The Weak Lensing Peak Statistics in the Mocks by the inverse-Gaussianization Method

Zhao Chen,1 Yu Yu,1 Xiangkun Liu,2 and Zuhui Fan2,3

1Department of Astronomy, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China
2South-Western Institute for Astronomy Research, Yunnan University Chenggong District, Kunming 650500, Yunnan Province, China
3Department of Astronomy, School of Physics, Peking University, Beijing 100871, China

ABSTRACT

We apply the inverse-Gaussianization method proposed in Yu et al. 2016 to fast produce weak lensing convergence maps and investigate the peak statistics, including the peak height counts and peak steepness counts, in these mocks. We find that the distribution of peak height and steepness is in good agreement with the simulation. The difference is ≲ 20% for these peak statistics in the maps at source redshift \( z_s = 1 \). Besides, the loss of off-diagonal elements in peak covariance motivates us to consider the super sample variance in weak lensing peak statistics. We propose correction methods to effectively recover the (anti-)correlation among different bins by adding scatters in the mean value of these mocks. Finally, as an example of the application, we adopt the improved inverse-Gaussianization method with the above improvement to fast generate 40,000 mocks to calculate precision matrices between the power spectrum and peak statistics.

Keywords: weak lensing, cosmology, peak statistics

1. INTRODUCTION

The gravity induced deflection of the light from distant background galaxies provides an effective cosmological probe which is referred to weak gravitational lensing. The lensing effect is induced by the total matter along the line-of-sight, no matter it is dark or luminous. This makes weak gravitational lensing a promising tool to probe the matter distribution and even the geometry of the universe (e.g. Mellier 1999; Bartelmann & Schneider 2001; Munshi et al. 2008; Kilbinger 2015). Various weak lensing observations have improved our knowledge on the universe. (e.g. CFHTLS: Semboloni et al. 2006; Fu et al. 2008; CFHTLenS: Van Waerbeke et al. 2013; Hildebrandt 2014; KiDS-450: Köhlinger et al. 2017; Hildebrandt et al. 2017; DES Y1 results: Troxel et al. 2018; Abbott et al. 2018). Ongoing and upcoming projects, such as DES (Abbott et al. 2018), KiDS (De Jong et al. 2015), HSC (Aihara et al. 2018), LSST (LSST Science Collaboration et al. 2009), Euclid (Laureijs et al. 2011), will obtain a huge amount of weak lensing observational data with unprecedented precision. To match this statistical power, the control of various weak lensing systematics in the analysis is of great importance.

The weak lensing field is a non-Gaussian field due to the nonlinear evolution of the late Universe (e.g. Scoccimarro et al. 1999; Lee & Pen 2008; Semboloni et al. 2010; Takada & Jain 2010; Joachimi et al. 2011). The second order statistics, i.e. the correlation function and the power spectrum, only capture a part of the cosmological information. Thanks to the improved statistical power, extracting the information in high order statistics helps to break the degeneracy in the cosmological parameters (e.g. Fu et al. 2014; Petri et al. 2015; Gatti et al. 2019). The weak lensing bispectrum and peak statistics are widely used in the analysis beyond the second order (e.g. LSST Science Collaboration et al. 2009; Fan et al. 2010; Yang et al. 2011; Liu & Haiman 2016; Liu et al. 2016; Kacprzak et al. 2016; Shan et al. 2017; Giocoli et al. 2018; Hall & Mead 2018; Martinet et al. 2018; Coulton et al. 2019; Munshi et al. 2019).

In weak lensing peak analysis, a peak is defined as a local maximum whose value is greater than the eight neighbors in the simulated/observed convergence field. Usually smoothing is adopted to suppress the noise in observation and simulation. High peaks are mainly dominated by the lensing contribution from a single massive halo along the line-of-sight, while the...
low peaks are dominated by the contributions of multiple haloes close to each other in angular position (e.g. Yuan et al. 2018; Wei et al. 2018). Separating the peaks into high and low type helps in peak modeling due to the different origins of these peaks. Both of them contain cosmological information (e.g. Hamana et al. 2004; Dietrich & Hartlap 2010; Kratochvil et al. 2010; Lin & Kilbinger 2015a; Liu & Haiman 2016). The number of peaks with different peak values is an important statistic called peak (height) counts which is sensitive to the non-Gaussianity of the lensing field. Yang et al. 2011 found that the joint analysis of the power spectrum and peak height distribution is approximately twice sensitive in the cosmological parameter $\Omega_m$ and $\sigma_8$ compared to using power spectrum alone. Liu et al. 2015a used the CFHTLenS observations to constrain cosmology in a three-dimensional parameter space, $\Omega_m$, $\sigma_8$ and the dark energy equation-of-state $\omega$, and the constraint from peak height counts is comparable to those from the second order statistics alone. Shan et al. 2017 indicated the peak height counts have the potential to break the degeneracy between $\Omega_m$ and $\sigma_8$ from the recent KiDS-450 observations (e.g. Hildebrandt et al. 2017; Martinet et al. 2018). Counting the weak lensing peaks is one of the most promising statistics to extract information beyond Gaussian case.

