XXIst Moriond Astrophysics Meeting - Les Arcs March 10-17, 2001

THE SUNYAEV-ZEL'DOVICH EFFECT

M. BIRKINSHAW
Department of Physics, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, U.K.

This article reviews the physics behind the Sunyaev-Zel'dovich effect, and the use of the effect to measure the gas contents of clusters of galaxies and as a tracer of the large-scale structure of the Universe. The current observational status of the effect is briefly reviewed, and some new dedicated instruments are briefly described.

1 The physics of the Sunyaev-Zel'dovich effect

The Sunyaev-Zel'dovich effect is produced by the inverse-Compton scattering of the cosmic microwave background radiation by hot gas, especially gas in the atmospheres of clusters of galaxies [41]. The physical process involved has been reviewed recently by Rephaeli [38] and Birkinshaw [6], and leads to two intensity effects with different spectra — the dominant thermal Sunyaev-Zel'dovich effect and the smaller kinematic effect.

1.1 The thermal Sunyaev-Zel'dovich effect

The larger effect arises through the upscattering of the low-energy photons of the microwave background radiation by the thermal electrons in the hot gas in clusters of galaxies. The average fractional frequency change of a scattered photon is proportional to the dimensionless electron temperature, $T_{\rm e}$,

$$\frac{\Delta\nu}{\nu} \propto \frac{k_{\rm B}\, T_{\rm e}}{m_{\rm e}\, c^2} \tag{1}$$

where c, $k_{\rm B}$, and $m_{\rm e}$ are the speed of light, the Boltzmann constant, and the electron mass. The optical depth to electron scattering is

$$\tau_{\rm e} \approx n_{\rm e}\, \sigma_{\rm T}\, l_{\rm z} \tag{2}$$

where $n_{\rm e}$ is the electron number density, $\sigma_{\rm T}$ is the Thomson scattering cross-section, and $l_{\rm z}$ is the line-of-sight distance through the gas distribution.

Combining these factors, the brightness temperature of the microwave background radiation seen through a cluster of galaxies is changed, relative to lines of sight which miss the gas, by an amount

$$\Delta T_{\rm RJ} \approx -2\, \tau_{\rm e} \left(\frac{\Delta\nu}{\nu}\right) T_{\rm rad} \quad . \tag{3}$$

Here $T_{\rm rad} = 2.728 \pm 0.002$ K is the radiation temperature of the microwave background [13] and the factor -2 is appropriate for the Rayleigh-Jeans part of the spectrum.

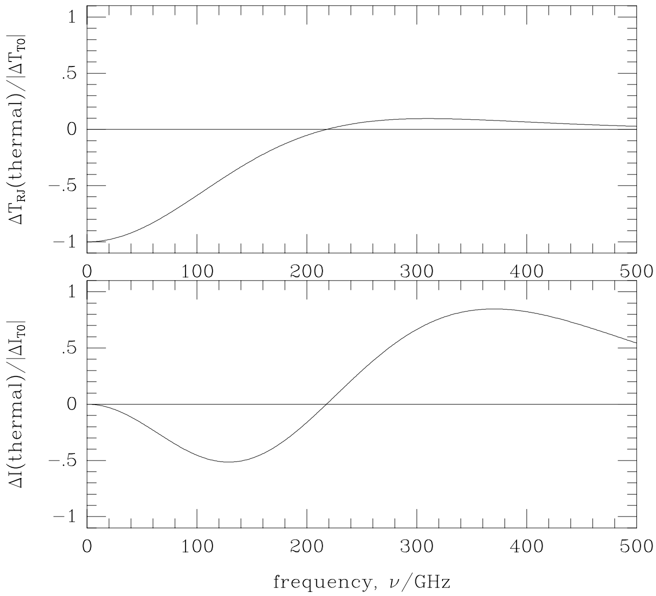

Figure 1: The spectrum of the thermal Sunyaev-Zel'dovich effect for non-relativistic electrons ($k_{\rm B}T_{\rm e} \lesssim 10$ keV) expressed as brightness temperature ($\Delta T_{\rm RJ}$) and intensity (ΔI) changes.

Since X-ray luminous clusters of galaxies may have $k_{\rm B}T_{\rm e} \approx 10$ keV, and average electron densities $n_{\rm e} \approx 10^3\ {\rm m}^{-3}$ over path lengths $l_{\rm z} \approx 3$ Mpc, both $\Delta\nu/\nu$ and $\tau_{\rm e}$ can be of order 10^{-2}, and so $\Delta T_{\rm RJ}$ can exceed 1 mK. For a cluster with an angular size of 1 arcmin, the corresponding flux density is about -1 mJy at 30 GHz.

When calculated in detail, using the non-relativistic Kompaneets [23] equation, or the relativistic formalism of Rephaeli[37] or Itoh *et al.*[20], the spectrum of the thermal Sunyaev-Zel'dovich effect is found to have an unusual form, with a negative flux density at low frequency and positive flux density at high frequency (Fig. 1). The detailed shape of the spectrum is independent of $T_{\rm e}$ in the Kompaneets limit, but varies significantly with electron temperature at $k_{\rm B}T_{\rm e} \gtrsim 10$ keV because a significant fraction of the electrons have velocities $> 0.2c$.

