The Sunyaev–Zel’dovich effect as a cosmological discriminator

J. M. Diego,1,2* E. Martínez-González,1 J. L. Sanz,1 N. Benitez3 and J. Silk2

1Instituto de Física de Cantabria, Consejo Superior de Investigaciones Científicas, Universidad de Cantabria, Santander, Spain
2Department of Physics & Astronomy, The Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218-2686, USA
3University of Oxford, Astrophysics Department, Keble Road, Oxford OX1 3RH

Abstract

We show how future measurements of the Sunyaev–Zel’dovich effect (SZE) can be used to constrain the cosmological parameters. We combine the SZ information expected from the Planck full-sky survey, $N(S)$, where no redshift information is included, with the $N(z)$ obtained from an optically identified SZ-selected survey covering less than 1 per cent of the sky. We demonstrate how with a small subsample ($\approx$300 clusters) of the whole SZ catalogue observed optically it is possible to reduce the degeneracy among the cosmological parameters drastically. We have studied the requirements for performing the optical follow-up and we show the feasibility of such a project. Finally, we have compared the cluster expectations for Planck with those expected for Newton–XMM during their lifetimes. It is shown that, owing to its larger sky coverage, Planck will detect a factor of $\sim$5 times more clusters than Newton–XMM and also provide a larger redshift coverage.

Key words: galaxies: clusters: general – cosmology: observations.

1 Introduction

Clusters of galaxies have been widely used as cosmological probes. Their modelling can be easily understood as they are the final stage of linearly evolved primordial density fluctuations (Press & Schechter 1974, hereafter PS). Consequently, it is possible to describe, as a function of the cosmological model, the distribution of clusters and their evolution, the mass function, which is usually used as a cosmological test (Bahcall & Cen 1993; Carlberg et al. 1997a; Bahcal & Fan 1998; Girardi et al. 1998; Rahman & Shandarin 2001; Diego et al. 2001a). Therefore, a detailed study of the cluster mass function will provide us with useful information concerning the underlying cosmology.

Following this idea, several groups have tried to constrain cosmological models by using the information contained in the mass function. They compare the observational mass function with the theoretical one given by the PS formalism (Carlberg et al. 1997; Bahcall & Fan 1998) or with simulated ones from $N$-body simulations (Bahcall & Cen 1993; Bode et al. 2001). This method has been shown to be very useful but is limited by the quality of the data. Cluster masses are not very well determined for intermediate to high-redshift clusters and even for low-redshift ones the error bars are still significant. The standard methods for determining cluster masses involving velocity dispersions, cluster richness, lensing and X-ray surface brightness deprojection usually give different answers. It is believed that the best mass estimator is that based on lensing (e.g. Wu & Fang 1997; Allen 1998), but the number of clusters with masses measured with this technique is too low to make this method a reliable technique to build a statistically complete mass function although several attempts have been made (Dahle 2000).

Instead of using the mass function, it is possible to study the cluster population via other functions such as the X-ray flux function, the X-ray luminosity function or the temperature function. The advantage of these distribution functions compared with the mass function is that estimates of X-ray fluxes, luminosities or temperatures of the clusters are less affected by systematics than are the estimates of mass from optical data. The drawback, however, is that to build these functions from the mass function, a relation between the mass and the X-ray luminosity, flux or temperature is needed. These relations are known to suffer from important uncertainties that should be taken into account. These uncertainties have their origin in the intrinsic scatter of these relations and in the quality of the observational data used to build them.

There are three basic wavebands in which galaxy clusters are observed; optical (galaxies bound to the cluster), X-ray (bremsstrahlung emission from the hot intracluster gas) and most recently the millimetre waveband (Sunyaev–Zel’dovich effect, SZE).

The first clusters were observed using optical telescopes and also the first cluster catalogues were based on optical observations [Abell 1958; Abell, Corwin & Olowin (ACO) 1989; Lumsden et al. 1992 (EDCC); Postman et al. 1996 (PDCS); Carlberg et al. 1997b (CNO)]. However, the optical identification of a cluster is not a trivial task. First, it is not easy to define the cluster limits or cluster...
size. When observed in the optical band, a cluster appears as a
group of galaxies that are bound by a common gravitational
potential. However, in the outer parts of the cluster there can be
some galaxies for which the extent to which they are bound to the
cluster is uncertain, or they may even be field galaxies.

Optical identification of clusters has other more important
shortcomings. They specifically suffer from projection effects, that
is, there can be some galaxies in the line of sight that are not bound
gravitationally but they can appear as a bound system because their
images are projected on to a small circle centred on the cluster
position. The best way to reduce this effect is by means of spectral
identification. However, this kind of observation is time consuming
and, when observing distant clusters, spectral identification is only
feasible for the most luminous galaxies in the cluster.

These problems are reduced when the cluster is observed in the
X-ray band. In this band clusters appear as luminous sources due
predominantly to the continuum bremsstrahlung emission coming
from the hot \( T \sim 10^7 \) K intracluster gas. The same intracluster
emission can be used to determine the size of the cluster gas cloud.
Moreover, projection effects are weaker in this case, thus X-ray
surveys are very efficient in detecting clusters. However, with
X-ray surveys it is difficult to detect clusters further than \( z \approx 1 \).
There are two reasons for this. One lies in the fact that the X-ray
cluster emissivity is proportional to \( n_e^2 \), where \( n_e \) is the electron
density, that is, the emissivity drops very quickly from the centre of
the cluster to the edge and only the densest central parts of the
cluster will generate substantial X-ray emission. The detection of
distant clusters for which the X-ray emission is concentrated in the
central parts of the cluster is therefore very difficult since the
apparent angular size will be small and it is more difficult to
distinguish the clusters from point sources. Moreover, the X-ray
flux declines as \( D_l^{-2} = [D_m/(1+z)]^{-2} \) (\( D_m \) is the comoving
distance and \( D_l \) is the luminosity distance). This selection function
limits the redshift at which one cluster can be observed by actual
X-ray detectors that are blind to the earlier stages of cluster
formation \( (z \gtrsim 2) \).

In the millimetre waveband the situation is quite different. The
SZE surface brightness varies as \( n_e \). Therefore, the SZE brightness
profile drops more slowly than in the X-ray case. This normally
makes clusters observed through the SZE show a larger angular
size since we can more easily observe the outer parts of the cluster
where the X-ray emission is too low to be detected. The second
reason why SZE surveys can be more efficient than X-ray surveys
in detecting high-redshift clusters is that in the X-ray band, the total
flux decays as \( D_l^{-2} \) with redshift while the integrated SZE flux
varies as \( D_l^{-2} = [D_m/(1+z)]^{-2} \) (where \( D_m \) is the angular
diameter distance). \( D_l \) grows more slowly than \( D_m \) and even decreases after
a certain redshift \( z \approx 1 \). Therefore, the SZE flux drops more slowly
with distance (or even increases) in the case of the SZE than in the
X-ray case. Another advantage is that, as in the X-ray situation, the
identification of galaxy clusters through the SZE is less affected by
projection effects. All of these reasons make the SZE the preferable
way of observing distant clusters.

Our interest on the SZE is twofold. First, it can be considered as
a contaminant of the cosmological signal (cosmic microwave
background, CMB) and therefore good knowledge of this effect is
required in order to perform an appropriate analysis of the CMB
data. However, it can also be considered as a very sensitive tool to
measure the mass–space cluster distribution. In this paper we will
concentrate our effort on this second aspect.

Planned CMB surveys will also be sensitive to the SZE
distortion induced by galaxy clusters (MAP, Planck). These
surveys will cover a wide area of the sky and they are expected to
detect the SZE signature for thousands of clusters. Furthermore,
proposed and current millimetre experiments will measure the SZE
for hundreds of clusters in the near future (e.g. AmiBa, Lo et al.
2000; LMT-BOLOCAM, López-Cruz & Gaztaña 2000; CBI,
Udomprasert, Mason & Readhead 2000; AMI, Kneissl et al. 2001;
or the interferometer proposed in Holder et al. 2000). The
cosmological possibilities for these new data sets are very relevant
as they will have a statistically large number of clusters and they
will significantly improve the redshift coverage. For a realistic
prediction of the power of future SZE surveys we should consider
the detector characteristics of one of these planned experiments. In
this work we will focus our attention on the expected SZE detec-
tions for the Planck mission and we will study the possibilities of
the anticipated data for probing cosmological models. Planck will
observe the whole sky at nine frequencies in the millimetre range
(including those frequencies where the SZE is most relevant) and
with angular resolutions ranging from 5- to 33-arcmin FWHM (see
Fig. 1 below). Observation of the SZE at different frequencies will
make the task of identifying clusters easier because of the peculiar
spectral behaviour of the SZE signal that can be very well
recognized with the nine Planck frequencies. As can be seen from
Fig. 1, the best channels are those at 100 GHz \( (x = 1.76) \), 143 GHz
\( (x = 2.5) \) and 353 GHz \( (x = 6.2) \) together with the channel at
217 GHz \( (x = 3.8) \) where the thermal SZE vanishes. A cross-
correlation of these channels, including knowledge of the spectral
shape factor, will allow us to discriminate between the SZE,
foregrounds and CMB (Diego et al. 2001b).

