UNCERTAINTIES IN THE SUNYAEV-ZEL’DOVICH–SELECTED CLUSTER ANGULAR POWER SPECTRUM

J. D. COHN\textsuperscript{1} and KENJI KADOTA\textsuperscript{2,3}

Received 2004 September 24; accepted 2005 June 13

ABSTRACT

Large Sunyaev-Zel’dovich–selected galaxy cluster surveys are beginning imminently. We compare the dependence of the galaxy cluster angular power spectrum on cosmological parameters, different modeling assumptions, and statistical observational errors. We quantify the degeneracies between theoretical assumptions such as the mass function and cosmological parameters such as $\sigma_8$. We also identify a rough scaling behavior of this angular power spectrum with $\sigma_8$ alone.

Subject headings: cosmic microwave background — cosmological parameters — galaxies: clusters: general

Online material: color figures

1. INTRODUCTION

The Sunyaev-Zel’dovich (SZ) effect is the upscattering of cosmic microwave background (CMB) photons by the hot electrons in galaxy clusters (Sunyaev & Zel’dovich 1972, 1980). As the SZ surface brightness of a cluster is redshift independent and the signal is relatively insensitive to the poorly known cluster core structure, the SZ effect is a powerful tool for selecting and studying galaxy clusters. The field of SZ measurements has progressed rapidly from a handful of SZ detections for already-known galaxy clusters to SZ maps of several clusters to the current stage: cluster surveys designed exclusively for SZ detection. For reviews, see, for example, Birkinshaw (1999) and Carlstrom et al. (2002). Some surveys are already in progress, such as those with the Arcminute Cosmology Bolometer Array Receiver (ACBAR)\textsuperscript{6} and the Sunyaev-Zeldovich Array (SZA).\textsuperscript{5} Others are on the verge of taking data, such as the Arcminute Microkelvin Imager (AMI)\textsuperscript{6} and the Atacama Pathfinder Experiment (APEX),\textsuperscript{7} and the Atacama Cosmology Telescope (ACT)\textsuperscript{8} is expected to be ready by the end of 2006. Thousands of previously unknown clusters will be observed by these experiments, starting with APEX. Future surveys such as the South Pole Telescope (SPT)\textsuperscript{9} and Planck\textsuperscript{10} will take place within the next five years and should observe tens of thousands of clusters. For a full list of current and upcoming experiments, see, for example, the CMB experiment page at LAMBDA (Legacy Archive for Microwave Background Data Analysis).\textsuperscript{11}

The prospect of combining these theoretical predictions and observational data on galaxy clusters to constrain fundamental cosmological parameters has led to much excitement (e.g., Holder et al. 2001; Majumdar & Mohr 2003 and references therein). With large survey data on the horizon and the detailed design of later experiments taking shape, it is timely to identify which theoretical assumptions are the most crucial to pin down in order to use the science harvest from these experiments. The initial galaxy cluster information that these surveys will produce will be number counts and the angular correlation function/cluster power spectrum.

Here we consider the angular power spectrum of clusters found by an SZ survey. We calculate the effect of cosmological, modeling, and observational parameters on this power spectrum. The cosmological parameters that we consider are $\Omega_m$ and $\sigma_8$. In the modeling, analytic predictions (based on the extremely good fit of analytic models to dark matter simulations) can predict the expected galaxy cluster counts, masses, and positions. These fits often differ at about the 10% level. In addition, cluster observations generally measure some proxy for the cluster mass (e.g., temperature, X-ray flux, SZ flux, velocities, or shear in images). These proxies generally depend on more complex astrophysical properties (gas dynamics, for instance, or the state of relaxation of the cluster) and/or are difficult to connect cleanly to the mass (e.g., weak-lensing masses; Metzler et al. 2001). We consider different analytic fits and assumptions for mass proxies. For the observational parameters we consider sensitivity and area of survey, using those of APEX for illustration. We quantify and compare the effects on the power spectrum of these uncertainties. Many of these uncertainties are degenerate; we identify which parameter changes are roughly equivalent, and which are not.

The SZ properties of clusters in upcoming surveys are of great interest, and many aspects have been studied in previous work. Constraints from the angular correlations of SZ clusters have been studied by Moscardini et al. (2002), Diaferio et al. (2003), and Mei & Bartlett (2003, 2004). These papers each varied a combination of different quantities (e.g., $\sigma_8$, normalization of SZ flux as a function of mass, biases, and scaling of SZ flux with mass and redshift). Majumdar & Mohr (2004) consider the two-dimensional power spectrum, and Wang et al. (2004) consider the three-dimensional power spectrum in conjunction with other measurements. A similar quantity, the angular correlation of the SZ temperature spectrum, has been considered by Komatsu & Kitayama (1999), Komatsu & Seljak (2002), and Majumdar & Mohr (2004). Battye & Weller (2003) studied systematics similar to those considered here for number counts as a function of redshift, $dN/dz$. Uncertainties that
were treated separately in (different) previous works can be degenerate, leading to potentially misleading interpretations. We quantify the changes in cosmological parameters that mimic these other theoretical assumptions.

In §2 we discuss the selection requirement for the SZ cluster catalog; §3 defines the angular power spectrum and angular correlation function. Our new results are presented in §4, in which we consider the cosmological modeling, and observational uncertainties in the power spectrum and compare them. We discuss many of the uncertainties in detail and then compare the changes caused by varying many different assumptions. We both find degeneracies and find which assumptions are not degenerate. Some of our assumptions have been considered separately in other works mentioned above (often for the correlation function rather than the power spectrum). By combining previously considered changes in assumptions and newer ones together in a homogeneous manner and comparing to the same relation function rather than the power spectrum). By combining both find degeneracies and find which assumptions are not degenerate. Some of our assumptions have been considered separately in other works mentioned above (often for the correlation function rather than the power spectrum). By combining previously considered changes in assumptions and newer ones together in a homogeneous manner and comparing to the same relation function rather than the power spectrum).

2. SZ-SELECTED GALAXY CLUSTER CATALOGS

An SZ-selected galaxy cluster catalog is one that includes all clusters above a certain minimum SZ flux or (equivalently) above some minimum Y-parameter $Y_{\text{min}}$. (The Y-parameter is defined in detail below.) For general reviews of the SZ effect, see, for example, Birkinshaw (1999), Rephaeli (1995a), and Sunyaev & Zel’dovich (1980). There are two SZ effects, thermal and kinetic. The (frequency dependent) thermal SZ effect is the change in the CMB spectrum due to random thermal motion of the intracluster electrons, and the (frequency independent) kinetic SZ effect is the change in the CMB spectrum due to bulk peculiar motions. The kinetic SZ effect is negligible compared to the thermal SZ effect for the purpose of cluster selection under study here; consequently, we restrict our attention to the thermal SZ effect; “SZ effect” hereafter means the thermal SZ effect.

The SZ effect can be described in terms of a CMB flux increment or decrement using the dimensionless Comptonization parameter $y(\theta)$,

$$\frac{\Delta T}{T_{\text{CMB}}} = g(x)y(\theta),$$

where the prefactor $g(x) = \{x[(e^x + 1)/(e^x - 1)] - 4\}[1 + \delta_{\text{rel}}(x)]$ includes frequency dependence, $x = h\nu/(k_B T_{\text{CMB}}) = \nu/56.84$ GHz; we neglect the small relativistic correction $\delta_{\text{rel}}$, which gives only a few percent effect for the hottest clusters (Rephaeli 1995b; Itoh et al. 1998; Nozawa et al. 2000; Fan & Wu 2003). We are interested in the integrated Y-parameter,

$$Y \equiv \int d\Omega y(\theta),$$

where the integral is over the solid angle subtended by the cluster. This can be written in terms of cluster properties, as $y(\theta)$ is the integration of pressure along the line of sight that passes at an angle $\theta$ away from the center of the cluster:

$$y(\theta) = \frac{k_B \sigma_T}{m_e c^2} \int dl n_e T_e.$$ 

Here $h$, $\sigma_T$, $k_B$, $m_e$, $n_e$, and $T_e$ are the Planck cross section, Boltzmann constant, electron mass, intracluster electron density, and temperature, respectively. The number of electrons along the line of sight to the cluster mass is taken to be

$$d^2 \int d\Omega dl n_e = \frac{M_{\text{vir}} f_{\text{gas}}}{\mu_e m_p},$$

where $d = [1/(1+1)](c/H_0) \int_0^R 1/(E(z')) dz'$, $E(z) = [\Omega_m(1+z)^3 + \Omega_{\Lambda}]^{1/2}$, $M_{\text{vir}}$ is the virial mass, $f_{\text{gas}}$ is the intracluster gas fraction $\frac{1}{12}$ [Lin et al. 2003, where $M_{\text{vir}} = M_{\text{vir}}(10^{15} h^{-1} M_\odot)$]

$$f_{\text{gas}} = \frac{0.1h^{-3/2}M_{\text{vir}}^{148}}{1 + 0.1M_{\text{vir}}^{0.25}}.$$ 

