

Likelihood Methods for Cluster Dark Energy Surveys

Wayne Hu¹ and J.D. Cohn²

¹*Kavli Institute for Cosmological Physics, Department of Astronomy and Astrophysics,
and Enrico Fermi Institute, University of Chicago, Chicago IL 60637*

²*Space Sciences Laboratory and Theoretical Astrophysics Center, University of California, Berkeley CA 94720*

Galaxy cluster counts at high redshift, binned into spatial pixels and binned into ranges in an observable proxy for mass, contain a wealth of information on both the dark energy equation of state and the mass selection function required to extract it. The likelihood of the number counts follows a Poisson distribution whose mean fluctuates with the large-scale structure of the universe. We develop a joint likelihood method that accounts for these distributions. Maximization of the likelihood over a theoretical model that includes both the cosmology and the observable-mass relations allows for a joint extraction of dark energy and cluster structural parameters.

I. INTRODUCTION

Upcoming high redshift cluster surveys have the potential to provide precision constraints on the evolution of the dark energy density due to the exponential sensitivity of the abundance of massive dark matter halos to the growth of structure (e.g. [1, 2]). The wealth of information contained in even the cluster counts alone partially offsets the fundamental problem that the masses of the clusters are not directly observable. Techniques that utilize this information to relate cluster observables such as Sunyaev-Zel'dovich (SZ) flux, X-ray temperature and surface brightness, or optical richness and shear to cluster masses and hence the cosmology are known in the literature as “self-calibration” methods [3, 4, 5].

Extracting the full information from the number counts requires a joint analysis of the mean abundance and spatial distribution. It is also required in that the spatial clustering of clusters is the source of sampling errors for the counts [6]. Likewise, cluster power spectrum estimators must account for the non-Gaussian distribution of their errors even in the linear regime. This fact distinguishes cluster likelihood methods from their galaxy counterparts [7, 8] in requiring a full treatment of the Poisson shot noise [9, 10, 11].

Likelihood analyses of the local cluster abundance in the literature conversely include the Poisson shot noise but omit the power spectrum source of error and information (see e.g. [12, 13]). This omission is justified as they typically utilize only a handful of the very rarest clusters where shot noise dominates and the likelihood is Poisson. Fisher matrix studies show that for sufficiently deep surveys the likelihood deviates significantly from the Poisson limit [6] and that these deviations contain useful information for self-calibration [14, 15]. For example, $\sim 10^4$ high redshift clusters are expected in the upcoming South Pole Telescope (SPT) survey [20], allowing the sample to be divided both spatially and in SZ flux.

In this *Brief Report*, we develop the likelihood techniques required to extract this information. We show that the joint Poisson-Gaussian likelihood can be reduced to a closed-form expression that is applicable to the next generation high- z cluster surveys. We begin in §II with

a general description of the likelihood function and specialize it to the case of rare high- z clusters in §III. We conclude in §IV.

II. CLUSTER LIKELIHOOD FUNCTION

Let us take the data to be the number of clusters $N_{i\mu}$ in bins of some observable proxy for mass ($\mu = 1, \dots, n_m$) and pixels delineated by angle and redshift ($i = 1, \dots, n_p$). We will assume that the clusters are Poisson distributed with a mean of $m_{i\mu}$. This mean fluctuates from pixel to pixel due to the large scale structure of the universe. We will further assume that these pixels are sufficiently large that the fluctuations are in the linear regime. Denoting the ensemble means with overbars,

$$m_{i\mu} = \bar{m}_{i\mu}(1 + b_{i\mu}\delta_i), \quad (1)$$

where δ_i is the overdensity within the pixel and $b_{i\mu}$ is the linear bias of the selected clusters.

The distribution of Poisson means is then given by a Gaussian with covariance [6]

$$\langle (m_{i\mu} - \bar{m}_{i\mu})(m_{j\nu} - \bar{m}_{j\nu}) \rangle = \bar{m}_{i\mu}\bar{m}_{j\nu}b_{i\mu}b_{j\nu}S_{ij}, \quad (2)$$

where

$$S_{ij} = \int \frac{d^3k}{(2\pi)^3} W_i^*(\mathbf{k}) W_j(\mathbf{k}) P(k; z_{ij}). \quad (3)$$

Here W_i is the Fourier transform of the pixel window, normalized such that $\int d^3x W_i(\mathbf{x}) = 1$ and $P(k; z_{ij})$ is the linear power spectrum at the mean redshift of the pixels z_{ij} .

