Mon. Not. R. Astron. Soc. **309**, 672–680 (1999)

Wavelets applied to cosmic microwave background maps: a multiresolution analysis for denoising

J. L. Sanz,[1] F. Argüeso,[2] L. Cayón,[1] E. Martínez-González,[1] R. B. Barreiro[1,3] and L. Toffolatti[4,5]

[1]*Instituto de Física de Cantabria, Fac. Ciencias, Av. los Castros s/n, 39005 Santander, Spain*
[2]*Departamento de Matemáticas, Universidad de Oviedo, c/ Calvo Sotelo s/n, 33007 Oviedo, Spain*
[3]*Departamento de Física Moderna, Universidad de Cantabria, 39005 Santander, Spain*
[4]*Departamento de Física, Universidad de Oviedo, c/ Calvo Sotelo s/n, 33007 Oviedo, Spain*
[5]*Osservatorio Astronomico di Padova, vicolo dell'Osservatorio n5, 35122 Padova, Italy*

Accepted 1999 June 16. Received 1999 April 26; in original form 1999 February 3

ABSTRACT

Analysis and denoising of cosmic microwave background (CMB) maps are performed using wavelet multiresolution techniques. The method is tested on $12.8 \times 12.8\,\mathrm{deg}^2$ maps, with the resolution resembling the experimental one expected for future high-resolution space observations. Semi-analytic formulae of the variance of wavelet coefficients are given for the Haar and Mexican Hat wavelet bases. Results are presented for the standard cold dark matter (CDM) model. Denoising of simulated maps is carried out by removal of wavelet coefficients dominated by instrumental noise. CMB maps with a signal-to-noise ratio, $S/N \sim 1$, are denoised with an error improvement factor between 3 and 5. Moreover, we have also tested how well the CMB temperature power spectrum is recovered after denoising. We are able to reconstruct the C_ℓs up to $l \sim 1500$ with errors always below 20 per cent in cases with $S/N \geqslant 1$.

Key words: methods: data analysis – cosmic microwave background.

1 INTRODUCTION

Future CMB space experiments will provide very detailed all-sky maps of CMB temperature anisotropies; NASA MAP Mission (Bennett et al. 1996) and the ESA Planck Mission (Mandolesi et al. 1998; Puget et al. 1998). The high sensitivity of these experiments will result in unique data to constrain fundamental cosmological parameters. Moreover, future CMB maps will enable a distinction to be made between competing theories of structure formation in the early Universe and will provide very fruitful data on astrophysical foregrounds.

The cosmological signal in CMB maps is hampered by instrumental noise and by foreground emissions. Therefore, a necessary step in analysing CMB maps is to separate the foreground emissions from the CMB signal. Several linear and non-linear methods have already been tested on simulated data (Bouchet, Gispert & Puget 1996; Tegmark & Efstathiou 1996; Hobson et al. 1998b, 1999). An alternative method can be one based on wavelets. Wavelets are known to be very efficient in dealing with problems of data compression and denoising. The development of wavelet techniques applied to signal processing has been very fast in the last 10 yr (see Jawerth & Sweldens 1994 for an overview). These techniques have already been applied to a variety of astrophysical problems. For example, regarding cosmology, Slezak, de Lapparent & Bijaoui (1993) have applied wavelet analysis to the detection of structures in the CfA redshift survey. They have also been introduced to study the Gaussian character of CMB maps (Pando et al. 1998; Hobson, Jones & Lasenby 1998a). A study using spherical Haar wavelets to denoise CMB maps has just appeared (Tenorio et al. 1999).

We consider small patches of the sky where a flat 2D approach is valid. We apply wavelet multiresolution techniques, known to be computationally very fast, taking only $O(N)$ operations to reconstruct an image of N pixels. In the 2D flat wavelet analysis a single scale and two translations are usually introduced, where the basis is generated by 4 tensor products of wavelets and scaling functions. Therefore, three *detail* images plus an *approximation* image appear at each level of resolution. Different wavelet bases are characterized by their location in space. The bases considered in this work are Mexican Hat, Haar and Daubechies. The first one is the most localized though, as unlike the other two, it does not have a compact support. As a first approach to the application of these techniques to CMB data we only consider maps with cosmological signal plus instrumental Gaussian noise. CMB experiments are contaminated by galactic (dust, free–free and synchrotron emission) and extragalactic foregrounds (infrared and radio galaxies, S–Z effects from clusters,…) in addition to noise. Obviously in view of this, the contents of this paper only cover a

small fraction of the work to be done. Our aim is to shed light on the wavelet characterization of the different components, the late phase being to build up a wavelet-based framework to disentagle all of them. In this line, a Bayesian method (maybe incorporating entropy or other constraints) dealing with wavelet components at different scales and integrating the different channels will be the final goal. Knowing the efficiency of wavelets in removing noise as shown by many works in other fields, the first step to take in the application of wavelet techniques to the CMB is to study the denoising of temperature maps.