At zero frequency, the thermal Sunyaev-Zel'dovich effect has amplitude

$$\Delta T_{\rm T0} = 2\,T_{\rm rad}\,y_{\rm e} = 2\,T_{\rm rad} \int n_{\rm e}\,\sigma_{\rm T}\,\frac{k_{\rm B}\,T_{\rm e}}{m_{\rm e}\,c^2}\,dl \qquad (4)$$

which is proportional to the line-of-sight integrated electron pressure. $y_{\rm e}$, the Comptonization parameter, is a convenient dimensionless quantity that describes the strength of scattering on a particular line of sight.

1.2 The kinematic Sunyaev-Zel'dovich effect

A moving scattering medium produces a change in the brightness of radiation passing through it. The brightness temperature change has the same spectral form as primordial fluctuations in the background radiation (Fig. 2). The amplitude of the brightness temperature change at zero frequency is

$$\Delta T_{\rm K0} = \frac{v_{\rm z}}{c}\,\tau_{\rm e}\,T_{\rm rad} \qquad (5)$$

where $v_{\rm z}$ is the peculiar velocity of the scattering medium along the line of sight away from the observer. This effect is smaller than the thermal Sunyaev-Zel'dovich effect by a factor

$$\frac{\Delta T_{\rm K0}}{\Delta T_{\rm T0}} \approx 0.085\left(v_{\rm z}/1000\ {\rm km\,s^{-1}}\right)\left(k_{\rm B}T_{\rm e}/10\ {\rm keV}\right)^{-1} \qquad (6)$$

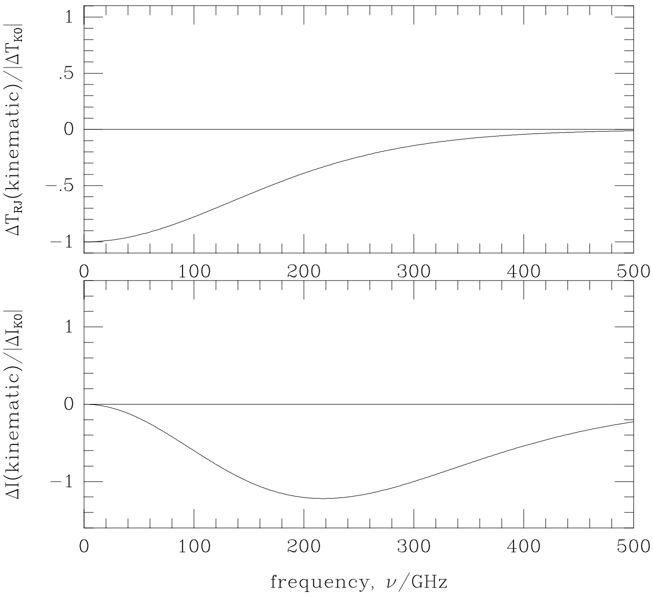

Figure 2: The spectrum of the kinematic Sunyaev-Zel'dovich effect, expressed in brightness temperature ($\Delta T_{\rm RJ}$) and intensity (ΔI) terms. Primordial fluctuations in the microwave background radiation have the same spectrum and so constitute a source of noise for measurements of the kinematic effect.

but has the same angular structure if the cluster is isothermal. It is therefore hard to detect in the presence of the larger thermal effect: separation of the two signals relies on their different spectra, with the kinematic effect dominating near the frequency at which the thermal effect changes sign

$$\nu_{\rm zero} = (217.7 \pm 0.2) \times \Big(1 + 0.022\,(k_{\rm B} T_{\rm e}/10\ {\rm keV})\Big) \quad {\rm GHz}. \tag{7}$$

Even at $\nu_{\rm zero}$ the kinematic effect will be difficult to detect unambiguously, because of confusion from primordial fluctuations in the microwave background. The kinematic effect is therefore easiest to measure for clusters of galaxies with small angular sizes ($\lesssim 4$ arcmin), since the confusion from primordial fluctuations begins to decrease rapidly at $l \gtrsim 1500$.

1.3 Polarization and other terms

Non-thermal electrons, such as those in radio galaxies or cluster radio halo sources, also have an effect on the microwave background radiation. The amplitude of this non-thermal Sunyaev-Zel'dovich effect is proportional to $\tau_{\rm e}$, rather than $y_{\rm e}$, and so a detectable Sunyaev-Zel'dovich effect can only be produced if the energy spectrum of the non-thermal electrons extends to unexpectedly low energies [6].

Clusters induce other secondary fluctuations in the microwave background radiation, including polarization associated with the thermal and kinematic Sunyaev-Zel'dovich effects (but weaker by factors $\sim \frac{v_{\rm z}}{c}$ or $\tau_{\rm e}$ [20,10]), and intensity effects from the time-varying gravitational field experienced by a photon passing through a non-static cluster.

This latter effect may arise because the cluster is contracting [35] or moving across the line of sight [34]. The fractional brightness change in the microwave background is of order $\frac{\Phi}{c^2}\frac{v}{c}$, where Φ is the gravitational potential, and v is the speed of structure change. While these effects would provide exciting information, they are obscured by primordial effects arising near recombination, of similar type and identical spectrum.

2 Uses of the effect

The amplitudes of the thermal and kinematic Sunyaev-Zel'dovich effects, $\Delta T_{\rm T0}$ and $\Delta T_{\rm K0}$, are independent of redshift. Clusters of any redshift with the same $\tau_{\rm e}$, $y_{\rm e}$, and $v_{\rm z}$ produce the same Sunyaev-Zel'dovich effects. Measurements of these effects can provide direct information about cluster atmospheres at any redshift where sufficient scattering is present for a detectable signal to be produced. Thus the effects can be regarded as good *mass finders* in the distant Universe, while the thermal effect can also be regarded as a *leptometer*, since it counts hot electrons, and the kinematical effect is a *radial speedometer*, measuring the radial components of the peculiar velocities of clusters.