Through N -body simulations, a cosmological model predicts the comoving spatial number density of clusters as a function of their mass M (e.g. [16, 17, 18, 19]). Given a selection in mass $p_{i\mu}(M)$ and the comoving volume of the pixel V_i , the ensemble mean becomes

$$\bar{m}_{i\mu} = V_i \int d \ln M p_{i\mu}(M) \frac{dn}{d \ln M}. \quad (4)$$

Likewise, simulations predict the bias of clusters as a function of their mass $b(M)$ and hence

$$b_{i\mu} = \frac{V_i}{\bar{m}_{i\mu}} \int d \ln M p_{i\mu}(M) b(M) \frac{dn}{d \ln M}. \quad (5)$$

For a completely fixed selection $p_{i\mu}$, the cosmology uniquely determines both the set of ensemble mean counts $\bar{m}_{i\mu}$ and their distribution through $b_{i\mu}$ for each pixel i . This fact allows a measure of “self-calibration” of the selection function $p_{i\mu}(M)$ through the relationships between $b_{i\mu}$ and $\bar{m}_{i\mu}$ in a given pixel [4, 5]. Note that the selection function can include effects such as instrument noise and point source contamination as well as the observable-mass relation itself.

To exploit this information, one can maximize the likelihood of the data $N_{i\mu}$ over a theory defined by both cosmological and selection parameters. Fisher matrix techniques estimate that recovery of the dark energy equation of state at the $\sim 10\%$ level for next generation surveys like the SPT is possible, even in the presence of moderate uncertainties in the selection function [4, 14, 15].

In practice, extracting this information will require an exploration of the likelihood space. Combining the Poisson likelihood of drawing $N_{i\mu}$ from $m_{i\mu}$ and the Gaussian likelihood of drawing $m_{i\mu}$ from $\bar{m}_{i\mu}$, we obtain [14]

$$\mathcal{L}(\mathbf{N}|\bar{\mathbf{m}}, \mathbf{b}, \mathbf{S}) = \left[\prod_{i=1}^{n_p} \int_{-b_{\max}^{-1}}^{\infty} d\delta_i \left(\prod_{\mu=1}^{n_m} \frac{m_{i\mu}^{N_{i\mu}}}{N_{i\mu}!} e^{-m_{i\mu}} \right) \right] \times \frac{1}{\sqrt{(2\pi)^{n_p} \det \mathbf{S}}} e^{-\frac{1}{2} \delta^T \mathbf{S}^{-1} \delta}. \quad (6)$$

Here and throughout we use vector and matrix notation where no confusion will arise. The lower limit on the integral is required to prevent negative Poisson means. As long as $b_{i\mu} S_{ii}^{1/2} \ll 1$, the likelihood is insensitive to the exact value of b_{\max} and so we set it to the maximum bias for clusters obtainable in the model. In more marginal cases, the missing probability from $-1 \leq \delta \leq -1/b_{\max}$ can be restored to the $N = 0$ likelihood (see §III for further discussion). For $S_{ii}^{1/2} \gtrsim 1$, the density field itself is not Gaussian.

Since $m_{i\mu}$ depends on δ_i through Eqn. (1), the full likelihood is a convolution of the Poisson and Gaussian distributions. To evaluate this convolution efficiently, let us rewrite it as

$$\mathcal{L}(\mathbf{N}|\bar{\mathbf{m}}, \mathbf{b}, \mathbf{S}) = \left[\prod_{i=1}^{n_p} \prod_{\mu=1}^{n_m} \frac{(-1)^{N_{i\mu}}}{N_{i\mu}!} \lim_{\lambda_{i\mu} \rightarrow 1} \left(\frac{\partial}{\partial \lambda_{i\mu}} \right)^{N_{i\mu}} \right] \times e^{-\lambda_{i\mu} \bar{m}_{i\mu}} F(\bar{\mathbf{m}}, \mathbf{b}, \mathbf{S}, \lambda), \quad (7)$$

where the generating function

$$F(\bar{\mathbf{m}}, \mathbf{b}, \mathbf{S}, \lambda) = \int_{-b_{\max}^{-1}}^{\infty} d^{n_p} \delta e^{-\sum_{i\mu} \lambda_{i\mu} \bar{m}_{i\mu} b_{i\mu} \delta_i} \times \frac{1}{\sqrt{(2\pi)^{n_p} \det \mathbf{S}}} e^{-\frac{1}{2} \delta^T \mathbf{S}^{-1} \delta}. \quad (8)$$

