The Astrophysical Journal, 656:27–41, 2007 February 10


# CROSS-CORRELATION LENSING: DETERMINING GALAXY AND CLUSTER MASS PROFILES FROM STATISTICAL WEAK-LENSING MEASUREMENTS

David E. Johnston,[1,2] Erin S. Sheldon,[3] Argyro Tasitsiomi,[4] Joshua A. Frieman,[5,6,7]
Risa H. Wechsler,[5,6,8] and Timothy A. McKay[9]
*Received 2005 July 20; accepted 2006 October 9*

## ABSTRACT

We present a new nonparametric method for determining mean 3D density and mass profiles from weak-lensing measurements around stacked samples of galaxies or clusters, that is, from measurement of the galaxy-shear or cluster-shear correlation functions. Since the correlation function is statistically isotropic, this method evades problems, such as projection of large-scale structure along the line of sight or halo asphericity, that complicate attempts to infer masses from weak-lensing measurements of individual objects. We demonstrate the utility of this method in measuring halo profiles, galaxy-mass and cluster-mass cross-correlation functions, and cluster virial masses. We test this method on an *N*-body simulation and show that it correctly and accurately recovers the 3D density and mass profiles of halos. We find no evidence of problems due to a mass sheet degeneracy in the simulation. Cross-correlation lensing provides a powerful method for calibrating the mass-observable relation for use in measurement of the cluster mass function in large surveys. It can also be used on large scales to measure and remove the halo bias and thereby provide a direct measurement of $\Omega_m\sigma_8$.

*Subject headings:* galaxies: clusters: general — gravitational lensing — large-scale structure of universe

## 1. INTRODUCTION

Gravitational lensing has become a powerful tool for studying cosmology and especially for studying the unseen matter that dominates the universe. Weak gravitational lensing produces an image of the universe that is slightly distorted: these distortions reveal the quantity and clustering properties of the elusive dark matter. One major application of weak lensing has been measurement of the mass distribution around rich galaxy clusters (Fahlman et al. 1994; Tyson & Fischer 1995; Luppino & Kaiser 1997; Clowe et al. 1998; Joffre et al. 2000; Irgens et al. 2002; Cypriano et al. 2004) This has allowed the reconstruction of two-dimensional (2D) projected cluster mass density maps that, combined with optical and X-ray measurements, have enabled measurement of mass-to-light ratios and determination of the cluster baryon fraction.

Another important use of weak lensing is the measurement of the power spectrum of mass fluctuations through cosmic shear, which has now been detected by several groups (Van Waerbeke et al. 2000; Wittman et al. 2000; Bacon et al. 2000; Maoli et al. 2001; Rhodes et al. 2001; Hoekstra et al. 2002; Jarvis et al. 2003; Bacon et al. 2003; Hamana et al. 2003). Cosmic shear, or shear-shear correlations, probes cosmology through its dependence on both the growth of structure and the expansion history of the universe (Hu 2002; Abazajian & Dodelson 2003; Takada & Jain 2004; Song & Knox 2004; Refregier et al. 2004); since it is sensitive only to the matter distribution, cosmic shear is theoretically clean.

While cluster lensing allows the study of the mass distribution of individual objects, cosmic shear is a purely statistical quantity analogous to the angular power spectrum of the cosmic microwave background. There is another method, called galaxy-galaxy lensing or galaxy-shear correlations, which is a hybrid of the two. Galaxy-galaxy lensing measures the mean weak-lensing shear produced by a sample of lens galaxies by averaging the shear signal over large numbers of them. It is thus a statistical measure that can be applied to different samples of galaxies to measure the average mass distribution around different galaxy types. After early null results (Kristian 1967; Tyson et al. 1984), the first detections using this technique were made by Brainerd et al. (1996), dell'Antonio & Tyson (1996), and Griffiths et al. (1996); these measurements have quickly improved (Hudson et al. 1998; Fischer et al. 2000; Wilson et al. 2001; McKay et al. 2001; Smith et al. 2001; Hoekstra et al. 2003). The Sloan Digital Sky Survey (SDSS; York et al. 2000) has provided an excellent sample for measuring galaxy-galaxy lensing: Sheldon et al. (2004) measured the weak-lensing signal around the SDSS main spectroscopic sample to high accuracy (see also Seljak et al. 2005). The galaxy-galaxy lensing signal is best interpreted, in the language of large-scale structure (LSS), as an estimate of the galaxy-mass correlation function. Combined with the galaxy autocorrelation function, it gives us information about galaxy biasing; on large scales, it provides a direct measurement of $\sigma_8\Omega_m$ that assumes only the physics of general relativity.

Another arena in which such statistical weak-lensing methods can be applied is the study of galaxy clusters. The number counts of galaxy clusters as a function of mass and redshift is a sensitive probe of cosmological parameters. The cluster mass function has been measured by several groups using various proxies for cluster mass (Bahcall & Cen 1992; White et al. 1993; Viana & Liddle 1999; Reiprich & Böhringer 2002; Bahcall et al. 2003). The measurement of cluster number counts at low redshift is usually used to constrain the combination $\sigma_8\Omega_m^\beta$ with $\beta \sim 0.5$, although the exact value of $\beta$ really depends on which range of mass scales is being probed (Rozo et al. 2004). Measurement of the evolution of the mass function with redshift allows one to break the degeneracy between these two parameters (Bahcall & Bode 2003). Galaxy clusters are strongly biased compared to the dark matter, and this bias as a function of halo mass is a prediction of theoretical models

[1] Jet Propulsion Laboratory, Pasadena, CA; davej@astro.caltech.edu.
[2] California Institute of Technology, Pasadena, CA.
[3] Physics Department, New York University, New York, NY.
[4] Princeton University Observatory, Princeton, NJ.
[5] Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL.
[6] Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL.
[7] Fermi National Accelerator Laboratory, Batavia, IL.
[8] Hubble Fellow; current address: Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, Stanford, CA.
[9] Physics Department, University of Michigan, Ann Arbor, MI.

(Kaiser 1984; Mo et al. 1996; Sheth & Tormen 1999; Seljak & Warren 2004). A measurement of cluster bias as a function of mass via cluster correlations thus provides an independent constraint on $\sigma_8$ (Rozo et al. 2004). In the future, planned large-area surveys will aim to determine the properties of the dark energy (its equation of state and energy density) by counting clusters and measuring their clustering to redshifts $z \sim 1$–2 (Haiman et al. 2001).

One of the main challenges to using galaxy clusters to constrain cosmology is uncertainty in the determination of cluster masses. Typically one selects clusters and/or infers their masses based on some observable proxy for mass, such as X-ray temperature or luminosity (Borgani & Guzzo 2001; Del Popolo 2003), Sunyaev-Zel'dovich effect (SZE) flux decrement (Grego et al. 2001), optical richness, or galaxy velocity dispersion (van der Marel et al. 2000; McKay et al. 2002; Łokas & Mamon 2003). In all of these cases, the mass-observable relation, as well as its scatter and possible evolution with redshift, must be calibrated (Pierpaoli et al. 2003) either by simulations, by combining number counts with spatial clustering information (Majumdar & Mohr 2004; Lima & Hu 2004, 2005), or by "direct" mass measurements of a subsample of the clusters. X-ray and SZE measurements are sensitive to varying degrees to the physics of the intracluster medium, such as radiative transfer, cooling, accretion, and feedback mechanisms associated with star and galaxy formation, possibly leading to scatter in the mass-observable relation. Dynamical mass measurements through the galaxy velocity dispersion suffer from uncertainties in the anisotropy of the velocity distribution function, possible bias of the galaxy velocities relative to the dark matter, and possible violation of the required assumption of dynamical equilibrium (e.g., due to recent mergers).

Weak lensing provides a method for measuring cluster masses directly and thereby calibrating mass-observable relations. The advantage of lensing is that it is sensitive to the total mass regardless of its makeup or dynamical state; the only physics involved is the deflection law given by general relativity. However, lensing measurements have limitations of their own. Weak-lensing mass reconstruction of individual clusters can only obtain a 2D surface mass density. There is no information about where the mass along the line of sight is located. Typically these measurements come from pointed observations of known massive clusters, selected by X-ray luminosity or some other observational property. Because the observations need to go deep enough to obtain the necessary density of background source galaxies to make spatially resolved mass maps, they are time-consuming and therefore tend to be limited in area, usually not extending much beyond the virial radius. Because of this limited range of physical scales covered, there are uncertainties arising from the mass sheet degeneracy (Bradač et al. 2004). In addition, LSS along the line of sight and chance projections with clusters at different redshifts can influence individual cluster mass determinations, causing ~30% scatter around the true mass and possibly a bias in the mass determination (Cen 1997; Metzler et al. 1999, 2001; Hoekstra 2001, 2003; White et al. 2002; Clowe et al. 2004; Dodelson 2004).

In this paper we demonstrate a new way to measure the average density and mass profile of clusters using stacked samples that, in principle, can be selected by any observable proxy. We introduce a completely nonparametric method to deproject the average shear profile and obtain the mean three-dimensional (3D) density and mass profiles for a sample. This method only assumes the statistical isotropy of the cluster-mass correlation function. We perform tests on an $N$-body simulation to show that the method correctly recovers the average 3D density and mass profiles. From the derived mass profile of the clusters, one can measure a virial radius and virial mass without assuming a model for the profile. Unlike individual cluster lensing reconstructions, these mass estimates are not adversely affected by LSS or chance projections along the line of sight. We show that these methods will allow one to constrain the density profiles of galaxy clusters, calibrate mass-observable relations, and thereby measure the cluster mass function and bias.

In § 2 we derive the relations between tangential shear measurements and the 2D surface mass density. In § 3 we invert the relation to obtain the 3D density profile. In § 4 we present methods of deriving 3D mass profiles, and in § 5 we describe some of the practical steps (interpolation, extrapolation, and error propagation) needed to implement the method. In § 6 we apply the method to a large $N$-body simulation, showing that it accurately recovers the correct 3D density and mass profiles statistically, and in § 7 we discuss why the method is insensitive to asphericity of halos and projected LSS along the line of sight. In § 8 we present our conclusions and outlook for the future. In future papers we will apply this method to measurement of the cluster-mass correlation function for optically selected clusters in the SDSS and in $N$-body simulations, derive scaling relations between optical richness and virial mass, and constrain the bias and the mass power spectrum by comparison with the cluster-cluster correlation function.

## 2. TWO-DIMENSIONAL DENSITY INVERSION

The primary measurement in cross-correlation lensing is of the average tangential distortion of background source galaxy images by foreground lens objects, typically galaxies or clusters. This distortion is measured in several radial annuli around the center of the lens and repeated for all lenses in the sample. For the purpose of this paper, we assume that we have redshift measurements for the lenses, as is the case with the SDSS and some other current and planned surveys. We also assume that we have an unbiased estimate of the lensing strength, $\langle \Sigma_{\rm crit}^{-1} \rangle$, which is to say we have a good understanding of the redshift distribution of the source galaxies, possibly from using photometric redshifts. Any bias in these redshift estimations, as well as shear calibration, will propagate into all else that follows.

For any set of lenses, we can define the lens-mass correlation function, $\xi_{lm}$, and write the average 3D mass density of the lens population as

$$\rho(r) = \bar{\rho}[1 + \xi_{lm}(r)] = \bar{\rho} + \Delta\rho(r), \tag{1}$$

where the cosmic mass density $\bar{\rho} = \Omega_m \rho_{\rm crit}$ and $\Delta\rho(r) \equiv \bar{\rho}\xi_{lm}(r)$. The lensing deflection and the shear are determined by the excess mass density projected along the line of sight, $\Sigma(R) = \int dz\, \Delta\rho(r)$, where we have dropped the constant term for reasons described below. Here and below, $R$ denotes the 2D separation from the lens center in the lens plane, $r$ denotes the 3D radial separation from the lens, and $z$ is the distance from the lens along the line of sight, so that $r^2 = R^2 + z^2$. The average tangential shear, $\gamma_T$, in a thin annulus of radius $R$ is given by (Miralda-Escude 1991)

$$\Sigma_{\rm crit}\gamma_T(R) = \bar{\Sigma}(< R) - \Sigma(R) \equiv \Delta\Sigma(R), \tag{2}$$

where $\bar{\Sigma}(<R)$ is the mean surface density interior to radius $R$, the critical surface density is

$$\Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{D_s}{D_l D_{ls}}, \tag{3}$$

and $D_s$, $D_l$, and $D_{ls}$ are the angular diameter distances to the source, to the lens, and from lens to source, respectively. Defining

the convergence $\kappa(R) \equiv \Sigma(R)/\Sigma_{\rm crit}$, the expression for the shear can be put in the dimensionless form

$$\gamma_T(R) = \bar{\kappa}(< R) - \kappa(R). \quad (4)$$

Equation (2) defines $\Delta\Sigma(R)$, the difference between the mean surface density inside a disk of radius $R$ and $\Sigma(R)$ (azimuthally averaged), which is the primary observable in cross-correlation lensing. This assumes that you have a good redshift estimate for the lens and have a good estimate for the source galaxy redshift distribution so that you can estimate $\Sigma_{\rm crit}$. From this definition, one can see that a uniform-density mass sheet of transverse size $>R$ produces no shear; this is referred to as the mass sheet degeneracy (Bradač et al. 2004) and is the reason why we dropped the constant term in the definition of $\Sigma(R)$ above. In the non-weak regime one can see that it is actually $\ln(1-\kappa)$ that is determined up to a constant by weak-lensing measurements. One can see this directly by noting that in equation (8), the integral is only determined up to a constant without explicitly imposing a boundary condition.