$\mu_e = 1.143$ is the mean mass per electron, and $m_p$ is the proton mass. Then the electron density-weighted average temperature

$$\langle T_e \rangle = \int dl n_e T_e \int dl n_e$$

is given, using virialization arguments (see, e.g., Battye & Weller 2003) by

$$\langle T_e \rangle = T_e \left(\frac{M_{\text{vir}}}{10^{15} h^{-1} M_\odot}\right)^{2/3} \left[\Delta_e(E(z))^{1/3} \left(1 - 2\frac{\Omega_{\Lambda}}{\Delta_e}\right)\right].$$

where $\Delta_e(z) = 18\pi^2 + 82[\Omega_m(z) - 1] - 39[\Omega_m(z) - 1]$. We have taken the z-dependence from Pierpaoli et al. (2003). Note that this assumes that only electrons within the virial radius are contributing to the SZ effect.

Although this form and a specific $T_e$ can be “derived” using virialization arguments, one can also just define $T_e$ as the constant of proportionality in the above. The above mass-temperature relation seems to work well for X-ray temperatures of high-mass clusters, and most measurements of $T_e$ for the above relation are done in the X-ray range; we call the X-ray value $T_{\text{X-ray}}$. Simulations find $T_{\text{X-ray}}$ to be $\sim 1.2$ keV, while observations tend to prefer higher values, $T_{\text{X-ray}} \sim 2.0$ keV. (For a recent compilation see, e.g., Huterer & White [2002], and for detailed discussion of subtleties in X-ray temperature definitions see, e.g., Mathiesen & Evrard [2001], and more recently, Borgani et al. [2004], Mazzotta et al. [2004], and Rasia et al. [2005] and references therein; convergence is improving steadily [A. Kravtsov 2005, private communication].) For the calculations here we need $T_{\text{SZ}}$. There is no a priori reason why we would expect the X-ray temperature and the SZ temperature normalizations to be identical, as they get the bulk of their signals from different parts of the cluster. We discuss this in more detail in §4.2 when we consider modeling uncertainties. In addition, the above is in terms of $M_{\text{vir}}$, several mass definitions are in use in the literature. If these differences are not taken into account correctly (White 2001) via mass conversion, an apparent (but incorrect) change in $T_{\text{SZ}}$ will result;

Note that for a $M_{\text{vir}} = 10^{14} h^{-1} M_\odot$ cluster we get $f_{\text{gas}} = 0.06 h^{-1/2}$. The factor of $h$ is included to make connection with other definitions; as Lin et al. (2003) point out, the variation with $h$ is not actually a simple scaling. We keep $h$ fixed in this paper. Note that measurements of $f_{\text{gas}}$ implicitly require gas physics theoretical modeling.
we discuss this issue further in § 4.2. Combining these definitions and again using $M_{15} = M_{\text{vir}}/(10^{15} \, h^{-1} \, M_\odot)$, we get

$$Y = \int d\Omega \frac{k_\beta \sigma_T}{m_e c^2} \int d\ell n_T \epsilon$$

$$= \frac{k_\beta \sigma_T f_{\text{gas}} M_{15}}{m_e c^2 \mu_\text{H}_2 m_p} \frac{1}{d_s} \frac{T_{s \text{SZ}} M_{15}^2}{Y_{\text{min}}^2} \left(1 - 2 \frac{\Omega_\Lambda}{\Delta} \right)$$

$$= 1.69 \times 10^4 f_{\text{gas}} T_{s \text{SZ}} \frac{\text{keV}}{M_{15}^2 Y_{\text{min}}^2} \left(1 - 2 \frac{\Omega_\Lambda}{\Delta} \right) \left(\frac{h^{-1} \text{Mpc}}{d_d} \right)^2 \text{arcmin}^2.$$

(8)

The resulting SZ effect is a small distortion of the CMB of the order of $\sim 1$ mK. Results are often quoted in terms of flux, with a conversion

$$F_{\nu} = 2.28 \times 10^{-4} \frac{x^4 e^x}{(e^x - 1)^2} \left(x e^x + 1 - e^x + 1 - 4\right) \frac{Y_{\text{arcmin}}}{\text{mJy}}.$$

(9)

For 143 GHz, the x-dependent factor is $\sim 4$ (which translates into $Y = -\Delta T/T$ for the Y-parameter), for 90 GHz it is $-3.3$, and for 265 GHz it is $+3.4$. The SZ effect switches from a decrement to an increment in the CMB spectrum at 218 GHz. Thus, one way of distinguishing the thermal SZ effect from other sources, such as primary anisotropy or noise, is to see whether it changes at 218 GHz.

Specifics in going from this flux or corresponding Y-value to a cluster detection depend on the particulars of each experiment. We consider the idealized case in which an SZ experiment will detect all clusters above some minimum Y-value, $Y_{\text{min}}$. Experiment-specific analysis and follow-up will be necessary to make reliable cluster identifications. Interferometer experiments will “resolve out” some of the power and thus will effectively have a higher $Y_{\text{min}}$. In addition, false clusters detected due to alignments of low-mass SZ sources will need to be discarded via some sort of follow up. The end result of this processing for an SZ cluster survey will be a catalog of clusters (with angular positions) with an SZ decrement above some minimum threshold value $Y_{\text{min}}$.

We show in Figure 1 a plot of the minimum cluster mass for a given $Y_{\text{min}}$ as a function of redshift, for some representative $Y_{\text{min}}$ values expected with APEX and SPT. As mentioned above, the minimum mass depends on the mass-temperature normalization $T_{s \text{SZ}}$, which is not well known and is discussed in our section on modeling uncertainties, § 4.2. For illustration we have taken representative values for $T_{s \text{SZ}}$ from X-ray measurements, which we might expect to be close to $T_{s \text{SZ}}$. For $Y_{\text{min}}$ we have taken the APEX’s quoted $10 \, \mu$K sensitivity and multiplied by a factor of 5, which would be a naive 5 σ detection; $f_{\text{gas}}$ is taken to be 0.10$^{+3.25}_{-1.15} M_{15}^{0.148}(1 + 0.10 M_{15}^{0.25})$ (Lin et al. 2003), and in practice it also has a scatter.\textsuperscript{14}

The slow change in $M_{\text{min}}$ with redshift is a feature of SZ selection, which in principle allows clusters of similar masses to be observed at all distances (Bartlett & Silk 1994; Barbosa et al. 1996; Holder et al. 2000; Bartlett 2002; Kneissl et al. 2001; Diaferio et al. 2003). We can easily calculate observable quantities based on $M_{\text{lim}}(Y_{\text{lim}}, z)$. This gives an advantage over X-rays where the flux dims rapidly at higher redshifts. An additional advantage of SZ over X-rays is that the SZ signal strength depends on the density, while the X-ray signal strength depends on the density squared. Thus, X-ray measurements boost the weight of the cluster core, which has poorly understood physics, in the detection. Conversely, the detection of SZ based on density means that the SZ signal is much more sensitive to line-of-sight contamination (White et al. 2002). Specifically, SZ effects are proportional to the total (hot) gas mass in the cluster along the line of sight ($\Delta T \propto \int n_e T_{\text{gas}} \, dl$). In theoretical models the dominant contributions come from the region within $\sim 0.2-0.4$ of the cluster virial radius (Komatsu & Seljak 2002). Note that at low redshifts the SZ selection probes very low mass objects, where the poorly understood gas physics dominates; as a result we take a minimum redshift cut of $z = 0.2$. Such a cut could be imposed experimentally in the follow-up.