Motivated by the recent deep learning study on weak lensing, another peak statistic, the peak steepness counting, was inferred. Ribli et al. 2018 found that the constraining power from the peak steepness counts is much greater than the peak height counts in both noiseless and noisy case. Now the peak steepness is of great interest by the lensing community.

To obtain the precise cosmological parameter constraints, an accurate estimation of the covariance matrix is crucial. The number of realizations to ensure 1% error in the power spectrum covariance is estimated to be $\sim O(10^4)$ (e.g. Dodelson & Schneider 2013; Taylor et al. 2013; Percival et al. 2014). Constructing such a large amount of realizations is a challenge especially for weak lensing cosmology, which requires both the large volume coverage and the small scale precision. Fast simulations are widely adopted in galaxy clustering measurement (e.g. Chuang et al. 2015; Agrawal et al. 2017; Lippich et al. 2019; Blot et al. 2019; Colavincenzo et al. 2019). Some of them are further developed for weak lensing (e.g. Izard et al. 2018). Recently, Peel et al. 2016 studied the application of a fast model, CAMELUS, for peak height counts prediction in the Euclid-like survey. Machine Learning technique was also adopted to fast produce the weak lensing maps (e.g. Mustafa et al. 2019). Based on the fact that the weak lensing field can to effectively Gaussianized by a local transform (Yu et al. 2011), Yu et al. 2016 proposed the inverse-Gaussianization method to fast generate lensing mocks. This method successfully produces the lensing power spectrum and a reasonable good power spectrum covariance. The bispectrum amplitude is also recovered although the detailed dependence on the wavenumber configuration is missing.

In this work, we extend the analysis in Yu et al. 2016. We adopt the inverse-Gaussianization method to fast generate weak lensing convergence maps and quantify the statistics of the convergence peaks including peak height and peak steepness counts. This paper is organized as follows. In Section 2, we review some basic knowledge about weak gravitational lensing and the inverse-Gaussianization method. Then, weak lensing convergence mocks are generated by the fast generating method and the peak counts results are presented in Section 3. We briefly introduce an application of the mocks produced by the inverse-Gaussianization method in Section 4. We present the conclusion and discussion in Section 5.

Note that Shirasaki 2017 adopted a similar local transform method to produce weak lensing maps and concluded that the local-Gaussianized model cannot explain the simulated peak counts in the range of $\kappa/\sigma_\kappa < 3$. The difference is about 10% and increases towards low peaks. We compare our results to Shirasaki 2017 in Section 3.

2. THE INVERSE-GAUSSIANIZATION METHOD

2.1. Weak Lensing Basics

Weak lensing is one of the important cosmic probes at the late-time Universe. From the distortion of the background galaxy images, we can probe the lensing effect from the total matter along the line-of-sight. The lensing effect is described by the Jacobi matrix

$$A = \begin{pmatrix} 1 - \kappa - \gamma_1 & \gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}.$$ (1)

Here, $\kappa$ is the lensing convergence field which describes the amplification of the observed object. While the complex shear, $\gamma = \gamma_1 + i\gamma_2$, gives the shape change of the image. They are not independent observables and can be converted to each other in the ideal case. In this work, we focus on the lensing convergence, which is a weighted projection of the matter density along the line-of-sight,

$$\kappa(\hat{n}, z_s) = \int_0^{\chi_s} W(z, z_s) \delta_m(\hat{n}, z) d\tilde{\chi}.$$ (2)

Here, $\delta_m$ is the matter density contrast. $\chi(z)$ is the comoving angular diameter distance to the lens redshift $z$. 
We conventionally express $\chi$ in units of the Hubble radius, $\tilde{\chi} \equiv \chi/(c/H_0)$, in which $H_0$ is the Hubble constant today. The lensing kernel for a source at redshift $z_s$ and a lens at redshift $z$ is given by

$$W(z, z_s) = \frac{3}{2} \Omega_m (1 + z) \tilde{\chi}(z) \left[ 1 - \frac{\chi(z)}{\chi(z_s)} \right].$$

(3)

In observation, a common way to gain the convergence field is to reconstruct it from the observed shear field (e.g. Schneider et al. 2017). Because of the intrinsic shape noise, the reconstructed convergence is noisy, and usually we smooth the convergence field by using a Gaussian smoothing filter (e.g. Hamana et al. 2004).

$$K_{\text{smoothed}}(\ell) = \int d^2\theta' \kappa(\theta - \theta') U(\theta'),$$

(4)

where $U(\theta)$ is the filter function:

$$U(\theta) = \frac{1}{\pi \theta_G^2} \exp\left(-\frac{\theta^2}{\theta_G^2}\right).$$

(5)

Here, we choose the smoothing scale $\theta_G = 1.2\text{arcmin}$, which is a suitable scale for the source at $z_s = 1.0$.