Practical observations always cause $\Delta T_{\rm T0}$ and $\Delta T_{\rm K0}$ to be averaged over some solid angle, and this introduces a redshift dependence into the observable brightness temperature change (e.g., Fig. 6). That is, although the central brightness temperature effects are independent of redshift, the flux density effects are not, but decrease rapidly as redshift increases. Nevertheless, the ready detectability of massive clusters at high redshift [22] in the Sunyaev-Zel'dovich effects has led to much investigation of their use in cosmology.

2.1 Hubble constant

A major reason for the interest in the thermal Sunyaev-Zel'dovich effect is that it can be used to measure the Hubble constant. The method relies on the different dependences on cluster properties and distance of the thermal Sunyaev-Zel'dovich effect and the X-ray surface brightness. The quantity $\Delta T_{\rm T0} \propto n_{\rm e0}\, T_{\rm e0}\, r_{\rm c}$ (Eq. 4), where $n_{\rm e0}$, $T_{\rm e0}$, and $r_{\rm c}$ are the central electron density, central temperature, and core radius of the cluster. Similarly, the central X-ray surface brightness, $\Sigma_{\rm X0} \propto n_{\rm e0}^2\, \Lambda(T_{\rm e0})\, r_{\rm c}$, where $\Lambda(T_{\rm e})$ is the emissivity of the gas as a function of temperature. If the temperature of the gas is known from X-ray spectroscopy, then the combination of observables

$$\Lambda(T_{\rm e0})\, T_{\rm e0}^{-2} \left(\frac{\Delta T_{\rm T0}^2}{\Sigma_{\rm X0}} \right) \propto r_{\rm c} \quad , \tag{8}$$

and so the core radius can be inferred from measurements of $\Delta T_{\rm T0}$, $\Sigma_{\rm X0}$ and $T_{\rm e0}$. When compared with the angular size of the cluster, $\theta_{\rm c}$, the angular diameter distance $d_{\rm A} = r_{\rm c}/\theta_{\rm c}$ can be used to measure the Hubble constant

$$H_0 \propto \frac{\Sigma_{\rm X0}}{\Delta T_{\rm T0}^2}\, T_{\rm e0}^2\, \left[\Lambda(T_{\rm e0})\right]^{-1}\, \theta_{\rm c} \tag{9}$$

and other cosmological parameters in the angular diameter distance/redshift relationship, if clusters can be detected in their Sunyaev-Zel'dovich effects and X-ray fluxes at redshifts $\gtrsim 1$. This technique applies to each cluster as an individual, and can be applied to obtain an absolute measurement of distance at high redshift without using a cosmic distance ladder.

A representative nine-cluster distance/redshift diagram based on a recent compilation of distance estimates made in this way [6] is shown in Fig. 3. Although a formal analysis of this diagram implies a Hubble constant $H_0 = 55 \pm 10\ \rm km\, s^{-1}\, Mpc^{-1}$, this ignores several important issues.

First, the result for H_0 depends critically on the absolute flux calibrations of the radio and X-ray telescopes, and few independent telescopes are involved in Fig. 3. Likely systematic errors of 5 per cent in each scale would lead to a systematic error $\pm 8\ \rm km\, s^{-1}\, Mpc^{-1}$ in the result for H_0. Excellent absolute calibrations are essential for the method to be effective.

Second, the method requires a good model for the structure of the cluster atmosphere. Eq. 9 relies on measurements of the amplitudes of the central Sunyaev-Zel'dovich effect and X-ray brightnesses. These amplitudes are obtained by fitting a model structure for the cluster, which

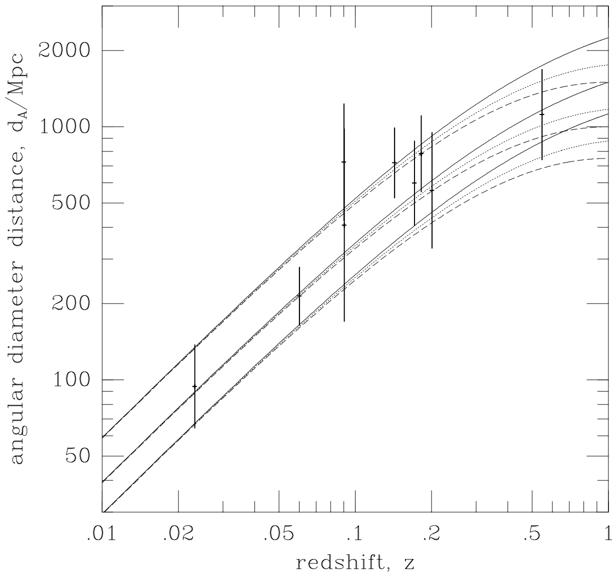

Figure 3: A plot of angular diameter distance as a function of redshift for nine clusters of galaxies for which $d_{\rm A}$ has been measured by comparing their Sunyaev-Zel'dovich effects and X-ray surface brightnesses. The curves show the angular diameter distance as a function of redshift for $H_0 = 50$, 75, and 100 $\rm km\,s^{-1}\,Mpc^{-1}$ (upper to lower groups of curves) and $q_0 = 0$, 0.5, and 1 (solid, dotted, dashed curves), with zero cosmological constant.

will depend on several parameters, and which will also provide the constant of proportionality in Eq. 9. But the choice of model is non-unique (although an isothermal beta model is commonly used), and different choices, like different fit parameters, lead to different results for H_0.