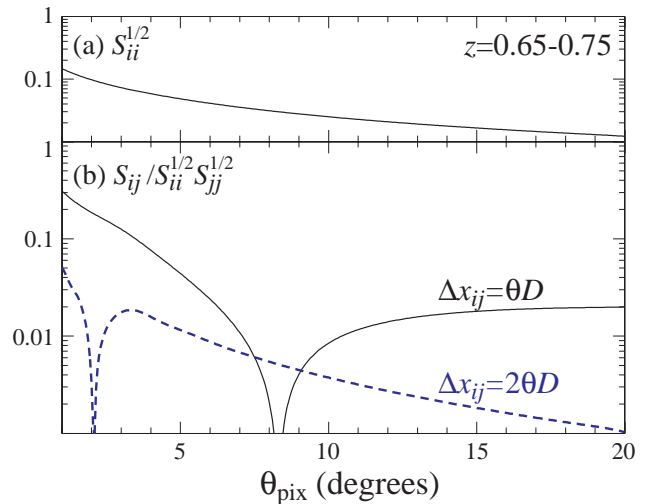


FIG. 1: Covariance matrix as a function of pixel size θ_{pix} at $z = 0.7$, $\Delta z = 0.1$: (a) rms density fluctuation per pixel $S_{ii}^{1/2}$; (b) correlation $S_{ij} / (S_{ii} S_{jj})^{1/2}$ for neighboring pixels. For $\theta_{\text{pix}} \gtrsim 1^\circ$ the covariance matrix is nearly diagonal with negligible correlations beyond the nearest neighbor pixel.

This integral may be transformed into one over a Gaussian by completion of squares through the diagonalization of the pixel covariance matrix $\mathbf{S} = \mathbf{A} \mathbf{\Sigma} \mathbf{A}^T$ where $\mathbf{\Sigma} = \text{diag}(\sigma_k^2)$ and \mathbf{A} is an orthonormal matrix of eigenvectors. If the covariance matrix is already nearly diagonal, the integration range will likewise transform to a semi-infinite interval such that

$$F(\bar{\mathbf{m}}, \mathbf{b}, \mathbf{S}, \lambda) = \prod_{k=1}^{n_p} e^{\frac{1}{2} R_k^2 \sigma_k^2} \left[1 - \frac{1}{2} \text{Erfc}(y_k / \sqrt{2}) \right], \quad (9)$$

where

$$R_k = \sum_{i\mu} \lambda_{i\mu} \bar{m}_{i\mu} b_{i\mu} A_{ik},$$

$$y_k = \sum_i \frac{1}{b_{\max} \sigma_k} A_{ik} - R_k \sigma_k. \quad (10)$$

The Erfc term appears because of the boundary at $-b_{\max}^{-1}$ and must be included for cases where the Poisson and Gaussian variance are comparable even if $b_{\max} \sigma_k \ll 1$ as required by Eqn. (6).

III. COARSE PIXEL LIMIT

Clusters at high redshift are extremely rare objects and hence coarse pixelization is required to have a finite expectation value $\bar{m}_{i\mu}$ per pixel. The likelihood function simplifies considerably in this limit.

For definiteness, let us assume that a survey like the SPT counts clusters above an observable threshold that

corresponds to a mean mass of $M_{\text{th}} = 10^{14.2} h^{-1} M_{\odot}$ with a $\sigma_{\ln M} = 0.25$ Gaussian scatter in $\ln M$ (see [14] for calculational details). We will take as the fiducial cosmology a flat Λ CDM model with $\Omega_{\Lambda} = 0.73$, baryon density $\Omega_b h^2 = 0.024$, matter density $\Omega_m h^2 = 0.14$ and a scale invariant initial curvature spectrum with amplitude $\delta_{\zeta} = 5.07 \times 10^{-5}$ corresponding to $\sigma_8 = 0.91$. Near the median of the redshift distribution at $z = 0.65 - 0.75$, the angular number density of these clusters is 0.86 deg.^{-2} in this fiducial cosmology and they have a linear bias of $b \approx 4$. To have an expectation value of at least 1 cluster per pixel, the pixels must be at least $1^{\circ} \times 1^{\circ}$ in size.