The outline of the paper is as follows. A theoretical continuous wavelet analyis of CMB data is presented in Section 2. General semi-analytical formulae are given for the variance of the *detail* wavelet coefficients as a function of the temperature power spectrum C_ℓ. Section 3 introduces the discrete wavelet technique, which is applied to denoising of simulated CMB maps in Section 4. Discussion and conclusions are presented in Section 5.

2 CONTINUOUS WAVELET ANALYSIS

2.1 One-dimensional transform

The Fourier transform is a powerful tool in many areas but in dealing with local behaviour shows a tremendous inefficiency. For instance, a large number of complex exponentials must be combined in order to produce a spike. The wavelet transform solves this problem, introducing a good space-frequency localization. It is conceptually simple and it constitutes a fast algorithm. Let $\psi(x)$ be a one-dimensional function satisfying the following conditions: (a) $\int_{-\infty}^{\infty} \mathrm{d}x\, \psi(x) = 0$, (b) $\int_{-\infty}^{\infty} \mathrm{d}x\, \psi^2(x) = 1$ and (c) $C_\psi \equiv \int_{-\infty}^{\infty} \mathrm{d}k\, |k|^{-1} \psi(k) < \infty$, where $\psi(k)$ is the Fourier transform of $\psi(x)$. So, according to condition (a), the wavelet must have oscillations. Condition (b) is a normalization and (c) represents an admissibility condition in order to reconstruct a function $f(x)$ with the basis ψ [see equation (2) for such a synthesis].

We define the *analysing* wavelet as $\Psi(x; R, b) \equiv R^{-1/2}\psi(x-b)/R$, dependent on two parameters: dilation (R) and translation (b). It operates as a mathematical microscope of magnification R^{-1} at the space point b. The wavelet coefficients associated to a one-dimensional function $f(x)$ are:

$$w(R, b) = \int \mathrm{d}x f(x) \Psi(x; R, b). \tag{1}$$

It is clear from the above definition that such coefficients represent the analysing wavelet at x_o for a delta distribution peaked at this point, i.e. for $f(x) = \delta(x - x_o)$. For $R = 1$, $w(R,b)$ is the convolution of the function f with the analysing wavelet ψ.

The reconstruction of the function f can be achieved in the form

$$f(x) = (2\pi C_\psi)^{-1} \int\int \mathrm{d}R\, \mathrm{d}b\, R^{-2} w(R, b) \Psi(x; R, b). \tag{2}$$

Examples of wavelet functions are: (i) Haar, $\psi = 1(-1)$ for $0 < x < 1/2$ $(1/2 < x < 1)$, (ii) Mexican hat, $\psi = 2/(9\pi)^{1/4}$ $(1 - x^2)\mathrm{e}^{-x^2/2}$.

2.2 Two-dimensional transform

Regarding the two-dimensional case, we introduce a one-dimensional *scaling* function ϕ normalized in the form: $\int_{-\infty}^{\infty} \mathrm{d}x\, \phi(x) = 1$. Examples of scaling functions are: (i) Haar, $\phi = 1(0)$ for $0 < x < 1$ $(x < 0, x > 1)$, (ii) Mexican Hat, $\phi = [2/(9\pi)^{1/4}]\mathrm{e}^{-x^2/2}$. The analysing *scaling* function $\Phi(x; R, b) \equiv R^{-1/2}\phi(x-b)/R$, allows the definition of *details* of an image, $f(\boldsymbol{x})$, with respect to the tensor products

$$\Gamma_\mathrm{d}(\boldsymbol{x}; R, \boldsymbol{b}) \equiv \Psi(x_1; R, b_1)\Psi(x_2; R, b_2), \tag{3}$$

$$\Gamma_\mathrm{h}(\boldsymbol{x}; R, \boldsymbol{b}) \equiv \Phi(x_1; R, b_1)\Psi(x_2; R, b_2), \tag{4}$$

$$\Gamma_\mathrm{v}(\boldsymbol{x}; R, \boldsymbol{b}) \equiv \Psi(x_1; R, b_1)\Phi(x_2; R, b_2). \tag{5}$$