Particular difficulties arise because the X-ray flux of a cluster is more dominated by its central regions than is the thermal Sunyaev-Zel'dovich effect. The structure model must be a good description of the regions of the cluster which provide the bulk of the $\Delta T_{\rm T0}$ and $\Sigma_{\rm X0}$ signals. For single-dish measurements of the Sunyaev-Zel'dovich effect, this may require a high-fidelity description of the cluster structure from $r_{\rm c}$ to $10 r_{\rm c}$. Interferometric data have an advantage here, since they are relatively insensitive to the outer structure of the Sunyaev-Zel'dovich effect.

The deprojection implied by model fitting will certainly be wrong if the cluster is non-spherical. Non-spherical clusters increase the noise on the Hubble diagram [6,33,28]: if a cluster is extended along the line of sight, then its apparent angular size will be an underestimate of $r_{\rm c}/d_{\rm A}$, and so it will be assigned too large a distance and too small a Hubble constant. This effect leads to biases if clusters are selected for observation based on their X-ray or Sunyaev-Zel'dovich effect brightnesses, or even optical richness, because such selections prefer objects extended along the line of sight [5,40]. This is an important reason for selecting an orientation-independent sample of objects by total X-ray flux, or, better, by total Sunyaev-Zel'dovich effect flux density. For similar reasons, upper limits on $d_{\rm A}$ (arising from upper limits on $\Delta T_{\rm T0}$), as well as measurements, should be included in Fig. 3.

Unmodelled structures in the cluster atmosphere, such as clumping or thermal gradients, can lead to significant errors for individual clusters, or to systematic errors in the Hubble diagram as a whole [6,28]. Thus, for example, no sample of clusters can remove the effects of clumping, which cause a one-sided error in the Hubble constant, because clumping always increases the X-ray intensity relative to an unclumped atmosphere.

The number of distance measurements made using this technique is increasing rapidly, with several programmes of parallel X-ray and microwave background studies of clusters in progress. Recent measurements approximately double the number of clusters with distances measured using this technique (e.g. [19,25,32,30,36,1]), and tend to raise the overall best value of the Hubble constant to about 65 $\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

2.2 Cluster velocities

The kinematic effect can measure the peculiar radial velocity, v_z, of a cluster only if good spectral separation from the thermal effect has been achieved. However, an irreducible noise on v_z arises from underlying primordial fluctuations with an identical spectrum (Fig. 2). On angular scales of about 2 arcmin, the primordial signal is expected to be about 10 μK. Eq. 5 shows that this corresponds to a velocity noise $\sim$ 100 $\mathrm{km\,s^{-1}}$ if the cluster has central scattering optical depth $\tau_{e0} \sim 10^{-2}$. Thus individual rich clusters at $z \gtrsim 0.2$ could be detected (at the 3σ level) in the kinematic effect if their velocities exceed 300 $\mathrm{km\,s^{-1}}$. In practice this measurement is difficult, and the best errors in v_z obtained so far are about $\pm$700 $\mathrm{km\,s^{-1}}$ [17].

If this technique could be applied to a large sample of clusters, then a statistical measure of the distribution function of peculiar radial velocities, $f(v_z)$, will be possible. Such a measurement of the noisiness of the peculiar velocity field on the largest mass scales would be of considerable interest for studies of the development of large-scale structure. The distance independence of ΔT_{K0} should allow the evolution of $f(v_z)$ to be followed from the earliest redshift at which clusters develop atmospheres with sufficient optical depth. However many clusters will have to be measured to accuracies of $\pm$10 μK in ΔT_{K0} for this measurement to be made, unless the distribution function $f(v_z)$ is much wider than current structure formation models predict.

While measurement of ΔT_{K0} to $\pm$10 μK is not impossible, such a study would be subject to systematic errors in v_z arising from thermal substructure within the cluster. Simultaneous understanding of the kinematic and thermal effects is necessary to eliminate the τ_e factor in Eq. 5 and hence measure v_z. At a level of 100 $\mathrm{km\,s^{-1}}$, large-scale gas flows within the cluster connected with accretion events or cooling are likely, and will make it difficult to extract the cluster peculiar velocity even from excellent measurements of ΔT_{K0}.

2.3 Cluster baryon fraction

The thermal Sunyaev-Zel'dovich effect provides a more direct measure of the total hot electron content of a cluster than the X-ray flux, since ΔT_{T0} is proportional to n_e rather than n_e^2. Thus for an isothermal cluster, the total Sunyaev-Zel'dovich effect flux density is proportional to the total electron count in the cluster atmosphere [29]. Since the atmospheres of rich clusters of galaxies contain more mass than the member galaxies, the Sunyaev-Zel'dovich effect flux density can be converted into a good estimate of the total baryonic mass of a cluster if the electron temperature is known.

A comparison of this mass with the total mass obtained by gravitational shear mapping yields a clean measurement of the baryonic mass fraction of a cluster. Indeed, this comparison can be made at image level: a Sunyaev-Zel'dovich effect image is a measure of the baryonic column densities across the face of a cluster, while a gravitational shear map is a measure of total column densities. The ratio provides a map of the baryonic mass fraction in a cluster.