We have shown how to predict the lensing observable $\Delta\Sigma$ from the theoretical quantity of interest, the correlation function $\Delta\rho = \Omega_m \rho_{\rm crit} \xi_{lm}$. We would like to invert this relation, to infer $\Omega_m \xi_{lm}$ from measurements of $\Delta\Sigma$. One way to proceed would be to use a parametric model for $\xi_{lm}$, such as a power-law or NFW profile (Navarro et al. 1997). If a model profile is assumed, one can project the model to 2D, compute $\Delta\Sigma(R)$, fit the shear data, and determine the best-fit model parameters. This method has the advantages of simplicity and of small formal error bars on the derived model parameters. Of course, the problem with any model-dependent method is that an incorrect model will result in incorrect interpretation of the best-fit parameters. In addition, parametric methods force smoothness on the data and thus do not allow one to see any features in $\xi_{lm}$ or detect fine-scale deviations from the model. Since we would like to constrain the full density profile and not just determine its best-fit parameters for a particular model, in this paper we develop a nonparametric approach that inverts the process above and allows a direct estimate of the 3D density profile $\Delta\rho(R)$. We start by deriving a 2D reconstruction for the mean convergence profile, $\kappa(R)$.

Taking the derivative of equation (4) with respect to $R$, one can show that

$$-\kappa'(R) = \gamma_T'(R) + 2\gamma_T(R)/R, \quad (5)$$

where the prime denotes $d/dR$. This is just the azimuthally symmetric version of Kaiser's equation (Kaiser 1995)

$$\nabla\kappa = \begin{pmatrix} \gamma_{1,1} + \gamma_{2,1} \\ \gamma_{2,1} - \gamma_{1,2} \end{pmatrix}, \quad (6)$$

where $\gamma_i$ denotes the two-component shear field and $\gamma_{i,j} = \partial\gamma_i/\partial\theta_j$ denotes the derivative in the lens plane. Equation (6) is the starting point for most finite-field nonlinear inversion methods for weak lensing (Kaiser 1995; Lombardi & Bertin 1999; Seitz & Schneider 2001). Note that these equations relate the shear to the density in a *local* way, in contrast to equation (4) for the tangential shear, which is nonlocal. As shown below, both the local and nonlocal aspects of these equations are useful when inferring the mass profile.

Schneider & Seitz (1995) pointed out that, except in the limit of very weak fields, $\gamma,\ \kappa \ll 1$, the shear, $\gamma$, is not directly observable. The average shape of background objects is not determined by $\gamma$ or $\kappa$ but rather by the degenerate combination $g \equiv \gamma/(1-\kappa)$, which is sometimes referred to as the reduced shear. Furthermore, the average ellipticities of background galaxies are equal to the complex distortion, $\delta \equiv 2g/(1+|g|^2)$. This is true when ellipticity is defined as $e \equiv (1-r^2)/(1+r^2)$, with $r$ being the *axis ratio* or ratio of major to minor axes. Solving this quadratic equation for $g$ results in $g = [1 \pm (1-|\delta|^2)^{1/2}]/\delta^*$. Since there are two roots to this equation, one cannot tell from purely local measurements which root to choose. This reflects an inherent local degeneracy to any weak-lensing measurements. For weak fields, one should take the negative root, for which $\gamma \simeq g \simeq \delta/2$. The choice of sign is related to image parity; it changes when crossing critical curves in the lens plane, where the magnification diverges and arcs are present. In practice, it is only in the cores of massive clusters where one might make the wrong choice of parity, and this is unlikely to present a major problem for the statistical method of weak lensing we pursue here. These rare, extreme regions are not the places where weak lensing will be the most useful probe of the mass distribution.

The other continuous degeneracy is not as problematic: one can show that if $g$ is measured, one can invert for $\kappa$ up to the mass sheet degeneracy (Kaiser 1995). For our azimuthally symmetric case, we can simply substitute $g(1-\kappa)$ for $\gamma$ in equation (4), resulting in

$$-\frac{\kappa'}{1-\kappa} = \frac{d}{dR}\ln(1-\kappa) = \frac{g' + 2g/R}{1-g} \equiv G(R), \quad (7)$$

so that $\ln(1-\kappa)$ is determined up to a constant. Imposing the boundary condition $\kappa(\infty) = 0$, one can integrate this ordinary differential equation:

$$\kappa(R) = 1 - \exp\left[-\int_R^\infty dp\, G(p)\right], \quad (8)$$

which in the weak-field limit becomes

$$\kappa(R) = \int_R^\infty dp\, \left[\gamma_T(p)' + 2\gamma_T(p)/p\right]. \quad (9)$$

In § 5 we discuss the truncation of integrals of this kind to a finite region. One can see from these two equations that in fact $\ln(1-\kappa)$ is determined only up to a constant, which in the weak approximation becomes the statement that $\kappa$ is determined only up to a constant.

Since equation (9) is linear in $\gamma_T$, one could simply multiply it by $\Sigma_{\rm crit}$ to obtain an expression for $\Sigma(R)$. This is suitable for an analysis such as that in Sheldon et al. (2004) where $\Delta\Sigma \equiv \Sigma_{\rm crit}\gamma$ is averaged. Beyond the weak-field regime, the situation is more complicated because of the $1-g$ term in the denominator of $G(R)$; in this case, an average of the lensing distortion does not simply yield an average of $\Sigma(R)$. Nonetheless, as noted above, this should only be a problem for small scales and high densities, i.e., for the cores of very massive clusters, and in that regime one could apply a correction or an iterative process to determine $\Sigma(R)$. Moreover, there is only a small spatial range where this correction is important and where one can still safely assume that $\gamma = g \simeq \delta/2$ is the correct root. The method of averaging shear profiles may simply not be the most prudent method for probing the non-weak regime; strong lensing can likely tell us more about the density profiles of massive clusters at such small scales. Alternatively, other methods, such as the maximum likelihood approach of Schneider & Rix (1997), might be able to better handle this regime, although at the cost of introducing explicit model dependence.

## 3. INVERTING TO THREE-DIMENSIONAL DENSITY PROFILES

We can rewrite equation (7) (after multiplying by $\Sigma_{\rm crit}$) as

$$-\Sigma'(R) = \Sigma_{\rm crit}(1-\kappa)G(R) = \Sigma_{\rm crit}G(R)\exp\left[-\int_R^\infty dp\, G(p)\right], \quad (10)$$

or in terms of $\Delta\Sigma$ (multiplying eq. [5] by $\Sigma_{\rm crit}$),

$$-\Sigma'(R) = \Delta\Sigma'(R) + \frac{2}{R}\Delta\Sigma(R). \quad (11)$$

An equation for $\Sigma'(R)$ is useful because of the existence of the Von Zeipel (or sometimes called Abel) inversion formula for the 3D density profile:

$$\Delta\rho(r) = \frac{1}{\pi}\int_r^\infty dR\,\frac{-\Sigma'(R)}{\sqrt{R^2-r^2}}. \quad (12)$$

The inverse of this equation is just the usual projection

$$\Sigma(R) = 2\int_R^\infty dr\, r\frac{\Delta\rho(r)}{\sqrt{r^2-R^2}}, \quad (13)$$

where, as before, $\Delta\rho(r) \equiv \rho(r) - \bar{\rho}$ because we do not recover the average mass density of the universe: lensing is not sensitive to mass sheets. That is, we have arbitrarily chosen the boundary condition $\rho(\infty) = \bar{\rho}$ or $\Delta\rho(\infty) = 0$; for further discussion of the mass sheet degeneracy, see § 5.2.

Inversions of this type have a long history in astronomy. Von Zeipel (1908) derived this inversion formula and used it to determine the 3D density profile of globular clusters from imaging data. Von Zeipel's proof of this inversion formula rests on reducing it through substitutions to Abel's formula (Abel 1826), which in turn is usually proved through the use of the Laplace transform. Plummer (1911), also concerned with globular clusters, derived a similar formula that uses intensities in long parallel strips rather than circular annuli. Recently Kaastra (1989) has shown that Von Zeipel's formula can be derived from Plummer's formula much more easily without using Abel's formula. The interested reader should consult Kaastra (1989), as well as Bremer (1995) and references therein. These inversions have been used for astronomical deprojections in many different contexts, including globular clusters, galaxy clusters, elliptical galaxies, supernova remnants, planetary nebulae, and galaxy correlation functions.

Formally this inversion formula assumes spherical symmetry. When used on the stacked shear profile of a large sample of clusters, spherical symmetry should be a valid approximation. This simply follows from the isotropy of the universe and the linearity of the inversion. This is not to say that the clusters individually must be spherically symmetric or have common spherically averaged radial profiles. They may individually be elliptical or irregularly shaped; however, the resultant average of a large number of independent clusters will approach spherical symmetry. This inverted radial profile will be the *sample average* of the *spherically averaged* individual profiles. This method of averaging is blind to any incoherent differences between the individual clusters. It would be best to split up the cluster sample by various measured quantities (optical or X-ray luminosity, richness, etc.) into various subsamples. If these measured quantities correlate strongly with mass and density profile shape, then one might hope that within the various subsamples, the profiles are sufficiently common that the inverted profile can be simply interpreted as being representative of the sample as a whole. This of course never guarantees that the individual profiles all resemble the average. Cosmological *N*-body simulations indicate that dark matter halos of a given mass seem to have near universal profiles (Navarro et al. 1997), and so this simple interpretation should be quite reasonable. Detecting profile differences within a subsample is an interesting problem that we do not address in this paper.

Since we are computing an average mass density of many lenses, it can be written as a correlation function, $\Delta\rho(r) = \bar{\rho}\xi_{lm}(r)$. Equation (12) combined with equation (11) results in the equation used in Sheldon et al. (2004) to measure the galaxy-mass correlation function:

$$\Delta\rho(r) = \Omega_m\rho_{\rm crit}\xi_{lm} = \frac{1}{\pi}\int_r^\infty dR\,\frac{\Delta\Sigma'(R) + 2\Delta\Sigma(R)/R}{\sqrt{R^2-r^2}}. \quad (14)$$

Again, one must truncate this integral since one only has measurements out to some finite projected separation $R_{\rm max}$. We discuss this truncation in § 5.

## 4. FROM DENSITY TO MASS PROFILES

Now that we can reconstruct mean 3D density profiles, it would seem straightforward to calculate the 3D mass profiles, $M(r) = 4\pi\int_0^r dy\, y^2\Delta\rho(y)$. [This mass, the second moment of the correlation function, is sometimes referred to as $J_3(r)$ in the LSS literature, although not usually with cross-correlation functions in mind.] However, the density profile cannot simply be integrated directly, since one only has reliable density information from shear measurements down to some minimum scale, $R_{\rm min}$. One could extrapolate the derived density profile to smaller scales, but that would introduce model-dependent assumptions. As we show, one can do better: the mass inside $R_{\rm min}$ is in fact constrained by the lensing data. This is because equations (2) and (4) for the tangential shear are nonlocal. To illustrate the point, consider that a point mass creates a shear, $\gamma(r) \sim r^{-2}$. However, substituting this power-law expression into equation (5) results in a complete cancellation. Thus, adding a point mass does not alter the derived 2D and 3D densities, as should be expected. The mass inside $R_{\rm min}$ is in fact the piece of information that we lost when taking the derivative of equation (4) to get the local equation (5). Hence, any mass estimator should use both the local and nonlocal equations.