The *diagonal, horizontal* and *vertical* wavelet coefficients are defined by ($\alpha \equiv d, h, v$)

$$w_\alpha(R, \boldsymbol{b}) = \int \mathrm{d}\boldsymbol{x} f(\boldsymbol{x}) \Gamma_\alpha(\boldsymbol{x}; R, \boldsymbol{b}). \tag{6}$$

Scaling functions act as low-pass filters whereas wavelet functions single out one scale. Therefore, *detail* coefficients provide local information about symmetrical (diagonal) and elongated/filamentary structure (vertical and horizontal).

Let us now assume an homogeneous, isotropic random field $f(\boldsymbol{x})$, i.e. the correlation function $C(r) \equiv \langle f(\boldsymbol{x}) f(\boldsymbol{x}+\boldsymbol{r})\rangle$, $r \equiv |\boldsymbol{r}|$, where $\langle\rangle$ denotes an average value over realizations of the field. The Fourier transform of the field $f(\boldsymbol{k})$ satisfies $\langle f(\boldsymbol{k}) f(\boldsymbol{k}')\rangle = P(\boldsymbol{k})\delta^2(\boldsymbol{k} - \boldsymbol{k}')$, where $k \equiv |\boldsymbol{k}|$ and $P(k)$ is the power spectrum [the Fourier transform of $C(r)$]. In this case we can calculate the correlation and variance of the wavelet coefficients: $C_\alpha(r; R) \equiv \langle w_\alpha(R, \boldsymbol{b}) w_\alpha(R, \boldsymbol{b} + \boldsymbol{r})\rangle$, $\sigma_\alpha^2(R) \equiv C_\alpha(0; R)$ and we find the following equations

$$C(0) \equiv \sigma^2 = C_{\Gamma_\alpha}^{-1} \int \mathrm{d}R\, R^{-3} \sigma_\alpha^2(R),$$

$$C_{\Gamma_\alpha} \equiv (2\pi)^2 \int \mathrm{d}\boldsymbol{k}\, k^{-2} |\tilde{\Gamma}_\alpha|^2(\boldsymbol{k}), \tag{7}$$

where $\tilde{\Gamma}_\alpha(\boldsymbol{k})$ is the Fourier transform of $R\Gamma_\alpha$.

On the other hand, we calculate the Fourier transform of the wavelet coefficients $w_\alpha(R, \boldsymbol{b})$ with respect to the $\boldsymbol{b}$ parameters:

$$\langle w_\alpha(R, \boldsymbol{k}) w_{\alpha'}(R', \boldsymbol{k}')\rangle = w_{\alpha\alpha'}(R, R'; \boldsymbol{k})\delta^2(\boldsymbol{k} - \boldsymbol{k}'), \tag{8}$$

$$w_{\alpha\alpha'} = (2\pi)^2 R R' P(k) \tilde{\Gamma}_\alpha^*(R\boldsymbol{k}) \tilde{\Gamma}_{\alpha'}(R'\boldsymbol{k}), \tag{9}$$

which allows us to get the *detail* wavelet variances as

$$\sigma_\alpha^2(R) = \int \mathrm{d}\boldsymbol{k}\, P(kR^{-1}) |\tilde{\Gamma}_\alpha(\boldsymbol{k})|^2. \tag{10}$$

The diagonal variance corresponds to the tensor product of two one-dimensional wavelets. If $|\psi(k)|^2$ is a function that is strongly peaked near $k \simeq 1$ then $\sigma_\mathrm{d}^2(R) \simeq P(k \simeq R^{-1})$, taking into account the normalization of the wavelet function, which allows an estimation of the power spectrum in terms of the diagonal component. This is what happens for the Mexican Hat: $|\psi(k)|^2 \propto k^4 \mathrm{e}^{-k^2}$, with a maximum at $k = 2^{-1/2}$, whereas the Haar wavelet is not localized in Fourier space: $|\psi(k)|^2 \propto (k/4)^{-2} \sin^4(k/4)$. We can also deduce that $C_\mathrm{h} = C_\mathrm{v}$ and $\sigma_\mathrm{h}^2 = \sigma_\mathrm{v}^2$ taking into account the symmetry of the equations. Moreover, the temperature power spectrum $P(k)$ can be obtained from the *detail* wavelet power spectrum $w_{\alpha\alpha}(R, R; \boldsymbol{k})$ as follows

$$P(k) = \frac{1}{C_{\Gamma_\alpha}} \int \frac{\mathrm{d}R}{R^3} \int \mathrm{d}\theta\, w_{\alpha\alpha}(R, R; k\boldsymbol{n}), \qquad \boldsymbol{n} = (\cos\theta, \sin\theta). \tag{11}$$

