Secondary anisotropies of the CMB

Nabila Aghanim1, Subhabrata Majumdar2 and Joseph Silk3

1 Institut d’Astrophysique Spatiale (IAS), CNRS, Bât. 121, Université Paris-Sud, F-91405, Orsay, France
2 Department of Astronomy & Astrophysics, Tata Institute of Fundamental Research (TIFR), Homi Bhabha Road, Mumbai, India
3 Denys Wilkinson Building, University of Oxford, Keble Road, Oxford, OX1 3RH, UK

E-mail: nabila.aghanim@ias.u-psud.fr and subha@tifr.res.in

Received 20 June 2007, in final form 21 December 2007
Published 12 May 2008
Online at stacks.iop.org/RoPP/71/066902

Abstract

The Cosmic Microwave Background fluctuations provide a powerful probe of the dark ages of the universe through the imprint of the secondary anisotropies associated with the reionization of the universe and the growth of structure. We review the relation between the secondary anisotropies and the primary anisotropies that are directly generated by quantum fluctuations in the very early universe. The physics of secondary fluctuations is described, with emphasis on the ionization history and the evolution of structure. We discuss the different signatures arising from the secondary effects in terms of their induced temperature fluctuations, polarization and statistics. The secondary anisotropies are being actively pursued at present, and we review the future and current observational status.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

In the post WMAP era for cosmic microwave background (CMB) measurements and in preparation for the Planck and post-Planck era, attention is now shifting towards small angular scales of the order of a few arc-minutes or even smaller. At these scales, CMB temperature and polarization fluctuations are no longer dominated by primary effects at the surface of last scattering but rather by the so-called secondary effects induced by the interaction of CMB photons with the matter in the line of sight.

Current and future CMB experiments have two main goals: (i) measuring small angular scale temperature fluctuations (below a few arc minutes), and (ii) measuring...
the CMB polarization power spectrum. These goals are fundamental for our understanding of the universe. The small-scale anisotropies are directly related to the presence of structures in the universe whereas the two types of polarization (E and B-modes, which we discuss later) probe both the reionization of the universe, i.e. the formation of the first emitting objects, and the inflationary potential. In this review, we focus on the end of the dark ages and the astrophysical probes of reionization.

There have been rapid and important advances in the recent past. We already have on the one hand measurements, by ACBAR, CBI, BIMA, VSA, of the temperature power spectrum for 2000 < ℓ < 4000 with CBI and BIMA data showing an excess of power as compared with the predicted damping tail of the CMB (figure 1). On the other hand, DASI, Archeops, Boomerang, Maxipol, CBI, QUaD and WMAP have made direct measurements of the E-mode polarization. The situation will change even more in the near future with anticipated results from experiments currently taking data or in preparation (QUaD, BICEP, EBEX, CLOVER, QUIET, SPIDER, Planck).

All of this experimental activity is motivated by what now amounts to the standard model of cosmology. The CMB temperature fluctuations which are generated prior to decoupling are measured on scales from 90′ to several arc minutes. This has led to a model of precision cosmology. The basic infrastructure is the Friedmann–Lemaître model with zero curvature, a cosmological constant (or dark energy), a baryonic content and non-baryonic dominant cold dark matter component (with fractions given by the recent WMAP data (Spergel et al 2007)). Superimposed on the cosmological background are the primordial adiabatic density fluctuations, described by a nearly scale-invariant power spectrum $|\delta_k|^2 \propto k^{-1}$, at horizon crossing (in the comoving gauge), that generated the large-scale structure via gravitational instability of the cold matter component. However it has become increasingly apparent that to further refine these parameters, and to face the more intriguing challenge of establishing possible deviations from the concordance model one has to address the degeneracies between cosmological parameters with those from the secondary anisotropies as well as the extragalactic astrophysical foregrounds.

The primary CMB anisotropies are due to the gravitational redshift at large angular scales (Sachs and Wolfe 1967) and to the evolution of the primordial photon–baryon fluid evolution under gravity and Compton scattering at lower scales (Silk 1967, Peebles and Yu 1970, Sunyaev and Zel’dovich 1970) to which one adds photon diffusion damping at small scales (Silk 1967). Primary fluctuations have provided us with an unparalleled probe of the primordial density fluctuations that seeded large-scale structure formation. Indeed on large angular scales, greater than the angular scale subtended by the sound horizon at recombination, one can directly view the approximately scale-invariant spectrum of primordial quantum fluctuations.

On their way towards us, the photons interact with cosmic structures and their frequency, energy or direction of propagation are affected. These are the secondary effects that involve the density and velocity fields and incorporate Compton scattering off electrons. This review is devoted to a study of these secondary anisotropies.

The CMB photons we observe today have traversed the universe from the last scattering surface to us and have thus interacted with matter along their path through the universe. These interactions generate the secondary anisotropies that arise from two major families of interactions. The first family includes the gravitational effects (figure 2, panel a), including gravitational lensing, the Rees–Sciama effect (RS), moving lenses and decaying potentials usually referred to as the integrated Sachs–Wolfe effect (ISW). These anisotropies arise from the interactions of the photons with gravitational potential wells. The second family incorporates the effects of scattering between CMB photons and free electrons (figure 2, panel b) such as inverse Compton interaction (the Sunyaev–Zel’dovich (SZ) effect) and velocity-induced scatterings such as the Ostriker–Vishniac (OV) effect and inhomogeneous reionization.

We define secondary anisotropies in the CMB to include all temperature fluctuations generated since the epoch of matter–radiation decoupling at z ~ 1100. The following contributions may be distinguished.

(i) The integrated Sachs–Wolfe (ISW) effect is due to CMB photons traversing a time-varying linear gravitational potential. The relevant scale is the curvature scale freeze-out in concordance cosmology: the horizon at $1 + z \sim (\Omega_m/\Omega_\Lambda)^{1/3}$. This corresponds to an angular scale of about 10′.

(ii) The Rees–Sciama (RS) effect is due to CMB photons traversing a non-linear gravitational potential, usually associated with gravitational collapse. The relevant scales are those of galaxy clusters and superclusters, corresponding to angular scales of 5–10 arc minutes.
Polarization is primarily a secondary phenomenon. Discrete sources provide an appreciable foreground, Gravitational lensing of the CMB by intervening large-scale structure does not change the total power in fluctuations, but power is redistributed preferentially towards smaller scales. The effects are significant only below a few arc minutes. Its effects may be significant on large scales when the observable of interest is the $B$-mode power spectrum.

The Sunyaev–Zel’dovich (SZ) effect from hot gas in clusters is due to the first-order correction for energy transfer in Thomson scattering. It is on the scale of galaxy clusters and superclusters, although it may be produced on very small scales by the first stars in the universe. There is a spectral distortion, energy being transferred from photons in the Rayleigh–Jeans tail of the cosmic blackbody radiation to the Wien tail.

The kinetic Sunyaev–Zel’dovich effect is the Doppler effect due to the motion of hot gas in clusters that scatters the CMB. It causes no spectral distortion.

The Ostriker–Vishniac (OV linear) effect is also due to Doppler boosting. It is the linear version of the kinetic Sunyaev–Zel’dovich effect. It is proportional to the product of $\Delta n_{\gamma}$ and $\Delta v$. This is effective on the scale of order 1 arc minute.

Discrete sources provide an appreciable foreground, especially at lower frequencies for radio sources and high frequencies for infra-red and submillimetre sources.

Polarization is primarily a secondary phenomenon. The primary effect from last scattering is induced by out-of-phase velocity perturbations and provides evidence for the acausal nature of the fluctuations. The secondary polarization is associated with the reionization of the universe and on large scales corresponds to the horizon at reionization. Inhomogeneous reionization and scattering at the galaxy cluster scale leads to smaller scale polarization. The reionization signal is weak, amounting to no more than 10% of the primary signal.

$B$-mode polarization can be induced by shear perturbations. One source is gravitational lensing of primary CMB fluctuations. A second is relic gravity waves from inflation. These are pure $B$-modes, and fall off rapidly on scales smaller than the horizon at recombination, corresponding to about half a degree. Mixing by Faraday rotation in the intra-cluster medium also contributes to $B$-mode generation on small angular scales. The $B$-mode polarization amplitude only amounts to about a per cent of the primary signal, and its discovery will pose the major challenge for future experiments.

2. Reionization

2.1. Basics of physics

In dealing with secondary CMB anisotropies at reionization or arising from ionized structure like the hot gas in galaxy clusters, we are concerned with the scattering of the CMB photons by the plasma. A specific example of this is the Sunyaev–Zel’dovich effect, which is discussed in detail in section 5, where the intra-cluster gas up-scatters the cold microwave photons.

Since the secondary anisotropies are distortions of the CMB, which is the radiation field, we start by looking at the properties of an isotropic and thermal radiation background. The distribution function, $f_\alpha(r, \mathbf{p}_\alpha, t)$, of any radiation field is defined such that $f_\alpha d^3r d^3p_\alpha$ is the number of photons in the real-space volume $d^3r$ about $r$ and the momentum space volume $d^3p_\alpha$ about $\mathbf{p}_\alpha$ ($\nu$ being the frequency) at time $t$ with polarization $\alpha = 1, 2$. This distribution can be related to the photon occupation number, $n_\alpha(r, \mathbf{p}_\alpha, t)$, by

$$n_\alpha(r, \mathbf{p}_\alpha, t) = h^3 f_\alpha(r, \mathbf{p}_\alpha, t).$$

For polarization a description in terms of the pure polarization states pre-supposes fully polarized radiation. For CMB
radiation, the occupation number has a Planckian distribution given by
\[ n_\alpha = (e^{h\nu/k_BT_{\text{cmb}}} - 1)^{-1} \quad \text{for} \quad \alpha = 1, 2, \]
where \( T_{\text{cmb}} \) is the temperature of the CMB photons. The specific intensity of radiation is related to the distribution function by
\[ I_\nu(\hat{k}, r, t) = \sum_{\alpha=1}^2 \left( \frac{h^4}{c^6} \right) f_\alpha(r, p_\alpha, t). \]

Commonly, the specific intensity is described in units of brightness temperature, \( T_{\text{R-J}} \), which is defined as the temperature of the thermal radiation field which in the Rayleigh–Jeans (R–J) limit (i.e. low frequency) would have the same brightness as the radiation that is being described. In the R–J limit, the specific intensity reduces to
\[ T_{\text{R-J}}(\nu) = \frac{c^2 I_\nu}{2kB^2}. \]

Now let us consider the scattering between two species (namely, photons and electrons). For an ensemble of particles, if the motion of one particle is completely independent of all other particles, then to describe the state of the particles, one can specify the single particle distribution function given by
\[ f(r, p, t) \]d3r d3p, which is the probability of finding a single particle in the phase space volume \( d^3r \) \( d^3p \) around the point \( (r, p) \) at time \( t \). If there are no interactions between the particles and if they are non-relativistic, then the distribution obeys the Liouville equation
\[ \frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{p}{m} \frac{\partial f}{\partial r} + F(r, p, t). \frac{\partial f}{\partial p}, \]
where \( F \) is any force that may be present, and \( m \) is the mass of a particle, assumed to be the same for all particles.

In the case of inter-particle interactions being random and statistical in nature, one cannot describe the system by a particle, assumed to be the same for all particles.

For elastic collisions, one ends up with the Boltzmann equation
\[ \frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{p}{m} \frac{\partial f}{\partial r} + F(r, p, t). \frac{\partial f}{\partial p} \]

where the scattering solid angle \( d\Omega \) is determined by the conservation of momentum and energy and \( d\sigma \) is the scattering cross section. Moreover, the collisions take place between particles with momenta \( p \) and \( p_1 \) and produced particles with momenta \( p' \) and \( p'_1 \). The Boltzmann equation, being integro-differential, is difficult to solve analytically. However, it can be tackled under some approximations which can be made when \( p \) is close to \( p' \) and \( p_1 \) is close to \( p'_1 \). It is then possible to expand the right hand side of equation (6) in powers of \( \Delta p = p' - p \) and carry out the integral. The result can be expressed in terms of a Taylor series to give the Fokker–Planck equation. A simplification of the Fokker–Planck equation yields the Kompaneets equation, whose solution for the case of photon–electron collisions in astrophysical situations gives the Sunyaev–Zel’dovich distortion (section 5).

At matter–radiation decoupling, the free electrons are non-relativistic and the scattering between them and the photons is simply Thomson scattering. The incident electromagnetic radiation with linear polarization \( \epsilon_1 \) is scattered by an electron at rest in a radiation field of polarization \( \epsilon_\perp \) into a solid angle \( d\Omega \) with a probability:
\[ d\sigma = \frac{3\sigma_T}{8\pi} |\epsilon_1 \cdot \epsilon_\perp|^2 d\Omega. \]

In the plane perpendicular to the scattering direction there is no variation of the polarization. In the scattering plane, however, there is a net polarization. As a consequence, if the incident radiation propagating along the \( z \)-axis comes from two orthogonal directions there will be no polarization transmitted along the \( z \)-axis. Isotropic non-polarized incident radiation will induce the same identical polarization along \( x \)- and \( y \)-axis. If the incident radiation is anisotropic and quadrupolar the scattered radiation shows an excess of energy and thus a non-zero polarization oriented according to the quadrupole orientation. As a result, the Thomson scattering induces a linear polarization under the condition that the incident radiation has at least a quadrupolar geometry. In the cosmological context, anisotropies are induced by density perturbations and the velocity gradients are responsible for the quadrupole moment. We therefore expect a Thomson scattering-induced polarization for the primary anisotropies.

The polarization intensity is governed by the Boltzmann equation (Peebles and Yu 1970, Sunyaev and Zel’dovich 1972, Bond and Efstathiou 1984, Ma and Bertschinger 1995, Hu and White 1997). This yields
\[ \Delta Q_{\pm iU} + \frac{\dot{n}_e}{a} \Delta Q_{\pm iU} = n_e \sigma_T \epsilon_\perp(\eta) \times \left( \Delta Q_{\pm iU} + \sqrt{\frac{6\pi}{5}} \sum_{m=-2}^{m=2} \hat{r}_{2m} (\hat{n} \Pi^{(m)}) \right), \]
where \( Q \) and \( U \) are the two Stokes parameters, \( \eta = \int dt/a \) is the conformal time, \( a \) is the expansion factor, and \( \hat{n} \) the direction of photon propagation, \( \hat{r}_{2m} \) are the spherical harmonics with spin-weight \( s \), and \( \Pi^{(m)} \) is defined in terms of the quadrupole components of the temperature \((\Delta T^2_{(r, \eta)})\) and polarization perturbations
\[ \Pi^{(m)}(r, \eta) \equiv \Delta T^2_{(r, \eta)} + 12 \sqrt{6} \Delta \Delta^2_{(r, \eta)}, \]
where
\[ \sigma_T \] is the Thomson cross section, and \( n_e \) is the free electron number density which can be written as \( n_e(r, \eta) = \tilde{n}_e(\eta)(1 + \delta_e(r, \eta)) \), with \( \delta_e \) and \( \tilde{n}_e \) the fluctuation and the background of the electron number
density, respectively. The electron density fluctuations can be due to matter density perturbations or to spatial variations of the ionization fraction. Replacing $n_e(r, \eta)$ in equation (7) by its full expression allows us to separate first-order effects (proportional to $n_e$) from second-order effects (proportional to $\delta$). Finally, the polarization perturbations at present can be obtained by integrating the Boltzmann equation (7) along the line of sight.

Assuming that primary temperature fluctuations dominate over polarization perturbations, the polarization at reionization is due to coupling between the electron density and the quadrupole moment. The solution for a single Fourier mode, $\Delta Q_{\pm \ell k}(\hat{\eta}, \eta_0)$, of the Boltzmann equation (7) is then given (e.g. Ng and Ng 1996) by

$$\Delta Q_{\pm \ell k}(\hat{\eta}, \eta_0) = \sqrt{\frac{6\pi}{5}} \int_0^{\eta_0} d\eta \exp^{ik(n_\eta-\mu)} g(\eta)$$

$$\times \sum_m \pm 2Y_m^2(\hat{\eta}) X_m^{(m)}(\hat{k}, \eta),$$

where $X_m^{(m)}(\hat{k}, \eta)$ equals $\Pi^{(0)}(\hat{k}, \eta)$ for the first-order contribution and $S_m^{(m)}(\hat{k}, \eta) = \delta_0(\hat{k}, \eta) Q(\eta)$ for the second-order contribution, with $Q(\eta)$ being the radiation quadrupole. The visibility function $g(\eta)$:

$$g(\eta) = -\frac{d\tau}{d\eta} e^{-\tau(\eta)},$$

provides us with the probability that a photon had its last scattering at $n_\eta$ and reached the observer at the present time, $\eta_0$. In equation (10), $\tau(\eta) = \int_{\eta_0}^{\eta} d\eta' a(\eta_0)n_e\sigma_T$ is the optical depth and $\mu = \hat{k} \cdot \hat{n}$.