As we will show in the following sections, once the cluster

Figure 1. Spectral shape of the thermal SZE. Planck channels (vertical
lines) are between 30 and 800 GHz, also including the 217-GHz channel
where the thermal SZE vanishes. The amplitudes of the lines are
proportional to the signal-to-noise ratio expected for the SZE in that
channel. The S/N ratio depends on the spectral shape factor \( f(x) \) since the
signal is proportional to \( f(x) \) (see equation 5 below) and the noise is
proportional to the sensitivity per resolution element \( \sigma \); \( S/N = \infty f(x)/\sigma \).
The best channels to detect the SZE decrement/increment will be those of
the HFI instrument at 100 GHz \( (x = 1.7) \), 143 GHz \( (x = 2.5) \) and 353 GHz
\( (x = 6.2) \), respectively, together with that at 217 GHz \( (x = 3.8) \). At
100 GHz \( (x = 1.7) \) there are two channels; the LFI (solid line) and the HFI
(dotted line).
population has been normalized at low redshift \((z < 0.2)\), the information concerning the cluster distribution at higher redshifts is crucial for determining the underlying cosmology and the SZE can be the tool to obtain that information.

The structure of this paper is the following; in Section 2 we review some basics of the SZE and some useful definitions, which will be of interest in subsequent sections. In Section 3, we show how through the SZE we can investigate the cluster population and evolution. We also show in Section 4 how future SZE surveys should be complemented with optical observations of a small subsample of SZE-selected clusters in order to provide redshift information of those clusters; in this way, it is possible to reduce the degeneracy in the cosmological models describing the data. We apply this idea in Section 5 to simulated \(\text{Planck} \) SZ data. In this section we obtain some interesting results on the possibilities of those future data. Finally, we discuss our results and present our conclusions in Section 6.

2 THE SUNYAEV–ZEL’DOVICH EFFECT

Since Sunyaev & Zel’dovich predicted that clusters would distort the spectrum of CMB photons when they traverse a galaxy cluster (Sunyaev & Zel’dovich 1972), several detections of this effect have been made. At present, the number of clusters with measured SZE is small because they are limited by the detector sensitivity (a typical SZE signal is of the order of \(10^{-4}\) in \(\Delta T / T\) with \(T = 2.73\) K). However, the sensitivity in the detectors is getting better and better and in a few years time the number of detected clusters through the SZE will make this effect an important tool for exploring clusters at high redshift.

When CMB photons cross a cluster of galaxies, the spectrum suffers a distortion caused by inverse Compton scattering. The net distortion in a given direction can be quantified by the cluster Comptonization parameter, \(\gamma_c\), which is defined as

\[
\gamma_c = \frac{\sigma_T k_B}{m_e c^2} T n_e(l) dl.
\]  

The integral is performed along the line of sight through the cluster. \(T\) and \(n_e\) are the intracluster temperature and density, respectively; \(\sigma_T\) is the Thomson cross-section, which is the appropriate cross-section in this energy regime and \(k_B\) is the Boltzmann constant.

The flux and the temperature distortion are more widely used than the Compton parameter because these are the quantities that are determined directly in any experiment when observing clusters at millimetre frequencies. The distortion in the background temperature is given by

\[
\frac{\delta T}{T} = g(x)\gamma_c,
\]  

where \(x = h v k T_{\text{CMB}}\) is the frequency in dimensionless units and \(g(x) = 1 \cot h(x/2) - 4\) is the spectral shape factor. There is a second kind of distortion owing to the bulk velocity of the cluster. It is known as the kinetic SZE. However, this contribution is much smaller than the thermal contribution and we will not consider it in this work.

As can be seen from equation (2), there is no redshift dependence on the temperature distortion, i.e. the same cluster will induce the same distortion in the CMB temperature, independent of the cluster distance (except for relativistic corrections). The only redshift dependence of the total SZE flux is caused by the fact that the apparent size of the cluster changes with redshift.

The total SZ flux is given by the integral

\[
S_{\text{total}}(\nu) = \int \Delta l(\nu, \theta) d\Omega,
\]

where the integral is performed over the solid angle subtended by the cluster. \(\Delta l(\nu, \theta)\) is the change in intensity induced by the SZE and is given by \(\Delta l(\nu, \theta) = I_0 f(x)\gamma_c\) where \(I_0 = (2hc^2) / (k_B T_{\text{CMB}})\) and

\[
f(x) = \frac{x^4e^x}{(e^x - 1)^2} \left[ \frac{x(e^x + 1)}{(e^x - 1)} - 4 \right],
\]

is the spectral shape (see Fig. 1).

In Fig. 1 we show the spectral shape of the SZE, \(f(x)\). In the same plot the nine frequency channels from \(\text{Planck}\) are shown as vertical lines. The amplitude of the lines are proportional to the signal-to-noise ratio per resolution element in each channel (assuming that the clusters are unresolved).

Equation (3) can be easily integrated assuming that the cluster is isothermal. Then the integral can be reduced to \(\int d\Omega \int n_e(\theta) d\ell\), which can also be transformed into \(D_s^2(z) \int dV n_e(\theta) = D_s^2(z) M_{\text{gas}} / m_p = D_s^2(z) (f_b/m_p) M\), where \(M, f_b\) and \(m_p\) are the total mass, baryon fraction and the proton mass, respectively. In this approach, we have assumed that the gas is only composed of ionized hydrogen. Finally, we obtain

\[
S_{\text{SZE}} = \frac{3.78 f(x)f_b T M_{15}}{D_s^3(z)}.
\]  

In the previous expression \(T\) is given in kelvin, \(M_{15}\) in \(10^{15} M_\odot\) and \(D_s(z)\) in Mpc. In these units, the flux is given in mJy. If we now introduce the \(h\) dependence in \(M_{15}\) \((10^{15} h^{-1} M_\odot)\), \(D_s(z)\) \((h^{-1} \text{Mpc})\) and \(f_b \approx 0.06 h^{-1}\) then the flux is given also in mJy units.

Equation (5) tells us that the total flux is given basically by three terms: temperature, mass and angular-diameter distance to the cluster. In contrast with the X-ray flux, where in the integral over the cluster volume one should consider \(n_e^2\) instead of just \(n_e\), the total SZE flux does not depend on the cluster profile. Therefore, for an isothermal cluster, no assumption concerning the electron density profile is needed when computing the total SZE flux. This aspect is very important since for the \(\text{Planck}\) resolution, only a small percentage of the clusters (= 1 per cent) will be resolved and as a first approximation we can consider that all the fluxes can be computed from equation (5). This fact will simplify our calculations significantly and, simultaneously, will reduce the uncertainty owing to the lack of precise knowledge of the cluster density profile.

The second important consequence that we can note from the previous equation concerns the angular-diameter distance dependence of the total flux. In X-rays, the total flux depends on \(D_s^2(z)\). If we compute the ratio \(D_s^2(z) / D_L^2(z)\) it goes as \((1 + z)^{-4}\), that is, the \(S_{\text{SZE}}\) drops by a factor of \((1 + z)^{-4}\) slower than \(S_L\). Thus for high-redshift clusters there is a considerable advantage in using millimetre surveys (SZE) as compared with X-ray surveys.