The two point correlation function and its Fourier pair power spectrum are the basic statistical tools in cosmology. The power spectrum is defined as the correlation in the Fourier space:

$$\langle \kappa(\vec{\ell}_1) \kappa(\vec{\ell}_2) \rangle = (2\pi)^2 \delta_D(\vec{\ell}_1 + \vec{\ell}_2) C_\kappa(\ell_1).$$

(6)

Here, Dirac function $\delta_D$ represents the homogeneity in statistics. The isotropy is implied by the fact that the power spectrum $C_\kappa(\ell)$ only depends on the mode $\ell = |\vec{\ell}|$. By using the Limber approximation and Eq.2, the convergence power spectrum is related to the matter power spectrum:

$$C_\kappa(\ell) = \int_0^{\chi_s} d\chi \frac{W(\chi, \chi_s)}{\chi^2} P_\delta \left( k = \frac{\ell}{\chi} ; \chi \right),$$

(7)

and contains important cosmological information.

The weak lensing two point statistics have the most constrain power for the cosmological parameter combination $S_8 = \sigma_8 \Omega_m^{1/2}$ with some power index $\alpha$ depending on the observation configuration and fiducial cosmology. To break the degeneracy, high-order statistics are helpful. The weak lensing skewness and kurtosis have been already measured and studied for more than ten years (e.g. Zhang et al. 2003; Jarvis et al. 2004; Pires et al. 2009, 2012; Gatti et al. 2019). Analog to the power spectrum, the bispectrum $B_\kappa$ is defined by the three point correlation function of convergence:

$$\langle \kappa(\vec{\ell}_1) \kappa(\vec{\ell}_2) \kappa(\vec{\ell}_3) \rangle = (2\pi)^2 \delta_D(\vec{\ell}_1 + \vec{\ell}_2 + \vec{\ell}_3) B_\kappa(\ell_1, \ell_2, \ell_3).$$

(8)

The joint analysis of the two point statistics and three point statistics can tighten the constrain on $S_8 = \sigma_8 (\Omega_m/0.27)^{\alpha}$ (Fu et al. 2014). Other than the high-order moments and the correlation hierarchy, the weak lensing peak is considered as one of the most powerful statistics beyond the 2nd order (e.g. Yang et al. 2011; Lin & Kilbinger 2015b; Liu & Haiman 2016). In this paper, we focus on the peak statistics including the well studied peak height counts and the new developed peak steepness counts.

Hereafter, we only use the smoothed lensing convergence maps. Thus, we still use the notation $κ$ to denote the smoothed lensing convergence field.

2.2. Fast Mock Generation

To capture more information from the non-Gaussian weak lensing field, several Gaussianization methods were proposed (e.g. Yu et al. 2011; Joachimi et al. 2011; Seo et al. 2012; Yu et al. 2012). Based on the fact that the weak lensing field could be effectively Gaussianized (Yu et al. 2011), the inverse-Gaussianization method was proposed to fast generate the lensing convergence maps (Yu et al. 2016). This method includes the following steps.

i Obtain the local transform function. Given a non-Gaussian lensing convergence map, we can find a monotonic local transform to map the $κ$ field into a new $y$ field. We require that the $y$ field has a Gaussian one-point PDF by design and the local transform function is obtained by

$$y = \text{erf}^{-1}(2\text{CDF}(κ) - 1).$$

(9)

Here, $\text{CDF}(κ)$ is the cumulative distribution function of the lensing convergence field. The Gaussian error function $\text{erf}(x)$ is given by the following expression:

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-\eta^2} d\eta.$$  

(10)

The resulting $y$ field has the mean of zero and $\sigma_y = \sqrt{2}/2$. For multiple realizations, we obtain the mean local transform first and use it as the Gaussian transform function to obtain the $y$ fields.

ii Obtain the realizations of $y$ fields.

From the Gaussianized $y$ maps, we can obtain the averaged power spectrum among the realizations. New Gaussian $y$ realizations are generated according to this measured power spectrum.

iii Inverse transform. Because we have obtained the local transform function ($y - κ$ relation) in Step
i, it’s straightforward to obtain the realizations of the weak lensing convergence maps by the inverse transform the Gaussian $y$ map. These independent realizations are what we need.

This method relies on the assumption that the weak lensing field could be effectively Gaussianized. However, the residual non-Gaussianity is observed in the Gaussianized $y$ fields. This residual non-Gaussianity definitely has some impact on some of the statistics. We mainly focus on the peak statistics in this work.