Throughout this calculation we assume that the mean density of the universe has been subtracted out, as above. This is just a statement of boundary conditions, which does not affect the final 3D profiles, and allows us to ignore the effects of the lensing kernel $\Sigma_{\rm crit}$. We begin with equation (2),

$$\Delta\Sigma(R) = \bar{\Sigma}(<R) - \Sigma(R) \quad (15)$$

$$= \frac{m_{\rm cyl}(R)}{\pi R^2} - \Sigma(R), \quad (16)$$

which defines the mass, $m_{\rm cyl}(R)$, inside a cylinder of radius $R$ oriented along the line of sight:

$$m_{\rm cyl}(R) = \pi R^2[\Delta\Sigma(R) + \Sigma(R)]. \quad (17)$$

Using equation (13) to replace $\Sigma(R)$, we have

$$m_{\rm cyl}(R) = \pi R^2\left[\Delta\Sigma(R) + 2\int_R^\infty dr\, r\frac{\Delta\rho(r)}{\sqrt{r^2-R^2}}\right]. \quad (18)$$

Now we want to relate this cylindrical mass, $m_{\rm cyl}(R)$, to the 3D spherical mass $M(R)$. It is a slight abuse of notation to write

$M(R)$, since $R$ is the 2D radius in the plane of the sky, but this avoids the bulkiness of the more precise notation, $M(r = R)$. Again, assuming spherical symmetry, the 3D mass can be written as the total mass inside the cylinder minus the mass inside the cylinder but outside the sphere. Using cylindrical polar coordinates $x$, $z$, $\phi$ and spherical radius $r = (x^2 + z^2)^{1/2}$, with the $z$-axis being the axis of projection, we have

$$M(R) = m_{\rm cyl}(R) - 2\pi \int_0^R dx\, x 2 \int_{\sqrt{R^2-x^2}}^{\infty} dz\, \Delta\rho(r) \tag{19}$$

$$= m_{\rm cyl}(R) - 4\pi \int_R^{\infty} dr\, r \Delta\rho(r) \left(r - \sqrt{r^2 - R^2}\right). \tag{20}$$

Substituting equation (18) for $m_{\rm cyl}$ into equation (20) gives an expression for $M(R)$ in terms of $\Delta\Sigma$ and $\Delta\rho$. Since $\Delta\rho$ is a functional of $\Delta\Sigma$, this is a mapping: $\Delta\Sigma \longrightarrow M$. We define this particular form for $M(R)$ as $M_{\rm out}(R)$, since it only uses data outside of $R$:

$$M_{\rm out}(R) = \pi R^2 \Delta\Sigma(R) + 2\pi \int_R^{\infty} dr\, r \Delta\rho(r) \times \left[\frac{R^2}{\sqrt{r^2 - R^2}} - 2\left(r - \sqrt{r^2 - R^2}\right)\right]. \tag{21}$$

Since we assumed above that the average density is zero, equation (21) actually gives only the mass overdensity, not the total mass; to properly correct for this, a term $4/3\pi R^3 \bar{\rho}$ should be added, but this formula is typically used on scales where this extra contribution is negligible.

We now derive another formula for the mass profile. As we mentioned, the simplest idea is just to integrate the density. We call this particular form $M_{\rm in}(R)$, which is given by

$$M_{\rm in}(R) = 4\pi \int_0^R dr\, r^2 \Delta\rho(r) \tag{22}$$

$$= M(R_{\rm min}) + 4\pi \int_{R_{\rm min}}^R dr\, r^2 \Delta\rho(r). \tag{23}$$

Replacing the first term with equation (21) evaluated for $M_{\rm out}(R_{\rm min})$ gives

$$M_{\rm in}(R) = \pi R_{\rm min}^2 \Delta\Sigma(R_{\rm min}) + 2\pi \int_{R_{\rm min}}^R dr\, r \Delta\rho(r) \frac{2r^2 - R_{\rm min}^2}{\sqrt{r^2 - R_{\rm min}^2}} + E(R, R_{\rm min}), \tag{24}$$

where the last term is

$$E(R, R_{\rm min}) = 2\pi \int_R^{\infty} dr\, r \Delta\rho(r) \times \left[\frac{R_{\rm min}^2}{\sqrt{r^2 - R_{\rm min}^2}} - 2\left(r - \sqrt{r^2 - R_{\rm min}^2}\right)\right]. \tag{25}$$

For $R \gg R_{\rm min}$ the term $E(R, R_{\rm min})$ is very small compared to the rest of the terms. Note that although we have referred to this expression as $M_{\rm in}$, the last term does depend on $\Delta\Sigma$ outside of $R$. There are thus two seemingly different expressions for the 3D mass profile. In fact, one could replace $R_{\rm min}$ in the above expression with any radius between $R_{\rm min}$ and $R$: all of these expressions are equivalent, and they do not add any extra information by combining them. They simply add up the information from different radii in a different order, much like changing the order of integration on a 2D integral.

For clusters, one can measure a mean virial mass from these mass profiles quite easily. The virial mass is defined as the mass within the virial radius, $r_{\rm vir}$, where the virial radius is the radius at which the average enclosed density is $\Delta_{\rm vir}$ times either the average or critical density of the universe. Often, $\Delta_{\rm vir} = 200$ (and critical) is used and this virial radius is often denoted $r_{200}$. One must then interpolate between points to find the radius at which $M(r_{200}) = 4\pi/3 r_{200}^3 200 \rho_{\rm crit}$. These nonparametric virial masses should prove very useful in comparing data to $N$-body simulations, which make frequent use of the virial mass.

As mentioned above, the mass profiles derived from $\Delta\Sigma$ actually contain more information than the density profiles. The mass profile at $R_{\rm min}$ tells us about the average density at $R < R_{\rm min}$, so it can be used to probe these smaller scales. This might be useful in answering the question of whether halos have cusps or cores at small radii. It is equally true that $\Delta\Sigma$ contains all of the information. However, projection and nonlocality make it somewhat harder to use, except through parametric modeling. If one just wants to fit a model to the data, one might as well use $\Delta\Sigma$. For visualizing the actual shape of the density profiles without a model and obtaining virial mass estimates, these inversions are most useful.

Finally, we note that there has been some work on deriving aperture mass estimates in 2D from the shear profile (Schneider 1996; Schneider & Bartelmann 1997). The aperture mass $m = \int d^2x\, \kappa(x) w(|x|)$, where $w$ is an azimuthally symmetric weight function. Schneider (1996) suggests choosing a normalized weight function satisfying $\int d^2x\, w(|x|) = 0$, i.e., a compensated filter, so that the mass estimate is not affected by the mass sheet degeneracy. Obviously this would restrict the choice of weight functions so that, for example, the top-hat weight function cannot be used. We have implicitly restricted ourselves above to top-hat apertures and derived a 3D mass profile from the lensing measurements. We discuss the mass sheet degeneracy by showing (in § 5.2) that, for stacked samples, this unknown constant can be predicted fairly well and only creates real uncertainty at the very largest scales.

## 5. PRACTICAL APPLICATION

With expressions in hand for inverting 3D density and mass profiles from the lensing data, $\Delta\Sigma$, we turn to a discussion of the application of these formulae to real data. The two main issues to consider are interpolation to evaluate the various integrals and handling the parts of the integrals that extend beyond the scales for which we have measurements. Improper handling of the binning and interpolation issues typically results in errors at the few percent level for all radii. The issue of extrapolating the integrals to infinity usually only affects the recovered profiles at the largest measured scales.

### 5.1. *Binning and Interpolation*

Typically, one will construct a $\Delta\Sigma$ profile by measuring the average weighted lensing distortion in a set of radial bins. For example, Sheldon et al. (2004) measured $\Delta\Sigma(R)$ in 18 logarithmically spaced radial bins from 0.02 to 11 $h^{-1}$ Mpc, giving a bin width $R_{i+1}/R_i = 1.41$. One would like to assign one effective radius to each bin rather than just taking the midpoint of $R_{i+1}$ and $R_i$. The radius assigned should be that radius at which we believe the true $\Delta\Sigma(R)$ is equal to our measured average, which can be calculated by considering the local logarithmic slope. For the

Sheldon et al. (2004) binning, this results in a shift of order 5% from the midpoint.

After selecting the proper effective radii, $R_i$, corresponding to each $\Delta\Sigma(R_i)$ bin, one can apply the inversion formulae of § 4. This requires a scheme for interpolating $\Delta\Sigma$ between the bins and for interpolating $\Delta\rho$ between the points to infer the mass. Saunders et al. (1992) applied Von Zeipel's inversion formula, essentially our equation (12), to the projected galaxy autocorrelation function, $w_p(R)$, to obtain the 3D autocorrelation function, $\xi(r)$. They used a piecewise linear model for $w_p(R)$ and hence a piecewise constant model for $w'(R)$, which enters into the Von Zeipel integral. This zeroth-order interpolation allows calculation of the integral over each bin analytically, which results in simply a weighted sum over the data points (details can be found in Saunders et al. 1992). However, the data for both $w_p$ and $\Delta\Sigma$ look roughly like power laws, $\Delta\Sigma \sim R^{-0.8}$, so this simple interpolation scheme biases the inversion low by about 10% for Sheldon's bin size. If one takes this scheme to first order by representing the derivative, $\Delta\Sigma'(R)$, as piecewise linear rather than constant over a bin, this biases the result by about the same amount, but high instead of low. One can in fact do $N$-point Lagrangian interpolation to any order and still evaluate the integrals analytically over each bin. Such methods are completely linear in the data, which is another advantage. In principle, higher order interpolation will reduce this bias, but only at the expense of correlating many neighboring points. We have found in practice that the best method is to use a power-law interpolation, equivalent to linear interpolation of the log-log values. Any two positive data points define a unique power law between them. The integrals for each bin need to be evaluated numerically, so this is slower computationally, and care needs to be taken to handle negative data points that may be present for noisy data. We therefore do not take logarithms directly but rather model the function as a power law plus constant, where the constant is only used for negative values. Data for which there are negative values are generally very noisy; for that case, interpolation is never the dominant source of errors. Tests of the inversion show that for power-law density profiles, this interpolation scheme is exact (unsurprisingly); for NFW profiles and halo model type profiles, the power-law interpolation is good to a fraction of a percent for Sheldon et al. (2004) sized bins (i.e., $R_{i+1}/R_i = 1.41$) and very robust for noisy data.

### 5.2. *Endpoint Corrections*

The inversion integrals in §§ 3 and 4 should in principle be evaluated from the minimum radial bin, $R_{\rm min}$, to $\infty$. In practice, one only has shear data covering a range from $R_{\rm min}$ to $R_{\rm max}$. We refer to the parts of the integrals from $R_{\rm max}$ to $\infty$ as endpoint corrections; they must be added to the parts of the integrals that we can actually perform by interpolating the data. Consider equation (12) as an example; we can rewrite it as

$$\Delta\rho(r) = \frac{1}{\pi}\int_r^{R_{\rm max}} dR\,\frac{-\Sigma'(R)}{\sqrt{R^2-r^2}} + \frac{1}{\pi}\int_{R_{\rm max}}^{\infty} dR\,\frac{-\Sigma'(R)}{\sqrt{R^2-r^2}} \equiv T_D + T_C, \tag{26}$$

where $T_D$ and $T_C$ stand for data and correction, respectively. To estimate the magnitude of each term, consider an isothermal sphere, $\Delta\Sigma \sim R^{-1}$; in this case, the integrals can be computed analytically, leading to

$$\frac{T_C}{T_D + T_C} = 1 - \sqrt{1 - \left(\frac{r}{R_{\rm max}}\right)^2}. \tag{27}$$

For the outermost data point, $r = R_{\rm max}$, the correction term gives the whole result, and one cannot learn anything about the density at this point. However, for smaller scales, $r \ll R_{\rm max}$, the ratio $T_C/(T_D + T_C) \simeq (r/R_{\rm max})^2/2$. For example, for Sheldon et al. (2004) sized bins, this implies that the correction is only about 14% at the value of $R_i = R_{16}$, two radial bins inward of $R_{\rm max} = R_{18}$. The correction terms at the next bins inward (proceeding from larger to smaller scale) are 7%, 3%, and 2%, while the remaining inner 12 bins are affected by less than 1%. In fact, it is only the error in estimating this correction term that concerns us, so we expect that even at the second bin inward of $R_{\rm max}$, one could estimate $\Delta\rho$ to an accuracy of a few percent. Around the virial radius of clusters, $R \sim 1\ h^{-1}$ Mpc, the correction term itself is only half a percent of the total for $R_{\rm max} = 10\ h^{-1}$ Mpc.

Beyond $10\ h^{-1}$ Mpc, the halo density profile is expected to drop off more steeply than isothermal, so the correction factor in this case should be less than the estimate in equation (27). At these large scales, it is probably a good assumption to use the shape of the linear mass correlation function (i.e., the Fourier transform of the linear power spectrum), which is well understood theoretically for a given cosmology, as the model with which to compute the endpoint correction. This is the expected large-scale behavior from the halo model. We describe this in detail in § 8. This requirement of providing endpoint corrections is the one part of this method that requires a model. However, this model dependence only affects the results at the largest scales for any reasonable extrapolation, guided by theoretical expectation and constrained by the large-scale galaxy autocorrelation function, which is now measured to $175\ h^{-1}$ Mpc (Eisenstein et al. 2005). It is left up to the experimenter to decide which points to omit and how to take the error in the correction into account.

In this context, for lensing data that extend to scales of order a few megaparsecs, the endpoint correction can be considered a manifestation of the mass sheet degeneracy. In the current paradigm of structure formation, the universe is not filled with uniform-density mass sheets, but it does contain large-scale density perturbations that are correlated with dark matter halos. For a cluster positioned near the peak of a large-scale overdensity with scale $\gg R_{\rm max}$, the latter can be thought of to first approximation as a mass sheet for the cluster shear measurement. Using the theoretical mass correlation function on large scales to estimate the endpoint correction, as described above, therefore provides an estimate of the effective mass sheet degeneracy for this method. As we have argued above and show below in an $N$-body simulation, the effect of this degeneracy on the estimate of cluster virial masses is below 1%.

