Cosmology with the largest galaxy cluster surveys: Going beyond Fisher matrix forecasts

Satej Khedekar[1] **and Subhabrata Majumdar**[2]

[1]MPI für Astrophysik, Karl-Schwarzschild str. 1, Garching, 85741, Germany
[2]Tata Institute of Fundamental Research
Homi Bhabha Road, Colaba, Mumbai - 400005, India.

E-mail: satej@mpa-garching.mpg.de, subha@tifr.res.in

Abstract. We make the first detailed MCMC likelihood study of cosmological constraints that are expected from some of the largest, ongoing and proposed, cluster surveys in different wave-bands and compare the estimates to the prevalent Fisher matrix forecasts. Mock catalogs of cluster counts expected from the surveys – eROSITA, WFXT, RCS2, DES and Planck, along with a mock dataset of follow-up mass calibrations are analyzed for this purpose. A fair agreement between MCMC and Fisher results is found only in the case of minimal models. However, for many cases, the marginalized constraints obtained from Fisher and MCMC methods can differ by factors of 30-100%. The discrepancy can be alarmingly large for a time dependent dark energy equation of state, $w(a)$; the Fisher methods are seen to underestimate the constraints by as much as a factor of 4–5. Typically, Fisher estimates become more and more inappropriate as we move away from ΛCDM, to a constant-w dark energy to varying-w dark energy cosmologies. Fisher analysis, also, predicts incorrect parameter degeneracies. There are noticeable offsets in the likelihood contours obtained from Fisher methods that is caused due to an asymmetry in the posterior likelihood distribution as seen through a MCMC analysis. From the point of mass-calibration uncertainties, a high value of unknown scatter about the mean mass-observable relation, and its redshift dependence, is seen to have large degeneracies with the cosmological parameters σ_8 and $w(a)$ and can degrade the cosmological constraints considerably. We find that the addition of mass-calibrated cluster datasets can improve dark energy and σ_8 constraints by factors of 2–3 from what can be obtained from CMB+SNe+BAO only . Finally, we show that a joint analysis of datasets of two (or more) different cluster surveys would significantly tighten cosmological constraints from using clusters only. Since, details of future cluster surveys are still being planned, we emphasize that optimal survey design must be done using MCMC analysis rather than Fisher forecasting.

Contents

1   Introduction

Galaxy cluster counts have been recognized as potentially important probes of precision cosmology [1–6], and have yielded in recent years, the first set of independent constraints on Ω_m, the matter density of the Universe and σ_8, the RMS density fluctuations on a scale of 8 h^{-1} Mpc [7–10]. There are many large scale structure (LSS) surveys that are either operational or being built, which are expected to yield a much larger dataset of cluster observations. This would provide strong constraints not only on Ω_m and σ_8, but also on the dark energy equation of state (w_0) and its evolution (w_a/w_z). There have also been recent works which demonstrate that cluster data can also be used to place constraints on the extensions of the standard cosmological model, for e.g. probing non-Gaussianities in primordial density perturbations [11–14], modifications to general relativity [15, 16] and to probe neutrino properties [17]. The cosmological constraints obtained from galaxy clusters are complementary to those obtained from other cosmological probes such as cosmic microwave background (CMB), supernovae

type Ia observations and baryon acoustic oscillations. This is because clusters have the dual advantage of not only probing the expansion history but also the growth of perturbations in the universe. Using datasets from various different probes helps in breaking degeneracies between various cosmological parameters, and gives tighter constraints [3, 18–23].

In recent years there have been a large number of cluster surveys running, funded or proposed with the aim of detecting between thousand to hundreds of thousands of clusters up to high redshifts ($z \gtrsim 1$) [1]. With the newer surveys being proposed, it is important to asses the cosmological relevance of such observations. Fisher based estimates of the constraints have been widely used in making forecasts because of both simplicity and lower computational costs [24]. However, Fisher estimates are reliable only when the underlying probability distribution function (pdf) is Gaussian shaped, while MCMC forecasts are always correct irrespective of the pdf of the likelihood. MCMC techniques can be used on real data whereas Fisher analysis is limited to making forecasts. Detailed comparisons of forecasts from Fisher and MCMC have been made in the context of CMB in [25] and SNe, BAO and weak lensing in [26]. However, no such comparison has been made for cluster surveys and in this work, we make a first detailed comparison for cluster surveys. Unlike the CMB where there are a few nuisance parameters like τ (optical depth) and $Y_{\rm He}$ (Helium abundance), analysis of cluster data requires a larger number of such parameters to arrive at unbiased cosmological parameter values. This is due to the significant non-linear astrophysical processes occurring in the intra-cluster medium (ICM) (much of which are still not well understood), as compared to the CMB physics which is both linear and well understood.

In this paper, we compare results on parameter constraints from both Fisher and MCMC methods for various cluster surveys in x-rays, SZE and optical. We make detailed comparison on a case by case basis for various cosmological models like ΛCDM[2], wCDM[3] and also w_aCDM(w_zCDM)[4]. The rest of the paper is organized as follows. In section 2 we begin by describing how clusters are used to probe cosmology. In section 3 and 4 we summarize the methodology of MCMC and Fisher analysis in the context of cluster data and also mention the advantages and disadvantages of each method. In section 5 we describe our fiducial cosmological models, and the priors used in this work; we also outline here our procedure for constructing the mock mass follow up catalog. Next, in section 6 we move on to briefly describe the six surveys across multiple wavelength bands that have been considered in the paper. Our results are described in detail in Section 7 and are also highlighted in Tables 2 - 6, which also offer a detailed comparison between MCMC and Fisher forecasts. Here, we also provide some examples of how the synergies of two independent cluster surveys may be utilized to break the parameter degeneracies to obtain significantly better cosmological constraints. We discuss some technical points relevant to our work in section 10 before concluding in section 11.

---

[1]www.darkenergysurvey.org   www.rcs2.org   www.mpe.mpg.de/eROSITA   wfxt.pha.jhu.edu www.rssd.esa.int/index.php?project=planck   pole.uchicago.edu   www.princeton.edu/act

[2]Cold dark matter with the cosmological constant as the dark energy component.

[3]Cold dark matter with a single parameter dark energy equation of state, w_0.

[4]Cold dark matter with an evolving dark energy equation of state, described by two parameters – w_0 and $w_a(w_z)$, see section 5 for more details.

2 Redshift distribution of clusters as a probe of cosmology

The redshift distribution of clusters is given by,

$$\frac{dN}{dz}(z) = \Delta\Omega \frac{dV(z)}{dzd\Omega} \int_{M_{\text{lim}}}^{\infty} \frac{dn(M,z)}{dM} dM \tag{2.1}$$

where $dV/dzd\Omega$ is the comoving volume element and $\Delta\Omega$ is the solid angle of the survey. This $dV/dzd\Omega$ is related to the expansion history of the background cosmology and depends on cosmological parameters – $\Omega_{tot}, \Omega_m, h$ and $w(z)$. The halo mass function $\frac{dn}{dM}$ shows exponential sensitivity to the growth of perturbations in the matter density field through σ_8 and the growth function $G(z)$, which again depends on the cosmological parameters. There have been a number of fitting forms [27–31] available in literature obtained from both semi-analytic as well as large N-body simulations; in this work we use the fitting form as given in reference [32].

As the mass, M, is not directly observed for most of the detected clusters in the sample, it is convenient to use a more readily observable proxy for mass, $\mathcal{O}$, such as luminosity, temperature or the product of gas mass and temperature in x-rays, Y_X [33]; integrated Y parameter in SZE or max S/N in matched filtering of SZ maps [34]; and richness of red galaxies in optical [35, 36]. Often, the relationship between mass and its observable proxy is described a simple power law form as, $\log M = A + \alpha \log \mathcal{O} + \gamma \log(1+z)$. Here A is the amplitude, α is the slope while the parameter γ captures any non-standard (departure from self-similar) cluster evolution [37]. Since the cluster mass function decreases steeply with mass, a knowledge of the distribution of clusters about this relation (or scatter) becomes very important. One frequently uses a log-normal model to account for this scatter between mass and its proxy. While in some cases, the values of these cluster scaling parameters may be known from past observations to within some error, a slightly different choice of the cluster parameters A, α and γ from their *true* values could bring in a significant bias in the derived cosmology. This occurs due to the strong correlation between the cluster and cosmology parameters. Thus, these scaling parameters should preferably not be fixed, but derived from the data along with the cosmology. However, attempts at marginalizing over these cluster parameters to determine the cosmology from only cluster counts, results in a significantly weakened set of cosmological constraints. There have been several ideas that have been proposed to break this cluster physics – cosmology degeneracy through self-calibration; for e.g. adding cluster power-spectrum [38, 39], counts-in-cells [40], binning in mass proxy [2] as well as redshift, planning a survey with both deep and wide component [41], etc. Even a simple mass follow-up of clusters can be very effective at breaking this degeneracy and can give a significant improvement in the cosmological constraints [42]; here, we construct a mass follow-up [37, 39], to calibrate the mass – observable relation. Since both weak lensing as well as x-ray observations can measure the mass like quantity $M_f(\theta) = M(\theta)/d_A(z)$, where $M(\theta)$ is the halo mass within an angle θ, we use this quantity as the observable for constructing our mock catalogs; henceforth we shall refer to it as the mass follow-up. We assume a flat 30% error estimate on M_f in all our mass catalogs constructed for each of the surveys considered in section 6. Note, that this is very conservative since with better observations one would be able to achieve more robust estimates of cluster masses [43].

3   Monte Carlo analysis of mock data

An accurate and computationally feasible mapping of the underlying multivariate distribution in a high dimensional parameter space is achieved by constructing Markov Chain Monte Carlo (MCMC) chains. A MCMC chain generates a set of points in the parameter space which have the same distribution as the target distribution (posterior likelihood). Our MCMC chains are obtained by computing the Bayesian likelihood at random points selected using the Metropolis Hastings (MH) algorithm [44, 45]. For other sampling algorithms used in the context of cosmological parameter estimation we refer the reader to [46–52]. At each point $\Theta = (\theta_1, \theta_2, ...\theta_n)$ the Bayesian posterior likelihood, $\mathcal{L}(\Theta \mid \mathcal{D})$ is computed given the data $\mathcal{D}$ as, $\mathcal{L}(\Theta \mid \mathcal{D}) \propto \mathcal{L}(\mathcal{D} \mid \Theta)\mathcal{L}(\Theta)$. Here $\mathcal{L}(\Theta)$ is the prior likelihood and $\mathcal{L}(\mathcal{D} \mid \Theta)$ is the likelihood of getting the data $\mathcal{D}$ given the parameter θ. For number count observations of clusters this is given by the Poisson statistic,

$$\mathcal{L}(\mathcal{D} \mid \Theta) = \prod_m \frac{N_m(\Theta)^{N_m^{(\mathcal{D})}} e^{-N_m(\Theta)}}{N_m^{(\mathcal{D})}} \tag{3.1}$$

Here, the product is over the redshift bins; $N_m^{(\mathcal{D})}$ corresponds to the observed number of clusters in the redshift m, while $N_m(\Theta)$ is the expected value for a model described with the parameter vector Θ. We add to this the chi-sq likelihoods $\mathcal{L} = \exp(-\chi^2/2)$ constructed from the mass follow up data.

