The observed growth of massive galaxy clusters – I. Statistical methods and cosmological constraints

A. Mantz,[1,2]★ S. W. Allen,[1,2] D. Rapetti[1,2] and H. Ebeling[3]

[1]*Kavli Institute for Particle Astrophysics and Cosmology at Stanford University, 452 Lomita Mall, Stanford, CA 94305-4085, USA*
[2]*SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA*
[3]*Institute for Astronomy, 2680 Woodlawn Drive, Honolulu, HI 96822, USA*

ABSTRACT

This is the first of a series of papers in which we derive simultaneous constraints on cosmological parameters and X-ray scaling relations using observations of the growth of massive, X-ray flux-selected galaxy clusters. Our data set consists of 238 cluster detections from the *ROSAT* All-Sky Survey, and incorporates follow-up observations of 94 of those clusters using the *Chandra* X-ray Observatory or *ROSAT*. Here we describe and implement a new statistical framework required to self-consistently produce simultaneous constraints on cosmology and scaling relations from such data, and present results on models of dark energy. In spatially flat models with a constant dark energy equation of state, w, the cluster data yield $\Omega_m = 0.23 \pm 0.04$, $\sigma_8 = 0.82 \pm 0.05$ and $w = -1.01 \pm 0.20$, incorporating standard priors on the Hubble parameter and mean baryon density of the Universe, and marginalizing over conservative allowances for systematic uncertainties. These constraints agree well and are competitive with independent data in the form of cosmic microwave background anisotropies, type Ia supernovae, cluster gas mass fractions, baryon acoustic oscillations, galaxy redshift surveys and cosmic shear. The combination of our data with current microwave background, supernova, gas mass fraction and baryon acoustic oscillation data yields $\Omega_m = 0.27 \pm 0.02$, $\sigma_8 = 0.79 \pm 0.03$ and $w = -0.96 \pm 0.06$ for flat, constant w models. The combined data also allow us to investigate evolving w models. Marginalizing over transition redshifts in the range 0.05–1, we constrain the equation of state at late and early times to be respectively $w_0 = -0.88 \pm 0.21$ and $w_{et} = -1.05^{+0.20}_{-0.36}$, again including conservative systematic allowances. The combined data provide constraints equivalent to a Dark Energy Task Force figure of merit of 15.5. Our results highlight the power of X-ray studies, which enable the straightforward production of large, complete and pure cluster samples and admit tight scaling relations, to constrain cosmology. However, the new statistical framework we apply to this task is equally applicable to cluster studies at other wavelengths.

Key words: cosmological parameters – cosmology: observations – large-scale structure of Universe – X-rays: galaxies: clusters.

1 INTRODUCTION

Clusters of galaxies have a long history as cosmological laboratories, beginning with the discovery of dark matter (Zwicky 1937) through studies of cluster galaxy orbits, and providing early evidence for a low matter density universe (White et al. 1993) through studies of their baryonic and dark matter content. These breakthroughs rested on observations of the properties of individual systems, but the population of clusters as a whole also contains a great deal of cosmological information. Clusters represent the most massive gravitationally bound systems in the Universe, and as such their abundance probes the amount of structure in the Universe and its growth over cosmic time.

The local cluster population has been used to jointly constrain the average density of matter, Ω_m, and the amplitude of density perturbations, σ_8 (recently by Henry et al. 2009; Rozo et al. 2010, but see also Reiprich & Böhringer 2002; Allen et al. 2003; Schuecker et al. 2003 and additional references in Mantz et al. 2008, hereafter M08). Most recently, the construction of X-ray flux-limited cluster samples out to redshift $z = 0.5$ and beyond has enabled studies of the growth of structure (M08; Vikhlinin et al. 2009b). Such

★E-mail: amantz@slac.stanford.edu

studies provide an important new window on cosmology, as various dark energy and modified gravity models designed to explain the acceleration of the cosmic expansion are potentially distinguishable by their effects on structure formation (e.g. Rapetti et al. 2009, 2010; Schmidt, Vikhlinin & Hu 2009).

The key ingredients for investigations of the cluster population are (1) a cluster survey with a well-understood selection function,[1] and (2) scaling relations that link cluster mass to observable quantities. Currently, the most successful approach to finding massive clusters over a range of redshifts is through the X-ray emission of the hot intracluster gas, although Sunyaev–Zel'dovich and optical surveys are also making significant progress. Because Malmquist and Eddington biases are ubiquitous in current X-ray flux-limited samples, great care is warranted in their analysis. In particular, precise, robust cosmological constraints can only be obtained by simultaneously fitting the X-ray luminosity–mass relation, as we do here. Conversely, rigorous analysis of the scaling relations must take into account the cluster mass function and the selection function of the data set, as discussed by Mantz et al. (2010a, hereafter Paper II; see also Stanek et al. 2006; Pacaud et al. 2007).

While a flux-limited sample, combined with more detailed follow-up observations of a subset of clusters in that sample, may contain the necessary information to provide simultaneous constraints on cosmology and scaling relations, a fully self-consistent statistical framework for this analysis has been lacking to this point. By 'self-consistent', we mean that a single likelihood function applying to the full data set (survey + follow-up observations) and encompassing the entire theoretical model (cosmology + scaling relations) should be derived from first principles, ensuring that the covariance among all the model parameters is fully captured and that the effects of the mass function and selection biases are properly accounted for throughout.

This is the first of a series of papers in which we address these issues. Paper II contains details of the follow-up X-ray observations and their reduction, presents the constraints on scaling relations from our simultaneous analysis, and discusses their astrophysical implications. In this paper, we present the statistical methods applied to the problems described above and the resulting cosmological constraints. The dark energy models we address here include the simple cosmological constant model; models with a constant dark energy equation of state, w; and simple evolving w models. As we will show, our analysis produces some of the tightest constraints on dark energy parameters of any experiment to date. In Papers III (Rapetti et al. 2010) and IV (Mantz, Allen & Rapetti 2010b), we respectively apply our analysis to investigations of modified gravity and neutrino properties.

In Section 2, we briefly review the cluster data set, which is more fully described in Paper II. Section 3 contains a full description of the theoretical model fit to the data, including the cosmology, cluster mass function and scaling relations. The new, self-consistent analysis method is presented in Section 4.1. Sections 5 and 6 respectively contain the cosmological results of our analysis and a discussion of relevant systematics.

In this paper, we adopt the conventional definition of cluster radius in terms of the critical density of the Universe; thus, r_Δ is the radius within which the mean density of the cluster is Δ times the critical density at the cluster's redshift, $\rho_{cr}(z)$. An alternative convention, used particularly in the literature relevant to the mass function, is to define the overdensity with respect to the mean matter density at redshift z. We will consistently use the former convention, and so, for example, an overdensity of 300 with respect to the matter density will be written as $\Delta = 300\Omega_m(z)$, where $\Omega_m(z)$ is the ratio of the mean matter density to the critical density.

2 DATA

The galaxy cluster data used in this work, as well as their selection and reduction, are discussed in detail in Paper II. Using three wide-area cluster samples drawn from the *ROSAT* All-Sky Survey (RASS; Trümper 1993) – the *ROSAT* Brightest Cluster Sample (BCS; Ebeling et al. 1998), the *ROSAT*-ESO Flux-Limited X-ray sample (REFLEX; Böhringer et al. 2004) and the bright subsample of the Massive Cluster Survey (Bright MACS; Ebeling, Edge & Henry 2001; Ebeling et al. 2010) – we select a statistically complete sample of 238 X-ray luminous clusters covering the redshift range $z < 0.5$. Of these 238 clusters, 94 have follow-up *Chandra* or *ROSAT* observations that we incorporate into the analysis. To distinguish it from cluster data used in other cosmological work (e.g. optically selected clusters), we refer to this data set as the cluster X-ray luminosity function (XLF), although in fact the data set contains a great deal more information than the luminosity function alone. From the follow-up observations, we measure X-ray luminosity, average temperature and gas mass within r_{500}. The gas mass is used as a proxy for total mass, using the finding of Allen et al. (2008, hereafter A08) that the gas mass fraction, $f_{gas} = M_{gas}/M_{tot}$, is a constant for the hot, massive clusters being studied.[2] The systematic uncertainty associated with the f_{gas} measurement is accounted for by simultaneously constraining the full f_{gas} model of A08, which includes generous systematic allowances for instrument calibration, non-thermal pressure support, the depletion of baryons in clusters relative to the cosmic mean and evolution in the baryon depletion and stellar content of clusters. We use only the six lowest redshift clusters from A08 ($z < 0.15$), which are sufficient to constrain the gas mass fraction at low redshift without directly producing a constraint on dark energy by themselves. In our analysis of the XLF data, we also use Gaussian priors to constrain the Hubble parameter and the mean baryon density, based on the results of the Hubble Key project ($h = H_0/100 \, \text{km s}^{-1} \, \text{Mpc}^{-1} = 0.72 \pm 0.08$, Freedman et al. 2001) and big bang nucleosynthesis studies ($\Omega_b h^2 = 0.0214 \pm 0.002$, Kirkman et al. 2003).

In addition, we compare and combine results from our own analysis with those of independent cosmological data, including the full cluster f_{gas} data set (A08, 42 clusters at $z < 1.1$), as well as cosmic microwave background (CMB), type Ia supernova (SNIa) and baryon acoustic oscillation (BAO) data. Our analysis of the CMB anisotropies uses 5-yr *Wilkinson Microwave Anisotropy Probe* (*WMAP*) data (Hill et al. 2009; Hinshaw et al. 2009; Nolta et al. 2009) with the March 2008 version of the *WMAP* likelihood code[3] (Dunkley et al. 2009). The SNIa results are derived from

---

[1] Throughout this paper, we use the term 'selection function' to refer to the probability that a cluster be detected by a survey and included in the resulting data set, as a function of that cluster's physical properties (redshift, flux, etc.). See Section 4.1.

[2] We prefer to use M_{gas} as a proxy for total mass rather than the thermal energy, $Y_X = M_{gas}kT$, because M_{gas} can be measured more precisely than temperature for a given exposure time, and because M_{gas} measurements at r_{500} are minimally affected by background and emission weighting uncertainties. In addition, M_{gas} displays exceptionally small scatter with mass (see A08 and Paper II) in the mass range considered here, $M > 3 \times 10^{14} \, M_\odot$.
[3] http://lambda.gsfc.nasa.gov

the Union compilation (Kowalski et al. 2008), which includes data from a variety of sources (307 SNIa in total; Hamuy et al. 1996; Garnavich et al. 1998; Riess et al. 1998, 1999, 2004, 2007; Schmidt et al. 1998; Perlmutter et al. 1999; Krisciunas et al. 2001, 2004a,b; Knop et al. 2003; Tonry et al. 2003; Barris et al. 2004; Astier et al. 2006; Jha et al. 2006; Miknaitis et al. 2007; Wood-Vasey et al. 2007), including the treatment of systematic errors employed by Kowalski et al. Our analysis of BAO data uses the constraints on the ratio of the sound horizon to the distance scale at $z = 0.25$ and $z = 0.35$ derived by Percival et al. (2007) from the galaxy correlation function in 2dF (Colless et al. 2001, 2003) and Sloan Digital Sky Survey (Adelman-McCarthy et al. 2007) data.

