RE-CAPTURING COSMIC INFORMATION

**Hee-Jong Seo**¹,⁷, Masanori Sato², Scott Dodelson¹,³,⁴, Bhuvnesh Jain⁵, and Masahiro Takada⁶

¹ Center for Particle Astrophysics, Fermi National Accelerator Laboratory, Batavia, IL 60510, USA; hee-jongseo@lbl.gov
² Department of Physics, Nagoya University, Nagoya 464-8602, Japan
³ Department of Astronomy & Astrophysics, The University of Chicago, Chicago, IL 60637, USA
⁴ Kavli Institute for Cosmological Physics, Chicago, IL 60637, USA
⁵ Department of Physics & Astronomy, Center for Particle-Cosmology, University of Pennsylvania, Philadelphia, PA 19104, USA
⁶ Institute for the Physics and Mathematics of the Universe (IPMU), University of Tokyo, Chiba 277-8582, Japan

Received 2010 October 19; accepted 2010 December 8; published 2010 February 10

**ABSTRACT**

Gravitational lensing of distant galaxies can be exploited to infer the convergence field as a function of angular position on the sky. The statistics of this field, much like that of the cosmic microwave background (CMB), can be studied to extract information about fundamental parameters in cosmology, most notably the dark energy in the universe. Unlike the CMB, the distribution of matter in the universe which determines the convergence field is highly non-Gaussian, reflecting the nonlinear processes that accompanied structure formation. Much of the cosmic information contained in the initial field is therefore unavailable to the standard power spectrum measurements. Here we propose a method for re-capturing cosmic information by using the power spectrum of a simple function of the observed (nonlinear) convergence field. We adapt the approach of Neyrinck et al. to lensing by using a modified logarithmic transform of the convergence field. The Fourier transform of the log-transformed field has modes that are nearly uncorrelated, which allows for additional cosmological information to be extracted from small-scale modes.

**Key words:** cosmology: theory – gravitational lensing: weak – large-scale structure of universe

**Online-only material:** color figures

### 1. INTRODUCTION

Gravitational lensing has emerged as a powerful tool to probe the distribution of matter in the universe (Bartelmann & Schneider 2001). Observations of the ellipticities of background galaxies can be transformed into estimates of the convergence field $\kappa(\theta)$. Along a given line of sight $\theta$, the convergence measures a weighted integral of the total mass density field. Thus, by carefully studying $\kappa$ as a function of position on the sky, we can learn about the underlying density field directly, without relying on the traditional assumption that every galaxy corresponds to an overdense region.

By measuring the convergence to sources at multiple background redshifts, cosmologists can infer not only the density field as a function of two-dimensional (2D) position (Kaiser 1992; Blandford et al. 1991; Jain & Seljak 1997; Jain et al. 2000), but also the evolution of this density field with time (Hu 1999). This information will be particularly valuable as a tool to study both dark matter and dark energy, which affect the growth of structure in the universe (Huterer 2002; Hu 2002). A number of wide-area surveys have been planned with the goal of mapping out the cosmic convergence field, and ultimately measuring properties of the dark energy (Aldering 2005; Miyazaki et al. 2006; Tyson 2006; Abbott et al. 2005; Refregier & the DUNE Collaboration 2009).

This goal appears attainable, as it is reminiscent of another cosmological success story: measurement of anisotropies in the cosmic microwave background (CMB; Hu & Dodelson 2002). In both cases, the values of the measured quantities—temperature in the case of the CMB and convergence from lensing—at any particular spot on the sky are not important. Rather, it is the statistics of the field that carries all the important information. The two-point function of the temperature of the CMB, the power spectrum of the anisotropies, is sensitive to a number of cosmological parameters, and some of these have now been measured to percent level accuracy (Komatsu et al. 2011). Similarly, the power spectrum of the convergence depends on cosmological parameters, and one can hope to extract information about these parameters from lensing surveys (Hu & Tegmark 1999; Abazajian & Dodelson 2003; Refregier et al. 2004; Hoekstra & Jain 2008).