2.3. Simulation Setups

The weak lensing convergence maps we use are from the simulation suite described in Liu et al. 2015b. The dark matter only simulations were run by Gadget-2 code (Springel 2005), in the flat $\Lambda$CDM framework, with $\Omega_m = 0.28$, $\Omega_{\Lambda} = 0.72$, $\Omega_b = 0.046$, $\sigma_8 = 0.82$, $n_s = 0.96$ and $h = 0.7$. The simulation suite contains 8 independent realizations with a box size of 320 $h^{-1}$Mpc and 4 independent realizations with a box size of 600 $h^{-1}$Mpc. For both cases, the particle number is $640^3$. The lensing convergence maps at $z = 1$ are constructed by the ray-tracing technique on the small box realizations. Consider the different Cartesian directions, there are 24 sets in total and every set can be used to construct 4 maps. Totally, 96 lensing convergence maps at $z = 1$ are constructed with map area of $3.5 \times 3.5$ deg$^2$, pixelized into 1024 $\times$ 1024 grids. We refer the readers to Liu et al. 2015b for the details of the lensing simulation.

3. RESULTS

We quantify the performance of inverse-Gaussianization method against simulations through a series of tests. Hereafter, we call the maps generated by the inverse-Gaussianization as lensing convergence mocks.

3.1. Transform function

The local transform functions ($\kappa$–$y$ relations) obtained from the 96 lensing convergence maps at $z = 1$ are shown in grey lines in Fig. 1. The black line represents the average transform function. We can see the divergence at the high end where the function is dominated by the rare extreme $\kappa$ values. This phenomenon is caused by the sample variance.

3.2. Power spectrum

First, we investigate the power spectrum of the lensing convergence mocks produced by the inverse-Gaussianization method. The power spectrum from the simulated lensing maps is presented in blue line in the right panel of Fig. 2. The error bars are the r.m.s. among the 96 realizations. Note that we applied Gaussian smoothing with $\theta_G = 1.2$ arcmin on these lensing maps. The power at $\ell > 3000$ is suppressed. For the $y$ maps, we present the power spectrum in blue line in the left panel of Fig. 2. The power spectrum amplitude for these Gaussianized fields is determined by $\sigma_y$ in the Gaussianization process. However, the choice of $\sigma_y$ value has no impact on the mock power spectrum amplitude since the scaling is cancelled in the inverse-transform process. We produce the Gaussian random fields with the measured power spectrum $C_y(\ell)$ for the same number of realizations. The power is shown in red line in the left panel of Fig. 2, which is consistent with the input power. The power for the inverse-transformed fields is plotted in red line in the right panel. We find the power spectrum is successfully reproduced by the inverse-Gaussianization method. The slight difference at the largest scale comes from the cosmic variance.

In Yu et al. 2016, the performance of the inverse-Gaussianization method is investigated on the projected density field over the thickness of 300 $h$Mpc$^{-1}$ at several redshifts. The maps used in this work is constructed from the ray-tracing method, which includes all the contribution from the nonlinear density field from $z = 0$ to $z = 1$ weighted by the lensing kernel $W(z, z_s = 1)$. The result here is consistent with the finding in Yu et al. 2016.

3.3. Peak Height

Peak height counts of the weak lensing convergence map is considered as a promising tool to probe nonlinear structure evolution at late times, which can provide additional cosmological information beyond power spec-
The missing of the extremely high peaks and the overproduction of high peaks of the lensing field. The inverse-Gaussianization method cannot capture the map-to-map variance for these extremely high peaks and as a consequence they accumulate as the overproduction of high peaks.

Figure 2. The blue line in the left plot represents the power spectrum \( C_y(\ell) \) of Gaussianized \( y \) fields which are produced at Step i in Sec.2.2. The error bars are the r.m.s. among realizations. Convolving this measured \( C_y(\ell) \) with white Gaussian Random Fields, we can obtain the Gaussian realizations whose power spectrum is shown as a red line in the left panel. In the right panel we compare the power spectrum of simulated maps (blue line) with the \( \kappa \) mocks produced by locally transforming the Gaussian Random Fields back (red line). They are consistent with each other except at the very small scale.

Figure 3. The peak height counts of the mocks and simulation realizations are shown in red and blue line in the top panel, respectively. The error bars are generated by the bootstrap method. The ratio of the two is shown in the bottom. The peaks are reproduced well for \( \kappa_p \in [-0.015, 0.065] \). The mocks from the inverse-Gaussianization method contain less peaks of extreme low value (\( \kappa_p < -0.03 \)), more peaks with \( \kappa_p \sim 0.1 \), and miss the extremely high peaks (\( \kappa_p > 0.13 \)). The missing of the extremely high peaks and the overproduction of the high peaks can be understood. Adopting the averaged transform function \( \kappa-y \) can not capture the map-to-map variance for these extremely high peaks and as a consequence they accumulate as the overproduction of high peaks.

We compare the peak height counts \( N(\kappa_p) \) with peak value \( \kappa_p \) between the mocks and simulations in Fig. 3. We divide the peaks into linear \( \kappa_p \)-bins covering \(-0.03 < \kappa_p < 0.20 \). In the upper panel, the blue line represents the peak counts of mocks while the simulation result is drawn by the red line. The error bars come from the non-parametric bootstrap method. We plot the ratio of the mock peak counts to the simulated one in the bottom panel. The error bars come from the fluctuations in the mocks. We find that the peak height counts are produced reasonably good by the inverse-Gaussianization method. For \(-0.015 < \kappa_p < 0.065 \) \((-1.23 < \nu \equiv \kappa_p/\sigma_{\kappa} < 5.37 \)), the difference is within 20%. The peaks with value \( \kappa_p > 0.065 \) are overproduced.