When fitting CMB data, we allow the scalar spectral index, n_s, and optical depth to reionization, τ, to vary as free parameters, and marginalize over a plausible range in the amplitude of the Sunyaev–Zel'dovich signal due to galaxy clusters ($0 < A_{SZ} < 2$; introduced by Spergel et al. 2007). The combination of CMB and f_{gas} data places tight constraints on both h and $\Omega_b h^2$ in addition to other parameters of interest (see A08), so the Hubble Key project and big bang nucleosynthesis priors are unnecessary in analyses of the combined data sets. This applies as well to the combination of CMB and XLF data, since we always use the $z < 0.15$ subset of the f_{gas} data to calibrate the cluster mass scale for the XLF analysis.

3 MODEL

In this section, we detail the various pieces of the model being fit to the XLF data. Section 3.1 reviews the important cosmological parameters and the dark energy models to be addressed, while Section 3.2 focuses on the predicted mass function and its cosmological dependence. In Section 3.3, we outline the model for the cluster scaling relations. Sections 3.4 and 3.5 describe the sampling model, i.e. the model for how a set of parameters predicts the measured data, including measurement errors and their correlation, and implicit dependencies on reference cosmological values. The model parameters are summarized in Table 1.

3.1 Cosmological models

The simplest cosmological model considered in this study contains dark energy in the form of a spatially uniform and non-evolving energy density, i.e. a cosmological constant (ΛCDM). In this model, the cosmological parameters relevant for the growth of structure are the mean baryon density, Ω_b; the mean total matter density, Ω_m; the Hubble parameter, h; and the matter power spectrum normalization, σ_8. Here the mean densities refer to redshift zero, since their values at other times are then determined by the Friedmann equation, and σ_8^2 is the $z = 0$ variance in the

Table 1. Parameters and priors used in the analysis. The nuisance parameters associated with the f_{gas} model (A08) are not shown, although they are also marginalized over. When no entry appears in the prior column, the prior was uniform and significantly wider than the marginal posterior for that parameter. $\mathcal{N}(\mu, \sigma)$ represents the normal distribution with mean μ and variance σ^2, and $\mathcal{U}(x_1, x_2)$ the uniform distribution with endpoints x_1 and x_2. For brevity, $\mathcal{N}_4(\mu, \ldots)$ represents the multivariate normal prior used for the Tinker et al. (2008) mass function parameters, where μ is the marginal mean for each parameter and the covariance matrix is not explicitly shown.

Type	Symbol	Meaning	Prior
Cosmology (Section 3.1)	h	Hubble parameter[a,b]	$\mathcal{N}(0.72, 0.08)$
	$\Omega_b h^2$	Baryon density[b]	$\mathcal{N}(0.0214, 0.002)$
	$\Omega_c h^2$	Cold dark matter density	
	$\ln(10^{10} A_s)$	Scalar power spectrum amplitude	
	n_s	Scalar power spectrum slope[b,c]	$= 0.95$
	τ	Optical depth to reionization[c]	
	A_{SZ}	Sunyaev–Zel'dovich signal amplitude[c]	$\mathcal{U}(0, 2)$
	w	Constant dark energy equation of state	
	w_0	Evolving w: current value	
	w_a, w_{et}	Evolving w: value at early times	
	a_t	Evolving w: transition scale factor	$\mathcal{U}(0.5, 0.95)$
Mass function (Section 3.2)	A	Global amplitude	$\mathcal{N}_4(0.20, \ldots)$
	a	Low mass amplitude	$\mathcal{N}_4(1.52, \ldots)$
	b	Low mass slope	$\mathcal{N}_4(2.25, \ldots)$
	c	Exponential cut-off scale	$\mathcal{N}_4(1.27, \ldots)$
	ε	Evolution strength	$\mathcal{N}(1.0, 0.1)$
Scaling relation (Section 3.3)	$\beta_0^{\ell m}, \beta_1^{\ell m}$	Nominal luminosity–mass relation	
	$\beta_0^{tm}, \beta_1^{tm}$	Nominal temperature–mass relation	
	$\sigma_{\ell m}, \sigma_{tm}$	Marginal scaling relation scatters	
	$\rho_{\ell tm}$	Scaling relation scatter correlation	
Other (Sections 3.4 and 3.5)	—	*Chandra* temperature calibration	10% Normal
	η_g	Cluster gas mass profile logarithmic slope	$\mathcal{N}(1.092, 0.006)$
	η_L	Cluster luminosity profile logarithmic slope	$\mathcal{N}(0.1135, 0.0005)$
	κ	r_{500}-to-survey flux conversion[d]	
	ξ	Completeness/purity[d]	$\mathcal{U}(0.95, 1.05)$

[a]When using CMB data, the angular size of the sound horizon at last scattering replaces h as a free parameter; [b]Indicates priors that are not used when CMB data are included in the analysis; [c]Indicates parameters that are free only when CMB data are included; [d]Indicates that there is an independent parameter of this type for each cluster sample.

density field at scales of $8\,h^{-1}$ Mpc, defined explicitly in equation (4). Because the mass range of our data corresponds to a small range in scale, we do not simultaneously fit for the spectral index of scalar density perturbations, n_s, but rather fix its value at 0.95 (e.g. Komatsu et al. 2009), except when simultaneously fitting CMB data (see Section 2 and Table 1). This assumption does not significantly affect our results; see Section 5.1. We assume that the Universe is spatially flat on large scales throughout.

We additionally consider models in which dark energy is a fluid parametrized by a constant equation of state, w (constant w models). Unlike the cosmological constant scenario, such a fluid will not in general have uniform density, and thus it contributes to some degree to the evolution of density perturbations, in addition to influencing the expansion history of the Universe. Due to theoretical uncertainties on the behaviour of the dark energy fluid on the non-linear scales that determine the mass function, numerical simulations of the mass function have to date been done only for models in which dark energy is uniform, even when $w \neq -1$. Our approach is to straightforwardly propagate the influence of non-uniform dark energy on linear scales but to leave the mass function unaltered, while continuing to use our standard systematic allowances on the mass function and its evolution (Section 3.2). We have verified that the value of the dark energy sound speed (assumed to be constant with time) has no effect on our results. Preliminary theoretical work indicates that the effect of dark energy perturbations on the mass function might be readily measurable (Abramo, Batista & Rosenfeld 2009; Park et al. 2009; Alimi et al. 2010; Creminelli et al. 2010), in which case our approach likely underestimates the ability of the data to discriminate among these models.

Finally, we consider models in which the dark energy equation of state is a function of time, according to two parametrizations. The first is the commonly used model of Chevallier & Polarski (2001) and Linder (2003),

$$w(a) = w_0 + w_a(1-a), \tag{1}$$

where $a = 1/(1+z)$ is the scalefactor, in which the equation of state makes a smooth transition from value w_0 at $z = 0$ to $w_0 + w_a$ at high redshift. A generalization due to Rapetti, Allen & Weller (2005),

$$w(z) = \frac{w_{et}z + w_0 z_t}{z + z_t}, \tag{2}$$

has the advantage that the transition redshift, z_t, can be marginalized over. (Equation 1 is a special case of equation 2 with $z_t = 1$ and $w_a = w_{et} - w_0$.) This model has greater applicability to current data, which primarily constrain w at $z < 1$, resulting in more commensurate constraints on the current and early-time equation of state, w_0 and w_{et}. In practice, we marginalize over the scale factor of the transition, a_t, within the range $0.5 < a_t < 0.95$.

3.2 Mass function

Cosmological analyses of the kind presented here are enabled by the fact that, to a good approximation, the expected number density of dark matter haloes as a function of mass, M, can be expressed as a relatively simple function of cosmological parameters,

$$\frac{dn(M, z)}{dM} = \frac{\bar{\rho}_m}{M} \frac{d \ln \sigma^{-1}}{dM} f(\sigma). \tag{3}$$

Here $\bar{\rho}_m$ is the mean comoving matter density and σ^2 is the variance of the linearly evolved density field, smoothed by a spherical top-hat window of comoving radius r, enclosing mass $M = 4\pi \bar{\rho}_m r^3/3$,

$$\sigma^2(M, z) = \frac{1}{2\pi^2} \int_0^\infty k^2 P(k, z) |W_M(k)|^2 \, dk, \tag{4}$$

where $P(k, z)$ is the linear power spectrum evolved to redshift z and $W_M(k)$ is the Fourier transform of the window function. In the formulation of equation (3), the mass function depends on cosmological parameters and redshift only through $\sigma^2(M, z)$. The function $f(\sigma)$ may be an analytic or semi-analytic approximation (Press & Schechter 1974; Bond et al. 1991; Sheth & Tormen 1999) or a fit to cosmological N-body simulations.

The applicability of this 'universal' form of the mass function was first demonstrated in numerical dark matter simulations of flat ΛCDM and open ($\Omega_\Lambda = 0$) cosmologies by Jenkins et al. (2001) and confirmed by Evrard et al. (2002). It has since been verified that the fitting function provided by Jenkins is approximately accurate (within $\sim$20 per cent) among models with constant $w \neq -1$ and some evolving w models (Klypin et al. 2003; Linder & Jenkins 2003; Łokas et al. 2004; Kuhlen et al. 2005).[4] Other authors have studied the dependence of $f(\sigma)$ on redshift beyond that implicit in $\sigma^2(M, z)$ (Lukić et al. 2007; Reed et al. 2007; Cohn & White 2008), replacing $f(\sigma)$ with $f(\sigma, z)$. The most recent and relevant work is that of Tinker et al. (2008), which we adopt here.

The Tinker fitting function has the form

$$f(\sigma, z) = A \left[\left(\frac{\sigma}{b}\right)^{-a} + 1 \right] e^{-c/\sigma^2}, \tag{5}$$

where each of the fitted parameters has a redshift dependence of the form

$$x(z) = x_0(1+z)^{\varepsilon \alpha_x}, \quad x \in \{A, a, b, c\}. \tag{6}$$

The various parameters x_0 and α_x are given in Tinker et al. (2008) as a function of the spherical overdensity, Δ, used to define the cluster radius. Unlike the $z = 0$ mass function, this additional redshift dependence has not been tested in simulations of cosmologies beyond the simple, flat ΛCDM model. We therefore introduce the parameter ε, which controls the overall strength of the evolution given by the α_x, in order to marginalize over the remaining uncertainties in the redshift dependence of the mass function in exotic cosmologies. We also choose to work with the $\Delta = 300\Omega_m(z)$ fit to $f(\sigma, z)$ (a relatively large cluster radius) because the evolution parametrized by the α_x becomes more pronounced with increasing overdensity (smaller radius).

To address the uncertainty in the normalization and shape of $f(\sigma, z = 0)$, we marginalized over each of the fitted parameters in equation (5) using the covariance matrix of the fit (Jeremy Tinker, private communication). The statistical error of this fit is <5 per cent; however, this figure does not reflect systematic uncertainties due to the presence of baryons (e.g. Stanek, Rudd & Evrard 2009), evolving dark energy, etc. We therefore scaled the covariance matrix when defining this prior on the mass function parameters such that the marginal uncertainty at fixed $\log_{10}(\sigma^{-1}) = 0.2$ ($M \sim 10^{15}\,M_\odot$ in the concordance model) is a conservative 10 per cent.[5]

---

[4] We reiterate that these works include the effects of dark energy on the mass function through the cosmic expansion rate, but not the effects of dark energy density perturbations (Section 3.1).