\textsuperscript{13} One issue is the effect of beam size. For wide beams, confusion from point sources is a significant source of noise (White & Majumdar 2004; Knox et al. 2004). For small beams, several pixels must be combined to produce the total cluster signal (see, e.g., Batty & Weller 2003) above the $Y_{\text{min}}$ threshold, and there may be errors inherent to the corresponding cluster finding. These can be dealt with both in the data acquisition (e.g., by having more frequencies and an appropriate scanning strategy to help identify the point sources) and in the analysis; the effects particular to an experiment will depend strongly on the details of that experiment. An early example finding clusters in a noisy map was done by Schulz & White (2003); a more recent start-to-finish analysis of N-body simulations, including cluster finding and noise modeling, has been done for \textit{Planck} SZ clusters by Geisb"ubs et al. (2005). See also Melin et al. (2005), Pierpaoli et al. (2005), and Vale & White (2005) for more on cluster finding.

\textsuperscript{14} The scatter found by Lin et al. (2003) gives $f_{\text{gas}} = 0.10(\pm 3\%) h^{3/2} M_{15}^{0.148^{\pm 27\%}}[1 + 0.10(\pm 6\%) M_{15}^{0.25^{\pm 28\%}}]^{-1}$.  

### 3. ANALYTIC CALCULATIONS

In order to go from this cluster catalog to cosmologically useful mass counts, theoretical processing and assumptions are needed. In this section we review and set the notation for the angular power spectrum/correlation function in terms of analytic quantities (which will be varied in the next section).
Theoretical inputs to the analytic calculations include the choice of mass function, transfer function, biasing scheme, mass-temperature relations (in particular $T^2$), initial power spectrum, and, of course, cosmological parameters. Some of these, such as the mass function, are well tested; other quantities, such as $T^2$, the mass-temperature normalization appropriate for calculating the $Y$-parameter, are not determined well at all, either theoretically or observationally, and widely varying approximations are in use. We compare these approximations in the following. The angular power spectrum/correlation function can be calculated with analytic prescriptions for the mass function $dn/dM(z)$ and the bias relating the dark matter correlations to the correlations for the galaxy clusters. The two-dimensional correlation function is (Moscardini et al. 2002; Diaferio et al. 2003; Mei & Bartlett 2003)

$$w(\theta) = \left[ \int_0^\infty r_1^2 dr_1 \int_0^\infty r_2^2 dr_2 \int_{\text{Max}(\gamma_{\text{min}})} \text{d}M_1 \right] \times \left[ \int_{\text{Max}(\gamma_{\text{min}})} \text{d}M_2 \phi(M_1, r_1) \phi(M_2, r_2) \xi_0(M_1, M_2, r_1 - r_2) \right] \times \left[ \int_0^\infty r^2 \int_{\text{Max}(\gamma_{\text{min}})} \text{d}M \phi(M, r) \text{d}r \right]^{-2}, \quad (10)$$

where $r_1$ and $r_2$ are the radial distances of the clusters, which have three-dimensional positions $r_1, r_2$ and $r_1 - r_2 = r_1 r_2 \cos \theta$; $\xi(M_1, M_2, r_1 - r_2)$ is the three-dimensional cluster correlation function, and $\phi(M, r)$ is the selection function (as a function of radial distance). In particular,

$$\phi(M, r) = \frac{dn}{dM} [z(r)],$$

$$\xi(M_1, M_2, r_1 - r_2) = b(p_{\text{col}}(r_1) b(M_2, z_2(r_2)) \xi_{\text{dim}}(r_1 - r_2), \quad \xi_{\text{dim}}(r_1 - r_2) = \int \Delta^2(k) \sin kr \text{d}k, \quad \Delta^2(k) = \frac{V}{(2\pi)^2} 4\pi k^3 P(k), \quad P_{\text{lin}}(k) = D^2(z) P_k \sigma_8^2 T^2(k). \quad (11)$$

In the above, the selection function is the number of clusters as a function of redshift (with the normalization included explicitly). The linear power spectrum $P_{\text{lin}}$ comes from an initial power spectrum with slope $n$, normalization $P_0$ (implied by the choice of $\sigma_8$), and transfer function $T^2(k)$. The nonlinear power spectrum is derived from $P_{\text{lin}}$ using the method of Smith et al. (2003). The linear bias $b$ is defined above, and the nonlinear bias is considered in the next section.

The full two-dimensional correlation function thus becomes

$$w(\theta) = \int_0^\infty r_1^2 dr_1 \int_0^\infty r_2^2 dr_2 \int_{\text{Max}(\gamma_{\text{min}})} \text{d}M_1 \frac{dn}{dM} [z(r_1)] \times \left[ \int_{\text{Max}(\gamma_{\text{min}})} \text{d}M_2 \frac{dn}{dM} \frac{[z(r_2)] b(M_1, z_1(r_1)) b(M_2, z_2(r_2)) \xi_{\text{dim}}(r_1 - r_2)]}{[z(r_2)]} \right] \times \left[ \int_0^\infty r^2 \int_{\text{Max}(\gamma_{\text{min}})} \text{d}M \frac{dn}{dM} [z(r)] \text{d}r \right]^{-2} \times \left[ \int_0^\infty r_1^2 \Phi(r_1) \int_0^\infty r_2^2 dr_2 \Phi(r_2) \xi_{\text{dim}}(r_1 - r_2) \right], \quad (10)$$

where we have defined a generalized selection function (Moscardini et al. 2002; Mei & Bartlett 2003)$^{15}$

$$\Phi(r) = \int_{m_{\text{min}}}^{\infty} \text{d}M \frac{dn}{dM} [z(r)] b(M, z(r)) \times \left[ \int_0^\infty r^2 \int_{m_{\text{min}}}^{\infty} \text{d}M \frac{dn}{dM} [z(r)] \text{d}r \right]^{-1}. \quad (13)$$

The above angular correlation function (and corresponding power spectrum) can be simplified via the Limber (1953) approximation. As $\Phi$ varies slowly relative to the correlation function, the integration can be rewritten as an integration over an average distance $y (r_1 + r_2)/2$ and one over relative separations $x = (r_1 - r_2)$. The integration over $x$ then gives a Bessel function $J_0$:

$$w(\theta) \sim \int dy y^4 \Phi(y)^2 \int dx \xi_{\text{dim}}(\sqrt{y^2 \theta^2 + x^2}) = \int dy y^4 \Phi(y)^2 \int d\ln k \pi J_0(k y \theta) \frac{\Delta^2(k)}{k}. \quad (14)$$

Then the power spectrum

$$C_l = 2\pi \int dy y^5 \Phi(y)^2 \frac{\Delta^2(l/y)}{l^3} \quad (15)$$

can be read off in the small-angle approximation

$$w(\theta) = \frac{1}{4\pi} \sum_{l=-l}^{l} \sum_{m=-l}^{l} \left| d_{l,m}^l \right|^2 P_l(\cos \theta) = \sum_{l} \frac{2l + 1}{4\pi} C_l P_l(\cos \theta) \approx \int d\l \frac{l}{\pi} C_{\theta} J_0(l \theta) \quad (16)$$

because $P_l(\cos \theta) \approx J_0(l \theta)$ for small angles. We can also define the inverse via

$$C_l = 2\pi \int_0^{\infty} w(\theta) J_0(l \theta) \theta d\theta. \quad (17)$$

The correlation function $w(\theta)$ and its power spectrum $C_l$ can be transformed to each other by the above equation, and therefore they encode the same information. However, to understand possible measurements and errors, the use of $C_l$ is usually preferable because the errors for different $l$-values are uncorrelated for small $l$. We use $C_l$ for the most part in the following.

For reference, we show in Figure 2 the angular power spectrum for a representative model. The dependence on the possible reasonable choices for the analytic and cosmological model parameters is the subject of a later section. The choices taken here will be our “vanilla” model: $^{16}$ for the dark matter we use the Evrard mass function, the Sheth-Tormen bias, and the Eisenstein-Hu transfer function; for cluster parameters we take

$^{15}$ Note that our selection function differs from that in Mei & Bartlett (2003) by the factor $r^2 dr/dz$; their expression for $w(\theta)$ is equivalent.