However, the convergence field differs in an important way from the anisotropy maps. CMB anisotropies provide a snapshot of the universe when it was very young, and hence all deviations from homogeneity are very small (temperature differences in the maps are of order several parts in a hundred thousand). The physics describing these perturbations is linear. Further, the perturbations were drawn from a Gaussian distribution, so the two-point function captures all of the information in the field. On the other hand, the cosmic density field today is nonlinear and non-Gaussian, increasingly so on smaller scales, so some of the information initially stored in the two-point function when the fields were linear is no longer present.

Before quantifying this notion that information has left the two-point function, it is worthwhile to review some approaches to this problem. Takada & Jain (2004) pointed out that including information from both the two- and three-point functions significantly reduces the errors on cosmological parameters. This makes intuitive sense: the nonlinear process of gravitational instability transforms the initially Gaussian field into one with appreciable non-Gaussianity, one hallmark of which is a nonzero skewness. The goal of measuring both sets of functions may work, but it suffers from the drawback of requiring non-trivial covariance matrices (which involve the challenge of computing five- and six-point functions; Takada & Jain 2009).

---

² Current address: Berkeley Center for Cosmological Physics, LBL and Department of Physics, University of California, Berkeley, CA 94720, USA.
A series of papers devoted to the three-dimensional (3D) density field \( \delta (\vec{x}) \equiv (\rho (\vec{x}) - \bar{\rho})/\bar{\rho} \) (Rimes & Hamilton 2005, 2006; Neyrinck & Szapudi 2007; Neyrinck et al. 2006, 2009) have noted that information in the power spectrum of \( \delta \) saturates at high wavenumbers \( k \) (or small length scales). That is, the power spectrum at high \( k \) is highly correlated, apparently due to the coupling of modes induced by nonlinear gravitational clustering. The most recent of these papers offered a useful proposal (Neyrinck et al. 2009) for re-capturing information about the 3D density field by pointing out that \( \ln(1 + \delta) \) has properties similar to the initial, linear density field. Its probability distribution is close to a Gaussian, the broadband power spectrum is shown by the dotted curve. The skewed PDF of \( \kappa \ln \) reduces to the standard convergence, but the log transform restores the field to a PDF that is nearly Gaussian. The latter has smaller amplitude at high \( l \). The linear power spectrum is shown by the dotted curve.

\( \kappa \ln \) is nearly Gaussian. A first glimpse into the advantages of the log transform can be seen from Figure 1 which shows the probability distribution function (PDF) of both \( \kappa \) and \( \kappa_{\ln} \), compared to the (linear) Gaussian PDF. The new field is much closer to Gaussian, a promising sign since the loss of information in \( \kappa \) is attributed to gravity transforming the initially Gaussian random field into one that is highly non-Gaussian.

To evaluate the log transform quantitatively, we take the Fourier transform of the three different convergence fields (linear, \( \kappa \), and \( \kappa_{\ln} \)) in each of the simulations. The angular power spectrum is estimated from the Fourier transforms (denoted \( \bar{\kappa} (\vec{l}) \)) by summing over all modes with wavenumber \( |\vec{l}| \) in a given bin \( \Delta \bar{\kappa}_{\ln}. \)

Figure 2 shows these spectra. As expected, the power spectrum of the nonlinear \( \kappa \) field is much larger than the linear field on small scales (large \( l \)). This excess power on small scales is suppressed when \( \kappa_{\ln} \) is used. Again the result is not surprising, as the high-density regions are smoothed out: \( \kappa_{\ln} \ll \kappa \) for large \( \kappa \).

3. RECOVERY OF COSMOLOGICAL INFORMATION

Although the power spectrum of \( \kappa_{\ln} \) is smaller than that of \( \kappa \), it contains more cosmological information. To see this, consider alter the field. The log-mapping described above is motivated by our goal to de-correlate the Fourier modes of the convergence field. Although the mapping is local on the sky, it is nonlinear, so in Fourier space it has the potential to undo some of the correlations introduced by nonlinear clustering.

