Optimizing the yield of Sunyaev-Zel’dovich cluster surveys

Richard A. Battye\textsuperscript{1⋆}, Jochen Weller\textsuperscript{2,3†}

\textsuperscript{1}Jodrell Bank Observatory, School of Physics & Astronomy, University of Manchester, Macclesfield, Cheshire SK11 9DL, UK.
\textsuperscript{2}NASA/Fermilab Astrophysics Group, Fermi National Accelerator Laboratory, Batavia IL 60510-0500, USA.
\textsuperscript{3}Department of Physics and Astronomy, University College London, London WC1E 6BT, UK.

Received **insert**; Accepted **insert**

ABSTRACT

We consider the optimum depth of a cluster survey selected using the Sunyaev-Zel’dovich effect. By using simple models for the evolution of the cluster mass function and detailed modeling for a variety of observational techniques, we show that the optimum survey yield is achieved when the average size of the clusters selected is close to the size of the telescope beam. For a total power measurement, we compute the optimum noise threshold per beam as a function of the beam size and then discuss how our results can be used in more general situations. As a by-product we gain some insight into what is the most advantageous instrumental set-up. In the case of beam switching observations one is not severely limited if one manages to set the noise threshold close to the point which corresponds to the optimum yield. By defining a particular reference configuration, we show how our results can be applied to interferometer observations. Considering a variety of alternative scenarios, we discuss how robust our conclusions are to modifications in the cluster model and cosmological parameters. The precise optimum is particularly sensitive to the amplitude of fluctuations and the profile of the gas in the cluster.

Key words: cosmic microwave background, Sunyaev-Zel’dovich effect, galaxy clusters

1 INTRODUCTION

The Sunyaev-Zel’dovich (SZ) effect is due to the re-scattering of cosmic microwave background (CMB) photons by the hot electrons situated in intracluster media (Sunyaev & Zel’dovich 1972). It has been realized that it should provide a reliable mechanism for selecting a sample of clusters (Haiman, Mohr & Holder 2001; Holder, Haiman & Mohr 2001) and a number of instruments sensitive to the microwave frequency range (15GHz-350GHz) are under construction to perform this task (Kneissl et al. 2001; Holder et al. 2000; Lo et al. 2002; Mauskopf et al. 2000; Browne et al. 2000; Mohr et al. 2002; Kosowsky 2003). Since the cluster mass function is sensitive to cosmological parameters (Thomas & Carlberg 1989; Oukbir & Blanchard 1992; Scaramella, Cen & Ostriker 1993; Barbosa et al. 1996; Viana & Liddle 1996; Eke, Cole & Frenk 1996) such surveys are likely to be an important part of the program to probe the late-time history of the universe, searching for the origin of the apparently accelerating expansion of the Universe.

The number of clusters as a function of redshift and flux is a strong function of $\Omega_m$, the mass density of the universe in units of the critical density, and $\sigma_8$, the normalization of the matter power spectrum on scales of $8h^{-1}$Mpc, and is a weaker function of $w_0$, the equation of state of the dark energy, $Q$, and its evolution $w_1$. Here, $w_0$ and $w_1$ are defined by $P_Q/\rho_Q = w = w_0 + w_1 z$. It has been shown that the degeneracies within the space of cosmological parameters are such that cluster surveys are complementary to those which might be obtained using observations of Type Ia Supernovae at high redshift (Weller, Battye & Kneissl 2002; Battye & Weller 2003).

Since the distribution of the clusters obeys Poisson statistics the errorbars deduced on the cosmological parameters will be very sensitive to the number of clusters found in a given survey. Therefore, we anticipate that finding the largest number of clusters possible for a given integration time will be close to giving the best constraints on cosmological parameters, although this may also weakly depend on their distribution with redshift. Clearly a survey which will find $\sim 1000$ clusters will constrain cosmological parameters much better than one which finds $\sim 100$. Having a much larger sample would also seem desirable in terms of reducing the systematic errors due to the poorly understood gas...
physics. Fortunately, we will show that, by assuming we can model the gas physics well, one can get to within a factor of two of the optimal situation for a total power measurement. This is more tricky in more complicated experimental situations such as beam switching whereby an ill-judged selection of the survey depth can lead to one finding very few clusters at all.

The goal of this paper, therefore, is to discuss issues related to optimizing the yield of a given survey; the optimum being defined as the survey in which the largest number of objects are found in a given amount of observing time. This is a question one is always faced with in any kind of astronomical measurement: how long does one integrate for on a given patch of sky? Obviously the shorter the required integration time for individual patches, the more sky area can be covered. However, if one doesn’t integrate long enough only the very brightest objects will be found. Clearly, one is always guided in such situations by one’s expectations of what is there to be observed.

For the present case of clusters, one can attempt to make a simple estimate by assuming (1) the clusters are point-like objects in the telescope beam, (2) their comoving number density is constant and (3) the Universe is Euclidean. In this case the number of objects with flux density, $S$, greater than the limiting flux density of a survey $S_{\text{lim}}$ is $N(S > S_{\text{lim}}) \propto \Delta \Omega / S_{\text{lim}}^{3/2}$, where $\Delta \Omega$ is the angular coverage.

Since for fixed total integration time $S_{\text{lim}} \propto \sqrt{\Delta \Omega}$, one can deduce that $N \propto \Delta \Omega^{1/4} \propto S_{\text{lim}}^{1/2}$, and hence a survey with largest possible area (that is, the whole sky) would be optimal in this case. Under these assumptions the survey would be flux density limited and hence an instrument with the largest possible collecting area would be optimal in terms of making the largest number of detections in the shortest period of time.

Unfortunately, in most cases clusters are unlikely to be point-like, nor is their comoving density likely to be constant due to evolution and, of course, the Universe is not Euclidean. As we will describe in section 2, the distribution of clusters as a function of mass and redshift can be modelled in a sensible way for a given set of cosmological parameters. However, the key to understanding the issue at hand here is how the limiting mass of a survey is related to the instrumental noise and the angular resolution of the telescope.

Shallow surveys, as suggested by the simple estimate above, will only find clusters which are very large and close by, and such clusters cover many beam areas. Since the same signal is spread over a number of beam areas, such objects can dip below the detection threshold even though they are very large. It is clear that this effect will compete with that due to Euclidean source counts in our initial estimate, leading to an optimum at a particular value of the noise and angular coverage. We will show that this optimum will depend on the size of the beam: larger beam sizes, say $8'$, will require a deeper survey to achieve the optimum, whereas a much shallower survey will be optimal for much smaller beam sizes, say $1'$. Essentially the optimal situation is for most of the clusters in the survey to be the same size as the beam as explained in Fig. 1.

Once the beam is sufficiently small to resolve all the clusters, the survey becomes surface brightness temperature limited, rather than being flux density limited, and increasing the collecting area no longer acts to increase the rate of detection. This illustrates an interesting feature of cluster survey, which is in contrast to intuition one might have from considering standard point sources and, as we shall see, leads to some other somewhat counter-intuitive conclusions.

In section 3 we will discuss how to derive a relation for the limiting mass of the survey as a function of redshift and the limiting flux, incorporating the effects of a Gaussian beam profile. Our initial motivation is to model idealized total power measurements using a single dish with a circular aperture. We will allow the observations to be smoothed using a window function which is exactly tuned to every possible cluster size above the maximum resolution of the telescope. Under these assumptions we show in section 4 that for a particular mimimum beam size there exists a unique value of the noise which corresponds to the optimum in terms of the total of clusters detected, and for all beam sizes less than around $2'$ (FWHM) this is equivalent in terms of the number of clusters found in the survey, although the precise optimum noise will be different for different beam sizes.

Accurate total power measurements are only feasible in space or at a very dry site (for example, the South Pole) on the earth’s surface, mainly due to the effects of water vapour. In section 5 we discuss how our calculations can be adapted to situations closer to those experienced in reality.

---

\*\* We define deep to mean a low noise level in terms of temperature.
In particular, we will illustrate how our calculations can be applied to techniques which minimize the effects of the atmospheric contamination which is ubiquitous at the relevant frequencies at most sites. These techniques involve either use of a beam switch or an interferometric array. For the most part our conclusions are similar to those for a total power experiment, although more care is required in interpreting them.