[5] There is no reason, a priori, that the systematic uncertainty in the mass function should have a similar form to the statistical covariance. At minimum, however, this procedure provides a straightforward way to marginalize over a family of functions that are similar to the mass function, while still allowing differences in the shape as well as the normalization. As we show

3.3 Scaling relations

To perform the cosmological analysis, we need to relate cluster mass to the observable that determines cluster detection, in this case X-ray flux. Given a redshift, z, a cluster's unabsorbed, soft X-ray flux, F, is determined by its intrinsic X-ray luminosity, L, temperature, kT, and metallicity, Z, as

$$F(z, L, kT, Z) = \frac{L}{4\pi d_L^2(z) K(z, kT, Z)}, \tag{7}$$

where $d_L(z)$ is the luminosity distance to the cluster, and $K(z, kT, Z)$ is the required K-correction. For intracluster medium temperatures $kT > 3\,\mathrm{keV}$ and luminosities and fluxes in the soft (*ROSAT*) X-ray band (0.1–2.4 keV), K has only a weak dependence on temperature and negligible dependence on metallicity; we hereafter fix Z to the typical value of 0.3 times the solar value.

As discussed in Paper II, a simple prescription for how L and kT are related to the total mass, M, is given by the self-similar model (Kaiser 1986). We define nominal luminosity–mass and temperature–mass relations at r_{500} of the form

$$\langle \ell(m) \rangle = \beta_0^{\ell m} + \beta_1^{\ell m} m,$$
$$\langle t(m) \rangle = \beta_0^{tm} + \beta_1^{tm} m, \tag{8}$$

where

$$\ell = \log_{10}\left(\frac{L_{500}}{E(z)10^{44}\,\mathrm{erg\,s^{-1}}}\right),$$

$$m = \log_{10}\left(\frac{E(z)M_{500}}{10^{15}\,\mathrm{M_\odot}}\right),$$

$$t = \log_{10}\left(\frac{kT_{500}}{\mathrm{keV}}\right), \tag{9}$$

and $E(z)$ is the normalized Hubble parameter, $H(z)/H_0$. The factors of $E(z)$ appearing explicitly in equation (9) follow from the definition of cluster radius using a fixed overdensity with respect to the critical density (Bryan & Norman 1998). To describe the intrinsic scatter in ℓ and t given m about the nominal relations, we adopt a simple, bivariate normal distribution parametrized by the marginal luminosity–mass and temperature–mass lognormal scatters, $\sigma_{\ell m}$ and σ_{tm}, and a coefficient of correlation, $\rho_{\ell tm}$.

There are various ways of adding complexity to this scaling relation model, including departures from self-similar evolution in the normalization of the nominal relations, evolution in the scatter, and asymmetry in the scatter. In Paper II, we show that the data are consistent with the simple model defined above, and do not require or prefer any such additions, even when the cosmological parameters are extremely restricted by external data. We therefore adopt the simple model above in this work.

Motivated by the results of Evrard et al. (2008), we have defined the scaling relations for quantities within r_{500}, which also corresponds to our measurements of mass, luminosity and temperature from follow-up observations (Paper II). To convert the mass definition used by the mass function, $\Delta = 300\Omega_\mathrm{m}(z)$, to $\Delta = 500$, we use the procedure of Hu & Kravtsov (2003), assuming a Navarro, Frenk & White (1997, hereafter NFW) mass distribution with concentration parameter $c = 4$. This conversion is negligibly sensitive to the assumed concentration parameter, since both r_{500} and $r_{300\Omega_\mathrm{m}(z)}$ are well beyond the NFW scale radius for reasonable values ($c > 3$; Zhao et al. 2003, 2009; Gao et al. 2008).

3.4 Sampling model: follow-up observations

The next component of the model connects quantities predicted from the cosmology, mass function and scaling relations to quantities measured from the follow-up X-ray observations. Because some of our measurements are made with respect to a reference cosmology (see Paper II), this procedure is not entirely trivial.

In particular, the mass, luminosity and temperature determined from the follow-up X-ray observations are measured within r_{500}, itself determined via the implicit equation

$$M(r_{500}) = \frac{M_\mathrm{gas}(r_{500})}{f_\mathrm{gas}(r_{500})} = \frac{4\pi}{3}(500)\rho_\mathrm{cr}(z)r_{500}^3, \tag{10}$$

which can be rewritten as

$$M_\mathrm{gas}(r) \propto \rho_\mathrm{cr}(z)r^3 f_\mathrm{gas}(r) \propto r^{\eta_g}, \tag{11}$$

where η_g is the logarithmic slope of the gas mass profile at large radius. Using that $\rho_\mathrm{cr}(z) \propto H^2(z)$, the expression

$$r_{500} \propto \left[f_\mathrm{gas}(r_{500})H^2(z)\right]^{1/(\eta_g - 3)} \tag{12}$$

relates the 'true' value of r_{500} predicted by a set of model parameters to the value of r_{500} that we would have inferred assuming our reference cosmology and reference f_gas value. The gas mass profiles measured in Paper II are self-similar, consistent with a constant value of η_g; for simplicity, we therefore adopt $\eta_g = 1.092 \pm 0.006$, determined from a fit to the entire sample from 0.7–$1.3r_{500}$, and marginalize over the uncertainty.

To see how the measurements of total cluster mass depend on model parameters, we first write

$$\frac{M^\mathrm{ref}(r)}{M(r)} = \frac{M_\mathrm{gas}^\mathrm{ref}(r)/f_\mathrm{gas}^\mathrm{ref}(r)}{M_\mathrm{gas}(r)/f_\mathrm{gas}(r)} R_\mathrm{NFW}$$

$$= \frac{d_A^\mathrm{ref}(z)^{2.5}}{d_A(z)^{2.5}} \frac{f_\mathrm{gas}}{f_\mathrm{gas}^\mathrm{ref}} R_\mathrm{NFW}, \tag{13}$$

where M^ref is the prediction for what mass would be measured for a cluster of true mass M using our assumed reference parameter values, and $d_A(z)$ is the angular diameter distance to redshift z. $d_A^\mathrm{ref}(z)$ and $f_\mathrm{gas}^\mathrm{ref}$ are not predictions for measured values, but they depend directly on the reference cosmology and f_gas value. The first term in this expression accounts for the dependence of the mass measured within a fixed angular aperture on distance and f_gas, given that the mass is estimated via the gas mass and gas mass fraction, and where we have used the scaling $M_\mathrm{gas} \propto d_A(z)^{2.5}$. The dependence on the angular size corresponding to physical radius r is handled by R_NFW, which we evaluate assuming that the shape of the total mass distribution near r_{500} is well approximated by the NFW profile; in this case the scaling factor is straightforward to compute using the scaling of r_{500} given in equation (12):

$$R_\mathrm{NFW} = \frac{\ln(1 + xc_{500}) - xc_{500}/(1 + xc_{500})}{\ln(1 + c_{500}) - c_{500}/(1 + c_{500})}, \tag{14}$$

where $x = r_{500}^\mathrm{ref}/r_{500}$. To evaluate this factor, we assume a concentration parameter $c = 4$,[6] although we note that R_NFW is extremely insensitive to this assumption provided that $c > 3$ (i.e. provided r_{500} is well beyond the scale radius).

Similarly, we can write the scaling of the luminosity as

$$L_{500}(r) \propto d_L^2(z)\left(\frac{r_{500}}{d_A(z)}\right)^{\eta_L}, \tag{15}$$

in Section 6.2, uncertainty on the mass function at this level has negligible impact on our results, in any case.

----

[6]Note that, by convention, c without a subscript refers to c_{200}, defined as r_{200} in units of the NFW scale radius; $c = 4$ corresponds to $c_{500} \approx 2.6$.

where $d_L(z)$ is the luminosity distance and the first term is the scaling of the luminosity measured within a fixed angle; the second term is simply the angular size of r_{500} raised to the power of η_L, the logarithmic slope of the integrated luminosity profile. As with the gas mass profiles, we fit this slope between 0.7 and $1.3r_{500}$, finding $\eta_L = 0.1135 \pm 0.0005$ from the entire sample, and marginalize over the uncertainty.

Since the emission-weighted, average temperature within r_{500} is a weak function of r_{500} compared to a typical statistical error bar, and considering the relative insensitivity of the cosmological analysis to the precise kT values, we do not similarly model any dependence of the measured temperatures on model parameters. However, we marginalize over a global 10 per cent Gaussian systematic uncertainty on the temperature measurements to account for residual uncertainty in instrument calibration.

Since the masses, luminosities and temperatures from the follow-up observations are determined from the same data, their statistical error bars are correlated. The masses and luminosities are integrated within r_{500}, so their errors are primarily correlated via the uncertainty in r_{500}. The luminosity and temperature errors are also somewhat correlated, since the K-corrections that convert flux to luminosity have a weak temperature dependence. (There is no correlation between mass and temperature errors using our method; see Paper II.) Since these measurement errors were propagated via Monte Carlo, it is straightforward to compute the error correlations from the samples; we find $\rho_{\hat{m}\hat{\ell}} = 0.29 \pm 0.19$ and $\rho_{\hat{\ell}\hat{t}} = -0.4 \pm 0.3$, averaged over the cluster sample. Note that the latter quantity is not related to $\rho_{\ell t m}$, defined in Section 3.3, which quantifies the correlation between departures in luminosity and temperature from the nominal scaling relation, not the correlation of measurement errors.

3.5 Sampling model: survey

The final ingredient is a sampling model for the selection of our clusters from the RASS, incorporating the selection function of each cluster sample, the survey flux measurement errors, and the relation between the flux measurements reported in each cluster sample and F_{500}, the flux within r_{500} from follow-up observations (in the same energy band). The fluxes reported for each cluster sample are measured differently (see Ebeling et al. 1998, 2001; Böhringer et al. 2004) and none corresponds to a flux measurement within a fixed-overdensity radius such as r_{500}. Fortunately, we find a good empirical correlation between the reported survey fluxes and our fluxes measured from the follow-up observations for clusters well above the flux limit.[7] We include a simple linear conversion between survey flux, F_s, and follow-up observation flux, F_{500}, in the model,

$$F_s = \kappa F_{500}, \tag{16}$$

where the slope, κ, is allowed to be different for each cluster sample. Because the survey fluxes suffer from Malmquist and Eddington biases, the κs must be constrained simultaneously with the full model, including the cosmology and scaling relation. These parameters also trivially account for any residual uncertainties in the flux cross-calibration between *ROSAT* and *Chandra* (see Paper II).

The flux measurement error as a function of reported flux for each sample is consistent with Poisson scaling,

$$\sigma_F \propto \sqrt{F_s}, \tag{17}$$

although the normalization varies by sample due to the different algorithms used to measure the flux. We approximate the survey flux sampling distribution as Gaussian, centred on the true flux, with width determined via equation (17), fitting a separate normalization for each cluster sample directly from the reported fluxes and errors.

Finally, the selection function of each cluster sample, the probability that a cluster is included in the sample as a function of redshift and measured survey flux, is modelled by interpolating the look-up tables provided for each sample (Ebeling et al. 1998, 2010; Böhringer et al. 2004). Conservatively, we include an additional nuisance parameter to marginalize over possible uniform incompleteness or impurity at the 5 per cent level,

$$P_{\text{sel}}(z, \hat{F}_s) \propto \xi, \tag{18}$$

where each sample can have a different value of ξ.

