDF.

The effect of mass resolution on the measured MFs is displayed in Figure 10. As with the PDF, a GRF model does not fit the impact of changes in mass resolution as well as it did for changes in lens-plane thickness. This may be partly due to the relatively larger change in power, as well as the $\ell$-dependence of that change. In general, measurements performed on simulations are less sensitive than what would be expected from GRFs, and their changes remain within errors for most cases.

Table 4 collects the statistical significance of changes in the mass resolution of the N-body simulations used to build synthetic $\kappa$ maps on the observables under study. Most of the considerations discussed in Section 3.1 about differences in the covariance matrices for the observables apply here as well. For example, the differences for the PS covariance are modest, with differences in the dominant, diagonal terms of less than 3% between the covariances for the 256$^3$ and 1024$^3$ particle simulations. For analyses in the presence of shape-noise levels commensurate with the ones considered in this study, 256$^3$ particles suffice. If the multipole range is limited to $\approx$3500, even lower resolutions can be used for the PS (or alternatively the same 256$^3$ resolution can be used for negligible levels of shape noise). The non-Gaussian statistics that are the most sensitive to resolution effects are the MFs.

As with the results for the thickness of the lensing planes, the figures in Table 4 do not incorporate the possible effect of errors in the covariance matrices used in the calculation of $\chi^2$. In this case, the Minkowski functionals are also the most sensitive statistic. For their measurements on the 256$^3$ particles’ simulation to be statistically different from the measurements on the 1024$^3$ particles’ maps, the true covariance would need to be overestimated by more than 117%.
4. Conclusions

We performed a series of numerical experiments to test the influence of the lens-plane thickness and the mass resolution of ray-traced $N$-body simulations on commonly used WL statistics: the convergence PS, the one-point PDF, lensing peak counts, and Minkowski functionals. While our simulations cannot be used directly to analyze survey data, they can serve to guide design choices in studies of non-Gaussian statistics on
small scales. They set some minimal requirements to avoid significant biases in predictions obtained from the application of the multiple-lens-plane algorithm.

We have found that in the multi-plane ray-tracing algorithm, lens planes in the range $60 - 120 \, h^{-1} \text{ Mpc}$ are safe to use in analyses of WL data with galaxy densities commensurate with

Figure 10. Same as Figure 2 for the three Minkowski functionals (MFs). Each pair of rows shows a different MF. The top row of each pair corresponds to noiseless data, and the bottom row corresponds to data in the presence of shape noise (in all cases, maps were smoothed at 1′ resolution). Thick lines correspond to measurements over simulated data, and thin lines correspond to predictions for Gaussian random fields with the same power spectrum as the simulated maps.
those in LSST-like surveys. Thinner planes will result in a loss of power across a wide range of multipoles, which can in principle be accounted for analytically for the PS but not for non-Gaussian statistics. On the thick-plane end, there are no biases induced in the observables, and the computational time for ray-tracing past light cones can be reduced (e.g., by $\approx 33\%$ when $120 \, h^{-1} \text{ Mpc}$ planes are used instead of $80 \, h^{-1} \text{ Mpc}$). However, such a choice implies a larger number of $N$-body simulations to generate the same number of pseudo-independent realizations of the $\kappa$ field, increasing the overall computation budget.

In order to analyze WL data sets at angular resolutions of $1'$ with LSST levels of shape noise, simulations with mass resolutions of $7.2 \times 10^{11} \, M_\odot$ per DM particle are sufficient (corresponding to $256^3$ particles in a $240 \, h^{-1} \text{ Mpc}$ simulation box), even if non-Gaussian statistics are included in the analysis. Moving, for instance, from $512^3$ to $256^3$ particles, could bring computational time savings of $\approx 77\%$.

We thank the anonymous referee for providing insightful feedback that has resulted in an improved manuscript. We also thank Alexandre Refregier for useful comments on the original manuscript and Takashi Hamana for pointing us to an analytic model for the loss of power due to finite-thickness lens planes. The simulations and data analysis for this work were performed on NSF’s XSEDE facility. This work was also supported by NASA ATP grant 80NSSC18K1093.

**ORCID iDs**

José Manuel Zorrilla Matilla [https://orcid.org/0000-0002-6267-716X](https://orcid.org/0000-0002-6267-716X)

Zoltán Haiman [https://orcid.org/0000-0003-3633-5403](https://orcid.org/0000-0003-3633-5403)

**References**

Barreira, A., Krause, E., & Schmidt, F. 2018, JCAP, 10, 053

Bartelmann, M., & Schneider, P. 2001, PhD, 340, 291

Bond, J. R., & Efstathiou, G. 1987, MNRAS, 226, 655

Chisari, N., Codis, S., Laigle, C., et al. 2015, MNRAS, 454, 2736

Chisari, N. E., Alonso, D., Krause, E., et al. 2019, ApJS, 242, 2

DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, arXiv:1611.00056

Dietrich, J. P., & Hartlap, J. 2010, MNRAS, 402, 1049

Feng, Y., Chu, M.-Y., Seljak, U., & McDonald, P. 2016, MNRAS, 463, 2273

---

### Table 4

Same Goodness of Fit as in Table 3, but for Different Mass-resolution Configurations

| Particle Mass (number or particles) | Model-dependent Noiseless | Fiducial | Model-dependent Noisy | Fiducial |
|-------------------------------------|---------------------------|----------|-----------------------|----------|
| $5.7 \times 10^{12} \, h^{-1} \, M_\odot$ (128$^3$) | 11494.43 (25.03) | 44822.28 (25.69) | 42.44 (0.85) | 44.26 (0.85) |
| $7.2 \times 10^{11} \, h^{-1} \, M_\odot$ (256$^3$) | 48.97 (0.63) | 50.48 (0.62) | 0.17 (0.10) | 0.17 (0.10) |
| $9.0 \times 10^{10} \, h^{-1} \, M_\odot$ (512$^3$) | 1.68 (0.05) | 1.66 (0.04) | 0.03 (0.01) | 0.03 (0.01) |

Note. Configurations with $\chi^2 < 1$, implying that they are indistinguishable from the fiducial case, are highlighted in bold. Power spectrum: values for a range of $\ell \in [200, 12000]$ and in parentheses $\ell \in [200, 3532]$. PDF: values for a range of $k \in [-3.0, 5.0]$ in units of the shape-noise rms, and in parentheses $k \in [-3.0, 3.1]$. Peak counts: values for a range of $\kappa \in [-0.5, 5.0]$ in units of the shape-noise rms, and in parentheses $\kappa \in [-0.5, 4.0]$. MFs: values for a range of $\kappa \in [-2.0, 5.0]$ in units of the shape-noise rms, and in parentheses $\kappa \in [-2.0, 3.0]$. 

---