Another obvious feature is that the inverse-Gaussianization method cannot produce the extremely high peaks \((\kappa_p > 0.13, \nu > 11)\). Two facts could induce this difference. One is that the Gaussianization process is not perfect. In Yu et al. 2016, residual non-Gaussianity is observed in the power spectrum covariance for the Gaussianized fields. The residual locates at small scales and implies that the peaks in the Gaussianized field are different from the one in a true Gaussian random field. The inverse-Gaussianization method cannot capture these peaks, leading to the missing high end in the peak height counts. The other fact is that we use the averaged transform function to perform the inverse-transform, partially neglecting the variance among the realizations. This also fails to convert some high peaks in \( y \) mocks into extremely high peaks in \( \kappa \) mocks. Those underestimated peaks are accumulated around \( \kappa_p \approx 0.1 \) and this partially causes the overproduction of the peaks at this range.

This result is slightly different from the result in Shirasaki 2017. We list several differences in the process. Shirasaki 2017 adopts the convergence maps reconstructed from the ray-tracing shear fields, while our convergence maps are directly from the ray-tracing output. Shirasaki 2017 adopts the truncated Gaussian filter to account for an undetermined constant in the convergence construction, while we use a simple Gaussian filter for theoretical motivation. The power spectrum of the
Gaussianized $y$ field in Shirasaki 2017 is theoretically obtained by its relations with the local transform function and the power spectrum of the $\kappa$ field. To numerically solve for $C_y(\ell)$, a continuous $C_\kappa(\ell)$ is required and obtained by modifying the theoretical convergence power spectrum from the halo model with two tuning factors accounting for the ray-tracing resolution effects. In our method, we directly measure $y$ power spectrum from the Gaussianized fields and use this measured $C_y(\ell)$ to fast produce new $y$ maps. In principle, all the numerical and resolution effects are automatically included in our process. Our method tends to slightly overproduce the number of peak for $0 < \nu < 5$, while Shirasaki 2017 produces less peaks in this range. For the absolute deviation from the simulated one, our method has a slightly better performance in the range $-1 < \nu < 2$. Both methods fail in the number of the extremely low peaks.

The normalized covariance matrix of peak height counts from simulation is shown in the left panel. We can find obvious structures which indicate the strong (anti-)correlation among different peak height bins. The right matrix is produced by inverse-Gaussianization method and is approximately diagonal. This result implies the absence of cosmological information caused by the super sample variance.

3.4. Peak Height Covariance

From the 96 realizations for both the simulated and mock lensing convergence, we estimate the covariance of the peak height counts, $C_{ij} = \langle \Delta N(\kappa_{p,i})\Delta N(\kappa_{p,j}) \rangle$. Here, $i,j$ is the index for the peak value bin and $\Delta N(\kappa_{p,i}) = N(\kappa_{p,i}) - \bar{N}(\kappa_{p,i})$. We choose 20 $\kappa_p$-bins linearly covering the range $-0.02 < \kappa_p < 0.1$. The normalized peak height counts covariance

$$\hat{C}_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}} \quad (11)$$

is shown in Fig. 4. It presents the correlation among counts at different convergence peak values, which is important for constraining cosmological parameters using the peak height counts. The left panel is the peak covariance matrix of the simulated maps. We observed strong off-diagonal features in this covariance. The number of low peaks is strongly correlated with themselves, and the same happens for the high peaks. The low peaks and high peaks have strong negative correlations.

However, the covariance of the mocks, which is shown in the right plot, is almost diagonal. This sharp contrast implies that there must be some important ingredient missing in the generation of the mocks. Another clue is that this missing ingredient has almost no impact on the power spectrum.

3.5. Super Sample Variance

![Figure 5](image.png)

**Figure 5.** The distribution of the mean for convergence fields are shown in the left panels while the right panels represent the distribution of mean for $y$ fields. The original simulations and Gaussianized $y$ fields have broad scatter in the mean of fields. However, the Gaussian Random Fields (newly generated $y$ fields, detailed in Section 2.2, Step ii) has exact zero mean. It leads to the very narrow distribution of mean for these mocks, which differs from the simulation.

By construction, the $y$ maps generated with the given power spectrum have the mean of zero. This is shown in the bottom right panel of Fig. 5. The PDF of the $\bar{y}$ is a delta function at zero. While the Gaussianized maps have a broad scatter of the mean, which is shown in the upper right panel, with $\sigma(\bar{y}) = 0.1327$. After the inverse-transform, we find that the mock $\bar{\kappa}$ is not exactly zero (bottom left panel), but the distribution is still very narrow, which is very different from the simulated maps (upper left panel).