For the Haar and Mexican wavelets we can calculate:

$$\text{Haar}: |\tilde{\Gamma}_\mathrm{d}|^2 = \frac{1}{(2\pi)^2} \left(\frac{k_1 k_2}{4}\right)^{-2} \left[\sin\frac{k_1}{4} \sin\frac{k_2}{4}\right]^4, \tag{12}$$

$$|\tilde{\Gamma}_{\rm h}|^2 = \frac{1}{(2\pi)^2}\left(\frac{k_1 k_2}{4}\right)^{-2}\left[\sin\frac{k_2}{4}\right]^4 \frac{1}{4}\sin^2 k_1, \tag{13}$$

$$\text{Mexhat}: |\tilde{\Gamma}_{\rm d}|^2 = \frac{16}{9\pi}(k_1 k_2)^4\, {\rm e}^{-k^2},$$

$$|\tilde{\Gamma}_{\rm h}|^2 = \frac{4}{3\pi}k_2^2\, {\rm e}^{-k^2}, \tag{14}$$

where $k^2 = k_1^2 + k_2^2$ and $\tilde{\Gamma}_{\rm v}$ can be obtained from $\tilde{\Gamma}_{\rm h}$, swapping k_1 and k_2. The variance of the *detail* wavelet coefficients for the Haar and Mexican Hat systems, assuming the standard CDM model, is presented in Fig. 1. As one can see, the acoustic peaks can be clearly noticed, being more pronounced for the Mexican Hat basis. This last result is a consequence of being a more localized wavelet system. For a more detailed discussion see Sanz et al. (1998, in preparation).

3 DISCRETE WAVELET ANALYSIS

3.1 One-dimensional multiresolution analysis

An orthonormal basis of $L^2(\Re)$ can be constructed from a wavelet ψ through dyadic dilations j and translations k

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k). \tag{15}$$

In addition, a scaling function ϕ can be defined associated to the *mother* wavelet ψ. Such a function gives rise to the so called multiresolution analysis. A multiresolution analysis of $L^2(\Re)$ is defined as a sequence of closed subspaces V_j of $L^2(\Re)$, $j \in Z$. Properties can be seen in Ogden (1997).

Subspaces V_j are generated by dyadic dilations and translations of the scaling function ϕ (this function forms an orthonormal basis of V_o, $\{\phi_{o,k}(x) = \phi(x-k)\}$). Moreover, each V_j can be expressed as the orthogonal sum $V_j = V_{j-1} \oplus W_{j-1}$, where W_{j-1} is created from wavelets $\psi_{j-1,k}$. Taking into account the properties of the scaling function, together with this last expression, we can construct *approximations* at increasing levels of resolution. These *approximations* are linear combinations of dilations and translations of a scaling function ϕ. The difference between two consecutive *approximations*, i.e. the *detail* at the corresponding resolution level, is given by a linear combination of dilations and translations of a wavelet function ψ.

3.2 Two-dimensional multiresolution analysis

The analysis performed in this work assumes equal dilations in the two dimensions involved. At a fixed level of resolution, subspaces in a 2D multiresolution analysis are the tensor products of the corresponding 1D ones $\mathbf{V}_{j+1} = V_{j+1} \otimes V_{j+1}$. The 2D basis is therefore built by the product of two scaling functions (*approximation*), the product of wavelet and scaling functions (horizontal and vertical *details*) and the product of two wavelets (diagonal *details*):

$$\mathbf{V}_{j+1} = (V_j \oplus W_j) \otimes (V_j \oplus W_j) \tag{16}$$

$$= (V_j \otimes V_j) \oplus [(V_j \otimes W_j) \oplus (W_j \otimes V_j) \oplus (W_j \otimes W_j)]. \tag{17}$$

Horizontal, vertical and diagonal *detail* coefficients represent the variations in these directions relative to a weighted average at a lower resolution level (given by the *approximation* coefficients).