The transverse size of the pixel is $L_{\text{pix}} = D\theta_{\text{pix}} \approx 31h^{-1} \text{ Mpc}$ ($\theta_{\text{pix}}/1^{\circ}$) where D is the comoving angular diameter distance to redshift $z \approx 0.7$. Since a pixel of several degrees on the side receives its fluctuations from k modes near the peak of the power spectrum, the spectrum is nearly white and the pixel covariance matrix is nearly diagonal.

To quantify this expectation, let us take square pixels on the sky of angular extent θ_{pix} with a depth Δz centered around a redshift z . The covariance in the radial direction is negligible for $\Delta z = 0.1$ and so we calculate only the angular pixel covariance at each redshift slice. The Fourier window function is then

$$|W(\mathbf{k})| = j_0(k_x L_{\text{pix}}/2) j_0(k_y L_{\text{pix}}/2) j_0(k_z \Delta z/2H), \quad (11)$$

where H is the Hubble parameter. The pixel covariance matrix becomes

$$S_{ij} = \int \frac{d^3 k}{(2\pi)^3} |W(\mathbf{k})|^2 \cos(k_x \Delta x_{ij}) \cos(k_y \Delta y_{ij}) P(k; z) \quad (12)$$

where $\Delta x_{ij} = L_{\text{pix}} n_{xij}$ and n_{xij} is the number of pixels separating i and j in the x direction and likewise for Δy_{ij} .

In Fig. 1, we plot the correlation of the nearest 2 neighboring pixels. Even for $\theta_{\text{pix}} = 1^{\circ}$, the covariance matrix beyond the nearest neighbor is negligible. Therefore S_{ij} is extremely sparse and the diagonalization can be efficiently implemented. In fact, for $\theta_{\text{pix}} > 3^{\circ}$ a strictly diagonal covariance matrix is a good approximation.

Now let us examine the likelihood function under this diagonal covariance matrix approximation. For simplicity, we first also assume that there is only one bin in observable mass such that the selection function is $p_i(M) = \frac{1}{2} \text{Erfc}[\ln(M_{\text{th}}/M)/\sqrt{2\sigma_{\ln M}^2}]$, i.e. a smooth function of the mass that increases from 0 to 1 across the threshold. The likelihood then simplifies to $\mathcal{L} = \prod_{i=1}^{n_{\text{pix}}} \mathcal{L}_i$, where

$$\mathcal{L}_i = \frac{(-1)^{N_i}}{N_i!} \lim_{\lambda_i \rightarrow 1} \left(\frac{\partial}{\partial \lambda_i} \right)^{N_i} e^{-\lambda_i \bar{m}_i + \frac{1}{2} \lambda_i^2 \bar{m}_i^2 b_i^2 \sigma_i^2} \quad (13)$$

$$\times \left\{ 1 - \frac{1}{2} \text{Erfc} \left[\frac{1}{\sqrt{2}} \left(\frac{1}{b_i \sigma_i} - \lambda_i \bar{m}_i b_i \sigma_i \right) \right] \right\}.$$

In Fig. 2, we plot this likelihood as a function of N_i and compare it with the Poisson likelihood. Here we take two representative cases: $\theta_{\text{pix}} = 2^{\circ}$ where $\bar{m}_i = 3.45$,