2.2. Constraints on reionization

As the CMB radiation possesses an rms primary quadrupole moment $Q_{\text{rms}}$, Thomson scattering between the CMB photons and free electrons generates linear polarization. This is the case at recombination but in particular it is true at reionization. Re-scattering of the CMB photons at reionization generates a new polarization anisotropy at larger angular scale because the horizon has grown to a much larger size by that epoch (Ng and Ng 1996). The location of the anisotropy (a bump), $\ell_{\text{peak}}$, relates to the horizon scale at the new ‘last scattering’ and thus depends on the ionization redshift $z_{\text{ion}}$. A fitting formula was given by Liu et al (2001):

$$\ell_{\text{peak}} = 0.74(1 + z_{\text{ion}})^{0.73} \Omega_m^{0.11}.$$

The height of the bump relates to the optical depth or in other words to the duration of the last scattering. Such a signature (bump at large scales) has first been observed by WMAP (Kogut et al 2003, Spergel et al 2003) by correlating the temperature and the polarization power spectra. The first year WMAP observations constrained the optical depth at reionization to a high value $\tau \sim 0.17$ and provided a simple model for the reionization, the ionization redshift was found to be $z_{\text{ion}} \sim 17$. The optical depth is degenerate with the tilt of the primordial power spectrum.

The WMAP first year result came as a surprise, in the context of earlier studies of the Gunn–Peterson effect inferred to be present in the most distant quasars at $z \sim 6$ (e.g. Fan et al 2003) and of the high temperature of the intergalactic medium at $z \sim 3$ (Theuns et al 2002). The situation was soon rectified with the WMAP 3 year data release (Spergel et al 2007). The improved data included an $E$-mode polarization map. The power spectrum is proportional to $\tau^2$ and the new constraints on polarization yielded an optical depth $\tau = 0.09 \pm 0.03$. Together with a better understanding of polarization foregrounds, the improved measurements enabled the degeneracy with the tilt to be reduced. The new tilt value of $n = 0.95 \pm 0.02$ lowers the small-scale power. Despite the reduced WMAP 3 year normalization, $\sigma_8 = 0.74 \pm 0.06$, the lower optical depth implies that the constraints on the possible sources of reionization remain essentially unchanged (Alvarez et al 2006). Precise measurements (cosmic variance-limited) of the $E$-mode polarization power spectrum will eventually allow us to phenomenologically reconstruct the reionization history (e.g. Hu and Okamoto 2004). This will help constrain the reionization models and enable us to explore the transition between partial and total reionization (e.g. Holder et al 2003).

Reionization must have occurred before $z \sim 6$ and the universe is now generally considered to have become reionized at a redshift between 7 and 20. The major question now is to identify the sources responsible for the reionization of the universe. The ionizing sources cannot be a population of normal galaxies or known quasars. Optical studies of the bright quasar luminosity function (Haiman et al 2001, Wyithe and Loeb 2003), as well the associated x-ray background (Dijkstra et al 2004) rule out the known quasar population as a reionization source. However miniquasars with correspondingly softer spectra could evade this constraint. Recourse must therefore be had to Population III stars or to miniquasars, both of which represent hypothetical but plausible populations of the first objects in the universe that are significant sources of ionizing photons.

We discuss theoretical issues in section 2.3.2. Here we ask whether one can observationally distinguish between the alternative hypotheses of stellar versus miniquasar ionization sources. The most promising techniques for probing reionization include 21 cm emission and absorption, Lyman-alpha absorption against high redshift quasars, and the statistics of Lyman-alpha emitters. One distinguishing feature is the intrinsic source spectrum, which is thermal for stars but with a cut-off at a few times the Lyman limit frequency, whereas it is a power-law for miniquasars with a spectrum that extends to higher energies with nearly equal logarithmic increments in energy per decade of frequency. One can also explore the evolution of the intergalactic medium during reionization through the study of the redshifted 21 cm hyperfine triplet-singlet level transition of the ground state of neutral hydrogen (H I). This line allows the detection of the H I gas in the early universe. It thus represents a unique way to map the spatial distribution of intergalactic hydrogen (e.g. Madau et al 1997, Ciardi and Madau 2003). Therefore it permits, in principle, a reconstruction of the reionization history as governed by the first luminous sources. The size of the ionized structures that could be detected depends on
the design of future radio telescopes. The forthcoming radio telescope, in the frequency range 80–180 MHz, LOW Frequency ARray (LOFAR)\(^4\) should have the sensitivity and resolution (~3 arc minutes) needed. Using cosmological radiative transfer numerical computations with an idealized LOFAR model, Valdes et al (2006) have simulated observations of the reionization signal for both early and late reionization scenarios. They show that if reionization occurs late, LOFAR will be able to detect individual HI structures on arc minute scales, emitting at a brightness temperature of \(\approx 35 \text{ mK}\) as a 3-\(\sigma\) signal in about 1000 h of observing time. Zaroubi and Silk (2005) showed that we could even distinguish between stars and miniquasars as sources of reionization since there is a dramatic difference between these two cases in the widths of the ionization fronts. Only the miniquasar model translates to scale-dependent 21 cm brightness temperature fluctuations that should be measurable by forthcoming LOFAR studies of the 21 cm angular correlation function (Zaroubi et al 2007). A hitherto undetected population of Lyman-alpha-emitting galaxies is a possible reionization source and may be visible during the pre-reionization era. One can hope to detect such objects to \(z \sim 10\) relative to the damping wing of the Gunn–Peterson absorption from the neutral intergalactic medium outside their \(\text{H}^\text{II}\) regions (Gnedin and Prada 2004).

### 2.3. Secondary anisotropies from reionization

When reionization is completed, the scattering between CMB photons and electrons moving along the line of sight generates secondary anisotropies through the Doppler effect. The amplitude of the fluctuations is given by

\[
\frac{\Delta T}{T}(\theta) = \int d\eta a(\eta)g(\eta)v_\eta(\theta, \eta) = -\int dt \sigma_T e^{-(\theta, t)} n_\eta(\theta, t) v_\eta(\theta, t)
\]

with \(v_\eta(\theta, t)\) the velocity along the line of sight (i.e. radial velocity). The electron density can be written as \(n_\eta(\theta, t) = n(\theta, t) \times \chi_\delta(\theta, t)\) the product of the matter density \(n(\theta, t)\) and the ionization fraction \(\chi_\delta(\theta, t)\). Both quantities vary around their average values. We can finally write the electron density as \(n_\eta(\theta, t) = \bar{n}_\eta(\theta, t)[1 + \delta + \bar{\delta}_\eta]\), with \(\bar{n}_\eta(\theta, t)\) the average number of electrons and \(\delta\) and \(\bar{\delta}_\eta\), the fluctuations of density field and ionization fraction respectively.

By replacing the electron density expression in equation (12), we can see that there is a first-order effect which suffers from cancellations, and two second-order effects which affect the probability of scattering of the CMB photons (e.g. Dodelson and Jubas 1995). They both generate secondary anisotropies. They are sometimes referred to as modulations of the Doppler effect (i.e. the velocity field) by density and ionization spatial variations.

#### 2.3.1. Density-induced anisotropies

These are produced when the ionization fraction is homogeneous, i.e. reionization is completed, and when the Doppler effect is modulated by spatial variations of the density field. The computation in the linear regime first appeared in Sunyaev and Zel’dovich (1970), was revisited by Vishniac and Ostriker (Ostriker and Vishniac 1986, Vishniac 1987), and is known as the Ostriker–Vishniac (OV) effect (see also Dodelson and Jubas (1995), Hu and White (1996), Jaffe and Kamionkowski (1998), Scannapieco (2000), Castro (2003)). The OV effect is a second order effect which weights as density squared (\(\propto \delta^2\)) and peaks at small angular scales (arc minutes) with an \(\text{rms}\) amplitude of the order of \(\mu\text{K}\). The computation of the density-induced anisotropies can be generalized to mildly non-linear and non-linear regimes. Because these regimes are difficult to describe analytically, a more appropriate tool is numerical simulations (e.g. Gnedin and Jaffe 2001, Zhang et al 2004). However, one can also use the halo model (see the review by Cooray and Sheth (2002)) to model analytically the mildly non-linear regime as done for example by Santos et al (2003) or Ma and Fry (2000, 2002). These studies showed that reionization-induced anisotropies are dominated by the OV effect at large angular scales. The contribution from non-linear effects only intervenes at smaller scales with amplitudes of \(\sim\) a few \(\mu\text{K}\) at \(\ell > 1000\) (figure 3). The non-linear contributions from collapsed and fully virialized structures such as galaxy clusters is historically known as the kinetic Sunyaev–Zel’dovich effect and will be discussed separately in section 5.

#### 2.3.2. Sources of patchy reionization

Before reionization is completed, ionized and neutral regions of the universe co-exist. This is called the inhomogeneous reionization (IHR) regime. In that case, the Doppler effect is modulated by variations of the ionization fraction \(\chi_e\). Aghanim et al (1996) computed the first estimate of the power spectrum of secondary anisotropies induced by early QSOs ionizing the universe from \(z = 12\) to complete reionization at \(z \sim 6\). They predicted a large contribution from such fluctuations whose amplitude and distribution depended on the number density of sources, their luminosities and their lifetimes. The model was revisited by Gruzinov and Hu (1998) and Knox et al (1998) who added the effect of spatial correlations between sources. The effect of an IHR on the CMB has been recently revisited in the context of a reionization scenario compatible with WMAP data. In this work, Santos et al (2003) found that secondary fluctuations from IHR dominates over density-modulated (OV) anisotropies. IHR is intimately linked to the nature of the ionizing sources, to their formation and evolution history and to their spatial distribution. As a result, predictions of the IHR effect span a large range of amplitudes and angular scales. A precise forecast of the effects of IHR on the CMB anisotropies requires a precise treatment of the reionization history of the universe together with the formation of the first ionizing sources including radiative transfer this is usually done using numerical simulations (e.g. Salvaterra et al 2005, Zahn et al 2005, Iliev et al 2007a).

Stellar ionizing sources have been studied by many authors (e.g. Cen 2003, Ciardi et al 2003, Haiman and Holder 2003, Wyithe and Loeb 2003, Sokasian et al 2003, Somerville and Livio 2003). The first cosmological 3D simulations incorporating radiative transfer of inhomogeneous reionization by protogalaxies were performed by Gnedin (2000). He found

---

\(^4\) www.lofar.org.
that reionization by protogalaxies spans the redshift range from \( z \sim 15 \) until \( z \sim 5 \). H\textsc{ii} regions gradually expand into the low-density intergalactic medium, leaving behind neutral high-density protrusions, and within the next 10% of the Hubble time, the \( \text{H\textsc{ii}} \) regions merge as the ionizing background rises by a large factor. The remaining dense neutral regions are gradually ionized. Sources as luminous as protogalaxies are too rare at these redshifts and recourse must be had to a population of galactic building blocks that are plausibly associated with dwarf galaxies or miniquasars.

Recent studies find in general that in order to provide enough ionizing flux at, or before, \( z = 15 \), for the usual scale-invariant primordial density perturbation power spectrum, one needs Population III stars, which provide about 20 times more ionizing photons per baryon than Population II (Bromm et al 2001, Schaerer 2002), or an IMF that is initially dominated by high-mass stars (Daigne et al 2004). This is in agreement with recent numerical simulations of the formation of the first stars from primordial molecular clouds suggesting that the first metal-free stars were predominantly very massive, \( m_{\text{star}} \gtrsim 100 M_\odot \) (Abel et al 2000, 2002, Bromm 2002). In general, possibly unrealistically high ionizing photon escape fractions are required for a stellar reionization source (Sokasian et al 2004).

Miniquasars have also been considered as a significant ionizing source (e.g. Oh 2001, Dijkstra et al 2004, Madau et al 2004, Ricotti and Ostriker 2004, Ricotti et al 2005). In view of the correlation between central black hole mass and spheroid velocity dispersion (Ferrarese and Merritt 2000, Gebhardt et al 2000), miniquasars are as plausible ionization sources as are Population III stars, whose nucleosynthetic traces have not yet been seen even in the most metal-poor halo stars nor in the high \( z \) Lyman-alpha forest. The observed correlation suggests that seed black holes must have been present before spheroid formation. Recent observations of a quasar host galaxy at \( z = 6.42 \) (Walter et al 2004) (and other AGN) suggest that super massive black holes were in place and predated the formation of the spheroid. Theory suggests that the seeds from which the super massive black-holes formed amounted to at least \( 1000 M_\odot \) and were in place before \( z \sim 10 \) (Madau and Rees 2001, Islam et al 2003, Volonteri et al 2003).

Decaying particles remain an option for reionization that is difficult to exclude. One recent example is provided by a decaying sterile neutrino whose decay products, relativistic electrons, result in partial ionization of the smooth gas (Hansen and Haiman 2004). A neutrino with a mass of \( \sim 200 \text{MeV} \) and a decay time of \( \sim 10^8 \text{yrs} \) can account for an electron scattering optical depth as high as \( 0.16 \) without violating existing astrophysical limits on the cosmic microwave and gamma-ray backgrounds. In this scenario, reionization is completed by subsequent star formation at lower redshifts. Dark matter annihilation during hydrogen recombination (at \( z \sim 1000 \)) can modify the recombination history of the Universe (Padmanabhan and Finkbeiner 2005). The residual ionization after recombination is enhanced. The surface of last scattering is broadened, partially suppressing the small-scale primary temperature fluctuations and enhancing the polarization fluctuations. In addition, the extended recombination phase weakens some of the cosmological parameter constraints, most notably on the scalar spectral index (Bean et al 2007).

2.4. Second-order polarization at reionization

In this section, we focus on the polarization signal at small scales induced at reionization by the coupling between primary quadrupole and fluctuations in the electron density at the new last scattering surface. These electron density

---

**Figure 3.** Left panel, from Zhang et al (2004): the Doppler effect induced temperature anisotropies (kinetic SZ) from numerical simulations. The results include non-linear regime and are obtained by assuming universe was reionized at \( z = 16.5 \) and remained ionized after that. The contribution from the linear regime, OV effect, (dashed line) is plotted for comparison, together with primary power spectrum and thermal SZ spectrum in the R–J region (see section 5). Reprinted with permission from Wiley-Blackwell Publishing. Right panel, from Santos et al (2003): Analytic computation of the secondary anisotropies produced by reionization. Top thick lines are for the inhomogeneous reionization-induced fluctuations. Bottom lines are for density-induced fluctuations where the solid thin line is for the linear OV effect and the dashed for the non-linear contribution to OV.
fluctuations can again have two origins: they are either due to density fluctuations in a homogeneously ionized universe (Seshadri and Subrahmanian 1998, Hu 2000a), or they can be associated with fluctuations of the ionizing fraction in an inhomogeneously ionized universe (Hu 2000a, Mortonson and Hu 2007). Additional polarization fluctuations from collapsed and virialized structures, such as galaxy clusters, will be treated separately in section 9.

The dominant second-order polarization fluctuations are due to coupling between primary quadrupole anisotropy $Q_{\text{rms}}$ and electron density fluctuations $\delta_e$ and are given by:

$$\Delta Q_{\text{alt}} \propto \int d\tau g(\tau) Q_{\text{rms}} \delta_e \propto \kappa Q_{\text{rms}} \delta_e.$$  \hspace{1cm} (13)

The quadrupole considered for generating polarization through Thomson scattering is in general the primary quadrupole. However in the rest frame of the scattering electrons, a quadrupole moment is also generated from quadratic Doppler effect (Sunyaev and Zel’dovich 1980). The amplitude of the polarization induced by coupling with electron density fluctuations in this case is smaller than those produced by the primary quadrupole as discussed by (Hu 2000a).

In all cases, the polarization signal from secondary anisotropies takes place at small angular scales, and has quite a small amplitude (figure 5). Liu et al (2001) found a typical amplitude of $\sim 10^{-2} \mu K$ in a pre-WMAP reionization model using numerical simulations to describe reionization (figure 5, left panel). More recently, this result was confirmed by Doré et al (2007) who also used numerical simulation compatible with current cosmological constraints. In a model reproducing the high optical depth suggested by 1st year WMAP observations, Santos et al (2003) generalized the computations to the non-linear regime using the halo model. They conclude that the modulation by ionizing fraction inhomogeneities, i.e. patchy reionization, dominates over the modulation by density fluctuations but the amplitudes remain small (figure 5, right panel).