The MH algorithm requires a proposal pdf to be specified, which we choose to be a multivariate Gaussian distribution for simplicity. For a quicker convergence, the proposal distribution should be as close to the actual distribution. In practice, this is achieved by making repeated short runs, each time using the computed covariance matrix from the data to be the new proposal matrix. One may also invert the Fisher matrix computed for the same observations in order to obtain the proposal covariance matrix, see section 4. The number of iterations required for convergence roughly scales as the number of parameters of the model and also depends on the nature of the actual distribution. Longer chains need to be constructed when the posterior likelihood departs from a Gaussian distribution. On having sampled sufficiently enough points in the parameter space the MCMC chain equilibrates to the target distribution, i.e. $\mathcal{L}(\Theta \mid \mathcal{D})$, and is said to have reached convergence. In practice, convergence may be tested thorough one of the many convergence tests, see section 10. We refer the reader to the references [53, 54] for further details on the MCMC technique.

4   Fisher matrix analysis

The Fisher information matrix [55, 56] is often used to obtain a quick estimate of the constraints on the parameters of a model for a given experiment. The Fisher information matrix is defined about a fiducial cosmological model as,

$$\mathcal{F}_{ij} \equiv -\left\langle \frac{\partial^2 \ln \mathcal{L}(\Theta \mid \mathcal{D})}{\partial \theta_i \partial \theta_j} \right\rangle \tag{4.1}$$

For a large dataset the likelihood function may be approximated as a multivariate Gaussian pdf of the model parameters, in which case the covariance matrix of the model parameters may be constructed as the inverse of the Fisher matrix, $\mathcal{C}_{ij} \equiv \langle \theta_i \theta_j \rangle - \langle \theta_i \rangle \langle \theta_j \rangle = \mathcal{F}_{ij}^{-1}$. Using the Cramer-Rao inequality one estimates the marginalized standard deviation (s.d.) of the

parameters as follows, $\Delta\theta \equiv \langle\theta_i^2\rangle - \langle\theta_i\rangle\langle\theta_i\rangle \geq \mathcal{C}_{ii}^{1/2}$. For galaxy cluster surveys the Fisher matrix may be constructed as [1, 37, 39],

$$\mathcal{F}_{ij} = \Sigma_m \frac{\partial N_m}{\partial p_i} \frac{\partial N_m}{\partial p_j} \frac{1}{N_m} + \Sigma_k \frac{\partial M_f^{(k)}}{\partial p_i} \frac{\partial M_f^{(k)}}{\partial p_j} \frac{1}{\sigma_{M_f^{(k)}}^2} + \frac{\delta_{ij}}{\sigma^2(\theta_i)} \qquad (4.2)$$

where N_m is the number of observed clusters in each redshift bin m; k is the summation index over all the follow-up masses; and $\sigma(\theta_i)$ is the Gaussian prior on the parameter θ_i.

Fisher matrix analysis has the advantage of being computationally simpler as it is calculated just about a single point, i.e. for the fiducial model parameters. However it must be emphasized that, this technique provides a reasonably accurate estimate of parameter constraints only when the likelihood function is close to being Gaussian. When 3rd or higher order moments of the parameters start becoming important the Fisher analysis may fail. In practice this happens when the model is described by a large number of degenerate parameters, leading to weakened constraints due to extended degeneracy between the parameters. However, the likelihood distribution of the combination of many independent parameters is always more Gaussian than the one of a single parameter [for e.g. ref. 56]. Non-Gaussian likelihoods are also seen when the data is not sufficiently large; or for non-Gaussian distribution of errors in the data. Another disadvantage of Fisher analysis is that it can handle only Gaussian priors; many a times, there are parameters which can take values only within a certain range, as other values may not be physical, for e.g., in a flat universe, $0 \leq \Omega_m \leq 1$. Such flat priors can only be imposed through an MCMC analysis. In our results we have taken care to see that the comparisons between Fisher and MCMC results are not affected by such issues.

5   Fiducial cosmology, priors and mass follow-up

We adopt our fiducial cosmology from the WMAP 7-year results (Table 6 of [57]). For simplicity, we choose a flat Universe since for an open wCDM model, WMAP7+BAO+H0 tightly constraints $\Delta\Omega_{\rm tot} \leq 0.007$. However, including the flatness as an additional parameter Ω_k with the WMAP prior does not change our results significantly.

We vary the following cosmological parameters - $\Omega_m, \Omega_b, w(z), h, n_s, \sigma_8$ along with cluster parameters A, α, γ. Since cluster counts by themselves do not constrain the cosmological parameters h, Ω_b and n_s very well, we impose WMAP Gaussian priors with s.d. 0.013, 0.0016 and 0.013 on these parameters respectively. We consider the following cosmological models for making comparisons between Fisher and MCMC – ΛCDM (dark energy equation of state is fixed to be -1), wCDM (with a single parameter, w_0 dark energy equation of state). We also consider the dark energy models with a redshift dependent equation of state for which we use the two common parametrizations – the Linder parametrization [58] $w(z) = w_0 + w_a \frac{z}{1+z}$ and the other simple form $w(z) = w_0 + w_z z$. In the rest of the paper, for the wCDM model we shall always use w_0 even to denote the single parameter dark energy equation of state, to ensure notational consistency.

We model a log-normal scatter [59] in the mass-observable relations parametrized as [11], $\eta(z) = \eta_0(1 + \eta_1 z)$. Due to the steep form of the cluster mass function, $\frac{dn}{dM}$ cluster counts are very sensitive to the amount of scatter. Thus when analysing the cluster count data the scatter should preferably not be fixed, but determined jointly with the cosmology. This is especially true for cluster data in optical/IR where the scatter is seen to be large

[8, 10]. However, this scatter need not be assumed to be completely unknown; it is expected that the mass follow-up observations would yield some constraints on the amount of scatter in the $M - \mathcal{O}$ relation. Thus, we place Gaussian priors with s.d. of 0.1 and 0.12 on the log-normal scatter of 0.45 and 0.58 in DES and RCS2 respectively. In our analysis of mock cluster data from the x-ray and SZE surveys, the relatively small scatter η is kept fixed.

For computing the $\frac{dN}{dz}$ likelihoods we first generate a mock dataset cluster distribution in redshift bins of width 0.1. We then use eq. 3.1 to compute the likelihoods for each point Θ in the Markov chain. As mentioned in section 2, just cluster counts cannot place tight constraints on cosmology by themselves. To break the cosmology-cluster physics degeneracy we consider a simple mass follow-up, M_f, of 100 clusters for all our surveys except the WFXT. For the WFXT we consider a larger follow-up of 1,000 clusters. Our mass catalogs are uniformly sampled over mass (from $2 - 8 \times 10^{14}$ M$_\odot$) and redshift ($0.3 \leq z \leq 0.9$) and consists of a table of redshift, mass, and observable mass-proxy. The likelihood of M_f is then added to the likelihoods from $\frac{dN}{dz}$.

6 Cluster surveys

6.1 Large Yield Surveys

We examine in detail, on a case by case basis, the model dependent cosmological parameter constraints from various surveys. We focus mainly on upcoming surveys, especially in x-ray and optical, which promise to provide survey yields of ten-to-hundreds of thousands of clusters of galaxies. These are the surveys which will have enough number of clusters so as to self-calibrate multiple cluster specific nuisance parameters to give *unbiased* cosmological parameter constraints. We consider two survey cases each for x-ray and optical based observations. In x-ray we consider the very large surveys, the upcoming eROSITA and the proposed WFXT. For optical surveys we consider the RCS2 survey whose observation runs have been recently completed and the DES survey that will yield clusters in the near future .

Table 1 summarizes the details like area coverage, expected cluster detections, redshift range, etc. that were assumed for each of the cluster surveys investigated in this paper.

Survey	wavelength	N_{cl} (in k)	area (deg.^{2})	flux/B_{gc}/N_{200} cut	redshift range	$N_{follow-up}$ clusters
eROSITA	x-ray	120	27,000	4×10^{-14} erg-cm^{-2}-s^{-1}	0.1 - 1.3	100
WFXT	x-ray	300	20,000	5×10^{-15} erg-cm^{-2}-s^{-1}	0.1 - 1.3	1,000
RCS2	optical	15	1,000	$B_{gc} = 300$	0.1 - 1.0	100
DES	optical	155	5,000	$N_{200} = 17$	0.1 - 1.3	100
Planck	SZE	2	32,000	300 mJy	0.1 - 1.3	100

Table 1. A summary of the six cluster surveys that we consider in this work along with the relevant parameters.

6.1.1 X-ray surveys

Clusters are detected in x-ray due to the thermal Bremsstrahlung emission from the hot ICM. X-ray telescopes like ROSAT, XMM-Newton and Chandra along with others have observed a large number of clusters at various redshifts. Some of these observations have already given interesting independent constraints on cosmological models [7, 9].

For a given survey the limiting mass $M_{\text{lim}}(z)$ in eq. 2.1 is found from the flux limit, f_{lim} of the survey. For x-ray survey, we adopt luminosity-mass relations from [33] given by the following expression[5] with fiducial parameter values: $A = -4.24$, $\alpha = 1.61$ and $\gamma = 0$ with a log-normal scatter of 0.246.

$$L_X = 10^A \left(\frac{M_{500}}{1 \times 10^{15}} \right)^\alpha E^{1.85}(z)(1+z)^\gamma \tag{6.1}$$

eROSITA : The extended Roentgen Survey with an Imaging Telescope Array (eROSITA) is the next big full sky x-ray survey, expected to be launched in the near future [60]. The primary scientific goal would be to study the nature of dark energy using about 100,000 x-ray galaxy clusters. We model the survey with a limiting flux of $f_{\text{lim}} = 4 \times 10^{-14}$ erg cm^{-2} s^{-1} in the [0.5-2.0 keV] band with a sky coverage of $\sim 27,000$ deg^2.

WFXT : The Wide Field X-ray Telescope (WFXT) is a proposed x-ray mission which is expected to be 2 orders of magnitude more sensitive than any previous x-ray mission [61]. It would survey a large fraction of the sky at a high angular resolution with a deep, wide and a medium survey. The proposed survey would be in 3 parts – a deep survey of 100 sq. deg, a medium survey of 3,000 sq. deg and a wide survey of 20,000 sq. deg. We examine the constraints from the wide survey with a flux limit of $f_{\text{lim}} = 5 \times 10^{-15}$ erg cm^{-2} s^{-1} in the [0.5-2.0 keV] band.

In both the x-ray surveys we impose an additional flux cut-off corresponding to a mass limit of $8 \times 10^{13} h^{-1}$ M$_\odot$. This is to prevent selection of very low mass halos like groups of galaxies, especially at lower redshifts.