4 ANALYSIS METHOD

4.1 Likelihood

Early cosmological work using galaxy cluster samples used an approach based around binning the detected clusters in redshift and flux (or luminosity) and either adopting external priors on the luminosity–mass relation (e.g. Borgani et al. 2001) or simultaneously fitting an external luminosity–mass data set without accounting for selection effects (e.g. Allen et al. 2003).[8] In M08, we took this approach to its logical limit by deriving the likelihood for bins of infinitesimal volume; however, that work still suffered from the fact that it used an external data set to constrain the luminosity–mass relation without explicitly accounting for selection bias. Consequently, it was necessary to restrict that external data set (Reiprich & Böhringer 2002) to low redshifts and high fluxes in order to minimize the effects of selection bias, making it impossible to test for departures from self-similar evolution in the scaling relation. More recently, Vikhlinin et al. (2009a, b) binned their detected clusters in redshift and mass (again with infinitesimally small bins) and used the same cluster sample to constrain the scaling relations; however, their procedure still does not produce a self-consistent fit for both scaling relations and cosmology.

In this section, we show that our procedure from M08 can be generalized to allow such a simultaneous and self-consistent fit, using follow-up observations of flux-selected clusters to constrain the scaling relations over the full redshift range of the data, and accounting fully for the presence of Malmquist and Eddington biases. We also show that the corresponding likelihood function can be derived from first principles, beginning with a simple Bayesian regression model. For simplicity, we derive the likelihood for the general problem of counting sources as a function of their properties. This general picture includes the following components.

(i) A population function,[9] $\langle dN/dx \rangle$, which provides a theoretical prediction for the underlying distribution of sources (i.e. their number, N) as a function of properties, x (see below).

---

[7] For true fluxes near or below the flux limit, the Malmquist bias distorts the relation.

[8] We note that some authors have followed the (in principle) equivalent approach of using a scaling relation to transform the predicted mass function into, e.g. the baryonic mass function (Voevodkin & Vikhlinin 2004) or X-ray temperature function (Henry 2004; Henry et al. 2009). Although the statistical formalism discussed in this section is considered as an extension of the flux–redshift binning approach, it also underlies these methods.

[9] Note that this nomenclature is not widely used. We adopt the term 'population function' to imply an analogy to the mass function used in this study.

(ii) Population variables, x, on which the population function depends.

(iii) Response variables, y, which obey a stochastic scaling relation as a function of x.

(iv) The stochastic scaling relation, $P(y|x)$.

(v) Observed values, $\hat{x}$ and $\hat{y}$. Note that all of the x and y need not actually be measured, the exception being those which determine whether a source is included in the sample (detected; see also Section 4.1.3).

(vi) Sampling distributions for the observations as a function of population and response variables, $P(\hat{x}, \hat{y}|x, y)$.

(vii) A selection function, $P(I|x, y, \hat{x}, \hat{y})$. I represents the inclusion of a source in the sample. In full generality, the selection function could be a function of true variables as well as observations; for example, if it is calibrated using reliable simulations.

Throughout the following, we make the approximation that the clustering of sources is unimportant compared with the pure Poisson nature of their occurrence, which is justified for current all-sky surveys of very massive clusters (i.e. where the survey dimensions are large compared with the correlation length), as in this study (Hu & Kravtsov 2003; Holder 2006). Provided this assumption remains valid, the approach described here can be applied equally well to cluster studies based on optical or Sunyaev–Zel'dovich (SZ) surveys. We note that this is only a simplifying assumption; in principle, the approach detailed below can be straightforwardly generalized to include the spatial correlation of sources.

4.1.1 Binning derivation

Divide the space $(\hat{x}, \hat{y})$ into bins (indexed by j) of very small volume $\Delta\hat{x}_j\Delta\hat{y}_j$, such that no bin contains more than one detected source, and both the population function and scaling relation are approximately constant over the bin volume. The expected number of detected sources in bin j, located at $(\hat{x}_j, \hat{y}_j)$, is

$$\langle N_{\mathrm{det},j}\rangle = (\Delta\hat{x}_j\Delta\hat{y}_j)\int \mathrm{d}x \int \mathrm{d}y \left\langle \frac{\mathrm{d}N}{\mathrm{d}x}\right\rangle P(y|x)$$
$$\times P(\hat{x}_j, \hat{y}_j|x, y)P(I|x, y, \hat{x}_j, \hat{y}_j). \quad (19)$$

The likelihood, $\mathcal{L}$, is a product of independent Poisson likelihoods for each bin,

$$\mathcal{L}(\{N_j\}) = \prod_j \frac{\langle N_{\mathrm{det},j}\rangle^{N_j}\mathrm{e}^{-\langle N_{\mathrm{det},j}\rangle}}{N_j!}$$
$$= \mathrm{e}^{-\langle N_{\mathrm{det}}\rangle}\prod_{j:N_j=1}\langle N_{\mathrm{det},j}\rangle, \quad (20)$$

where N_j is the actual number of clusters detected in bin j, $\langle N_{\mathrm{det}}\rangle$ is the predicted total number of detected sources (summed over all bins) and the second equality follows from the fact that $N_j \in \{0, 1\}$ by construction.

Since the binning scheme must be invariant under changes to the values of the model parameters, the derivative of the log-likelihood is independent of the bin volumes, $\Delta\hat{x}_j\Delta\hat{y}_j$. Defining $\langle\tilde{n}_{\mathrm{det},j}\rangle = \langle N_{\mathrm{det},j}\rangle/(\Delta\hat{x}_j\Delta\hat{y}_j)$, we can write

$$\mathcal{L}(\{N_j\}) \propto \mathrm{e}^{-\langle N_{\mathrm{det}}\rangle}\prod_{j:N_j=1}\langle\tilde{n}_{\mathrm{det},j}\rangle \quad (21)$$

and ignore the constant of proportionality for the purposes of maximizing $\mathcal{L}$.

4.1.2 Regression derivation

In the regression of truncated data (where some sources are not detected and the total number of sources is thus unknown), the total number of detected plus undetected sources, N, becomes a parameter of the model and must be marginalized over (for background, see Gelman et al. 2004; Kelly 2007). We define $\langle N\rangle$, $\langle N_{\mathrm{det}}\rangle$ and $\langle N_{\mathrm{mis}}\rangle$ to be respectively the predictions for the total number of sources, number of detected sources and number of undetected (missed) sources as follows:

$$\langle N\rangle = \int \mathrm{d}x \left\langle\frac{\mathrm{d}N}{\mathrm{d}x}\right\rangle,$$
$$\langle N_{\mathrm{det}}\rangle = \int \mathrm{d}x \left\langle\frac{\mathrm{d}N}{\mathrm{d}x}\right\rangle \int \mathrm{d}y\, P(y|x)$$
$$\times \int \mathrm{d}\hat{x} \int \mathrm{d}\hat{y}\, P(\hat{x}, \hat{y}|x, y)P(I|x, y, \hat{x}, \hat{y}),$$
$$\langle N_{\mathrm{mis}}\rangle = \langle N\rangle - \langle N_{\mathrm{det}}\rangle. \quad (22)$$

The joint likelihood of the observations and N is

$$\mathcal{L}(\hat{x}, \hat{y}, N) = \left[\frac{\langle N\rangle^N \mathrm{e}^{-\langle N\rangle}}{N!}\right]\left[\frac{N!}{N_{\mathrm{det}}!N_{\mathrm{mis}}!}\right]$$
$$\times \prod_{i=1}^{N_{\mathrm{det}}}P_{\mathrm{det}}(\hat{x}_i, \hat{y}_i, I)\prod_{j=1}^{N_{\mathrm{mis}}}P_{\mathrm{mis}}(\bar{I}). \quad (23)$$

The first term is a Poisson likelihood for the model parameter N, the total number of sources. The second factor is a binomial coefficient enumerating the number of ways of selecting N_{det} detected sources from N. This term is necessary because, in the Bayesian treatment, the true source properties (x, y) are random variables; as such, sources are distinguishable only in terms of their detection (I) or non-detection ($\bar{I}$). The probability that the N_{det} detected sources have measurements $(\hat{x}, \hat{y})$ is accounted for by the product over detected objects. The second product accounts for the likelihood of not detecting an additional N_{mis} sources whose properties are distributed according to the population function and scaling relation.

Noting that the marginal probability for a source to have properties x is

$$P(x) = \frac{\langle \mathrm{d}N/\mathrm{d}x\rangle}{\langle N\rangle}, \quad (24)$$

the functions P_{det} and P_{mis} can be related to quantities defined previously (cf. equations 19 and 22):

$$P_{\mathrm{det}}(\hat{x}_i, \hat{y}_i, I) = \int \mathrm{d}x \int \mathrm{d}y \frac{\langle \mathrm{d}N/\mathrm{d}x\rangle}{\langle N\rangle}P(y|x)$$
$$\times P(\hat{x}_i, \hat{y}_i|x, y)P(I|x, y, \hat{x}_i, \hat{y}_i)$$
$$= \frac{\langle\tilde{n}_{\mathrm{det},i}\rangle}{\langle N\rangle} \quad (25)$$

and

$$P_{\mathrm{mis}}(\bar{I}) = \int \mathrm{d}x \int \mathrm{d}y \frac{\langle \mathrm{d}N/\mathrm{d}x\rangle}{\langle N\rangle}P(y|x)$$
$$\times \int \mathrm{d}\hat{x} \int \mathrm{d}\hat{y}\, P(\hat{x}, \hat{y}|x, y)P(\bar{I}|x, y, \hat{x}, \hat{y})$$
$$= \frac{\langle N_{\mathrm{mis}}\rangle}{\langle N\rangle}. \quad (26)$$

Substituting these expressions, the likelihood simplifies to

$$\mathcal{L}(\hat{x}, \hat{y}, N) = \left[\frac{\langle N\rangle^N}{\langle N\rangle^{N_{\mathrm{det}}}\langle N\rangle^{N_{\mathrm{mis}}}}\right]\left[\frac{1}{N_{\mathrm{det}}!}\right]$$
$$\times \left[\frac{\langle N_{\mathrm{mis}}\rangle^{N_{\mathrm{mis}}}\mathrm{e}^{-\langle N_{\mathrm{mis}}\rangle}}{N_{\mathrm{mis}}!}\right]$$
$$\times \mathrm{e}^{-\langle N_{\mathrm{det}}\rangle}\prod_{i=1}^{N_{\mathrm{det}}}\langle\tilde{n}_{\mathrm{det},i}\rangle. \quad (27)$$

The first factor above is unity, and the second is a normalization term independent of model predictions. The third term is a Poisson density which sums to unity when $\mathcal{L}$ is marginalized over N from N_{det} to infinity (equivalently, N_{mis} from zero to infinity). The remaining terms are exactly equivalent to those in equation (21).[10]

4.1.3 Application to the growth of structure

The formalism of the last sections can be straightforwardly applied to observations of the growth of cosmic structure. The population variables describing clusters are redshift, z, and mass, m, and the population function is simply the product of the mass function and the comoving volume element, $\mathrm{d}N/\mathrm{d}z\,\mathrm{d}m = (\mathrm{d}n/\mathrm{d}m)(\mathrm{d}V/\mathrm{d}z)$ (see definitions in Section 3). The response variables are luminosity, ℓ, and temperature, t, which scale with mass and redshift as described in Section 3.3. The observations include spectroscopically determined redshifts (for which we neglect the measurement errors in practice), the masses, luminosities and temperatures determined from follow-up observations, and the survey fluxes. Their sampling distributions, which account for the correlations among m, ℓ and t measurements from the same follow-up data, are discussed in Sections 3.4 and 3.5. The selection functions of the cluster samples are given as functions of redshift and observed survey flux, and incorporate the appropriate sky coverage fraction.