If the cluster is unresolved by the antenna, then the total flux from the cluster (equation 5) can be considered to be contained within the resolution element of our antenna. If in contrast, the cluster is resolved, then the previous approximation is no longer adequate and we should integrate the surface brightness over the cluster solid angle.
The gas density profile can be well fitted by a $\beta$-model.
\[ n_g(r) = n_0 (1 + (r/r_c)^2)^{-3\beta/2}, \]
where $n_0$ is the central electron density ($n_0 \approx 2.0 \times 10^{-3}$ $(1 + z)^{2/3} h^{1/2} \, e^{-cm^{-3}}$) and $r_c$ is the core radius of the cluster. The observed values of $\beta$, obtained from X-ray surface brightness profiles, typically range from 0.5 to 0.7 (Jones & Forman 1984; Markevitch et al. 1998).

3 SZE AS A PROBE OF THE COSMOLOGICAL MODEL

As we mentioned in the introduction, an estimate of the cluster mass function can be used to constrain the cosmological parameters. However, such estimates are affected by the systematics in the mass estimators. It is possible to use other functions to explore the cluster population such as, for instance, the temperature function $d^2N/dTdz$ (Henry & Arnaud 1991; Eke et al. 1998; Viana & Liddle 1999; Donahue & Voit 1999; Blanchard et al. 2000; Henry 2000). The connection between the mass function and the temperature function is the $T-M$ relation,
\[ \frac{d^2N}{dTdz} = \frac{d^2N}{dMdz} \frac{dT}{dM}. \]

Temperature estimates are more reliable than mass estimates. Therefore, the temperature function should be less affected by the systematics than the mass function. However, the scatter in the $T-M$ relation introduces new uncertainties, which should be taken into account. Similar problems occur with the X-ray luminosity function and the X-ray flux function that have been widely used in the literature to constrain the cosmological parameters (Mathiesen & Evrard 1998; Borgani et al. 1999; Diego et al. 2001a). In these cases new uncertainties appear owing to the scatter in the $Lx-T$ relation. In a previous paper (Diego et al. 2001a), we considered these ideas and applied them to constrain the cosmological parameters by fitting the model to different data sets: mass function, temperature function, X-ray luminosity and flux functions. That work was innovative in the sense that we considered all the previous data sets simultaneously in our fit (and not just one as is more usual) and we looked for the best-fitting model to all the data sets considered. It is also important to remark that in that work we considered the cluster scaling relations as free-parameter relations. These two points are very relevant because when fitting cluster data sets it is important to check that the best-fitting models agree well with other data sets. Otherwise, we should re-examine the assumptions made in the model. One of these assumptions is, for instance, the cluster scaling relations $T-M$ or $Lx-T$. These relations are commonly fixed a priori. However, they are known to suffer from important scatter as we have shown in our paper. In fact, we found that not all the scaling relations considered previously in earlier works were appropriate to conveniently describe several cluster data sets in a simultaneous fit. By fitting our free-parameter model to all the data sets, we obtained, not only the cosmological parameters, but also the best scaling parameters.

In our work we found as our best-fitting parameters those given in Table 1. In that paper the authors have connected the millimetre flux function with the observed X-ray luminosity function through the $Lx-T$ and $T-M$ relations. If we consider a were free parameters. We found that the two models in Table 1 were indistinguishable when they were compared with the data. In fact, both models agree well with the data considered and there was no way to determine which one was the best. In order to distinguish them, more and better-quality data are needed. Ongoing X-ray experiments (Chandra, Newton—XMM) will help to do that, as will current and proposed optical surveys (e.g. Gladders et al. 2000). The models which were indistinguishable in the low-redshift interval will show different behaviour at higher redshift. However, as we mentioned in the introduction, the waveband in which distant clusters will be detected most easily, is not the optical nor the X-ray band, but the millimetre band. The usefulness of the SZE as a tool to measure cosmological parameters has motivated much previous work (e.g. Silk & White 1978; Bartlett & Silk 1994; Markevitch et al. 1994; Barbossa et al. 1996; Aghanim et al. 1997) and current work (Holder et al. 2000; Grego et al. 2001; Mason et al. 2001, this work).

Using SZE data it is possible to estimate the cosmological parameters by looking at some SZE-derived function related to the mass function. We just need a measurement of a cluster quantity related to its mass. As we have shown in equation (5), the total SZE flux can be such a quantity. From that equation it is possible to build a millimetre flux function ($dN/dS_{\text{mm}}$) from a mass function ($dN/dM$). An interesting alternative to this approach can be found in Xue & Wu (2001). In that paper the authors have connected the millimetre flux function with the observed X-ray luminosity function through the $Lx-T$ relation.

Table 1. Best $\Lambda$CDM (flat) and OCDM ($\Lambda = 0$) models. Both models are indistinguishable in the redshift regime of the data considered in Diego et al. (2001a).

| Model      | $\sigma_8$ | $\Gamma$ | $\Omega_m$ | $T_0$ ($10^8 h^{-2} K$) | $\alpha$ | $\psi$ |
|------------|------------|----------|------------|-------------------------|----------|--------|
| $\Lambda$CDM | 0.8        | 0.2      | 0.3        | 1.1                     | 0.75     | 1.0    |
| OCDM       | 0.8        | 0.2      | 0.3        | 1.1                     | 0.8      | 1.0    |

Figure 2. The SZE integrated number counts of the cluster population predicted at millimetre wavelengths (353 GHz) for the best flat (solid) and open (dotted) models in Table 1. In the plot three redshift shells are represented: top $z < 1$, middle $z \in [1, 2]$ and bottom $z > 2$.
future millimetre experiment (such as Planck) where thousands of clusters are expected to be detected through the SZE, probably a fit to the millimetre flux function would be able to distinguish the models in Table 1. In Fig. 2 we show the expected number of clusters with total flux above a given flux for the two previous models. These are integrated fluxes, i.e. we assume that the clusters are not resolved and that their total flux falls into the antenna beam size. This assumption is appropriate when the antenna size is above several arcmin as is the case for the Planck satellite where its best resolution is $\text{FWHM} = 5 \text{ arcmin}$.

In Fig. 2 we did not show the whole $0 < z < \infty$ integrated $N(S)$ curve since it looks quite similar to the two curves on the top of the figure and they still look indistinguishable. To compare them in a more realistic way we have performed a Kolgomorov–Smirnov test where we have compared the $N(S)$ curve for a realization (over two-thirds of the sky) of the $\Lambda$ cold dark matter (LCDM) model with the mean expected values of both OCDM and CDM models. The KS test was unable to distinguish between the models.

As can be seen from the figure, both models predict the same number of clusters in the $z < 1$ redshift interval (top). That situation was similar in Diego et al. (2001a) where most of the cluster data were at low $z$ and it was not possible to discriminate between both models.

However, the situation changes at redshifts $z > 1$. In the redshift bin $z \in (1, 2)$ (middle) the differences between both models are significant. These differences are increased when we compare the models in the redshift bin $z > 2$ (bottom) where in the OCDM model two orders of magnitude more clusters are expected than in the LCDM case above $S = 30 \text{ mJy}$. This flux ($S = 30 \text{ mJy}$ at 353 GHz) is expected to be the flux limit of Planck (see Appendix A where we estimate that limit based on MEM residuals). However, this limiting flux will depend on the method used to identify such clusters in the maps. The final method to be used with Planck is still under development. MEM methods work very well (Hobson et al. 1998, 1999) and at present are preferred.

In Fig. 3 we show the number of clusters with fluxes above that limiting flux of Planck as a function of redshift. Again the differences are not significant below $z = 1$, but they are quite relevant above that redshift. By looking at Figs 2 and 3, we see that the differences in the cosmological models are more evident at high redshifts. In order to discriminate between the cosmological models one should consider not only the cluster population at low redshift (normalization), but also the cluster population at high redshift (evolution). A recent application of this idea can be found in Fan & Chiu (2001) where the authors study the function $r = N(z < 0.5)/N(z > 1)$ as a function of the limiting flux. That work suggests that a limited knowledge of the redshift of the cluster would give a constraint on the cosmological parameters.

Both Figs 2 and 3 suggest that with future SZE data it will be possible to improve the determination of the cosmological parameters. From all of these plots, it is evident that Planck (together with redshift information for a small subsample, see the next section) would be able to discriminate between the models that previously were indistinguishable when they were compared with present X-ray and optical data. Furthermore, the accuracy in the free parameters obtained in previous works will be increased with these new data.

3.1 Monte Carlo simulations: Planck versus Newton–XMM

The previous discussion was based on theoretical expectations of the mean number of SZE detections expected at different redshifts. We want to go further by computing Monte Carlo (Press–Schechter) realizations of the expected SZE on a specific patch of the sky. In order to compare the millimetre and X-ray bands we will also compare the expected SZ detections in this area for Planck with those based on X-ray expectations for Newton–XMM in the same patch of sky.

