FERMILAB-PUB-10-294-A-T

Re-capturing cosmic information

Hee-Jong Seo[1], Masanori Sato[2], Scott Dodelson[1,3,4], Bhuvnesh Jain[5], and Masahiro Takada[6]

[1]*Center for Particle Astrophysics, Fermi National Accelerator Laboratory, Batavia, IL 60510*

[2]*Department of Physics, Nagoya University, Nagoya 464-8602, Japan*

[3]*Department of Astronomy & Astrophysics,*

The University of Chicago, Chicago, IL 60637

[4]*Kavli Institute for Cosmological Physics, Chicago, IL 60637*

[5]*Department of Physics & Astronomy, Center for Particle-Cosmology,*

University of Pennsylvania, Philadelphia PA 19104 and

[6]*Institute for the Physics and Mathematics of the Universe (IPMU),*

University of Tokyo, Chiba 277-8582, Japan

(Dated: August 3, 2010)

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 which 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 (non-linear) convergence field. We adapt the approach of Neyrinck et al. (2009) 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.

PACS numbers:

arXiv:1008.0349v1 [astro-ph.CO] 2 Aug 2010

Operated by Fermi Research Alliance, LLC under Contract No. De-AC02-07CH11359 with the United States Department of Energy.

I. INTRODUCTION

Gravitational lensing has emerged as a powerful tool to probe the distribution of matter in the Universe [1]. Observations of the ellipticities of background galaxies can be transformed into estimates of the *convergence* field $\kappa(\vec{\theta})$. Along a given line of sight $\vec{\theta}$, the convergence measures a weighted integral of the total mass density field. Thus by carefully studying κ 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 2D position [2–5], but also the evolution of this density field with time [6]. 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 [7, 8]. 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 [9–13].

This goal appears attainable as it is reminiscent of another cosmological success story: measurement of anisotropies in the CMB [14]. 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 [15]. Similarly, the power spectrum of the convergence depends on cosmological parameters, and one can hope to extract information about these parameters from lensing surveys [16–19].

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 non-linear 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 worth-

while to review some approaches to this problem. Takada & Jain [20] 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 non-zero 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) [21].

A series of papers devoted to the 3D density field $\delta(\vec{x}) \equiv (\rho(\vec{x}) - \bar{\rho})/\bar{\rho}$ [22–26] have noted that information in the power spectrum of δ 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 [26] 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 shape is close to that of the linear power spectrum, and finally, the information content is close to the Gaussian case. Practically this transform may be of limited utility because the 3D density field is typically estimated by using galaxies as tracers, and it is unlikely that the log transform of the *galaxy* density will be a useful tracer of the linear *matter* density field. However, we now show that 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.

II. LOG-MAPPING FOR LENSING

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

$$\kappa_{\ln}(\vec{\theta}) \equiv \kappa_0 \ln\left[1 + \frac{\kappa(\vec{\theta})}{\kappa_0}\right] \tag{1}$$

where κ_0 is a constant with a value slightly larger than the absolute value of the minimum value of κ in the survey – this keeps the argument of the logarithm positive. In the limit of small κ, $\kappa_{\ln}$ reduces to the standard convergence, but the log alters it in very high or low density regimes. The parameter κ_0 tunes the degree of the alteration: the smaller κ_0, the more we alter the field[1]. The

[1] For our fiducial maps with 0.15 arcmin pixel scale, we use $\kappa_0 = 0.0482$ based on the minimum value of measured κ.

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.