A discrete orthonormal basis, $\Gamma_\alpha(\boldsymbol{x}; j, \boldsymbol{k})$, can be defined by setting $R = 2^{-j}$ and $\boldsymbol{b} = 2^{-j}\boldsymbol{k}$ in equations (3)–(5), then $(\Gamma_\alpha(\boldsymbol{x}; j, \boldsymbol{k})\boldsymbol{\Gamma}_{\boldsymbol{\alpha}'}(\boldsymbol{x}; j', \boldsymbol{k}')) = \delta_{\alpha\alpha'}\delta_{jj'}\delta_{\boldsymbol{k}\boldsymbol{k}'}$, where () denotes the scalar product in $L^2(\Re^2)$. If we define the discrete wavelet coefficients associated to any *detail* by equation (6)

$$w_\alpha(j, \boldsymbol{k}) = \int {\rm d}\boldsymbol{x} f(\boldsymbol{x})\Gamma_\alpha(\boldsymbol{x}; j, \boldsymbol{k}), \tag{18}$$

we can thus reconstruct the image with all the *details*

$$f(\boldsymbol{x}) = \sum_{\alpha, j, \boldsymbol{k}} w_\alpha(j, \boldsymbol{k})\Gamma_\alpha(\boldsymbol{x}; j, \boldsymbol{k}). \tag{19}$$

In particular, we get the following expression for the second-order moment of the image

$$(f^2(\boldsymbol{x})) = \sum_{\alpha, j, \boldsymbol{k}} w_\alpha^2(j, \boldsymbol{k}), \tag{20}$$

which expresses how the energy of the field is distributed locally at any scale and *detail*.

For a finite image, $R_{\max} \times R_{\max}$, in order to reconstruct it we must add to equation (19) an approximation $w_a(\boldsymbol{k})\Gamma_a(\boldsymbol{x};\boldsymbol{k})$ with $\Gamma_a(\boldsymbol{x};\boldsymbol{k}) \equiv \Phi(x_1; R_{\max}, k_1)\ \ \Phi(x_2; R_{\max}, k_2)$ and $w_a(\boldsymbol{k}) \equiv \int {\rm d}\boldsymbol{x} f(\boldsymbol{x})\Gamma_a(\boldsymbol{x}, \boldsymbol{k})$, representing the field at the lower resolution. If $f(\boldsymbol{x})$ represents the temperature fluctuation field then the variance is given by $\langle(\Delta T/T)^2\rangle = ((\Delta T/T)^2)/N_{\rm p}$, $N_{\rm p}$ being the number of pixels.

The orthonormal bases that we are going to use are the standard Daubechies N (Haar corresponds to $N = 1$), which have been extensively used in the literature because of their special properties: they are defined in a compact support, have increasing regularity with N and vanishing moments up to order $N - 1$ (Daubechies 1988). On the contrary, the Mexican Hat wavelet is not defined in a compact support and it is not appropriate for this multiresolution analysis.

For discrete wavelet analysis of the CMB maps we use the

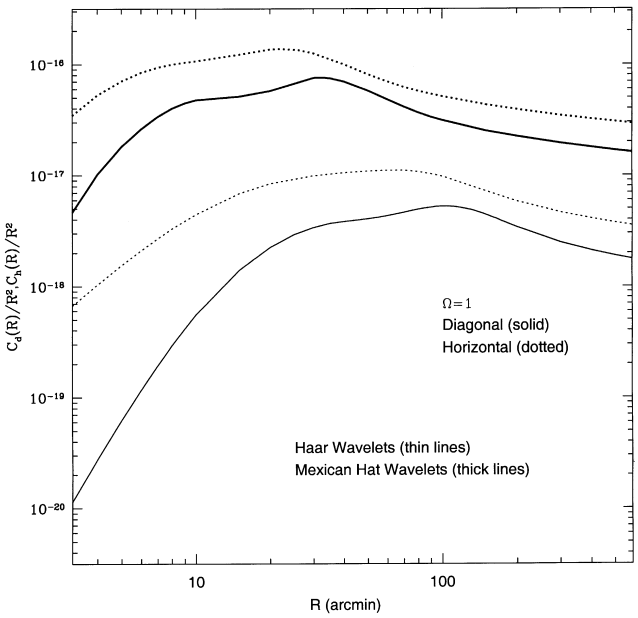

Figure 1. Variance of diagonal (solid lines) and horizontal/vertical (dotted lines) *detail* wavelet coefficients $C_{\rm d}$, $C_{\rm h}$ versus scale R. A standard CDM cosmological model is assumed. Thin lines outline the result obtained for the Haar wavelet system and the thick lines correspond to the Mexican Hat wavelet basis.