Obviously our goal has some clear dangers since one is, by necessity, required to assume a specific distribution of clusters. In situations where the source population is completely unknown, attempts at optimizing observations in this way are completely useless! However, the distribution and precise characteristics of clusters are becoming much better understood; something which is only likely to improve as observations progress. We have, therefore, used the so-called concordance cosmological parameters (Ostriker & Steinhardt 1995) and have attempted to model most of the relevant features of the distribution of clusters. In section 6 we test the robustness of our predictions to variations in the cosmological parameters and in the cluster model. Small variations do not affect the optimum significantly but, not surprisingly, large variations substantially change our predictions. We quantify this in terms of the parameters. Of particular concern is the uncertainty in the detailed modelling of the poorly understood gas physics within the cluster. It is clear that more sophisticated modelling of this physics would be required before we can have full confidence in our predicted optima; they should, nonetheless, provide a useful guide.

For the purposes of this paper, we will ignore the effects of confusion due to radio and infra-red (IR) point sources, the CMB anisotropies and that due the distribution of clusters themselves which are below the flux limit. These will be discussed in a future publication. Suffice to say that the limiting flux densities and temperatures that we suggest here are compatible with confusion being negligible except for the most extreme cases. Very much deeper surveys run the risk of becoming confusion limited. We note that some of our conclusions concerning the most advantageous instrumental setup do not take into account possible strategies to mitigate problems which might occur due to confusion backgrounds.

We should note that our optimization are only of relevance to ground-based experiments specifically designed to perform high resolution CMB measurements and SZ surveys at a single frequency. For general purpose, multi-frequency instruments such as PLANCK, the observational strategy (in the case of PLANCK, an all sky survey) has been designed to take into account a wide range of possible applications, not just the ability to detect the SZ in clusters.

Throughout this paper we will work in natural units where \( c = h = k_B = 1 \). Standard conversion factors can be found, for example, in Kolb & Turner (1989).

2 MODELLING THE CLUSTER DISTRIBUTION

Except in section 6 where we test the robustness of our predictions, we will use a fiducial cosmological model which is a flat \( \Lambda \)CDM model with \( \left( \Omega_m, \sigma_8, h, \omega_0, \omega_1 \right) = (0.3, 0.9, 0.72, -1, 0) \). This model is compatible with recent CMB (Spergel et al. 2003), Large-Scale Structure (LSS) (Percival et al. 2001; Percival et al. 2002), Hubble constant (Freedman et al. 2001) and weak lensing measurements (Contaldi, Hoekstra & Lewis 2003). The precise value of \( \sigma_8 \) is currently under close scrutiny and is very important in making predictions of cluster observations since the number of objects found is extremely sensitive to its value. We should note that our chosen value is larger than some other values quoted in the literature (Allen, Schmidt & Fabian 2002) and we will return to this issue in section 6. Where necessary we use a baryon density which satisfies current estimates (Spergel et al. 2003) and a spectral index of initial density fluctuations \( n_s = 1 \). Only very extreme changes from these two parameters can lead to significantly different results.

We will use the mass function deduced from the suite of N-body simulations (Evrard et al. 2002) performed by the VIRGO collaboration. The comoving number density of objects at redshift \( z \) in the mass range \( M \rightarrow M + dM \) is given by

\[
\frac{dn}{dM}(z, M) = -0.22 \frac{\rho_{m}(t_0)}{M \sigma_M} \frac{d\sigma_M}{dM} \exp \left\{ -|A(z, M)|^{3.86} \right\}, \tag{1}
\]

where \( A(z, M) = 0.73 - \log[D(z)\sigma_M], \rho_{m}(t_0) \) is the matter density at the present day, \( \sigma_M \) is the over density in a virialized region containing mass \( M \) and \( D(z) \) is the normalized growth factor which is dependent on cosmological parameters. Here, \( M \) is defined to be \( M_{200} \), that is inside a region with an overdensity of 200, see Battye & Weller (2003) for more details on how to apply this to cluster calculations.

The clusters are assumed to be spherical and in hydrodynamical equilibrium, in which case the virial mass, \( M_{vir} \), and virial radius, \( R_{vir} \), are related by

\[
M_{vir} = \frac{4\pi}{3} \rho_{crit}(z) \Delta_c(z) R_{vir}^3, \tag{2}
\]

where \( \rho_{crit}(z) \) is the critical density at redshift \( z \) and \( \Delta_c(z) \) is the relevant overdensity which can be computed from the spherical collapse model (Partridge & Peebles 1967; Gunn & Gott 1972; Lahav et al. 1991). This can be adapted to apply to models where \( w \neq -1 \) (Wang & Steinhardt 1998; Battye & Weller 2003).

We will assume that the electron density weighted temperature \( \left( T_e \right)_{\rho} \) can be related to \( M_{vir} \) under the assumption that the cluster is isothermal via

\[
\frac{M_{vir}}{10^{15} h^{-1} M_{\odot}} = \left( \frac{T_e}{T_0 \text{ keV}} \right)^{3/2} \left[ \frac{\Delta_c(z) E(z)}{\Omega_c(z)} \right]^{1-1/2} \times \left[ 1 + (1 + 3w) \Omega_c(z) \Delta_c(z) \right]^{-3/2}, \tag{3}
\]

where the Hubble parameter \( H(z) = H_0 E(z) \) and \( \Omega_c(z) \) is the density of the dark energy component relative to the critical density at redshift \( z \). We will use \( T_0 = 1.6 \) as the normalization factor for most of our discussion, but will consider the effects of other values of \( T_0 \) in section 6.

The profile of the hot gas \( \zeta(\theta) \), whose mass is related to the total mass of the cluster by \( M_{gas} = f_{gas} M_{vir} \), is modelled using the isothermal \( \beta \)-model (Cavaliere & Fusco-Femiano 1976) with \( \beta = 2/3 \) as suggested by observations,
where $\theta_{\text{vis}}$ is the apparent angular size of the cluster core, $\alpha = R_{\text{vir}}/d$ and $\theta_{\text{vis}}$ is related to $R_{\text{vir}}$ via the angular diameter distance $d_A(z)$. We will assume for simplicity that $d_A(z)$ is universal and redshift independent with value $10^{20}$ Mpc. The corresponding beam area is given by

$$A = \Delta \Omega \frac{dV}{dzd\Omega} \frac{dM}{dn}.$$

where $\Delta \Omega$ is the angular coverage and $dV/(dzd\Omega)$ is the FRW comoving volume element. $M(z)$ is the limiting mass of the survey which one can relate to the limiting flux density $S_{\text{lim}}$ as discussed in the section 3. We will use $z_{\text{max}} = 1.5$ since we are interested in clusters whose redshifts can be found using standard techniques. At present it is difficult to accurately deduce cluster redshifts which are $\sim 1$. We optimistically assume that, by the time these SZ surveys have been performed, techniques will have been developed to at least deduce photometric redshifts routinely out to $z = 1.5$. There are expected to be very few clusters with $z > 1.5$. In principle our calculations could easily be modified to optimize the number of clusters between any two redshifts as desired.

3 COMPUTING THE LIMITING MASS

The flux density per beam due to the Sunyaev-Zel’dovich effect is given by

$$S = 2f^2 T_{\text{CMB}} h(x) Y,$$

where $T_{\text{CMB}}$ is the CMB temperature, $Y$ is the Compton $y$-distortion integrated over the beam profile and $h(x)$ is given by

$$h(x) = \frac{x^2 e^x}{(e^x - 1)^2} \left[ \frac{x}{\tanh(x/2)} - 4 \right] = -2 + O(x),$$

where $x = 2\pi f/T_{\text{CMB}} \approx f/56.8$ GHz. $f$ is the frequency of observation and $\lambda$ is the corresponding wavelength. The limiting virial mass of the survey can be related to the flux density per beam (Battye & Weller 2003) by

$$S = \frac{2f^2 T_{\text{CMB}} (T_e)_{\nu} \sigma_T h(x)}{m_e d_A^2} \frac{f_{\text{gas}} M_{\text{vir}}}{\mu_e m_p} I(B, \zeta),$$

where

$$I(B, \zeta) = \int \frac{d\Omega B(\theta) \zeta(\theta)}{d\Omega \zeta(\theta)},$$