In the absence of a gravitational shear map, X-ray data can be used to calculate a mass for a cluster under the usual assumption that the atmosphere is in hydrostatic equilibrium. This can already be done for every cluster being used for Hubble diagram work. The current result for the baryonic mass fraction shows considerable scatter, but appears to be $(0.07 \pm 0.02)\,h_{100}^{-1}$ at all redshifts [29,36,15], which is consistent with the results from the power spectrum of the microwave background radiation [18] and primordial nucleosynthesis [31], which imply $\Omega_b \gtrsim 0.019 h_{100}^{-2}$ and $\Omega_m \lesssim 0.42 h_{100}^{-2}$, so that the mass fraction in clusters ($= \Omega_b/\Omega_m$) should exceed 0.045.

In the future this work could measure the baryonic mass as a function of radius within a cluster, and test whether clusters constitute a fair sample of the mass content of the Universe.

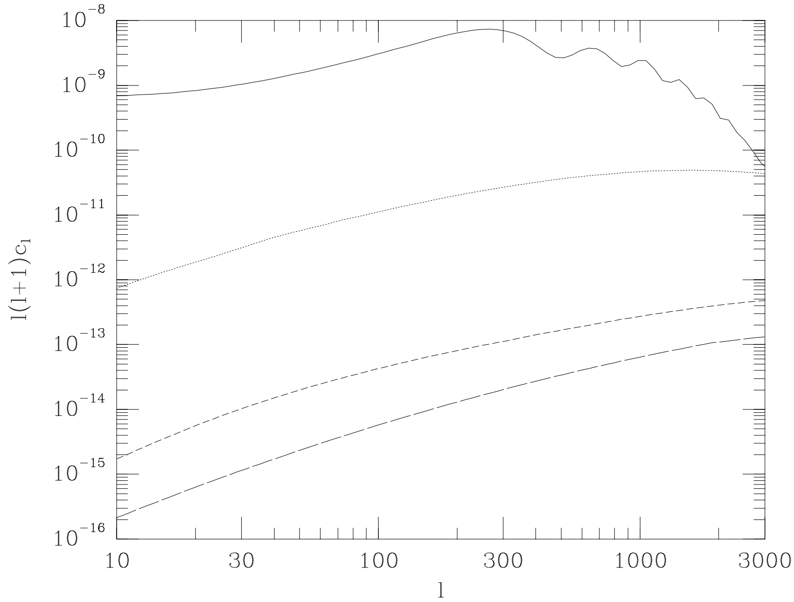

Figure 4: The power spectrum of primordial fluctuations in the microwave background radiation (upper line), the thermal Sunyaev-Zel'dovich effect (dotted line), the kinematic Sunyaev-Zel'dovich effect (dashed line), and the Rees-Sciama effect from moving clusters (long dashed line) for a ΛCDM cosmology.

Alternatively, it can be used to estimate the contribution of matter to Ω_0 [15].

2.4 Contributions to the power spectrum

Since clusters of galaxies can produce large microwave background structures in the form of the thermal and kinematic Sunyaev-Zel'dovich effects, and smaller effects of Rees-Sciama type, it is of interest to calculate their contribution to the power spectrum of brightness fluctuations. A representative simulation is shown in Fig. 4 [26]. It can be seen that cluster thermal Sunyaev-Zel'dovich effects make a small ($<$ 1 per cent) contribution to the power spectrum on large angular scales ($l \lesssim 500$), but become important on scales $l \gtrsim 3000$ (corresponding to angular scales of a few arcmin). The kinematic Sunyaev-Zel'dovich effect effect is never important, and neither is the effect of clusters moving across the line of sight [27].

The relative powers in primordial fluctuations and the thermal Sunyaev-Zel'dovich effect are a strong function of cosmology [27,2], depending particularly on the redshifts at which clusters acquire significant atmospheres, and the run of angular diameter distance with redshift. Measurements of the shape and amplitude of the power spectrum in the thermal Sunyaev-Zel'dovich effect at $l \gtrsim 1000$ will lead to constraints on models of cluster evolution and cosmological parameters [16,2]). This type of data will be provided by large-area blind surveys which are a major aim of the next generation of instruments (Sec. 5).

3 Observational status

The Sunyaev-Zel'dovich effects from X-ray luminous clusters are now relatively easily detected, provided that the observing technique is able to separate them from contaminating astrophysical or terrestrial signals. The major issue for many clusters is of distinguishing Sunyaev-Zel'dovich effects in the presence of non-thermal and often variable radio sources. Extended radio emission from cluster radio halo sources may also be a problem, but the steep radio spectra of such sources usually means that their contribution is small. If we seek to measure the typical 100 μK thermal Sunyaev-Zel'dovich effect signal to 5σ accuracy, we require good control of any extraneous signal at the level of 20 μK. For a cluster at $z \approx 0.2$, observed at 30 GHz, this corresponds to a contaminating flux density of less than about 20 μJy, perhaps distributed over a number of sources in or near the cluster. The difficulty in achieving this has meant that many of the cluster measurements to date have been for the high X-ray luminosity clusters where larger Sunyaev-Zel'dovich effects are expected.

The background of primordial structure in the microwave background radiation is also a confusing signal. The strongest thermal effects overwhelm the primordial signal, but typical cluster effects ≈ 100 μK may be confused by primordial fluctuations. In this case, multiple frequency bands must be observed, and spectral separation attempted. However, this is not possible for the kinematic effect (Sec. 1.2), and controlling contaminating sources at the 20 μJy level at several observing frequencies is even more difficult.