3. Secondary effects from large-scale structure

3.1. The ISW effect

After decoupling, as the universe continues to expand, seeds of cosmic structures that scattered the CMB at the last scattering surface grow due to gravitational instability giving rise to large-scale structure. The gravitational potential evolves with evolution of the structure and the CMB photons are influenced once again by the change in the gravitational potential which they traverse. One can subdivide the gravitational secondaries broadly into two classes, one arising from the time-variable metric perturbations and the other due to gravitational lensing. The former is generally known as the integrated Sachs–Wolfe effect (Sachs and Wolfe 1967) in the linear region and goes by the names of Rees-Sciama effect and moving-halo effect (sometimes called the proper-motion effect) in the non-linear regime. The integrated Sachs–Wolfe (ISW) effect is further divided in the literature into an early ISW effect and a late ISW effect. The early ISW effect is only important around recombination when anisotropies can start growing and the radiation energy density is still dynamically important. The final anisotropy for these gravitational secondaries depends on the parameters of the background cosmology and is also tightly coupled to the clustering and the spatial and temporal evolution of the intervening structure.

In general, the temperature anisotropies, along any direction $n$, associated with the gravitational potential and proper motions can be written in the form (Sachs and Wolfe 1967, Hu 1994, see Martinez-González et al (1990), for a simple derivation)

$$\Delta T(n) / T = (\phi_{\text{rec}} - \phi_0) + \int_{\eta_0}^{\eta_{\text{rec}}} 2\phi \, d\eta,$$ \hspace{1cm} (14)

where $\eta_{\text{rec}}$ is the recombination time, $\eta_0$ the present time and $\phi$ is the gravitational potential. The first term represents the Sachs–Wolfe effect due to different gravitational potentials at recombination and present. The second term is the integrated ISW effect and depends on the time derivative of $\phi$ with respect
to the conformal time. The numerical factors multiplying each term in the equation depend on the choice of gauge and hence differ among various authors. A point to note is that the temperature change due to the gravitational redshifting of photons is frequency independent (in contrast to the SZ effect) and cannot be separated from the primary anisotropies using spectral information only.

The origin of the late ISW effect lies in the decay of the gravitational potential (Kofman and Starobinsky 1985, Mukhanov 1992, Kamionkowski and Spergel 1994). When the CMB photons pass through structures they are blue and redshifted when they respectively enter and exit the gravitational potential wells of the cosmic structures. The net effect is zero except in the case of a non-static universe. This can happen naturally in a low matter density universe and at the onset of dark energy (or spatial curvature) domination typically occurring at late times. The increased rate of expansion of the universe reduces the amplitude of gravitational potential. The differential redshift of the photons climbing in and out of the potential gives rise to a net temperature anisotropy. There is one qualitative difference between the early ISW and the late ISW effects. For the late ISW effect, the potential decays over a much longer time (of the order of the present day Hubble time). Thus the photons have to travel through multiple peaks and troughs of the perturbations and the chances of cancellation of the coherence in gravitational redshifts becomes greater leaving, little net perturbation to the photon temperature (Tuluie et al 1996).

To study the amplitude of the late ISW effect, we start by constructing its power spectrum. We expand the potential time derivative, \( \dot{\phi} \), in spherical basis to get the expression for the power spectrum in a flat universe as

\[
C_\ell = (4\pi)^2 \int \frac{k^2}{2} P_\phi(k) \frac{d}{d\eta} \left[ F(\eta) j_\ell(k r) \right] d\eta, \tag{15}
\]

where \( j_\ell(x) \) is the spherical Bessel function and \( r \) is the comoving distance between the photon at a conformal time \( \eta \) and the observer. \( F(k, \eta) = D/a \) is the growth rate of potential, where \( a \) is the expansion factor normalized to have \( a_0 = 1 \) and \( D \) is the linear growth factor. \( D(z) \) governs the growth of amplitude of density perturbation with time. It is simply equal to unity for \( \Omega_m = 1 \) flat universe. For universe with both matter and vacuum energy (i.e. \( \Omega_m, \Omega_\Lambda \)), one has accurate fitting formulae for the growth function (Carroll et al 1992).

The power spectrum of the potential, \( P_\phi \), is given by

\[
\langle \phi(\vec{k})\phi(\vec{k}') \rangle = (2\pi)^3 \delta_D(\vec{k} + \vec{k}') P_\phi(k). \tag{16}
\]

The main assumption in writing equation (15) is that in the linear regime the mode does not change in phase and so the change in its amplitude with time is simply described through the growth factor. The equation also ignores gravitational lensing to be discussed later.

In the small angular scale limit and under the assumptions that the correlations at a distance \( k^{-1} \) are slowly changing on a timescale \( (ck)^{-1} \), the radial integral in equation (15) can be broken into a product of the spherical Bessel function \( j_\ell(kr) \) and a slowly changing function of time. Taking out the slowly varying part outside the radial integral and using the large \( \ell \) approximation for the Bessel function, we can use the Limber approximation to get

\[
C_\ell = 32\pi^3 \int_0^{\infty} \frac{\hat{F}^2(k = \ell/r, \eta) P_\phi(k = \ell/r) d\eta}{r^2}. \tag{17}
\]
From the above equation, we can define the power spectrum of the potential term derivative as $P_{\phi}(k, \eta) = \dot{F}^2(k, \eta)P_{\phi}(k)$. Note that in the non-linear regime, the growth factor depends on the wavenumber $k$. However, equation (17) is still valid due to the slow time dependence of $P_{\phi}$.

To calculate $\phi$ as a function of time and scale, we relate the potential to the matter density via the Poisson equation. In $k$-space, this can be written as

$$\frac{\kappa}{\Omega_m} \left( \frac{\Delta_m}{a} \right) = \frac{\dot{\phi}}{a} = 0,$$

where $\Delta_m$ is the present day matter density parameter and $\delta$ is the matter density perturbation.

At this point it is straightforward to calculate the late ISW effect once we put in an appropriate expression for $F(\eta)$. After we do this, the first thing to notice is that for a flat matter-dominated $\Omega_m = 1$ universe, $D(\eta) \propto a(\eta)$, and so in the linear regime there is no ISW effect. Until non-linear effects are considered, the late ISW effect occurs only in open and lambda dominated universes. The linear ISW effect, the non-linear ISW effect and gravitational lensing effect are shown in figure 6.

The ISW effect is seen mainly in the lowest $\ell$-values in the power spectrum (Tulue et al 1996). Its importance comes from the fact that it is very sensitive to the amount, equation of state and clustering properties of the dark energy. Detection of such a signal is, however, limited by cosmic variance. The time evolution of the potential that gives rise to the ISW effect may also be probed by observations of large-scale structure. One can thus expect the ISW to be correlated with tracers of large-scale structure. This idea was first proposed by Crittenden and Turok (1996) and has been widely discussed in the literature (Kamionkowski 1996, Kinkhabwala and Kamionkowski 1999, Cooray 2002b, Afshordi 2004, Hu and Scarratt 2004). The ISW detection was attempted using the COBE data and radio sources or the x-ray background (Boughn et al 1998, Boughn and Crittenden 2002) without much success. The recent WMAP data (Spiegel et al 2003, 2007) provide for the first time all-sky high quality CMB measurements at large scales. Those data were used recently in combination with many large scale-structure tracers to detect the ISW signal. The correlations are presently performed mainly using galaxy surveys (2MASS, SDSS, NVSS, SDSS, APM, HEAO), see figure 7 for a recent result. However, despite numerous attempts both in real space (Diego et al 2003, Boughn and Crittenden 2004, Fosalba and Gaztanaga 2004, Hernandez-Monteagudo and Rubiono-Martin 2004, Nolta et al 2004, Afshordi et al 2005, Padmanabhan et al 2005, Gaztanaga et al 2006, Rassat et al 2007) or in the wavelet domain (e.g. Vietl et al 2006), there is very weak (or null) detection of the ISW effect through correlations. The ISW effect provides and offers a promising new way of inferring cosmological constraints (e.g. Cooray et al 2005, Pogosian 2006).

### 3.2. The Rees–Sciama and the moving-halo effects

As mentioned in the previous section, the ISW is linear in first-order perturbation theory. Cancellations of the ISW on small spatial scales leave second-order and non-linear effects. In hierarchical structure formation, the collapse of a structure can present a changing gravitational potential to passing photons. If the photon crossing time is a non-negligible fraction of the evolution time-scale, the net effect of the blue and redshift is different from zero and the path through the structures leaves a signature on the CMB. This was first pointed out by Rees and Sciama (1968) for evolving density profiles of any individual large-scale structures (see also Dyer 1976). This goes by the name of the Rees–Sciama (RS) effect. Subsequently, there have been many studies of the RS effect from isolated structures using the ‘Swiss–Cheese’ model (Kaiser 1982, Thompson and Vishniac 1987, Martinez–Gonzalez et al 1990, Chodorowski 1992, 1994), Tolman–Bondi solutions (Panek 1992, Lasenby et al 1999) and from clustering (Fang and Wu 1993). Calculations have also been done for non-linear regimes, both analytically (e.g. Cooray 2002a) and using numerical simulation (e.g. Seljak 1996a, Dahbrowski et al 1999). Much of this work was concerned with the possible contamination of primary anisotropies by the RS effect, since both are present at similar angular scales and cannot be distinguished using multi-frequency observations. As we shall see below, the RS effect is negligibly small at all angular scales (figure 6). The non-linear evolution of primordial scalar fields generates some vector and tensor modes, inducing, in turn, $B$ mode polarization anisotropies (Mollerach et al 2004). This secondary signal although smaller than the one associated with gravitational lensing effects (see section 4) might constitute a limiting background for future CMB polarization experiments.

For an isolated collapsed structure, there can be a change in the gravitational potential along the line of sight due to

---

**Figure 6.** From Cooray (2002a): The power spectrum of the ISW effect, including non-linear contribution. The Rees–Sciama effect shows the non-linear extension. The curve labeled ‘nl’ is the non-linear contribution while the curve labeled ‘lin’ is the contribution from the momentum field under the second-order perturbation theory. The primary anisotropy power spectrum accounting for the lensing effect is shown for comparison. Reprinted with permission. Copyright 2002 by the American Physical Society.
its bulk motion across the line of sight. For clusters of galaxies, this was first shown by Birkinshaw and Gull (1983); (see also Birkinshaw (1989) for a correction to the original results) as a way to measure their transverse velocities and is known as the ‘moving-halo’ effect. At the same time, these authors pointed out to the fact that CMB anisotropies should be gravitationally lensed by such moving halos. A similar proposal for temperature anisotropies due to the presence of cosmic string wakes was proposed by Kaiser and Stebbins (1984) (see also Stebbins 1988). The CMB photons entering ahead of a moving structure (galaxy cluster or super cluster) traversing the line of sight will be redshifted, while those entering the structure wake are blueshifted. The transverse motion induces a bipolar imprint in CMB whose amplitude is proportional to the velocity $v_t$ and to the depth of the potential $\phi$. The matter crossing time $d/v_t$ is taken as the evolution time $t_c$. From energy balance arguments, we get $\phi \sim v_t^2$. Thus, we have $t_c \sim d/\phi^{1/2}$. Putting all these together in equation (19), we can write
\[
\frac{\Delta T}{T} \sim \phi^{3/2} + v_t \phi.
\]

The above estimate is rather crude since we have used linear perturbation theory to describe non-linear regions. Moreover, it only applies to an isolated structure (for which the RS effect is independently treated from the velocity effect) and a proper justification can be only done using simulations where
the phase dependence of the growth factor is naturally taken into account. The non-linear ISW effect can also be calculated using the halo model which allows us to describe both the density and velocity fields of the large-scale structure in a coherent way (for details see Cooray and Sheth 2002). In such an approach, the basic idea is to take the time derivative of the Poisson equation (i.e. equation (18)) and using the continuity equation in $k$-space given by

$$\delta + i\vec{k} \cdot \vec{p} = 0,$$

(22)

where the momentum density field $\vec{p}(\vec{r}) = (1 + \delta)\vec{v}(\vec{r})$; one then obtains the following expression:

$$\phi = \frac{3}{2} \frac{\Omega_m}{a} \left( \frac{H_0}{k} \right)^2 \left( \frac{\dot{a}}{a} + i \vec{k} \cdot \vec{v} \right).$$

(23)

This relation connects the potential time derivative to the density and the momentum density. One can now obtain the power spectrum of $\phi$ by averaging over all the $k$-modes. It is easy to see from equation (23) that the power spectrum will involve correlation between density fields and time derivatives of density fields, as well as cross-correlation between density and momentum fields. Thus the general result has information about both the classical RS effect as well as the moving-halo effect. Numerical simulations capture an important point that is often missed in analytical perturbation theory calculations which is that in the strongly non-linear regime the power spectrum of $\phi$ is dominated by the momentum density.

The angular power spectrum including the non-linear ISW effect is shown in figure 6. For all cases, the temperature anisotropy $\Delta T / T$ is that in the strongly non-linear regime the power is often missed in analytical perturbation theory calculations about both the classical RS effect as well as the moving-halo and momentum fields. Thus the general result has information of density fields, as well as cross-correlation between density fields and time derivatives by other sources of secondary anisotropies such as the thermal effect becomes equal to the primary anisotropy at $\ell \approx 5000$. However, well before this equality is reached, it is overtaken by other sources of secondary anisotropies such as the thermal SZ effect.

4. Lensing of the CMB

4.1. Lensing by large-scale structure

As the CMB photons propagate from the last scattering surface, the intervening large-scale structure can not only generate new anisotropies (as shown in the last section, section 3) but can also gravitationally lens the primary anisotropies (Blanchard and Schneider 1987, Kashlinsky 1988, Linder 1988, Cayon et al 1993, Seljak 1996b, Metcalf and Silk 1997, Hu 2000b). For a detailed description of the process we refer the reader to a recent and thorough review by Lewis and Challinor (2006) (figure 9). Formally, lensing does not generate any new temperature anisotropies. There are indeed no new anisotropies generated if the gravitational potential is not evolving (see previous section for this case), whereas lensing occurs whenever there is a gravitational potential. Since lensing conserves surface brightness, the effect of gravitational lensing of the primary CMB can only be observed if the latter has anisotropies. In this case lensing magnifies certain patches in the sky and demagnifies others (figure 9). If the primary CMB were completely isotropic, one would not be able to differentiate between the different (de)magnifications. Gravitational lensing of the CMB has remained for a long time unobserved until very recently where Smith et al (2007) have indeed achieved a 3.4σ detection of this effect by cross correlating maps of the lensing potential inferred from WMAP temperature anisotropy with radio galaxy counts in the NVSS survey.

For gravitational lensing, the absolute value of the light deflection does not matter. What matters is the relative deflection of close-by light rays. If all the adjacent CMB photons are isotropically deflected, there would only be a coherent shift relative to the actual pattern. However, if they are not isotropically deflected, then the net dispersion of the deflection angles would change the intrinsic anisotropies at the relevant angular scales. The net result of gravitational lensing is to transfer power from larger scales (thus smoothing the initial peaks in the CMB power spectrum) to smaller scales. In the following, we detail the lensing effects on both the temperature anisotropies as well as on the polarized signal.