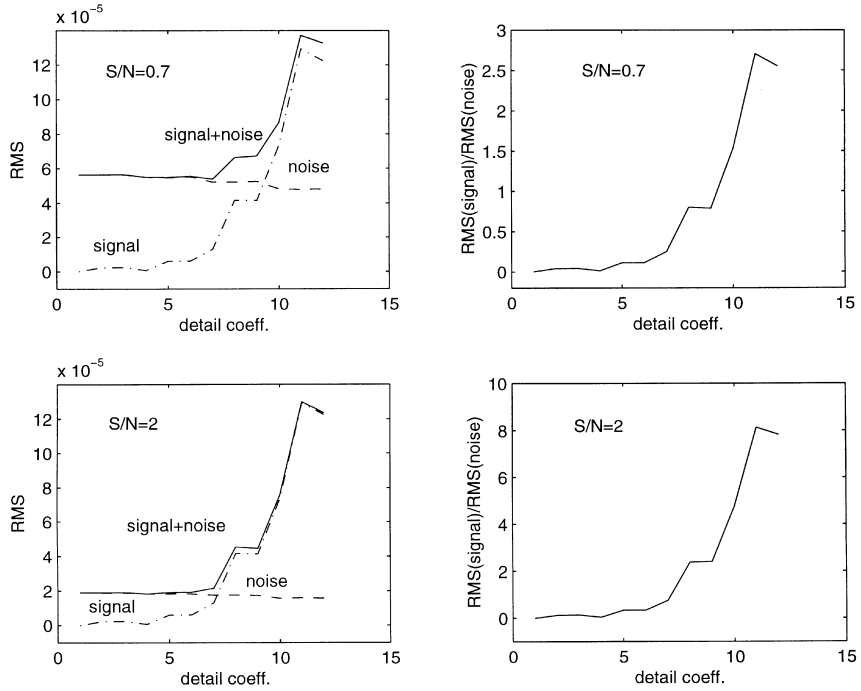

Figure 2. Rms deviation of wavelet *detail* coefficients obtained from CMB maps (signal maps, dashed-dotted lines), CMB plus noise maps (signal plus noise maps, solid lines) and pure noise (dashed lines) are presented in the left-hand panels. The right-hand panels show the ratio of the rms deviation of the *detail* coefficients from signal maps divided by the rms deviation of the *detail* coefficients from noise maps. Top panels correspond to simulated maps with $S/N = 0.7$; bottom panels correspond to $S/N = 2$.

Matlab Wavelet Toolbox (Misiti et al. 1996). This toolbox is an extensive collection of programs for analysing, denoising and compressing signals and 2D images. Discrete Wavelet decomposition is performed as described above to obtain the *approximation* and *detail* coefficients of the 2D CMB maps at several levels.

4 DENOISING OF CMB MAPS

Future CMB space experiments will provide maps with resolution scales of a few arcmin. In this work we analyse simulated maps of $12.8 \times 12.8\,\mathrm{deg}^2$ with pixel size of 1.5 arcmin. Simulations are made assuming the standard CDM, $\Omega = 1$ and $H_o = 50\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. The maps are filtered with a 4.5-arcmin full width at half-maximum (FWHM) Gaussian beam to reproduce approximately the filtering scale of the High Frequency Channels of the Planck Mission. Simulated maps have a rms signal of $\Delta T/T = 3.7 \times 10^{-5}$. Gaussian noise is added to these maps at different *S/N* levels between 0.7 and 3. A non-uniform noise is also considered to account for the non-uniform sampling introduced in satellite observations. As an extreme case, we have assigned the signal-to-noise ratio at each pixel from a truncated (at the 2σ level) Gaussian distribution with a mean value of 2 and a dispersion of 0.5. We use the set of Matlab Wavelet 2D programs with the corresponding graphical interface to analyse and denoise those maps. Suitable bases of wavelets are studied. Daubechies' four wavelets are the ones used in this analysis. No significant changes are observed when the analysis is carried out using other higher order Daubechies bases. On the other hand, the Haar system is not so efficient for denoising CMB maps since it produces reconstruction errors much larger than when using high-order Daubechies systems.