It is important to note that this analysis method does not require all the potentially observed values $\{\hat{x}, \hat{y}\}$ to be available for each source, apart from those which determine whether a source is included in the sample.[11] This is perhaps most apparent from considering how equation (25) would be modified if a particular observable, $\hat{y}_0$, were unavailable: an additional integration over $\hat{y}_0$ would be required, which would result in $P(\hat{y}_0|x, y)$ being replaced by unity.[12] In the binning derivation, this is analogous to the bin containing that cluster being unbounded rather than constrained in the $\hat{y}_0$ direction. This result is intuitive, as it simply means that the integral over y_0 for that cluster is not weighted by the information gained from a measurement.

Thus, our analysis makes use of the available information for all the 238 clusters meeting our selection criteria. For the 144 clusters without follow-up data, only the redshift and survey fluxes are available; following the argument above, the terms associated with follow-up luminosity, mass and temperature measurements for each of these clusters integrate to unity. In effect, the likelihood for these clusters (here we refer to the terms $\langle \tilde{n}_{\mathrm{det},j} \rangle$) simplifies to the form used in M08, which depends only on observations of redshift and survey flux. For the 94 clusters with follow-up measurements, the analysis uses the redshift and survey flux, in addition to the (independently measured) follow-up luminosity, mass and temperature. The latter are, of course, crucial for obtaining simultaneous constraints on the scaling relations.

The ability to use partial follow-up in a statistically rigorous way has significance for the planning of large, future cluster surveys. In particular, it reveals as overly simplistic the choice between strategies employing exhaustive follow-up observations to calibrate scaling relations over a relatively small mass range and those using 'self-calibration', which use many more sources but rely exclusively on the shape of the mass function to provide information about the scaling relations. Indeed, obtaining follow-up data for a fair subsample of the detected clusters potentially allows robust calibration of the scaling relations over a wide mass range, without unnecessarily limiting the size of the sample used for cosmology.[13] However, we note that this attractively simple strategy is not necessarily the best; a more complete analysis, including an optimization procedure similar to that employed by Wu, Rozo & Wechsler (2010), is warranted in planning the follow-up of future surveys.

Finally, we notice that the procedure described above is a conceptually simple generalization of that of M08. In fact, equation (21) is identical to equation (18) in M08, with the exception that our bins are now defined in the higher-dimensional space including all observable quantities, rather than only survey flux and redshift.

4.2 Mechanics and priors

Using the likelihood function detailed in Section 4.1, cosmological constraints were obtained via Markov chain Monte Carlo (MCMC), employing the Metropolis sampler embedded in the COSMOMC code[14] of Lewis & Bridle (2002). The 2008 May COSMOMC release includes the 5-yr *WMAP* and Union supernova data and analysis codes; an additional module implementing the f_{gas} analysis has also been publicly released (Rapetti et al. 2005, A08).[15] Further modifications were made to include the likelihood codes for the XLF and BAO data. The CMB and matter power spectrum calculations were performed using the CAMB package of Lewis, Challinor & Lasenby (2000),[16] suitably modified to incorporate the evolving w models of Section 3.1 (Rapetti et al. 2005).

The set of parameters describing the complete model is summarized in Table 1, along with the priors used in our analysis.

4.3 Improvements over M08

Compared to our previous work in M08, the constraints on flat ΛCDM and constant w models reported below are tighter by a factor of 2–3. Given that we have used essentially the same cluster samples in the two studies, it is sensible to address the specific differences in the two analyses that result in these improvements. We attribute the change to several closely related improvements in our analysis.

(i) *Cluster mass determination*. In M08, we calibrated the X-ray luminosity–mass relation using the HIFLUGCS data of Reiprich & Böhringer (2002), for which masses were estimated from X-ray data, assuming hydrostatic equilibrium. This procedure is known to underestimate the mass by tens of per cent (e.g. Nagai, Vikhlinin & Kravtsov 2007; Mahdavi et al. 2008), since the intracluster medium is not supported solely by thermal pressure. However, the overall size of this effect is not well established, nor is the variation in non-thermal pressure from cluster to cluster. Consequently, that earlier work was forced to incorporate a large systematic uncertainty in the

---

[10]The normalization factors presented in equations (20) and (27) are different because the two likelihood functions are formally defined over different domains. What is significant is that both functions have the same dependence on model predictions.

[11]However, the observations do have to span the range of values present in the sample reasonably well, in order to adequately probe the scaling relations.

[12]This simplification occurs because the selection function is necessarily not dependent on $\hat{y}_0$. Similarly, in practice, equation (26) needs to be integrated only over observables that influence selection.

[13]Sahlén et al. (2009) demonstrate the improvement in cosmological constraints available to future surveys by incorporating follow-up observations.
[14]http://cosmologist.info/cosmomc/
[15]http://www.stanford.edu/~drapetti/fgas_module/
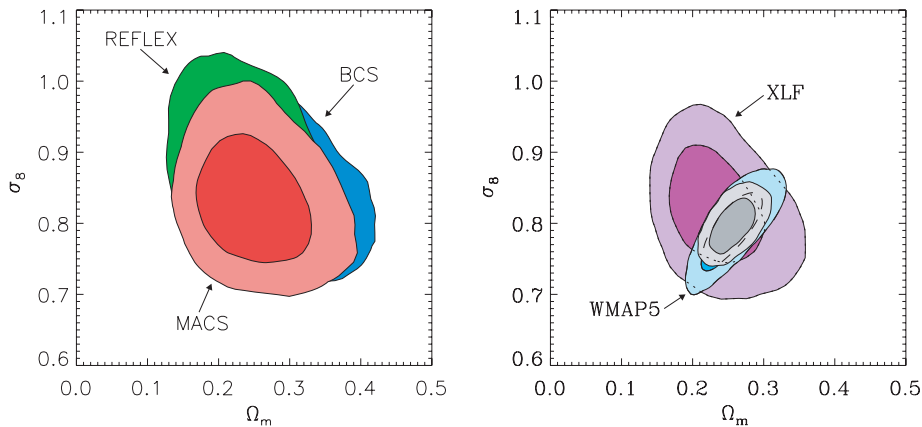[16]http://www.camb.info

Figure 1. Left: joint 68.3 and 95.4 per cent confidence regions for parameters of the ΛCDM cosmology from the BCS (blue), REFLEX (green) and Bright MACS (red) cluster samples individually, including all systematic allowances in Table 1. Note that only the 95.4 per cent confidence regions are visible for BCS and REFLEX. Right: constraints from the full XLF data set (purple) and 5-yr *WMAP* data (blue; Dunkley et al. 2009, marginalized over the SZ signal). Results from the combination of the XLF and *WMAP*5 data are shown in grey.

mass measurements, which significantly degraded constraints on the scaling relation and its intrinsic scatter, resulting in a corresponding degradation in cosmological constraints.

This work avoids this issue by using gas mass as a proxy for total mass (see details in Paper II). Unlike total mass, gas mass can be measured from X-ray data with very little bias. The relation between gas mass and total mass is provided by A08; ultimately, the cluster mass scale is set using the hydrostatic method for the clusters in that work, i.e. the class of hot, dynamically relaxed clusters for which the bias due to non-thermal pressure is minimal.

(ii) *Follow-up data over a range of redshifts*. The statistical methods employed in M08 lack the internal consistency of our new procedure (Section 4.1), making it impossible to rigorously incorporate follow-up data at redshifts much greater than zero.[17] As a result, we could not directly constrain the evolution of the scaling relations, and instead marginalized over a range of possibilities, using conservative priors. With our new method, follow-up data spanning the full redshift range of cluster detections can be included, allowing evolution in the scaling relations to be tested directly (Paper II). Relatedly, the distribution of follow-up data over a range in redshift improves the constraints on the dark energy equation of state.

(iii) *Use of f_{gas} data*. As discussed by Vikhlinin et al. (2009b), the method used to estimate masses has an effect on the constraints obtained because different mass proxies have different dependencies on distance, with gas mass ($M_{gas} \propto d^{2.5}$) being more sensitive than temperature or Y_X. Our results are therefore sensitive to this choice at some level.

Another way to look at this is to notice that the A08 data that we use to constrain f_{gas} inherently contain information about Ω_m in addition (given a bound on the baryonic depletion of clusters). Thus, our results on Ω_m are primarily, though not entirely, driven by these f_{gas} data. With the value of f_{gas} constrained, the XLF data then determine σ_8 and w. Note that, as mentioned in Section 2, we use only six clusters from A08 at redshifts $z < 0.15$; due to this redshift restriction, these f_{gas} data do not produce a constraint on w by themselves.

Although it is not stated explicitly above, we emphasize that the new statistical method outlined in this section is ultimately the most important advantage over M08, since the improvements in follow-up data and mass determination cannot be fully or fairly exploited without it.

5 RESULTS

5.1 Constraints on the ΛCDM model

For a spatially flat, cosmological constant ($w = -1$) model, the joint constraints on Ω_m and σ_8 obtained from the BCS, REFLEX and Bright MACS cluster samples individually are displayed in the left panel of Fig. 1. Results from combining the three samples appear as purple contours in the right panel; the constraints, marginalized over the systematic allowances listed in Table 1, are $\Omega_m = 0.23 \pm 0.04$ and $\sigma_8 = 0.82 \pm 0.05$ (Table 2). These results agree well with the tight constraints obtained from *WMAP*5 data (Dunkley et al. 2009) for this model (blue contours in the right panel). Results from the combination of *WMAP*5 and the XLF (grey contours) are somewhat improved: $\Omega_m = 0.26 \pm 0.02$ and $\sigma_8 = 0.80 \pm 0.02$. Our XLF results are in agreement with recent estimates based on other X-ray selected (Henry et al. 2009; Vikhlinin et al. 2009b) and optically selected (Rozo et al. 2010) cluster samples, and a variety of independent cosmological data (Percival et al. 2007, 2010; A08; Fu et al. 2008; Ho et al. 2008; Kowalski et al. 2008; Hicken et al. 2009; Reid et al. 2010, see also references in M08).

As mentioned in Section 3.1, the scalar spectral index, n_s, is fixed at 0.95 in our analysis of the XLF data alone. The only parameter that is significantly degenerate with n_s is σ_8; our results for values of n_s other than 0.95 can be adequately described by shifting the σ_8 constraints along the linear relation $\sigma_8 = 0.82 + 0.25(n_s - 0.95)$.

The marginalized posterior distribution for Ω_m obtained from the XLF data is compared with those obtained from the analysis of *WMAP*5, f_{gas} (A08), SNIa (Kowalski et al. 2008) and BAO (Percival et al. 2007, also using our standard priors on h and $\Omega_b h^2$) data in Fig. 2. The agreement among all the data sets is good, though note that the XLF and f_{gas} results are not independent (Sections 2 and 4.3).