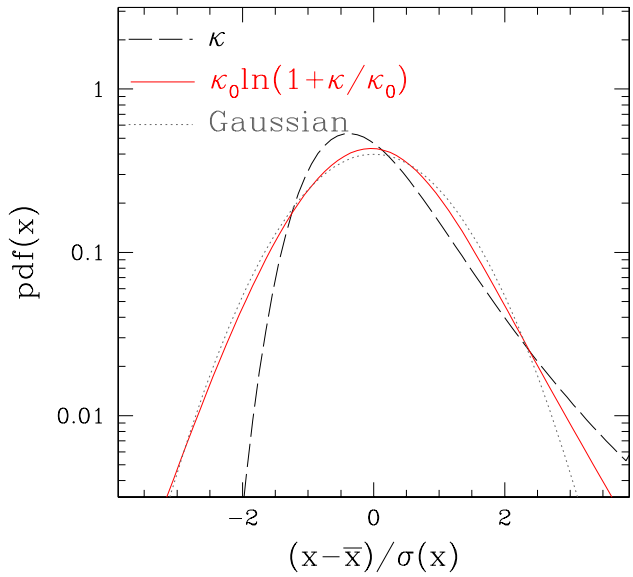

FIG. 1: The probability distribution function of the two fields κ (black, dashed) and $\kappa_{\rm ln}$ (red, solid) in comparison with a Gaussian Probability Distribution Function (dotted). The skewed PDF of κ reflects the distribution of structure in the Universe: large underdense regions separated by some very overdense regions. The log transform restores the field to a PDF that is nearly Gaussian.

To study the properties of $\kappa_{\rm ln}$, we use a suite of numerical simulations: 100 convergence fields, each $5^\circ \times 5^\circ$ (a total of 2500 square degrees) with 2048^2 pixels (i.e., 0.15 arcmin per pixel) were generated using N-body simulations as described in [27]. All source galaxies are taken to be at redshift $z_s = 1$ for all the results shown below, though we have also checked other source redshifts.

A first glimpse into the advantages of the log transform can be seen from Fig. 1 which shows the probability distribution function (PDF) of both κ and $\kappa_{\rm ln}$, compared to the (linear) Gaussian PDF. The new field is much closer to Gaussian, a promising sign since the loss of information in κ 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, κ, and $\kappa_{\rm ln}$) in each of the simulations. The angular power spectrum is estimated from the Fourier transforms (denoted $\tilde{\kappa}(\vec{l})$) by summing over all modes with wavenumber $|\vec{l}|$ in a given bin $l_{\rm bin}$.

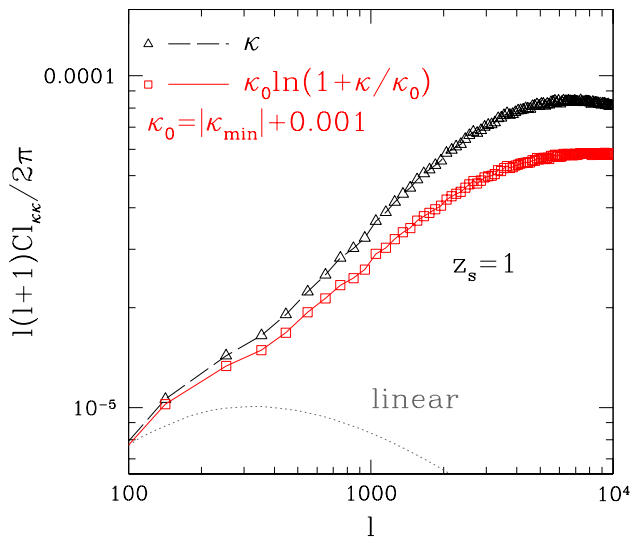

FIG. 2: The measured power spectra of the convergence field κ and the log transformed field $\kappa_{\rm ln}$. The latter has smaller amplitude at high l. The linear power spectrum is shown by the dotted curve.

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

III. RECOVERY OF COSMOLOGICAL INFORMATION

Although the power spectrum of $\kappa_{\rm ln}$ is smaller than that of κ, it contains more cosmological information. To see this, consider 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 defined as

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

where C_l is the power spectrum of multipole l before and after the transform, Cov is the covariance matrix describing correlations between the power spectra of multipoles l and l' ($l, l' < l_{\rm max}$), and the summation runs over all the multipoles l and l' subject to $l, l' < l_{\rm max}$ [27, 28]. We follow [26] and call the square of the S/N ratio 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 [29]), 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[2], it would be 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 $l_{\rm max}^2$.