In order to understand the effects of gravitational lensing on the CMB power spectrum we have to write its effect on a single temperature and polarization anisotropy. Gravitational lensing modifies the CMB anisotropies, which are then measured as an angular displacement in the following way

$$T_{\text{obs}}(\theta) = T(\theta + \xi(\theta)), $$

(24)

where $\theta$ is the original undistorted angle. However, it is inaccurate to approximate the observed temperature by a truncated expansion in the deflection angle. This is only a good approximation on scales where the CMB is very over the relevant lensing deflection, i.e. on large scales, or very small (see Challinor and Lewis 2005). In the weak lensing limit, the regime of interest for CMB studies, we can use the perturbative approach and write the lensed CMB anisotropies as:

$$T_{\text{obs}}(\theta) \sim T(\theta) + \xi^\parallel(\theta) \cdot T_i + \frac{1}{2} \xi^\perp(\theta) \cdot T_{ij} $$

(25)

with $\xi$ given by

$$\xi_i(\theta) = -\frac{3}{2} \Omega_0 \int \frac{dz'}{H(z')} - \frac{1}{a} \frac{D_0(z')D_0(z)\varphi^{(1)}_{ij}(\theta, z')}{D_0(z)}, $$

(26)

$D_0(z,z')$ is the angular diameter distance between redshifts $z$ and $z'$ and $\varphi^{(1)}_{ij}(\theta, z)$ is the perpendicular gradient of the Newtonian potential in the direction $\theta$. In the same way, the modified polarization anisotropy is written as $P_{\text{obs}}(\theta) = P(\theta_{\text{obs}}) = P(\theta + \xi)$ which similarly gives second order:

$$P_{\text{obs}}(\theta) \sim P(\theta + \xi^\parallel(\theta) \cdot P_i + \frac{1}{2} \xi^\perp(\theta) \cdot P_{ij} $$

(27)
4.1.1. Lensed CMB power spectrum. In order to have an idea of the effect of gravitational lensing on the CMB power spectrum, we can use the perturbative approach to second order and obtain:

$$\langle T_{\text{obs}}(0) T_{\text{obs}}(\theta) \rangle = \langle T(0 + \xi(0)) T(\theta + \xi(\theta)) \rangle$$

$$= \langle T(0) T(\theta) + \langle T(0) \xi(\theta) \rangle \langle T(\theta) \xi(0) \rangle \rangle$$

$$+ \frac{1}{2} \langle \xi(0) \xi(\theta) \rangle \langle T_{ij} T_{ij}(\theta) \rangle$$

$$+ \frac{1}{2} \langle \xi(\theta) \xi(0) \rangle \langle T_{ij} T_{ij}(\theta) \rangle .$$

The lensed power spectrum $C^\text{lens}_\ell$ as a function of the unaltered power spectrum $C_\ell$ is obtained after Fourier transformation, which is directly associated with multipole decomposition. It is given by

$$C^\text{lens}_\ell = C_\ell \left[ 1 - \int \frac{d^2 k}{(2\pi)^2} \frac{(\ell \cdot k)^2}{k^4} \tilde{P}(k) \right]$$

$$+ \int \frac{d^2 k}{(2\pi)^2} \frac{(\ell \cdot k)^2 - k^4}{k^4} \tilde{P}(k) C_{|\ell-k|},$$

(28)

where $\tilde{P}$ is the projected power spectrum of the of the lensing convergence. A generalization of the computation (Hu 2000b) shows that the errors introduced by the flat sky approximation are negligible as shown in figure 8.

The expression of $C^\text{lens}_\ell$ clearly shows the effect of the gravitation lensing on the CMB:

- The first term is a renormalization due to the second-order effect introduced by the lenses in the perturbative formulae.
- The second term is a mode coupling due to the convolution of the unperturbed spectrum by the projected power spectrum $\tilde{P}$. Both cause the smoothing of the acoustic peaks at small scales.

Weak lensing does not introduce any characteristic scale in the CMB. Its effects are mostly noticeable at small scales where they modify the CMB damping tail through power transfer from large to small scales. This increase in power at large $\ell$s is significantly smaller than the modifications due to scattering effects (e.g. SZ effect). To identify the effects of gravitational lensing on the CMB it is necessary to explore not only the power spectrum but also higher order moments that possibly reveal the induced non-Gaussian signatures left by the non-linear coupling (Bernardeau 1997, 1998, Zaldarriaga 2000, Cooray 2002c, Kesden et al 2003). The projected mass distribution from $z \sim 1000$ to present and hence the lensing effect can be reconstructed in principle via maximum likelihood estimators or quadratic statistics in the temperature and polarization (e.g. Goldberg and Spergel 1999, Hu 2001, Okamoto and Hu 2003, Cooray and Kesden 2003, Hirata and Seljak 2003). However, as shown for example in Amblard et al (2004), lensing reconstruction is affected by other secondary effects indistinguishable from lensing such as the KSZ effect or residual foreground contaminations. In addition to providing the projected mass density, the weak lensing effect on the CMB is a potentially powerful tool to probe the neutrino mass and dark energy equation of state (e.g. Kaplinghat et al 2003, Lesgourgues et al 2006).

4.1.2. Effects of lensing on CMB polarization. A curl-free vector field does not remain scalar if it is distorted. Consequently in the case of CMB polarization vector field, we expect that gravitational lensing will mix the $E$ and $B$ components of the polarization. Computing equation (27) for $E$ and $B$ components implies second derivatives of a distorted field (e.g. Benabed et al 2001) and gives

$$\Delta E_{\text{obs}} = (1 - 2\xi) \Delta E + \xi \cdot \nabla (\Delta E)$$

$$- 2\gamma_j \Delta P_j + \nabla \gamma_j \cdot \nabla P_j$$

(29)

and

$$\Delta B_{\text{obs}} = (1 - 2\xi) \Delta B + \xi \cdot \nabla (\Delta B) - 2\epsilon^{ij} (\gamma_j \Delta P_i)$$

$$+ \nabla \gamma_i \cdot \nabla P_i ,$$

(30)
where $\Delta$ denotes the Laplacian, $\kappa$ and $\gamma$ are the convergence and shear of the gravitational field, and $\delta$ and $\epsilon$ the identity and the anti-symmetric tensors.

These two expressions already show the three major effects of gravitational lensing on polarization:

- A displacement shown by the term $(1 + \xi \cdot \nabla (\Delta E/B))$.
- An amplification expressed by $-2\chi (\Delta E/B)$ and controlled by the convergence of the lensing.
- A mixing term representing the coupling between the shear of the gravitational lensing and its gradient, with the polarization vector $P$.

From the previous set of equations we immediately note that if gravitational waves are negligible as it is the case for scalar density perturbations the equation for the $B$modes is written as:

$$\Delta B = -2\epsilon^{ij}(\gamma_i^j \Delta P_j + \nabla \gamma_i^j \nabla P_j).$$

This means that the convolution of the primary polarization, of initially scalar type (from Thomson scattering) with the shear of the gravitational lensing generates a $B$ mode polarization.

The observed or lensed power spectrum of the CMB polarization can be computed in the flat sky approximation (e.g. Zaldarriaga and Seljak 1997, 1998). It gives

$$(C^E_\ell)_{\text{obs}} = C^E_\ell [1 - l^2 \sigma] + \int \frac{d^2k}{(2\pi)^2} \frac{(\ell \cdot k)^2 - k^4}{2k^4} \tilde{P}(k) \times \left[ (C^{E}_{[\ell-k]} + C^{B}_{[\ell-k]}) + \cos(4\phi_{\ell-k}) (C^{E}_{[\ell-k]} - C^{B}_{[\ell-k]}) \right]$$

and

$$(C^B_\ell)_{\text{obs}} = C^B_\ell [1 - l^2 \sigma] + \int \frac{d^2k}{(2\pi)^2} \frac{(\ell \cdot k)^2 - k^4}{2k^4} \tilde{P}(k) \times \left[ (C^{E}_{[\ell-k]} + C^{B}_{[\ell-k]}) - \cos(4\phi_{\ell-k}) (C^{E}_{[\ell-k]} - C^{B}_{[\ell-k]}) \right].$$

where $\sigma = \int (d^2k/(2\pi)^2)((\ell \cdot k)^2/k^4)\tilde{P}(k)$. In the case of no or negligible primary $B$ mode polarization the first term in the $(C^E_\ell)_{\text{obs}}$ is neglected and we left with the coupling term. A comparison of the full sky approach (Hu 2000b) and a flat sky computation shows that the error introduced by the simplification are negligible.

Weak lensing induced-$B$ mode polarization, in addition to galactic emission, is one of the major contamination for the future post-Planck polarization-devoted CMB experiments (see figure 10), whose main scientific goal will be to detect inflation-generated gravitational waves. If the inflation potential is such that $V \lesssim 4 \times 10^{15}$ GeV, cleaning for lensing-induced polarization is a requirement. However for larger potentials, deep integrations of moderately large patches of the sky at low resolution should suffice to account for the noise induced by lensing (this is the case for, e.g. Planck, QUAD, BICEP, B-POL). However, if the inflation potential is much smaller, lensing-induced polarization will be the dominant foreground in the range $\ell \sim 50$–100, once the galactic contamination is removed (figure 10). The lensing signal can be separated from gravitational wave-$B$ modes using high order statistics as the case for the temperature anisotropies (e.g. Hu and Okamoto 2002, Kesden et al 2003, Kaplinghat et al 2003). The separation between primordial $B$ modes and lensing-induced $B$ polarization depends on the reconstruction of the lensing signal. For the secondary polarization signal to be reduced by a factor 10 in power spectrum amplitude, a full sky measure of temperature and polarization with a resolution of a few arc minutes and a noise of $1\mu$K-arc minutes is needed.

5. The Sunyaev–Zel’dovich effect

The best known and most studied secondary contribution due to cosmic structure is definitively the Sunyaev–Zel’dovich (SZ) effect (Sunyaev and Zel’dovich 1972, 1980; see also Rephaeli 1995, Birkinshaw 1999, Carlstrom et al 2002). It is caused by the inverse Compton interaction between the CMB photons and the free electrons of a hot ionized gas along the line of sight. The SZ effect can be broadly subdivided into: the thermal SZ (TSZ) effect where the photons are scattered by the random motion of the thermal electrons and the kinetic SZ (KSZ) effect which is due to the bulk motion of the electrons. In the former case, the resultant CMB photons have a unique spectral dependence, whereas the final spectrum remains Planckian in the case of KSZ effect since it only Doppler shifts the incident spectrum.

5.1. The thermal SZ effect

The TSZ effect describes comptonization, the process by which electron scattering brings a photon gas to equilibrium. The term Comptonization is used if the electrons are in thermal equilibrium at some temperature $T_e$, and if both
where $n(x, y) = (e^x - 1)^{-1}$, since in the absence of distortions (i.e., $y = 0$), the photon spectrum is a black body. If $x^2 y < 1$, then one can expand the exponential in equation (34) around $n(x, 0)$ for small $y$.

The final distortion can then be written as

$$
\Delta n \over n = {n(x, y) - n(x, 0) \over n(x, 0)} = y {xe^x \over (e^x - 1)} \left[ x \coth(x/2) - 4 \right].
$$

(35)

Since the change in radiation spectrum $\Delta I(x)$ at frequency $x$ is given by $\Delta I(x) = x^3 \Delta n(x) I_0$, where $I_0 = (2\pi/\nu)^2 \left( (k_B T_{\text{CMB}}/h\nu)^3 \right)$, we obtain the distinct spectral signature of the TSZ effect:

$$
\Delta I(x) = I_0 \left[ x \coth(x/2) - 4 \right].
$$

(36)

This signature assumes an incident Planckian spectrum and is valid in the single-scattering approximation. The temperature anisotropy due to inverse Compton scattering of CMB photons is given by

$$
\Delta T \over T = \Delta I(x) \ln I(x) \over I(x) \ln T = y \left[ x \coth(x/2) - 4 \right].
$$

(37)

In the non-relativistic limit, the frequency dependence of the distortion, shown in figure 11, is characterized by three distinct frequencies: $x_0 = 3.83$, when TSZ effect vanishes, $x_{\text{min}} = 2.26$ which gives the minimum decrement of the CMB intensity and $x_{\text{max}} = 6.51$ which gives the maximum distortion. In the Rayleigh–Jeans (R–J) limit (i.e., when $x \to 0$) and in the Wien region we have $\Delta T_{\text{CMB}}/T_{\text{CMB}} = -2y$ and $x^2 y$, respectively. Thus at low frequencies we would see an apparent decrease in the sky brightness of the CMB sky sometimes referred to as ‘holes in the sky’ (Birkinshaw and Gull 1978).

5.2. The kinetic SZ effect

The KSZ effect occurs, along with TSZ effect, if the scattering plasma has a bulk motion relative to the CMB. In that case, the CMB photons appear anisotropic in the reference frame of the scatterer and KSZ effect tends to isotropize the radiation. This, however, makes the radiation anisotropic in the reference frame of the observer, and there is a distortion towards the scatterer with amplitude proportional to the radial peculiar velocity $v_r$ of the scattering gas (Sunyaev and Zeldovich 1972, Rephaeli and Lahav 1991). To derive the expression for the CMB temperature distortion due to KSZ effect, one can either start with the Boltzmann equation (for example, see Nozawa et al. 1998) or use the radiative transfer equation (see Birkinshaw 1999).

In the limit of non-relativistic plasma moving with $v_r \ll c$, the change in the flux and temperature of the CMB in the direction of an object giving rise to KSZ effect is given by

$$
\Delta I(x) \over I(x) = -v_r \over c \left[ x e^x \over e^{x^2} - 1 \right].
$$

(38)

$$
\Delta T \over T = -{v_r \over c} \tau_{\text{clus}}.
$$

where $\tau_{\text{clus}}$ is the optical depth of the intra-cluster medium. Unlike TSZ effect, the spectral distribution of the kinetic SZE,
Figure 11. Frequency dependence of TSZ and KSZ effects. The thick line shows the frequency dependence of $\Delta T/T_{\text{cmb}}$ from TSZ effect, whereas the thin solid line shows the same for the change in spectral intensity $\Delta I (x)$. The thin dashed lines show the change in spectral intensity for KSZ effect, the upper one for an approaching source and the lower one for a receding source. The vertical dotted line shows the scaled frequency at which TSZ is zero and KSZ effect is maximum. Here, $x_0$, $I_0$ and $T_{\text{cmb}}$ are all scaled to unity.

Since typical peculiar velocities are around a few hundred kilometres per second and typical temperatures a few keV, the kinetic effect comes out to be at least an order of magnitude less than the thermal effect. However, there can be cases when the kinetic distortion is larger than the corresponding thermal distortion. This is so when the plasma is either too tenuous or relatively cool or both and the peculiar velocity is large. We discuss such scenarios in section 8.

5.3. Corrections to the SZ effect

When one deals with hotter and denser scattering media there are corrections to the simple expressions of TSZ effect (equation (37)) and KSZ effect (equation (39)) derived in the previous sections, which become important. These issues have been addressed in detail in many studies (Challinor and Lasenby 1998, 1999, Itoh et al 1998, Nozawa et al 1998, Molnar and Birkinshaw 1999, Nozawa et al 2000, Itoh et al 2001, Colafrancesco et al 2003, Shimon and Rephaeli 2004). The first step in most procedures, to calculate such corrections, is to expand the Kompaneets equation in a power series in $\theta_e = k_0 T_e / m_e c^2$. This can be done for many choices of parameters such as $p/m$, $v = E/p$. The convergence of such expansions which are, in general, asymptotic expansions in nature and converge slowly is then an important issue. For scattering media having high temperatures, corrections up to 3–5 orders in $\theta_e$ are sufficient and match fully relativistic numerical calculations well. The relativistic corrections modify the frequency dependence of the SZE (see figure 12, left panel). There is an associated correction to the cross-over frequency (Itoh et al 1998) well approximated by a linear function in $\theta_e$ for $k_0 T_e < 20$ keV and a quadratic function in $\theta_e$ up to 50 keV.

The numerical fit is given by:

$$x_0 = 3.830 \left(1 + 1.1674 \theta_e - 0.8533 \theta_e^2\right).$$

(40)

The relativistic corrections to the KSZ effect can be obtained by starting again from a generalized Kompaneets equation and applying a Lorentz boost to the direction of the peculiar velocity. The electron distribution functions are connected between the cluster frame and the CMB frame by Lorentz transformations. One expands the Kompaneets solution in powers of $\theta_e$ including cross terms such as and $\beta\theta_e^2$. The $\beta\theta_e^2$ term can give rise to a correction of the order of 10% for a typical electron temperature of 10 keV (figure 12). The other higher order terms lead to negligible corrections for temperatures of interest. The relativistic correction in the R–J limit is written as

$$\frac{\Delta n (X)}{n_0 (X)} \approx -2 y \beta_0 \left[1 - \frac{17}{10} \theta_e + \frac{123}{40} \theta_e^2 \right] + y \beta \left[1 - \frac{2}{5} \theta_e + \frac{13}{5} \theta_e^2 \right].$$

(41)

where we considered cluster moving along the line of sight such that $\beta = v_\parallel / c$ and neglected all $\beta^3$ and higher order terms. Similarly, there is a correction to the cross-over frequency which is very small.

Finally, one can relax the assumption of low optical depth and look at multiple scatterings of the incident photon spectrum. In general, the multiple scattering contribution is found out to be rather small compared with single scatterings. As an example, for a 15 keV cluster, the multiple scattering affects the final result by $\sim 0.3%$ in the Wien region and $\sim 0.03%$ in the R–J region. One can think of other effects that will add further corrections to the SZ distortion. The presence of magnetic fields would give rise to magnetic pressure which will add to the gas pressure in determining the hydrostatic equilibrium of the gas. Simple calculations that incorporate
this effect show a net decrease in the SZ effect distortion (Koch et al 2003, Zhang 2004). It has been proposed that the presence of magnetic fields would lead to an anisotropic velocity distribution such that one ends up with a two-temperature relativistic Maxwellian distribution of the thermal electrons. This can lead to a net enhancement of the SZ effect. Finally, it has been shown that the presence of a temperature gradient in the cluster temperature would lead to corrections to the electron momentum distribution thereby leading to corrections of the TSZ spectrum (Hattori and Okabe 2004). Unfortunately, the expected amplitude of the corrections is almost two orders of magnitude smaller than that of the TSZ effect. Presence of a significant amount of non-thermal population of electrons can also lead to deviations from the thermal SZ spectrum. A self-consistent treatment of several corrections to the thermal SZ effect in the presence of both thermal and non-thermal populations of electrons is given in Colafrancesco et al (2003) (see figure 12, right panel).