In Paper II, we show that this simple cosmological model (and the simple scaling relation model of Section 3.3) provides an acceptable

---

[17] In detail, the follow-up data were restricted to low redshifts ($z < 0.11$) where the HIFLUGCS discovery space was large, i.e. where many clusters were found well above the flux limit. This procedure reduces the biasing effects of flux selection, although it does not eliminate them.

Table 2. Marginalized 68.3 per cent confidence intervals on cosmological parameters from our analysis, including all systematic uncertainties in Table 1. The various cosmological models are discussed in Section 3.1. 'All' refers to the combination of XLF, cluster f_{gas}, *WMAP*5, SNIa and BAO data.

Data	Model	Ω_m	σ_8	w, w_0	w_a, w_{et}
XLF	ΛCDM	0.23 ± 0.04	0.82 ± 0.05	–	–
XLF+*WMAP*5	ΛCDM	0.26 ± 0.02	0.80 ± 0.02	–	–
All	ΛCDM	0.257 ± 0.015	0.80 ± 0.02	–	–
XLF	Constant w	0.23 ± 0.04	0.82 ± 0.05	-1.01 ± 0.20	–
XLF+*WMAP*5	Constant w	0.27 ± 0.04	0.78 ± 0.04	-0.95 ± 0.14	–
XLF+f_{gas}	Constant w	0.22 ± 0.04	0.83 ± 0.05	-1.06 ± 0.15	–
All	Constant w	0.272 ± 0.016	0.79 ± 0.03	-0.96 ± 0.06	–
XLF+*WMAP*5	Evolving w (w_a)	0.26 ± 0.04	0.79 ± 0.05	-0.77 ± 0.31	$-0.34^{+0.72}_{-1.42}$
All	Evolving w (w_a)	0.256 ± 0.016	0.80 ± 0.03	-0.93 ± 0.16	$-0.13^{+0.47}_{-0.73}$
XLF+*WMAP*5	Evolving w (w_{et})	0.26 ± 0.04	0.79 ± 0.05	-0.73 ± 0.40	$-1.10^{+0.59}_{-0.39}$
All	Evolving w (w_{et})	0.257 ± 0.016	0.80 ± 0.03	-0.88 ± 0.21	$-1.05^{+0.20}_{-0.36}$
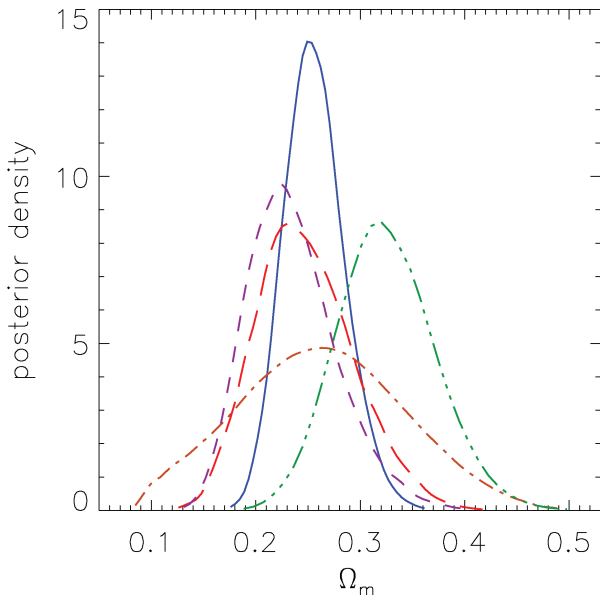
Figure 2. Marginalized posterior distributions for Ω_m in the ΛCDM model from analysis of the XLF data (purple, short dashed line; including systematic allowances in Table 1); 5-yr *WMAP* data (blue, solid line; Dunkley et al. 2009, marginalized over the SZ signal); cluster f_{gas} (red, long dashed line; A08, including conservative systematic allowances); SNIa (green, 3-dot-dashed line; Kowalski et al. 2008, including their treatment of systematics); and BAO (brown, dot-dashed line; Percival et al. 2007, also using our standard priors on h and $\Omega_b h^2$). Note that the XLF and f_{gas} results are not independent (Sections 2 and 4.3).

fit to the XLF data. Constraints on the parameters describing the scaling relations are also presented in that work.

5.2 Constraints on constant *w* models

For models in which the dark energy equation of state, w, is constant in time, the joint constraints on Ω_m and w from the XLF are shown in the left panel of Fig. 3 (purple contours). The constraint on the equation of state from the XLF alone, which is driven by the evolution in the cluster mass function, is $w = -1.0 \pm 0.2$ (including systematics), a factor of 3 improvement over M08. Our results on Ω_m and σ_8 are nearly identical to the ΛCDM case (Table 2); these constraints are driven by the cluster data at low redshift and are thus largely insensitive to the properties of dark energy.

Fig. 3 also demonstrates the agreement of our results with independent constraints from other cosmological data (Section 2), including *WMAP*5 (blue), SNIa (green), cluster f_{gas} (red) and BAO (brown). The combination of these data sets appears as gold contours in the figure; the improved, marginalized constraints from this combination are $\Omega_m = 0.27 \pm 0.02$, $\sigma_8 = 0.79 \pm 0.03$ and $w = -0.96 \pm 0.06$. Our results are also consistent with those derived independently by Vikhlinin et al. (2009b) from X-ray flux-selected clusters.

The constraints in the σ_8-w plane from the XLF and *WMAP*5 data are shown in the right panel of Fig. 3. The complementarity of the two data sets is evident, with the XLF constraint on σ_8 breaking an important degeneracy affecting the CMB data. The combination of only the XLF and *WMAP*5 data yields the grey contours in the figure, corresponding to the one-dimensional constraints $\Omega_m = 0.27 \pm 0.04$, $\sigma_8 = 0.78 \pm 0.04$ and $w = -0.95 \pm 0.14$.

We note that the combination of the two types of cluster data used here, XLF and f_{gas} (with standard priors on h and $\Omega_b h^2$), also places competitive constraints on dark energy: $w = -1.06 \pm 0.15$ (Fig. 4). Interestingly, this result is comparable to that of the combination of XLF with *WMAP*5 data, or f_{gas} with 3-yr *WMAP*, Cosmic Background Imager, Arcminute Cosmology Bolometer Array Receiver and BOOMERanG data ($w = -1.00 \pm 0.14$; A08).

As discussed in Section 3.1, our evaluation of the linear matter power spectrum for models with $w \neq -1$ includes the effects of dark energy density perturbations. For comparison with other works where dark energy perturbations are neglected, we also performed an analysis where dark energy affects only the expansion of space (i.e. where it has uniform density) despite not being a cosmological constant. In this case, we obtain from the XLF the slightly larger constraint $w = -0.97 \pm 0.25$, including all systematic uncertainties in the usual way.

5.3 Constraints on evolving *w* models

Constraints on the two evolving w models discussed in Section 3.1 (equations 1 and 2), (w_0, w_a) and (w_0, w_{et}), are shown respectively in the left and right panels of Fig. 5. Results from the combination of XLF and *WMAP*5 data are shown as grey contours, and the combination of those two with f_{gas}, SNIa and BAO data is shown in gold. In both cases, there is no evidence for evolution in the dark energy equation of state, and the results are consistent with the cosmological constant model (indicated by a cross in the figure). Marginalized constraints are listed in Table 2. For the more general (w_0, w_{et}) model, we find $w_0 = -0.88 \pm 0.21$ and
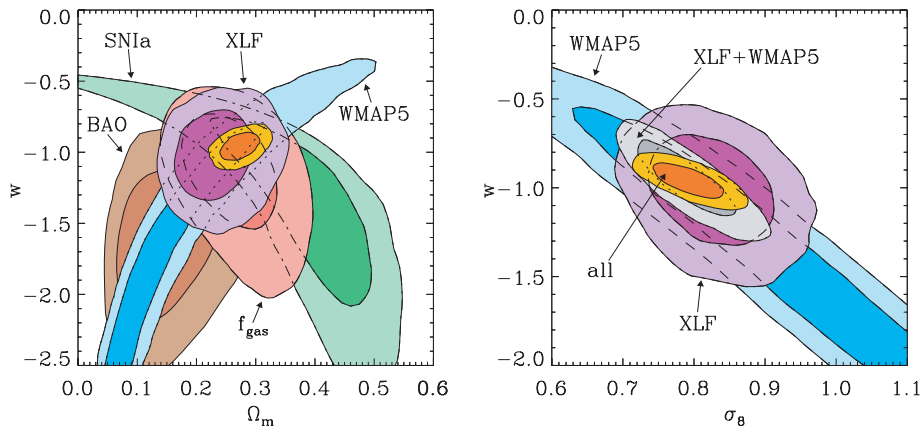
Figure 3. Joint 68.3 and 95.4 per cent confidence regions for parameters of the constant w model. Left: constraints on Ω_m and w from the XLF (purple, including all systematic allowances in Table 1) are compared with those from cluster f_{gas} data (red; A08, including conservative systematic allowances); 5-yr *WMAP* data (blue; Dunkley et al. 2009, marginalized over the SZ signal); SNIa data (green; Kowalski et al. 2008, including their treatment of systematics); and BAO observations, (brown; Percival et al. 2007, also using our standard priors on h and $\Omega_b h^2$). Results from combining these five data sets are shown in gold. Right: constraints on σ_8 and w from XLF and *WMAP*5 data. The combination of the XLF and *WMAP*5 yields the grey contours; adding the other data listed above produces the gold contours.
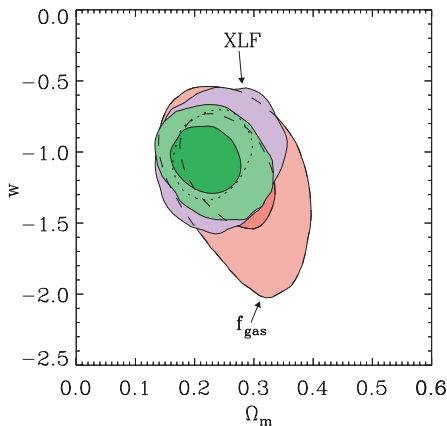
Figure 4. Joint 68.3 and 95.4 per cent confidence regions for parameters of the constant w model, including conservative systematic allowances. Purple contours indicate constraints from the XLF (which includes six $z < 0.15$ f_{gas} clusters; see Sections 2 and 4.3), while constraints from all 42 f_{gas} clusters alone are shown in red. Results combining the XLF data with all 42 f_{gas} clusters appear in green. The combination of these two types of cluster data with standard priors on h and $\Omega_b h^2$ yields a competitive constraint on dark energy, $w = -1.06 \pm 0.15$.

$w_{et} = -1.05^{+0.20}_{-0.36}$ from the combination of all the data (Fig. 6). These results are a significant improvement over previous constraints on this model from the combination of f_{gas}, CMB and SNIa data ($w_0 = -1.05^{+0.31}_{-0.26}$, $w_{et} = -0.83^{+0.48}_{-0.43}$; A08).

Adding spatial curvature as a free parameter and fixing the transition redshift to 1 (equation 1), we obtain constraints equivalent to $\left[\sigma(w_p)\sigma(w_a)\right]^{-1} = 15.5$, using the notation and definitions of the Dark Energy Task Force (DETF; Albrecht et al. 2006).