$\sigma_T$ is the Thomson scattering cross-section, $\mu_e \approx 1.143$ is the mean molecular weight of the electrons and $m_e, m_p$ are the masses of the electron and the proton, respectively. The corresponding beam area is given by

$$\Omega_b = \int d\Omega B(\theta).$$

The corresponding thermodynamic temperature can be computed using the standard formula

$$T_{\text{THERM}} = \frac{g(x) S}{2f^2 h(x) \Omega_b},$$

and

$$g(x) = \frac{x}{\tanh(x/2)} - 4,$$

is the frequency dependence of the SZ brightness temperature. We note that this quantity is not the peak of the SZ effect due to the cluster, but the average integrated over a region whose size is set by the instrumental beam size. For the purposes of our discussion we will define a non-standard temperature unit which includes this frequency dependence, allowing us, where desirable, to quote frequency independent limits in the subsequent discussion,

$$T = \frac{T_{\text{THERM}}}{g(x)} = \frac{T_{\text{CMB}} (T_e)_{\nu} \sigma_T f_{\text{gas}} M_{\text{vir}}}{\mu_e m_p m_p d_A^2} I(B, \zeta) \Omega_b.$$
In the absence of beam smoothing, much smaller beams give an altogether different shape since they find it more difficult to detect clusters which are very large, that is, those which are very close by. This led Bartlett (2000) to suggest that surveys performed with very high resolution instruments should be very deep, contrary to the argument based on Euclidean source counts we presented earlier. However, one must also attempt to take into account the effects of beam smoothing and be careful to make comparisons with configurations of the same temperature sensitivity, as we have done here, rather than flux density sensitivity which is related to the overall collecting area. As we will see this leads to a conclusion somewhere in between these two extremes, strongly dependent on the resolution of the telescope.

In a diffraction limited observation, the highest resolution beam size possible is $\theta_{\mathrm{FWHM}}^{\min} \propto \lambda/D$ where $D$ is the diameter of the telescope. Coarser resolution is possible by either physically under-illuminating the dish using appropriate feed horns, or by beam smoothing the resultant map. Let us consider the latter possibility of an increased beam size $\theta_{\mathrm{FWHM}} = V N_b\sigma_{\mathrm{FWHM}}$ whose area covers $N_b$ copies of the minimum beam size, that is, $\Omega_b = N_b \Omega_{\theta_{\mathrm{FWHM}}^{\min}}$.

By adding the noise together for the $N_b$ beams, we see that $T_{\text{lim}} = T_{\text{lim}}^{\min}\sqrt{N_b}$ where $T_{\text{lim}}^{\min}$ is the noise temperature corresponding to the minimum beam size, and the noise level in terms of flux density increases to $\sqrt{N_b}$. Hence, the limiting flux density becomes $S_{\text{lim}} = \sqrt{N_b} S_{\text{lim}}^{\min}$ where $S_{\text{lim}}^{\min}$ is the flux density limit equivalent to $T_{\text{lim}}^{\min}$. This would correspond to taking a map with resolution determined by $\theta_{\mathrm{FWHM}}^{\min}$ and convolving it with a Gaussian window function with $\mathrm{FWHM} \sqrt{\lambda/D}$.

A simple, and unrealistic, assumption is that one can smooth the resulting map to the size of any cluster larger than the minimum beam size which might contribute to it, that is, $N_b$ can take any value $\geq 1$ adjusted to the size of the cluster in question. This represents the most one could hope to achieve from a particular observation. The number of beam areas covered by a cluster weighted by the gas profile function is given by

$$N_b = \frac{\int \theta d\theta \zeta(\theta)}{\int \theta d\theta B(\theta)}. \quad (15)$$

This can be used in conjunction with the Newton Raphson method to compute the minimum limiting mass. This is shown in Fig. 3 as well as curves of fixed resolution with $N_b = 1, 4, 16, 256$, which correspond to $\theta_{\mathrm{FWHM}} = (2^k)^{1/2}$ and $N_b = 2^k$ with $k = 0, 1, 2, 3, 4$. We have used $S_{\text{lim}}^{\min} = 500\mu$Jy, $f = 30\mathrm{GHz}$ and $\theta_{\mathrm{FWHM}}^{\min} = 1'$. The minimum limiting mass clearly represents the envelope of the family of curves with $N_b$ taking any value between 1 and $\infty$. This will be the basis for our idealized total power measurements discussed in section 4.
resolution; one which can then be iterated further to give an even lower resolution map. In this case $N_b = 4^k$ for some integer $k$.

We investigate the effects of these two more realistic scenarios on the limiting mass in Figs. 4 and 5. In Fig. 4 we reduce $\theta_{\text{FWHM}}^{\text{max}}$ while still allowing any real value for $N_b$, whereas in Fig. 5 we only allow $N_b = 4^k$, while keeping $\theta_{\text{FWHM}}^{\text{max}} = \infty$. We see that if $\theta_{\text{FWHM}}^{\text{max}} > 4'$, the limiting mass and, hence, the number of objects found in the survey, is not significantly affected by a finite value of $\theta_{\text{FWHM}}^{\text{max}}$. Moreover, it is clear that the discreetness effects introduced by considering only values of $N_b$ which are multiples of 4 are unlikely to have a particularly strong effect on the number of clusters found in the survey since the solid curve in Fig. 5 is very close to that in Fig. 3.

4 IDEALIZED TOTAL POWER MEASUREMENTS

We will consider, initially, a simple single telescope observation characterized by the system temperature, $T_{\text{sys}}$, the receiver bandwidth, $\Delta f$, the number of simultaneous beams, $n_b$, generated by a focal plane array, and the diameter of the telescope $D$. The expected $1 - \sigma$ instantaneous sensitivity is $T_{\text{inst}} \approx T_{\text{sys}} / (\eta \sqrt{\Delta f})$ where $\eta$ is the aperture efficiency of the telescope. In the rest of our discussion we shall use $\theta_{\text{FWHM}}$ to refer to the minimum beam size $\theta_{\text{FWHM}}^{\text{min}} \approx \lambda / D$.

The question which we would like to answer is how this sensitivity can be deployed most effectively. The extremes are to point the telescope at the same piece of sky for the whole period of integration, or to cover the whole sky with a very shallow survey. Of course, we are likely to do neither, but the noise temperature per beam of any survey,

![Figure 4](image1.png)

Figure 4. The solid line is the optimum limiting mass allowing for beam smoothing with any value of $N_b$ for $S_{\text{lim}}^{\text{min}} = 500 \mu$Jy, $f = 30$GHz, $\theta_{\text{FWHM}}^{\text{max}} = \infty$ and $\theta_{\text{FWHM}}^{\text{min}} = 1'$ and is the same as the solid line in Fig. 3. The other curves represent the effects of imposing a finite value of $\theta_{\text{FWHM}}^{\text{max}}$ still allowing for any value of $N_b$ up to this limit. The dotted line has $\theta_{\text{FWHM}}^{\text{max}} = 1'$ which is equivalent to the 1' curve in Fig. 3, the short-dashed line it is 2', and for the long-dashed line it is 4'. It is clear that extending the beam much above 4' will have little effect on the number of clusters found in the survey. Using the fiducial model $N_{\text{tot}}(1.5) = 0.33, 0.57, 0.67, 0.70$ per deg$^2$ for $\theta_{\text{FWHM}}^{\text{max}} = 1'$, 2', 4', and $\infty$ respectively.

![Figure 5](image2.png)

Figure 5. The solid line is the exact envelope of the dashed lines which are the same as the $\theta_{\text{FWHM}} = 1'$, 2', and 4' from Fig. 3 which was computed by allowing $N_b = 4^k$ as described in the text. Notice that it is similar to, but not the same as, the solid line in Fig. 3 which allows any value of $N_b$. It is unlikely that accurate inclusion of these effects of realistic beam smoothing will significantly alter the conclusions of this paper.

![Figure 6](image3.png)

Figure 6. Contours of the predicted number of clusters computed using (18) found per unit $C$ as a function of $\theta_{\text{FWHM}}$ and $T_{\text{lim}}$ per beam in the case of optimal beam smoothing. We have used $z_{\text{max}} = 1.5$. The contours represent 95%, 90%, 80%, 50% and 20% of the maximum number that could be found with particular fixed values of $T_{\text{sys}}, \eta, n_b, t$ and $\Delta f$. Notice that, for $\theta_{\text{FWHM}} < 2'$ there is a degenerate line $T_{\text{lim}} = 4T_{\sigma} \approx 55\mu K/(\theta_{\text{FWHM}}/1')$ along which the survey would have optimal yield.