6.1.2 Optical surveys

Clusters are detected in optical surveys by searching for over-densities of galaxies with known color properties. Compared to other wavelengths, cluster surveys in optical provide the advantage of a high S/N ratio and a wide field of view. Also, because of a greater depth, these surveys give the largest yield in terms of cluster detection per sq. deg. Optical surveys are also essential to provide photometric redshift information of clusters detected in other wavelengths and for the measurement of galaxy power spectrum.

RCS2 : The Red Sequence Cluster Survey 2 (RCS2) is the sequel to RCS1 which was a 78 sq. deg. survey which detected $\sim 1,000$ clusters. The RCS2 is a larger survey with 1,000 sq. deg of sky coverage carried out using the one-square-degree MegaCam on the CFHT. As with the RCS1, it uses a red-sequence of cluster early-type galaxies to identify galaxy clusters [62]. It is expected to detect 15,000 clusters in the redshift range 0.3 - 1.0. We model the RCS2 on the lines of RCS1 (see [10]) [36] with a uniform $B_{gcR} = 300$ cut-off as follows, with $A = 10.29$, $\alpha = 1.70$, $\gamma = 0.64$ and $\eta_0 = \sigma_{M_{200}|Bgc} = 0.58$.

$$M_{200} = 10^A B_{gcR}^\alpha (1+z)^\gamma \tag{6.2}$$

---

[5]An alternate way is to take γ directly as the exponent of $E(z)$ without the need for the $(1+z)^\gamma$ term; however we avoid this form as we find that it induces extra correlations of γ_X with the cosmological parameters - Ω_m, h and others.

DES :   The Dark Energy Survey (DES) is an upcoming optical imaging survey project that would map out 5,000 sq. deg. of area in 5 years using the Blanco 4-meter telescope at the Cerro Tololo Inter-American Observatory in Andes. One of the primary goal of this survey is to make precision measurements of dark energy using various cosmological probes like Supernovae, baryon acoustic oscillations, weak lensing and galaxy clusters [63]. We model this survey with a maxBCG selection function and a mass cut-off corresponding to $N_{200} = 17$ as used in ref. [64], with $A = 2.34$ and $\alpha = 0.757$ and $\gamma = 0$ and a log-normal scatter, $\eta_0 = \sigma_{M_{200}|N_{200}} = 0.45$ in the mass-richness relation. Larger datasets give the opportunity to play with greater number of nuisance parameters for more robust cosmological constraint forecasts. Thus, for the DES we also examine the effect of introducing an extra nuisance parameter η_1 (as described in section 5) that can capture the redshift dependent scatter.

$$\ln(N_{200}) = A + \alpha \ln \left(\frac{M_{200}}{1.09 \times 10^{14}} \right) + \gamma \ln(1 + z) \tag{6.3}$$

6.2   Smaller Yield SZE Surveys

For the sake of completeness and to explore survey complementarity, we also consider the ongoing space based Planck mission. The Planck survey has already detected clusters [65] through their SZE signal and have obtained initial mass calibration [66]. For the Planck survey, we use the integrated Compton y parameter – Y as the mass proxy and write the SZE flux-mass relation as in [37, 39],

$$Y d_{\mathrm{A}}^2 = f(\nu) 10^A M_{200}^{\alpha} E^{2/3}(z) (1 + z)^{\gamma} \tag{6.4}$$

Here, $f(\nu)$ is the well known frequency dependence of the SZ effect and d_{A} is the angular diameter distance in Mpc. The conversion from Y to SZ flux is done as in ref. [67]. The values for the SZE scaling parameters are: $A = -28.07$, $\alpha = 1.61$, $\gamma = 0$ which are taken from recent observations [68]. A log-normal scatter of 0.2 is assumed in the flux-mass relation.

Planck :   Planck covers the entire sky, with good sensitivity and at a high angular resolution over 9 frequency channels from 30–857 GHz. Planck was launched into orbit in 2009, and is expected to complete its observations by 2012. Among other goals such as measurements of intensity and polarization of the primordial and lensed CMB, Planck would also create a catalog of galaxy clusters through SZE. The first all sky SZ catalog from Plank was released recently which was from six months of survey. Projecting the initial cluster count to the full Planck mission time, one would expect a total yield of $1000 - 2000$ clusters over the entire sky. We take the flux limit to be 160 mJy (at 353 GHz) which gives $\lesssim 2,000$ clusters [69] in $\sim 32,000$ deg^2. The higher flux limit means that Planck would be able to detect only the most massive clusters.

7   Results

We now compare the forecasts for parameter constraints from each of the surveys considered in section 6, obtained from Fisher and MCMC analysis of mock cluster data.

7.1   X-ray surveys

eROSITA :   The eROSITA survey places strong constraints on the ΛCDM cosmology, with $\Delta\Omega_m = 0.028$ and $\Delta\sigma_8 = 0.016$. To compare, these constraints are slightly better than

model	ΛCDM		wCDM		w_aCDM	
parameter	MCMC	Fisher	MCMC	Fisher	MCMC	Fisher
Ω_m	0.028	0.026	0.035	0.028	**0.080**	**0.028**
w_0	—	—	0.191	0.145	**0.374**	**0.172**
w_a	—	—	—	—	**0.685**	**0.165**
σ_8	0.016	0.018	0.034	0.025	**0.059**	**0.026**
A	0.048	0.036	0.052	0.039	0.096	0.055
α	0.031	0.035	0.037	0.040	0.066	0.047
γ	0.100	0.089	0.179	0.177	0.204	0.185

Table 2. A comparison of the standard deviations of marginalized parameter constraints from analysis of the mock cluster data from the eROSITA x-ray survey using MCMC and Fisher methods. Numbers in bold highlight the significant discrepancy in constraint estimates between the two methods.
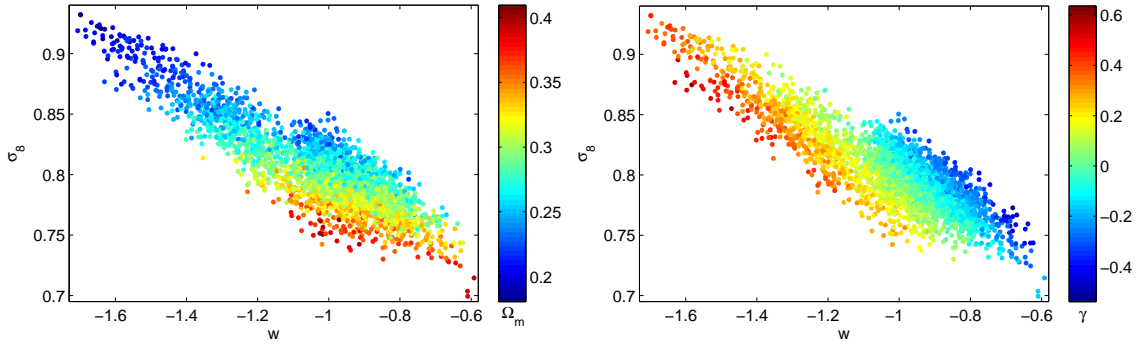

Figure 1. The degeneracy between the cosmological parameters Ω_m and σ_8 for the eROSITA survey is indicated by a few samples drawn from the posterior likelihood distribution of the wCDM model. The correlation with a third parameter – Ω_m (left) and γ (right) is indicated through the color or value of the third parameter.
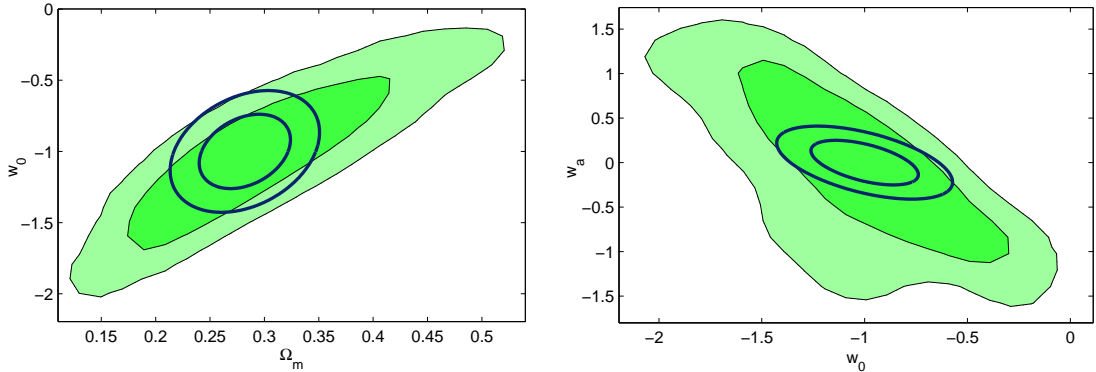

Figure 2. The dark(light) green regions indicate the projected 1-σ (2-σ) confidence regions obtained from the analysis of mock cluster data for the eROSITA survey using MCMC methods for a w_aCDM model. The blue ellipses are the corresponding constraints from a Fisher analysis.

those obtained from just the CMB (WMAP7 results) for which $\Delta\Omega_m = 0.029$ and $\Delta\sigma_8 = 0.030$. For the wCDM model, the constraints get relaxed to 0.035 and 0.034 on Ω_m and σ_8
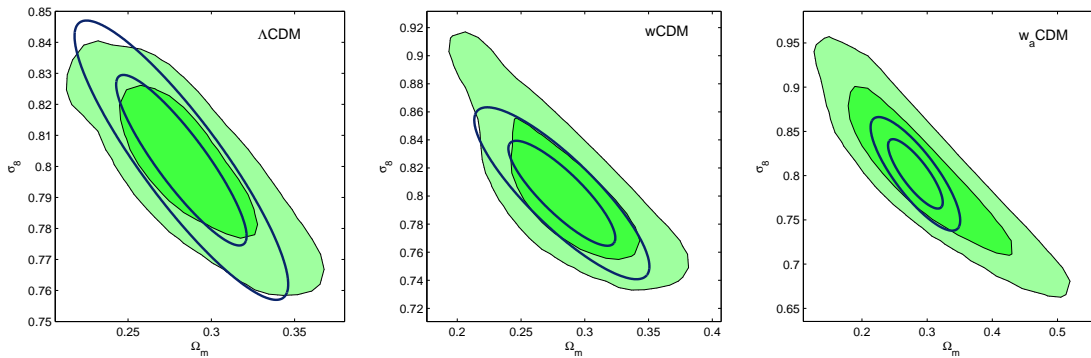
Figure 3. The dark (light) green regions indicate the projected 1-σ (2-σ) confidence regions obtained from the analysis of mock data for eROSITA using MCMC methods for the three cosmological models – ΛCDM, wCDM and w_aCDM in the $\Omega_m - \sigma_8$ plane. The blue ellipses are the corresponding constraints from a Fisher analysis.