Fig. 3 shows the $(S/N)^2$ as a function of $l_{\rm max}$. The linear κ field is shown by the dotted gray line. The information obtained from the nonlinear κ 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 $l_{\rm max}^2$ Gaussian limit. The log transform undoes a large portion of this damage. The left panel of Fig. 3 shows that the information in $\kappa_{\rm ln}$ is well above that in κ 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 ~ 1.3 improvement in $(S/N)^2$ at $l_{\rm max} \sim 250$, a factor of ~ 2.6 at $l_{\rm max} \sim 1000$, a factor of 4 at $l_{\rm max} \sim 2000$, and a factor of 8 at $l_{\rm max} \sim 5000$.

The restored information in the $\kappa_{\rm ln}$ field can be understood by examining the covariance matrix of the power spectra. Fig. 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 κ 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 $\kappa_{\rm ln}$, 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 $\kappa_{\rm ln}$.

Another way of understanding the gain in information in the log field is to consider the Taylor expansion of the log transform $\kappa_{\rm ln}$. For $-1 < \kappa/\kappa_0 \leq 1$, one sees that $\kappa_{\rm ln}$ contains the standard convergence field, but also a piece that scales as κ^2 (and higher orders). Considered perturbatively, then, the spectrum of $\kappa_{\rm ln}$ will depend not only on the 2-point function of κ, C_l, but also on the 3-point function, the bispectrum, as well as higher-point functions. Effectively, then this rather

[2] The ellipticity of a single galaxy is, in the absence of any distortion by the intervening density field, randomly distributed on the sky with an RMS of about 0.3. This corresponds to noise in the measurement of the cosmic convergence field, a noise which decreases as the square root of the number of galaxies in a pixel. The resulting noise is called shape noise.

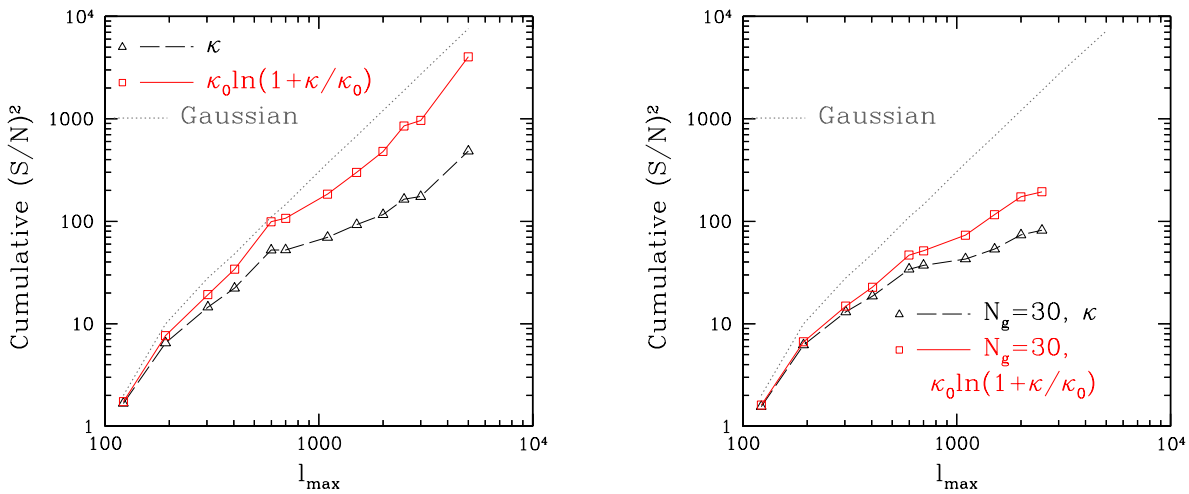

FIG. 3: Left: the information, represented by $(S/N)^2$, contained in the two fields κ and $\kappa_{\rm ln}$ in comparison with the Gaussianized field. The information in the Gaussianized κ field (dotted curve) increases as $l_{\rm max}^2$ as smaller scales are included. The actual nonlinear convergence κ (dashed black line/triangles) loses much of the $(S/N)^2$ at large l, while the log transform (solid red line/squares) recovers it. Right: the effect of the log transform in the presence of shape noise: we assume a galaxy number density $N_g = 30/{\rm arcmin}^2$ at $z_s = 1$ and increase the pixel size to 2.4 arcmin (accordingly we use $\kappa_0 = 0.112$). We find an improvement of 1.7 (2.4) in the information content for $l_{\rm max} \sim 1000$ (2000) even in the presence of shape noise.