5.4. SZ observations

The first observations of the SZ effect were targeted in nature and looked at specific x-ray selected clusters. The SZ flux from these observations was used along with x-ray modeling of the clusters to estimate, in particular, the value of the Hubble constant (section 5.5). The major instruments responsible for such measurements were the OVRO 5 m telescope at 32 GHz, the IRAM 30 m telescope at 140 GHz, the Nobeyama 45 m telescope at 21 GHz, 43 GHz and 150 GHz, the SuZIE array at 140 GHz and the BOLOCAM 151 element array. Additionally, interferometers such as the BIMA array at 30 GHz, the Ryle Telescope at 15 GHz, CBI working between 25 and 36 GHz, ACBAR and AMI at 15 GHz have also been used (see Carlstrom et al (2002), Birkinshaw and Lancaster (2004) for a recent review). In figure 13, the SZ image of the cluster A3266 taken by ACBAR, with a beam of ~4.5 arc minutes, is shown for three frequencies 150, 220 and 275 GHz including the cross-over frequency of 217 GHz. Note that for 150 GHZ, the cluster SZ effect appears as a decrement while for 275 GHz it is an increment. Note also that the combination of the three frequencies permits to subtract the CMB contamination which remains important at the beam scale.

5.5. Hubble constant from the SZ effect

After the first TSZ observations of clusters started in the seventies, it was pointed out by Cavaliere et al (1977, 1979) and Silk and White (1978) that the distance to a cluster can be estimated from the SZ and x-ray cluster observations. If we put in plausible values for the matter and energy budget of the universe, then one can estimate the value of $H_0$ from this distance. This has been attempted or performed using SZ effect observations: single dish at radio wavelengths (Birkinshaw and Hughes 1994, Hughes and Birkinshaw 1998), millimetre wavelengths (Holzapfel et al 1997, Pointecouteau et al 1999), submillimetre wavelengths (Komatsu et al 1999) and also using interferometers (Jones et al 1993, Grego et al 2001, Reese et al 2002, Bonamente et al 2006).

The gist of the method can be understood simply: the SZ temperature decrement $\Delta T/T$ and the x-ray surface brightness $S_X(r)$ depend on the cluster gas structure differently; $\Delta T/T \propto n_e T_e L_{\text{cluster}}$ and $S_X(r) \propto n_e^2 T_e^{1/2} L_{\text{cluster}}$, where $n_e$, $T_e$ and $L_{\text{cluster}}$ are the characteristic density, temperature and extent of the cluster gas. Eliminating the gas density, one can obtain the cluster size in terms of the SZ and x-ray observables and the gas temperature. Once the angular size of the cluster $\theta_{\text{cluster}}$ is measured, we are able to obtain the cosmologically sensitive angular diameter distance $d_A = L_{\text{cluster}}/\theta_{\text{cluster}}$. For nearby ($z \ll 1$) clusters, $d_A$ can be approximated in terms of the
Figure 13. From Gomez et al (2003): mosaic of the 150, 220, 275 GHz, and CMB spectrally subtracted colourscale images of Abell 3266 (convolved with a Gaussian with FWHM $\sim$4.5 arc minutes) overlaid onto the ROSAT contours. The rms noise level of $\sim$25 $\mu$K beam$^{-1}$. Most of the CMB present in the 150, 220 and 275 GHz channels have been minimized in the CMB subtracted map. As expected, the SZ signal at 220 GHz is minimum.

deceleration parameter $q_0 = \Omega_0/2 - \Omega_\Lambda$ as

$$d_A = \frac{c}{H_0(1+z)} \left[ z - \frac{1 + q_0}{2} z^2 \right]. \quad (42)$$

The derived value of the Hubble constant depends on the other cosmological parameters, $\Omega_0$ and $\Omega_\Lambda$. As long as the redshift is less that 0.2, $d_A$ does not change significantly with small variation in the presently acceptable values of the cosmological parameters. For example, changing $q_0$ from 0 to 0.5 for clusters A665 or A2218 (having $z \sim 0.17–0.18$) leads to a change in $H_0$ by $\sim$3%. For a high redshift cluster, the changes in $H_0$ due to different cosmology can be higher by $\sim$5–10% (Kobayashi et al 1996, Reese et al 2002). More generally, the combination of SZ effect and x-ray observations can be used to probe dark energy. This is done especially when searching for violations of the duality relation between the angular diameter distance and the luminosity distance. The test of the reciprocity relation (between the source angular distance and the observer area distance) and the distance duality relation that derives from it was proposed as an additional test of dark energy (Bassett and Kunz 2004). While the reciprocity relation holds when
photon's follow null geodesic and that the geodesics deviation equation is valid, the distance duality relation will hold if the reciprocity relation is valid and the number of photons is conserved. Violations can thus occur if the number of photons is not conserved (e.g. in the case of absorption by dust) or if gravity is not described by a metric theory, i.e. photons do not follow null geodesic. Uzan et al (2004) tested for the duality relation. Using a data set of SZ effect and x-ray clusters, they found no significant departure from the reciprocity.

The procedure described here both for $H_0$ determinations or for distance duality tests, in general, interprets the SZ effect and x-ray observations with simple modeling of the cluster gas as spherical, unclumped and isothermal distribution such as the $\beta$-model (Cavaliere and Fusco-Femiano 1978). All clusters, however, show departures in error in the determination of $H_0$ as well as limitations to the distance duality test are observed in nature and come from the uncertainty in the cluster parameters such as its core radius and temperature, the intracluster parameters such as the central electron density $n_e,0$ and the central values of the SZ effect and x-ray measurement. Errors can also be due to contamination of the SZ effect from point sources or a poor knowledge of their spectra (e.g. Holder 2002, Aghanim et al 2005). There can also be systematic errors due to overall flux and brightness temperature calibration uncertainties and from improper subtraction of a zero level offset to the SZ data.

Without proper accounting for the many systematics present in the observations, the estimates made from this method are biased (Birkinshaw et al 1991, Inagaki et al 1995, Majumdar and Nath 2000, Reese et al 2002). They seem, in particular, to favor a low value of $H_0$ compared with other methods. However, with a more careful treatment of the systematics the discrepancies appear to be much less (Reese et al 2002, Ameglio et al 2006). Combining recent Chandra x-ray data with BIMA/OVRO SZ effect data on a large sample of clusters at $0.14 < z < 0.89$ yields $H_0 = 75 \pm 10$ km s$^{-1}$ Mpc$^{-1}$ for the standard LCDM model (Bonamente et al 2006). This result holds whether or not virial equilibrium is assumed and is consistent with the HST determination of $H_0$.

6. SZ cluster counts

Counts of galaxy clusters, detected through their SZ effect, can be used as major probe of cosmological as well as cluster properties. The frequency dependence of the SZ effect can be used to extract the clusters from a radio survey of the sky, making SZ cluster catalog possible. Once the clusters are detected, follow-up redshift measurements can be carried out to get the cluster redshift counts. The abundance of the clusters $N_M$, their redshift distribution $dN/dz$, as well as their clustering $\xi(r)$, are governed by the geometry of the universe and the power spectrum of the initial density perturbations. Gas physics related to cluster structure and evolution also enters through mapping of the cluster SZ flux relative to the true mass of the cluster.

6.1. Cluster mass and cluster mass-function

The fundamental quantity that goes into calculating the observed cluster counts is the cluster mass function, $dn/dM$, which predicts the multiplicity function of clusters having mass in the range $[M, M + \Delta M]$ at a given redshift for a choice of cosmology. One starts by calculating the variance of the linear density field, extrapolated to the redshift $z$ at which halos (i.e. clusters) are identified, after smoothing the mean density field with a spherical top-hat filter so as to enclose the mass $M$. This variance can be expressed in terms of the power spectrum $P(k)$ of the linear density field extrapolated to redshift zero as

$$\sigma^2(M,z) = \frac{D^2(z)}{2\pi^2} \int_0^\infty k^2 P(k) W^2(k,M) dk,$$

where $D(z)$ is the growth factor of linear perturbations normalized to unity at $z = 0$ and $W(k,M)$ is the Fourier transform of a real-space top-hat filter

$$W(k) = \frac{3}{(kR_h)^3} [\sin(kR_h) - (kR_h) \cos(kR_h)].$$

In equation (44), the mass $M$ is enclosed within a comoving radius $R_h$. An important cosmological parameter, related to the amplitude of fluctuations, is the mass variance at $R_h = 8h^{-1}$ Mpc denoted by $\sigma_8$.

One can define the mass function for a particular cosmological model in terms of the quantity $\ln(\sigma^{-1}(M,z)$ instead of $M$ as given by Jenkins et al (2001):

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

where $n(M,z)$ is the abundance of halos with mass less than $M$ at redshift $z$ and $\rho(z)$ is the mean density of the universe at that time. This implies that the mass function depends only on $\sigma(M,z)$, which in turn depends on the background cosmology. Further, the mass function is normalized to have $\int_0^\infty f(\sigma) d\ln(\sigma^{-1}) = 1$. In the following paragraphs we list the four most commonly used mass functions.

The first mass function was based on theoretical considerations (Press and Schecter 1974). The number density of clusters is derived by applying the statistics of peaks in a Gaussian random field (Bond et al 1991, Lacey and Cole 1993, Sheth et al 2001) to the initial density perturbations. It is assumed that the fraction of matter residing in objects of a mass $M$ can be traced to a portion of the initial density lying at an overdensity over a critical threshold value, $\delta_c$. This mass function is given by

$$f(\sigma)_{PS} = \sqrt{\frac{2}{\pi}} \frac{\delta_c}{\sigma} \exp \left[ -\frac{\delta_c^2}{2\sigma^2} \right].$$

where $\delta_c$ is the extrapolated linear overdensity of a spherical perturbation at the time of collapse and is a weak function of $\Omega_m$ and $\Omega_\Lambda$ (Eke et al 1996). Notice that the abundance of objects is exponentially sensitive to their masses at a particular redshift. Recently, other mass functions, mainly from fits to dark matter simulations, have been proposed in the literature.
For example, the mass function of Sheth and Tormen (1999) can be written as
\[ f(\sigma)_{ST} = A \sqrt{\frac{2a}{\pi}} \left[ 1 + \left( \frac{\sigma^2}{a^2} \right)^p \right] \frac{\delta_c^2}{\sigma} \exp \left[ - \frac{a\delta_c^2}{2\sigma^2} \right]. \] (47)

with \( A = 0.3222 \), \( a = 0.707 \) and \( p = 0.3 \) and the masses were estimated with a spherical overdensity algorithm, by computing the mass within the radius encompassing a mean overdensity equal to the virial one. Jenkins et al. (2001), using a much larger simulation and a friend-of-friend algorithm for cluster finding, proposed
\[ f(\sigma)_{Jenkins} = 0.315 \exp(-|\ln \sigma^{-1} + 0.61|^3). \] (48)

This fit, which is widely used, has a fractional accuracy better than 20\% for \(-1.2 \leq \ln \sigma^{-1} \leq 1\). Recently, Warren et al. (2006) have come up with an improved fit given by
\[ f(\sigma)_{Warren} = A(\sigma^{-a} + b) \exp (-c/\sigma^2) \] (49)

with \( A = 0.7234 \), \( a = 1.625 \), \( b = 0.2538 \) and \( c = 1.1982 \).

In spite of the progress in obtaining the mass function from dark matter simulations, there have been considerable differences between different simulation fits, especially at high redshifts. Precision cosmology with clusters may, ultimately, be limited by our understanding of these differences. The first step in this direction has already been taken recently by Lukic et al. (2007).

6.2. Cluster abundance and redshift distribution

Once the halo mass function and its evolution are quantified, the cluster redshift distribution, \( d^2N/dzd\Omega \) (i.e. the number of clusters per unit redshift per unit solid angle), can be estimated by multiplying the number density or abundance of clusters \( n(z) \) with the volume \( d^2V/dzd\Omega \) surveyed. This volume depends on the angular diameter distance \( d_A(z) \) and the Hubble parameter \( H(z) \) at that redshift. The abundance can then be easily computed by integrating the mass function over the limiting mass of a survey which depends on the selection function \( f_{survey}(M, z) \) of the survey,
\[ \frac{d^2N}{dz d\Omega} = \frac{c}{H(z)} d_A^2(z) \left( 1 + z \right)^2 \times \int_0^{\infty} f_{survey}(M, z) \frac{dN}{dM}(M, z) dM, \] (50)

where \( dN/dM \) is calculated using equation (45). The cosmological information contained in the observed cluster counts comes through its dependence on the expansion history of the universe and on the growth rate of structures (Haiman et al. 2001).

Once the mass function is written in the universal form (equation (45)), its evolution is completely governed by the growth factor \( D(z) \). The difference in the cluster counts with varying energy density of different components is explained as follows: small-amplitude density perturbations grow as \( D(z) = (1 + z)^{-1} \) when \( \Omega_m(z) \approx 1 \), but perturbation growth stalls at around \( z \sim (1/\Omega_m) - 1 \) when \( \Omega_m(z) \ll 1 \). For a fixed \( \Omega_m \) at \( z = 0 \), its behavior at a higher redshift depends on the Hubble expansion factor \( H(z) \) which in turn depends on the different energy densities including parameters for the amount of dark energy, \( \Omega_\Lambda \), and its equation of state \( w \). Dark energy starts to dominate the universe at a later time for large \( \Omega_\Lambda \) and a more negative value of \( w \). The different growth histories are manifest most strongly in high-mass clusters where the exponential dependence of the mass function on \( \sigma(M, z) = D(z)\sigma(M, 0) \) has a dramatic effect on the abundance of clusters. These cosmological sensitivities of cluster redshift distribution (see figure 14) have led clusters to be considered as probes of precision cosmology.

6.3. Precision cosmology with cluster counts

Galaxy cluster surveys of the nearby universe (Abell 1958) have been done for many decades. However, with the beginning of SZ surveys, ambitious plans to detect clusters in the faraway universe have started to take place. In the recent past, a 12 deg\(^2\) interferometric SZ survey of the high redshift universe (Holder et al. 2000) has been carried out. Future surveys covering many hundreds to thousands of degrees capable of detecting tens of thousands of clusters are already being attempted. The goal of all these surveys is to use the sensitivity of the cluster redshift distribution to the cosmological parameter as cosmological tools. In particular, it has been demonstrated by many that a suitably large cluster survey can be used as a strong discriminator of dark energy models (Haiman et al. 2001, Levine et al. 2002, Weller et al. 2002, Majumdar and Mohr 2003, 2004). All these authors forecast few per cent level constraints on cosmological parameters, including those of dark energy.

At this point, let us stress the fact that the exponential sensitivity of the cluster mass function to the cluster...
mass (equations (46)–(49)) is both the boon and bane for cosmological studies with clusters. Any systematic error in the estimation of cluster mass, including conversion between one definition of mass to another, is exponentially magnified by the steep slope of the mass function. Numerous techniques have been proposed in the last few years to tackle this complication, as described below.

In spite of the cosmological usefulness of cluster number counts, there are several theoretical and observational requirements needed to achieve precise cosmological constraints. These include advances in understanding the formation and evolution of cluster size halos (in practice, a tighter fitting for the mass function), a well understood/controlled cluster selection function as well as a robust observational proxy for the cluster mass. Additionally, one would need a follow-up programme to estimate the redshifts of these clusters. Finally, one needs to calibrate and control any instrumental systematics. For a SZ survey detecting 10 000–20 000 Poisson distributed clusters in roughly 10 redshift bins with equal weights, the statistical uncertainty is $\sim 10^{-7}\%$. This gives a ballpark number at which systematic uncertainties need to be controlled.