$^{16}$ Note that this is not identical to the vanilla model of Tegmark et al. (2004); we merely use the term to denote a model that we take to be the simplest in some rough sense. For example, our vanilla model includes a choice of $T^2$, and of mass function.
scatter in the relation of minimum mass and $Y$ (White 2001).

Mass conversions and the origin transfer function. We also differ from Mei & Bartlett in using the Eisenstein-Hu transfer function. The basic expression above was used by Diaferio et al. (2003). Figure 2 has the same axes for (almost) all the plots, so that the relative impact of different effects are easily comparable. For the vanilla parameters one gets about 10 clusters per square degree.

We have made some changes from earlier similar works in this review section. The basic expression above was used by Mei & Bartlett (2003), and a variant (to be discussed below) was derived and used by Diaferio et al. (2003). Figure 2 has the following changes from this earlier work: it gives the power spectrum rather than the correlation function, implements mass conversions (which can change masses by 30%), and includes a 10% scatter in the relation of minimum mass and $Y_{\text{min}}$ due to the cluster parameter uncertainties. Mass conversions and the origin for the amount of scatter that we have chosen are discussed in § 4. We also differ from Mei & Bartlett in using the Eisenstein-Hu transfer function.

4. UNCERTAINTIES

We now compare the effects of cosmological, modeling, and observational uncertainties/unknowns on the cluster angular power spectrum. Some aspects of these, with different assumptions, have been considered previously for the cluster angular correlation function: Diaferio et al. (2003) consider two cosmological models and vary $Y_{\text{min}}$ and the bias, and Mei & Bartlett (2003, 2004) vary $Y_{\text{min}}$, $\sigma_g$, $\Omega_m$, and $T^S_{\text{SZ}}$. For the spatial power spectrum in a Fisher matrix analysis, Wang et al. (2004) vary these and the primordial fluctuation spectrum, the dark energy density and equation of state, the baryon density. Both they and Mei & Bartlett (2003, 2004) allow extra $(1 + z)$ and $M$ factors to appear in the $Y(M)$ relation. Mei & Bartlett find a small effect; note that this is at large (relative to our work here) values of the $Y$-parameter, corresponding to large masses, where the gas physics is not as important. Majumdar & Mohr (2003, 2004) vary most of these factors, as well, in finding their constraints. Except for the Mei & Bartlett (2004) paper, which considers APEX, these other experiments are primarily concerned with experiments in the far future such as SPT.

In this work we consider other theoretical uncertainties, such as changing the mass function and the gas fraction. Another new aspect of this work is that we compare all these recognized uncertainties to each other, which helps identify which modeling and observational uncertainties need to be reduced the most. In addition, as mentioned earlier, we are primarily concerned with the angular power spectrum, as in Majumdar & Mohr (2004) and Mei & Bartlett (2004), rather than the correlation function. Unlike the angular power spectrum in the Gaussian regime, the errors in the correlation function $w(\theta)$ are correlated, which makes it more difficult to understand how well measurements at given separations can determine various quantities.

4.1. Cosmological Model Dependence

We start by showing what changes to the cosmological model do to the vanilla model, for later comparison with the modeling and experimental uncertainties. For instance, the current published joint Wilkinson Microwave Anisotropy Probe/Sloan Digital Sky Survey cosmological parameters and errors are $\Omega_m = 0.30 \pm 0.04$ and $\sigma_g = 0.86^{+0.11}_{-0.10}$ (Tegmark et al. 2004). We compare our vanilla model with changes in $\Omega_m$ and $\sigma_g$, in Figure 3.

The correlation function decreases with increasing $\sigma_g$, as a higher $\sigma_g$ means that the generalized selection function broadens and moves its peak to higher redshift. The broadening gives more noncorrelated clusters nearby any given cluster, and in addition, the biasing is weaker for high $\sigma_g$. The

17 Mei & Bartlett (2003) have illustrated this effect in their paper.
18 We thank the referee for emphasizing the second point.

Fig. 3.—Dependence of the power spectrum on varying $\sigma_g$ and $\Omega_m$ separately. The solid line shows our reference vanilla model with $\sigma_g = 0.9$ and $\Omega_m = 0.3$. The long-dashed line is for $\sigma_g = 0.8$, the short-dashed–long-dashed line is for $\sigma_g = 1.0$, and the dotted line is for $\Omega_m = 0.25$. [See the electronic edition of the Journal for a color version of this figure.]
cluster power spectrum scales quite differently with $\sigma_8$ than the temperature correlation function. If one parameterizes $C_l$ as

$$C_l(\sigma_8) = \sigma_8^{\alpha(l)} f(l),$$  \hspace{1cm} (18)$$

for $Y_{\text{min}} = 1.7 \times 10^{-3}$ and $\sigma_8 = 0.9$, $\alpha(l)$ ranges from around 2.7 for $l \sim 100$ to about 1.6 for $l \sim 6000$. Doubling $Y_{\text{min}}$ changes this range to 3.5 and 1.7, respectively (note the negative power of $\sigma_8$, which is different from the positive power scaling behavior of the SZ temperature power spectrum; Komatsu & Seljak 2001; Sadeh & Rephaeli 2004).

In fact, the effect of changing $\sigma_8$ is somewhat smaller, as there are other constraints that must be satisfied when changing $\sigma_8$, for instance, the observed and hence fixed number counts of clusters. One can include this constraint by requiring the number of clusters to remain fixed along with $\Omega_m$, for instance, and find the required scaling between $T_{50}^{\text{SZ}}$ and $\sigma_8$ (Sadeh & Rephaeli 2004). Allowing $\Omega_m$ to vary still gives a constraint using arguments identical to those of Huterer & White (2002; see also Evrard et al. 2002), who were considering the X-ray temperature. Fixing the number of clusters with a given $Y$ parameter (an observable) and allowing $T_{50}^{\text{SZ}}$, $\sigma_8$, and $\Omega_m$ to vary gives the following (directly analogous to X-ray) scaling relation:

$$T_{50}^{\text{SZ}} \sim (\sigma_8 \Omega_m^{0.6})^{-1.1}. \hspace{1cm} (19)$$

More precisely, it will be $f_{\text{gas}} T_{50}^{\text{SZ}}$ on the left. The modeling parameter $T_{50}^{\text{SZ}}$ is discussed in detail in the next section. This relation already suggests a strong degeneracy between the effects of changing $\sigma_8$, $T_{50}^{\text{SZ}}$, and $\Omega_m$ (see § 4.3 for more discussion on degeneracies).

4.2 Modeling Uncertainties

There are several parameters and functions that go into the theoretical predictions besides the cosmological parameters. These can be divided into those related to cluster properties independent of the SZ effect and those related to the transformation from cluster mass to the observable $Y$ (eq. [8]). We treat both of these in turn. For an SZ-selected survey, uncertainties in modeling affect results by moving objects into or out of the survey. As a result, uncertainties that have a large effect on the observational properties of the lowest mass clusters in the survey ($M_{200} \sim 10^{14} h^{-1} M_{\odot}$) have the most impact.

Dark matter cluster properties.—The cluster properties independent of the SZ effect used in the analytic prediction of $C_l$ (eq. [15]) are the correlation function of the dark matter, the mass function, and the bias.

Dark matter correlation function.—The dark matter correlation function is quite well known and can be obtained from the initial power spectrum via a transfer function (as in Bardeen et al. [1986, hereafter BBKS] or Eisenstein & Hu [1999]) and then by implementing a nonlinear power spectrum prescription (as in Peacock & Dodds [1996] or Smith et al. [2003]). The vanilla model uses the more accurate recent fits, i.e., the Smith et al. nonlinear power spectrum fit and the Eisenstein & Hu transfer function (the latter is within 3% of the exact CMB power spectrum; M. White 2004, private communication). We compare the case with the BBKS transfer function to the vanilla model in Figure 5, as it is still in use by many groups.