The simulations were done over a patch of the sky of $12.8 \times 12.8 \text{ deg}^2$ and with a pixel size of $1.5 \times 1.5 \text{ arcmin}^2$ filtered with a FWHM of $5 \text{ arcmin}$ and at a frequency of 353 GHz, following the characteristics of the 353-GHz channel of Planck. The parameters for this simulation corresponds to the CDM model of Table 1.

The total number of clusters is about 20,000 in one of these simulations. The mean Comptonization parameter is well below the FIRAS limit ($\sim 10^{-6}$ compared with $2.5 \times 10^{-5}$, Fixsen et al. 1996). The resulting distribution of clusters is shown in Fig. 4. The allowed range of simulated masses was $2.0 \times 10^{13} < M < 1.0 \times 10^{16} \text{ h}^{-1} \text{ M}_\odot$. The lower limit is at the boundary between a small cluster and a galaxy group and it is the mass corresponding to a cluster with temperature $T \sim 1 \text{ keV}$ which can be considered as the minimum temperature for a virialized cluster.

In Fig. 4 each point corresponds to one cluster with mass $M$ and redshift $z$. This distribution of points follows the Press–Schechter mass function used to make the simulation. The Press–Schechter model is in agreement with the observational constraints given in Bahcall & Cen (1993) and Bahcall & Fan (1998) as shown in Diego et al. (2001a).

The dotted lines should be considered as the maximum mass expected in that simulation at each redshift. These lines were computed by searching for the mass at which the mean number of clusters given by Press–Schechter was equal to 1 at each redshift. This maximum-mass limit depends on the surveyed area of the sky, since the probability of having a cluster with a given mass increases with the solid angle. This fact is shown in Fig. 4 where we compare the maximum-mass limits for two different solid angles, $(12.8^\circ)^2$ and two-thirds of the sky. In a given simulation it is
Figure 4. $\Lambda$CDM model. Each cluster is represented as a dot in this plot. Clusters marked with a big black dot have fluxes (in the 0.5–2 keV band) above the limiting flux $S_X = 1.4 \times 10^{-14}$ erg s$^{-1}$ cm$^{-2}$ and they are the clusters expected to be detected by Newton–XMM in the XCS (Romer et al. 1999). Cluster marked with an open circle are those with a SZE flux above 30 mJy at 353 GHz and they are the clusters expected to be detected by Planck. Solid lines represent the masses for which the flux equals the limiting flux of Planck (30 mJy), and the XCS. Dotted lines show the maximum cluster mass expected in a solid angle equal to 2/3 of the sky (top) and (12.8)$^2$ (bottom). Finally, the heavy solid line indicates the selection function for a hypothetical optical survey made with a 10-m class telescope and 2 h of observation (eight bands and 900 s per band, see Section 4.1).

Although Newton–XMM will see many more clusters at low redshift, Planck will observe a greater proportion of high-redshift clusters. The reason for this is the different selection functions used. The X-ray flux varies as $S_X \propto D_L^2(z)$ and in order to detect clusters deeper in redshift they must be more luminous (more massive) in order to compensate the monotonically increasing function $D_L(z)$. In comparison, the SZE flux goes as $S_{\text{SZE}} \propto D_L^2(z)$ and at redshift $\approx 1$, the angular diameter distance reaches a maximum and after that it starts to drop to smaller values. Therefore, the masses needed to provide a particular flux, $S_{\text{SZE}} = 30$ mJy, can be smaller at redshift $z > 1$ than the required masses at $z \approx 1$.

The point where the selection function crosses the maximum-mass limit corresponds to the maximum redshift expected in the sample. For a survey covering (12.8)$^2$ both Planck and XMM reach the same redshift ($z \approx 1.3$) in the $\Lambda$CDM case and XMM goes slightly deeper in the OCDM case. However, Planck will cover a much larger solid angle than XMM does.

As noted in Romer et al. (2001), the XMM Cluster Survey (XCS) will cover $\sim 800^2$ and will contain more than 8000 clusters. In contrast Planck will observe the full sky, that is the sky coverage will be $\sim 50$ times larger than that of the XCS. In order to compare Newton–XMM with Planck this difference in sky coverage should be considered. In Fig. 6 we show the number of detected clusters expected by Planck and Newton–XMM (XCS) where we have taken into account the sky coverage and limiting fluxes expected for both missions.

Now, the differences between Planck and Newton–XMM are possible to have masses above or below these maximum masses owing to the random nature of the Monte Carlo simulations. The smaller the solid angle of the survey, the larger the scatter around the maximum-mass limit.

Clusters marked with an open circle correspond to those with a total flux above 30 mJy and according to our criterion, these clusters would be detected by Planck (see Appendix A). There are 185 clusters above this limit and only one of them would be observed above $z = 1$ in this simulation. As a comparison we show the same picture but corresponding to the OCDM model in Table 1 (Fig. 5). In this case $\approx 15$ clusters are detected above redshift $z = 1$. This comparison demonstrates how a small region of the sky can show up the differences between the two models in Table 1. Both models are nearly indistinguishable below $z = 0.7$, but they differ significantly above that redshift. We will return to this point in Section 4.

As a comparison with the number of detected clusters expected with Planck, we show in Figs 4 and 5 the clusters expected to be detected by Newton–XMM (big solid circles, $S_X > 1.5 \times 10^{-14}$ erg s$^{-1}$ cm$^{-2}$) in the 0.5–2 keV band, Romer et al. 2001).

Solid lines represent the minimum mass as a function of redshift for which the corresponding flux is above the limiting flux in Planck and Newton–XMM, respectively, that is, they are the selection functions for both missions. The selection function for XMM is better than the selection function for Planck at low to intermediate $z$ but above $z \approx 1.3$ the selection function for Planck is better in the $\Lambda$CDM case. In the OCDM case the selection function for XMM is better at almost all redshifts.
evident. The large sky coverage of Planck together with the constancy of the SZE surface brightness with redshift will allow this satellite to detect a more significant number of clusters than Newton–XMM in the XCS. Also, we can conclude from this plot that, for the best-fitting models in Table 1, no clusters are expected to be detected above $z < 3$, neither for the OCDM nor the LCDM. However, the cluster abundance between $z \approx 1$ and $2$ will provide a definitive probe of the cosmological parameters and the cluster scaling relations.

An important consequence of the previous plots (Figs 4 and 5) is that only a small portion of the whole sky would be needed to distinguish between the two models $\Lambda$CDM and OCDM. This is an important point because only spectral identification of $\approx 200$ (not resolved) random selected clusters from the whole catalogue would be needed. However, it is important to answer the question of what is the minimum number of clusters needed to distinguish the models in Table 1 at, for instance, the $3\sigma$ level? We will try to answer this question in the next section.

4 AN OPTICALLY IDENTIFIED SZE-SELECTED CATALOGUE

As we have seen in Fig. 2, the information provided by a hypothetical $N(S)$ curve (even if this curve corresponds to a full-sky survey) will be insufficient to distinguish the two models considered in this figure. Redshift information will be needed in order to make the distinction. Different cosmological models predict different cluster populations as a function of redshift. If we analyse the evolution of the cluster population with $z$, then we should be able to discriminate between these models. However, to study the evolution of the cluster population we need spectral identification of the clusters (or at least optical observations in several bands in order to obtain photometric redshifts) since the SZE does not provide any estimate of $z$. Performing these observations for a hypothetical full-sky SZ catalogue would be a huge task, but if only a small number of unresolved clusters need to be identified then the work is significantly simplified. Now, we should ask the question, how small can our optically identified sample be if we want to distinguish between, for instance, the two models in Table 1? To answer that question we have compared the curves $N(S_{\text{mm}} > 30 \text{ mJy} > z)$ for the two models in which we are
By determining the redshift of \( \approx 300 \) clusters randomly selected from the Planck catalogue we could distinguish both models. The differences would show up in the small subsample of clusters having \( z > 0.6 \).

We require that at a given redshift, both curves must be distinguished at a 3\( \sigma \) level, that is, we require the condition

\[
N^{\Omega}(z) - 3\sqrt{N^\Omega(z)} = N^A(z) + 3\sqrt{N^A(z)} \quad \text{(assuming Poissonian statistics),}
\]

where \( N^{\Omega}(z) \) and \( N^A(z) \) are the number of detected clusters above a given \( z \) in the two cases OCDM and ACDM, respectively.