Single-dish radiometer measurements of the microwave background anisotropies towards clusters of galaxies, as used for the first detections of the thermal Sunyaev-Zel'dovich effect [3,4], are now relatively routine. While particularly well suited to the measurement of the integrated Sunyaev-Zel'dovich effect signal, and hence total baryonic content (Sec. 2.3), such observations suffer from systematic errors to do with imperfect subtraction of radio sources near the line of sight, atmospheric noise, and spillover, so particularly careful beam-switched and position-switched measurements are needed. Nevertheless, this remains one of the fastest methods of surveying a sample of clusters (e.g. Sec. 4 and [30]).

Many difficulties with systematic errors are avoided by observing with an interferometer, although care has to be taken that the shortest baselines adequately sample the Fourier components that contain most of the Sunyaev-Zel'dovich effect signal. Well-designed configurations have many short baselines for detecting the Sunyaev-Zel'dovich effect with good angular dynamic range, and a similar number of longer baselines which resolve out the cluster but provide contemporaneous measurements of small-scale contaminating radio sources, so that they can be subtracted accurately. Sources which are clustered, or which have complicated angular structures, are harder to remove, and imperfect removal will lead to an error in the measured Sunyaev-Zel'dovich effect. A major advantage of interferometers is that they produce an image of the cluster, although the raw visibilities are more useful for model fitting. The first interferometric detection of a cluster [21] has now been followed by extensive mapping campaigns, for example [8,9,39,22]. However, since the published images of cluster Sunyaev-Zel'dovich effects generally involve data with a small angular dynamic range, they must be treated cautiously since they show only the most compact parts of the Sunyaev-Zel'dovich effect structures.

Bolometric observations of clusters, using arrays on mm-wave telescopes, can also produce good results, although they have to contend with difficult atmospheric conditions. These observations are performed at high frequency (> 100 GHz), and so must be interpreted using the relativistic form of the spectrum, but by using multiple bands are able to separate the thermal Sunyaev-Zel'dovich effect from the kinematic effect or primordial fluctuations [17]. While confusion from radio sources is not likely to be significant at these bands, high-luminosity dusty galaxies do provide a possible contaminating signal which must also be separated from thermal or kinematic effects.

4 Samples and surveys

Current Sunyaev-Zel'dovich effect work is usually restricted to the hottest and most X-ray luminous clusters because of the expected correlation between $\Delta T_{\rm T0}$ and X-ray luminosity. This strategy may lead to misinterpretation of the detections, for example if they are used to estimate the value of the Hubble constant and intrinsic structural properties of clusters correlate with X-ray luminosity. Blind surveys for the thermal Sunyaev-Zel'dovich effect are now feasible using improved ground-based array receivers and bolometers, although considerable time is required

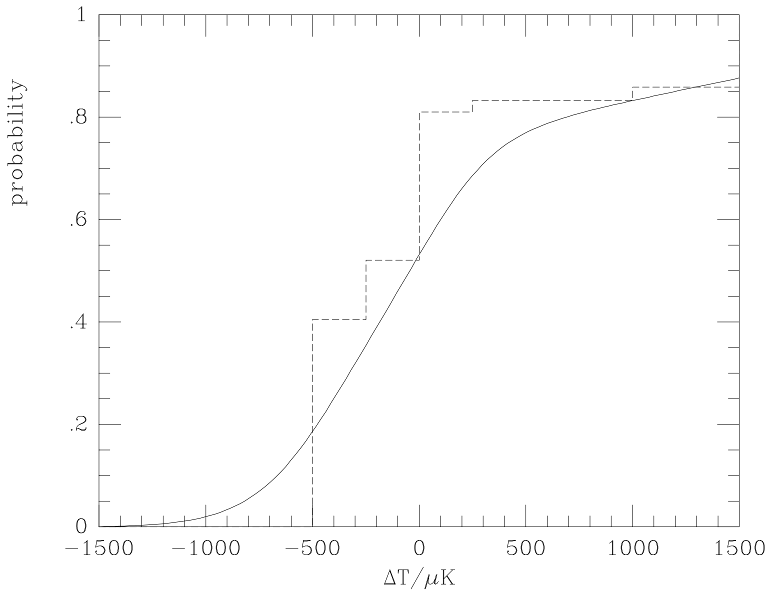

Figure 5: The distribution of brightness temperature differences for the combined north ecliptic pole and X-ray brightest Abell cluster samples, as measured with the OVRO 40-m telescope at 18 GHz. The solid line shows the probability density of the observed cumulative distribution, and the dotted line the underlying distribution expressed as a histogram.

to cover a useful sky area ($10-100$ deg^2). Balloon-borne telescopes are reaching the sensitivities and angular resolutions needed, and the MAP and Planck satellites will certainly have sufficient sensitivity to detect many effects.

As a trial, I have used the Owens Valley Radio Observatory (OVRO) 40-m telescope at 18 GHz to survey 26 Abell clusters in the north ecliptic pole region, and 12 Abell clusters which appear in the sample of X-ray brightest Abell clusters [12]. Although these surveys are not blind, nor selected in an orientation-independent fashion, the optical selection differs from the anecdotal selections that have characterised much Sunyaev-Zel'dovich effect work to date. No attempt was made to exclude clusters with known radio sources. The measured distribution of apparent Sunyaev-Zel'dovich effect signals from the clusters, and the inferred underlying distribution of intrinsic signals (contaminated by radio sources) is shown in Fig. 5.