We have argued that this mass sheet uncertainty is only a problem at the largest scales, relative to the size of the survey. This method is ideally suited to a *panoramic* survey such as the SDSS where many decades in scale can be measured, in which case the uncertainty in the recovered density profiles in the last two logarithmic bins (out of 18) is no great loss. This method is less well suited to a *pencil beam* survey with a small field of view around a given set of clusters. This is because one might only be able to make measurements over a single decade in scale and with perhaps four or five logarithmically spaced bins, and in this case, throwing away the two endpoint bins is a large fraction of your data. Other methods such as profile fitting do not solve this problem, as they still have to make an assumption about how the density extrapolates to scales beyond which measurements exist.

### 5.3. *Propagation of Errors*

In performing these inversions, it is important to correctly propagate the errors from $\Delta\Sigma(R)$ to $\Delta\rho(r)$ and $M(r)$. In particular, the

covariance matrix between radial bins $R_i$ needs to be calculated. We assume that we have already calculated the covariance matrix for $\Delta\Sigma_i$, which we denote by $C_{ij}^{\Delta}$. Usually this is just diagonal from shot noise (i.e., lensing shape noise), but at larger scales there can be off-diagonal terms from sample variance and systematics such as imperfect point-spread function (PSF) anisotropy correction. These off-diagonal terms can be estimated with jackknife resampling (Sheldon et al. 2004). We denote the covariance matrix for $\Delta\rho_i$ and $M_i$ as $C_{ij}^{\rho}$ and $C_{ij}^{M}$.

As we have mentioned, with the simplest polynomial interpolations, the mapping $\Delta\Sigma(R) \longrightarrow \Delta\rho(r)$ is linear, so there exists a matrix $\mathbf{A}$ such that $\Delta\rho_i = A_{ij}\Delta\Sigma_j$; likewise, since $M$ is a linear combination of $\Delta\Sigma$ and $\Delta\rho$, and $\Delta\rho$ is computed linearly from $\Delta\Sigma$, there exists a matrix $\mathbf{B}$ such that $M_i = B_{ij}\Delta\Sigma_j$. For the polynomial interpolation schemes, these matrix elements can be computed analytically, without requiring numerical integration. For the power-law interpolation that we recommend using, the mapping is not *exactly* linear, but it is well approximated by its first-order Taylor expansion,

$$A_{ij} = \frac{\partial\Delta\rho_i}{\partial\Delta\Sigma_j}, \quad B_{ij} = \frac{\partial M_i}{\partial\Delta\Sigma_j}. \tag{28}$$

Although these derivatives can be calculated analytically in principle, in practice we estimate them numerically with a finite difference method.

Alternatively, one could use the linear interpolations just for the covariance matrices, where the matrix elements can be computed analytically. This would be the fastest method computationally. Once the linearization has been computed (by any method), the covariance matrices can be propagated simply with matrix multiplication, $\mathbf{C}^{\rho} = \mathbf{A}\mathbf{C}^{\Delta}\mathbf{A}^{T}$ and $\mathbf{C}^{M} = \mathbf{B}\mathbf{C}^{\Delta}\mathbf{B}^{T}$. Generally, we find that the covariance matrix $\mathbf{C}^{\rho}$ is mostly diagonal if $\mathbf{C}^{\Delta}$ is diagonal. The covariance matrix $\mathbf{C}^{M}$, however, has large off-diagonal terms since $M(r)$ is a cumulative statistic and therefore neighboring bins are correlated.

## 6. TESTS OF THE METHOD: *N*-BODY SIMULATIONS

As a proof of principle, we have performed tests of these inversion methods on an *N*-body cold dark matter (CDM) simulation. In the simulation, we can measure the 3D density and mass profiles directly and check that the inversion of projected quantities correctly recovers the true values.

The simulation we use has $512^3$ particles in a periodic cube of length 300 $h^{-1}$ Mpc. The simulation is evolved from $z = 60$ to 0 using a TreePM code (White 2002, 2003);[10] we use only the $z = 0$ output. The cosmological parameters used are $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, $h = 0.7$, $n = 1$, $\Omega_b h^2 = 0.02$, and $\sigma_8 = 1$. The simulation has an effective Plummer force softening scale of 20 $h^{-1}$ kpc, which is fixed in comoving coordinates. The mass of each dark matter particle is $1.7 \times 10^{10}\ h^{-1}\ M_\odot$. Dark matter halos are identified using a friends-of-friends (FoF) algorithm (Davis et al. 1985) with a linking length of 0.2 in units of the mean interparticle separation. Specific details of the simulation such as resolution, cosmology, and halo finding are not crucially important, since we are only interested in whether the inversion methods recover the 3D quantities. The simulation box was chosen to have high enough resolution to resolve the inner regions of clusters measurable by SDSS and is large enough to have relatively low cosmic variance; nevertheless, it is about 5 times smaller in volume than the SDSS cluster sample currently in preparation.

[10] The simulation is publicly available at http://mwhite.berkeley.edu/Sim1/.

We select all halos of mass $M_{\rm vir} > 10^{14}\ h^{-1}$ Mpc, measure $\rho(r)$ and $M(r)$, and average these 3D quantities for all such massive halos. For $\rho(r)$, we correct for the effects of binning in a similar way to the methods described in § 5.1. Next, we project the box separately along each of its three axes ($x$, $y$, $z$). For each of these projections, we measure the average $\Sigma(R)$ and $\Delta\Sigma(R)$ from equation (2). We do not perform ray tracing to determine the real shear, so we are only testing the inversion method, assuming that one can measure $\Delta\Sigma$ through measurement of background galaxy shear (see Sheldon et al. 2004). In particular, we are not testing the shear nonlinearities, the details of galaxy shape measurement, photometric redshift estimation and calibration, or any kind of cluster finding. In a real survey, one would find clusters with techniques such as optical red sequence methods or by observing X-ray fluxes or SZ decrements instead of having the luxury of selecting halos directly by mass. These details are important for any analysis of real data, but they are beyond the scope of this paper (and will be discussed further in future papers in this series).

By studying the differences in these projected 2D quantities for the three different projections of the simulation box, we can crudely measure a kind of sample variance that we call anisotropic sample variance (ASV). If the halo-mass correlation function were exactly isotropic, all three projections would agree precisely. Since this is only true in the infinite volume limit, there will be differences between the three projections of the simulation, as well as small departures from isotropy in any real measurement. On small scales, these differences arise from the asphericity of halos: if halo orientations are not strongly correlated, the residual ellipticity will result in a small random error that will decrease as $\sim 1/(N_{\rm halos})^{1/2}$. More explicitly, we can define the halo ellipticity along the $z$-direction as $e_z = (3\sigma_z^2 - \sigma^2)/\sigma^2$, where $\sigma_z^2$ is the second moment of the cluster density distribution along the projected direction and $\sigma^2 = \sigma_x^2 + \sigma_y^2 + \sigma_z^2$ defines the characteristic cluster size.

The ensemble average of $e_z$ is zero by symmetry, but the rms for a finite sample will be zero within $\pm(\langle e_z^2\rangle/N_{\rm halos})^{1/2}$. One can show that this rms halo ellipticity induces a multiplicative correction factor in the projected density and all derived quantities ($\rho$ and $M$) of $1 \pm (\langle e_z^2\rangle/N_{\rm halos})^{1/2}$; this factor will vary randomly from realization to realization (or from projection to projection within a single *N*-body realization). The rms halo ellipticity $(\langle e_z^2\rangle)^{1/2}$ appears to be about 0.5 for the massive halos in the *N*-body simulation described above. Since large survey samples will have very many clusters, this correction will usually be a very small effect. For example, in the sample used in this study, there are 1226 halos above the mass threshold in the simulation, resulting in an expected multiplicative variance of $\pm 1.4\%$ in the inferred profiles. This estimate agrees very well with the observed scatter between the three different projections on small and intermediate scales, as shown below in Figures 2 and 3. The SDSS cluster samples are typically larger than this, so the effect will be correspondingly reduced. Less massive halos (these are $M = 10$ $14\ h^{-1}\ M_\odot$), being more common, will have a larger $N_{\rm halos}$ and thus this will be reduced further.

On larger scales, where the halo-mass correlation function is dominated by contributions from mass elements in halos different from the lens halo (i.e., where, in halo model terms, the two-halo term dominates; Seljak 2000; Mandelbaum et al. 2005), the asymmetric sample variance arises predominantly from the shapes of larger scale structures such as filaments and superclusters. Since there are fewer structures contributing to this part of the ASV (fewer filaments than halos), and since these large structures are more asymmetric than the halos, the variance between projections is expected to be larger on large scales than small scales, in

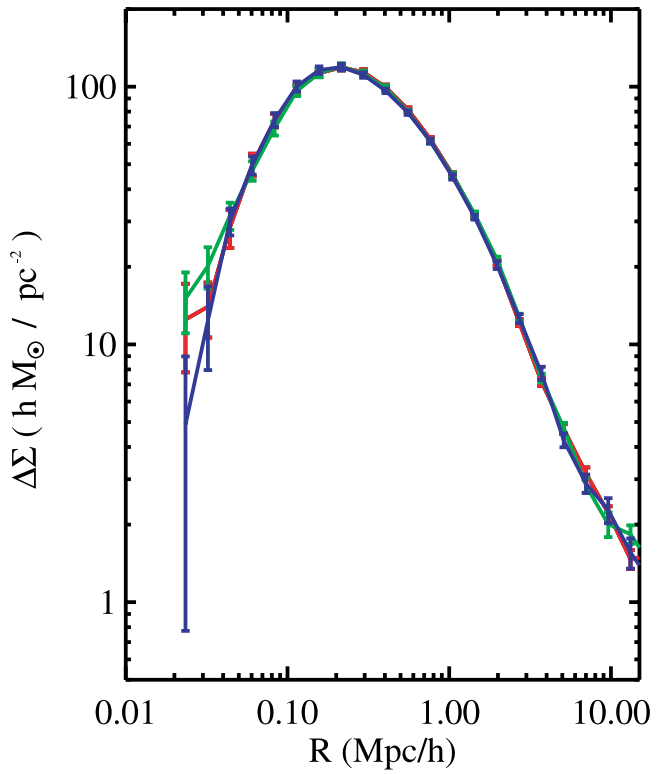


FIG. 1.—Measurements of $\Delta\Sigma(R)$ (proportional to the mean tangential shear) for halos with virial masses greater than $10^{14}\ h^{-1}\ M_{\odot}$ for three orthogonal projections of a ΛCDM *N*-body simulation. The turnover at around $R \sim 200\ h^{-1}$ kpc is due to the finite resolution of the simulation. The scatter between the lines at the smallest scale is likely due to low-number statistics since there are very few dark matter particles at these scales due to the resolution. At large radius, we attribute the differences to ASV caused by LSS along the line of sight.

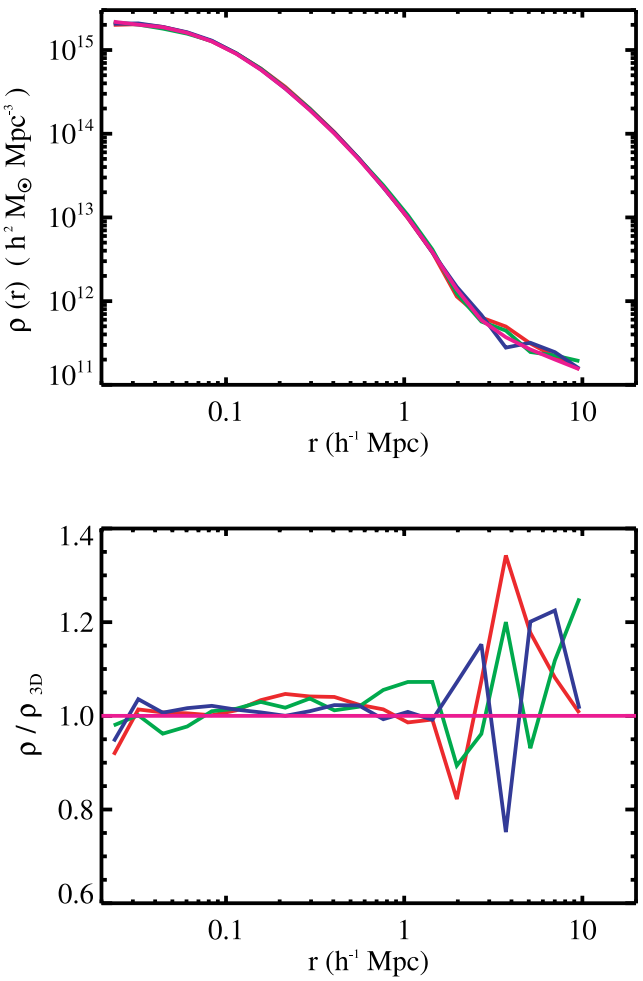


FIG. 2.—*Top*: Inverted 3D density profiles for the three projections of the simulation. The red, green, and blue lines indicate the inversions, while the magenta line denotes the true density profile measured by averaging the dark matter particle density in spherical shells. The inverted profiles are only distinguishable from each other and from the true profile at large scales. *Bottom*: Ratio of the inverted density to the true 3D density for the three simulation projections as a function of radius; lines correspond to the same projections as in the top panel. For each projection, the inverted and true density profiles agree to about 5% or better within the virial radius ($r_{200} = 0.585\ h^{-1}$ Mpc). On larger scales, there is significantly more scatter due to ASV in this relatively small simulation box.

qualitative agreement with Figure 2 below. This large-scale asymmetric sample variance is expected to scale as $1/\sqrt{V}$, where $V$ is the volume of the survey. When the shot noise from galaxy shapes (i.e., the shape noise) is sufficiently small, the large-scale errors will be dominated by the ASV. Judging by the volume of our simulation and the scatter that we see at scales of a few $h^{-1}$ Mpc, we expect this large-scale multiplicative error to scale roughly as $2\%(V/\mathrm{Gpc}^3)^{-0.5}$, although it may vary with scale $R$.