Since we know \( N^{\Omega}(z) \) and \( N^A(z) \) at each \( z \) for the two models, then we can compute the required total number of clusters for which we should determine \( z [N_{\text{obs}}(z)] \) in order to satisfy the previous condition at each redshift. In Fig. 7 we show this calculation for three different selection criteria of the clusters. In each one of the lines we show the total number of clusters randomly selected from the catalogue for which we should determine their redshifts in order to distinguish (at a 3\( \sigma \) level) \( N^{\Omega}(z) \) and \( N^A(z) \) at redshift \( z \). Each one of the lines represent a different criteria selection: select only Planck clusters with total flux \( S_{\text{mm}} > 100 \) mJy, top, \( S_{\text{mm}} > 30 \) mJy, middle, and \( 30 < S_{\text{mm}} < 40 \) mJy, bottom. As can be seen from the plot, randomly about \( 300 \) clusters with \( S_{\text{mm}} > 30 \) mJy from the full Planck catalogue and determining the redshift for each of them will allow us to distinguish the two models at a 3\( \sigma \) level above \( z = 0.6 \) just by looking at the \( N(z = 0.6) \) curve. That is, in the sample of 300 clusters, most of them will be at low redshift but a small subsample will be above \( z = 0.6 \). Then, by fitting the portion of clusters above \( z = 0.6 \) to the ACDM and OCDM models we can distinguish them.

Compare, for instance, the cluster population detected by Planck above \( z \approx 0.6 \) in Figs 4 and 5. In both simulations there are \( \approx 200 \) clusters above 30 mJy, but only a few of them are above \( z = 0.6 \) in the ACDM model, while many more are above the same redshift in the OCDM model. The explanation for this fact is given by the different evolution of the cluster formation in both models.

In the ACDM case there is a cooling phase (or inflection point in the acceleration parameter) at \( z \approx 0.7 \) that helps to form structure at this redshift. This phase is not present in the OCDM case where there is a redshift \( z \approx 2 \) below which the collapse of linear fluctuations is inhibited by the fast expansion of the Universe.

Choosing the selection criteria \( 30 < S_{\text{mm}} < 40 \) mJy, it is possible to reduce slightly the number of clusters to be observed optically. If in contrast, only the brightest clusters with \( S_{\text{mm}} > 100 \) mJy are identified optically, then we would need a significantly larger number of clusters in order to make the distinction between the models, although we could distinguish them at lower redshifts (\( \approx 700 \) clusters to be observed (with all redshifts) for distinguishing the models by looking at the portion of clusters with redshifts above \( \approx 0.4 \)).

### 4.1 Cluster optical detection simulations

Probably the most cost efficient way of identifying the galaxy clusters detected by Planck in the optical range is using photometric redshifts. Even a rough estimate of the photo-z allows one to considerably reduce the background galaxy contamination and enhance the contrast density of the cluster. In addition, although the error in the photo-z of an individual galaxy is usually \( \approx 0.1 \) at \( z < 1 \) (Benitez 2000), the total error in the cluster redshift will be \( 0.1/\sqrt{N_{\text{cl}}} \) (where \( N_{\text{cl}} \) is the number of galaxies in the cluster). To estimate the feasibility of detecting Planck SZ cluster candidates using optical imaging, we perform simulations based on empirical information. Since extensive data are only available for relatively low-redshift clusters, this has the disadvantage of ignoring evolution effects. However, it has been shown that the evolution of the cluster early-type galaxies is not dramatic up to \( z \approx 1.3 \) (Rosati et al. 1999). Therefore, at worst, this makes the results obtained here a conservative, lower limit on the detectability of high-redshift clusters, since any reasonable luminosity evolution would tend to make the cluster galaxies bluer and brighter, increasing the number of detected galaxies with respect to the non-evolution case.

Wilson et al. (1997) represent the luminosity function of bright galaxy clusters as the superposition of a Schechter function and a Gaussian for the brightest cluster galaxies. The spectral fraction for the cluster galaxies can be derived by comparing the \( V-I \) colours of galaxies in A370 and CL 1447+23 (Small et al. 1997) with those expected for El, Sbc, Scd and Im spectral templates (Colemam, Wu & Weedman 1980). With the above luminosity function and spectral fractions (extended to 8-mag fainter than \( M_v \)) we generate a bright galaxy cluster at \( z = 0.18 \). By redshifting this cluster, we generate a series of mock cluster catalogues containing the \( I \)-band magnitude and the spectral type \( T \) from \( z = 0.2 \) to 2.0. The \( I \)-band magnitudes are transformed using \( K \)-corrections derived from the templates of Coleman et al. To model the surface number counts distribution of the clusters, we use an \( n(R) \propto R^{-0.3} \) law, found by Vlchez et al. (in preparation) to agree well with the projected galaxy density in the central regions of galaxy clusters. To link the optical results with X-ray and SZ quantities, we integrate the luminosity function of the cluster in the \( V \) band, and assume \( MIL = 300 \), which leads to \( M \approx 10^{15} M_{\odot} \) for an A1689-like cluster.

We simulate the background galaxy distribution \( n(z, T, I) \) using the Hubble Deep Fields (Williams et al. 1996, 2000). For the redshifts, we use the spectroscopic results of Cohen et al. (2000) for the HDFN and photo-z obtained using the BPZ code (Benitez 2000) for the rest of the HDFN galaxies and those of the HDFS.

Once we have \( I, z, T \) catalogues for all the galaxies contained in the field, we generate \( UBVRIJKH \) magnitudes using the above-mentioned template library, enlarged for the HDF galaxies using two starburst templates from Kinney et al. (1996). Gaussian
photometric errors are added to these ideal magnitudes using empirical relationships derived from real observations with 10-m class telescopes, scaled by the square root of the exposure time needed to simulate a 900 s per band observation.

For cluster detection, we look for an overdensity of ellipticals with respect to the expected background population. The reason for doing this instead of using the whole cluster population is that the density contrast is much higher for this type of object. Even a moderate cluster at \( z \approx 1 \) (Benitez et al. 1999) stands out conspicuously against the relatively sparse numbers of field ellipticals (see also Gladders & Yee 2000). Therefore, we estimate photometric redshifts for all the galaxies within an angular aperture corresponding to \( \approx 1 \) h\(^{-1}\) Mpc, classify them into different spectral types and construct a redshift histogram for the early types. The presence of a ‘spike’ in the redshift histogram will indicate the existence of a cluster. The signal-to-noise ratio of such a detection is

\[
SIN = \frac{[N(z) - \langle N(z) \rangle] \sigma_g(z)}{\sigma(z)},
\]

where \( \sigma_g(z) \) is the expected fluctuation in the galaxy numbers within a redshift slice centred on \( z \). Its value can be estimated as

\[
\sigma_g(z)^2 = \frac{\langle N(z) \rangle^2}{S(z)} \left[ w(\theta_{12}) D S_1 D S_2 \right],
\]

where \( N(z) \) is the detected number of galaxies with area \( S \), \( \langle N(z) \rangle \) is the expected average density and \( w(\theta_{12}) \) is the two-point correlation function within the redshift slice. For most of the redshifts considered, 1 Mpc \( h^{-1} \) corresponds to \( \approx 4 \) arcmin. Taking the amplitude of \( w(\theta_{12}) \) to be \( A \approx 7.6 \times 10^{-3} \) within a \( z = 0.2 \) slice (Brunner, Szalay & Connolly 2000), which is approximately the same redshift interval used here to detect the clusters, the value of \( \sigma(z) \) is roughly \( \sigma_g(z)^2 \approx \langle N(z) \rangle [1 + 3.27 \times 10^{-3} \langle N(z) \rangle] \). This number may be an underestimate since the clustering strength of the early types is known to be higher than for the general galaxy population. The numbers below are based on a 3\( \sigma \) detection limit as defined by the above equation. There are plenty of other methods, parametric and non-parametric, which will probably be more efficient in finding clusters (see, e.g., Nichol et al. 2000).