Mass conversions.—Before defining the mass functions, it should be recalled that there are several different mass definitions in use. For SZ calculations, the mass-temperature relation usually involves the virial mass, but the popular mass functions usually are instead for some linking length or parameter $\Delta$ such that the mass inside a radius $r_\Delta$ is $\Delta$ times the critical density:

$$M_\Delta = \frac{4\pi}{3} \Delta \rho_{\text{crit}} r_\Delta^3 = \int_0^{r_\Delta} 4\pi r^2 \, dr \rho(r). \hspace{1cm} (20)$$

This mass can differ significantly from the virial mass (White 2001), which enters into the definition of $Y$; and a conversion must be made between the virial mass and the mass appearing in the mass function. The way suggested in White (2001) is to start by assuming an NFW (Navarro et al. 1997) mass profile (a fitting formula for this method is given in Hu & Kravtsov 2003). For instance, in an $\Omega_m = 0.3$ universe, $M_{\text{crit}}/M_{200} \sim 1.3$ for a cluster with a concentration of 5. A 30% difference in mass is significant when one is attempting to make percentage-level predictions! Masses enter the calculations in the $Y(M)$ relation, the mass function, and the bias; consistent definitions or the relevant conversions are needed. We show in Figure 5 the effect on the vanilla model of using $M_{200}$ rather than the appropriate masses in the $Y(M)$ relation and bias.

Mass functions.—In Figure 4 we show three commonly used mass functions; in Figure 5 we compare the effects on the correlation function and angular power spectrum of these three choices, as well as the effects of using a less accurate transfer function and neglecting the mass conversion.

The three mass functions taken here can be viewed as generalizations of the heuristic mass function of Press & Schechter (1974). These are based on simulations with large enough volume to accurately sample the number of rare objects such as galaxy clusters: the Jenkins et al. (2001) mass function, the Sheth-Tormen (1999) mass function, and the Evrard et al. (2002) mass function. The first two are for masses in terms of $M_{180k}$, while the last is in terms of $M_{200}$. These mass functions are expressed in terms of $\sigma(M, z)$ or $\nu(M, z)$. Here $\sigma(M, z)$ is the rms of the mass density field smoothed on a scale $R = 3M/(4\pi\rho_0)^{1/3}$ with

\[^{19}\text{Note that some people use } \Delta \text{ to refer to the density relative to the mean density.}\]
Using the Peacock & Dodds (1996) nonlinear prescription a top-hat window function, \( \sigma(M, z) = \int_0^\infty (dk/k) \Delta^2(k) \Psi^2_M(k, z) \), where \( \Psi^2_M(k, z) \) is the Fourier transform of the window function, \( \nu = \delta_c/\sigma \), and \( \rho_b = \Omega_m \rho_{crit} \). We ignore the weak cosmological dependence of \( \delta_c \) and set \( \delta_c = 1.686 \). We then can write

\[
\frac{dn}{dM} = \frac{\rho_m(z)}{M} f(\ln \sigma^{-1}) \frac{d\ln \sigma^{-1}}{dM}
\]

(21)

to get the three mass functions described in Table 1.

The mass density field and the correlation function are derived from the primordial power spectrum, which we take to be scale-free, i.e., \( n = 1 \), and normalized by \( \sigma_8 \). To get a sense of the effects, the (less accurate) earlier BBKS transfer function decreases \( C_l \) by 5% at \( l = 214 \), neglecting the mass conversion between \( M_{200} \) and \( M_{vir} \) decreases \( C_{214} \) by 11%, and the Sheth-Tormen mass function (Sheth & Tormen 1999) decreases it by 19%. Using the Peacock & Dodds (1996) nonlinear prescription gives no noticeable change; hence, we did not show it in the figure. An extensive comparison of different cases is given in Table 3.

Bias.—There are also several different possibilities for bias. The (linear) bias \( b(M, z) \) is defined via

\[
\xi(M, M, r, z) = b^2(M, z) \xi(r, z),
\]

(22)

where \( \xi(r, z) \) is the dark matter power spectrum and \( \xi(M, M, r, z) \) is the power spectrum of halos of mass \( M \). The original idea of peak biasing by Kaiser (1984) has been improved on with fits to simulations. Table 2 shows the Sheth-Tormen (1999) bias, fit to simulations and motivated by a moving wall argument; the bias found by Sheth et al. (2001, hereafter SMT); and the bias more recently found by Seljak & Warren (2004). The Seljak & Warren bias was found for small masses but has the best statistics currently available. It overlaps closely with the Sheth-Tormen bias where it is valid but is systematically lower and does not extend very far into the high-mass range needed for clusters. Thus, using a combination of the two biases would result in a bias that does not integrate to 1 when combined with the Sheth-Tormen mass function. Consequently, we have taken the Sheth-Tormen bias as our default.

For the SZ-selected power spectrum one integrates over all masses greater than some \( M_{min} \), so that what one is actually probing is an integral of the bias, i.e., \( \Phi(Y_{min}, z) \) in equation (13). One can define a related (rescaled by the number density) quantity:

\[
b_{\text{method, eff}} = \frac{\int_{M_{\text{min}}}^{\infty} dM [dn/\langle dM \rangle] [z(r)] b_{\text{method}}(M, z(r))}{[dn/\langle dM \rangle] [z(r)]},
\]

(23)

where \( b_{\text{method}} \) is one of the above biases. The linear biasing prescription above does not work as well for short distances, and for this a “scale-dependent bias” has been calculated for the cluster correlation functions by Hamana et al. (2001) and Diaferio et al. (2003). This scale-dependent bias is also a function of the separation \( r \) of the objects of interest and for Diaferio et al. is

\[
b_{\text{eff}}(r, z, Y_{\text{min}}) = b_{\text{ST, eff}}(z, Y_{\text{min}}) [1 + b_{\text{ST, eff}}(z, Y_{\text{min}}) \sigma(r, z)]^{0.35},
\]

(24)

the corresponding expression for Hamana et al. has an exponent 0.15. Diaferio et al. have shown that this bias works well for cluster correlation functions for a range of redshifts. We use the

![Fig. 5.—Left: Angular correlation function with same models and line labeling as in Fig. 4, plus the vanilla model except for the BBKS transfer function (dotted line) and vanilla model without using mass function conversion for mass-temperature (thick dotted line). Right: Corresponding angular power spectra. [See the electronic edition of the Journal for a color version of this figure.]]
Diaferio et al. case for illustration. The bias of Hamana et al. is midway between the linear biasing case and the Diaferio et al. (2003) case. With the Diaferio et al. bias, the correlation function

\[ w(\theta) = \int_0^\infty r_1^2 \Phi(r_1) \int_0^\infty r_2^2 dr_2 \Phi(r_2) \tilde{\xi}_{dm}(|r_1 - r_2|), \]

(25)

where

\[ \tilde{\xi}_{dm}(|r_1 - r_2|) = \left[ 1 + b_{\text{ST,eff}}(z, Y_{\text{min}}) \sigma(r, z) \right]^{0.70} \xi_{dm}(|r_1 - r_2|). \]

(26)

In the Limber approximation one then finds

\[ w(\theta) = \int dy y^A \Phi(y) \int dx \left[ 1 + b_{\text{ST,eff}}(z, Y_{\text{min}}) \right] \sigma \left( \sqrt{y^2 \theta^2 + x^2}, z \right) \xi_{dm} \left( \sqrt{y^2 \theta^2 + x^2} \right). \]

(27)

As this does not easily allow a rewriting in terms of \( J_0(\theta) \), obtaining the power spectrum \( C_\ell \) values is somewhat more difficult. In Figure 6 we show the correlation function \( w(\theta) \) at left and the power spectrum \( C_\ell \) at right for the vanilla model and then the SMT linear bias and the Diaferio et al. nonlinear bias. (Although the Hamana et al. bias is also between the Diaferio et al. bias and the vanilla bias, it is not what is shown in this figure.) We transformed the difference of \( w(\theta) \) values between the vanilla model and Diaferio et al. model to get the \( C_\ell \) values for the former. We use the vanilla model parameters for the Diazferio et al. bias (e.g., \( T^{25}_s = 1.2 \) rather than \( T^{25}_s = 2.0 \), as they did).