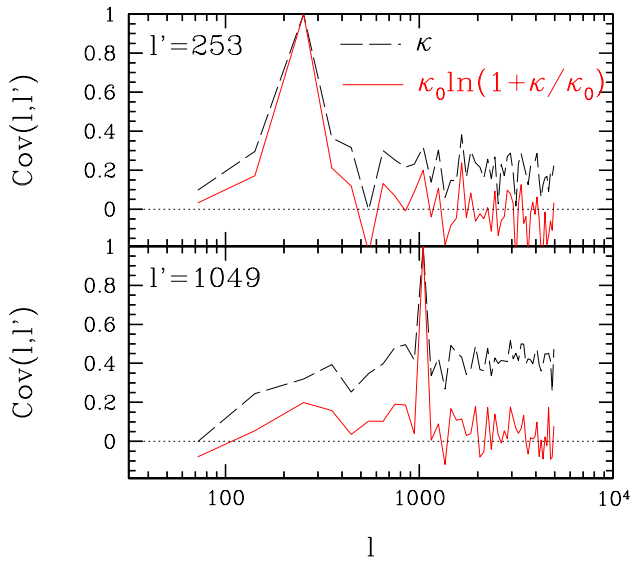

FIG. 4: Slices of the covariance matrix. The off-diagonal elements are normalized relative to the diagonal elements, i.e., ${\rm Cov}(l,l')/({\rm Cov}(l,l){\rm Cov}(l',l'))^{1/2}$ are shown as a function of l for two choices of l'. The off-diagonal covariances between different scales have been substantially decreased by the log transform.

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 Fig. 3 suggests that using $\kappa_{\rm ln}$ 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 $\kappa'_{\rm ln} \equiv \kappa - \kappa^2/(2\kappa_0)$ and found that, once we apply an appropriate cutoff[3] on high κ 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 κ and $\kappa - \kappa^2/\kappa_0$ which involves only up to bispectrum, with an appropriate cutoff, replicates most of the improvement up to $l \sim 1000$. This implies that the bispectrum 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 $\kappa_{\rm ln}$ approach those depicted in the left panel of Fig. 3. For a galaxy number density of 30 per square arcminutes at $z_s = 1$, as expected for the planned Subaru Hyper SuprimeCam survey [10], we find an improvement of 1.7 (2.4) in the information content for $l_{\rm max} \sim 1000$ (2000) (right panel in Fig. 3). The gain is larger for more ambitious surveys like LSST or DUNE [11, 13] and smaller for shallower surveys like the Dark Energy Survey [12].

Second, although $\kappa_{\rm ln}$ has some of the advantages of the linear κ 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 [21]. 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 affects our main point: that the log transform $\kappa_{\rm ln}$ recovers cosmological information.

[3] We remove the high κ values by replacing κ larger than 0.1 with 0.1.

IV. CONCLUSION

We have found that the log transform of Eq. (1) alters the nonlinear convergence field to one that mimics the properties of a Gaussian field. It returns a PDF that is close to a Gaussian – analogous to the findings of [26] for the 3D density field. The signal-to-noise (i.e., precision) of the measurement of the amplitude of the power spectrum is greatly improved over a wide range of angular scales, $200 \lesssim l \lesssim 10^4$. Even in the presence of shape noise, this improvement holds, to a greater or lesser extent depending on the galaxy number density. The improvement arises from the effect on the covariance matrix: the off-diagonal elements of the covariance matrix are substantially reduced for the log transform. We find that the bispectrum that is embedded in the log transform is the dominant contributor to this improvement. Therefore the log transform appears to bundle much of the information from higher order statistics into the power spectrum.

Upcoming imaging surveys will collect data on the shapes of galaxies at an unprecedented rate, with an eye towards understanding the physics which drives the acceleration of the Universe. It is imperative that we use algorithms to analyze this data which extract as much of the cosmological information as possible: the log transform $\kappa_{\rm ln}$ is a step in this direction.

Acknowledgement: We would like to thank David Johnston, Nick Gnedin, Mark Neyrinck and Alex Szalay for useful comments.