However, again, we think that using a relatively simple procedure is a good idea concerning the practicality of this approach. A reasonable observing strategy will be to start with those clusters not detected in shallower imaging surveys, e.g., the SDSS catalogue, and depending on the redshift/luminosity range of interest (e.g., low-mass/low-redshift clusters or high-mass/high-redshift ones), to use only a small subsample of the BVRIJHK filter set mentioned above, which brackets the 4000-Å break at the required redshift, and which would be enough to detect the early types. If one desired to reach a higher precision in the photo-\( z \) estimate, or wanted to reach the limits shown in Figs 4 and 5 at all the redshift intervals, then the whole filter set should be used. The optical selection function in Figs 4 and 5 is quite jagged. Apart from relatively smooth effects such as the cosmological dimming or the K-corrections that determine the general trend, the detectability of the clusters is affected significantly, at least at \( z < 1 \), by the relative placement of the redshifted 4000-Å break with respect to the filter set, especially to the R- and I-band filters, i.e., those which go deeper for a fixed exposure time. When the break falls almost exactly in between these two filters, the photo-\( z \) precision is improved, whereas if the break is close to the central position of a filter the redshift error increases and the estimation is more easily affected by colour/redshift degeneracies. Therefore, the relative exposure times and the filter choice should be optimized depending on the redshift interval that it is being targeted.

### 5 Estimating the cosmological model from \( N(S) \) and \( N(Z) \)

Previous work (Kitayama & Suto 1997; Eke et al. 1998; Mathiesen & Evrard 1998; Borgani et al. 1999; Diego et al. 2001a) has shown the power of X-ray surveys for constraining the cosmological model. From the previous sections we conclude that the SZE data can also be used for the same purpose (see also Markevitch et al. 1994; Barbosa et al. 1996; Aghanim et al. 1997; Holder & Carlstrom 1999; Majumdar & Subrahmanyan 2000; Fan & Chiueh 2001).

As we have seen in the previous sections, both the \( N(S) \) and the \( N(> 30 \) mJy, \( z \)) curves can be used to study the cluster population, \( N(S) \) being the curve with the larger number of clusters (although with no \( z \) information) and \( N(> 30 \) mJy, \( z \)) the curve which is more sensitive to the evolution of the cluster population. Following Diego et al. (2001a) we will combine both curves to reduce the degeneracy. Some models, which are compatible with the first curve, will be incompatible with the second one and vice versa. Thus, these models will be rejected.

Since this kind of data are not available yet, we will check the method with two simulated data sets following the characteristics of the Planck satellite (Section 3) for \( N(S) \). For the second curve we will assume that a randomly selected subsample of 300 clusters from the whole Planck catalogue has been observed optically and that we know the redshift for each cluster in this subsample (see Section 4). The input model used to simulate both data sets was the \( \Lambda \)CDM model in Table 1. That model is compatible with the mass function given in Bahcall & Cen (1993), the temperature function of Henry & Arnaud (1991), the luminosity function of Ebeling et al. (1997), and the flux function of Rosati et al. (1998) and De Grandi et al. (1999), as was shown in Diego et al. (2001a).

We have compared both simulated data sets with \( \sim 2 \times 10^6 \) different flat \( \Lambda \)CDM models where the six free parameters of our model have been changed on a regular grid. The first three parameters correspond to the cosmological ones (\( \sigma_8 \), \( \Omega \) and \( \Gamma \)) which control the cluster population in the Press–Schechter formalism. The other three parameters correspond to the \( T–M \) relation (equation 8), whose free parameters will enter in the fitting procedure at the same level as the cosmological ones. This relation is needed to build the total flux from the mass of the cluster (equation 5). By considering the \( T–M \) as a free parameter relation, we will check the influence of the uncertainty in the scaling relation on the determination of the cosmological parameters.

In Fig. 8 we show the results of our fit. We have marginalized the probability over each one of the six free parameters. The probability was defined as in Diego et al. (2001a) using the Bayesian estimator given in Lahav et al. (2000).

\[
-2 \ln P_L = \chi_L^2,
\]

where,

\[
\chi_L^2 = \sum N_i \ln (\chi_i^2).
\]

\( \chi_i^2 \) is the ordinary \( \chi^2 \) for each data set and \( N_i \) represents the number of data points for the data set \( i \). Based on a Bayesian approach with the choice of non-informative uniform priors on the log, those authors have seen that this estimator is appropriate for the case
when different data sets are combined together, as in our case. The factor \( N_i \) plays the role of a weighting factor. Larger data sets are considered more reliable for the parameter determination.

As shown in that figure, the estimate of the cosmological parameters is unbiased (compare with the input model, black dots). They are also very well constrained with a small degeneracy between the parameters. This result shows how with future SZE data it will be possible to discriminate among several scenarios of cluster formation. The situation is different in the parameters of the \( T - M \) relation. In this case we do not obtain any spectacular result. Only the \( \psi \) parameter has been well located. There is a degeneracy between the amplitude \( T_0 \) and the exponent \( \alpha \) which will be discussed in the next section.

In Figs 9 and 10 we present the simulated data sets and the two indistinguishable models given in Table 1. From the first figure it is evident that both models would remain indistinguishable if only that data set is used in the fit, but in the second figure we can see that it is possible to distinguish the two models at a high significant level owing to their different evolution in redshift.

### 6 Discussion and Conclusions

In previous sections we have seen that the SZE will be a very powerful tool to study the cluster population at different redshifts. Up to now, no cluster has been detected above \( z < 2.0 \). Previous X-ray surveys have been limited in redshift and current experiments (Chandra, Newton–XMM) are not expected to detect clusters at higher \( z \). Only through the SZE do we have the possibility to observe clusters above that redshift (or maybe to conclude that no cluster has been formed above that redshift).

These high-redshift clusters are fundamental for understanding the physics of cluster formation and also for establishing the evolution of the cluster scalings such as \( T - M \) or \( L_x - M \).

We have seen that Planck will be able to detect distant clusters, which will provide very useful information concerning the cluster population and the underlying cosmology.

However, we have seen that with only the \( dN(S)/dS \) curve, it will be difficult to discriminate among models which were previously indistinguishable. To distinguish them, we need redshift information. We have seen that for the whole SZE sky catalogue, only a relatively small number (\( \approx 300 \)) of optically observed clusters randomly selected from the whole Planck catalogue is needed in order to discriminate between the \( \Lambda \)CDM and OCDM models just by looking at the different cluster population as a function of redshift. If we want to discriminate among the \( \Lambda \)CDM models, we have shown that by combining the statistically large data set

---

**Figure 8.** Marginalized probability in each parameter. The black dot represents the input model (\( \Lambda \)CDM model in Table 1).

**Figure 9.** \( dN/dS \) curve for the \( \Lambda \)CDM (solid) and OCDM (dotted) models in Table 1. Data points represent a Monte Carlo realization of mock data for the \( \Lambda \)CDM model. This curve includes all the Planck detected clusters (2/3 of the sky).

**Figure 10.** \( dN(S > 30 \text{ mJy}, z)/dz \) curve for the \( \Lambda \)CDM (solid) and OCDM (dotted) models in Table 1. As in Fig. 9 the data were simulated assuming a \( \Lambda \)CDM model. The simulation was over a solid angle \( d\Omega = 0.6675 \) per cent of the sky where a total of \( \approx 300 \) clusters were found with fluxes greater than 30 mJy.
\( \frac{dN}{dS} \) with the cosmological sensitivity of \( dN(> 30 \text{ mJy}, z)/dz \) it is possible to reduce significantly the degeneracy in the cosmological parameters as can be seen in Fig. 8.

One important conclusion is that this result is almost independent of the assumed \( T-M \) relation. In fact, our method is practically insensitive to the \( T_0 \) amplitude and the \( \alpha \) exponent in equation (8). We have marginalized the probability by assuming different fixed values for \( T_0 \) and \( \alpha \). The resulting marginalized probabilities were very similar in all the cases considered, showing the small dependence on the assumed values of \( T_0 \) and \( \alpha \). The almost null dependence on \( \alpha \) can be easily understood by looking at equation (5). In the computation of the temperature function (see equation 7) the derivative \( \frac{dM}{dT} \) is inversely proportional to \( aM^{a-1} \). The X-ray-derived functions (such as the temperature function) are sensitive to the \( \alpha \) exponent through the previous derivative. In contrast, the flux function, \( \frac{d^2N}{dSdz} \) is inversely proportional to the derivative \( \frac{dS}{dT} \) of equation 7. Therefore, a change of 0.1 units in \( \alpha \), represents a change in the \( \frac{dM}{dT} \) derivative of 14 per cent, while in the flux function the same change in \( \alpha \) implies a variation of only 6 per cent in the derivative \( \frac{dS}{dT} \), both percentages assume \( M = 1 \times 10^{15} h^{-1} \text{ M}_\odot \). This explains why the SZ flux function is less sensitive to \( T_0 \) and \( \alpha \) than the X-ray derived functions. The uncertainty in \( T_0 \) is a bit more difficult to understand. From equation (5) the total flux is directly proportional to \( T_0 \) and therefore we should expect some dependence of our fit on this parameter. However, if a change in \( T_0 \) is compensated by a change in \( \alpha \) then we would have a degeneracy on these two parameters \( [S \propto T_0 M^{1+\alpha} (1+z)^{\phi}/D_L(z)^2] \). In fact, from Fig. 8 we can see that those models with a low value for \( T_0 \) and a high value of \( \alpha \) are slightly favoured, indicating the fact that there is some kind of compensation between these two parameters.