© RAS, MNRAS 000, 1–16
can be considered as a figure of merit for a single dish instrument if $\theta_{\text{FWHM}} < 2'$. Ignoring possible systematic effects it quantifies the number of objects an instrument will find for a fixed integration time, or conversely the time taken to find a fixed number of objects is $\propto M_{\text{SD}}^{-1}$. We see that once this limit on $\theta_{\text{FWHM}}$ is saturated, the rate of making detection is independent of the overall collecting area, in contrast to the case for point sources. We also see that the survey goes from being flux density limited to surface brightness temperature limited at this saturation point. Given that currently available receivers have values of $T_{\text{sys}}$ and $\Delta f$ which are close physical limits, one sees that the only way to improve the detection rate is by increasing $n_b$. We note that the largest value of $n_b$ is limited by the size of the focal plane of the telescope.

An important consequence of an SZ survey being surface brightness temperature limited is that, for fixed values of $T_{\text{sys}}$ and $\Delta f$, the optimal frequency to observe at would be the lowest possible, that is, $x \approx 0$, where the SZ effect is strongest in terms of brightness temperature. In reality, the best available values for $T_{\text{sys}}$ and $\Delta f$ also are dependent on $f$ (typically the best available instantaneous sensitivity, $T_{\text{sys}}/\sqrt{\Delta f}$, is a very weak function of frequency), and this is also mitigated by, both, the fact that point source confusion is greatest at the lowest frequency, and probably more significantly, that achieving the saturation resolution requires a larger diameter telescope for lower frequencies. Nonetheless, it questions the prejudice that the best frequency at which to design an instrument to perform a blind SZ search is near the maximum in the flux density of the SZ effect at $f \approx 150\,\text{GHz}$.

In Fig. 6 we have plotted (18) as a function of $T_{\lim}$ and $\theta_{\text{FWHM}}$ assuming that one can tune any beam smoothing to the size of any cluster larger than the beam size, that is, $N_b$ can take any value between 1 and $\infty$ (that is, $\theta_{\text{FWHM}}^\text{lim} = \infty$) and using $z_{\text{max}} = 1.5$. The computed values have been normalized so that the maximum value is one, and hence the precise value plotted quantifies the efficiency of a particular observational setup.

We see that there is a degenerate line with $T_{\lim} = 4T_\sigma \approx 55\mu\text{K}/(\theta_{\text{FWHM}}/1')$ for $\theta_{\text{FWHM}}$ less than the saturation resolution of $2'$, which can also be phrased in terms of the integrated $y$-distortion limit $Y_{\lim} \approx 2 \times 10^{-5}/(\theta_{\text{FWHM}}/1')$. Therefore, we have shown that for sufficiently high resolution surveys, it should be possible to perform a survey which is optimal in the sense of finding the largest number of clusters for a given receiver sensitivity. For lower resolution surveys, there is still an optimum for a given value of $\theta_{\text{FWHM}}^\text{lim}$, but this would be sub-optimal relative to performing a survey with the same receivers, but using a telescope with higher resolution.

It is simple to understand the degenerate line in Fig. 6: one has to achieve a given brightness temperature sensitivity at the saturation resolution of $2'$ in order to detect the clusters with masses $\sim 10^{14}M_\odot$ which is most efficient to detect. The required sensitivity at a higher resolution is that, which when smoothed, is optimal at the saturation resolution, that is, $T_{\lim} \theta_{\text{FWHM}} \propto S_{\text{lim}}/\theta_{\text{FWHM}}$ is resolution independent. This corresponds to it being independent of $N_b$ as discussed earlier and this is an important property of a single dish, total power measurement.

We have done the same calculation, but this time allow-
Table 1. The optimum limiting flux densities (in µJy) for different set ups, which correspond to the maxima of the graphs curves in Fig. 8. We have assumed the idealized total power setup allowing for beam smoothing described in section 4. The numbers are quoted to the nearest 10µJy except for $\theta_{\text{FWHM}} = 0.5'$ and $f = 15$GHz where the flux is quoted to the nearest 5µJy.

| $\theta_{\text{FWHM}}$ | $f = 15$GHz | $30$GHz | $90$GHz | $150$GHz |
|------------------------|-------------|---------|---------|---------|
| 0.5'                   | 35          | 140     | 870     | 1030    |
| 1'                     | 70          | 280     | 1730    | 2030    |
| 2'                     | 160         | 620     | 3830    | 4500    |
| 4'                     | 480         | 1870    | 11520   | 13470   |
| 8'                     | 1500        | 5810    | 35690   | 41620   |

Table 2. The thermodynamic temperature limits (in µK) quoted which correspond to the flux density limits presented in Table 1 to the nearest 5µK. Note that these can also be computed approximately from $T_{\text{lim}}^{\text{therm}} \approx 55\mu K |g(x)| / (\theta_{\text{FWHM}}/1')$ for $\theta_{\text{FWHM}} \leq 2'$.

| $\theta_{\text{FWHM}}$ | $f = 15$GHz | $30$GHz | $90$GHz | $150$GHz |
|------------------------|-------------|---------|---------|---------|
| 0.5'                   | 230         | 225     | 185     | 110     |
| 1'                     | 110         | 110     | 90      | 55      |
| 2'                     | 60          | 60      | 50      | 30      |
| 4'                     | 45          | 45      | 40      | 25      |
| 8'                     | 35          | 35      | 30      | 20      |

Figure 8. The relative number of clusters with $z < 1.5$ that would be found in a survey with $f = 15$GHz (bottom left), $f = 30$GHz (bottom right), $f = 90$GHz (top left) and $f = 150$GHz (top right). In each case $\theta_{\text{FWHM}} = 0.5'$ (short dashed line), 1' (dotted line), 2' (solid line), 4' (dashed-dotted line) and 8' (long dashed line). The optimal survey strategy would be to choose $\theta_{\text{lim}}$ to take the value where the curves are maximum. They are normalized by the largest possible number of clusters that could be found in survey with that frequency and hence we see that surveys with $\theta_{\text{lim}} > 2'$ are not optimal. The curves represent a slice through the contour plot in Fig. 6 expressed in flux density units which naturally introduces a dependence on $f$.

5 MORE REALISTIC OBSERVATIONS

5.1 Systematics in total power measurements

The ability to perform beam smoothing is not a given, but we have shown that it is an essential feature of an SZ survey instrument with resolution $< 2'$. It appears that a number of ground based single dish experiments have been able to make accurate measurements of the angular power spectrum of the primary CMB anisotropies over a wide range of scales. However, the reliable detection of extended objects like clusters requires the maps to not only have the power distributed correctly, but also requires the phases to be accurate.

A variety of instrumental systematics, such as baseline offsets caused by the receivers and the atmosphere, could lead to artifacts which might destroy the integrity of a map once it is smoothed to low resolution. In terms of the design of an SZ survey this leads to the possibility that $\theta_{\text{lim}} < \infty$. In what follows we will assume that some beam smoothing is possible, but that this is limited to some constant multiple of $\theta_{\text{lim}}$. In particular, we consider the effects of $\theta_{\text{lim}} = 3\theta_{\text{lim}}$ and $\theta_{\text{lim}} = 5\theta_{\text{lim}}$ in Figs. 9 and 10. We see that in both cases the degenerate line seen in Fig. 6 is truncated at small values of $\theta_{\text{FWHM}}$ (higher values for $\theta_{\text{lim}} = 3\theta_{\text{lim}}$). This can be understood by realizing that instruments with very small value of $\theta_{\text{FWHM}}$
are unable to achieve critical value of $2'$ by beam smoothing when these constraints are placed on $\theta_{\text{FWHM}}$. We conclude from this, that although extra resolution might be desirable (for example, to allow subselection within the sample), excessive amounts can be damaging to an instrument’s survey potential.

5.2 Beam switching

Beam switching can be used to subtract out the effects of the atmosphere in single dish observations when it varies on timescales shorter than the characteristic integration time. To investigate its effect, we will consider two Gaussian beams separated by angle $\theta_s$ on the sky. It is likely that the properties of the atmosphere will prevent $\theta_s$ from being greater than some value (probably somewhere between 5' and 10'), but from the point of view of our discussion we will assume that it can take any value between $\theta_{\text{FWHM}}^{\text{min}}$ and $\infty$.