Significant Sunyaev-Zel'dovich effects were detected from 8 clusters in the combined sample. A statistical analysis indicates that 40 per cent of all clusters of the type sampled show Sunyaev-Zel'dovich effects exceeding 100 μK, but that about 10 per cent are contaminated by bright radio sources (more than 1000 μK at 18 GHz). As might be expected, the X-ray brightest Abell clusters tend to show larger Sunyaev-Zel'dovich effects and stronger radio sources than clusters in the north ecliptic pole sample, but the same fraction of clusters with Sunyaev-Zel'dovich effects exceeding 100 μK is detected in each sample.

These results suggest that a larger targetted or blind survey is likely to be successful. A survey to a depth of ~ 30 μK should detect about 40 per cent of all rich clusters, even without simultaneous subtraction of contaminating radio sources.

5 New instruments and the future

Studies of the Sunyaev-Zel'dovich effect have begun to yield impressive results, but most of the work to date has used *ad hoc* instrumentation based on existing telescopes or instruments. The next phase of advances will come from dedicated microwave background facilities, designed with

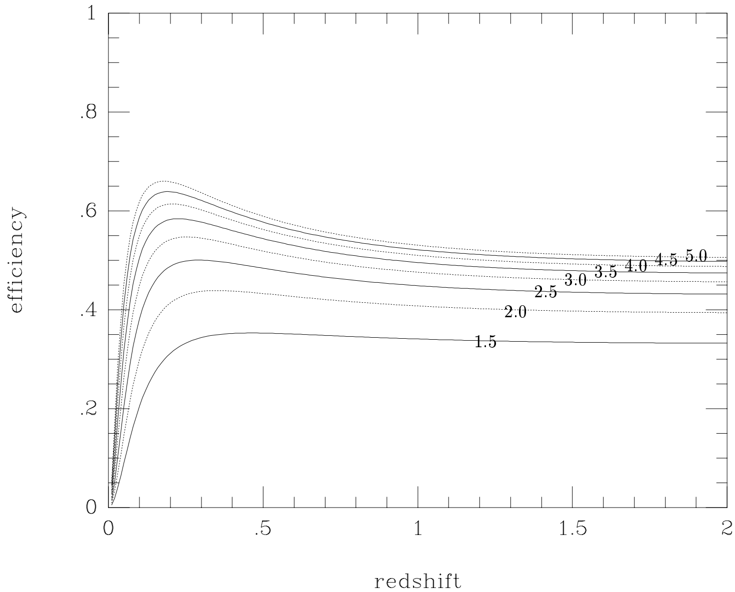

Figure 6: The efficiency of the planned OCRA focal plane array for observations of clusters of galaxies, as a function of redshift and beam separation. With a 2.5 arcmin separation between adjacent pairs of 1 arcmin beams, OCRA will have efficiency > 0.4 for the detection of cluster Sunyaev-Zel'dovich effects at $0.1 < z < 2$, and will permit sensitive blind surveys of large areas of sky.

a key science aim of excellent Sunyaev-Zel'dovich effect studies.

An ideal design for a dedicated facility would provide sensitivity of 20 μK or better on angular scales from 30 arcsec to 5 arcmin. The limitations imposed by the confusing source population, and by detector sensitivity, suggest operating bands between 30 and 345 GHz. Several bands, with matched angular resolution and spanning a factor of three or more in frequency, would be needed to distinguish between the thermal and kinematic effects.

MAP and Planck fulfil the requirements for raw sensitivity and frequency coverage, although not having the desired angular resolution. Despite the strong beam dilution that these missions will suffer for high-redshift clusters, their all-sky coverage will provide definitive surveys for strong Sunyaev-Zel'dovich effects. Planck might be expected to provide a sample of $\gtrsim 10^4$ clusters selected by their thermal Sunyaev-Zel'dovich effects alone [11]. Such large samples will allow strong statistical studies of the thermal effect, and should permit a measure of the dispersion of cluster peculiar velocities through the kinematic effect.

Because of the poor angular resolution of the satellite observatories, it will be ground-based telescopes, tailored for Sunyaev-Zel'dovich effect measurements, that will dominate the study of cluster evolution. Interferometers with baselines in the 100λ to 5000λ range and operating frequencies 30 to 230 GHz are expected to be particularly important, since they will be capable of high sensitivity and tuned angular resolution. They are therefore suitable for extremely deep survey work, and high-quality mapping of selected clusters. The AMiBA project [24] falls into this category: its 1 arcmin synthesized beam and 40 or 10 arcmin primary beam (depending on whether it is equipped with the 0.3 m or 1.2 m antennas) and 7 μK sensitivity in 1 hour of integration, make it an excellent match to the deep X-ray surveys possible with XMM, although its single operating band ($\sim$ 90 GHz) means that it cannot discriminate between the thermal Sunyaev-Zel'dovich effect and the kinematic effect or primordial structure in the microwave background radiation.

Fast, lower-resolution, surveys will be made with cm-wave receiver arrays (such as the One Centimetre Receiver Array, OCRA [7]) or mm-wave bolometer arrays (such as Bolocam [14]). It is now possible to make arrays of hundreds of receiver or bolometer elements which, mounted on a large telescope, achieve arcminute resolution and can survey to sensitivity $\sim 20\ \mu$K at a rate $\sim 1\ \mathrm{deg}^2$ per day. Arrays such as OCRA will be capable of detecting clusters at redshifts $z > 2$ with high efficiency (Fig. 6) and will provide the definitive counts of Sunyaev-Zel'dovich effect clusters that can set strong constraints on models of large scale structure formation. Arrays such as Bolocam, which can work at several mm-wave bands, will be able to discriminate between thermal and kinematic Sunyaev-Zel'dovich effects, and will also be able to survey many deg^2 of sky per year.