The first requirement depends on our ability to perform large simulations. This becomes more feasible with the increase in computational power. The second requirement translates into understanding the cluster selection function (the limiting mass and the completeness level of the survey) from as realistic as possible mock cluster catalogs. For a telescope beam larger than the cluster, a survey is limited by SZ flux. Moreover, since for SZ fluxes, the redshift dependence enters through the angular diameter distance rather than the luminosity distance, the mass-selection function is more uniform than that of x-ray surveys, except at nearby redshifts ($z < 0.2$) where the clusters will be partially resolved. The mass-selection function is directly linked with the cluster SZ-flux measurements through cluster scaling relations (Kaiser 1982, Borgani 2006). Finally, it has been pointed out (da Silva et al 2001, Motl et al 2005, Pfrommer et al 2006) that the SZ-flux of a galaxy cluster is a good proxy for its virial mass with a tight scatter in the scaling relation. The relation between the virial mass $M$ and the SZ-flux $f_{SZ}$ can be written as

$$f_{SZ}(z, \nu) d_A^2 = f(\nu) f_{gas} A_{SZ} M_{vir}^{\beta_{SZ}} E(z)^{2/3} F(\gamma, z),$$

where $H(z) = H_0 E(z)^2$, $f(\nu)$ is the SZ frequency dependence, $f_{gas}$ is the gas fraction of the cluster out to the virial radius, $A_{SZ}$ and $\beta_{SZ}$ are the amplitude and slope of the scaling relation and $F(\gamma, z)$ denotes any deviation from the standard evolution. For simplicity, we usually set $F(\gamma, z) = (1 + z)^\gamma$. The complexity of cluster structure is then encoded in the three parameters of the SZ flux–mass relation $A_{SZ}, \beta_{SZ}$ and $\gamma$. Uncertainties in the mass–observable relation can in principle be reduced by the use of ‘self-calibration’ techniques (Majumdar and Mohr 2004) where one uses additional observables such as the power spectrum of galaxy clusters (Lima and Hu 2004, Majumdar and Mohr 2004). Moreover, the distribution of clusters in observed flux at each redshift provides additional mass information (Hu 2003). Finally, direct mass measurements—through x-ray observations, optical spectroscopy or weak lensing—provide important additional leverage on cluster masses and hence on cosmology (Majumdar and Mohr 2003, 2004). Once parameter degeneracies are broken through use of multiple cluster information, it is possible to achieve strong cosmological constraints from upcoming cluster surveys. As an example, in figure 15, we show forecasts for $\Omega_m, \Omega_\Lambda$ and $w$ constraints from the South Pole Telescope (SPT) survey. It also shows that SZ clusters as probes are highly complementary to other experiments since each experiment

5 Only photometric redshifts are plausible for a sample of tens of thousands of clusters.

6 For application to actual data, see Gladders et al (2007).

7 The dark energy survey (DES) would be used to follow up SPT clusters to get their redshifts.
constrains a different combination of cosmological parameters and is subject to different systematics.

However, due to the very nature of the entanglement between gas physics and cosmology in using clusters as cosmological probes, there must be cosmology–cluster physics degeneracies (Majumdar and Mohr 2003, Majumdar and Cox 2007). These degeneracies can be broken by adding constraints from complementary information within cluster surveys (such as dN/dz + P_{\text{cluster}}(k)) or external information (e.g. mass follow-up).

7. The SZ power spectrum

The SZ effect from galaxy clusters is one of the major sources of secondary temperature anisotropies. A convenient way of describing its effect on the CMB is by computing its angular power spectrum. The fluctuations in the temperature background due to SZ effect from clusters of galaxies can be expressed in terms of correlations between the fluctuations along two lines of sight separated by an angle. The rms distortion can be quantified by the spherical harmonic coefficients $a_{lm}$, which is defined as $\Delta T(n) = T_0^{-1} \sum_{lm} a_{lm} Y_{lm}(n)$. The angular power spectrum of the SZ effect is then given by $C_\ell = \langle |a_{lm}|^2 \rangle$, the brackets denoting an ensemble average.

7.1. Modeling the SZ power spectrum

The SZ power spectrum can be derived from numerical simulation of structure formation and evolution or from analytical computations. In the first approach, hydrodynamical numerical simulations are the most appropriate way to describe both the dark matter of which the gravitational potential wells are made of and the baryonic gas which is responsible for the SZ effect. Various groups (Scaramella et al 1993, Refregier et al 2000, da Silva et al 2000, Seljak et al 2001, Refregier and Teyssier 2002, Zhang et al 2002) have performed such simulations and computed the associated power spectra. A compilation of the predictions from the different groups can be found in Springel et al (2001). The SZ power spectra computed from numerical simulations globally agree within a factor of 2. However, the results are quite sensitive to the resolution of the simulations which acts as an artificial damping effect at small angular scales and to the size of simulation which if not large enough underestimates the number of massive clusters and thus the power at large angular scales.

In the second approach, the SZ power spectrum can be computed analytically (Cole and Kaiser 1988, Makino and Suto 1993, Atrio-Barandela and Mucket 1999, Komatsu and Kitayama 1999, Molnar and Birkinshaw 2000, Cooray 2001, Majumdar 2001, Komatsu and Seljak 2002). The computation is based on two quantities:

- The cluster number counts or mass function dn/dM which provides us with the number of clusters of a given mass $M$ present at a redshift $z$ (see section 6).
- The cluster model or mass–SZ flux relation, i.e. its temperature and density profiles which give the spatial form factor of the associated SZ effect. To begin with, let us assume that the cluster cross-correlation function can be known (for details see Peebles (1980), Cole and Kaiser (1988)). The pattern of temperature anisotropy on the sky, induced by a population of clusters, is found by the convolution of the temperature anisotropy due to a single ‘template’ cluster of mass $M$ at redshift $z$ with the angular distribution of the clusters and then integrating over their mass and redshift distributions. If one takes an ensemble average and further assumes that $n(M, z)$ is constant over the range of comoving separations for which the cross-correlation function $\xi(M_1, M_2, z, \delta r)$ is non-zero, then the angular temperature power spectrum $C_\ell$ can be written as the sum of two terms, the ‘1-halo’ or the Poisson term and ‘2-halo’ or the clustering term, i.e.

$$C_\ell^{\text{total}} = C_\ell^{\text{Poisson}} + C_\ell^{\text{clustering}}. \tag{52}$$

The power spectrum for the Poisson distribution of objects can then be written as (Cole and Kaiser 1988)

$$C_\ell^{\text{Poisson}} = \int_0^{\ell_{\text{max}}} \frac{dV(z)}{dz} \int_{M_{\text{min}}}^{M_{\text{max}}} \frac{dn(M, z)}{dM} \left| y_e(M, z) \right|^2, \tag{53}$$

where $dV(z)/dz$ is the differential comoving volume, $dn/dM$ is the number density of objects and $y_e$ is the 2D Fourier transform of the projected Compton $y$-parameter. The mass range is chosen so as to cover from group scale to the largest cluster scales. Since these fluctuations occur at small angular scales, we can use the small angle approximation of the Legendre transformation and write $y_e$ as the angular Fourier transform of $y(\theta)$ as $y_e = 2\pi \int y(\theta) J_0[(\ell + 1/2)]\theta d\theta$ (Peebles 1980, Molnar and Birkinshaw 2000), where $J_0$ is the Bessel function of the first kind and zero order.

The clustering power spectra depend on lines of sight passing though an ensemble of correlated clusters. It can be estimated (Komatsu and Kitayama 1999) as

$$C_\ell^{\text{clustering}} = \int_0^{\ell_{\text{max}}} \frac{dV(z)}{dz} P(k) \times \left[ \int_{M_{\text{min}}}^{M_{\text{max}}} \frac{dn(M, z)}{dM} b(M, z) y_e(M, z) \right]^2. \tag{54}$$

where $b(M, z)$ is the time dependent linear bias factor. The matter power spectrum, $P(k, z)$, is related to the power spectrum of cluster correlation function $P_{\text{cluster}}(k, M_1, M_2, z)$ through the bias, i.e. $P_{\text{cluster}}(k, M_1, M_2, z) = b(M_1, z)b(M_2, z) D^2(z)P(k, z = 0)$. Convenient expressions for the bias at cluster scales are given by Sheth and Tormen (1999) and Jing (1999).

When one calculates the variance in beams of fixed size, the Poissonian model is a good approximation if the probability that a cluster has a neighbor is small inside the beam. This probability is the product of the number density and the volume integral of the cross-correlation function over the region probed by the beam. It can be shown that for beams comparable to the size of rich clusters ($R \sim 1.5h^{-1}\text{Mpc}$), the Poissonian approach is a valid approximation. Only for very large beams, the variance will increase due to positive correlation of the clusters. It can be shown that the Poisson power spectrum
dominates at all $\ell$ values greater than 100. However, by subtracting x-ray selected clusters of galaxies over a certain flux ($S_X > 10^{-12}$ erg cm$^{-2}$ s$^{-1}$), from both power spectra, one can make the clustering part of the spectrum dominant around $\ell \sim 700$ (Komatsu and Kitayama 1999, Majumdar 2001b).

To calculate the 2D profile of each cluster in the cluster ensemble, one needs a cluster gas density and temperature model. This can be either the empirical truncated $\beta$-profile (e.g. in Molnar and Birkinshaw (2000)), or it can be derived by solving the hydrostatic equilibrium equation of a gas within a NFW dark matter potential (e.g. in Komatsu and Seljak (2001)), or it can simply be obtained from fits to simulated cluster profiles as in Diego and Majumdar (2004).

7.2. Cosmological studies with SZ power spectrum

Both the volume element and the abundance of clusters depend on the cosmological model. As a consequence the power spectrum will also vary with cosmological parameters (Komatsu and Seljak (2002), see also figure 16 (left panel)). For example, SZ power spectrum is sensitive to the density parameter $\Omega_m$ which mainly affects the number of low and moderate redshift clusters, and the equation of state of the dark energy $w = P/\rho_{DE}$ which mainly affects the number of high redshift clusters. However in the range of allowed values for $\Omega_m$ and $w$ the effects are rather small. As for the other parameters the effects are quite negligible, except that of $\sigma_8$. The SZ power spectrum is strongly sensitive to the normalization of mass fluctuations at cluster scales, i.e $\sigma_8$. Numerical simulations of SZ clusters show $C_\ell \propto \sigma_8^2$. Using simple scaling analysis, one can show that $C_\ell \propto \sigma_8^{14/(3+n)}$ where $n$ is the effective spectral index of mass fluctuations at cluster scales. If the highest mass halos contribute to the power spectrum then $n \sim -1$ and we recover back the simulation results. Note that when smaller mass halos contribute significantly to the rms$^8$, then the effective $n$ is greater than $-1$ and $C_\ell$ can have a stronger dependence on $\sigma_8$. The cluster power spectrum depends, just like the mass function, on the cluster physics more specifically the mass-observable relation affects the power spectrum (figure 16, right panel).

Recently, the estimated power spectrum from the CBI-ACBAR-BIMA data interpreted as an SZ signal has been used to constrain the value of $\sigma_8$. The resulting value $\sigma_8 = 0.9$ was found to be higher than other estimates of $\sigma_8$ which are now converging to the range [0.7–0.8]. By forcing the clusters in the simulations or in the analytic calculations of SZ-$C_\ell$ to have scaling properties compatible with observed clusters, one can fit the CBI excess with a $\sigma_8$ closer to WMAP 3 year value. Note, however, that the excess $\sigma_8$ only appears if we believe that the resulting excess CMB fluctuations at the CBI scales is due to the SZ effect from galaxy clusters. It has indeed been pointed out (Toffolatti et al. 2005, Douspis et al. 2006) that SZ from clusters are only mildly needed if there are unremoved point sources below the detection limit. On a a more speculative note, non-Gaussianity in the primordial power spectrum can boost the SZ power spectrum at cluster scales up to CBI excess (Mathis et al. 2004, Sadeh et al. 2006).

7.3. Extraction of SZ effect from CMB data

The SZ contribution to the CMB power spectrum is dominant as compared with other sources of secondary anisotropies. It was shown in Douspis et al. (2006) that the SZ contribution, if coherently taken into account, affects the determination of cosmological parameters such as the normalization $\sigma_8$, the optical depth $\tau$ and the initial power spectrum index $n$. However, the TSZ contribution should in principle be

$^8$ For example, observed smaller mass clusters show increased entropy over simple self-similar predictions.
removed provided down to a given cluster mass from the measured power spectrum provided multi-frequency observations are conducted. To this aim many methods have been proposed and developed especially in the context of the Planck experiment (Sanz et al 2001, Vielva et al 2001, Diego et al 2002, Schäfer et al 2006a, 2006b, Pierpaoli et al 2006, Pires et al 2006). All of them benefit from the specific spectral signature of TSZ signal discussed in section 5. These methods also use additional spatial constraints based on adapted or matched filters, wavelets, etc. They are mainly aimed at providing us with SZ cluster catalogs that will be further used as cosmological probes. Consequently they help in cleaning out the primary CMB signal. In practice the sensitivity limits of the experiments, their frequency coverages as well as their finite beams prevent us from a complete cleaning of the TSZ effect. The TSZ effect is not the only source of power at small scales. One can also probe KSZ by its effect on the power spectrum at high $\ell$. From amplitude arguments, it is easy to find that the amplitude of the KSZ power spectrum is much smaller that that of the TSZ. However, patchy reionization has interesting implications for the KSZ effect power spectrum (Iliev et al 2007b). At fixed optical depth, patchy reionization approximately doubles the total KSZ power above $\ell = 3000$ by up to an order of magnitude compared with a uniform reionization scenario. The KSZ effect has the same dependence as the CMB anisotropies and the multi-frequency observations do not serve in removing this contribution. Optimized methods to extract a map of the KSZ temperature fluctuations from the CMB anisotropies can be developed (Forni and Aghanim 2004). Assuming that a map of Compton parameters for the TSZ effect can be obtained by multi-frequency separation, one can benefit from the spatial correlation between KSZ and TSZ effects which are due to the same galaxy clusters. This correlation allows us to use the TSZ map as a spatial template in order to mask, in the temperature anisotropy map encompassing both CMB and KSZ signals, the regions where clusters must have imprinted an SZ fluctuation. By further using the statistical properties of KSZ, which is a non-Gaussian, one can achieve good separation of the KSZ signal out of the primary CMB.

8. The SZ effect from other astrophysical sources

The SZ effect, as proposed originally by Sunyaev and Zel’dovich, represents the shift experienced by the CMB photons when they undergo inverse Compton interactions with the free electrons of the hot ionized intra-cluster gas. The SZ effect is thus historically associated with galaxy clusters. However, and more generally, inverse Compton scattering can take place in all astrophysical environments where both conditions of ionization and high temperature are fulfilled. As a result, the SZ effect was studied in a variety of redshift domains and astrophysical sources from early protogalaxies and galaxies to local galaxies such as M31, as suggested by Taylor et al (2003).

In the standard scenario of structure formation, baryonic matter is believed to lie in the potential wells formed by the DM. The baryonic matter in dynamical equilibrium with the DM can thus reach high temperatures at virialization. The baryonic gas can be heated by additional means (photoionization, mechanical heating, etc). Therefore, induced SZ anisotropies are expected to span a large range of amplitudes and angular scales.

At intermediate and large angular scales, a warm-hot gas ($T_o = 10^5–10^7$ keV) is likely to exist in the large structures of the cosmic web. This gas might account for a fraction of the missing baryons (e.g. Fukugita et al 1998). Due to its relatively large temperature, this warm medium is expected to exhibit an SZ signal but also an x-ray emission. Observing the SZ signal from this warm medium, which would contribute to the CMB signal at large scales, is certainly important from the CMB point of view for disentangling primary and secondary anisotropies. Observing the warm medium may also be a unique way to seek for, and find, the missing baryons in the universe. Several studies have aimed at studying and describing this contribution. Since it is associated with non-linear structures the ideal tool is numerical simulations (e.g. Springel et al 2001, da Silva et al 2001, Zhang et al 2002). The expected SZ signal was found to have quite low amplitudes, thus, making it difficult to detect directly. One way around the problem might be to target the correlations between the SZ signal and x-ray emission. Future experiments will tell us to what extent this will be possible.