First, let us consider a flat sky approximation with coordinates $x$ and $y$, and let $\tau_c(x, y)$ be the optical depth to the SZ effect along the line of sight parameterized by $(x, y)$. If a Gaussian beam, with width denoted by $\sigma$, is centred at the point $(x_s, 0)$ then the resultant flux will be proportional to

$$\int dx \int dy \tau_c(x, y) \exp \left[ -\frac{1}{2\sigma^2} \left( (x - x_s)^2 + y^2 \right) \right],$$

where the range of integration is $x^2 + y^2 < R_{\text{FWHM}}^2$. The optical depth $\tau_c(x, y) = \tau_c(0) \zeta(r)$ is only dependent on the radial coordinate $r = \sqrt{x^2 + y^2}$. By making the coordinate change $x = r \cos \phi$ and $y = r \sin \phi$, one can write this integral as

$$\tau_c(0) \int_0^{R_{\text{FWHM}}} dr r \zeta(r) \exp \left[ -\frac{r^2}{2\sigma^2} \right] K(r, x_s),$$

where

$$K(r, x_s) = \exp \left[ -\frac{x_s^2}{2\sigma^2} \right] \int_0^{2\pi} d\phi \exp \left[ \frac{rx_s \cos \phi}{\sigma^2} \right].$$

If $K(r, x_s) = 1$ then this expression is that of a Gaussian beam observing the centre of the object as used in our discussion of idealized total power measurements. This factor represents the effects of a non-centred beam.

One can re-write (23) in terms of the modified Bessel function $I_0(z)$ and, by considering the difference between a beam pointing at the origin and one pointing at $x = x_s$, an effective beam $B_{\text{eff}}(\theta)$ can be constructed which represents the effects of a beam switch. If we also make the transition from the flat-sky coordinate $r$ to the angular coordinate $\theta$, factor $I(B, \zeta)$ becomes

$$I(B_{\text{eff}}, \zeta) = \frac{\int d\Omega B_{\text{eff}}(\theta) \zeta(\theta)}{\int d\Omega \zeta(\theta)},$$

where

$$B_{\text{eff}}(\theta) = \exp \left( -\frac{\theta^2}{2\sigma^2} \right) \left[ 1 - \exp \left( -\frac{\theta^2}{2\sigma^2} \right) I_0 \left( \frac{\theta\theta_s}{\sigma^2} \right) \right].$$

In Appendix B, we show that, under sensible assumptions as the form of $\zeta(\theta)$, $I(B_{\text{eff}}, \zeta) \rightarrow 0$ as $\theta_s/\theta_{\text{FWHM}} \rightarrow \infty$, in which case it is clear that $M_{\text{lim}} \rightarrow \infty$ as $z \rightarrow 0$ in contrast to the case of a Gaussian beam profile for which $M_{\text{lim}}$ tends to a finite value. The reason for this is simple: as one approaches $z = 0$ the clusters become much larger, but the beam switching makes it impossible to find objects which are very much larger than the switching angle. This phenomenon will play a key role in our subsequent discussion.

We have computed $M_{\text{lim}}(z)$ for $\theta_{\text{FWHM}} = 1'$, $f = 30\text{GHz}$ and $S_{\text{lim}} = 500\mu\text{Jy}$ for a range of values for $\theta_s$ compared to that for no switching and the results are presented in Fig. 11. At this stage, we ignore the effects of beam smoothing. We see that for large values of $\theta_s \approx 10'$, the limiting mass is the same as for no switching when $z > 0.5$. However, as suggested above, we see that for all finite values of $\theta_s$, $M_{\text{lim}} \rightarrow \infty$ as $z \rightarrow 0$. As $\theta_s$ is reduced the value of $z$ above which the no switching case is a good approximation increases. For $\theta_s \approx 3'$ the limiting mass is affected ($\sim 50\%$) even at $z \approx 1$; given that the number of clusters found in the survey is very sensitive to the limiting mass this can lead to the number of clusters detected being significantly reduced.

The calculation discussed above does not take into account beam smoothing which is more subtle than in the case.
of a single-dish observation. In the case of switching it will be impossible to perform beam smoothing beyond the switching angle, that is, $\theta_{\text{FWHM}}^{\text{max}} = \theta_\nu$. We illustrate the effects of this on the limiting mass the same values of $f$, $\theta_{\text{FWHM}}^{\text{min}}$ for a range of values of $\theta_\nu$ in Fig. 12 compared to the case of no switching. Once again we see that for all finite values of $\theta_\nu$, $M_{\lim} \rightarrow \infty$ as $z \rightarrow 0$, but that for $\theta_\nu > 4'$ the limiting mass is only significantly different when compared to the case of no switching when $z < 0.2$. Since there are only very few clusters to be detected at these low redshifts, we conclude that for $\theta_\nu > 4'$ and $\theta_{\text{FWHM}}^{\text{min}} = 1'$ beam switching does not affect the number of clusters detected in a given survey.

Now that we have a method for computing the limiting mass in the case of beam switching, we can attempt to investigate how to make the optimum choice for the survey depth as we did for the single dish, total power case. We will assume that one can beam smooth up to $\theta_{\text{FWHM}}^{\text{max}} = \theta_\nu$ and will do this only for the case of $\theta_{\text{FWHM}}^{\text{min}} = 1'$ and $f = 30\,\text{GHz}$, and will work in terms of $S_{\lim}$ (this can be converted into $T_{\lim}$ using (17)). Fig. 13 is the equivalent to Fig. 8, but for a range of values of $\theta_\nu$. We see that, so long as the optimum choice of $S_{\lim}$ is made, the number of objects is only marginally reduced for $\theta_\nu > 5'$. For $\theta_\nu = 4'$ the number of objects found is only reduced to about 80% at the optimum, while for $\theta_\nu = 3'$ one can only achieve about 60% of that possible for no switching. The number of objects found is significantly reduced for lower values of $\theta_\nu$. Moreover, we see that the optimal value of $S_{\lim}$ is reduced from $280\,\mu\text{Jy}$ (see Table 1) for no switching to around $200\,\mu\text{Jy}$ ($T_{\lim}^{\text{THERM}} \approx 80\,\mu\text{K}$) for $\theta_\nu = 4'$ and $100\,\mu\text{Jy}$ ($T_{\lim} \approx 40\,\mu\text{K}$) for $\theta_\nu = 3'$. Although the optimal value of $S_{\lim}$ only changes by about a factor 3 in going from $\theta_\nu = \infty$ to $\theta_\nu = 3'$, the number of objects found for larger values of $S_{\lim}$ falls off much more quickly than in the no switching case. Both these effects can be explained by the fact that it is difficult to detect objects at very low redshifts for small values of $\theta_\nu$: as $\theta_\nu$ decreases one has to make surveys which are deeper and deeper in order to find anything, and very shallow surveys are unlikely to find anything at all since they are only sensitive to the largest, very nearby clusters which are effectively resolved out by the beam switch.

5.3 Aspects of interferometric observations

As we have already pointed out, making total power, single-dish measurements are not always possible. Interferometric methods using a telescope array provide an alternative to this, which can reduce the sensitivity of the measurements to a range of systematic effects notably those due to the atmosphere. The techniques involved (interferometers work in Fourier space, whereas single dishes work in real space) are very different. In this section we will discuss how, under certain circumstances, they can be thought of as being a special case of the total power measurements which we described in section 4 and we will outline how they can be incorporated into our framework. There are definitely some caveats to their application to an interferometer which we will ignore, for the most part, here.

Let us consider an array with $n_t$ receivers mounted on
Fig. 13. The relative number of clusters with \( z < 1.5 \) that would be found in a beam switch surveys with \( f = 30 \)GHz and \( \theta_{\text{FWHM}} = 1' \) for different values of the switching angle \( \theta_s \). The curves are labelled as in Figs. 11 and 12. The solid line is that for no switching and the curves are normalized to one at the maximum for this at around \( S_{\text{lim}} \approx 280 \mu\text{Jy} \). As the switching angle decreases, searching for clusters becomes less efficient, and the optimal limiting flux decreases to around \( S_{\text{lim}} \approx 100 \mu\text{Jy} \) for \( \theta_s = 3' \) which is about 60% efficient.

The relative number of clusters with any survey with a fixed integration time, one would be able to cover less area to the same depth.

The other important difference which needs to be taken into account can be very significant in terms of the optimum depth of a survey. The quantity \( T_{\text{lim}} \theta_{\text{FWHM}} \) need not be independent of resolution for an interferometer, as it is for a single dish; this is a consequence of the density of visibilities within the \( u-v \) plane generated by positions of the antennae in the interferometer. We will describe an interferometer for which \( T_{\text{lim}} \theta_{\text{FWHM}} \) is independent of resolution as being uniformly filled.