This work is supported in part by World Premier International Research Center Initiative (WPI Initiative), by Grant-in-Aid for Scientific Research on Priority Areas No. 467 “Probing the Dark Energy through an Extremely Wide and Deep Survey with Subaru Telescope”, MEXT, Japan, by the JSPS Research Fellows, by the DOE at Fermilab through grant DE-FG02-95ER40896, and by NSF Grants AST-0908072 and AST-0607667.

[1] M. Bartelmann and P. Schneider, Phys. Rept. **340**, 291 (2001), astro-ph/9912508.

[2] N. Kaiser, Astrophys. J. **388**, 272 (1992).

[3] R. D. Blandford, A. B. Saust, T. G. Brainerd, and J. V. Villumsen, Mon. Not. Roy. Astron. Soc. **251**, 600 (1991).

[4] B. Jain and U. Seljak, Astrophys. J. **484**, 560 (1997), astro-ph/9611077.

[5] B. Jain, U. Seljak, and S. D. M. White, Astrophys. J. **530**, 547 (2000), astro-ph/9901191.

[6] W. Hu, Astrophys. J. **522**, L21 (1999), astro-ph/9904153.

[7] D. Huterer, Phys. Rev. D **65**, 063001 (2002), astro-ph/0106399.

[8] W. Hu, Phys. Rev. D **65**, 023003 (2002), astro-ph/0108090.

[9] G. Aldering, New Astron. Rev. **49**, 346 (2005), astro-ph/0507426.

[10] S. Miyazaki, Y. Komiyama, H. Nakaya, Y. Doi, H. Furusawa, P. Gillingham, Y. Kamata, K. Takeshi, and K. Nariai, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series (2006), vol. 6269 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series.

[11] J. A. Tyson (LSST), AIP Conf. Proc. **870**, 44 (2006), astro-ph/0609516.

[12] T. Abbott et al. (Dark Energy Survey) (2005), astro-ph/0510346.

[13] A. Refregier and the DUNE collaboration, Exper. Astron. **23**, 17 (2009), 0802.2522.

[14] W. Hu and S. Dodelson, Ann. Rev. Astron. Astrophys. **40**, 171 (2002), astro-ph/0110414.

[15] E. Komatsu et al. (2010), 1001.4538.

[16] W. Hu and M. Tegmark, Astrophys. J. **514**, L65 (1999), astro-ph/9811168.

[17] K. N. Abazajian and S. Dodelson, Phys. Rev. Lett. **91**, 041301 (2003), astro-ph/0212216.

[18] A. Refregier et al., Astron. J. **127**, 3102 (2004), astro-ph/0304419.

[19] H. Hoekstra and B. Jain, Ann. Rev. Nucl. Part. Sci. **58**, 99 (2008), 0805.0139.

[20] M. Takada and B. Jain, Mon. Not. Roy. Astron. Soc. **348**, 897 (2004), astro-ph/0310125.

[21] M. Takada and B. Jain, Mon. Not. Roy. Astron. Soc. **395**, 2065 (2009), 0810.4170.

[22] C. D. Rimes and A. J. S. Hamilton (2005), astro-ph/0502081.

[23] C. D. Rimes and A. J. S. Hamilton, Mon. Not. Roy. Astron. Soc. **371**, 1205 (2006), astro-ph/0511418.

[24] M. C. Neyrinck and I. Szapudi, Mon. Not. Roy. Astron. Soc. **375**, L51 (2007), astro-ph/0610211.

[25] M. C. Neyrinck, I. Szapudi, and C. D. Rimes, Mon. Not. Roy. Astron. Soc. Lett. **370**, L66 (2006), astro-ph/0604282.

[26] M. C. Neyrinck, I. Szapudi, and A. S. Szalay, Astrophys. J. **698**, L90 (2009), 0903.4693.

[27] M. Sato et al., Astrophys. J. **701**, 945 (2009), 0906.2237.

[28] R. Takahashi, N. Yoshida, M. Takada, T. Matsubara, N. Sugiyama, I. Kayo, A. J. Nishizawa, T. Nishimichi, S. Saito, and A. Taruya, Astrophys. J. **700**, 479 (2009), 0902.0371.

[29] J. Lee and U. Pen, Astrophys. J. **686**, L1 (2008), 0807.1538.