The nonlinear bias has the strongest effect at short distances in the correlation function—its effect is strongly localized around \( \theta \approx 0 \). Therefore, its Fourier transform, the angular power spectrum, has additional contributions of almost constant magnitude at all \( l \) relative to the vanilla model. (The limit of this would be adding power at only \( \theta = 0 \), which translates into adding a constant to the power spectrum. Some of this difference is deceptive, as the errors for \( C_\ell \) are independent in the linear regime, while those for \( w(\theta) \) combine the \( w(\theta) \) measurements for different values of \( \theta \). For a visual comparison, the independence of the errors for \( C_\ell \) makes it easier to draw conclusions about relations between different parameter choices and uncertainties. In addition, even though the power spectrum and the correlation

![Table 2: Bias Function](image)

| Bias Function | \( \Delta \rho_{\text{eff}} \) | Bias |
|---------------|----------------|------|
| Sheth-Tormen (1999) | 180 \( \rho_b \) | \( 1 + \frac{1}{\delta_c(\nu/2 - 1)} + \frac{2p}{\delta_c(1 + \nu/2)} \) |
| SMT | 180 \( \rho_b \) | \( 1 + \frac{1}{\delta_c} \left[ \nu^2 + b(1 - c)(1 - \nu/2) \right] \) |
| Seljak & Warren (2004) | 200 \( \rho_{\text{eff}} \) | \( 0.53 + 0.39 \sqrt{0.42 + \frac{0.08}{40x + 1} + 10^{-4}x^{1.7}} \), \( x = M/M_{\text{crit}} \) |

![Fig. 6.—Effect of different bias prescriptions on the correlation function \( w(\theta) \) (left) and the power spectrum \( C_\ell \) (right). The vanilla reference model is shown in both cases (solid line), as well as the cases with the SMT bias in eq. (2) (dotted line) and the nonlinear bias of Diaferio et al. in eq. (25) (dashed line). [See the electronic edition of the Journal for a color version of this figure.]](image)
function are Fourier transform pairs, estimators for these differ in practice when real data is in hand and give different information when one does not have full $2\pi$ angular coverage in $\theta$. Ideally, one would use both given real data.

There is also intrinsic scatter around the bias and the mass functions. We do not put this in explicitly; however, the weak dependence on the scatter in the $Y(M)$ relation (mentioned below) leads us to suspect that it will not be a large effect.

$Y$-parameter.—The next step is relating the cluster mass to a $Y$-parameter, which involves more complicated gas physics. In addition, the increased sensitivities of upcoming experiments will allow smaller and smaller mass clusters to be detected, which are more and more easily disrupted by this gas physics.

There are actually three questions: the actual form of the mass-temperature/$Y$-parameter relation, the normalization of this relation (i.e., $T_*^{SZ}f_{\text{gas}}$), and the scatter around this normalization for a representative group of galaxy clusters. Simulations alone cannot determine these: the heating and cooling properties of clusters are not understood at an accuracy needed for precision cosmology, so these questions are intermingled by the assumptions used. For instance, the scaling relation was obtained by assuming an isothermal gas profile. Assuming hydrostatic equilibrium and using the total $Y$-parameter means that the details of the profile (many others have been suggested, e.g., Komatsu & Seljak [2001] and Loken et al. [2002]) get absorbed into the mass-temperature (or $Y$-parameter) normalization or form. If the parameters of the gas profile change with cosmology, it is possible that the normalization will also do so, or rather the form of the $Y(M)$ relation, a possibility that will need to be checked for carefully in the data. We consider the form of the mass-temperature/$Y$-parameter relation, the normalization, and the scatter in turn.

Mass-temperature/$Y$-parameter relation.—There are several different mass-temperature relations in the literature (see Sadeh & Rephaeli [2004] for a description of five common ones), usually based on X-ray mass-temperature relations. These relations have been tested with both observations and simulations (bear in mind again that the simulations do not seem to yet have all the necessary physics), and some generalizations of these relations have also been tested. For instance, equation (8) can be generalized to include different $z$ and $M$ dependence, such as multiplying by a factor of $(1+z)^{\gamma}M^\alpha$. The mass–$Y$-parameter relation also has dependence on $f_{\text{gas}}$, which can be generalized to change $f_{\text{gas}}$ with redshift or change it differently with mass. Observational data and simulation data have been used to search for these effects.

Most observational tests are of the X-ray mass-temperature relations and of the change of $f_{\text{gas}}$ with mass or redshift, rather than of the $Y(M)$ relation. For example, Ettori et al. (2004a) have found no evidence for additional evolution in redshift of the mass-temperature relation. For scaling of mass with temperature, Ettori et al. (2004b) and Ota & Mitsuda (2004) find $T_*^{X-ray} \sim M^{2/3} - 0.5$, but the departure from the $M^{2/3}$ relation is about 1.5 $\sigma$ for the Ettori et al. data and is marginal given the error bars for the Ota & Mitsuda data. The former group also finds marginal (less than 2 $\sigma$) evidence for clusters of a fixed temperature to have smaller gas mass at high redshift. We have already included the $f_{\text{gas}}$ dependence with mass found by Lin et al. (2003) but have not included any redshift dependence.

More simulations than observations have addressed the $z$ and $M$ dependences of the $Y(M)$ relation directly. For example, da Silva et al. (2004) find numerically that the $z$-dependence seems to be well represented by the simple scaling given in equation (8).

For lower mass objects da Silva et al. (2004) find that the $M$ dependence in the $Y(M)$ relation steepens from $M^{0.3}$. Our use of an $M$-dependent $f_{\text{gas}}$ produces an effect in the same direction. However, for clusters with $M_{500} > 5 \times 10^{13}$ $h^{-1}$ $M_\odot$ (or $M_{\text{vir}} > 6.5 \times 10^{13}$ $h^{-1}$ $M_\odot$), one finds that this can just be absorbed into a scatter of about 10% around the $M^{0.3}$ scaling (White et al. 2002) in the $Y(M)$ relation. The effects of adding an additional factor of $(1+z)^{\gamma}M^\alpha$ to $Y(M)$ have been considered for the angular correlation function by Moscardini et al. (2002) and Mei & Bartlett (2003), and Wang et al. (2004) and Majumdar & Mohr (2003, 2004) have considered the effects of this additional scaling on parameter estimation from the power spectrum (the three-dimensional one in the former case).

Normalizing of $Y(M)$ relation.—For normalization, we have combined all the mass-temperature conversion ignorance into the parameter combination $Y \propto T_*^{SZ}f_{\text{gas}}$. The simplest procedure would be to say that $Y = Y_{\text{vir}}$ (and that gas outside this radius does not contribute significantly) and to take $T_*^{SZ}$ to be the X-ray value,

$$T_*^{SZ} \equiv T_*^{X-ray}. \tag{28}$$

As noted in § 2, there is strong disagreement between simulations and observations for $T_*^{X-ray}$. Thus, the parameter $T_*^{SZ}$ is not well determined. Even if it were, using it to normalize the $Y(M)$ relation is not necessarily justified. Not only do X-ray measurements weight the center of the cluster more strongly, as mentioned before, but for the SZ effect the normalization has an additional contribution due to line-of-sight contamination from gas outside the cluster. If one takes simulation results (which should be taken with a grain of salt given the above-mentioned discrepancy), this projection effect on $Y(M)$ raises the normalization about 8% (White et al. 2002) above the normalization due to the cluster alone.

The differences between power spectra for different normalizations of the $Y(M)$ relation are shown in Figure 7. We also show that taking $f_{\text{gas}}$ fixed at $0.06$ $h^{-1}$ (the value for a cluster with $M_{\text{vir}} = 10^{14}$ $h^{-1}$ $M_\odot$) gives a power spectrum very close to the vanilla model. And we additionally have shown a model with $T_*^{SZ} = 2.2$ keV, which has the 8% increase from line-of-sight projection (from $T_*^{SZ} = 2$ keV). Of course, as changing $T_*^{SZ}$ just rescales the $Y$-parameter, raising $T_*^{SZ}$ is equivalent to lowering $Y_{\text{min}}$; i.e., one is probing clusters with smaller mass.