respectively, with almost a twofold increase in $\Delta\sigma_8$. This occurs due to the long degeneracy between w_0 and σ_8, that causes wider constraints on σ_8 after marginalization over w; the marginalised constraint on w_0 is 0.191. This degeneracy in the $w_0 - \sigma_8$ plane is indicated through figure 1; the correlation of σ_8 and w_0 with other parameters like Ω_m and γ is also shown here. The expected degeneracy of $\Omega_m - \sigma_8$ and $w - \gamma$ is clearly visible through the color coding. We find here that the Fisher results are underestimated by a factor of ≈ 1.3 for all Ω_m, σ_8 and w_0. Addition of priors from the WMAP7 results can significantly shrink the constraints on the wCDM model to $\Delta\Omega_m = 0.013$, $\Delta w_0 = 0.042$ and $\Delta\sigma_8 = 0.011$. The large number of clusters expected to be detected from a full sky survey like eROSITA will also place constraints on the possible evolution of the dark energy equation of state. We examine the constraints from w_aCDM model in which dark energy equation of state varies smoothly from w_0 at $z = 0$ to approach $w_0 + w_a$ at a large redshift ($z \gg 1$). For the w_aCDM model, with a two parameter dark energy equation of state there would be weaker constraints on cosmology. We also see that there is no agreement between Fisher and MCMC estimates whatsoever – $\Delta\Omega_m = 0.080(0.028)$, $\Delta w_0 = 0.374(0.172)$, $\Delta w_a = 0.685(0.165)$ and $\Delta\sigma_8 = 0.059(0.026)$ for the MCMC(Fisher) results respectively. There is also a noticeable difference in the degeneracy direction in the $w_0 - w_a$ plane, see figure 2. It is alarming to note how *the Fisher estimates mislead us into false estimates of tight constraints on the parameter w_a, where the two forecasts can differ by a factor up to 4 or more.* We find fairly good agreement between the Fisher and MCMC results only for the ΛCDM model. The results between the two methods are seen to progressively diverge for the wCDM and w_aCDM models, see figure 3. The detailed results for eROSITA are also listed in table 2.

WFXT : The WFXT is seen to place much stronger constraints on cosmology – i) By virtue of a large number of cluster detections due to its high flux sensitivity; ii) a 10 times larger mass follow-up of 1,000 clusters that we have considered to break the cosmology - cluster physics degeneracy. On the ΛCDM model we get very tight constraints with $\Delta\Omega_m = 0.012$ and $\Delta\sigma_8 = 0.027$. These constraints are about a factor of 2 or so better as compared to eRSOITA. Here, the Fisher case slightly overestimates for the constraints as compared to MCMC. For the wCDM model we find $\Delta\Omega_m = 0.029$, $\Delta\sigma_8 = 0.031$ and $\Delta w_0 = 0.133$; while the Fisher and MCMC give very similar results. However they turn out to be very

model	ΛCDM		wCDM		w_aCDM		w_zCDM	
parameter	MCMC	Fisher	MCMC	Fisher	MCMC	Fisher	MCMC	Fisher
Ω_m	0.012	0.017	0.029	0.026	**0.047**	**0.026**	0.034	0.043
w_0	—	—	0.133	0.120	**0.225**	**0.120**	0.181	0.206
w_a	—	—	—	—	**0.343**	**0.093**	0.172	0.125
σ_8	0.014	0.017	0.031	0.029	**0.041**	**0.029**	0.037	0.043
A	0.024	0.028	0.039	0.035	0.052	0.037	0.045	0.053
α	0.037	0.037	0.038	0.038	0.039	0.038	0.039	0.039
γ	0.048	0.057	0.136	0.137	0.140	0.153	0.145	0.154

Table 3. A comparison of standard deviations of marginalized parameter constraints obtained using MCMC and Fisher analysis of the mock cluster data from the WFXT x-ray survey. Numbers in bold highlight the significant discrepancy in constraint estimates.
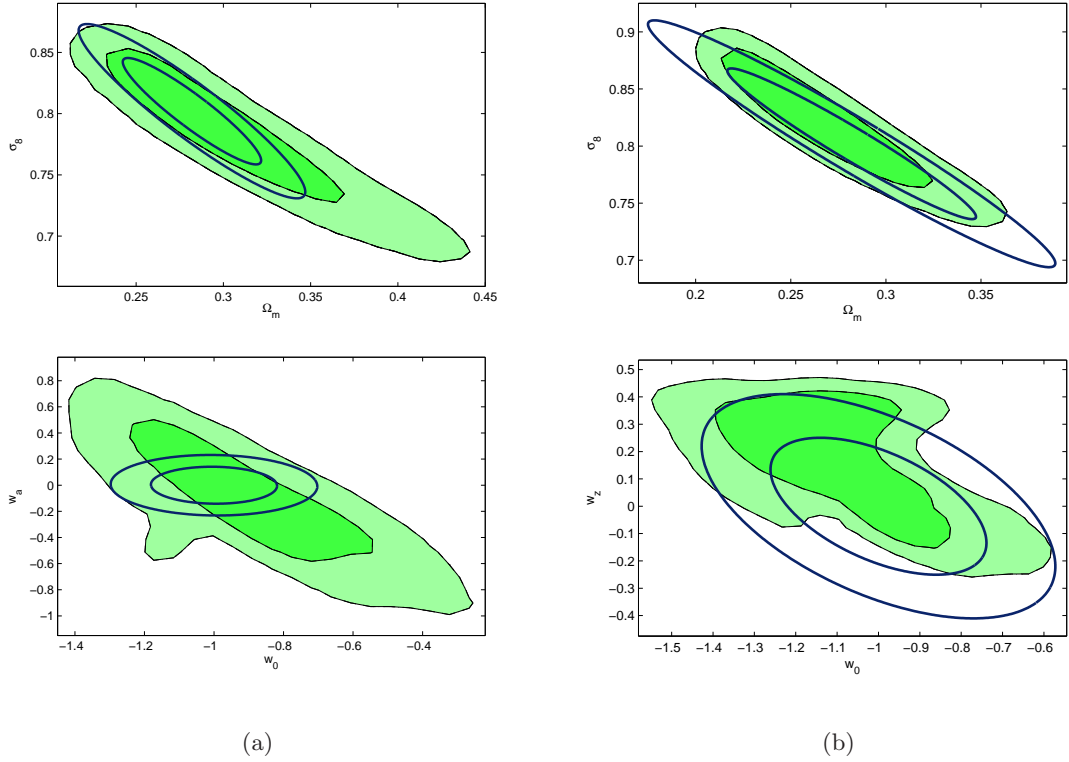

(a)   (b)

Figure 4. A comparison of parameter constraints from the WFXT survey for the two different parametrizations of the dark energy equation of state – $w(z) = w_0 + \frac{w_a z}{1+z}$ (left) and $w(z) = w_0 + w_z z$ (right). The dark(light) green regions indicate the projected 1-σ (2-σ) confidence regions obtained from the analysis of mock cluster data using MCMC methods. The blue ellipses indicate the corresponding constraints from a Fisher analysis.

different for the w_aCDM model, as seen in the eROSITA case before. Here, we find $\Delta\Omega_m$ = 0.047(0.026), Δw_0 = 0.225(0.120), Δw_a = 0.343(0.093) and $\Delta\sigma_8$ = 0.041(0.029) for the MCMC(Fisher) forecasts respectively. The only difference is that the estimates for σ_8 are not as divergent. We find again that introducing the parameter w_a for the redshift evolution

of dark energy equation of state makes the Fisher constraints seem much tighter than they actually are; also we find that the *degeneracy direction predicted by the Fisher method is incorrect*, see figure 4. Figure 5 shows the 2-dimensional likelihood contours for pairs of some of the relevant parameters, along with the 1-dimensional marginalized pdf's along the diagonal. In the figure, for the 2-dimensional likelihoods, the red and blue contours indicate the 1-σ and 2-σ confidence regions for MCMC and Fisher forecasts respectively.

For the WFXT survey we also examine the effect of a different parametrization for the dark energy equation of state: $w(z) = w_0 + w_z z$. We do so since both models, w_a and w_z have been considered in the literature, and here we point out the degeneracies resulting from the two parametrizations. The w_z parametrization goes linearly with z, clearly not realistic at very large z ($z \gg 1$). However as long as we are dealing with low redshift data, and we are interested in constraining deviations from a $w = -1$ equation of state, both models can be used.

The degeneracy direction of Fisher ellipses in the w_0-w_a plane is different from that in the w_0-w_z plane. Although the Fisher ellipses do not provide a very accurate description for the dark energy constraints for both the parametrizations (see Fig. 4), we find that the w_z parametrization compares much favorably as far as the discrepancy between Fisher/MCMC forecasts from cluster surveys are considered. For e.g. we find that $\Delta\Omega_m = 0.034(0.043)$, Δw_0 = 0.225(0.140), $\Delta w_z = 0.343(0.075)$ and $\Delta\sigma_8 = 0.041(0.031)$ for the MCMC(Fisher) forecasts respectively. For the w_zCDM model not only is the discrepancy significantly smaller, but also the degeneracy directions are more or less correctly predicted by the Fisher method, see figure 4. However, the marginalized 2-dimensional likelihoods in the dark energy plane are seen to be highly skewed. Table 3 summarises the results for the WXFT survey. Finally, we emphasize again that the computation of the Fisher matrix (and therefore the constraints resulting from it) are dependent on the derivatives computed just at the Fiducuial point in the parameter space. For a non-Gaussian distribution of the likelihood, the Fisher matrix may or may not provide a correct estimate of the actual (MCMC) constraints. This may be the reason for the different discrepancies in the Fisher/MCMC comparison for the two parametrizations, as seen here. See also the discussion under DES in section 7.2.

| model | ΛCDM | | wCDM | |
parameter	MCMC	Fisher	MCMC	Fisher
Ω_m	0.040	0.040	0.046	0.040
w_0	—	—	0.144	0.137
σ_8	0.141	0.140	0.153	0.151
A	0.299	0.300	0.302	0.301
α	0.117	0.117	0.117	0.117
γ	0.210	0.220	0.227	0.221
η_0	0.099	0.114	0.111	0.120

Table 4. A comparison of standard deviations of marginalized parameter constraints obtained using MCMC and Fisher analysis of the mock cluster data from the RCS2 optical survey for the wCDM model.

7.2 Optical surveys

For the optical surveys there is a large scatter between the mass-observable relation. To account for this we introduce two additional parameters to characterize this scatter and

Figure 5. Projected 1-σ and 2-σ confidence regions obtained from the analysis of mock cluster data for the WFXT survey for the w_aCDM model. The contours in red (blue) show results obtained using MCMC (Fisher) analysis. The diagonal plots in black show the corresponding marginalized pdf's for each parameter.

its evolution with redshift. As we shall see, the presence of an unknown scatter creates extra degeneracies, and weakens the cosmological constraints. In order to investigate this degeneracy, we consider here only the ΛCDM and wCDM model for simplicity.