At smaller angular scales, when structures collapse to form galaxies the temperature of the baryonic gas increases to larger values due to shock heating. Moreover, the thermal content of the collapsed structures increases due to feedback processes (star formation, AGN activity, etc). In both cases, an SZ signal is expected at the galaxy scale. Its detailed amplitude depends on the efficiency of the heating mechanisms and of the galactic environment (mainly the gas density). The predicted SZ signal has been computed in both cases for shock heating and feedback. In the first case, Valageas et al (2001) found that the contribution from collapsed objects dominates only at very small scales $\ell > 10^4$. The SZ signal from galaxies can be even larger if they host central supermassive black holes (BH). In that case the galactic outflows powered by the mechanical energy of accreting matter onto the BH induce an important SZ signal as large as the COBE limit (e.g. Natarajan and Sigurdsson 1999, Aghanim et al 2000, Lapi et al 2003). The SZ effect from quasar feedback (Chatterjee and Kosowsky 2007) has been predicted at the 1 $\mu$K level which is potentially detectable by ALMA. The feedback from stars and its associated SZ effect was also studied. Rosa-Gonzalez et al (2004) calculated the signal expected from star-formation activity during the formation of the most luminous bulges of normal galaxies. They found that the temperatures and densities were high enough to produce $y$ parameters comparable to those of galaxy clusters. The supernova driven galactic winds during the early stages of evolution of normal galaxies can also cause the distortion of the CMB radiation as proposed by (Majumdar and Nath 2001). Finally, the SZ effect can not only arise from forming or early formed objects but it can also be associated with relic objects such as hot regions, ‘cocoons’, around radio galaxies as proposed by Yamada et al (1999). In that case, the Compton parameter associated with...
the ensemble of cocoons was found to be of the same order as the COBE constraint. Radio galaxies in galaxy clusters can eject large quantities of energy which is either thermalized or remains as relativistic radio plasma (radio ghost) in the galaxy clusters. Ensslin and Kaiser (2000) estimated the Compton parameter from these two phases and found it too small to be detected. However a statistical estimate of the relativistic population in clusters can be envisaged by stacking the SZ signal from all the clusters detected by the future Planck satellite (Ensslin and Hansen 2004).

In some cases the ionized regions have too low temperatures or densities, or both, to exhibit significant Compton distortions. In these cases, the kinetic SZ effect (if the gas moves with respect to the CMB) becomes the dominant source of secondary anisotropies. This is, particularly, the case for the patchy reionization (Aghanim et al 1996). Here, the temperature of the ionized bubbles generated by emitting and ionizing sources is low (typically $\sim 10^4$ K) implying a negligible $y$ distortion, but the proper motion of the ionized bubbles causes significant KSZ fluctuations. The secondary anisotropies due to the Doppler effect have been the subject of quite a large number of studies, in the context of the reionization problem, especially in view of the first year WMAP constraints on optical depth (see section 2 and references therein for details). A large TSZ signal from the sources responsible for the reionization is however not excluded as yet. It might on the contrary contribute to the excess power measured by CBI and BIMA. An example for this is the case in which early massive stars have played an important role in the reionization history of the universe (Oh et al 2003).

9. Polarization from galaxy clusters

As shown by Sunyaev and Zel’dovich (1980) not only CMB intensity is altered by the presence of clusters, through the TSZ and KSZ effects, the CMB polarization is affected by the presence of galaxy clusters along the photon lines of sight. Polarization anisotropies are generated when the photons scatter off free electrons in the intra-cluster gas they are propagating through a plasma in the frequency $\nu$ of the propagation and the plasma frequency $\nu_p$. The polarization produced in this way is given by

$$P_{\nu} = \frac{1}{2} \frac{\nu^2 e^2 (e^2 + 1)}{2(e^2 - 1)^2} \left( \frac{\nu_p}{c} \right)^2 \tau.$$

The amplitude depends on the observed frequency. It is higher in the Wien part of the spectrum. The frequency-integrated polarization is simply proportional to $(\nu_p/c)^2 \tau$. The polarization vector is perpendicular to the plane formed by the velocity vector and the observing direction.

(ii) Double scattering-induced polarization. This process is also called the finite optical depth effect. When the CMB photons scatter off free electrons in the intra-cluster gas they can acquire anisotropies due to TSZ ($\propto \tau ((k_B T_e)/(m_e c^2))$) and KSZ ($\propto \tau (\nu_p/c)$) effects. A second scattering within the cluster induces polarization of the order of $(k_B T_e)/(m_e c^2) \tau^2$ and $(\nu_p/c) \tau^2$ without modifying the frequency dependences. This effect can be generalized to any other source of local anisotropy such as gravitational effects (moving gravitational lens effects were computed by Gibilisco (1997), bulk motions of moving gas clouds in the inner part of clusters (Diego et al 2003), collapse or expansion effects). The amplitude of the polarization depends on the gas distribution $\rho(r)$. For a homogeneous spherical cloud with gas density $\rho_0$, Sazonov and Sunyaev (1999) found that the maximal polarization degrees are $0.025 (\nu_p/c) r_0^3 g(x)$ and $0.014 ((k_B T_e)/(m_e c^2)) r_0^3 f(x)$, with $r_0 = 2 \sigma_T \rho_0$ and $g(x)$ and $f(x)$ the spectral dependences of the TSZ and KSZ, respectively. This results in a unique spectral signature displayed in figure 17.

(iii) Faraday rotation in magnetized intra-cluster medium. A radiation of frequency $\nu$ propagating through a plasma in the presence of a magnetic field $B$ along direction $n$ sees its linear polarization vector rotated by an angle $\Delta \phi$. This effect is the Faraday rotation (FR). Clusters show evidence for magnetic fields (e.g. Murgia et al (2004) and references therein). The CMB polarized radiation passing through magnetized galaxy clusters undergoes FR. This mixes the Stokes parameters $Q$ and $U$ and thus generates $B$ modes out of $E$ polarization. The
B-mode power spectrum depends on the details of the electron density distribution per individual cluster, on the mass function of clusters, as well as on the magnetic field distribution and evolution. Such a contribution was computed by Takada et al. (2001) and recently revisited by Tashiro et al. (2007). It is proportional to the product $B^2 \nu_{abc}$. At the frequencies typically used for CMB observations, the amplitude of the FR-induced $B$ polarization is small. However, the polarization observed at the cluster scale could be a powerful tool for probing the gas distribution (Ohno et al. 2003).

(iv) CMB quadrupole-induced polarization. The presence of a quadrupole component produces a polarization signal proportional to the cluster optical depth $\tau$. It has a maximum amplitude of $P_{\text{max}} \sim 2 \times 10^{-6} g(x) \tau$ which changes with frequency ($x = h_{\gamma 1} v / k_B T$) following $g(x) = x e^x / (e^x - 1)$. This effect should be the dominant source of polarization related to clusters. The primary quadrupole-induced polarization due to galaxy clusters and warm gas in filamentary structures has been investigated using hydrodynamical simulations (Liu et al. 2005). As shown in figure 18, this effect dominates at very small angular scales. On the larger scales the signal is dominated by the contributions from the filamentary structures. At the smallest scales it is the galaxy cluster contribution which dominates.

Galaxy clusters produce both $E$ and $B$ polarization but at a level that is much smaller than the primary signal. However, despite the relatively low signal amplitudes, the study of cluster-induced-polarization has gained new interest since it appears to be a potentially interesting cosmological probe (see e.g. Cooray and Baumann 2003). Cluster polarization measurement was proposed to probe the large-scale velocity fields through measuring the cluster transverse motions, as well as the galaxy cluster dynamics. However, the most promising application of cluster polarization measurement is associated with the dominant effect, the quadrupole-induced polarization. The polarized signal from numerous galaxy clusters has been suggested by Kamionkowski and Loeb (1997) as a method to probe the CMB quadrupole by reducing the cosmic variance uncertainty (see also Portsmouth (2004)). This in turn provides a new way to obtain accurate ISW measurements and probe the dark energy content of the universe (e.g. Cooray et al. 2004, Bunn 2006).

10. High-order statistics of secondary anisotropies

Most, and indeed the simplest, inflationary scenarios (e.g. Guth 1981, Sato 1981) predict that the temperature anisotropy field obeys Gaussian statistics to first order. In this case, the statistical distribution of the CMB anisotropies is fully described by its second moment, the power spectrum, given by

$$C_\ell = \sum_{m=\ell}^{\ell} |a_{\ell m}|^2,$$

where the $a_{\ell m}$ are the multipole coefficients in the spherical harmonic expansion $\Delta T(\theta)/T = \sum_{\ell, m} a_{\ell m} Y_{\ell m}(\theta)$. Nevertheless, other cosmological scenarios such as topological defects (e.g. Vilenkin and Shellard (1994), Landrieau and Shellard (2003) and references therein) and multi-field inflation (Gangui et al. 1994, Bernardeau and Uzan 2002) suggest departures from the Gaussian hypothesis. One example of a non-Gaussian model that lends itself to specific predictions is the so-called $\chi^2_{\mu}$ model for the multi-field inflaton potential (Koyama et al. 1999) implemented for $m = 2$ (2 fields) by Sadeh et al. (2006) to study SZ observables. The issue of testing, through higher order statistics, assumptions about the early universe is quite important and is becoming feasible in the context of present and future CMB experiments. Therefore, a battery of non-Gaussian (NG) estimators have been recently developed and tested. Among the most commonly used, there are the three and four-point functions and their harmonic analogues the bi- ($T_3$) and trispectrum ($T_4$) (e.g. Hu 2001, Komatsu and Spergel 2001, Kunz et al. 2001), respectively, given by

$$\langle T(\ell_1) T(\ell_2) T(\ell_3) \rangle_\ell = (2\pi)^3 \delta(\ell_{123}) T_3(\ell_1, \ell_2, \ell_3)$$

and

$$\langle T(\ell_1) T(\ell_2) T(\ell_3) T(\ell_4) \rangle_\ell = (2\pi)^4 \delta(\ell_{1234}) T_4(\ell_1, \ell_2, \ell_3, \ell_4),$$

where $\ell_{123} = \ell_1 + \ell_2 + \ell_3$ and $\ell_{1234} = \ell_1 + \ell_2 + \ell_3 + \ell_4$. Also widely used are the higher order moments of the wavelet coefficients (skewness and excess kurtosis) (e.g. Pando et al. 1998, Forni and Aghanim 1999, Hobson et al. 1999, Barreiro and Hobson 2001). The wavelet analysis, in the dyadic wavelet transform scheme, decomposes a signal $s$ in a series of the form

$$s(j) = \sum_k c_{j,k} (\phi_\lambda)_{j,l}(k) + \sum_k \sum_{j=1}^J (\psi_\lambda)_{j,l}(k) w_{j,k},$$

where $J$ is the number of decomposition levels, $w_{j,k}$ the wavelet (or detail) coefficients at position $k$ and scale $j$ (the indexing is such that $j = 1$ corresponds to the finest scale, i.e. highest-frequencies) and $c_{j,k}$ is a coarse or smooth version of the original signal $s$. Other tests of non-Gaussianity are the global Minkowski functionals such as the total area of excursion

---

**Figure 18.** From Liu et al. (2005): the power spectra of the polarized $E$ and $B$ modes from primary and secondary anisotropies due to quadrupole-induced interactions.
regions enclosed by isotemperature contours or total contour length and genus (e.g. Gott et al 1990, Schmalzing and Gorski 1998, Novikov et al 2000, Shandarin 2002), the harmonic space analysis (Hansen et al 2002), the peak statistics (e.g. Bond and Efstathiou 1987, Vittorio and Juszkiewicz 1987).

Not only departures from the simplest inflation model can generate non-Gaussian signatures. Systematic effects, point source and foreground-induced non-Gaussianities will inevitably arise at small scales from the secondary anisotropies, either through the non-linear growth of fluctuations or through the interactions of CMB photons with the potential wells or ionized matter along their lines of sight. Besides, the study of secondary non-Gaussianities is very interesting on its own, since it is related to the cosmic structures, their evolution and spatial distribution; it is also of great importance in order to go beyond the information provided by the power spectrum of the CMB primary anisotropies. In this context, the higher order statistics of the secondary anisotropies are used to predict the NG signatures of non-primordial origin in the CMB and to better detect and understand the structures themselves.

The NG signatures are of particular importance in the case of gravitational lensing since they allow us in theory to reconstruct the mass distribution of the lenses. As a matter of fact, the deflection angles are small compared with the scale of structures and the lensing effect is hardly seen directly in a CMB map. The effects on the power spectrum are generally small and sub-dominant, and the two-point-statistics is thus not sufficient to allow for the reconstruction of the mass distribution of lenses. To better identify the effects of gravitational lensing on the CMB, one has to consider the induced NG signatures, naturally arising from the second-order effects in the anisotropies (correlations between large-scale gradients and small-scale generated power), through higher order statistics. The week lensing of primary anisotropies produces a four-point signature (e.g. Bernardeau 1997, Zaldarriaga 2000, Kesden et al 2003). Quadratic statistics (such as the power spectrum of the squared temperature maps) permit us to recover the information in the four-point function about the mass distribution of the lens field (e.g. Zaldarriaga and Seljak 1999, Hu 2001, Takada 2001, Hu and Okamoto 2002, Cooray and Kesden 2003). These methods use lensed anisotropy maps only, or combine them with the polarization field (especially the B field) which is less contaminated by the primary signal. In all cases, mapping the lens, and thus dark matter, distribution requires high resolution, high signal-to-noise maps of the CMB temperature fluctuations and polarization fields.

The interactions of CMB photons with the free electrons along their lines of sight, through Compton or Doppler effects, also produce secondary NG signatures. These sources of secondary anisotropies are expected to be important; it was therefore necessary to forecast their NG signal and study its detectability. This was done mainly for the SZ effect and the inhomogeneous reionization through the trispectrum (Cooray 2001) and through the high-order moments of the wavelet coefficients (Aghanim and Forni 1999). For the SZ thermal effect, Cooray (2001) gave the expression for the trispectrum of the TSZ effect in the flat sky approximation

$$\langle y(\ell_1) y(\ell_2) y(\ell_3) y(\ell_4) \rangle = (2\pi)^4 \delta(\ell_{1234}) T^{TSZ}(\ell_1, \ell_2, \ell_3, \ell_4),$$

where c designates the connected part and $T^{TSZ}$ is given by

$$T^{TSZ} = \int dr \frac{W^{TSZ}(r)^4}{d_k^4} T_k \left( \frac{\ell_1}{d_k}, \frac{\ell_2}{d_k}, \frac{\ell_3}{d_k}, \frac{\ell_4}{d_k}, r \right).$$

where $T_k$ is the pressure trispectrum. The weight function $W^{TSZ}(r) = -2((k_0 \sigma_T^2)/(a r^2 m_e c^2))$ is given in the Rayleigh–Jeans regime. In all cases, the signal from the SZ effect dominates at small angular scales. For all vector-like fields such as the Ostriker–Vishniac effect, but also the mildly non-linear regime probed by the KSZ effect for large-scale structures, even moments were shown to dominate over odd moments, making the trispectrum a more sensitive estimator of non-Gaussianity than the bispectrum (Castro 2003). As a result while the bispectrum is most likely undetectable by future CMB experiments, the trispectrum of the OV effect could be measured by Planck or by arc-minute scale interferometric experiments.

The NG signatures associated with the secondary effects can be used to probe and trace the matter distribution; they can also be used as additional constraints to separate the secondary effects from the primary CMB signal (e.g. Forni and Aghanim 2004). However in all these cases, this signal at small angular scales is the sum of the CMB anisotropies and all the secondary contributions. This makes it harder to disentangle them and requires the use of the polarization field or cross-correlations and couplings between components.

Figure 19 shows an example of SZ angular power spectra for the galaxy cluster contribution, demonstrating that the non-Gaussian $\chi^2_{an}$ model can have a substantial

---

**Figure 19.** From Sadeh et al (2007): SZ power spectrum obtained for a $\Lambda$CDM model with $\sigma_8 = 0.74$ (solid line) and 0.8 (thick dashed–dotted line), an early dark energy model (dashed line) and a non-Gaussian $\chi^2_{an}$ model (dashed–dotted line). Shaded areas correspond to WMAP 1σ error on $\sigma_8$. The data points are those of BIMA (diamonds), ACBAR ($\times$ symbols) and CBI (+). Reprinted with permission from Wiley-Blackwell Publishing.
impact for $10^3 < \ell < 10^4$, especially in the case of WMAP 3 year normalization ($\sigma_8 = 0.74$). Also shown are examples of Gaussian models with different normalizations ($\sigma_8 = 0.74, 0.8$) and an early dark energy model (Bartelmann et al. 2006).

11. Discussion and conclusion

Secondary effects induce temperature and polarization anisotropies. These additional anisotropies contribute to the CMB signal and modify (at certain scales) both its amplitude and its statistical character. Such a contribution was not actually important within the context of the first generation of CMB experiments (e.g. COBE). Already now with WMAP, and even more so with the future Planck satellite, the aim of measuring the CMB signal with fundamental instrumental noise limits forces us to investigate with extreme care the effects of the secondary anisotropies. They might constitute in some cases an important limiting factor on the scientific objectives of future CMB studies such as constraining the energy scale of inflation through the $B$-mode polarization induced by the stochastic gravitational wave background, or constraining the inflationary field through the statistical nature of the temperature anisotropies. Present day CMB experiments are now reaching sensitivities and angular resolutions such that secondary effects can no longer be neglected. This is the case for lensing by large-scale structures which convert the $E$-mode primary polarization into a $B$-mode secondary contribution and is by far the largest contaminant. This is also the case for the example of the SZ effect from galaxy clusters which could explain the excess of power at high multipoles measured by ACBAR, BIMA and CBI.