In order to break the degeneracy in \( T_0-\alpha \) we should include in our fit information concerning the mass of the clusters just to make the fit sensitive to an independent change in \( T_0 \) and/or \( \alpha \) not only to a change in the quantity \( T_0 M^{1+\alpha} \). The previous situation was considered in Diego et al. (2001a), where we included in the fit the cluster mass function. In that case we found, in fact, that there was no degeneracy in these parameters.

The third parameter of the \( T-M \) relation (\( \psi \)) seems, however, to be very relevant for our fit. This is not surprising as we are using one data set, which is expressed as a function of redshift (Fig. 10). While both, \( T_0 \) and \( \alpha \) can be mutually compensated, the effect of changing \( \psi \) on the simulated data sets (Figs 9 and 10) can only be compensated with a change in some of the cosmological parameters [through their effect on the cluster population and in \( D_L(z) \)], but as the allowed range of variation of the cosmological parameters is small (see Fig. 8), the confidence interval for \( \psi \) will also be small.

In this work, the \( T-M \) relation was considered as a free relation just for consistency with our previous work. When fitting SZ data, we have shown that the choice of one specific value for \( T_0 \) and \( \alpha \) in the \( T-M \) relation is not very relevant, although it is important to include in the fit the possible dependence of this relation on \( z \).

The situation was different in Diego et al. (2001a) where the redshift dependence was not relevant (since most of the data was at low redshift), but the choice of \( T_0 \) and \( \alpha \) was important to obtain a good fit to the X-ray and optical data considered in this work. The specific form of the \( T-M \) relation will be more important in the case of fitting future X-ray data. \textit{Newton–XMM} will provide very relevant information, especially at low and medium redshift, concerning the cluster population and the scaling relations \( T-M \) and \( L_x-T \). However, we have seen that the expected number of detected clusters and the redshift coverage will be smaller for this mission compared with \textit{Planck} and therefore \textit{Planck} will provide key information that will allow understanding of cluster formation and evolution. For instance, as we have already seen, the information concerning the \( T-M \) relation can be complemented with studies of the SZE on clusters. Meanwhile, \( T_0 \) and \( \alpha \) can be determined through the study of low-redshift X-ray data, \( \psi \) could be constrained with the high-redshift SZE data. The best results will come, therefore, from the combination of data from X-ray and millimetre missions (see, e.g., Haiman, Mohr & Holder 2001).

With \textit{Newton–XMM} and its better selection function at low \( z \) we can obtain a good sample of the cluster population at low to intermediate redshift with their corresponding temperatures and X-ray fluxes (also detecting the low-mass population) and with \textit{Planck} we will explore the cluster population further in redshift.

A very interesting possibility has been analysed by Xue & Wu (2001). The authors suggested the use of the X-ray luminosity function as a starting point to derive the millimetre (SZE) flux function. In the process, several assumptions concerning the \( L_x-T \) and \( T-M \) relation is to be made. These assumptions could be tested with future SZE data, opening up the possibility of studying these relations at redshifts where no clusters can be observed in the X-ray band.

**ACKNOWLEDGMENTS**

This work has been supported by the Spanish DGESIC Project PB98-0531-C02-01, FEDER Project 1FD97-1769-C04-01, the EU project INTAS-OPEN-97-1192 and the RTN of the EU project HPRN-CT-2000-00124. JMD acknowledges support from a Spanish MEC fellowship FP96-20194004 and financial support from the University of Cantabria.

**REFERENCES**

Abell G. O., 1958, ApJS, 3, 211
Abell G. O., Corwin H. G., Olowin R. P., 1989, ApJS, 70, 1
Aghanim N., de Luca A., Bouchet F. R., Gispert R., Puget J. L., 1997, A&A, 325, 9
Allen S. W., 1998, MNRAS, 296, 392
Bahcall N. A., Cen R., 1993, ApJ, 407, L49
Bahcall N. A., Fan X., 1998, ApJ, 504, 1
Barbosa D., Bartlett G., Blanchard A., Oukbir J., 1996, A&A, 314, 13
Barlett J. G., Silk J., 1994, ApJ, 423, 12
Benitez N., Broadhurst T., Rosati P., Courbin F., Squires G., Lidman C., Magain P., 1999, ApJ, 527, 31
Benitez N., 2000, ApJ, 536, 571
Blanchard A., Sadat R., Bartlett J. G., Le Dour M., 2000, A&A, 362, 809
Bode P., Bahcall N. A., Ford E. B., Ostriker J. P., 2001, ApJ, 551, 15
Borgani S., Rosati P., Tozzi P., Norman C., 1999, ApJ, 517, 40
Brunner R. J., Szalay A. S., Connolly A. J. J., 2000, ApJ, 541, 527
Carlberg R. G., Morris S. L., Yee H. K. C., Ellingson E., 1997a, ApJ, 479, L19
Carlberg R. G. et al., 1997b, in Hamilton D., ed., Ringberg Workshop on Large-Scale Structure, Kluwer, Dordrecht, in press (astro-ph/9704060)
Cohen J. G., Hogg D. W., Blandford R., Cowie L. L., Hu E., Songaila A., Shopbell P., Richberg K., 2000, ApJ, 538, 29
Coleman G. D., Wu C.-C., Weedman D. W., 1980, ApJS, 43, 393
Dahle H., 2000, in Bergvall N., Takolo L. O., Piirola V., eds, The NOT in the 2000’s. Univ. Turku, Tuorla Observatory, Finland
De Grandi S. et al., 1999, ApJ, 514, 148
Diego J. M., Martínez-González E., Sanz J. L., Cayón L., Silk J., 2001a, MNRAS, 325, 1533
Diego J. M., Vieyla P., Martínez-González E., Silk J., Sanz J. L., 2001b, MNRAS, submitted
Donahue M., Voit G. M., 1999, ApJ, 523, L137