Figure 1 shows the $\Delta\Sigma$ results for each of the three projections of the simulation (the three different colors correspond to the three different projections). We note that these lines turn over below about 200 $h^{-1}$ kpc, an effect almost certainly due to the finite resolution of the simulation: this scale is about 10 times the Plummer force softening scale. The turnover in $\Delta\Sigma$ typically occurs at a few times the 3D correlation function core radius, and the 3D correlation function core radius typically occurs at a few times the Plummer force softening scale. By correlation function core radius we have in mind something like a power law with a softening, $r \rightarrow (r^2 + r_{\mathrm{core}}^2)^{1/2}$. Smaller volume, higher resolution simulations do not show this turnover in $\Delta\Sigma$. For a pure NFW profile, $\Delta\Sigma$ flattens at small scale but does not decline. As noted above, despite this resolution issue, we should be able to use these small scales to test our inversion methods since the 3D values are equally affected by resolution. In fact, this gives us an opportunity to test the robustness of the methods by including a pathological regime instead of a perfect power-law or NFW profile on all scales. The large errors (estimated by a jackknife technique) at small scales indicate that there are not many particles in these weak cusps. The larger deviations at the largest scales indicate the larger ASV, as expected from the arguments above.

We invert each of these $\Delta\Sigma$ profiles to obtain the estimated density $\Delta\rho(r)$ and $M(r)$ using the formulae and algorithms described in previous sections. The endpoint corrections are carried out assuming power-law extrapolations with logarithmic slopes that are not too different from the expected continuation of the linear theory correlation function. The last two bins from the inversions are not plotted, since they are highly affected by the endpoint corrections.

Figure 2 shows results of the inversion for the density profile $\rho(r)$. As in Figure 1, the red, green, and blue lines denote the inversions from the three different projections. The magenta line shows the true 3D density as measured in spherical shells. On this log-log plot one can hardly see any difference between the true and inverted profiles. The bottom panel of Figure 2 shows the ratio of each of these inverted density profiles to the true 3D density with a linear axis. On scales $r < 1\ h^{-1}$ Mpc, these ratios

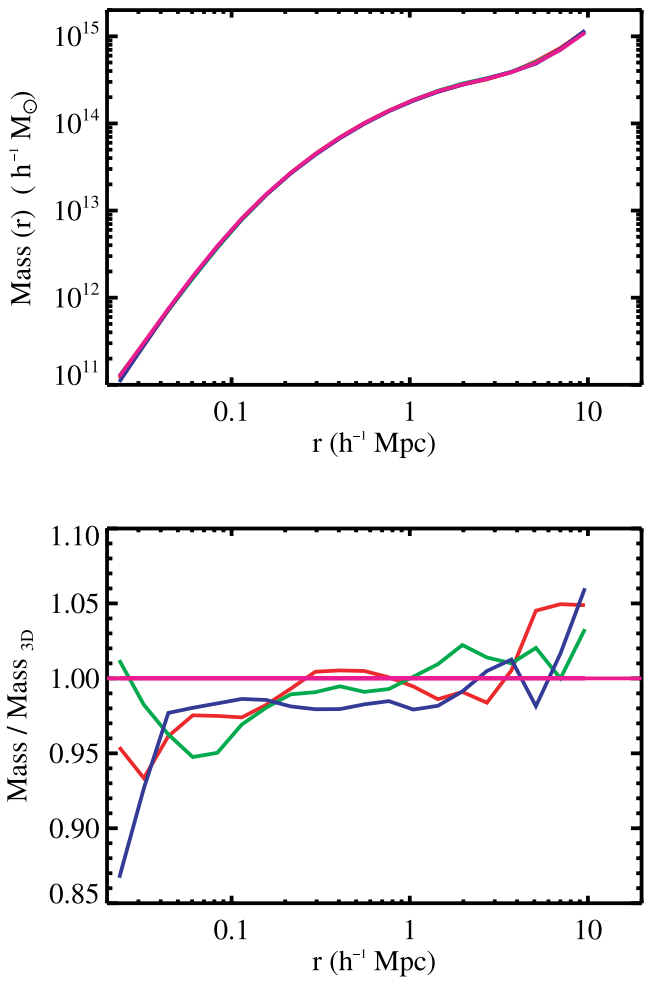


FIG. 3.—*Top*: Mass profiles inverted from the tangential shear for the three projections of the *N*-body simulation (*red, green, and blue lines*). The magenta line shows the true average mass profile measured in 3D for these halos. The differences are not evident on this scale. *Bottom*: Ratios of the three inverted mass profiles to the true 3D profile. The inverted masses are all correct to within 5% out to the largest scales and to within 2% near the virial radius. Because the points are correlated, the slight trend with radius that is apparent is not significant. The deviations at the smallest scale are most likely due to low-number statistics; the bins are logarithmically spaced, but the smoothed profile is noncuspy and so there are very few particles in these bins.

are all consistent with unity to within about 5%. On larger scales, the ASV becomes larger, and the three lines differ randomly from unity by about 20%. There do not appear to be any systematic biases or significant trends. Changing the endpoint correction moves these lines around by a few percent at the largest scales.

Figure 3 shows results of the inversions for the mass profile, $M(r)$. Again, the top panel shows all three inversions with the true $M(r)$ overplotted in magenta; the inversions and the true profile are indistinguishable on this plot. The bottom panel shows the ratios of inverted to true mass. Except for the smallest scales, where the scatter between the three inversions is a little larger, the mass ratios are consistent with unity to about 5%. Around the virial radius, $r \sim 1\ h^{-1}$ Mpc, which is the optimal location in terms of measurement signal-to-noise ratio, the inverted masses are correct to about 2%. There appears to exist a slight tilt to these ratio curves, but the points at different radii are strongly correlated [$M(r)$ being a cumulative quantity], so this trend is not significant. On the whole, we consider these tests a confirmation that inverting lensing measurements can correctly recover the density and mass profiles of galaxy and cluster halos.

## 7. EFFECTS OF HALO ASPHERICITY AND LARGE-SCALE STRUCTURE ALONG THE LINE OF SIGHT

In the case of weak-lensing mass reconstructions for individual clusters, one must consider the systematic effects of both cluster asphericity and contamination from structure along the line of sight, such as filaments or clusters at other distances seen in projection. Several studies (Cen 1997; Metzler et al. 1999, 2001; White et al. 2002) have concluded that LSS along the line of sight introduces significant scatter and bias in the resulting cluster mass estimates. Later studies (Hoekstra 2001, 2003; Clowe et al. 2004; de Putter & White 2005) agreed that LSS introduces scatter in the mass estimate, although Clowe et al. (2004) argued that the bias found in the previous studies was due to an oversimplification of the mass estimation. Dodelson (2004) suggested ways that the extra noise from LSS can be reduced, although not eliminated, through a particular type of filtering when making mass maps. Clowe et al. (2004), on the other hand, argued that cluster halo asphericity is a larger problem for determining cluster masses than LSS along the line of sight: if the major (minor) axis of the halo is aligned along the line of sight, the virial mass will be overestimated (underestimated). In general, one does not know the orientation of the halo in 3D, so this creates an uncertainty in the inferred mass at the ~30% level.

As demonstrated in § 6 and Figure 3, neither of these effects compromises the accuracy of our statistical mass inversion method, provided that the cluster sample is sufficiently large. By stacking many clusters to form a statistical sample, the effects of LSS uncorrelated with the lensing cluster must average out by statistical isotropy. On small scales, the mean orientation of aspherical halos for a large sample of lenses must be very close to random, provided that the method of *selecting* cluster lenses is not biased toward including those oriented, say, along the line of sight. These selection effects can be checked in *N*-body simulations that incorporate some prescription for assigning luminous galaxies to dark matter halos; as such, they are beyond the scope of this paper, but we plan to address them in a future publication. (Note that cluster samples selected by their shear signal may suffer from such a bias; see, e.g., White et al. 2002.) On intermediate scales, the orientation of nearby filamentary structure correlated with the cluster lens (as well as the angular distribution of nearby correlated halos) must also be random for a large statistical sample. Since such structure is correlated with the lens, it *does* contribute to the tangential shear on large scales; indeed, as we have argued above, it dominates the signal from clusters on scales above a few $h^{-1}$ Mpc and simply represents the large-scale cluster-mass correlation function. Again, as shown above, the effects of correlated structure along the line of sight appear to be negligible out to the virial radius of massive clusters.

The only residual problem for the statistical inversion method is ASV at large scales (discussed in § 6); this finite-volume effect creates a small scatter in the inferred mass but does not create a bias.

## 8. ESTIMATION OF SIGNAL-TO-NOISE RATIOS FOR REAL DATA

In this section we calculate the expected signal-to-noise ratios for $\Delta\Sigma(R)$ and derived quantities, $\Delta\rho(r)$ and $M(r)$, for a cluster survey resembling the SDSS. We argue that this method should be immediately applicable to SDSS data. Let us consider a sample of galaxy clusters of approximately $10^{14}\ h^{-1}\ M_{\odot}$. First, we calculate how many clusters are expected in a bin of mass close to this value. We do this using the Sheth & Tormen (1999) mass function and a flat ΛCDM cosmology with $\Omega_m = 0.3$, $\sigma_8 = 0.8$,

$h = 0.7$, and $n_s = 1.0$ with a matter transfer function from Eisenstein & Hu (1998). We assume a volume-limited sample of clusters in the redshift range $z = 0.05$–$0.2$ and an area coverage of 8000 deg$^2$, similar to the SDSS. This gives a fraction of the sky of 0.19 and an effective volume of 0.15 (Gpc $h^{-1}$)$^3$. We also assume that we have selected a fairly narrow mass bin: $9 \times 10^{13}\ h^{-1}\ M_\odot < M < 2 \times 10^{14}\ h^{-1}\ M_\odot$. We estimate that we will have approximately 800 clusters to use as lenses. This number is very sensitive to the assumed cosmology, as well as the effective width of the mass bin, which depends on the method of finding and selecting clusters. Any choice of binning will need to be done on some observable tracer such as X-ray luminosity, optical luminosity, galaxy richness or SZ decrement, etc. Due to the scatter between mass and the tracer, any hard cut on these observational quantities will necessarily result in a selection in mass that is not hard. However, we assume for the present purposes that one can select a sample of clusters with an effective range in mass as above.

We assume that these 800 clusters have masses of approximately $10^{14}\ h^{-1}\ M_\odot$ and have similar NFW profiles. The NFW density profile is parameterized by two numbers, a virial radius $r_{200}$ and a concentration parameter $c$, and is given by

$$\rho_{\rm NFW}(r,\, c,\, r_{200}) = \frac{\rho_{\rm crit}\delta_c}{r/r_s(1 + r/r_s)^2}, \tag{29}$$

with $r_s \equiv r_{200}/c$ and

$$\delta_c = \frac{(200/3)c^3}{\ln(1 + c) - c/(1 + c)}. \tag{30}$$

With this mass, they have a virial radius, $r_{200} = 0.72\ h^{-1}$ Mpc. We use an NFW concentration parameter of $c = 5.0$.

The 3D density profile will be the sum of the NFW profile and the large-scale or two-halo term

$$\rho(r) = \rho_{\rm NFW}(r,\, c,\, r_{200}) + \rho_{\rm crit} b(M)\Omega_m \xi(r), \tag{31}$$

where $\xi(r)$ is the linear correlation function and $b(M)$ is the linear bias for clusters of this mass. We calculate $\xi(r)$ by Fourier transforming the linear power spectrum using the same cosmological parameters above and normalizing it with the same $\sigma_8 = 0.8$. We use a linear bias of $b = 2.0$.