The SZ effect is more than a nuisance factor to cosmological parameter extraction. It is a potentially powerful tool for cosmology. SZ cluster counts can be used to probe the cosmological model and put constraints on the nature of dark energy. In combination with other observations, especially at x-ray energies, it allows us to measure cosmological parameters such as the Hubble constant and the cluster gas mass fraction (e.g. Grego et al. 2001). The SZ effect can also be used to characterize the clusters themselves as it potentially can measure their radial peculiar velocities (Lamarre et al. 1998, Benson et al. 2003). The non-relativistic corrections to the SZ effect can also be used to measure the gas temperature directly for massive clusters. This might be an important issue for future SZ surveys for which x-ray counterparts will not be available. The spectral signature of the SZ effect can in principle probe the electron gas distribution and constrain any non-thermal electron population in the intra-cluster medium. Moreover multi-frequency SZ measurements might provide a novel way of constraining the CMB temperature and its evolution with redshift (Battistelli et al. 2003, Horellou et al. 2005).

To achieve these goals, high precision measurements of the SZ effect will be needed over large areas of the sky. This requires a new generation of SZ telescopes that are already being built or designed. Following OVRO and BIMA, the Sunyaev–Zel’dovich Array (SZA) which consists of eight 3.5 m telescopes is operating at 26–36 GHz and 85–115 GHz. The SZA along with BIMA/OVRO forms the Combined ARray for Millimeter Astronomy (CARMA) telescope which aims to provide high resolution, detailed imaging of SZ clusters.

Several other telescopes are being commissioned (e.g. AMI), in 2007, with the aim of surveying large areas of the SZ sky for blind detection of clusters. These deep SZ cluster surveys will be performed by the AMIBA interferometer, the South Pole Telescope (SPT), the Atacama Cosmology Telescope (ACT) and the Atacama Pathfinder Experiment (APEX). Moreover, the Planck satellite scheduled to be launched in 2008 will detect thousands of SZ clusters over the whole sky.

Although both ACT and SPT are primarily designed for SZ cluster detection, the predicted KSZ signal, at a few arc minute scales, induced by the reionization might be sufficiently strong to be detected by these upcoming experiments (figure 20). These high $\ell$ measurements of the reionization-induced temperature anisotropies will however not suffice to unravel the ionization history. Polarization measurements at low $\ell$ are the optimal CMB tool to achieve Figure 20. From Iliev et al. (2007b): observability of the Doppler induced temperature anisotropies: the sky power spectrum of the reionization signal (black, solid; from two simulations) with the forecast error bars for ACT (left) and SPT (right). The primary CMB anisotropy (dotted) and the post-reionization KSZ signal (dashed) are also shown and are added to the noise error bars for the reionization signal. The TSZ component is assumed to be completely separated.
this. In the near future, Planck will provide all-sky E-mode polarization maps and will be sensitive to partial or double reionization models at the per cent level. In principle this could help discriminate between different models with identical optical depths (Kaplinghat et al. 2003), subject to our being able to understand, model and remove the relevant galactic foregrounds. In combination with a low frequency radio interferometer such as LOFAR and eventually SKA it should be possible to probe the onset of the reionization and the end of the dark ages by anti-correlating 21 cm emission and CMB temperature fluctuations (Alvarez et al. 2006).

The next generation of polarization-optimized satellites, such as B-POL or EPIC, is being designed to measure the primary B-modes from inflation. These experiments will inevitably have high enough sensitivity to actually reconstruct the ionization history of the universe. A new generation of moderate resolution ground-based and balloon-born CMB polarization detectors (CLOVER, at 97, 150 and 220 GHz, QUIET at 40 and 90 GHz, QUaD, EBEX, BICEP, SPIDER, BRAIN) are under operation, construction or design. The principal aim is to measure primordial B-modes. They will also measure weak lensing-induced B-modes with resolution over multipoles $20 < \ell < 1000$ and down to $\sim 0.1\,\mu K$ precision. This is an essential prerequisite to searching, at these scales, for the gravity-wave induced B-mode background from inflation. For $20 < \ell < 100$, current constraints on the scalar to tensor ration should allow the primordial signal to dominate lensing.

Secondary effects are not simply a ‘foreground’ that adds noise and limits our knowledge. They are by nature the best tools to probe structure formation and evolution providing a complementary picture of the late time universe to that obtained from traditional tools such as galaxy surveys.

**Acknowledgments**

The authors would like to thank an anonymous referee and Matthias Bartelmann for a careful reading of and comments on the paper. NA and SM wish to thank Oxford University for hospitality. SM would like to thank CITA where a large part of the work was done as well as IAS-Orsay for hospitality during the final stages of the review.

**References**

Abel T, Bryan G L and Norman M L 2000 *Astrophys. J.* **540** 39
Abel T, Bryan G L and Norman M L 2002 *Science* **295** 93
Abell G O 1958 *Astrophys. J. Suppl.* **3** 211
Abreu M E et al 2004 *Astrophys. J.* **605** 607
Afshordi N 2004 *Phys. Rev.* **D 70** 083536
Afshordi N, Lin Y-T and Sanderson A J R 2005 *Astrophys. J.* **629** 1
Aghanim N, Balland C and Silk J 2000 *Astron. Astrophys.* **357** 1
Aghanim N, Désert F X, Puget J L and Gispert R 1996 *Astron. Astrophys.* **311** 1
Aghanim N and Forni O 1999 *Astron. Astrophys.* **347** 409
Aghanim N, Hansen S H and Lagache G 2005 *Astron. Astrophys.* **439** 901
Aghanim N, Hansen S H, Pastor S and Semikoz D V 2003 *J. Cosmol. Astropart. Phys.* **JCAP05**/2003/007
Aghanim N, Prunet S, Forni O and Bouchet F R 1998 *Astron. Astrophys.* **334** 409

Alvarez M A, Shapiro P R, Ahn K and Iliiev I T 2006 *Astrophys. J.* **644** L101
Alvarez M A, Komatsu E, Dore O and Shapiro P R 2006 *Astrophys. J.* **647** 840
Amblard A, Vale C and White M 2004 Preprint astro-ph/0403075
Ameiglo S, Borgani S, Diaferio A, Dolag K 2006 *Mon. Not. R. Astron. Soc.* **369** 1459
Arnaud M, Pointecouteau E and Pratt G W 2005 *Astron. Astrophys.* **441** 893
Atro-Barandela F and Mucket J 1999 *Astrophys. J.* **515** 465
Audit E and Simmons J F L 1999 *Mon. Not. R. Astron. Soc.* **305** L27
Barreiro R B and Hobson M P 2001 *Mon. Not. R. Astron. Soc.* **327** 813
Bartelmann M, Doran M and Wetterich C 2006 *Astron. Astrophys.* **454** 27
Bassett B A and Kunz M 2004 *Astrophys. J.* **607** 661
Battistelli E S et al 2003 *Astrophys. J.* **598** L75
Bean R, Melchiorri A and Silk J 2007 *Phys. Rev. D* **75** 063505
Benabed K, Bernardeau F and van Waerbeke L 2001 *Phys. Rev. D* **63** 3501
Benson B A et al 2003 *Astrophys. J.* **592** 674
Bernardeau F 1997 *Astron. Astrophys.* **324** 15
Bernardeau F 1998 *Astron. Astrophys.* **338** 767
Bernardeau F and Uzan J-P 2002 *Phys. Rev. D* **66** 103506
Birkinshaw M 1989 *Moving Gravitational Lenses* ed J Moran et al (Berlin: Springer) p 59
Birkinshaw M 1999 *Phys. Rep.* **310** 97
Birkinshaw M and Gull S F 1978 *Nature* **274** 111
Birkinshaw M and Gull S F 1983 *Nature* **302** 315
Birkinshaw M, Hughes J P and Arnaud K A 1991 *Astrophys. J.* **379** 466
Birkinshaw M and Hughes J P 1994 *Astrophys. J.* **420** 33
Birkinshaw M and Lancaster K 2004 Background microwave radiation and intractable cosmology *Proc. Int. School of Physics Enrico Fermi* (Preprint astro-ph/0410336)
Blanchard A and Schneider J 1987 *Astron. Astrophys.* **184** 1
Bock J et al 2006 Preprint astro-ph/0604101
Bonamente M, Joy M, La Roque S, Carlstrom J, Reese E and Dawson K 2006 *Astrophys. J.* **647** 25
Bond J R and Efstathiou G 1984 *Astrophys. J.* **285** 409
Bond J R and Efstathiou G 1987 *Mon. Not. R. Astron. Soc.* **226** 655
Bond J, Kaiser N, Cole S and Efstathiou G 1991 *Astrophys. J.* **379** 440
Borgani 2006 Lectures for 2005 Guillermo Haro Summer School on Clusters (Lecture Notes in Physics) (Preprint astro-ph/0605575)
Boughn S P and Crittenden R G 2002 *Phys. Rev. Lett.* **88** 021302
Boughn S P and Crittenden R G 2004 *Nature* **427** 45
Boughn S P, Crittenden R G, Turok N G 1998 *New Astron.* **3** 275
Bromm V, Coppi P S and Larson R B 2002 *Astrophys. J.* **564** 23
Bromm V, Kudritzki R P and Loeb A 2001 *Astrophys. J.* **552** 464
Bunn E F 2006 *Phys. Rev. D* **73** L3171
Carlstrom J E, Holder G P and Reese E D 2002 *Annu. Rev. Astron. Astrophys.* **40** 643
Carroll S M, Press W H and Turner E L 1992 *Annu. Rev. Astron. Astrophys.* **30** 499
Castro P G 2003 *Phys. Rev. D* **67** 123001
Cavaliere A, Danese L and de Zotti G 1977 *Astrophys. J.* **217** 6
Cavaliere A, Danese L and de Zotti G 1979 *Astron. Astrophys.* **75** 677
Cavaliere A and Fusco-Femiano R 1978 *Astron. Astrophys.* **70** 677
Cayon L, Martinez-Gonzalez E and Sanz J 1993 *Astrophys. J.* **413** 10
Cen R 2003 *Astrophys. J.* **591** 5
Challinor A and Lasenby A N 1998 *Astrophys. J.* **499** 1
Challinor A and Lasenby A N 1999 *Astrophys. J.* **510** 930
Challinor A and Lewis A 2005 *Phys. Rev. D* **71** 103010
Chatterjee S and Kosowsky A 2007 Preprint astro-ph/0701759
Chluba J and Mannheim K 2002 *Astron. Astrophys.* **396** 419
Chodorowski M 1992 *Mon. Not. R. Astron. Soc.* **259** 218
Scannapieco E 2000 Astrophys. J. 540 20
Scaramella R, Cen R and Ostriker J 1993 Astrophys. J. 416 399
Schaerer D 2002 Astron. Astrophys. 24 337
Schafer B M, Pfirrmann C, Bartelmann M, Springel V and Hernquist L 2006a Mon. Not. R. Astron. Soc. 370 1309
Schafer B M, Pfirrmann C, Hell R M and Bartelmann M 2006b Mon. Not. R. Astron. Soc. 370 1713
Schmalzing J and Gorski K M 1998 Mon. Not. R. Astron. Soc. 297 355
Seljak U 1996a Astrophys. J. 460 549
Seljak U 1996b Astrophys. J. 463 1
Seljak U, Burwella J and Pen U 2001 Phys. Rev. D 63 063001
Seshadri T R and Subramanian K 1998 Phys. Rev. D 58 063002
Shandarin S F 2002 Mon. Not. R. Astron. Soc. 331 865
Sheth R K, Mo H J and Tormen G 2001 Mon. Not. R. Astron. Soc. 323 1
Sheth R K and Tormen G 1999 Mon. Not. R. Astron. Soc. 308 119
Shimon M and Rephaeli Y 2004 New Astron. 9 69
Shimon M, Rephaeli Y, O’Shea B W and Norman M L 2006 Mon. Not. R. Astron. Soc. 368 511
Silk J 1967 Nature 215 1155
Silk J and White S D M 1978 Astrophys. J. 226 103
Smith K M, Zahn O and Doré O 2007 Phys. Rev. D 76 043510
Sokasian A, Abel T, Hernquist L and Springel V 2003 Mon. Not. R. Astron. Soc. 344 607
Sokasian A et al 2004 Mon. Not. R. Astron. Soc. 350 47
Somerville R and Livio M 2003 Astrophys. J. 593 611
Spergel D N et al 2003 Astrophys. J. Suppl. 148 175
Spergel D N et al 2007 Astrophys. J. Suppl. 170 377
Springel V, White M and Hernquist L 2001 Astrophys. J. 549 681
Stebbins A 1988 Astrophys. J. 327 584
Sunyaev R A and Zel’dovich Ya B 1970 Astrophys. Space Sci. 7 20
Sunyaev R A and Zel’dovich Ya B 1972 Comments Astrophys. Space Phys. 4 173
Sunyaev R A and Zel’dovich Ya B 1980 Ann. Rev. Astron. Astrophys. 18 537
Takada M 2001 Astrophys. J. 558 29
Takada M, Ohno H and Sugiyama N Preprint astro-ph/0112412
Tashiro H, Aghanim N and Langer M 2007 Preprint arXiv:0705.2861
Taylor J E, Moodley K and Diego J M 2003 Mon. Not. R. Astron. Soc. 345 1127
Theuns et al 2002 Astrophys. J. 574 111
Thompson K L and Vishniac E T 2003 Astrophys. J. 582 66
Toshio L, Negrello M, González-Nuevo J, De Zotti G, Silva L, Granato G L and Argueso F 2005 Astron. Astrophys. 438 475
Tomi K 2005 Phys. Rev. D 72 103506
Tomi K 2006 Phys. Rev. D 73 029901
Tolstoy E, Lagana P and Anninos P 1996 Astrophys. J. 463 15
Uzan J-P, Aghanim N and Mellett Y 2004 Phys. Rev. D 70 083533
Valageas P, Balbi A and Silk J 2001 Astron. Astrophys. 367 1
Valdès M, Ciardi B, Ferrara A, Johnston-Hollitt M and Rottgering H 2006 Mon. Not. R. Astron. Soc. 369 66
Vale C 2005 Preprint astro-ph/0509039
Vielva P, Barreiro R B, Hobson M P, Martinez-González E, Lasenby A N, Sanz J L and Toffolatti L 2001 Mon. Not. R. Astron. Soc. 328 1
Vielva P, Martinez-González E and Tucci M 2006 Mon. Not. R. Astron. Soc. 365 891
Vilenkin A and Shellard E P S 1994 Cosmic Strings and Other Topological Defects (Cambridge Monographs on Mathematical Physics) (Cambridge: Cambridge University Press)
Vishniac E T 1987 Astrophys. J. 322 597
Vittorio N and Juszkiewicz R 1987 Astrophys. J. 314 29
Volonteri M, Haardt F and Madau P 2003 Astrophys. J. 582 559
Walter et al 2004 Astrophys. J. 615 L17
Warren M S, Abazajian K, Holz D E and Teodoro L 2006 Astrophys. J. 646 L881
Weller J, Battye R A and Kneissl R 2002 Phys. Rev. Lett. 88 231301
Wyithe and Loeb A 2003 Astrophys. J. 588 69
Yamada M, Sugiyaama N and Silk J 1999 Astrophys. J. 522 66
Zahn O, Zaldarriaga M, Hernquist L and McQuinn M 2005 Astrophys. J. 630 657
Zaldarriaga M 2000 Phys. Rev. D 62 063510
Zaldarriaga M and Seljak U 1997 Phys. Rev. D 55 1830
Zaldarriaga M and Seljak U 1998 Phys. Rev. D 58 023003
Zaldarriaga M and Seljak U 1999 Phys. Rev. D 59 123507
Zaroubi S and Silk J 2005 Mon. Not. R. Astron. Soc. 360 64
Zaroubi S, Thomas R M, Sugiyama N and Silk J 2007 Mon. Not. R. Astron. Soc. 375 1269
Zhang P J 2004 Mon. Not. R. Astron. Soc. 348 1348
Zhang P J, Pen U-L and Trac H 2004 Mon. Not. R. Astron. Soc. 347 1224
Zhang P J, Pen U-L and Wang B 2002 Astrophys. J. 577 555