The lack of the variance in the mean value comes from the lack of super sample variance, i.e. the long wave perturbation beyond the map size we investigated. The mock convergence maps have no power beyond the map size by design. In simulation, the light-cone is cut off from the stacked boxes. Although the simulated maps also miss some super sample variance beyond the sim-
We describe the power spectrum only shows negligible change and we do not present the result here. The normalized peak height counts covariance is presented in the top right panel of Fig. 6.

### iii Add simulated $\bar{\kappa}$.
We add the mean values taken from the simulated convergence maps to our mock $\kappa$ maps. This process does not assume the distribution of the mean is Gaussian, but the mocks have the exact $\bar{\kappa}$ distribution with the simulated ones. The result is presented in the bottom left panel of Fig. 6.

### iv Add simulated $\bar{y}$.
We add the $\bar{y}$'s taken from the Gaussianized convergence maps to our mock $y$ maps. Same as the previous one, the $\bar{y}$ distribution is exactly the same with the Gaussianized fields. The result is shown in the bottom right panel of Fig. 6.

All of these approaches can recover the features in the simulated peak height counts covariance to some extent. Careful readers can find that the strong correlations between the large $\kappa_p$ bins, between the small $\kappa_p$ bins, and the anti-correlation between large and small $\kappa_p$ bins are not perfectly reproduced. Compared to the simulated case, the (anti-)correlation is weaker for method i and iii. Method ii and iv produces stronger (anti-)correlation. The method ii produces the most similar result with the simulated one. These results tell us that the features in the simulated one are mainly from the super sample variance, and this effect can be reproduced by adding reasonable scatters in the mean value of these maps.

### 3.6. Peak Steepness
Ribli et al. 2018 proposed an improved lensing observable motivated by deep learning, which is called peak steepness. The peak steepness counts perform better than the peak height counts on constraining $\sigma_8$ and $\Omega_m$. So we also test the peak steepness counts and its covariance on the mocks.

Peak steepness means the difference of peaks and their neighbors. Naturally, the gradient at the peak position must be zero. As in Ribli et al. 2018, we describe the steepness of peaks by the magnitudes of gradients near the peaks, which can be calculated by the convolution with the isotropic discrete Laplace operator. The operator we use has the following form:

$$L = 4 \begin{bmatrix} 0 & -0.25 & 0 \\ -0.25 & 1 & -0.25 \\ 0 & -0.25 & 0 \end{bmatrix}.$$ (12)

---

**Figure 6.** The correlation matrices improved by method i - iv (detailed in Section 3.5) are presented in the top left, top right, bottom left and bottom right panel, respectively. The top two cases assume that the distribution of the mean $\kappa/y$ is Gaussian, and the bottom two methods directly add simulated mean value into corresponding mocks. We can find that adding mean value on $y$ maps pre to inverse-transform recover stronger features for the off-diagonal matrix elements than directly modifying $\kappa$ maps. Comparing with Fig. 4, the method ii produces the most similar pattern to the simulated peak height counts covariance.
Figure 7. Similar to Fig. 3, the distributions of peak steepness are shown in this plot. The blue line represents steepness counts of original simulated $\kappa$ maps, while the result of mocks by the inverse-Gaussianization method is shown by the red line. We find that the steepness counts are well produced for almost all the range we consider.

We use the notation $s_p$ to denote the peak steepness, and by design $s_p > 0$.

Similar to Fig. 3, we present the peak steepness counts $N(s_p)$ in Fig. 7. The blue line is the result of simulated lensing convergence maps and the red one is from the inverse-Gaussianization method. We find that the difference is within 20% for almost all the steepness range we investigated. An overall trend is that the number of low steepness peaks ($s_p \lesssim 1.7 \times 10^{-3}$) is smaller by 0–20% than the simulation and the high steepness peaks ($2 \times 10^{-3} < s_p < 4 \times 10^{-3}$) are overproduced by similar amounts. Notice that the steepness is sensitive to the pixel size and smoothing scale. To our knowledge, the dependence of the steepness measure on these factors has not been investigated in detail and it is out of the scope of this work. We expect that increasing the smoothing scale and pixel size will affect the shape of the local transform function, but will not affect the performance of the inverse-Gaussianization much. Our study may give some clues for modeling of the steepness statistic.

3.7. Peak Steepness Covariance

Similar to peak height counts, we also study the covariance of peak steepness counts which is defined as $C_{ij} = \langle \Delta N(s_{p,i}) \Delta N(s_{p,j}) \rangle$. We divide the peaks into $20$ $s_p$-bins linearly covering $3 \times 10^{-4} < s_p < 5 \times 10^{-3}$.

The weak correlation at different steepness in the simulated maps is shown in the left panel of Fig. 8. The right panel is the normalized steepness covariance of 96 mocks by the inverse-Gaussianization method. It is approximately a diagonal matrix. This situation is similar to the peak height counts covariance. The missing feature in the covariance is caused by the lack of super sample variance.