© 2002 RAS, MNRAS 331, 556–568

Downloaded from https://academic.oup.com/mnras/article-abstract/331/3/556/1021235 by guest on 29 July 2018
Ebeling H., Edge A. C., Fabian A. C., Allen S. W., Crawford C. S., 1997, ApJ, 479, L101
Eke V. R., Cole S., Frenk C. S., Henry J. P., 1998, MNRAS, 298, 114
Fan Z., Chiuhe T., 2001, ApJ, 550, 547
Fixsen D. J., Cheng E. S., Gales J. M., Mather J. C., Shafer R. A., Wright E. L., 1996, ApJ, 473, 576
Gladders M. D., Yee H. K. C., 2000, AJ, 120, 2148
Girardi M., Borgani S., Giuricin G., Mardirossian F., Mezzetti M., 1998, ApJ, 506, 45
Grego L., Carlstrom J. E., Reese E. D., Holder G. P., Holzapfel W. L., Joy M. K., Mohr J. J., Patel S., 2001, ApJ, 552, 2
Haiman Z., Mohr J., Holder G. P., 2001, ApJ, 553, 545
Henry J. P., Arnaud K. A., 1991, ApJ, 372, 410
Henry J. P., 2000, ApJ, 534, 565
Hobson M. P., Jones A. W., Lasenby A. W., Bouchet F. R., 1998, MNRAS, 300, 1
Hobson M. P., Barreiro R. B., Toffolatti L., Lasenby A. W., Sanz J. L., Jones A. W., Bouchet F. R., 1999, MNRAS, 306, 232
Holder G. P., Carlstrom J. E., 1999, in de Oliveira-Costa A., Tegmark M., Hobson M. P., Barreiro R. B., Toffolatti L., Lasenby A. W., Bouchet F. R., 1998, MNRAS, 301, 562
Holder G. P., Mohr J. J., Carlstrom J. E., Evrard A. E., Leitch E. M., 2000, ApJ, 544, 629
Jones C., Forman W., 1984, ApJ, 276, 38
Kinney A. L., Calzetti D., Bohlin R. C., McQuade K., Storchi-Bergmann T., Schmitt H. R., 1996, ApJ, 467, 38
Kitayama T., Suto Y., 1997, ApJ, 490, 557
Kneissl R., Jones M. E., Saunders R., Eke V. R., Lasenby A. N., Grainge K., Garret C., 2001, MNRAS, 328, 783
Lahav O., Bridle S. L., Hobson M. P., Lasenby A. N., Dodr’e L., Jr, 2000, MNRAS, 315, 45
Lo K. H., Chiuhe T. H., Martin R. N., Kin-Wang Ng, Liang H., Pen Ue-Li, Ma Chung-Pei, 2000, AAS, 197, 56.03
F. F., 1998, MNRAS, 295, 769
Lumsden S. L., Nichol R. C., Collins C. A., Guzzo L., 1992, MNRAS, 258, 77
Lubin L. M., Gunn J. E., Oke J. B., Hoessel J. G., Schneider D. P., Christensen J. A., 1996, ApJ, 111, 615
Press W. H., 1974, ApJ, 187, 425 (PS)
Rahman N., Shandrin S. F., 2001, ApJ, 555, L1
Mathiesen B., Evrard A. E., 1998, MNRAS, 295, 769
Nichol R. et al., 2000, in Banday A. J., Zaroubi S., Bartelmann M., eds, Proc. MPA/ESO/MPE Workshop, Mining the Sky. Springer-Verlag, Berlin, in press
Postman M., Lubin L. M., Gunn E. J., Oke J. B., Hoessel J. G., Schneider D. P., Christensen J. A., 1996, ApJ, 111, 615
Postman M., Lubin L. M., Gunn E. J., Oke J. B., Hoessel J. G., Schneider D. P., Christensen J. A., 1996, ApJ, 111, 615
Press W. H., Schechter P., 1974, ApJ, 187, 425 (PS)
Rahman N., Shandrin S. F., 2001, ApJ, 550, 121
Romer A. K., Viana P. T. P., Liddle A. R., Mann R. G., 2001, ApJ, 547, 594
Rosati P., Della Ceca R., Norman C., Giacconi R., 1998, ApJ, 492, L21
Rosati P., Stanford S. A., Eisenhardt P. R., Elston R., Spinrad H., Stern D., Dey A., 1999, AJ, 118, 76
Silk J., White S. D. M., 1978, ApJ, 226, L103
Smail I., Dressler A., Couch W. J., Ellis R. S., Oemler A., Jr, Butcher H., Sharples R. M., 1997, ApJS, 110, 213
Sunyaev R. A., Zel’dovich Ya. B., 1972, A&A, 20, 189
Udomprasert P. S., Mason B. S., Readhead A. C. S., 2000, in Durret F., Gerbal D., eds, IAP, 2000 Meeting, Constructing the Universe with Clusters of Galaxies. Available from http://www.iap.fr/Conferences/Colloque/collo2000/contributions, Reference 4.4.6
Viana P. T. P., Liddle A. R., 1998, MNRAS, 303, 535
Williams et al., 1996, AJ, 112, 1335
Williams et al., 2000, AJ, 120, 2735
Wilson G., Smail I., Ellis R. S., Couch W. J., 1997, MNRAS, 284, 915
Wu X.-P., Fang L.-Z., 1997, ApJ, 483, 62
Xue Y.-J., Wu X.-P., 2001, ApJ, 552, 452

APPENDIX A: PLANCK FLUX LIMIT

In this appendix we want to justify that Planck will be able to detect clusters with an integrated total flux above $\approx 30$ mJy.

Obviously, the number of detected clusters will depend on the technique used to detect them. We will focus our attention on the maximum-entropy method (MEM) (Hobson et al. 1998, 1999), where the authors have shown that with such a method they obtain a good recovery of the thermal SZE. In Hobson et al. (1999) it is shown that the rms residuals for a $4.5$-arcmin FWHM Gaussian beam for the MEM reconstruction is $\approx 6$ $\mu$K pixel$^{-1}$ when no power spectrum information is assumed and the point sources are included in the analysis.

Now we compute the flux (in mJy) corresponding to that rms temperature and therefore it should be considered as a flux per pixel.

The flux is defined as the integral of the specific intensity on the solid angle. When we compute the total flux of the cluster the solid angle is that subtended by the cluster (see equation 3 in Section 2). In equation (3), $\Delta T/\Delta T$ can be related to $\Delta T/T$ as $\Delta T/\Delta T = I_0(h\times(\Delta T/T))$, where $I_0$ is a constant given below, $h(\tau)$ is the spectral shape factor, $\Delta T = 6.0$ $\mu$K is the temperature we want to transform into a flux and $T = 2.73$ K.

$$I_0 = \frac{2 \left(k_B T_{\text{CMB}}\right)^3}{(h c)^2} = 2.7 \times 10^{14} \text{ mJy str}^{-1}$$

and the spectral dependence $h(\tau)$ is given by

$$h(\tau) = \frac{x^4 e^x}{(e^x - 1)^2}$$

where $x = h\tau(k_B T_{\text{CMB}}) \approx \tau(GHz)/56.8$ is the a dimensional frequency. Assembling terms in equation (3) we obtain the flux per pixel $[d\Omega = (1.5 \text{ arcmin}^2) = 1.9 \times 10^{-7} \text{sr}]$

$$N_{\text{rms}}(\nu) = 2.7 \times 10^{11} \frac{\Delta T}{T} h(\tau) d\Omega \text{ mJy str}^{-1}$$

$$= 0.45 \text{ mJy pix}^{-1} (300 \text{ GHz}).$$

This number has been calculated for the frequency $\nu = 300$ GHz. In the paper we presented our calculations for $353$ GHz. The flux at this frequency is a factor $\approx 1.3$ times higher than at $300$ GHz. Therefore,

$$N_{\text{rms}} = 0.58 \text{ mJy pix}^{-1} (353 \text{ GHz}).$$

This flux should be considered as the rms in the residual map when the SZE is recovered by MEM, that is the noise per pixel. We will consider a conservative limit for the detection of a cluster of signal-to-noise ratio $\approx 7.5\sigma$ on the FWHM. If we consider our antenna as a Gaussian beam and the cluster is not resolved (this will happen in most of the cases in Planck) then we can consider that the cluster profile follows a Gaussian pattern:

$$\text{signal} = S = A e^{-r^2/2\sigma^2}$$

Inside the FWHM the signal is just the integral of the previous equation from $r = 0$ to FWHM/2 which can be solved easily (by
using the change of variables $x = r^2/2\sigma^2$:

$$S(\text{FWHM}) = A \int_0^{\text{FWHM}/2} 2\pi r \, dr \, e^{-r^2/2\sigma^2} = 6.3A. \quad (A6)$$

Inside the FWHM the noise associated with MEM goes as

$$N(\text{FWHM}) = N_{\text{rms}} \sqrt{N_{\text{pix}}^{\text{FWHM}}}, \quad (A7)$$

where $N_{\text{rms}}$ is the noise per pixel found previously and $N_{\text{pix}}^{\text{FWHM}}$ is the number of pixels corresponding to the area enclosed by the FWHM, $N_{\text{pix}}^{\text{FWHM}} = \text{FWHM}^2 = 11.1$ (FWHM = 5 arcmin for the channel at 353 GHz).

Now if we require $S(\text{FWHM})/N(\text{FWHM}) = 7.5$ then by dividing equations (A6) by (A7) we find that $A$ must be $\approx 4N_{\text{rms}}$, that is, an unresolved cluster with a signal following equation (A5) should have an amplitude $A \approx 4N_{\text{rms}}$ in order to have $S/N = 7.5$ inside the FWHM.

The total flux inside the full antenna beam of such a signal is $S_{\text{Total}} = 12.6 \times A \approx 50 \times N_{\text{rms}}$. If now we substitute $N_{\text{rms}} = 0.58 \text{ mJy pixel}^{-1}$, then we finally obtain, $S_{\text{Total}} \approx 29 \text{ mJy}$ which approximates the value of 30 mJy used in this paper.

This paper has been typeset from a \TeX/\LaTeX file prepared by the author.