We now show how the log transform can be applied to the 2D lensing convergence field to de-correlate modes and obtain information from higher order correlations back in the two-point function.

2. LOG-MAPPING FOR LENSING

Using simulations, we study the statistics of a new field:

\[
\kappa_{\ln}(\vec{\theta}) \equiv \kappa_{0} \ln \left[ 1 + \frac{\kappa(\vec{\theta})}{\kappa_{0}} \right],
\]

where \( \kappa_{0} \) is a constant with a value slightly larger than the absolute value of the minimum value of \( \kappa \) in the survey—this keeps the argument of the logarithm positive. In the limit of small \( \kappa \), \( \kappa_{\ln} \) reduces to the standard convergence, but the log alters it in very high or low density regimes. The parameter \( \kappa_{0} \) tunes the degree of the alteration: the smaller \( \kappa_{0} \), the more we
a model with one free parameter, the amplitude of the observed, nonlinear power spectrum before and after the log transform. The projected fractional error on this parameter is the inverse of the signal-to-noise ratio (S/N) defined as

\[ \frac{S}{N}(l_{\text{max}}) \equiv \left[ \sum_{l,l'} C_l \text{Cov}^{-1}(l, l') C_{l'} \right]^{1/2} \tag{2} \]

where \( C_l \) is the power spectrum of multipole \( l \) before and after the transform, \( \text{Cov} \) is the covariance matrix describing correlations between the power spectra of multipoles \( l \) and \( l' \) (\( l, l' < l_{\text{max}} \)), and the summation runs over all the multipoles \( l \) and \( l' \) subject to \( l, l' < l_{\text{max}} \) (Sato et al. 2009; Takahashi et al. 2009). We measure the non-Gaussian covariance matrix from the dispersions of the 100 convergence fields before and after the transform. We follow Neyrinck et al. (2009) and call the square of the S/N the information content. Heuristically, then, “information” quantifies how accurately parameters will be determined. To compute the expected error on the chosen cosmological parameter (here the amplitude of the power spectrum; Lee & Pen 2008), one needs to know the covariance matrix of the spectra. If the field was Gaussian random, the covariance matrix would be diagonal. In the absence of shape noise, it would arise from sample variance and be equal to the spectrum squared divided by the number of independent modes in the bin. In that case, since the number of modes in a bin grows as \( l \) for log binning, the \( (S/N)^2 \) would grow as \( l_{\text{max}}^2 \).

Figure 3 shows the \( (S/N)^2 \) as a function of \( l_{\text{max}} \). The linear \( \kappa \) field is shown by the dotted gray line. The information obtained from the nonlinear \( \kappa \) field falls well below this ideal limit, as seen in the figure. This arises because the nonlinearities significantly affect the covariance matrix. Non-zero off-diagonal elements in the covariance matrix mean that many of the modes carry redundant information, so the total gain is significantly below the \( l_{\text{max}}^2 \) Gaussian limit. The log transform undoes a large portion of this damage. The left panel of Figure 3 shows that the information in \( \kappa_{\text{ln}} \) is well above that in \( \kappa \) and close to the Gaussian case. In other words, we measure the amplitude of the power spectrum with higher precision if we use the log-transformed field. We find a factor of \( \sim 1.3 \) improvement in \( (S/N)^2 \) at \( l_{\text{max}} \sim 250 \), a factor of \( \sim 2.6 \) at \( l_{\text{max}} \sim 1000 \), and a factor of 4 at \( l_{\text{max}} \sim 2000 \), and a factor of 8 at \( l_{\text{max}} \sim 5000 \).