The shear, $\gamma$, is calculated from the 3D density by projecting to 2D and using the formulae described in § 2. This involves a factor of $\Sigma_{\rm crit}$, and for this we need to consider the redshift distributions involved. In reality, both the lenses and sources are at many different redshifts and an effective $\Sigma_{\rm crit}$ needs to be calculated by integrating over the redshift distributions. This is described in detail in Sheldon et al. (2004). For simplicity, we fix the lens redshift at $z_l = 0.15$ and a source redshift at $z_s = 0.35$.

We choose 22 logarithmically spaced radial bins between $r = 20\ h^{-1}$ kpc and $30\ h^{-1}$ Mpc. This range spans 11″ (smallest bin) to 4.5° (largest bin) at a lens redshift of $z = 0.15$. We use a source galaxy surface density of one galaxy per square arcminute, which is appropriate for the SDSS. Given the bin sizes and this source galaxy density, we can calculate the mean number of source galaxies in each radial bin for each cluster. Since we are stacking the signal over all clusters, we can multiply this number by the number of clusters to get the expected number of lens-source pairs, which determines the noise on the lensing signal. The noise on the shear in each radial bin is given by $\sigma_\gamma = 0.3/(N_{\rm pair})^{1/2}$. Multiplying this by $\Sigma_{\rm crit}$ gives the noise for $\Delta\Sigma(R)$.

The top left panel of Figure 4 shows the predicted $\Delta\Sigma(R)$ for this stacked sample of 800 clusters where we have added the appropriate amount of Gaussian noise using the above analysis. The red line is the noiseless contribution to $\Delta\Sigma$ from the NFW halo or the one-halo term. The blue line is the two-halo contribution, and the magenta line is the full noiseless $\Delta\Sigma(R)$, which is the sum of the one- and two-halo terms. The top right panel shows the signal-to-noise ratio of $\Delta\Sigma(R)$ as a function of scale. Interestingly, the signal-to-noise ratio increases with scale, dips slightly around the one-halo to two-halo transition, and then increases thereafter. The linear correlation function is not very steep, and for this reason the shear can be measured with increasing accuracy at very large scales if one has a big enough panoramic survey, simply due to the large number of foreground-background pairs. This, however, neglects other sources of error that most likely will dominate on large scales such as sample variance and systematic errors due to PSF correction, shear calibration, and photo-$z$ estimation. We do not attempt to model these sources of error in this paper but believe that they will begin to dominate around $20\ h^{-1}$ Mpc for surveys such as the SDSS.

The bottom left panel of Figure 4 shows the 3D density profile obtained by inversion of the noisy $\Delta\Sigma$ profile. The solid lines, as before, show the contributions from each of the one-halo (*red*) and two-halo (*blue*) terms with their sum (*magenta*). The error bars are calculated by propagating the errors of $\Delta\Sigma$ as discussed in § 5.3. The bottom right panel shows the inverted mass profile. The mass profile appears to have smaller errors only because the points are more correlated since mass is a cumulative quantity. Despite the many assumptions we have made, this analysis shows that a survey like the SDSS can produce excellent measurements of the average density and mass profiles of a stacked sample of clusters. In fact, this analysis can be performed in many different bins of any observation tracer and measure the density and mass profiles for different mass ranges. Although the more massive clusters have a higher shear signal, they are more rare and thus have a higher noise level. Because of this, the signal-to-noise ratio should be about the same over a fairly wide range in masses from $10^{13}$ to $10^{15}\ h^{-1}\ M_\odot$. Mass scales larger than this are usually strong-lensing clusters and do not need to be stacked. Masses smaller than this are merging into the galaxy-galaxy lensing regime, which has been measured quite accurately already by Sheldon et al. (2004) among others. This method of stacking the lensing signal of groups and clusters should provide a bridge between these two regimes that have both been studied quite extensively.

## 9. THE EFFECTS OF SHEAR NONLINEARITY ON REAL DATA

For the most part, in this paper, we have been working in the weak-lensing approximation where we have assumed that we can average the shapes of galaxies and measure the shear, $\gamma(R)$. However, as we pointed out, one really measures the distortion parameter $\delta(R)$ from which you can calculate $g = \{1 \pm [1 - \delta(R)^2]^{1/2}\}/\delta(R)$, which is related to the shear by $g(R) = \gamma(R)/[1 - \kappa(R)]$. If you are confident that you have not included data inside the critical curves, then you can choose the negative sign for $g$ and avoid or ignore that discreet degeneracy. We want to investigate how important this weak-lensing assumption is and see what we can do to mitigate the problems that will arise with real data. For this purpose let us assume that we have a cluster survey like the SDSS where the source galaxies are at $z_s = 0.35$ and lensing clusters are at $z_l = 0.15$. We assume NFW profiles with concentration, $c = 5$, and consider three mass ranges, $M = (10^{13},\ 10^{14},\ 10^{15})\ h^{-1}\ M_\odot$. For each mass range we calculate the $\Sigma(R)$ and $\Delta\Sigma(R)$ profiles and use the redshifts to calculate $\Sigma_{\rm crit}$.

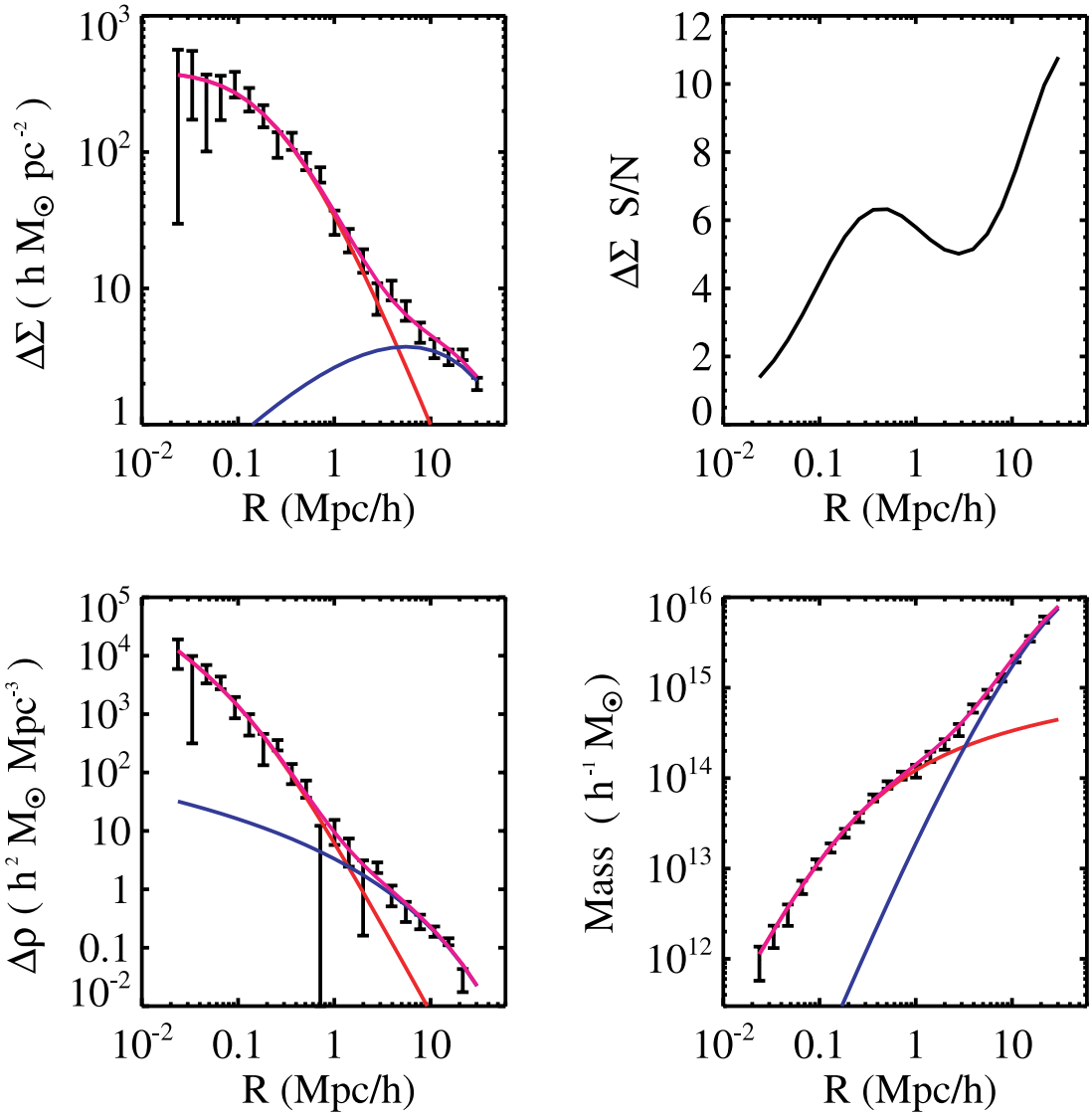


FIG. 4.—Predicted lensing signal and corresponding noise level around a sample of clusters in a narrow bin of mass $M \sim 10^{14}\ h^{-1}\ M_{\odot}$ from a survey similar to the SDSS. *Top left*: $\Delta\Sigma$ profile with added noise of the expected level. The magenta line shows the noiseless signal, with the red line indicating the contribution from the one-halo term or NFW profile. The blue line shows the contribution from the two-halo term. *Top right*: Signal-to-noise ratio for $\Delta\Sigma$ as a function of scale. *Bottom left*: Inverted density profile $\Delta\rho$. *Bottom right*: Inverted mass profile. The error bars for both $\Delta\rho(r)$ and $M(r)$ are propagated from $\Delta\Sigma$. In all cases the solid lines show the contribution from each of the one- and two-halo terms.

From this we can compute the reduced shear $g(R)$ and the distortion $\delta(R)$. These are shown in Figure 5. The three rows correspond to the three mass ranges. In the left panels, we have plotted $\kappa(R)$ (*solid black line*) and $\gamma(R)$ (*solid red line*). The solid blue line is $\delta(R)/2$, and the dashed magenta line is $g(R)$. The first thing to note is that all of these are subcritical on these scales; $\kappa < 1$ although the largest mass cluster has $\kappa \simeq 0.6$ on the smallest scales. For these NFW profiles, the shear is less than $\kappa$. Contrast this to the isothermal sphere where $\kappa(R) = \gamma(R)$. For the smallest mass, all three of $\gamma$, $g$, and $\delta/2$ are about equal. For the other two masses, the shear $\gamma$ deviates from the other two curves. The right panels show this in more detail. The solid blues lines are $\delta/(2\gamma)$, and the dashed magenta lines are $g/\gamma$. So, for the smallest mass clusters, the measured distortion only differs from the shear by a few percent: 8% at $R = 10\ h^{-1}$ kpc, which is probably smaller than one would want to push this method due to blending issues and foreground cluster galaxy contamination, etc. Sheldon et al. (2004) make measurements down to 26 $h^{-1}$ kpc. For the other two mass ranges, however, the shear is significantly different from the distortion. At small scales (20 $h^{-1}$ kpc) they differ by 20% for $M = 10^{14}\ h^{-1}\ M_{\odot}$ and by a factor of 2 for $M = 10^{15}\ h^{-1}\ M_{\odot}$. Notice, however, that $g$ and $\delta$ differ from each other by smaller amounts. It can be shown readily that the difference between these two quantities is quadratic in the distortion whereas the difference between $\gamma$ and $g = \gamma/(1 - \kappa)$ is better described as linear since $\kappa$ and $\gamma$ are linearly related.