Scatter.—Unlike the normalization, the effect of scatter on the $Y(M)$ relation was very weak. Taking 10%–12% intrinsic scatter in the mass-temperature relation (from Evrard et al. [1996] for X-ray simulations and similarly from White et al. [2002] for SZ simulations) had less than 1% effect on the $C_i$, even with APEX sensitivity and accompanying very low mass cuts. For a large scatter of about 30%, $C_i$ was roughly decreased by about 3%. A similar robustness to $M$-$T$ scatter was found in the Fisher matrix calculations (Levine et al. 2002) and in that of number counts (Battye & Weller 2003). Metzler (1998) also found from simulations that the scatter in the mass-temperature relation was larger than that for the mass–$Y$-parameter relation. Thus, for the bulk of the paper, we have used equation (8) and combined all our ignorance into the parameters $T_*^{SZ}f_{\text{gas}}$. We included the unnoticeable 10% scatter in the $Y(M)$ relation in all our calculations here.

We were concerned about sources of scatter that are not included in our analytic description, mergers in particular. One might expect that processes such as merging will disrupt the
clusters and thus invalidate the assumption of virialization used in some analytic calculations. The most massive clusters have the most recent mergers (as they are generally the most recently formed objects), but these tend to be included automatically in the catalog, as their estimated masses, even if inaccurate, are quite high relative to the mass cut. The selection for the catalog depends most sensitively on the least massive clusters included, where mergers are relatively rarer. (However, at high redshift the “low mass” clusters are recently formed, as they are the most massive collapsed objects at that time, so one might expect some effect from them.) Mergers are automatically included in the cosmological simulations, and the scatter in the $Y$-$M$ relation is not larger than that expected due to scatter in the $M$-$T$ relation (White et al. 2002), leading us to expect that merger-induced scatter is relatively small. However, as simulations cannot reproduce cluster properties precisely yet, observational data will be needed to calibrate this effect.

4.3. Degeneracies

We have considered several different choices for theoretical inputs to the calculations of the angular power spectrum. With a brief visual inspection it can be seen that there are many degeneracies, i.e., that several changes to parameters and modeling seem to have the same effect on the power spectrum. Note that this does not take into account other constraints at the same time, such as fixed number counts; only one parameter is varied at a time. We show some of these degeneracies more explicitly in Figure 8; there are other degeneracies in the parameters considered here that are not shown. For instance, choosing a constant $f_{\text{gas}}$ rather than having evolution with mass, as done in the previous subsection, is degenerate with changing $T_{\text{SZ}}$ from 1.2 to 1.3 keV; using the Sheth-Tormen mass function rather than the Evrard mass function is degenerate with changing $T_{\text{SZ}}$ from 1.2 to 1.3 keV; and neglecting the mass conversion from $M_{200}$ to $M_{\text{vir}}$ and $M_{180/\beta}$ in the $Y(M)$ relation and bias.

---

20 We thank the referee for encouraging us to add more discussion on this issue.
respectively, is roughly degenerate with rescaling $C_l$ overall by a factor of 0.88.

This can be compared to changing $\Omega_m$ or using the nonlinear bias of Diaferio et al. or using a less accurate transfer function such as BBKS, all of which change the “shape” of $C_l$ differently from those above. These are shown in Figure 8 (right). Values of $C_l$ for these models are compared quantitatively in Table 3.

The fact that $\Omega_m$ is not as degenerate with $\sigma_8$ as expected shows the limitations of the rough scaling estimate made in § 4.1. Some of these degeneracies can be broken by other measurements, for instance, $dN/dz$; however, the degeneracies are very similar, except for the bias. The bias has no effect on $dN/dz$ and thus can be taken out easily. In Figure 9, we show $dN/dz$ for the same models considered for Figure 8 (note that we are considering a full steradian).

As we have shown how choices such as the mass function and different biases are degenerate with other parameter choices (in most cases), we have demonstrated that the one is only as well measured as the other is known. In order to see how close these power spectra are in practice, we now compare to the inherent statistical measurement error.

4.4. Cosmic Variance, Sample Variance, and Shot Noise

There are three sources of inherent statistical measurement error for the power spectrum $C_l$ in the absence of any systematic errors: shot noise, cosmic variance, and sample variance. These can be combined to give the standard expression for overall error (see, e.g., Knox 1995):\footnote{Here the shot noise is considered in the Gaussian limit, for the full Poisson errors for the shot noise, which can be important (see Cohn 2005). In addition, there are corrections to the error due to the three- and four-point functions of galaxy clusters, which are usually not included, and we do not include them here.}

$$\delta C_l = \sqrt{\frac{2}{(2l + 1) f_{\text{sky}}}} \left( C_l + \frac{1}{N} \right).$$ (29)

The factor of $2l + 1$ is due to the $2l + 1$ independent measurements of the power for any $l$. For small $l$ (large scales) there are very few independent measurements of power in the sky, which is dubbed cosmic variance. Sample variance increases the error as the sky coverage $f_{\text{sky}} \leq 1$ decreases. Shot noise is determined by $N$, the number density of clusters per steradian.

As the depth of the survey goes up ($Y_{\min}$ decreasing), the power spectrum also decreases, as there are more and more clusters of lower and lower mass, and these are less correlated. However, the shot noise also goes down. On the other hand, as the depth of the survey decreases ($Y_{\min}$ increasing), the power spectrum becomes restricted to higher and higher mass objects and thus goes up. However, as these objects are rarer, the shot noise also increases. We can compare the error quantities for three examples, APEX, SPT, and Planck.

![Fig. 9.—Degeneracies in theoretical modeling for $dN/dz$: labeling is the same as in Fig. 8 (left), except that the upper and lower groups of lines have switched. In addition, changing the bias to that of SMT has no effect on $dN/dz$, so that line is now degenerate with the vanilla case. [See the electronic edition of the Journal for a color version of this figure.]](image1)

![Fig. 10.—Angular power spectrum with errors from shot noise, sampling, and cosmic variance error expected from Planck, for $Y_{\min} = 10^{-4}$. The bin size is the spacing between the error bars, equal spacing in log $l$. See text for more details.]
We use the vanilla model and again take $Y_{\text{min}} = 1.7 \times 10^{-5}$, corresponding to a $5\sigma$ detection for APEX with $\delta T \sim 10\mu K$ (at 150 GHz). APEX will survey 100/200 deg$^2$ at two frequencies (214 and 150 GHz, corresponding to 1.4 and 2 mm wavelengths), with 0.75 resolution. For SPT, we expect about 4000 deg$^2$ and similar sensitivity and resolution. For Planck, we have the entire sky and will take a rough estimate of $Y_{\text{min}} = 10^{-4}$ (sensitivity/resolution ranging from 5 $\mu K$ at 7'0 to 50 $\mu K$ at 33', depending on frequency). First we plot the errors due to shot noise, sampling, and cosmic variance for the Planck specifications above in Fig. 10.

In comparison, those for APEX and SPT are shown in Figure 11. The vanilla model has the same parameters as earlier (including $Y_{\text{min}} = 1.7 \times 10^{-5}$), and error bars are shown for 100, 200, and 4000 deg$^2$, representative areas for data sets expected from APEX and SPT. The largest error bars are for the smallest area. The plot on the top right has the same observing area, but with a minimum $Y_{\text{min}}$ value of $10^{-3}$. This might occur if, e.g., $T_S^Z$ were 2.0 keV rather than 1.2, a reasonable possibility given X-ray measurements and line-of-sight contamination effects. The sensitivity is directly proportional to $Y_{\text{min}}$, so “in principle” an experiment can fix area/$Y_{\text{min}}^2$ for a given observing time, producing a trade-off between wide and shallow or narrow and deep. The errors given in equation (29) are shown for the vanilla model correlation function and then compared to other possibilities with fixed observing time. Note that this ignores how the efficiency and difficulty of cluster identification changes with $Y_{\text{min}}$. The errors also neglect the Poisson nature of the shot noise and the three- and four-point functions of the clusters; thus, we call them “naive.” Figure 11 (bottom two graphs) shows the vanilla model plotted with a smaller $Y_{\text{min}}$ with error bars corresponding to the 100 and 200 deg$^2$ surveys with area up by the factor of 4 (left) or 16 (right). Note that as the sensitivity goes down the power spectrum and the shot noise go up, as mentioned earlier—only rarer and thus more clustered objects are included. The bin size is the spacing between the error bars, equal in log $l$. 