With the assistance of OCRA, AMiBA, Bolocam, Planck, MAP, and other instruments, we can expect, within the next decade, to have samples of $> 10^4$ clusters selected by their Sunyaev-Zel'dovich effects. For some hundreds of these clusters we should have high-quality interferometric maps suitable for detailed comparison with gravitational shear maps, and sensitive X-ray images with spatially resolved spectroscopy. The comparisons of the Sunyaev-Zel'dovich effect, X-ray, and gravitational shear images will yield information on the baryonic mass fraction in clusters, and its evolution with redshift, and provide a Hubble diagram populated with large numbers of clusters to $z \approx 1$. Further information on the development of large-scale structure should come from statistical measures of cluster peculiar velocities.

Studies of the Sunyaev-Zel'dovich effects have come a long way since the effects were first described in 1972, but much remains to be done.

References

1. P. Andreani *et al.*, *Ap. J.* **513**, 23 (1999).
2. F. Atrio-Barandela and J.P. Mücket, *Ap. J.* **515**, 465 (1999).
3. M. Birkinshaw, S.F. Gull and K.J.E. Northover, *Nature* **275**, 40 (1978).
4. M. Birkinshaw, S.F. Gull and H.E. Hardebeck, *Nature* **309**, 34 (1984).
5. M. Birkinshaw, J.P. Hughes and K.A. Arnaud, *Ap. J.* **379**, 466 (1991).
6. M. Birkinshaw, *Phys. Reports* **310**, 97 (1999).
7. I.W.A. Browne *et al.*, *Proc. SPIE* **4015**, 299 (2000).
8. J.E. Carlstrom, M. Joy and L. Grego, *Ap. J.* **456**, L75 (1996) and erratum *Ap. J.* **461**, L59 (1996).
9. J.E. Carlstrom *et al.*, *Physica Scripta* **T85**, 148 (2000).
10. R. Challinor, M. Ford and A.N. Lasenby, *Mon. Not. R. astr. Soc.* **312**, 159 (1999).
11. A.C. da Silva, B. Domingos, A.R. Liddle and P. Thomas, *Mon. Not. R. astr. Soc.* **317**, 37 (2000)
12. H. Ebeling *et al.*, *Mon. Not. R. astr. Soc.* **281**, 799 (1996) and erratum *Mon. Not. R. astr. Soc.* **283**, 1103 (1996).
13. D.J. Fixsen *et al.*, *Ap. J.* **473**, 576 (1996)
14. J. Glenn *et al.*, *Proc. SPIE* **3357**, 326 (1998).
15. L. Grego *et al.*, *Ap. J.* **552**, 2 (2001).
16. C. Hernández-Monteagudo, F. Atrio-Barandela and J.P. Mücket, *Ap. J.* **528**, L69 (2000).
17. W.L. Holzapfel *et al.*, *Ap. J.* **481**, 35 (1997).
18. W. Hu, M. Fukugita, M. Zaldarriaga and M. Tegmark, *Ap. J.* **549**, 669 (2001)
19. J.P. Hughes and M. Birkinshaw, *Ap. J.* **501**, 1 (1998).
20. N. Itoh, S. Nozawa and Y. Kohyama, *Ap. J.* **533**, 588 (1999).
21. M. Jones *et al.*, *Nature* **365**, 320 (1993).
22. M. Joy *et al.*, *Ap. J.* **551**, L1 (2001).
23. A.S. Kompaneets, *Zh. Eksp. Fiz. Teor.* **31**, 876 (1956).
24. K.Y. Lo *et al.*, in *Proc. IAU 201*, 31; eds. A. Lasenby and A. Wilkinson (Astronomical Society of the Pacific, 2001).

25. P.D. Mauskopf *et al.*, *Ap. J.* **538**, 505 (2000).
26. S.M. Molnar, PhD thesis, University of Bristol.
27. S.M. Molnar and M. Birkinshaw, *Ap. J.* **537**, 542 (2000).
28. S.M. Molnar, M. Birkinshaw and R.F. Mushotzky, *Ap. J.*, in preparation.
29. S.T. Myers *et al.*, *Ap. J.* **485**, 1 (1997).
30. S.T. Myers, in *Proc. IAU 201*, 19; eds. A. Lasenby and A. Wilkinson (Astronomical Society of the Pacific, 2001).
31. K.A. Olive, G. Steigman and T.P. Walker, *Phys. Reports* **333**, 389 (2000).
32. S.K. Patel *et al.*, *Ap. J.* **541**, 37 (2000).
33. D. Puy, L. Grenacher, P. Jetzer and M. Signore, *Astr. Ap.* **363**, 415 (2000) and erratum *Astr. Ap.* **368**, 15 (2001).
34. T. Pyne and M. Birkinshaw, *Ap. J.* **415**, 459 (1993).
35. M.J. Rees and D.W. Sciama, *Nature* **217**, 511 (1968).
36. E.D. Reese *et al.*, *Ap. J.* **533**, 38 (2000).
37. Y. Rephaeli, *Ap. J.* **445**, 33 (1995).
38. Y. Rephaeli, *Ann. Rev. Astr. Ap.* **33**, 541 (1996).
39. R. Saunders *et al.*, *Mon. Not. R. astr. Soc.*, in press (2001).
40. M. Sulkanen, *Ap. J.* **522**, 59 (1999)
41. R.A. Sunyaev and Ya.B. Zel'dovich, *Comm. Ap. Sp. Phys.* **4**, 173 (1972).