The restored information in the \( \kappa_{\text{ln}} \) field can be understood by examining the covariance matrix of the power spectra. Figure 4 shows two rows of the covariance matrix for the fields, with one of the wavenumbers fixed at \( l' = 253 \) and \( l' = 1049 \) in the two cases (upper and lower panel). The \( \kappa \) covariance matrix has large off-diagonal elements in adjacent bins—these carry redundant information and therefore do not add much to the S/N. The transformed \( \kappa_{\text{ln}} \), on the other hand, is much more nearly diagonal. A nearly diagonal covariance matrix implies another important advantage of the log transform: the approximation...
of a Gaussian covariance matrix for cosmological parameter estimation is more accurate for $\kappa_{\ln}$.

Another way of understanding the gain in information in the log field is to consider the Taylor expansion of the log transform $\kappa_{\ln}$. For $-1 < \kappa / \kappa_0 \lesssim 1$, one sees that $\kappa_{\ln}$ contains the standard convergence field, but also a piece that scales as $\kappa^2$ (and higher orders). Considered perturbatively, then, the spectrum of $\kappa_{\ln}$ will depend not only on the two-point function of $\kappa$, $C_\kappa$, but also on the three-point function, the bi-spectrum, as well as higher-point functions. Effectively, then this rather simple transform captures information in the power spectrum, bi-spectrum, tri-spectrum, etc., in a compact way. Of course, it does not contain all the information in these higher point functions, but the improvement seen in Figure 3 suggests that using $\kappa_{\ln}$ as a transform in future surveys may be a simple, powerful way to bundle much of this information into one simple spectrum. We have tested this by measuring the information contained in $\kappa_{\ln} \equiv \kappa - \kappa^2 / (2\kappa_0)$ and found that, once we apply an appropriate cutoff\(^{10}\) on high $\kappa$ values to make the polynomial expansion more sensible, the single extra term replicates most of the improvement observed in the log transform. Meanwhile, the cross-correlation of the $\kappa$ and $\kappa - \kappa^2 / \kappa_0$ which involves only up to bi-spectrum, with an appropriate cutoff, replicates most of the improvement up to $l \sim 1000$. This implies that the bi-spectrum is the dominant contributor to this improvement up to the scales.

There are several caveats to this analysis. So far, we have neglected noise, in particular shape noise due to the random orientations of galaxies on the sky. We have studied this issue for several survey parameters. Surveys with higher number density have lower shape noise and therefore the advantages of $\kappa_{\ln}$ approach those depicted in the left panel of Figure 3. For a galaxy number density of 30 arcmin$^{-2}$ at $z_g = 1$,\(^{11}\) as expected for the planned Subaru Hyper SuprimeCam survey (Miyazaki et al. 2006), we find an improvement of 1.7 (2.4) in the information content for $l_{\text{max}} \sim 1000$ (2000) (right panel in Figure 3). The gain is larger for more ambitious surveys like LSST or Euclid (Tyson 2006; Refregier et al. 2010) and smaller for shallower surveys like the Dark Energy Survey (Abbott et al. 2005). Further details will be presented in our future paper.

Second, although $\kappa_{\ln}$ has some of the advantages of the linear $\kappa$ field, it does not actually recover the initial field phase by phase since the cross-correlation between the initial and final fields, when tested for the density fields, does not improve by this transformation. Third, our analysis (and our definition of information) revolved around only one parameter, the amplitude of the power spectrum. Its shape and evolution certainly contain additional cosmological information, as discussed by Takada & Jain (2009). We therefore need to test that the improvement in $(S/N)^2$ in the amplitude translates to the improved precision on the parameters such as $\sigma_8$, matter density, and dark energy parameters. We plan to conduct a full Fisher matrix analysis using $N$-body simulations in our future paper to investigate this issue.

Finally, we have assumed that the convergence field, reconstructed from the shear, will be available over the entire survey area—in practice such a reconstruction adds additional noise. We are in the process of studying these issues, but they are not expected to affect our main point: that the log transform $\kappa_{\ln}$ recovers cosmological information.

---

\(^{10}\) We remove the high $\kappa$ values by replacing $\kappa$ larger than 0.1 with 0.1.

\(^{11}\) We subtract the shape noise contribution from $C_\ell$ when calculating the information content.