However, different from the peak height, peak steepness is immune to a constant shift by definition. Thus, we apply the method ii and iv described in Section 3.5 to recover the features in the steepness covariance. We find that both methods recover the features but overproduce the correlation and anti-correlation among $s_p$-bins.

Figure 8. The normalized covariance for peak steepness counts of the simulated and mock maps is presented in the left and right panel, respectively. The result from simulations has some weak off-diagonal elements while the matrix of mocks is diagonal. Both of matrix have the same steepness value range at $[3 \times 10^{-4}, 5 \times 10^{-3}]$. This result indicates the absence of super sample variance.

Figure 9. We mimic the super sample effect by two of the four methods detailed in Sec. 3.5. Identical to Fig. 8, the result of method ii is shown in the left panel, while the right normalized covariance matrix is calculated by the last method in Sec. 3.5. We find that both methods produce the matrices with stronger (anti-)correlation. This result is consistent with the result of peak height counts covariance.

4. APPLICATION

Joint analysis of the two-point statistics and the high-order one such as peak height counts can break the degeneracy in the cosmological parameter constraints. However, they are not independent observables. The precision of the covariance between these observables
is essential in this joint analysis to avoid the mis-
estimation of the constraints. A large number of mocks
are required to quantify the covariance, which is out
of the capability if the full-resolution simulation is used.
In this section, we will focus on an important appli-
cation of the inverse-Gaussianization method, calculating
the covariance matrix between the power spectrum and
peak statistics.

Assume the cosmological parameter we care is \( \lambda \) and
the observables are band power \( C(\ell_i) \) with \( i = 1, \cdots, m \)
and peak height counts \( N(\kappa_{p,j}) \) with \( j = 1, \cdots, n \). The
data vector \( \mathbf{p} = [C(\ell_i), N(\kappa_{p,j})] \). The likelihood of the
joint analysis is

\[
\ln \mathcal{L} \propto \frac{\partial \ln \mathbf{p}}{\partial \lambda} \hat{\mathbf{C}}^{-1} \frac{\partial \ln \mathbf{p}^T}{\partial \lambda}. \tag{13}
\]

Here the normalized joint covariance \( \hat{\mathbf{C}} \) has a size of
\((m + n) \times (m + n)\) and contains the correlation between
band power and peak counts on different bins, and the
correlation between the two.

We set the number of bins to be 10 for both the power
spectrum and peak height counts. So the size of full cov-
ariance is \( 20 \times 20 \). In order to obtain the precise covari-
ce matrix, we use 40,000 weak gravitational lensing
mocks generated by the inverse-Gaussianization method
with the inclusion of the scatter of \( \bar{y} \) in a Gaussian form
(method ii). The result is presented in Fig. 10. The left
plot is the result from 96 simulated maps, while the right
plot is the result of 40,000 mocks. The joint covariance
from the mocks is obviously more smooth and less noisy.
However, the inverse-Gaussianization method overpro-
duces the correlation among the peak height bins and
the cross-correlation between the two observables. We
argue that this discrepancy comes from the inaccuracy in
adding the super-sample effects.

We also present the joint covariance for power spec-
trum and peak steepness from the same 40,000 conver-
gence mocks in the right plot in Fig. 11 and the left
plot is the result of simulated maps. Compared to the
simulated one, our method overestimate the covariance,
and again, this discrepancy comes from the difference in
the super sample variance in two cases. In particu-
lar, we find the correlation between the power spec-
trum and peak steepness counts is significantly weaker
than the joint of the power spectrum and peak height
counts. Based on the findings in Ribli et al. 2018, the
peak steepness is more powerful than peak height on
the constraints of \( \Omega_m \) and \( \sigma_8 \). The weaker correlation
between the two observables means that they provide
more independent information. The joint of second or-
der statistics and peak steepness has great potential to
break the degeneracy of \( \sigma_8 \).

5. CONCLUSION AND DISCUSSION

The next stage weak lensing survey has great statis-
tical power, implying the ability to use non-Gaussian
statistics to break the cosmological parameter degen-
rency in the joint analysis. In this work, we focused on
the peak statistics and investigated whether the peak statis-
tics could be reproduced by the inverse-Gaussianization
method.