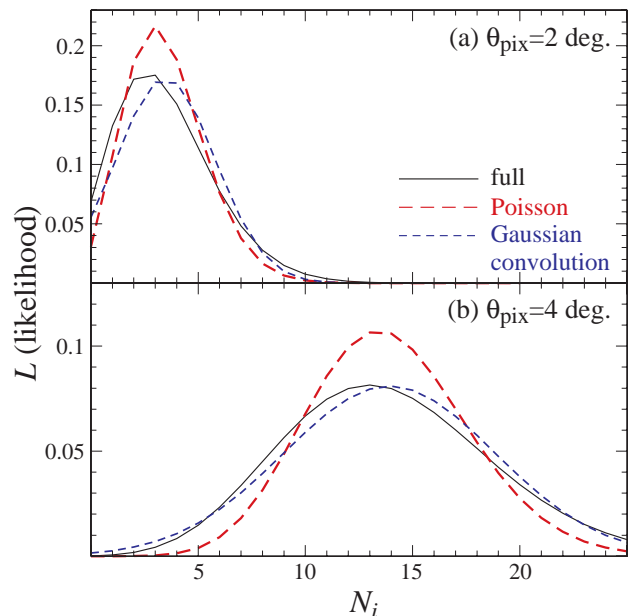


FIG. 2: Likelihood as a function of counts N_i above a mean threshold mass $M_{\text{th}} = 10^{14.2} h^{-1} M_{\odot}$ and $z = 0.7$, $\Delta z = 0.1$ for (a) $\theta_{\text{pix}} = 2^{\circ}$ and (b) $\theta_{\text{pix}} = 4^{\circ}$. Shown are the full likelihood (solid) from Eqn. (13) compared with a Poisson likelihood (long dashed) and the $\bar{m}_i \gg 1$ limiting form in Eqn. (14) of a convolution of Gaussian shot noise and sample variance (dashed).

$\sigma_i = 0.097$; and $\theta_{\text{pix}} = 4^{\circ}$ where $\bar{m}_i = 13.9$, $\sigma_i = 0.057$. The full likelihood is broader than the Poisson likelihood due to the fluctuating Poisson mean. This excess variance can be used to self-calibrate the mass selection through the bias in a manner that is stable to the choice of pixelization in this range [14].

This likelihood is normalized such that $\sum_{N_i} \mathcal{L}_i(N_i) = 1$ for a fixed model in the limit that $b_i \sigma_i \ll 1$. In practice, the finite $b_i \sigma_i$ will mean that there is a finite probability of drawing a Poisson mean of $m_i = 0$ and hence no clusters in the pixel. This probability has been omitted in Eqns. (6) and (13) but may be directly restored to the $N_i = 0$ likelihood. For $\theta_{\text{pix}} = 2^{\circ}$, this correction is 0.005 ; for 4° , this correction is a negligible 8×10^{-6} .

For the $\theta_{\text{pix}} = 4^{\circ}$ case, $\bar{m}_i \gg 1$ and the Poisson distribution of N_i around m_i should be nearly Gaussian. In this limit, the full likelihood is approximately the convolution of two Gaussians and becomes

$$\mathcal{L}_i \rightarrow \frac{1}{\sqrt{2\pi s_i^2}} \exp\left(-\frac{(N_i - \bar{m}_i)^2}{2s_i^2}\right), \quad (14)$$

where $s_i^2 = \bar{m}_i + \bar{m}_i^2 b_i^2 \sigma_i^2$ is the total variance of the convolution. Fig. 2 shows the accuracy of this approximation.

Finally, it is instructive to consider the case where the Poisson variance is much larger than the Gaussian sample variance, i.e. $\bar{m}_i \gg \bar{m}_i^2 b_i^2 \sigma_i^2$. In this case the Erfc boundary term is negligible and Taylor expanding the

likelihood gives a simple closed form

$$\mathcal{L}_i \rightarrow \frac{\bar{m}_i^{N_i} e^{-\bar{m}_i}}{N_i!} \left\{ 1 + \frac{1}{2} [(N_i - \bar{m}_i)^2 - N_i b_i^2 \sigma_i^2] \right\} \quad (15)$$

For the fiducial survey at $M_{\text{th}} = 10^{14.2} h^{-1} M_{\odot}$, $\bar{m}_i b_i^2 \sigma_i^2 = 0.52$ for $\theta_{\text{pix}} = 2^\circ$ and 0.76 for $\theta_{\text{pix}} = 4^\circ$. This implies that the Poisson variance is comparable to the sample variance and so this limiting case does not apply. For a finer pixelization, this limit can be reached but not before the linear theory assumption breaks down. Nevertheless, a survey that selects rarer clusters may reach this limit in the linear regime. For example with $M_{\text{th}} = 10^{14.4} h^{-1} M_{\odot}$, $\bar{m}_i = 4.6$ for $\theta_{\text{pix}} = 4^\circ$ pixels and $b_i = 4.7$, and the approximation deviates by less than 5% near the peak. It is also useful in understanding the qualitative behavior: the Gaussian variance of the Poisson mean m_i reduces the likelihood near the ensemble mean $N_i = \bar{m}_i$ and enhance it in the tails.