Despite the fact that the weak-lensing approximation $\delta \simeq 2\gamma$ may break down for massive clusters at smaller scales, the inversion methods can still be used. In §§ 2 and 3 we showed how to perform the inversions using the reduced shear, $g(R)$. We define $G(R) \equiv (g' + 2g/R)/(1 - g)$ (our eq. [7]) and can then use the unapproximated equation (10) rather than equation (11). Thus, as long as we are confident that we have not crossed over critical curves and can therefore choose the negative sign for $g$, we can still do the inversions exactly. The one complication is one of interpreting what we mean by the *average* profile. In the weak-lensing regime, the inversion equations are all completely linear in the lensing distortion so that the average of the distortion is just the

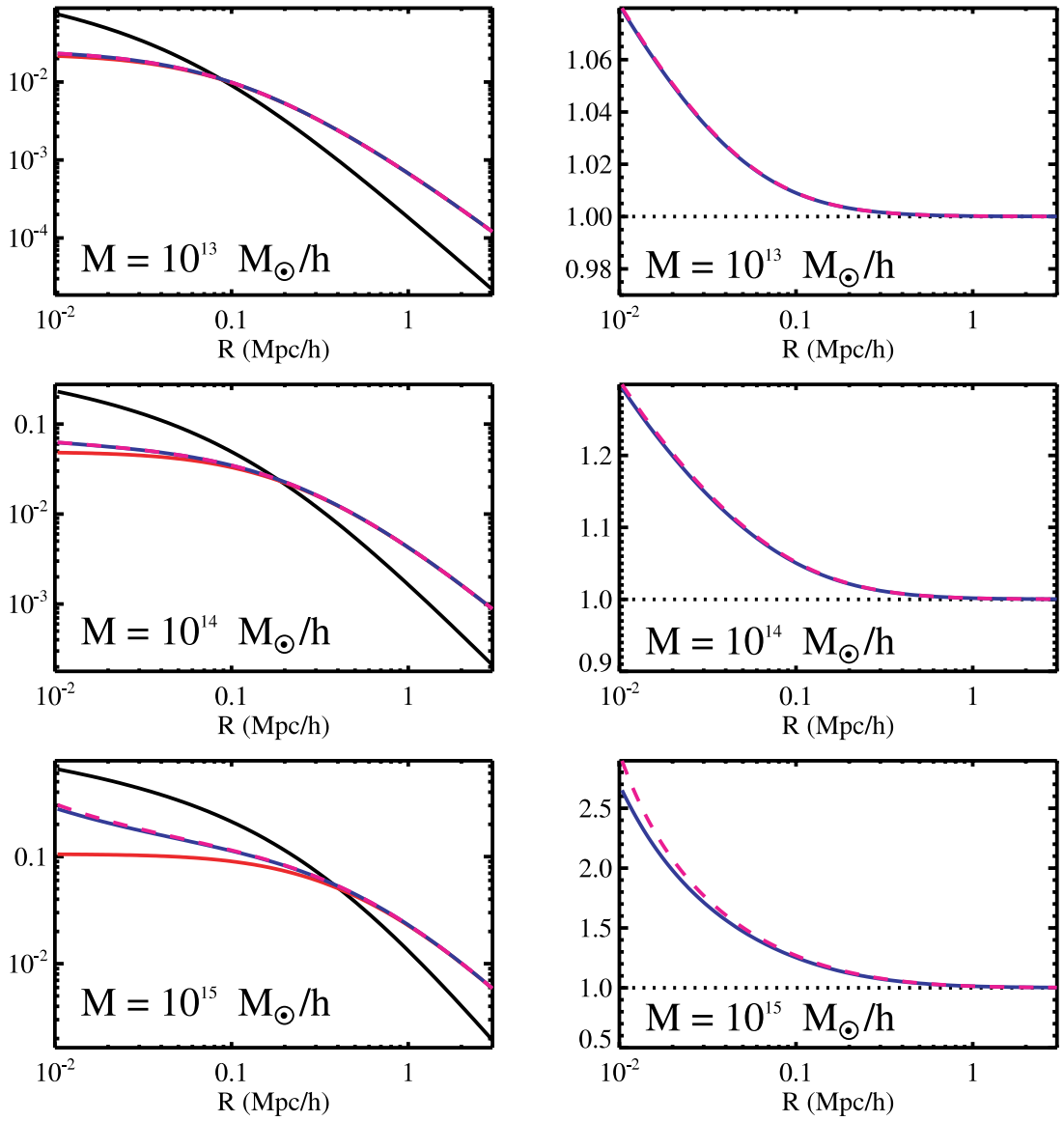


FIG. 5.—*Left panels*: For three mass ranges $M = (10^{13}, 10^{14}, 10^{15})\ h^{-1}\ M_{\odot}$ for NFW profiles (*top to bottom*) we show $\kappa(R)$ (*solid black line*), $\gamma(R)$ (*solid red line*), the distortion $\delta(R)/2$ (*solid blue line*), and $g$ (*dashed magenta line*). For the smallest mass, all three of $\gamma$, $g$, and $\delta/2$ basically coincide. For the larger masses, $\gamma$ is significantly less than the other two. The right panels show $\delta/2$ (*solid blue line*) and $g$ (*dashed magenta line*) both divided by $\gamma$. This shows more clearly the difference between these quantities. In the nonlinear regime, the difference between the measured distortion and $\gamma$ needs to be taken into account. These plots assume typical redshift ranges for SDSS data.

average shear and thus the inverted densities and mass profiles are simply averages of the individual density and mass profiles. However, this is no longer exactly true in the nonweak regime. The nonlinearities inherent in equations (7) and (10) make this more complicated. Normally you would average the $\delta(R)$ profiles and compute $g(R)$ from this average $\delta(R)$ and then $G(R)$ from $g(R)$ and finally your estimator for $-\Sigma'(R)$, which is $\Sigma_{\rm crit} G(R) \exp\left[-\int_R^{\infty} dp\, G(p)\right]$.

Note that this will not be equal to what you would get if you measure $\Sigma_{\rm crit} G(R) \exp\left[-\int_R^{\infty} dp\, G(p)\right]$ for each cluster and then average those. If you proceed this way instead, your inversion for $\Delta\rho$ would really be an unbiased estimator for the real average $\Delta\rho$. These issues of nonlinearities will not affect virial mass estimates or interpretation of the large-scale lensing signal in terms of the linear correlation function. Small-scale issues such as NFW cusp slope and concentration measurements may require more careful handling. As we stated earlier, clusters approaching the strong-lensing regime probably will not need to be stacked anyway. There may also be strong-lensing information available such as arcs or multiple images. However, stacking them will reduce projection effects that may bias any given cluster mass depending on its particular orientation.

## 10. THE EFFECTS OF CLUSTER MISCENTERING

To interpret the tangential shear pattern around clusters as the tangential shear from the associated dark matter halos, we are assuming that the centers that we are using coincide with that of the dark matter halos. Here we wish to relax this assumption and demonstrate how the tangential shear pattern is affected by any miscentering.

Let us assume that we have an offset of $R_s$ between the cluster center and the underlying dark matter halo. Let us further assume that when averaging over many clusters the direction of the offset is random and so we can azimuthally average over this angle.

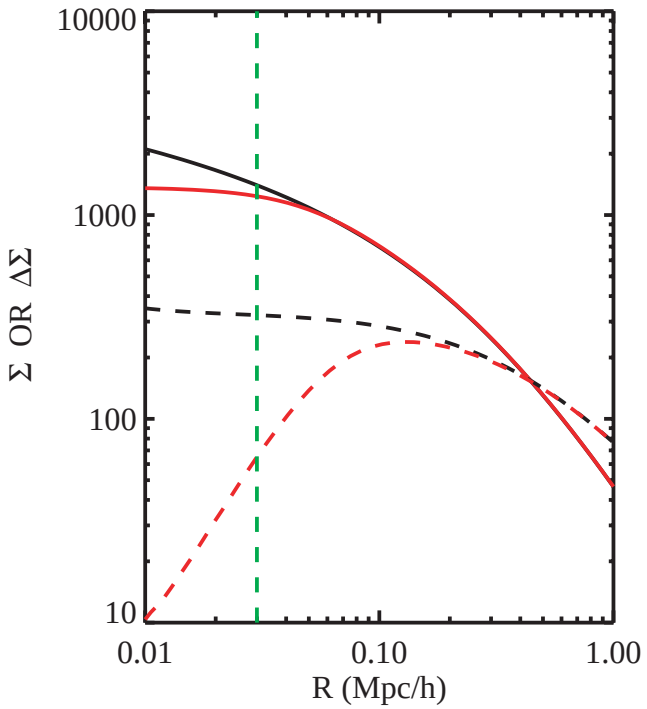


FIG. 6.—Demonstration of the effects of miscentering (intentional or not) on the tangential shear profile. The upper solid lines show the NFW $\Sigma(R)$ profile (*black*) and the convolved $\Sigma_c(R)$ profile (*red*). The lower dashed lines are the NFW $\Delta\Sigma(R)$ profile (*black*) and the resultant $\Delta\Sigma_c(R)$ profile (*red*) derived from the convolved $\Sigma_c$. The convolution kernel is a Gaussian of scale length $\sigma$ indicated by the vertical dashed green line. The relative effects of convolution on $\Delta\Sigma$ are larger than on $\Sigma$ and make a noticeable difference to about $R = 10\ \sigma$.

The resultant $\Delta\Sigma(R|R_s)$ is given in terms of the halos's profile $\Delta\Sigma(R)$ by the formula (Yang et al. 2006)

$$\Sigma(R|R_s) = \frac{1}{2\pi}\int_0^{2\pi} d\theta\, \Sigma\left(\sqrt{R^2 + R_s^2 + 2RR_s\cos\theta}\right). \quad (32)$$

If there is some distribution $P(R_s)$ of offsets, we can average over it to get the resultant convolved tangential shear profile, $\Sigma_c(R)$,

$$\Sigma_c(R) = \int_0^{\infty} dR_s\, P(R_s)\Sigma(R|R_s). \quad (33)$$

In general, for any distribution of offsets $g(\boldsymbol{x})$, the resultant $\Sigma_c(\boldsymbol{x})$ is given by convolution, $\Sigma_c = \Sigma g$, with the above equations applying to azimuthally symmetric $g$.

As an example, we can choose an azimuthally symmetric Gaussian for $g(\boldsymbol{x})$ with scale length $\sigma$ that gives normalized $P(R_s)$ given by

$$P(R_s) = \frac{R_s}{\sigma^2}\exp\left(-\frac{1}{2}\frac{R_s^2}{\sigma^2}\right). \quad (34)$$

Furthermore, we choose an NFW profile for $\Sigma(R)$ with concentration $c = 3$ and $r_{\rm vir} = 1\ h^{-1}$ Mpc and set $\sigma = 30\ h^{-1}$ kpc.

Figure 6 shows the resultant profiles. The top two lines show that the effect of the convolution on $\Sigma$ is to flatten the cusp where $R \le \sigma$. The vertical line indicates $\sigma$. The effect on $\Delta\Sigma$ is more pronounced. At $R = \sigma$, the two $\Sigma$ profiles only differ by 10%, but the $\Delta\Sigma$ differ by a factor of 4.5. The effect on $\Sigma$ is negligible after about $R > 2\ \sigma$ but not until $R > 10\ \sigma$ for $\Delta\Sigma$. This is not surprising since a flattening $\Sigma$ is nearly a mass sheet and so produces no shear.

Fitting model profiles such as NFW to real data should take into account possible miscentering. In practice, one would probably center clusters on the brightest cluster galaxy (BCG). Even if this is not at the center of the cluster's dark matter halo, it will probably have its own dark matter subhalo and at any rate contains some mass in its stars. Therefore, it may be expected that the measured $\Delta\Sigma$ profile will be cuspy at small scales even if the central BCG is offset from the main halo center. At small scales, the interpretation of a measured signal is complicated by both this effect and also the effects of shear nonlinearity discussed in the previous section.

By moving the center systematically to different locations, for example, from the BCG to the X-ray center, one should be able to measure which of these locations better traces the dark matter halo. One may also be able to measure the fraction of mass in subhalos by recentering on each galaxy in the cluster and subtracting off the contribution from the convolved cluster halo.

## 11. DISCUSSION

We have developed a nonparametric method for inverting cross-correlation lensing measurements to obtain average density and mass profiles of halos. We have demonstrated the method on an *N*-body simulation and shown that it successfully recovers the 3D profiles. We argued that asphericity of halos and LSS along the line of sight do not introduce bias or substantial uncertainty in the inferred mass estimates.

This method should find several useful applications in surveys. Applied to galaxy-galaxy lensing measurements, this method can be used to measure $\Delta\rho(r) = \Omega_m\rho_{\rm crit}\xi_{gm}(r)$ for samples of lens galaxies (e.g., Sheldon, et al. 2004). On small scales, this provides information about galaxy dark matter halos and should thereby constrain models of galaxy formation and evolution, including the details of hierarchical structure formation and merging. Statistical mass profiles around galaxies can provide virial mass measurements and, combined with the average light profile around these galaxies, determine mass-to-light ratios as a function of scale.

On larger scales, galaxy-galaxy lensing inversions should tell us more about cosmology. The autocorrelation function of galaxies allows us to measure $\xi_{gg}(r) = b^2(r)\xi_{mm}(r)$, where $b(r)$ is the scale-dependent bias and $\xi_{mm}(r)$ is the mass autocorrelation function. Lensing allows us to measure $\Omega_m\xi_{gm}(r) = \Omega_m b(r) r_\times(r)\xi_{mm}(r)$, where the cross bias, $r_\times(r)$, is sometimes referred to as stochastic bias (a term we do not recommend using in this context, since it is not bounded by $\pm 1$). The presence of these two bias functions, $b(r)$ and $r_\times(r)$, allows for the most general model of the relative clustering of galaxies and mass (see also Neyrinck et al. 2005). Simulations indicate that $r_\times(r)$ is consistent with unity on scales larger than $\sim 2\ h^{-1}$ Mpc for any range of halo masses (Tasitsiomi et al. 2004). In this case, measurement of large-scale galaxy-galaxy lensing and of the galaxy autocorrelation function determines the mass correlation function, $\Omega_m^2\xi_{mm} = (\Omega_m\xi_{gm})^2/\xi_{gg}$, and so fixing the *shape* of the linear correlation function determines the amplitude or normalization $\Omega_m\sigma_8$. It also determines the bias, $b(r)/\Omega_m = \xi_{gg}/\Omega_m\xi_{gm}$. The assumption that bias is scale independent on large scales is crucial to extracting cosmological information from the galaxy power spectrum; lensing can provide a way to test this important assumption. Other cosmological probes, such as cluster counts and cosmic shear, measure a different parameter combination, $\Omega_m^{0.5}\sigma_8$, so combining these probes with cross-correlation lensing helps break

this classic degeneracy. An advantage of this new method is that it is robust; the only assumptions used are that general relativity is correct and that $r_\times(r) \simeq 1$ on large scales.