We can attempt to model an interferometer within our framework by assuming that the baseline lengths can be grouped into \( m \) bins and that there is an approximate resolution \( \theta_{\text{FWHM}} \approx \lambda/b_i \) corresponding to each of the bins \( 1 \leq i \leq m \), where \( b_i \) is the average length of the baselines in bin \( i \). There should be no significant error introduced by binning of the baselines if \( m \) sufficiently large. We will assume that at each resolution the response of the interferometer can be modelled by a Gaussian beam with FWHM \( \theta_{\text{FWHM}} \) which should be a good approximation so long as the number of baselines with length \( \leq b_i \) is sufficiently large. By reconstructing the beam in this way, we are effectively ignoring the data from long baselines. The flux limit corresponding to resolution \( \theta_{\text{FWHM}} \) will be given by

\[
S_{\text{lim}}^i = \sqrt{\frac{B}{B^i}} S_{\text{lim}}^\min ,
\]

where \( B^i \) is the number in bins with length \( \leq b_i \).

There is no particular reason, a priori, why \( T_{\text{lim}}^i \theta_{\text{FWHM}} \propto S_{\text{lim}}^i/\theta_{\text{FWHM}} \) should be independent of \( i \), which is what would be required for the interferometer to be uniformly filled! Given a specific set of baseline lengths, it would be simple to compute the overall limiting mass as that which corresponds to the envelope of the individual values of \( i \).

One can rephrase this discussion in terms of a single dish with the same overall collecting area \( \propto n_s D^2 \), but the temperature sensitivity is diluted by the so-called ‘filling factor’ \( \eta_F = n_s D^2/(D + d)^2 \leq 1 \)

\[
T_{\text{sys}} = \frac{T_{\text{sys}}}{\eta_F \sqrt{\Delta f}}.
\]

from which we can see that any survey with a fixed integration time should satisfy

\[
\frac{T_{\text{sys}}}{\sqrt{\Delta f}} = \sqrt{\frac{4 \log 2}{\pi}} \frac{T_{\text{sys}}}{\theta_{\text{FWHM}}^\text{lim}} \eta_{\text{sys}} \eta_{\text{n}} \Delta f .
\]

and hence the corresponding instrumental figure of merit is given by

\[
\mathcal{M}_{\text{INT}} = \frac{\eta_{\text{sys}} \eta_{\text{n}} \eta^2 \Delta f}{T_{\text{sys}}^2} = \frac{\eta_{\text{sys}} \eta_{\text{n}} \eta^2 \Delta f}{(T_{\text{HERM}}/2) [g(x)]^2} .
\]

The limit \( \eta_{\text{sys}} = 1 \) corresponds to a fully-filled array and is equivalent in some sense to a single dish with focal plane array of \( n_b = n_r \) elements. We see that an interferometer which is fully-filled will be the most efficient. This is not particularly important in the context of our discussion here since it just means that it will take longer to achieve the same temperature sensitivity, or given a fixed integration time, one would be able to cover less area to the same depth.
Figure 14. A realization of the simple interferometer arrangement discussed in the test with \( n_r = 6 \), \( d \) and \( s \) is the longest and shortest baselines respectively. \( D \) is the dish size. In this case there are three different types of baseline which, in order of increasing length have solid, dotted and dashed lines.

where

\[
B^u = 2\pi \int_{u_{\text{min}}}^{u} u' \rho(u') du',
\]

(31)

is the number of visibilities \( < u \) and \( \rho(u) \) is the density of visibilities. The case of a uniformly filled interferometer can then be seen to correspond to \( \rho(u) = \rho_0 \), that is, a uniform density of visibilities in the \( u-v \) plane.

An interesting parameterization of interferometers in this context is to consider \( \rho(u) \propto u^\gamma \) where \( \gamma \geq 0 \). For \( u \gg u_{\text{min}}, B^u \propto u^{\gamma+2} \) and hence

\[
S_{\text{lim}}^u = \left(\frac{u_{\text{max}}}{u}\right)^{\frac{\gamma}{2}+\frac{3}{2}} S_{\text{lim}}^d.
\]

(32)

If we choose \( N_h = (u_{\text{max}}/u)^{1/2} \) in order detect clusters which cover \( N_h \) beams, then we find that

\[
S_{\text{lim}}^d = N_h^{\frac{2}{2}+\frac{3}{2}} S_{\text{lim}}^d.
\]

(33)

We will explore this parameterization in future work, but it is clear that unless \( \gamma = 0 \) the behaviour of an interferometer when smoothing will differ from that of the idealized single dish case discussed in the previous section.

In order to illustrate this further, we will consider a very specific arrangement of the telescopes. Let us assume that they are arranged in a circular configuration. In order to make our discussion simpler we will assume that \( n_r \) is even, in which case there \( n_r/2 \) different baseline lengths. If these baseline lengths are denoted \( b_i \) for \( 1 \leq i \leq n_r/2 \) then

\[
b_i = d \sin \left[ \frac{i\pi}{n_r} \right],
\]

(34)

where there are \( n_r \) of length \( b_i \) for \( i \leq n_r/2 - 1 \) and \( n_r/2 \) with length \( b_{n_r/2} = d \). A diagram illustrating the arrangement for \( n_r = 6 \) is presented in Fig. 14.

The shortest baseline is given by \( s = d \sin[\pi/n_r] \) and hence the ratio of the shortest to longest baseline, which is equal to \( \theta_{\text{FWHM}}^{\text{min}} / \theta_{\text{FWHM}}^{\text{max}} \) is given by

\[
\frac{s}{d} = \frac{\theta_{\text{FWHM}}^{\text{min}}}{\theta_{\text{FWHM}}^{\text{max}}} = \sin \left[ \frac{\pi}{n_r} \right].
\]

(35)

The largest possible filling factor achievable with such a configuration is when \( s = D \), that is, when all the dishes in the circular arrangement are touching. In this case \( \eta_r = n_r/(1 + \text{cosec}(\pi/n_r))^2 \) which tends to zero like \( 1/n_r \) for large \( n_r \).

A range of resolutions are possible between the minimum and maximum, given by

\[
\theta_{\text{FWHM}}^i = \frac{\lambda}{b_i} = \text{cosec} \left[ \frac{i\pi}{n_r} \right] \theta_{\text{FWHM}}^{\text{min}},
\]

(36)

which result in the limiting flux density of

\[
S_{\text{lim}}^i = \sqrt{\frac{n_r - 1}{2i}} S_{\text{lim}}^{\text{min}},
\]

(37)

for \( i \leq n_r/2 - 1 \) and limiting temperature of

\[
T_{\text{lim}}^i = \sqrt{\frac{n_r - 1}{2i}} \sin \left[ \frac{i\pi}{n_r} \right] T_{\text{lim}}^{\text{min}} \theta_{\text{FWHM}}^{\text{min}}.
\]

(38)

It is clear that this interferometer is not uniformly filled for finite values of \( n_r \) since

\[
\frac{S_{\text{lim}}^i}{\theta_{\text{FWHM}}^i} = \sqrt{\frac{n_r - 1}{2i}} \sin \left[ \frac{i\pi}{n_r} \right] T_{\text{lim}}^{\text{min}} \theta_{\text{FWHM}}^{\text{min}},
\]

(39)

which is clearly not independent of \( i \). Such a configuration is close to being uniformly filled for the longest baselines if
\( n_i \) is large. If we define \( j = n_i/2 - i \) and assume that \( n_i \) is large then
\[
T_{\text{lim}}^i \theta_{\text{FWHM}}^i \approx T_{\text{lim}}^{\text{min}} \theta_{\text{FWHM}}^{\text{min}} + O(j),
\] (40)
that is, the longest baselines are close to being uniformly filled. But even in this case, the filling factor is not uniform for small \( i \) where the discreteness takes hold.

We have plotted the mass limits for \( 1 \leq i \leq 4 \) for \( n_i = 8 \) in Fig. 15 for a range of values of \( \theta_{\text{FWHM}}^{\text{min}} \) along with the predictions of a single dish with the same maximum resolution. In this case the best filling factor possible is \( \eta_F = 8/(1 + \csc(\pi/8))^2 \approx 0.6 \). The actual mass limit that one would achieve with an interferometer is given by the envelope of the curves for \( i = 1 \) to 4, and it is clear to see that one always does worse in terms of detecting clusters with an interferometer when compared to an ideal single dish with the same resolution. We note that despite this advantage, systematics are nearly always worse for single dish observations.