6 INVESTIGATION OF SYSTEMATICS

6.1 Sensitivity to priors

Our analysis of the XLF data alone includes priors to constrain the Hubble parameter and mean baryon density, and to marginalize over systematic allowances on the mass function, cluster sample completeness and cluster gas mass fraction (Table 1). To investigate the influence of each prior individually, we importance sampled the Markov chains produced in the XLF analysis for the constant w model, reducing the width of each prior in turn by a factor of 2. If importance sampling a particular prior in this way changes the posterior distribution for parameters of interest, we can conclude that the prior is influential.

Of the priors listed above, only the systematic allowances associated with the determination of f_{gas} are significant. Since the gas mass fraction determines the overall mass scale when M_{gas} is used as a proxy for total mass, its systematic uncertainty affects primarily the constraints on Ω_m and σ_8, as shown in Fig. 7. Following A08, our standard analysis uses conservative systematic allowances for *Chandra* calibration (10 per cent), non-thermal pressure support (10 per cent), the depletion of baryons in clusters with respect to the cosmic mean (20 per cent) and evolution with redshift in the baryon fraction and stellar content of clusters (10 and 20 per cent). The most significant ones for this work are on the depletion of baryons in clusters, which determines the width of the Ω_m constraint, and the amount of non-thermal pressure and overall calibration of *Chandra* for temperature measurements, which determine the constraint on σ_8 at fixed Ω_m.

In contrast, none of the priors affects the constraint on w significantly. This was confirmed by an additional analysis in which all of the systematic uncertainties were eliminated, i.e. every nuisance parameter was fixed rather than being marginalized over. Ordinarily, this would indicate that our constraints on dark energy are statistically limited, and this may plausibly be the case; however, we note that our constraints may also be limited by the fact that the effect of dark energy perturbations on the mass function is not yet understood. If these perturbations produce changes in the high-mass tail of the mass function at the few tens of per cent level, there may be additional constraining power available from the current XLF data.

6.2 Choice of mass function

In our standard analysis, we have used the mass function of Tinker et al. (2008) with cluster radius defined by an overdensity of $300\Omega_m(z)$. One could also justify using the mass function determined at a cluster radius closer to the radius where the
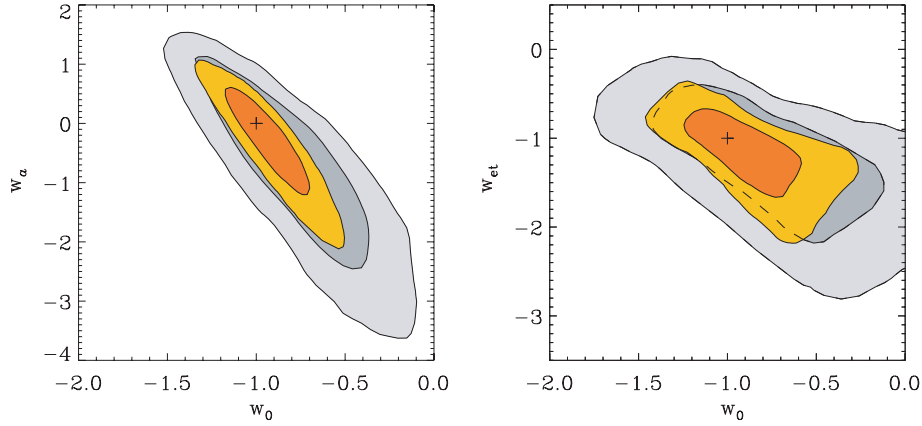
Figure 5. Joint 68.3 and 95.4 per cent confidence regions, including conservative systematic allowances, for parameters of the evolving w models. Left: constraints on the (w_0, w_a) model (equation 1) from the combination of XLF and 5-yr *WMAP* data are shown in grey, as well as the combination of XLF, *WMAP*5 (Dunkley et al. 2009), cluster f_{gas} (A08), SNIa (Kowalski et al. 2008) and BAO (Percival et al. 2007) data (gold). Right: constraints on the (w_0, w_{et}) model (equation 2). The transition scalefactor is marginalized over the range $0.5 < a_t < 0.95$. Crosses in each panel indicate the ΛCDM model, with constant $w = -1$.
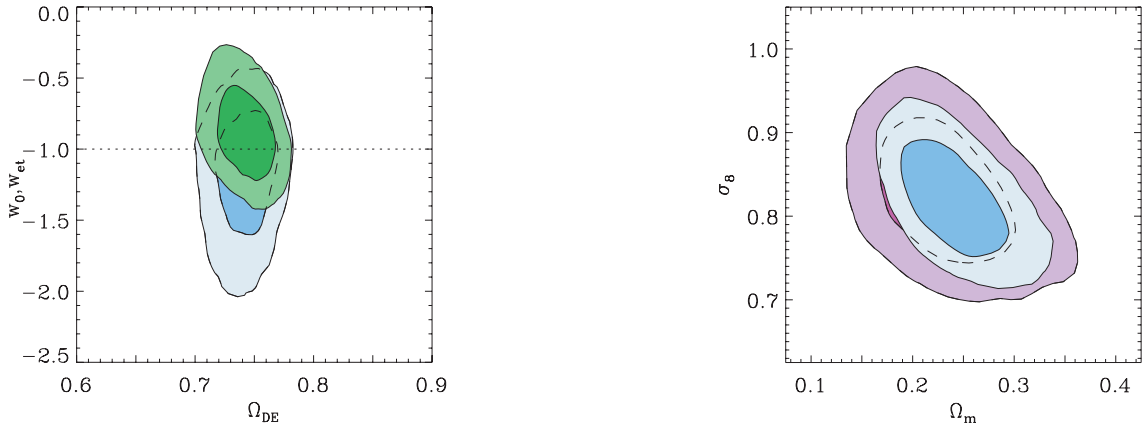
Figure 6. Joint 68.3 and 95.4 per cent confidence regions from the combination of XLF, f_{gas}, *WMAP*5, SNIa and BAO data, including systematic uncertainties, for parameters of the (w_0, w_{et}) model (equation 2). The results are the same as in the right panel of Fig. 5, but with both w_0 (green contours) and w_{et} (blue) shown on the y-axis, against the dark energy density, Ω_{DE}, on the x-axis. The transition scalefactor is marginalized over the range $0.5 < a_t < 0.95$. The horizontal, dotted line indicates the cosmological constant model ($w_0 = w_{et} = -1$).

Figure 7. Joint 68.3 and 95.4 per cent confidence regions for Ω_m and σ_8 in the constant w model from the XLF data, including the standard systematic allowances in Table 1, are shown in purple. The blue contours result from reducing the width of all of the allowances associated with the f_{gas} model of A08 by a factor of 2. The figure demonstrates that our results on Ω_m and σ_8 are limited by the systematic uncertainty in f_{gas}. (These allowances have no effect on the w constraint.) Improvements in hydrodynamical simulations and the incorporation of gravitational lensing mass measurements offer the possibility of significantly reducing the uncertainty in f_{gas}, and thus improving the constraints.

follow-up mass measurements are made, r_{500}. We therefore repeated the XLF analysis using the mass function given by Tinker et al. for overdensity $1600\Omega_m(z)$. For Ω_m values of 0.2–0.3 and $z < 0.5$, this corresponds to overdensities of 300–1000, compared to 50–200 for $300\Omega_m(z)$. Our results with this alternative mass function were virtually identical to the standard results.

Similarly, we repeated the analysis using the older mass function of Jenkins et al. (2001) at overdensity $324\Omega_m(z)$. Here we used a 20 per cent Gaussian systematic allowance on only the normalization of the mass function and fixed the shape, as in M08. Again, the cosmological results were essentially identical to our standard results, despite the fact that the Jenkins function does not include evolution with redshift, while the Tinker function does. This indifference to the details of the mass function is consistent with the observation in Section 6.1 that the current results are insensitive to the systematic allowances on the mass function and its evolution.

6.3 Note on *Chandra* calibration

A major, recent (2009 January 21) update to the *Chandra* ACIS effective area at soft energies is accounted for in our analysis of follow-up *Chandra* observations for the XLF data (see Paper II). However, the raw data analysis in A08 predates this calibration update, which in our tests typically results in an increase in the inferred gas fraction of ~10 per cent. We have accounted for this correction by shifting the centre of the '*Chandra* calibration' nuisance parameter in the A08 model such that the preferred value of f_{gas} increases correspondingly by 10 per cent. The Gaussian prior on this parameter also has a width of 10 per cent, so this systematic allowance encompasses both the old value and the new value expected from the calibration update.

The effect of this higher gas mass fraction is to shift the Ω_m constraint from the f_{gas} analysis to lower values by approximately 0.02. A corresponding shift appears in the XLF results, since the Ω_m constraints are partially driven by our use of a subset of the f_{gas} data to set the overall cluster mass scale. The correction additionally results in a shift of similar magnitude towards lower σ_8 values, but has no effect on the determination of w. While we expect that the 10 per cent Gaussian allowance adequately reflects the systematic uncertainty in the overall *Chandra* calibration, these trends should be kept in mind in interpreting our results. For example, if we have overestimated the effect that the calibration update has on the value of f_{gas}, our best-fitting results will shift slightly to higher Ω_m and σ_8.

7 CONCLUSION

We have presented cosmological constraints obtained from an X-ray flux-limited sample of 238 massive galaxy clusters spanning the redshift range $z < 0.5$. Follow-up *Chandra* or *ROSAT* X-ray observations of 94 of these clusters are incorporated, as detailed in Paper II. Our analysis accounts for all selection biases, includes conservative allowances for systematic uncertainties, and, for the first time, produces simultaneous constraints on cosmology and the cluster scaling relations using a rigorous and fully self-consistent statistical method. The incorporation of follow-up data and our improved analysis method results in cosmological constraints that are a factor of 2–3 better than our previous results in M08, which were based on the same flux-limited sample of clusters. The results presented here are among the tightest and most robust constraints on cosmological parameters available from current data.

The constraints on spatially flat, cosmological constant models from our XLF data are $\Omega_m = 0.23 \pm 0.04$ and $\sigma_8 = 0.82 \pm 0.05$. Introducing a constant dark energy equation of state, w, as a free parameter, we find $w = -1.0 \pm 0.2$, obtaining the same tight constraints on Ω_m and σ_8. These results confirm at higher precision the first constraints on the dark energy equation of state from the analysis of XLF data reported by M08, and are consistent with independent findings based on cluster gas mass fractions, CMB anisotropies, type Ia supernovae, baryon acoustic oscillations, galaxy redshift surveys and cosmic shear. We also find good agreement with recent, independent analyses of X-ray selected galaxy clusters (Henry et al. 2009; Vikhlinin et al. 2009b) and optically selected clusters (Rozo et al. 2010).

Using the combination of XLF, f_{gas}, *WMAP*5, SNIa and BAO data, we investigate evolving w models, including the standard (w_0, w_a) model as well as a (w_0, w_{et}) model in which the transition scalefactor, a_t, is marginalized over. In both cases, the best current cosmological data remain consistent with a constant value of $w = -1$. The constraints on the non-flat (w_0, w_a) model from the combination of current data are equivalent to DETF figure of merit $[\sigma(w_p)\sigma(w_a)]^{-1} = 15.5$. Compared to the figure of merit of 8.3 reported by Wang (2008) from the combination of *WMAP*5, SNIa and BAO data, this result highlights the contribution of cluster data (XLF and f_{gas}), and their complementarity to other cosmological information.