![Fig. 11.—Angular power spectrum with errors from shot noise, sampling, and cosmic variance error. Top: Error bars are for 100, 200, and 4000 deg$^2$ (left, vanilla model; right, $Y_{\text{min}} = 10^{-5}$), with the largest error bars corresponding to the smallest area. Bottom: Vanilla model except for the change $Y_{\text{min}} \rightarrow 2Y_{\text{min, vanilla}}$ (left) and $4Y_{\text{min, vanilla}}$ (right), with the area going up by factors of 4 and 16, respectively, relative to the vanilla model at top left. This naively keeps the observing time fixed. The error bars are equally spaced in log $l$. See the text for more details. [See the electronic edition of the Journal for a color version of this figure.](fig11.jpg)
However, the decrease in error bars due to increased sky coverage outweighs the increase in shot noise: shallow and wide appears naively preferable to deep and narrow.\footnote{22}

In fact, poorer sensitivity is not necessarily a bad thing, as the omitted clusters have low mass and thus the strongest dependence on the poorly understood gas physics. Exactly how deep is most profitable will depend very strongly on the normalization of the $Y(M)$ relation. In addition, the corresponding changes in the purity and completeness of the surveys have not been taken into account in this scaling argument; how these properties scale will depend on the cluster-finding algorithms in use as well. Fixing the sensitivity and increasing the area, as SPT will do, is of course a major improvement in any case.

In addition, even though the error bars are quite large, the measurement contains useful information. Theory gives smooth and known functions of $l$, $C_l$, so we can bin $N_l$ nearby values of $l$, reducing the error. For example, we can use the combined shape of the power spectrum and the errors in equation (29) to find $1\sigma$ contours in the $\sigma_8$-$\Omega_m$ plane. This is shown in Figure 12 for marginalizing over a 10% prior and a 30% prior on $T^\text{SZ}$ and holding all the other uncertainties fixed for our vanilla model. These error ellipses clearly illustrate the degeneracies inherent in the effects of changing these two cosmological parameters for the angular power spectrum.\footnote{23}

The closest error analysis to ours was done by Mei and Bartlett (2004), who used the measurement of $\omega(\theta)$ at 30$'$ and counts at the flux limit to get error contours for $\sigma_8$ and $\Omega_m$ of comparable size. Here we use the full power spectrum instead because our intent is to illustrate the degeneracies in using the power spectrum to constrain these parameters. As Mei and Bartlett’s degeneracy line in $\Omega_m$, $\sigma_8$ for $\omega(30')$ differs from the ellipse axes in Figure 12, the two sets of constraints are complementary.

5. CONCLUSIONS

FORTHCOMING SZ SURVEYS SUCH AS APEX ARE EXPECTED TO OBSERVE 1 OR 2 ORDERS OF MAGNITUDE MORE CLUSTERS THAN CURRENTLY IN HAND. THE ANGULAR POWER SPECTRUM WILL BE AN IMMEDIATE RESULT ONCE CLUSTERS HAVE BEEN IDENTIFIED. THERE STILL EXISTS UNCERTAINTY IN THEORETICAL PREDICTIONS: WE EXAMINED HOW THOSE UNCERTAINTIES AFFECT THE CLUSTER POWER SPECTRUM.

WE CALCULATED THE ANGULAR POWER SPECTRUM FOR DIFFERENT REASONABLE MASS FUNCTIONS, BIASES, MASS-TEMPERATURE NORMALIZATIONS $T^\text{SZ}$, AND GAS FRACTIONS $f_{\text{gas}}$ AND FOUND THAT THESE CHANGES ARE COMPARABLE TO CHANGES DUE TO COSMOLOGICAL PARAMETERS OF INTEREST SUCH AS $\sigma_8$ WITHIN CURRENT RANGES OF INTEREST. IN PARTICULAR, WE IDENTIFIED WHICH MODELING UNCERTAINTIES MIMICKED CHANGES IN THE COSMOLOGICAL PARAMETERS CONSIDERED (BY FINDING THE RELEVANT COSMOLOGICAL PARAMETER VALUES) AND WHICH DID NOT. SOME OF THESE MODELING AND COSMOLOGICAL DEPENDENCIES HAVE BEEN STUDIED FOR THE CORRELATION FUNCTION IN TWO DIMENSIONS OR THE THREE-DIMENSIONAL POWER SPECTRUM. DIFFERENT SUBSETS WITH DIFFERENT FIXED ASSUMPTIONS HAVE BEEN CONSIDERED IN PREVIOUS LITERATURE AT DIFFERENT TIMES. BY COMBINING ALL THESE VARIATIONS IN A HOMOGENEOUS MANNER AND INCLUDING SEVERAL MORE THAT HAVE BEEN IDENTIFIED SINCE, THE RELATIVE IMPORTANCE OF THE VARIOUS MODELING ASSUMPTIONS CAN BE MORE DIRECTLY ASSESSED.

WE FIND THAT PROGRESS ON SEVERAL FRONTS IS NEEDED BEFORE THE SCIENTIFIC HARVEST FROM THESE EXPERIMENTS CAN BE FULLY REALIZED. THE UNCERTAINTIES IN THE MASS FUNCTION OF CLUSTERS AND THE BIAS CAN BOTH BE IMPROVED WITH SIMULATIONS. WE SHOWED THAT THE DIFFERENCES BETWEEN MANY COMMONLY USED ONES ARE SIGNIFICANT IN COMPARISON TO THE UNCERTAINTIES FOR THE COSMOLOGICAL PARAMETERS OF INTEREST.

WE HAVE SUMMARIZED MUCH OF THE OBSERVATIONAL AND THEORETICAL WORK ON THE $Y(M)$ RELATION, BUT IMPORTANT UNCERTAINTIES

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure12.png}
\caption{The 1$\sigma$ contours using the full power spectrum and errors in eq. (29), varying $\Omega_m$ and $\sigma_8$, marginalized with a 10% prior on $T^\text{SZ}$ (left) and a 30% prior on $T^\text{SZ}$ (right). The larger contour is for 200 deg$^2$ (APEX), and the smaller one is for 4000 deg$^2$ (SPT). The true model is our vanilla model, shown by the point at the center.}
\end{figure}
still remain. The normalization of the $Y(M)$ relation needs considerable observational and/or simulation input (this point has also been emphasized previously by other authors as referenced in the text). Current experiments such as SZA might be able to perform this calibration when combined with other measurements for mass, as long as the normalization is not redshift dependent. Simulations with gas are known to have incomplete treatments of physics but can be used as a guide, e.g., to calibrate the effect of projection on the mass-temperature relation, which gives a systematic increase of the $Y(M)$ normalization of about 8%. We included scatter in the $Y(M)$ relation; however, this was a small effect: the angular power spectrum changed by less than 1% when the $Y(M)$ relation was taken to have a scatter (as recent hydrodynamic simulations suggest) of 10%. The effects of mergers seem to be small, once the normalization is fixed; we argue that the effect of most of the parameters is just to decide whether objects are included in the survey or not, and mergers have the largest effect on the largest clusters, which tend to be included in any case for the low values of $Y_{\text{min}}$ under consideration here.

We also studied the experimental uncertainties that exist even in the absence of systematic uncertainty. We compared area to sensitivity $Y_{\text{min}}$ and found that naively a shallower, wider survey is more powerful.

We focused on the angular power spectrum alone, although of course complementarity is key to progress. Complementary data will even be provided from a survey producing the power spectrum itself. A survey providing an angular power spectrum will also produce number counts per square degree and $dN/dY$, number counts as a function of $Y$. In addition, the temperature correlation function and perhaps $dN/dz$ will be available. Various combinations of these quantities have been analyzed in the literature. Mei & Bartlett considered number counts and the angular correlation function and combined it with the number counts from X-rays, for instance. This is one example of self-calibration (Levine et al. 2002); adding the angular power spectrum to other measurements will increase the leverage of all of them.

J. D. C. thanks G. Evrard, A. Kravtsov, E. Reese, and W. Hu for discussions and S. Mei for help with comparing with her work; K. K. thanks P. Zhang, W. Hu, and J. Weller for useful discussions. We both especially thank M. White for numerous discussions, early collaboration on this project, and comments on the draft and the anonymous referee for many constructive suggestions. This work was supported in part by DOE, NSF grant NSF-AST 02-05935, and NASA grant NAG5-10842.