RCS2 : For the RCS2 we see a reasonably good agreement between the Fisher and MCMC forecasts. For the ΛCDM model we find fairly wide constraints of 0.040 and 0.141 on Ω_m and σ_8 respectively, which get marginally relaxed to 0.046 and 0.153 in the wCDM model with Δw_0 being 0.144. It is interesting to note that the constraints on σ_8 are much wider than those from smaller yield SZ surveys like Planck. This is caused due to presence of large scatter in the mass - proxy relation which causes a large degeneracy between σ_8 and the scatter η_0; marginalizing over this degeneracy washes out the constraints on σ_8. Figure 6 shows the 2-dimensional likelihood contours for each pair of some of the relevant parameters,

Figure 6. Projected 1-σ and 2-σ confidence regions obtained from the analysis of mock data for the RCS2 survey for the wCDM model. The contours in red (blue) show results obtained using MCMC (Fisher) analysis. The diagonal plots in black show the corresponding marginalized pdf's for each parameter.

along with the 1-dimensional marginalized pdf's along the diagonal. For the 2-dimensional likelihoods, the red and blue contours indicate the 1-σ and 2-σ confidence regions for MCMC and Fisher forecasts respectively. The RCS2 results are listed in table 4.

DES : For a larger optical survey like DES we also examine the effect of introducing η_1 as the free parameter for the redshift dependent scatter. As expected, with an extra parameter the errors on Ω_m and σ_8 get larger from 0.025 to 0.030 and 0.070 to 0.113 respectively, for the ΛCDM model. For the wCDM case the increase is smaller, 0.033 to 0.037 for Ω_m and 0.094 to 0.116 for σ_8. The parameter w_0 is seen to be degenerate with η_1, see figure 7; and this causes a significant degradation in the constraints on w, from 0.071 to 0.149, with and without η_1, respectively. The Fisher results for w_0 are always underestimated; for σ_8 they are either overestimated (up to 50%) or remain same; while for Ω_m these are seen to be more or less consistent. This may be understood by observing the shape of the peaks in the marginalized single parameter pdf's – for a sharply peaked curve such as in w_0, the Fisher results would

model parameter	ΛCDM MCMC	ΛCDM Fisher	ΛCDM MCMC	ΛCDM Fisher	wCDM MCMC	wCDM Fisher	wCDM MCMC	wCDM Fisher
Ω_m	0.025	0.028	0.030	0.032	0.033	0.035	0.037	0.036
w_0	—	—	—	—	**0.071**	**0.051**	0.149	0.130
σ_8	**0.070**	**0.101**	0.113	0.118	0.094	0.109	0.116	0.123
A	0.082	0.080	0.095	0.099	0.102	0.100	0.104	0.100
α	0.055	0.050	0.055	0.052	0.055	0.052	0.056	0.052
γ	0.099	0.124	0.131	0.146	0.140	0.143	0.148	0.146
η_0	0.057	0.084	0.091	0.100	0.075	0.097	0.089	0.100
η_1	—	—	0.041	0.026	—	—	0.072	0.066

Table 5. Comparison of $1 - \sigma$ marginalized parameter constraints from MCMC and Fisher analysis of the mock cluster data from the DES optical survey. Numbers in bold highlight the significant discrepancy in forecasts.
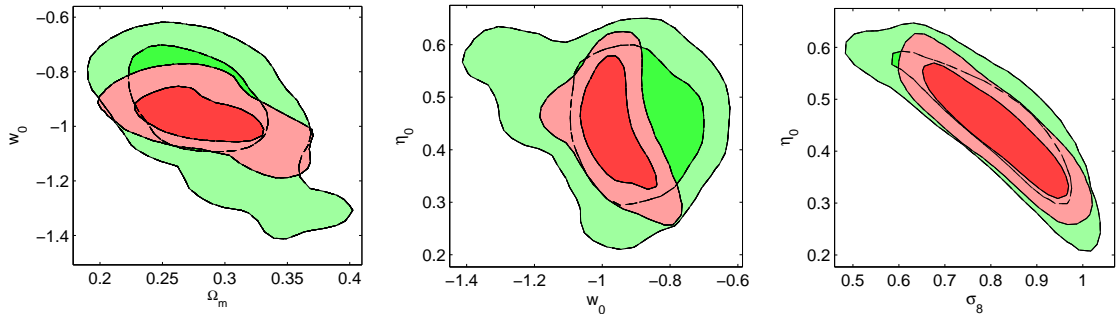
Figure 7. The dark (light) regions indicate the projected 1-σ (2-σ) confidence regions obtained from the analysis of mock data for the DES survey using MCMC methods for the wCDM model in various parameter planes. The green(red) region indicates constraints with (without) redshift dependent scatter.

be underestimated, while for a flatly peaked curve, they would be overestimated, see figure 8. The Fisher constraints on η_0 are overestimated, but underestimated for η_1. The marginalized pdf's in this figure show asymmetry as well as the presence of extended non-Gaussian tails. The degeneracy of σ_8 with η_1 is seen to be much weaker than with η_0, however the degeneracy shape is highly non-Gaussian and skewed, see figure 9 causing a factor of 1.5–2 discrepancy in $\Delta\eta_1$, in Fisher and MCMC forecasts. The above mentioned results are also listed in detail in Table 5.

8   Improved constraints from a combination of clusters with CMB+BAO+SNe observations

We also we consider the effect of adding the results from CMB+BAO+SNe to the $\frac{dN}{dz}$ constraints from cluster surveys. We add Gaussian priors of $\Delta\Omega_m$=0.015, Δw=0.053 and $\Delta\sigma_8$=0.038 obtained from the WMAP7 analysis. We find a significant improvement in the cosmological constraints on adding these priors. For the eROSITA and WFXT surveys $\Delta\sigma_8$ value can be improved by a factor of 2.5–3 over the WMAP7 constraints. The constraints on the dark energy equation of state w_0 shrink by half for WFXT and DES surveys, in compar-
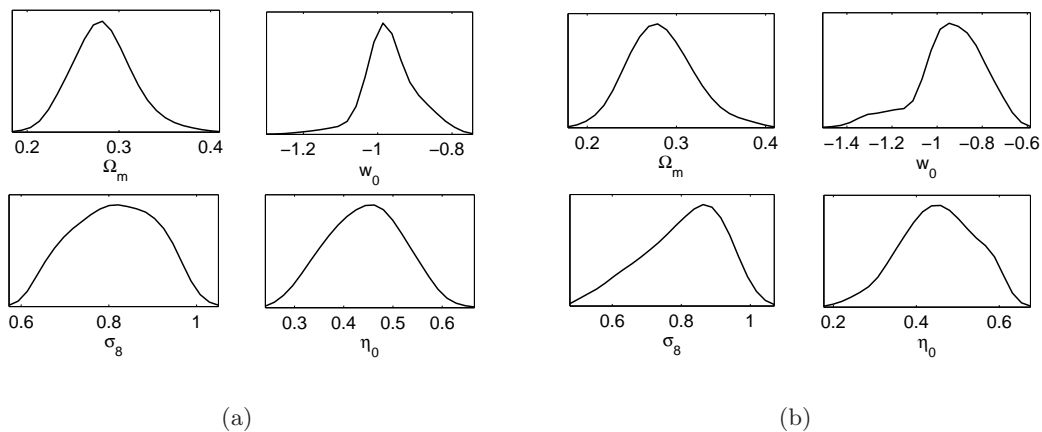
Figure 8. Marginalized pdf's for a few parameters obtained from the analysis of mock cluster data of the DES using MCMC methods. **(a)** wCDM model with a one parameter scatter; **(b)** wCDM model with a two parameter scatter
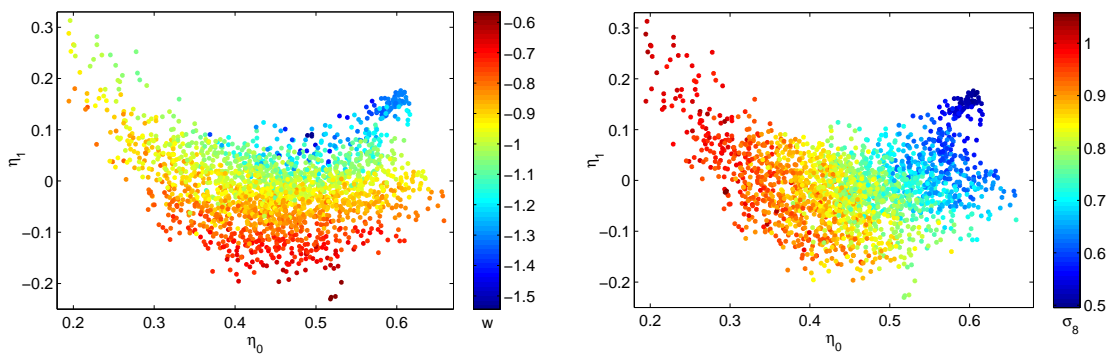

Figure 9. An analysis of mock data for DES using MCMC methods shows the degeneracy between the scatter parameters η_0 and η_1 as well as their correlation with a third parameter – w_0 (left) and σ_8 (right). The dots indicate a few samples drawn from the posterior distribution, colored according to the value of the third parameter, Ω_m or γ.

ison to the results from just CMB+BAO+SNe. These results are also summarized in table 6.

survey parameter	CMB +SNe +BAO	+eROSITA		+WFXT		+DES	
		$\frac{dN}{dz}$+FUP	$\frac{dN}{dz}$+FUP +WMAP7	$\frac{dN}{dz}$+FUP	$\frac{dN}{dz}$+FUP +WMAP7	$\frac{dN}{dz}$+FUP	$\frac{dN}{dz}$+FUP +WMAP7
Ω_m	0.015	0.035	0.013	0.029	0.011	0.033	0.013
w_0	0.053	0.191	0.042	0.133	0.025	0.071	0.026
σ_8	0.038	0.034	0.011	0.031	0.013	0.094	0.035

Table 6. Improvement in the $1-\sigma$ marginalized parameter constraints (from MCMC analysis) from adding clusters to WMAP7 analysis of CMB+BAO+SNe datasets. All constraints are for the wCDM model.

parameter	fiducial	mean	1-σ		2-σ		3-σ	
			MCMC	Fisher	MCMC	Fisher	MCMC	Fisher
Ω_m	0.282	0.346	0.408	0.339	0.502	0.396	0.608	0.453
			0.284	0.225	0.256	0.168	0.229	0.111
w_0	-1	-0.897	-0.700	0.811	-0.492	-0.622	-0.294	-0.433
			-1.092	-1.189	-1.293	-1.378	-1.557	-1.567
σ_8	0.802	0.760	0.802	0.846	0.823	0.890	0.843	0.934
			0.718	0.758	0.667	0.714	0.618	0.670