Applied to galaxy clusters, this inversion method should prove useful for both astrophysics and cosmology. On small scales, cross-correlation lensing probes the mean density profiles of clusters. Cosmological dark matter simulations indicate that clusters have universal dark matter halos (although baryonic physics presumably needs to be taken into account to understand the inner density structure in detail). Inverted density profiles can test this assumption quite directly. As we have demonstrated, cluster mass profiles can be determined, leading to virial mass estimates independent of a model for the density profiles. Virial masses allow a direct comparison between simulated and real clusters. In addition, in a cluster survey we can measure the cluster abundance, $n(M, z)$, binned by any observable proxy for mass, and use this inversion method to calibrate the virial mass-observable relation as a function of redshift. This should allow cluster surveys to more precisely probe cosmology, including the dark energy. We still need to gather information about the scatter in the mass-observable relation; this can be constrained by simulations, by self-calibration (Lima & Hu 2005), and internally by dividing the clusters into subsamples and separately estimating the mass-observable relation.

Before applying this method with confidence to real cluster data, the possible selection biases involved in a given cluster *selection* algorithm need to be explored. If the selection algorithm preferentially finds clusters aligned along the line of sight, the assumption of statistical isotropy will be violated and the mass inversions may be biased. Similarly, if the selection algorithm assigns cluster centroids that are displaced from the true halo centers, the resulting density profiles will be convolved with distribution of centroid errors. Additional complications for the cluster counting technique arise if clusters, however selected, are not isomorphic to massive dark matter halos. These issues can be tested with simulations.

Applying these lens inversion methods to clusters on large scales provides new cosmological information beyond the cluster counting technique. As noted for galaxies above, we can use the cluster-mass and cluster-cluster correlations to measure $\Omega_m\sigma_8$. A useful cross-check is provided by comparing the results for different cluster samples binned by some observable, e.g., richness. This method should be even more robust for clusters than for galaxies, since the halo bias as a function of mass is predictable from theory and can be studied in dissipationless $N$-body simulations (Seljak & Warren 2004; Tasitsiomi et al. 2004). One can use this information in two ways. Conservatively, one could measure the average cluster masses, $M$, and $b(M)/\Omega_m$; from these bias measurements one could further constrain cosmology through the theoretical bias predictions and the direct measurement of $\Omega_m\sigma_8$. Alternatively, one could use the lensing data only on small scales to determine masses and assume that the theoretical bias predictions are correct; in this approach, one would use the mass and an assumed cosmology to predict the bias. This would allow one to "debias" the cluster autocorrelation function and obtain a constraint on $\sigma_8$ that depends on other cosmological parameters as well. Since the scaling of bias with mass is somewhat weak [$b(M) \sim M^{0.2}$ in the vicinity of $M = 10^{13}\,h^{-1}\,M_\odot$] and the cluster autocorrelation function can be measured with high signal-to-noise ratio, this measurement of $\Omega_m\sigma_8$ could be quite precise. This method has been applied to galaxies by Seljak et al. (2005) utilizing the halo model. We believe that clusters are a more natural choice for this method in that they do not need an additional halo occupancy prescription.

The approach of "debiasing" the correlation function also makes possible the study of the growth of structure. In principle, the inversion method can be applied to data in redshift bins, allowing one to measure the large-scale mass power spectrum amplitude as a function of redshift. This is particularly useful because the linear density perturbation growth rate is sensitive to dark energy. Moreover, since the evolution of the halo mass function is also sensitive to the growth of structure, combining cluster counts and the cluster-mass correlation function should offer a powerful probe of cosmology.

We are currently applying these statistical lensing inversion techniques to optically selected clusters in the SDSS. In the future, ongoing and planned wide-area surveys, including RCS II, CFHTLS, the VST and VISTA surveys, the Dark Energy Survey, PanSTARRS, LSST, and SNAP/JDEM, will allow for higher signal-to-noise ratio lensing measurements to be made and will eventually probe much larger cosmological volumes. The cosmological constraints they supply from cross-correlation lensing will complement those from other probes.

D. E. J. would like to thank the KICP for its hospitality on many visits, as well as the Aspen Center for Physics. We would like to thank Martin White and Andreas Berlind for providing the $N$-body simulation that was used in this paper. R. H. W. was supported by NASA through Hubble Fellowship HF-01168.01-A awarded by Space Telescope Science Institute.

REFERENCES

Abazajian, K., & Dodelson, S. 2003, Phys. Rev. Lett., 91, 041301
Abel, N. H. 1826, J. Reine Angewandte Math., 1, 153
Bacon, D. J., Massey, R. J., Refregier, A. R., & Ellis, R. S. 2003, MNRAS, 344, 673
Bacon, D. J., Refregier, A. R., & Ellis, R. S. 2000, MNRAS, 318, 625
Bahcall, N. A., & Bode, P. 2003, ApJ, 588, L1
Bahcall, N. A., & Cen, R. 1992, ApJ, 398, L81
Bahcall, N. A., et al. 2003, ApJ, 585, 182
Borgani, S., & Guzzo, L. 2001, Nature, 409, 39
Bradač, M., Lombardi, M., & Schneider, P. 2004, A&A, 424, 13
Brainerd, T. G., Blandford, R. D., & Smail, I. 1996, ApJ, 466, 623
Bremer, M. 1995, A&AS, 112, 551
Cen, R. 1997, ApJ, 485, 39
Clowe, D., De Lucia, G., & King, L. 2004, MNRAS, 350, 1038
Clowe, D., Luppino, G. A., Kaiser, N., Henry, J. P., & Gioia, I. M. 1998, ApJ, 497, L61
Cypriano, E. S., Sodré, L. J., Kneib, J., & Campusano, L. E. 2004, ApJ, 613, 95
Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371
dell'Antonio, I. P., & Tyson, J. A. 1996, ApJ, 473, L17
Del Popolo, A. 2003, ApJ, 599, 723
de Putter, R., & White, M. 2005, NewA, 10, 676
Dodelson, S. 2004, Phys. Rev. D, 70, 023008
Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605
Eisenstein, D. J., et al. 2005, ApJ, 633, 560
Fahlman, G., Kaiser, N., Squires, G., & Woods, D. 1994, ApJ, 437, 56
Fischer, P., et al. 2000, AJ, 120, 1198
Grego, L., Carlstrom, J. E., Reese, E. D., Holder, G. P., Holzapfel, W. L., Joy, M. K., Mohr, J. J., & Patel, S. 2001, ApJ, 552, 2
Griffiths, R. E., Casertano, S., Im, M., & Ratnatunga, K. U. 1996, MNRAS, 282, 1159
Haiman, Z., Mohr, J. J., & Holder, G. P. 2001, ApJ, 553, 545
Hamana, T., et al. 2003, ApJ, 597, 98
Hoekstra, H. 2001, A&A, 370, 743
———. 2003, MNRAS, 339, 1155
Hoekstra, H., Franx, M., Kuijken, K., Carlberg, R. G., & Yee, H. K. C. 2003, MNRAS, 340, 609
Hoekstra, H., Yee, H. K. C., & Gladders, M. D. 2002, ApJ, 577, 595
Hu, W. 2002, Phys. Rev. D, 66, 083515

Hudson, M. J., Gwyn, S. D. J., Dahle, H., & Kaiser, N. 1998, ApJ, 503, 531
Irgens, R. J., Lilje, P. B., Dahle, H., & Maddox, S. J. 2002, ApJ, 579, 227
Jarvis, M., Bernstein, G. M., Fischer, P., Smith, D., Jain, B., Tyson, J. A., & Wittman, D. 2003, AJ, 125, 1014
Joffre, M., et al. 2000, ApJ, 534, L131
Kaastra, J. S. 1989, A&A, 224, 338
Kaiser, N. 1984, ApJ, 284, L9
———. 1995, ApJ, 439, L1
Kristian, J. 1967, ApJ, 147, 864
Lima, M., & Hu, W. 2004, Phys. Rev. D, 70, 043504
———. 2005, Phys. Rev. D, 72, 043006
Łokas, E. L., & Mamon, G. A. 2003, MNRAS, 343, 401
Lombardi, M., & Bertin, G. 1999, A&A, 348, 38
Luppino, G. A., & Kaiser, N. 1997, ApJ, 475, 20
Majumdar, S., & Mohr, J. J. 2004, ApJ, 613, 41
Mandelbaum, R., Tasitsiomi, A., Seljak, U., Kravtsov, A. V., & Wechsler, R. H. 2005, MNRAS, 362, 1451
Maoli, R., et al. 2001, A&A, 368, 766
McKay, T. A., et al. 2001, ApJ, submitted (astro-ph/0108013)
———. 2002, ApJ, 571, L85
Metzler, C. A., White, M., & Loken, C. 2001, ApJ, 547, 560
Metzler, C. A., White, M., Norman, M., & Loken, C. 1999, ApJ, 520, L9
Miralda-Escude, J. 1991, ApJ, 370, 1
Mo, H. J., Jing, Y. P., & White, S. D. M. 1996, MNRAS, 282, 1096
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Neyrinck, M. C., Hamilton, A. J. S., & Gnedin, N. Y. 2005, MNRAS, 362, 337
Pierpaoli, E., Borgani, S., Scott, D., & White, M. 2003, MNRAS, 342, 163
Plummer, H. C. 1911, MNRAS, 71, 460
Refregier, A., et al. 2004, AJ, 127, 3102
Reiprich, T. H., & Böhringer, H. 2002, ApJ, 567, 716
Rhodes, J., Refregier, A., & Groth, E. J. 2001, ApJ, 552, L85
Rozo, E., Dodelson, S., & Frieman, J. A. 2004, Phys. Rev. D, 70, 083008
Saunders, W., Rowan-Robinson, M., & Lawrence, A. 1992, MNRAS, 258, 134
Schneider, P. 1996, MNRAS, 283, 837
Schneider, P., & Bartelmann, M. 1997, MNRAS, 286, 696
Schneider, P., & Rix, H. 1997, ApJ, 474, 25
Schneider, P., & Seitz, C. 1995, A&A, 294, 411
Seitz, S., & Schneider, P. 2001, A&A, 374, 740
Seljak, U. 2000, MNRAS, 318, 203
Seljak, U., & Warren, M. S. 2004, MNRAS, 355, 129
Seljak, U., et al. 2005, Phys. Rev. D, 71, 043511
Sheldon, E. S., et al. 2004, AJ, 127, 2544
Sheth, R. K., & Tormen, G. 1999, MNRAS, 308, 119
Smith, D. R., Bernstein, G. M., Fischer, P., & Jarvis, M. 2001, ApJ, 551, 643
Song, Y., & Knox, L. 2004, Phys. Rev. D, 70, 063510
Takada, M., & Jain, B. 2004, MNRAS, 348, 897
Tasitsiomi, A., Kravtsov, A. V., Wechsler, R. H., & Primack, J. R. 2004, ApJ, 614, 533
Tyson, J. A., & Fischer, P. 1995, ApJ, 446, L55
Tyson, J. A., Valdes, F., Jarvis, J. F., & Mills, A. P. 1984, ApJ, 281, L59
van der Marel, R. P., Magorrian, J., Carlberg, R. G., Yee, H. K. C., & Ellingson, E. 2000, AJ, 119, 2038
Van Waerbeke, L., et al. 2000, A&A, 358, 30
Viana, P. T. P., & Liddle, A. R. 1999, MNRAS, 303, 535
Von Zeipel, M. H. 1908, Ann. Obs. Paris, 25, 1
White, M. 2002, ApJS, 143, 241
———. 2003, ApJ, 597, 650
White, M., van Waerbeke, L., & Mackey, J. 2002, ApJ, 575, 640
White, S. D. M., Efstathiou, G., & Frenk, C. S. 1993, MNRAS, 262, 1023
Wilson, G., Kaiser, N., Luppino, G. A., & Cowie, L. L. 2001, ApJ, 555, 572
Wittman, D. M., Tyson, J. A., Kirkman, D., Dell'Antonio, I., & Bernstein, G. 2000, Nature, 405, 143
Yang, X., Mo, H. J., van den Bosch, F. C., Jing, Y. P., Weinmann, S. M., & Meneghetti, M. 2006, MNRAS, 373, 1159
York, D. G., et al. 2000, AJ, 120, 1579