Currently, the most significant systematic limitation on our results for Ω_m and σ_8 is systematic uncertainty in the cluster gas mass fraction, f_{gas}. Specifically, the dominant uncertainty in the determination of Ω_m is in the depletion of baryons in clusters relative to the cosmic mean, for which a 20 per cent systematic allowance is adopted in the present work. The prospect of significant improvement here exists, as numerical simulations incorporate more complete treatments of baryonic physics, especially feedback and star formation processes in the centres of clusters. The results on σ_8 are limited by our knowledge of the non-thermal pressure in clusters and the overall calibration of X-ray temperature measurements; cluster mass measurements via gravitational lensing provide an observational means of circumventing this issue, by directly constraining the overall mass scale of the data.

XLF results on dark energy remain statistically limited. The incorporation of additional X-ray, SZ (e.g. *Planck*, the South Pole Telescope and the Atacama Cosmology Telescope) and optical/near-infrared (e.g. the Dark Energy Survey and Pan-STARS) cluster surveys could lead to significant near-term improvements, although we stress that these surveys must have well-understood selection functions, sampling models and scaling relations. The possibility also remains that cosmological simulations of the mass function including the effects of dark energy perturbations will reveal additional constraining power. It is therefore important to pursue a theoretical understanding of the effects of fluid dark energy on non-linear structure formation, even as much larger X-ray (eROSITA[18]), and optical (e.g. the Large Synoptic Survey Telescope) cluster surveys promise to increase the data available for cluster cosmology by orders of magnitude.

We note that the statistical approach described in this work has significance for the planning of follow-up observations for future cluster surveys. Our method permits the use of partial follow-up, allowing scaling relations to be directly and robustly constrained over a wide range in mass without raising the cost in exposure time prohibitively, as would be the case for a complete, exhaustive follow-up campaign. At the same time, the incorporation of follow-up data provides a significant advantage over pure self-calibration approaches. In detail, any plan for X-ray follow-up observations of future surveys should also account for the additional benefit of observing dynamically relaxed clusters suitable for the f_{gas} test; as we have shown, the combination of cluster f_{gas} data, which directly probe the expansion of the Universe, and the XLF, which measures the growth of structure, can provide precise cosmological constraints independent of CMB and SNIa observations.

Readers interested in the simultaneous constraints on cluster scaling relations produced in this work should refer to Paper II. MCMC samples encoding the results of these two papers will be made available for download on the web.[19]

[18]http://www.mpe.mpg.de/projects.html#erosita
[19]http://www.stanford.edu/group/xoc/papers/xlf2009.html

REFERENCES

Abramo L. R., Batista R. C., Rosenfeld R., 2009, J. Cosmology Astropart. Phys., 7, 40

Adelman-McCarthy J. K. et al., 2007, ApJS, 172, 634

Albrecht A. et al., 2006, preprint (astro-ph/0609591)

Alimi J., Füzfa A., Boucher V., Rasera Y., Courtin J., Corasaniti P., 2010, MNRAS, 401, 775

Allen S. W., Schmidt R. W., Fabian A. C., Ebeling H., 2003, MNRAS, 342, 287

Allen S. W., Rapetti D. A., Schmidt R. W., Ebeling H., Morris R. G., Fabian A. C., 2008, MNRAS, 383, 879 (A08)

Astier P. et al., 2006, A&A, 447, 31

Barris B. J. et al., 2004, ApJ, 602, 571

Böhringer H. et al., 2004, A&A, 425, 367

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440

Borgani S. et al., 2001, ApJ, 561, 13

Bryan G. L., Norman M. L., 1998, ApJ, 495, 80

Chevallier M., Polarski D., 2001, Int. J. Modern Phys. D, 10, 213

Cohn J. D., White M., 2008, MNRAS, 385, 2025

Colless M. et al., 2001, MNRAS, 328, 1039

Colless M. et al., 2003, preprint (astro-ph/0306581)

Creminelli P., D'Amico G., Noreña J., Senatore L., Vernizzi F., 2010, J. Cosmol. Astropart. Phys., 3, 27

Dunkley J. et al., 2009, ApJS, 180, 306

Ebeling H., Edge A. C., Bohringer H., Allen S. W., Crawford C. S., Fabian A. C., Voges W., Huchra J. P., 1998, MNRAS, 301, 881

Ebeling H., Edge A. C., Henry J. P., 2001, ApJ, 553, 668

Ebeling H., Edge A. C., Mantz A., Barrett E., Henry J. P., Ma C. J., van Speybroeck L., 2010, in press (doi:10.1111/j.1365-2966.2010.16920.x) (arXiv:1004.4683)

Evrard A. E. et al., 2002, ApJ, 573, 7

Evrard A. E. et al., 2008, ApJ, 672, 122

Freedman W. L. et al., 2001, ApJ, 553, 47

Fu L. et al., 2008, A&A, 479, 9

Gao L., Navarro J. F., Cole S., Frenk C. S., White S. D. M., Springel V., Jenkins A., Neto A. F., 2008, MNRAS, 387, 536

Garnavich P. M. et al., 1998, ApJ, 509, 74

Gelman A., Carlin J. B., Stern H. S., Rubin D. B., 2004, Bayesian Data Analysis. Chapman & Hall/CRC, Boca Raton, FL

Hamuy M. et al., 1996, AJ, 112, 2408

Henry J. P., 2004, ApJ, 609, 603

Henry J. P., Evrard A. E., Hoekstra H., Babul A., Mahdavi A., 2009, ApJ, 691, 1307

Hicken M., Wood-Vasey W. M., Blondin S., Challis P., Jha S., Kelly P. L., Rest A., Kirshner R. P., 2009, ApJ, 700, 1097

Hill R. S. et al., 2009, ApJS, 180, 246

Hinshaw G. et al., 2009, ApJS, 180, 225

Ho S., Hirata C., Padmanabhan N., Seljak U., Bahcall N., 2008, Phys. Rev. D, 78, 043519

Holder G., 2006, preprint (arXiv:0602251)

Hu W., Kravtsov A. V., 2003, ApJ, 584, 702

Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372

Jha S. et al., 2006, AJ, 131, 527

Kaiser N., 1986, MNRAS, 222, 323

Kelly B. C., 2007, ApJ, 665, 1489

Kirkman D., Tytler D., Suzuki N., O'Meara J. M., Lubin D., 2003, ApJS, 149, 1

Klypin A., Macciò A. V., Mainini R., Bonometto S. A., 2003, ApJ, 599, 31

Knop R. A. et al., 2003, ApJ, 598, 102

Komatsu E. et al., 2009, ApJS, 180, 330

Kowalski M. et al., 2008, ApJ, 686, 749

Krisciunas K. et al., 2001, AJ, 122, 1616

Krisciunas K. et al., 2004a, AJ, 127, 1664

Krisciunas K. et al., 2004b, AJ, 128, 3034

Kuhlen M., Strigari L. E., Zentner A. R., Bullock J. S., Primack J. R., 2005, MNRAS, 357, 387

Lewis A., Bridle S., 2002, Phys. Rev. D, 66, 103511

Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473

Linder E. V., 2003, Phys. Rev. Lett., 90, 091301

Linder E. V., Jenkins A., 2003, MNRAS, 346, 573

Łokas E. L., Bode P., Hoffman Y., 2004, MNRAS, 349, 595

Lukić Z., Heitmann K., Habib S., Bashinsky S., Ricker P. M., 2007, ApJ, 671, 1160

Mahdavi A., Hoekstra H., Babul A., Henry J. P., 2008, MNRAS, 384, 1567

Mantz A., Allen S. W., Ebeling H., Rapetti D., 2008, MNRAS, 387, 1179 (M08)

Mantz A., Allen S. W., Ebeling H., Rapetti D., Drlica-Wagner A., 2010a, MNRAS, in press (doi:10.1111/j.1365-2966.2010.16993.x) (arXiv:0909.3099) (Paper II, this issue)

Mantz A., Allen S. W., Rapetti D., 2010b, MNRAS, in press (doi:10.1111/j.1365-2966.2010.16794.x) (arXiv:0911.1788) (Paper IV, this issue)

Miknaitis G. et al., 2007, ApJ, 666, 674

Nagai D., Vikhlinin A., Kravtsov A. V., 2007, ApJ, 655, 98

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493 (NFW)

Nolta M. R. et al., 2009, ApJS, 180, 296

Pacaud F. et al., 2007, MNRAS, 382, 1289

Park C., Hwang J., Lee J., Noh H., 2009, Phys. Rev. Lett., 103, 151303

Percival W. J. et al., 2007, ApJ, 657, 51

Percival W. J. et al., 2010, MNRAS, 401, 2148

Perlmutter S. et al., 1999, ApJ, 517, 565

Press W. H., Schechter P., 1974, ApJ, 187, 425

Rapetti D., Allen S. W., Weller J., 2005, MNRAS, 360, 555

Rapetti D., Allen S. W., Mantz A., Ebeling H., 2009, MNRAS, 400, 699

Rapetti D., Allen S. W., Mantz A., Ebeling H., 2010, MNRAS, in press (doi:10.1111/j.1365-2966.2010.16799.x) (arXiv:0911.1787) (Paper III, this issue)

Reed D. S., Bower R., Frenk C. S., Jenkins A., Theuns T., 2007, MNRAS, 374, 2

Reid B. A. et al., 2010, MNRAS, 404, 60

Reiprich T. H., Böhringer H., 2002, ApJ, 567, 716

Riess A. G. et al., 1998, AJ, 116, 1009

Riess A. G. et al., 1999, AJ, 117, 707

Riess A. G. et al., 2004, ApJ, 607, 665

Riess A. G. et al., 2007, ApJ, 659, 98

Rozo E. et al., 2010, ApJ, 708, 645

Sahlén M. et al., 2009, MNRAS, 397, 577

Schmidt B. P. et al., 1998, ApJ, 507, 46

Schmidt F., Vikhlinin A., Hu W., 2009, Phys. Rev. D, 80, 083505

Schuecker P., Böhringer H., Collins C. A., Guzzo L., 2003, A&A, 398, 867

Sheth R. K., Tormen G., 1999, MNRAS, 308, 119

Spergel D. N. et al., 2007, ApJS, 170, 377

Stanek R., Evrard A. E., Böhringer H., Schuecker P., Nord B., 2006, ApJ, 648, 956

Stanek R., Rudd D., Evrard A. E., 2009, MNRAS, 394, L11

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709

Tonry J. L. et al., 2003, ApJ, 594, 1

Trümper J., 1993, Sci, 260, 1769

Vikhlinin A. et al., 2009a, ApJ, 692, 1033

Vikhlinin A. et al., 2009b, ApJ, 692, 1060

Voevodkin A., Vikhlinin A., 2004, ApJ, 601, 610

Wang Y., 2008, Phys. Rev. D, 77, 123525

White S. D. M., Navarro J. F., Evrard A. E., Frenk C. S., 1993, Nat, 366, 429

Wood-Vasey W. M. et al., 2007, ApJ, 666, 694

Wu H., Rozo E., Wechsler R. H., 2010, ApJ, 713, 1207

Zhao D. H., Jing Y. P., Mo H. J., Börner G., 2003, ApJ, 597, L9

Zhao D. H., Jing Y. P., Mo H. J., Börner G., 2009, ApJ, 707, 354

Zwicky F., 1937, ApJ, 86, 217

This paper has been typeset from a TEX/LATEX file prepared by the author.