Table 7. A comparison of 1-σ, 2-σ and 3-σ marginalized limits on the cosmological parameters obtained using MCMC and Fisher analysis of the mock cluster data from the Planck SZ survey for the wCDM model.
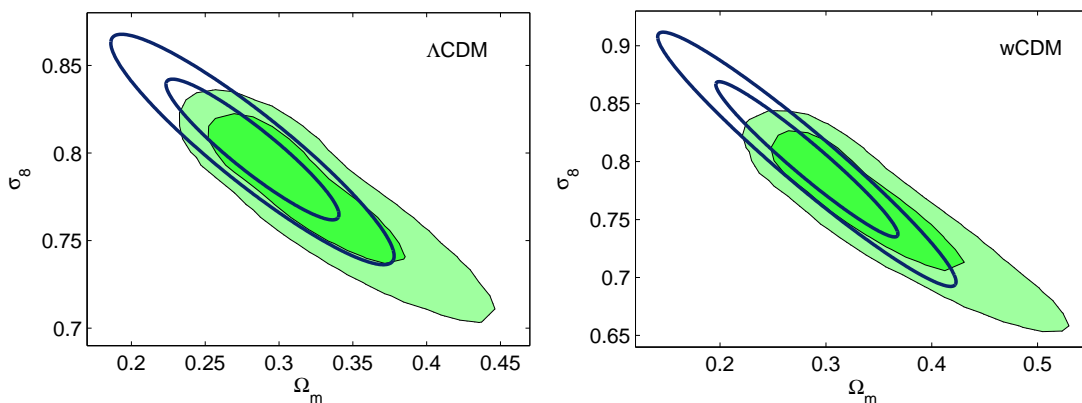
Figure 10. The dark (light) green regions indicate the projected 1-σ (2-σ) confidence regions obtained from the analysis of mock cluster data for the Planck SZE survey using MCMC methods for the ΛCDM and wCDM cosmological models in the $\Omega_m - \sigma_8$ plane. The blue ellipses are the corresponding constraints from a Fisher analysis.

9 Improved constraints from combining multiple cluster surveys

In the previous section we showed how having clusters together with CMB, SNe & BAO leads to percent level constraints on key cosmology parameters. In this section we point out a way to have tight cosmological constraints, than possible in section 7, but using cluster datasets only. The surveys that we have discussed till now along with SZE survey (like Planck) will have datasets which will include clusters common in overlapping regions of the sky. A subset of these clusters can be used innovatively to constrain cosmological constraints from using number counts only from any particular dataset [70]. However, we show below, that due to the different degeneracy directions among cosmological and cluster parameters in each of these surveys, a joint analysis of $\frac{dN}{dz}$ from any two surveys will greatly increase the constraining power of clusters as cosmological probes. We show this for three sets: (i) optical (DES) + x-ray (WFXT), (ii) optical (DES) + SZE (Planck), and (iii) x-ray (eROSITA) + SZE (Planck). Before proceeding further, we look at the prospects from Planck clusters.

Planck Cluster Cosmology : With an estimated yield of ~ 2000 clusters, Planck gives the following constraints - $\Delta\Omega_m = 0.044$ and 0.064, while $\Delta\sigma_8 = 0.028$ and 0.041 for the

ΛCDM and wCDM models respectively with $\Delta w_0 = 0.199$ in the latter case. We find that fairly good agreement exists between the Fisher and MCMC estimates for Planck. However, it is interesting to take note of the fact that the 2-dimensional projected likelihoods obtained using Fisher in the $\Omega_m - \sigma_8$ plane show a substantial offset as compared to the MCMC likelihoods, see figure 10. This is seen in both the ΛCDM as well as wCDM models, and arises due to the asymmetric distribution of the posterior likelihoods, especially so in the plane of these two parameters. Table 7 shows a comparison of 1-σ, 2-σ and 3-σ limits on the marginalized parameter constraints obtained using MCMC and Fisher analysis Planck SZ survey for the wCDM model.

It is worth emphasizing again at this point that the surveys considered in this work give rise to very different degeneracy directions in the parameter space. If these degeneracies could be broken through a joint analysis of various cluster datasets, we would obtain significantly tighter cosmological constraints *using clusters alone*. We now give a few examples to show how, and also to indicate the extent to which this could be possible. A detailed report of such improvements in the constraints from such a joint analysis will be reported in a future work.
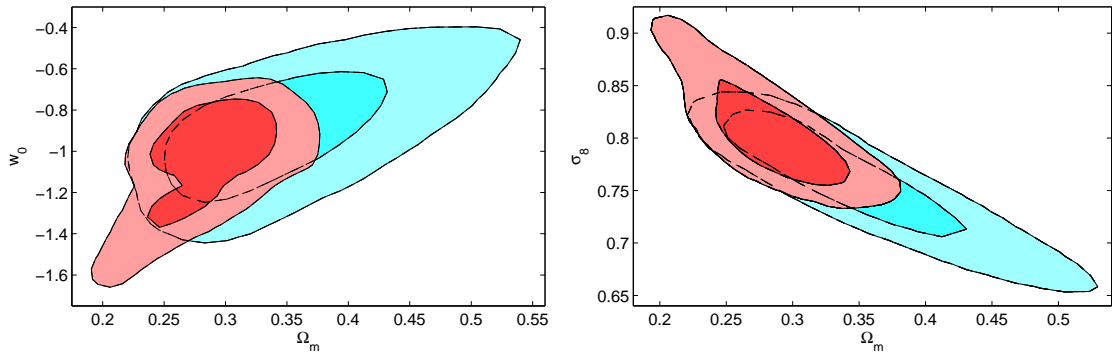

Figure 11. The dark(light) regions indicate the projected 1-σ (2-σ) confidence regions obtained from the analysis of mock cluster data using MCMC methods for a w_aCDM model. The cyan (red) regions correspond to constraints from the Planck (eROSITA) survey.
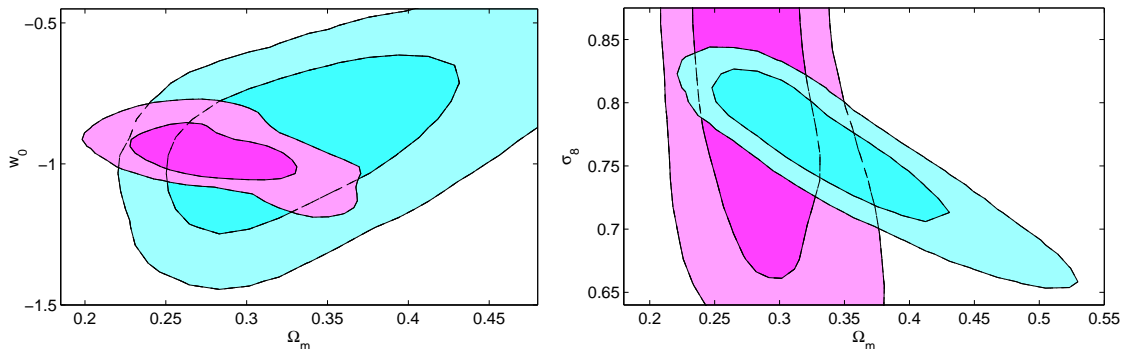

Figure 12. The dark(light) regions indicate the projected 1-σ (2-σ) confidence regions obtained from the analysis of mock cluster data using MCMC methods for a w_aCDM model. The cyan (magenta) regions correspond to constraints from the Planck (DES) survey.
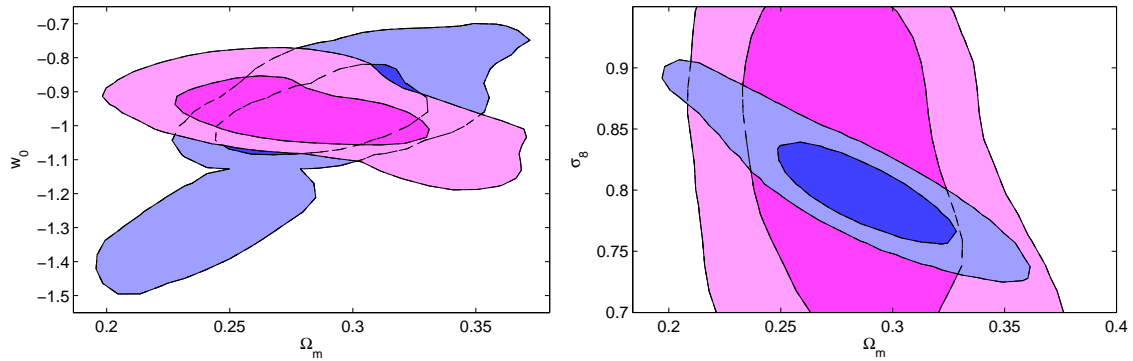
Figure 13. The dark(light) regions indicate the projected 1-σ (2-σ) confidence regions obtained from the analysis of mock cluster data using MCMC methods for a w_aCDM model. The blue (magenta) regions correspond to constraints from the WFXT (DES) survey.

The degeneracy directions of the parameters in the $\Omega_m - w$ and $\Omega_m - \sigma_8$ planes for Planck and eROSITA surveys are parallel, however the posterior likelihood distributions are seen to be asymmetric with the tails being exactly in opposite directions, see figure 11. Each of these surveys individually shows large degeneracy, however the overlap region is much smaller.

Next, the optical survey DES and the SZ survey Planck seem to be exactly complimentary. As seen before, the DES is expected to give only weak constraints on the parameter σ_8 due to a large scatter, but at the same time gives relatively strong constraints on both Ω_m as well as the dark energy parameter w. It turns out that Planck would have an almost orthogonal degeneracy with DES, see figure 12, which could kill the degeneracy to give much better joint constraints.

Finally, we consider the benefits of a joint WFXT and DES analysis. Again, we find that degeneracy directions are almost orthogonal between the two surveys, see figure 13. The WFXT survey gives somewhat extended posterior distribution for the dark energy parameter w, beyond the 1-σ region. Using these surveys in combination would produce better 2-σ constraints on the parameter w_0.

10 Discussions

In our assumptions made to model future galaxy cluster surveys, we followed a simple recipe for self-calibration using a mass follow-up. The parameter degeneracies and also the constraints certainly depend on the self-calibration methods used to improve upon the cosmological constraints. In another work [70], we examine the dependence of degeneracies for SZ surveys like ACT/SPT in the Ω_m-w_0 and w_0-w_a plane with various choices of calibration (see Fig. 2) such as distance measurements from supernovae type Ia, distances from joint X-ray+SZ observations of clusters and mass follow-up's with different errors on mass measurements. In [41] we propose a different method of self-calibration, using a combination of deep and wide SZ surveys to obtain significantly tighter cosmological constraints. Here we use the fact that the deep and wide surveys have slightly different degeneracies, and we show that it is possible to get competitive constraints without the need for a costly mass follow-up.

While estimating the parameter constraints $\Delta\theta_i$ from the Fisher technique it is important to check for numerical stability. Numerical instability may arise from - the numerical

derivatives not having converged; or the Fisher matrix $\mathcal{F}$ being unstable under inversion. The later occurs especially when the eigenvalues of $\mathcal{F}$ span a large range of values (This may be avoided to some extent in practice by re-parametrizing θ_i so that $\Delta\theta_i$ for different parameters are not too different in magnitude.). For a near-singular $\mathcal{F}$, methods of singular value decomposition should be used for the inversion. It is preferable to compute the numerical derivative symmetrically about a given value, i.e. for small h, $\frac{\partial f}{\partial \theta} \simeq \frac{f(\theta+h/2)-f(\theta+h/2)}{h}$. At the same time h should not be chosen to be too small, such that the changes in the numerator Δf become comparable to the accuracy in computing the function f. This may give rise to artificially inflated values in the parameter constraints.