The values of \( S_{\text{lim}}^{\text{min}} \) have, in each case, been chosen to be the values which are close to the optimal survey strategy in the case where the interferometer is uniformly filled (but this might not be optimal here). There are some interesting aspects of the problem which are raised by Fig. 15. We see that for \( \theta_{\text{FWHM}} = 1' \) if \( z < 1 \) the mass limit is ostensibly defined by that for \( i = 1 \) corresponding to the shortest baselines, whereas for \( \theta_{\text{FWHM}} = 8' \) it is that for \( i = 4 \) corresponding to the longest baselines which dominates except for the very smallest values of \( z \).

6 ARE OUR PREDICTIONS ROBUST?

6.1 Cosmological parameters

In making our calculations we found it necessary to make choice of cosmological parameters. It is unlikely that the total number of clusters found in a given survey is very strongly dependent on the parameters \( h, \omega_0 \) and \( \omega_1 \) although how they are distributed as a function of \( z \) should be. However, it is known that it will be strongly dependent on \( \sigma_8 \) and \( \Omega_m \). We have made sensible choices for these parameters in the light of presently available cosmological observations, but it is perfectly possible that our choices are incorrect. The value we have used, \( \sigma_8 = 0.9 \), is that preferred by measurements of the primary anisotropies in the CMB and weak lensing. However, measurements of the X-ray luminosity function of clusters and the gas fraction therein suggest that \( \sigma_8 \approx 0.7 \) (Allen, Schmidt & Fabian 2002).

In light of this discrepancy, we have performed our calculations for a range of different values of \( \sigma_8 \) between 0.7 and 1.1. We find that there is a degenerate line in the equivalent of Fig. 6 for every value of \( \sigma_8 \) we tried and that a survey can be optimal for \( \theta_{\text{FWHM}} < 2' \). However, the normalizing amplitude is modified by \( \sigma_8 \). We find that
\[
T_{\text{lim}} = 55\mu K \left( \frac{\sigma_8}{0.9} \right)^{4.1},
\] (41)
at \( \theta_{\text{FWHM}} = 1' \). The strong dependence on \( \sigma_8 \) is due to the exponential dependence of the number of clusters on it.

We have also investigated the effects of variations in \( \Omega_m \). Fig. 16 presents the equivalent of Fig. 6 for \( \Omega_m = 0.5 \).

![Figure 16](image)

**Figure 16.** Equivalent to Fig. 6 for \( \Omega_m = 0.5 \). There is the characteristic degenerate line, but with an amplitude of \( T_{\text{lim}} \approx 80\mu K \) at \( \theta_{\text{FWHM}} = 1' \).

Yet again there is a degenerate line for \( \theta_{\text{FWHM}} < 2' \) and for this case the amplitude for \( \theta_{\text{FWHM}} = 1' \) is \( T_{\text{lim}} \approx 80\mu K \).

6.2 Cluster model parameters

Another possible uncertainty of our calculation is the cluster model which we have used. The two important features of the model in this respect are the \( M-T \) relation (3) and the cluster profile function.

We have investigated the effects of varying the parameter \( T_\star \), although we note that it is perfectly possible that the relation between \( M \) and \( \langle T(\eta_i) \rangle \) as a function of \( z \) is not precisely of the form (3). In this case we find that the normalizing amplitude at \( \theta_{\text{FWHM}} = 1' \) is given by
\[
T_{\text{lim}} = 55\mu K \left( \frac{T_\star}{1.8} \right),
\] (42)
with the degenerate line and saturation value of \( \theta_{\text{FWHM}} \) as before. It is easy to understand this linear dependence of \( T_\star \) since modifying it just leads to a corresponding modification of the temperature of each cluster. It is clear that modification to the relationship between \( M \) and \( \langle T(\eta_i) \rangle \) to the \( z \) dependence can lead to more complicated modifications to the optimal value of \( T_{\text{lim}} \) dependent on how this affects the overall distribution of clusters. Once more information is available on this it might be worth investigating this further.

We have also varied the parameters of the isothermal \( \beta \)-model, \( \alpha \) and \( \beta \). For a general value of \( \beta \) the profile function is given by
\[
\zeta(\theta) = \left( 1 + \frac{\theta^2}{\theta_c^2} \right)^{-\frac{\alpha}{2} - \frac{\beta}{2} + \frac{1}{2}},
\] (43)
where
\[
J(a, b) = \int_0^a (1 + x^2)^{-\frac{3b}{2}} \, dx.
\] (44)

Fig. 17 shows the effect of varying \( \alpha = 5, 10 \) and 100 and \( \beta = 1/3, 2/3, 1 \) and 4/3 for a survey with \( \theta_{\text{FWHM}} = 1' \) and \( f = 30\text{GHz} \). We see that the optimal value of \( S_{\text{lim}} \) varies only by about a factor of 2 when \( \alpha \) varies by a factor of 20 between 5 and 100. We note also that larger values of \( \alpha \) lead...
to more clusters detected; about a factor of 2 between \( \alpha = 5 \) and 100. This is because for larger values of \( \alpha \) the clusters are more point-like, making a given mass easier to detect.

The variations as a function of \( \beta \) are more significant. The optimal value of \( S_{\text{lim}} \) for \( \beta = 2/3 \) is 280\( \mu \)Jy and the total number of clusters increases by a factor \( \sim 10 \) in going from \( \beta = 1/3 \) to \( \beta = 4/3 \). As before this is because as \( \beta \) increases the cluster becomes more point-like: for \( \beta = 1/3 \), the electron number density \( n_e \propto r^{-1} \), whereas for \( \beta = 4/3 \), \( n_e \propto r^{-4} \).

### 6.3 Mass function

The final major aspect of our calculation which could contain systematic errors is the comoving number of density or temperature limit is typically quite broad, which is fortun-ate since the calculations we have made contain

In Fig. 18 we have plotted the same quantity as in Fig. 8 for \( \theta_{\text{FWHM}} = 1' \) and \( f = 30 \)GHz using the mass function already used (Evrard et al. 2002) which is based on VIRGO consortium simulations, their alternative mass function (Jenkins et al. 2001), the ST (Sheth & Tormen 1999) mass function using \( M = M_{60} \) and the PS (Press & Schechter 1974) mass function using \( M = M_{500} \). We see that the maximum of the curve, and hence the optimal survey depth is almost independent of the mass function within the bounds of those functions which we have investigated. Rather worryingly, the overall number of clusters found for a given flux limit varies by a factor of two for the different mass functions. This point was highlighted in Battye & Weller (2003), and it is clearly something which needs further investigation. Nonetheless, we conclude that uncertainties in the mass function are not significant when considering the optimal depth of a survey.

### 7 CONCLUSIONS

We have discussed in detail the optimal depth of an SZ cluster survey and have as a by-product gained some insight in the optimal experimental configuration. Ignoring the effects of confusion noise, for every beamsize gained some insight in the optimal experimental configuration. Ignoring the effects of confusion noise, for every beamsize considered we have shown that there exists an optimum depth at which one will find the largest number of clusters for a fixed integration time. The number of clusters found as a function of the flux density or temperature limit is typically quite broad, which is fortun-ate since the calculations we have made contain
many uncertainties. We have shown that the most significant of these are those relating to the $M - T$ relation and the cluster profile. Even within the known uncertainties in these parameters it should be possible to get within 50% of the optimum.

For an instrumental setup with $\theta_{\text{FWHM}}^\text{min} < 2'$ and assuming that beam smoothing is possible, one is brightness temperature limited and it is inefficient to make very deep surveys searching for clusters which match a much smaller beam. One is always in a position to perform shorter integrations on many small beam areas and smooth them down to the lower resolution where the maximum number of clusters are available. These conclusions are particularly sensitive to some of the more sophisticated experimental techniques which are used to counteract the effects of the atmosphere. We have discussed two: in the case of beam switching, we have shown that one is not severely affected if the beam switch is not so small as to prevent one detecting the most abundant clusters which are which have a virial diameter of $\sim 2'$.

The discussion of the optimal design of an interferometer and the survey strategy is complicated by the extra freedom due to density of visibilities in the $u$-$v$ plane. We have discussed how our methods can be applied in that case, but, other than asserting that the most filled interferometer would have the optimal design, we have not made any statements on the optimal strategy. Given a particular situation it might be possible to do this in more detail.