We used 96 simulated weak lensing convergence maps
at \( z_s = 1 \) to obtain the Gaussianization transform func-
tion, and the power spectrum of the Gaussianized fields.
We fast generated the lensing mocks by generating the
Gaussian random fields according to the above power
spectrum and then inverse transform them. The result-

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{image1.png}
\caption{Both maps represent the covariance matrix between the power spectrum and peak heights. The 1-10 bins in both the x-axis and y-axis correspond with the power spectrum, while the 11-20 linear bins are the peak height counts for \( \kappa_p \in [-0.02, 0.1] \). The left panel is calculated by 96 original simulation fields, while we obtain the right map from 40,000 mocks generated by our improved inverse-Gaussianization method. We found that there is little difference between the two maps which is caused by the numbers of data. The matrix from mocks is more accurate than the left matrix, especially when compared with Fig. 4.
}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{image2.png}
\caption{Identical to Fig.10, but for the covariance matrix between power spectrum and peak steepness counts. The last 10 bins of both x-axis and y-axis represent the value of convergence peak steepness covering the range for \( 3 \times 10^{-4} < s_p < 5 \times 10^{-3} \).
}
\end{figure}
ing mocks have the same power spectrum with the simulated one. The peak height counts are also well reproduced. The difference from the simulated one is < 20% for \(-0.015 < \kappa_p < 0.065\) \((\kappa_p/\sigma_p \in [-1.2, 5.4])\) which covers the common value range of the recent weak lensing peak research. The peak steepness counts are also well reproduced and the difference from the simulation is < 20% for almost all the steepness range we investigated.

The covariance for the peak height counts and peak steepness counts is very close to diagonal, in contrast with the simulated cases. We found that the lack of super sample variance is the main reason for this difference. We proposed several methods to mimic the super sample effects and the (anti-)correlation among different bins is recovered. The remaining difference comes from the fact that we only mimic the super sample variance by tuning the mean value for these mocks.

We used this improved inverse-Gaussianization method with the inclusion of effective sample variance to obtain the covariance for the joint analysis of the power spectrum and the peak statistics. A greater potential of the combination between power spectrum and peak steepness was found, compared to the joint analysis between power spectrum and peak heights.

The inverse-Gaussianization method is to fast generate noiseless convergence fields based on a small patch of weak lensing simulations. They could be converted to cosmic shear fields, and further include the noise and mask effect to mimic the observation.

Our method produces slightly better distribution for the peak height counts than the result in Shirasaki 2017. Both methods use the local transform to obtain lensing mocks and they differ in details. The main difference is that Shirasaki 2017 uses theoretical power spectrum with corrections for resolution effects to obtain the power spectrum after Gaussianization. The new Gaussian random \(y\) fields are produced by this power spectrum with continuous \(\ell\). Our method directly measures the power spectrum from these Gaussianized fields, and the new \(y\) fields are produced from the measured band power. This process automatically includes the numerical/resolution effects.

The following details worth further investigation in the future. The missing of the rare extremely high peaks in the inverse-Gaussianization method could be partially solved by considering the scatter in the local transform function especially at the high end. However, these rare highest peaks correspond to very nonlinear small scale structures where the Gaussianization performs worse. It is interesting to study how well the inverse-Gaussianization method with further improvement could capture this complicated information.

The direct inverse-Gaussianization method cannot produce strong features in peak statistics covariance. This is due to the lack of super sample variance. Note that the simulated maps we use also miss some super sample modes since it is cut out from a light-cone constructed by stacking cubic simulation boxes. In principle, we could increase the map size to include the super sample variance. In the case that we do not have such a weak lensing simulation covering a very large portion of the sky, it is still doable thanks to the linearity at large scales. Given the cosmological parameters, the power spectrum at large scales is well predicted. Also at these scales, the power spectrum of the Gaussianized field is close to the power spectrum of the convergence field. We could adopt the theoretically predicted large scale power to construct even the full-sky weak lensing map.

6. ACKNOWLEDGMENTS
We thank Pengjie Zhang, Jun Zhang, Shuo Yuan for useful discussions. This work was supported by the National Key Basic Research and Development Program of China (No. 2018YFA0404504), and the National Science Foundation of China (grants No. 11773048, 11621303, 11890691). X.K.L. acknowledges the support from NSFC-11803028 and YNU Grant C176220100008.

REFERENCES
Abbott, T., Abdalla, F. B., Aleksi, J., et al. 2018, Monthly Notices of the Royal Astronomical Society, 460, 1270
Abbott, T. M. C., Abdalla, F. B., Alarcon, A., et al. 2018, PhRvD, 98, 043526
Agrawal, A., Makiya, R., Chiang, C.-T., et al. 2017, JCAP, 2017, 003
Aihara, H., Armstrong, R., Bickerton, S., et al. 2018, PASJ, 70, 88
Bartelmann, M., & Schneider, P. 2001, PhR, 340, 291
Blot, L., Crocce, M., Sefusatti, E., et al. 2019, MNRAS, 485, 2806
Chuang, C.-H., Kitaura, F.-S., Prada, F., Zhao, C., & Yepes, G. 2015, MNRAS, 446, 2621
Colavincenzo, M., Sefusatti, E., Monaco, P., et al. 2019, MNRAS, 482, 4883
Coulton, W. R., Liu, J., Madhavacheril, M. S., Böhm, V., & Spergel, D. N. 2019, JCAP, 2019, 043
Yuan, S., Liu, X., Pan, C., Wang, Q., & Fan, Z. 2018, ApJ, 857, 112

Zhang, T.-J., Pen, U.-L., Zhang, P., & Dubinski, J. 2003, ApJ, 598, 818