In our analysis we find that the degeneracies of the parameter w_a with other parameters often shows different directions in Fisher and MCMC analysis, for e.g. see Fig. 2 and 4. In our view the discrepancy may be explained in the following manner: The direction of the Fisher ellipse is related to the ratio of the marginalized errors on the parameters. Since the discrepancy between Fisher and MCMC (marginalized) constraints on the parameter w_a are very different, it should not be surprising that the degeneracy directions are also different. For example, in the case of eROSITA Fisher, the errors from Fisher and MCMC differ by factors of 2.2 and 4.2 for the parameters w_0 and w_a respectively. This causes the conflict when comparing with the degeneracy as seen from MCMC.

Our MCMC analysis is based on the data collected from 6-7 independently run chains. We use the convergence statistic R-1 = variance(chain mean)/mean(chain variance) [71], computed for the second half of the set of seven independently evolved MCMC chains. For the Metropolis sampling we rescale the proposal matrix further in order to get an acceptance rate of 20-25 %. For a higher acceptance rate the chain keeps jumping around and is more likely to land in regions of low probability, while for a low acceptance rate the chains mix very slowly; in both the cases convergence is slow. For good convergence, R-1 $\lesssim$ 0.03. We find that our analysis requires between 300,000 - 10,000,000 points sampled during the MCMC exploration of parameter space for the chains to be well converged. Sometimes, the convergence of chains can be improved by annealing the chains, i.e. instead of sampling from the distribution $\mathcal{L}$, one uses $\mathcal{L}^{1/T}$. Here the parameter T plays the role similar to temperature, in the sense of broadening the distribution for $T > 1$ and vice-versa. This is especially helpful for exploring the tails of distributions, discovering other local minima, and for getting more robust high-confidence error bars. The parameter likelihoods must be adjusted accordingly (or cooled) during the analysis to get back the correct target pdf. We find that a mild heating to $T = 1.5 - 2$ can boost the convergence significantly. However, heating to higher T may cause the chains to approach the limits set by allowed values of parameters, for e.g. $0 < \Omega_m < 1$, which may cause an artificial cut-off in the distribution on cooling back the chains. In addition, it is also useful to thin the chains by factors of 10-100 in order to reduce the auto-correlations in parameter values of the chains.

Finally, in our MCMC analysis we did not consider the effect of scatter in both $\frac{dN}{dz}$ and the follow-up data, as it is not possible to introduce scatter in the the simple Fisher analysis. However, real data is always expected to contain some scatter and the best fit parameter values are seen to move away from the fiducial values as scatter brings in some extra freedom in the fitting of data.

11 Conclusion

We have made a detailed study of the cosmological forecasts for upcoming cluster surveys showing that these cluster surveys would be able to place strong constraints on cosmology even with a simple mass follow-up of about a 100 clusters. In comparing forecasts from Fisher estimates to those from MCMC analysis, we strongly advocate the use of full MCMC likelihood in forecasting parameter constraints, especially on those related to dark energy. We find that the Fisher estimates are reliable only for predicting the constraints on minimal cosmological models. In many cases the results from the two methods are seen to diverge by factors of $1.5 - 2$. We show that the Fisher estimates can completely fail to correctly forecast the constraints as well as the degeneracy directions for cosmological models with a dark energy whose equation of state $w(z)$ evolves with redshift. For some surveys we find an asymmetric posterior distribution for the cosmological parameters. For example, the MCMC analysis of the mock cluster data expected from the Planck survey shows a large offset compared to the Fisher ellipses. In the case of optical cluster surveys we see that, a large unknown scatter in the mass-proxy relation degrades the constraints on σ_8. $\Delta\sigma_8$ further degrades if the scatter is also allowed to be redshift dependent. This occurs due to the large degeneracy between σ_8 and scatter. The redshift dependent scatter is seen to weaken the constraints on w_0 by a factor of 2 or so. The presence of scatter, especially a redshift dependent one, is seen to give a skewed distribution for the posterior likelihoods as well as non-Gaussian tails in the marginalized single parameter pdf's. Breaking the degeneracies between scatter and σ_8 would be particularly important for future cluster surveys in optical (where the scatter is estimated to be large) in order to get competitive constraints with other probes. In the end we show through some specific examples that even though the parameter constraints obtained from the analysis of some of the cluster survey data are relatively weak to begin with, a joint analysis of the datasets from two (or more) cluster datasets can be effectively utilised to kill the near orthogonal parameter degeneracies occurring from these surveys to obtain significantly tighter cosmological constraints.

In general we observe that the discrepancy between the two forecasting methods increases with the introduction of additional parameters like w_a and η_1 for the redshift dependence of dark energy and scatter respectively. Most of the forecasts on cluster cosmology presented in the literature have used the Fisher matrix analysis, mainly due to its simplicity. In our view, such forecasts, even though being agreeable with MCMC in many cases need to be accepted with a fair bit of caution. This is especially important when additional parameters relevant to cosmology or cluster physics are introduced in the analysis of data – parameters for which we are *ab initio* unaware about the Gaussianity of the likelihood distribution. Moreover, an accurate knowledge of parameter degeneracies, available from MCMC, is very important for joint analysis of datasets from two or more surveys to obtain tighter constraints on cosmological and cluster physics.

12 Acknowledgements

for providing several useful comments and suggestions that have helped in improving the presentation of this work.

References

[1] G. Holder, Z. Haiman, and J. J. Mohr, *Constraints on Ω_m, Ω_Λ, and σ_8 from Galaxy Cluster Redshift Distributions*, *ApJ Lett.* **560** (Oct., 2001) L111–L114, [astro-ph/0105396].

[2] W. Hu, *Self-consistency and calibration of cluster number count surveys for dark energy*, *Phys. Rev. D* **67** (Apr., 2003) 081304, [astro-ph/0301416].

[3] E. S. Levine, A. E. Schulz, and M. White, *Future Galaxy Cluster Surveys: The Effect of Theory Uncertainty on Constraining Cosmological Parameters*, *ApJ* **577** (Oct., 2002) 569–578, [astro-ph/0204273].

[4] L. Wang and P. J. Steinhardt, *Cluster Abundance Constraints for Cosmological Models with a Time-varying, Spatially Inhomogeneous Energy Component with Negative Pressure*, *ApJ* **508** (Dec., 1998) 483–490, [astro-ph/9804015].

[5] J. Weller, R. A. Battye, and R. Kneissl, *Constraining Dark Energy with Sunyaev-Zel'dovich Cluster Surveys*, *Phys. Rev. Lett.* **88** (June, 2002) 231301, [astro-ph/0110353].

[6] S. M. Molnar, Z. Haiman, M. Birkinshaw, and R. F. Mushotzky, *Constraints on the Energy Content of the Universe from a Combination of Galaxy Cluster Observables*, *ApJ* **601** (Jan., 2004) 22–27, [astro-ph/0309807].

[7] S. Borgani, P. Rosati, P. Tozzi, and C. Norman, *Cosmological Constraints from the ROSAT Deep Cluster Survey*, *ApJ* **517** (May, 1999) 40–53, [astro-ph/9901017].

[8] E. Rozo et al., *Cosmological Constraints from the Sloan Digital Sky Survey maxBCG Cluster Catalog*, *ApJ* **708** (Jan., 2010) 645–660, [arXiv:0902.3702].

[9] A. Vikhlinin et al., *Chandra Cluster Cosmology Project III: Cosmological Parameter Constraints*, *ApJ* **692** (Feb., 2009) 1060–1074, [arXiv:0812.2720].

[10] M. D. Gladders et al., *Cosmological Constraints from the Red-Sequence Cluster Survey*, *ApJ* **655** (Jan., 2007) 128–134, [astro-ph/0603588].

[11] B. Sartoris et al., *The potential of X-ray cluster surveys to constrain primordial non-Gaussianity*, *MNRAS* **407** (Oct., 2010) 2339–2354, [arXiv:1003.0841].

[12] C. Cunha, D. Huterer, and O. Doré, *Primordial non-Gaussianity from the covariance of galaxy cluster counts*, *Phys. Rev. D* **82** (July, 2010) 023004, [arXiv:1003.2416].

[13] C. Fedeli, L. Moscardini, and S. Matarrese, *The clustering of galaxy clusters in cosmological models with non-Gaussian initial conditions: predictions for future surveys*, *MNRAS* **397** (Aug., 2009) 1125–1137, [arXiv:0904.3248].

[14] M. Roncarelli et al., *Imprints of primordial non-Gaussianities in X-ray and SZ signals from galaxy clusters*, *MNRAS* **402** (Feb., 2010) 923–933, [arXiv:0909.4714].

[15] D. Rapetti, S. W. Allen, A. Mantz, and H. Ebeling, *The observed growth of massive galaxy clusters - III. Testing general relativity on cosmological scales*, *MNRAS* **406** (Aug., 2010) 1796–1804, [arXiv:0911.1787].

[16] D. Rapetti, S. W. Allen, A. Mantz, and H. Ebeling, *Constraints on modified gravity from the observed X-ray luminosity function of galaxy clusters*, *MNRAS* **400** (Dec., 2009) 699–704, [arXiv:0812.2259].

[17] A. Mantz, S. W. Allen, and D. Rapetti, *The observed growth of massive galaxy clusters - IV. Robust constraints on neutrino properties*, *MNRAS* **406** (Aug., 2010) 1805–1814, [arXiv:0911.1788].

[18] N. A. Bahcall, J. P. Ostriker, S. Perlmutter, and P. J. Steinhardt, *The Cosmic Triangle: Revealing the State of the Universe*, *Science* **284** (May, 1999) 1481, [`astro-ph/9906463`].

[19] D. Huterer and M. S. Turner, *Probing dark energy: Methods and strategies*, *Phys. Rev. D* **64** (Dec., 2001) 123527, [`astro-ph/0012510`].

[20] R. A. Battye and J. Weller, *Optimizing the yield of Sunyaev-Zel'dovich cluster surveys*, *MNRAS* **362** (Sept., 2005) 171–183, [`astro-ph/0410392`].

[21] S. Wang, J. Khoury, Z. Haiman, and M. May, *Constraining the evolution of dark energy with a combination of galaxy cluster observables*, *Phys. Rev. D* **70** (Dec., 2004) 123008, [`astro-ph/0406331`].

[22] A. Melchiorri, P. Bode, N. A. Bahcall, and J. Silk, *Cosmological Constraints from a Combined Analysis of the Cluster Mass Function and Microwave Background Anisotropies*, *ApJ Lett.* **586** (Mar., 2003) L1–L3, [`astro-ph/0212276`].