Likewise, the Poisson dominated limit is useful in understanding the effect of having multiple observable-mass bins per pixel. For example with 2 bins

$$\mathcal{L}_i \rightarrow \frac{\bar{m}_{i1}^{N_{i1}} e^{-\bar{m}_{i1}}}{N_{i1}!} \frac{\bar{m}_{i2}^{N_{i2}} e^{-\bar{m}_{i2}}}{N_{i2}!} \left(1 + \frac{1}{2} \{ [b_{i1}(N_{i1} - \bar{m}_{i1}) + b_{i2}(N_{i2} - \bar{m}_{i2})]^2 - b_{i1}^2 N_{i1} - b_{i2}^2 N_{i2} \} \sigma_i^2 \right). \quad (16)$$

The enhancement of the tails of the distribution occurs most strongly for joint fluctuations of N_{i1} and N_{i2} as

would be expected from the joint fluctuations of their Poisson means m_{i1} and m_{i2} .

The two cases of a nearly diagonal pixel covariance and a Poisson dominated likelihood considered here are the ones of practical importance for high redshift clusters.

IV. DISCUSSION

We have presented a closed-form expression for the full likelihood function for cluster number count surveys including both the Poisson shot noise and Gaussian sample variance of the Poisson means from the large-scale structure of the universe. We also allow for multiple bins in the observable mass per spatial pixel.

This treatment is especially useful for high redshift cluster surveys. Here the spatial covariance is nearly diagonal for pixels that are sufficiently large to have an expectation value of more than one cluster per pixel. A sparse covariance matrix allows the likelihood to be evaluated efficiently. Maximization of the likelihood over models for both the cosmology and the observable-mass relations will allow future cluster surveys to jointly determine dark energy and cluster structural parameters.

Acknowledgments: We thank T. Crawford, G. Holder, and M. White for useful discussions. J.D.C. was supported in part by NSF AST-0205935. W.H. was supported by the DOE, the Packard Foundation, and the KICP under NSF PHY-0114422.

-
- [1] L. Wang and P. Steinhardt, *Astrophys. J.* **508**, 483 (1998), astro-ph/9804015.
 - [2] Z. Haiman, J. Mohr, and G. Holder, *Astrophys. J.* **553**, 545 (2001).
 - [3] E. Levine, A. Schulz, and M. White, *Astrophys. J.* **577**, 569 (2002), astro-ph/0204273.
 - [4] S. Majumdar and J. Mohr, *Astrophys. J.* **585**, 603 (2003), astro-ph/0208002.
 - [5] W. Hu, *Phys. Rev. D* **67**, 081304 (2003), astro-ph/0301416.
 - [6] W. Hu and A. Kravtsov, *Astrophys. J.* **584**, 702 (2003), astro-ph/0203169.
 - [7] M. S. Vogeley and A. S. Szalay, *Astrophys. J.* **465**, 34 (1996).
 - [8] M. Tegmark, A. J. S. Hamilton, M. A. Strauss, M. S. Vogeley, and A. S. Szalay, *Astrophys. J.* **499**, 555 (1998).
 - [9] S. D. M. White, *Mon. Not. R. Astron. Soc.* **186**, 145 (1979).
 - [10] S. Dodelson, L. Hui, and A. H. Jaffe (1997), astro-ph/9712074.
 - [11] J. D. Cohn, *New Astronomy* **11**, 226 (2006).
 - [12] E. Pierpaoli, D. Scott, and M. White, *Mon. Not. R. Astron. Soc.* **325**, 77 (2001).
 - [13] Y. Ikebe, T. Reiprich, H. Boehringer, Y. Tanaka, and T. Kitayama, *Astron. Astrophys.* **383**, 773 (2002), astro-ph/0112315.
 - [14] M. Lima and W. Hu, *Phys. Rev. D* **70**, 043504 (2004), astro-ph/0401559.
 - [15] M. Lima and W. Hu, *Phys. Rev. D* **72**, 043006 (2005), astro-ph/0503363.
 - [16] R. Sheth and B. Tormen, *Mon. Not. R. Astron. Soc.* **308**, 119 (1999).
 - [17] A. Jenkins et al., *Mon. Not. R. Astron. Soc.* **321**, 372 (2001).
 - [18] A. E. Evrard et al., *Astrophys. J.* **573**, 7 (2002).
 - [19] K. Heitmann, P. Ricker, M. Warren, and S. Habib, *Astrophys. J.* submitted (2004).
 - [20] <http://spt.uchicago.edu>