We hope that our calculations can be used to assist observers in their choice of survey strategy. We acknowledge numerous times that our predictions are model dependent. One could imagine a method by which one could search adaptively for the optimum depth of a survey as part of the process of performing the observation. The basic idea is to take some small patch of sky, compared to the potential survey volume, and compute the number of sources as a function of $S_{\text{lim}}$. When the logarithmic gradient of this quantity is greater than -2, the optimum has been achieved, and that value of $S_{\text{lim}}$ should be used on the rest of the survey area. Clearly, this will only be possible in scenarios where one is expected to find a large number of objects since one will have to find a sufficiently large number in the initial test area to make the estimate of the gradient accurate enough.

**ACKNOWLEDGEMENTS**

RAB is funded by PPARC and thanks the members of the CMB group at Jodrell Bank Observatory, in particular Peter Wilkinson, Ian Browne and Clive Dickinson, for many useful discussions on the content of the is paper. JW is supported by the DOE and the NASA grant NAG 5-10842 at Fermilab.

**Appendix A**

Let us consider computing the integral $I(B, \zeta)$ when $\zeta(\theta) = F(y)$ where $y = \theta^2 / \theta_{\text{vir}}^2$ for $\theta \leq \theta_{\text{vir}}$ and zero for $\theta > \theta_{\text{vir}}$. The function is defined so that $F(0) = 1$. It is easy to see that for the isothermal $\beta$-model with $\beta = 2/3$

$$F(y) = (1 + \alpha^2 y^2)^{-1/3} \tan^{-1} \left[ \frac{3 y / (1 + \alpha^2 y^2)^{3/2}}{\tan^{-1} \alpha} \right],$$

(45)

and it is likely that most sensible profile functions will have this property.

If we define $R = 4(\theta_{\text{vir}} / \theta_{\text{FWHM}})^2 \log 2$ then

$$I(B, \zeta) = \frac{\int_0^1 e^{-Ry} F(y) dy}{\int_0^1 F(y) dy}.$$  

(46)

One can see that in the limit $R \to 0$ the function $I \to 1$ which corresponds to the point source limit. The opposite limit, $R \to \infty$, is more interesting since it corresponds to that when the cluster is much larger than the beam. One can show that the leading order term in the asymptotic expansion in $1/R$ is

$$I \sim \left[ R \int_0^1 F(y) dy \right]^{-1} \propto \theta_{\text{vir}}^{-2}.$$  

(47)

From (8) we see that $S \propto (T_s)_n M_{\text{vir}} \theta_{\text{vir}}^2$ and since $R_{\text{vir}} = \theta_{\text{vir}}^2 \theta_{\text{FWHM}}$, we can deduce that

$$S \propto \frac{(T_s)_n M_{\text{vir}}}{R_{\text{vir}}^2},$$

(48)

which is finite as $z \to 0$. Hence, we see that the limiting mass tends to a finite value at $z = 0$ in contrast to the case of a point source.

**Appendix B**

Let us consider $I(B_{\text{eff}}, \zeta)$ for $B_{\text{eff}}(\theta)$ given by that for a switch beam system (25). We will assume that $F(y)$ and $R$ are defined as in Appendix A, and will define $Q = 4(\theta_{\text{vir}} / \theta_{\text{FWHM}})^2 \log 2$. Using these definitions we find that

$$I(B_{\text{eff}}, \zeta) = \frac{\int_0^1 e^{-Qy} \left[ 1 - e^{-QI_0(2\sqrt{RQy})} \right] F(y) dy}{\int_0^1 F(y) dy}.$$  

(49)

Understanding the asymptotic behaviour of this integral is more difficult than in the standard case. In order to make some progress let us write $F(y)$ as a power series

$$F(y) = \sum_{n=0}^{\infty} A_n y^n,$$

(50)

where $A_0 = 1$ and $A_1 < 0$. From this we can deduce that

$$I(B_{\text{eff}}, \zeta) = \frac{\sum_{n=0}^{\infty} A_n (B_m - e^{-Q} C_m)}{\int_0^1 F(y) dy}.$$  

(51)

where

$$B_m = \int_0^1 e^{-Qy} y^m dy,$$

(52)

$$C_m = \int_0^1 e^{-Qy} I_0(2\sqrt{RQy}) y^m dy.$$  

(53)

One can write $C_m$ in terms of $B_m$ by writing $I_0(x)$ as a power series in $x$,

$$C_m = \int_0^1 \frac{1}{(n!)^2} (RQ)^n B_{m+n}.$$  

(54)

Moreover, it can be shown that $B_m \sim m! / R^{m+1}$ and hence $C_0 \sim e^Q / R$ and $C_1 \sim (Q + 1)e^Q / R^2$. From this we see that in the asymptotic expansion, the first term in (51) cancels and therefore
This cancellation is enough to ensure that $I \to 0$ as $z \to 0$ sufficiently fast so that

$$S \propto \frac{(T_c)_{\nu} M_{\text{vir}} d_{\text{A}}^2}{R_{\text{vir}}^2},$$

in this limit. It is easy to see that this forces the limiting mass to tend to infinity as $z \to 0$.

REFERENCES
Allen S.W., Schmidt R.W., Fabian A.C., 2002, MNRAS 334, L11
Barbosa et al., 1996, A&A, 314, 13
Bartlett J., 2000, astro-ph/0001267
Battye R.A., Weller J., 2003, Phys. Rev. D68, 083506
Browne I.W. et al., 2000, in ‘Radio Telescopes’ edited by H.R. Butcher, SPIE Proc Vol 4015, (SPIE-International Society for Optical Engineering, Bellingham, WA, 2000)
Bryan G.L., Norman M.L., 1998, ApJ 495, 80
Cavaliere A., Fusco-Femiano R., 1976, Astron. Astrophys., 49, 137
Contaldi C.R., Hoekstra H., Lewis A., 2003, Phys. Rev. Lett 90, 221303
Eke V., Cole S., Frenk C., 1996, MNRAS, 282, 323
Evrard A.E. et al, 2002, Ap. J., 573, 7
Freedman W.L. et al, 2001, Ap. J., 553, 47
Gunn J.E., Gott J.R., 1972, ApJ, 176, 1
Haiman Z., Mohr J.J., Holder G., 2001, ApJ 553, 545
Holder G., Haiman Z., Mohr J.J., 2001, ApJ 560, L11
Holder G. et al., 2000, ApJ 544, 629
Jenkins A. et al., 2001, MNRAS 321, 372
Kneissl R. et al., 2001, MNRAS 328, 783
Kolb E.W., Turner M.S., ‘The Early Universe’ (Addison Wesley, 1990)
Kosowsky A., 2003, New Astron. Rev. 47, 939
Lo K. et al., 2002, in ‘New cosmological data and the values of the fundamental parameters’, edited by A. Lasenby and A. Wilkinson (Astronomical Society of the Pacific, Salt Lake City, 2000)
Lahav O. et al, 1991, MNRAS 251, 128
Mauskopf P.D. et al, in ‘Imaging at radio through submillimeter wavelength’ edited by J. Mangum (Astronomical Society of the Pacific, Salt Lake City)
Mohr J.J., Mathiesen B., Evrard A.E., 1999, Ap. J. 517, 627
Mohr J.J. et al, 2003, Nucl. Phys. Proc. Suppl. 124, 63
Ostriker J.P., Steinhardt P.J., 1995, Nature 377, 600
Oukbir J., Blanchard A., 1992, Astro. Ap. 262, L210
Partridge R.B., Peebles P.J.E., 1967, ApJ 147, 868
Percival W.J. et al, 2001, MNRAS 327, 1297
Percival W.J. et al., 2002, MNRAS 337, 1068
Pierpaoli E., Scott D., White M., 2001, MNRAS 325, 77
Press W.H., Schechter P., 1974, ApJ 187, 425
Sunyaev R.A., Zel’dovich Ya., 1972, Comm. Astrophys. Space Phys., 4, 173
Sheth R., Tormen G., 1999, MNRAS 308, 119
Scaramella R., Cen R., Ostriker J.P., 1993, ApJ 416, 399
Silk J., 1968, ApJ, 529, 12
Spergel D.N. et al, 2003, ApJS 148, 175
Stolyarov A., Hobson M.P., Ashdown M.A.J., Lasenby A.N., 2002, MNRAS 336, 97
Thomas P., Carlberg R.G., 1989, MNRAS, 240, 1009
Wang L., Steinhardt P.J., 1998, ApJ 508, 483
Weller J., Battye R.A., Kneissl R., 2002, Phys. Rev. Lett. 88, 231301