[23] A. Mantz, S. W. Allen, D. Rapetti, and H. Ebeling, *The observed growth of massive galaxy clusters - I. Statistical methods and cosmological constraints*, *MNRAS* **406** (Aug., 2010) 1759–1772, [`arXiv:0909.3098`].

[24] A. Pillepich, C. Porciani, and T. H. Reiprich, *The X-ray cluster survey with eRosita: forecasts for cosmology, cluster physics and primordial non-Gaussianity*, *mnras* **422** (May, 2012) 44–69, [`arXiv:1111.6587`].

[25] L. Perotto et al., *Probing cosmological parameters with the CMB: forecasts from Monte Carlo simulations*, *JCAP* **10** (Oct., 2006) 13, [`astro-ph/0606227`].

[26] L. Wolz, M. Kilbinger, J. Weller, and T. Giannantonio, *On the validity of cosmological Fisher matrix forecasts*, *JCAP* **9** (Sept., 2012) 9, [`arXiv:1205.3984`].

[27] A. Jenkins et al., *The mass function of dark matter haloes*, *MNRAS* **321** (Feb., 2001) 372–384, [`astro-ph/0005260`].

[28] J. Tinker et al., *Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality*, *ApJ* **688** (Dec., 2008) 709–728, [`arXiv:0803.2706`].

[29] M. S. Warren, K. Abazajian, D. E. Holz, and L. Teodoro, *Precision Determination of the Mass Function of Dark Matter Halos*, *ApJ* **646** (Aug., 2006) 881–885, [`astro-ph/0506395`].

[30] R. K. Sheth and G. Tormen, *Large-scale bias and the peak background split*, *MNRAS* **308** (Sept., 1999) 119–126, [`astro-ph/9901122`].

[31] R. K. Sheth, H. J. Mo, and G. Tormen, *Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes*, *MNRAS* **323** (May, 2001) 1–12, [`astro-ph/9907024`].

[32] J. L. Tinker et al., *The Large Scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests*, *astro-ph/1001.3162* (Jan., 2010) [`arXiv:1001.3162`].

[33] A. Vikhlinin et al., *Chandra Cluster Cosmology Project. II. Samples and X-Ray Data Reduction*, *ApJ* **692** (Feb., 2009) 1033–1059, [`arXiv:0805.2207`].

[34] K. Vanderlinde et al., *Galaxy Clusters Selected with the Sunyaev-Zel'dovich Effect from 2008 South Pole Telescope Observations*, *ApJ* **722** (Oct., 2010) 1180–1196, [`arXiv:1003.0003`].

[35] B. P. Koester et al., *MaxBCG: A Red-Sequence Galaxy Cluster Finder*, *ApJ* **660** (May, 2007) 221–238, [`astro-ph/0701268`].

[36] H. K. C. Yee and E. Ellingson, *Correlations of Richness and Global Properties in Galaxy Clusters*, *ApJ* **585** (Mar., 2003) 215–226, [`astro-ph/0211096`].

[37] S. Majumdar and J. J. Mohr, *Importance of Cluster Structural Evolution in Using X-Ray and Sunyaev-Zeldovich Effect Galaxy Cluster Surveys to Study Dark Energy*, *ApJ* **585** (Mar., 2003) 603–610, [`astro-ph/0208002`].

[38] H. A. Feldman, N. Kaiser, and J. A. Peacock, *Power-spectrum analysis of three-dimensional redshift surveys*, *ApJ* **426** (May, 1994) 23–37, [`astro-ph/9304022`].

[39] S. Majumdar and J. J. Mohr, *Self-Calibration in Cluster Studies of Dark Energy: Combining the Cluster Redshift Distribution, the Power Spectrum, and Mass Measurements*, *ApJ* **613** (Sept., 2004) 41–50, [`astro-ph/0305341`].

[40] M. Lima and W. Hu, *Self-calibration of cluster dark energy studies: Counts in cells*, *Phys. Rev. D* **70** (Aug., 2004) 043504, [`astro-ph/0401559`].

[41] S. Khedekar, S. Majumdar, and S. Das, *Precision cosmology with a combination of wide and deep Sunyaev-Zel'dovich cluster surveys*, *Phys. Rev. D* **82** (Aug., 2010) 041301, [`arXiv:1005.0387`].

[42] H. Wu, E. Rozo, and R. H. Wechsler, *Annealing a Follow-up Program: Improvement of the Dark Energy Figure of Merit for Optical Galaxy Cluster Surveys*, *ApJ* **713** (Apr., 2010) 1207–1218, [`arXiv:0907.2690`].

[43] A. Mahdavi, H. Hoekstra, A. Babul, C. Bildfell, T. Jeltema, and J. P. Henry, *Joint Analysis of Cluster Observations: II. Chandra/XMM-Newton X-ray and Weak Lensing Scaling Relations for a Sample of 50 Rich Clusters of Galaxies*, *ArXiv e-prints* (Oct., 2012) [`arXiv:1210.3689`].

[44] N. Metropolis et al., *Equation of State Calculations by Fast Computing Machines*, *J. Chem. Phys.* **21** (1953), no. 6 1087.

[45] W. K. Hastings, *Equation of State Calculations by Fast Computing Machines*, *Biometrika* **57** (1970), no. 1 97–109.

[46] A. Hajian, *Efficient cosmological parameter estimation with Hamiltonian MonteCarlo technique*, *Phys. Rev. D* **75** (Apr., 2007) 083525, [`astro-ph/0608679`].

[47] J. Jasche and F. S. Kitaura, *Fast Hamiltonian sampling for large-scale structure inference*, *MNRAS* **407** (Sept., 2010) 29–42, [`arXiv:0911.2496`].

[48] D. Wraith et al., *Estimation of cosmological parameters using adaptive importance sampling*, *Phys. Rev. D* **80** (July, 2009) 023507, [`arXiv:0903.0837`].

[49] J. F. Taylor, M. A. J. Ashdown, and M. P. Hobson, *Fast optimal CMB power spectrum estimation with Hamiltonian sampling*, *MNRAS* **389** (Sept., 2008) 1284–1292, [`arXiv:0708.2989`].

[50] B. D. Wandelt, D. L. Larson, and A. Lakshminarayanan, *Global, exact cosmic microwave background data analysis using Gibbs sampling*, *Phys. Rev. D* **70** (Oct., 2004) 083511, [`astro-ph/0310080`].

[51] J. Jewell, S. Levin, and C. H. Anderson, *Application of Monte Carlo Algorithms to the Bayesian Analysis of the Cosmic Microwave Background*, *ApJ* **609** (July, 2004) 1–14, [`astro-ph/0209560`].

[52] D. L. Larson et al., *Estimation of Polarized Power Spectra by Gibbs Sampling*, *ApJ* **656** (Feb., 2007) 653–660, [`astro-ph/0608007`].

[53] N. Christensen, R. Meyer, L. Knox, and B. Luey, *Bayesian methods for cosmological parameter estimation from cosmic microwave background measurements*, *Classical and Quantum Gravity* **18** (July, 2001) 2677–2688, [`astro-ph/0103134`].

[54] A. Lewis and S. Bridle, *Cosmological parameters from CMB and other data: A Monte Carlo approach*, *Phys. Rev. D* **66** (Nov., 2002) 103511, [`astro-ph/0205436`].

[55] R. A. Fisher, *Equation of State Calculations by Fast Computing Machines*, *J. Roy. Stat* **98** (1935) 39.

[56] M. Tegmark, A. N. Taylor, and A. F. Heavens, *Karhunen-Loeve Eigenvalue Problems in Cosmology: How Should We Tackle Large Data Sets?*, *ApJ* **480** (May, 1997) 22, [`astro-ph/9603021`].

[57] E. Komatsu et al., *Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation*, *astro-ph/1001.4538* (Jan., 2010) [`arXiv:1001.4538`].

[58] M. Chevallier, and D. Polarski, *Accelerating Universes with Scaling Dark Matter*, *Int. J. Mod. Phys. D* **10** (2001) 213 [`gr-qc/0009008`].

[59] M. Lima and W. Hu, *Self-calibration of cluster dark energy studies: Observable-mass distribution*, *Phys. Rev. D* **72** (Aug., 2005) 043006, [`astro-ph/0503363`].

[60] A. Merloni et al., *eROSITA Science Book: Mapping the Structure of the Energetic Universe*, *ArXiv e-prints* (Sept., 2012) [`arXiv:1209.3114`].

[61] S. Borgani, P. Rosati, B. Sartoris, P. Tozzi, Giacconi, R. , and WFXT Team, *Astrophysics and cosmology with galaxy clusters: the WFXT perspective*, *Memorie della Societa Astronomica Italiana Supplementi* **17** (2011) 36, [`arXiv:1010.6213`].

[62] D. G. Gilbank, M. D. Gladders, H. K. C. Yee, and B. C. Hsieh, *The Red-sequence Cluster Survey-2 (RCS-2): Survey Details and Photometric Catalog Construction*, *aj* **141** (Mar., 2011) 94, [`arXiv:1012.3470`].

[63] B. Flaugher, *The Dark Energy Survey*, in *APS April Meeting Abstracts*, p. D7007, Mar., 2012.

[64] E. Rozo et al., *Cosmological Constraints from SDSS maxBCG Cluster Abundances*, *arXiv:0703571* (Mar., 2007) [`0703571`].

[65] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. Balbi, A. J. Banday, R. B. Barreiro, and et al., *Planck early results. VIII. The all-sky early Sunyaev-Zeldovich cluster sample*, *aap* **536** (Dec., 2011) A8, [`arXiv:1101.2024`].

[66] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. Balbi, A. J. Banday, R. B. Barreiro, and et al., *Planck early results. XI. Calibration of the local galaxy cluster Sunyaev-Zeldovich scaling relations*, *aap* **536** (Dec., 2011) A11, [`arXiv:1101.2026`].

[67] M. White and S. Majumdar, *Point Sources in the Context of Future SZ Surveys*, *ApJ* **602** (Feb., 2004) 565–570, [`astro-ph/0308464`].

[68] A. Chaudhuri and S. Majumdar, *Sunyaev-Zel'dovich Scaling Relations from a Simple Phenomenological Model for Galaxy Clusters*, *ApJ Lett.* **728** (Feb., 2011) L41+, [`arXiv:1005.0920`].

[69] A. Chamballu, J. G. Bartlett, and J. Melin, *The Planck SZ Cluster Catalog: Expected X-ray Properties*, *ArXiv e-prints* (July, 2010) [`arXiv:1007.3193`].

[70] S. Khedekar and S. Majumdar, *Using clusters in Sunyaev-Zel'dovich effect plus x-ray surveys as an ensemble of rulers to constrain cosmology*, *Phys. Rev. D* **82** (Oct., 2010) 081301, [`arXiv:1005.0388`].

[71] A. Heavens, *Statistical techniques in cosmology*, *ArXiv e-prints* (June, 2009) [`arXiv:0906.0664`].