First of all, three wavelet decompositions are performed obtaining wavelet coefficients corresponding to the CMB original map, to the signal plus noise map and to the pure noise map. Decompositions are carried out up to the fourth resolution level. Denoising of the signal plus noise maps is based on subtraction of certain sets of coefficients affected by noise. White noise is the most common in CMB experiments. The dispersion of wavelet coefficients of that type of noise is constant as can be seen from equation (10). On the contrary, CMB detail wavelet dispersions go to zero as R goes to zero. Therefore, first-level wavelet coefficients are dominated by noise and then, for a given signal plus noise map, it is possible to know the noise and consequently the CMB wavelet coefficient dispersions at all levels. CMB maps produced by typical experiments with a ratio between antenna and pixel size of ≈ 3 will have wavelet coefficients containing the relevant information on the signal at level 3 and above. As shown below, level 3 is the critical one for performing denoising as the noise can still be at a level comparable to the signal. Fig. 2 shows rms deviations and corresponding ratios for two simulations with $S/N = 0.7$ and $S/N = 2$. *Detail* coefficient numbering corresponds to the three directions, diagonal, vertical and horizontal at the three consecutive levels, i.e. numbers 1, 2, 3 correspond to diagonal, vertical and horizontal cofficients at the first resolution level, 4, 5, 6 to the second-level coefficients in the same order and 7, 8, 9, 10, 11, 12 to levels 3 and 4, respectively. As it can be seen, the first two levels are entirely dominated by noise as pointed out before. Therefore, all these coefficients can be removed to reconstruct a denoised map. This is equivalent to using a hard thresholding, assuming a threshold above all these coefficients. On the other side, level 4 is completely dominated by the CMB signal and is left untouched. Ratios between rms deviations of the signal and noise maps at the third resolution level are not always clearly dominated by noise or signal. Ratios of ≈ 1 are treated with a soft-thresholding technique (in practice we consider ratios in the range

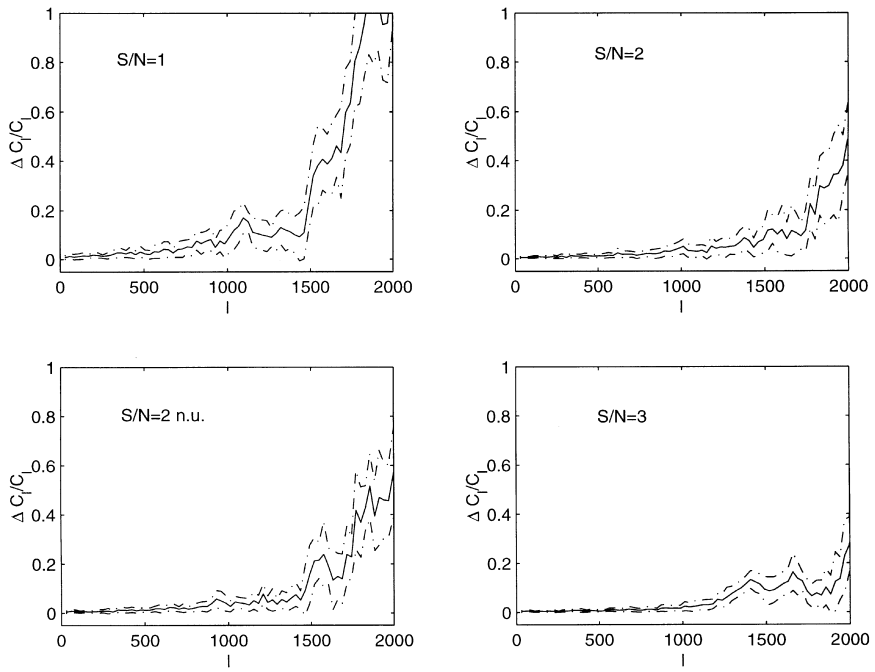

Figure 3. Mean value (solid line) and 1σ error (dashed-dotted lines) of the absolute value of the relative errors, $\Delta C_\ell/C_\ell$. Top-left panel corresponds to $S/N = 1.0$, top-right to $S/N = 2$, bottom-left to $S/N = 2$ with non-uniform noise and bottom-right to $S/N = 3$.

Table 1. Reconstruction errors vs *S/N*.

S/N	Per cent map error $\pm 1\sigma$
0.7	26.3 ± 0.4
1.0	20.7 ± 0.4
2.0	13.3 ± 0.2
2.0 (n.u.)	14.3 ± 0.3
3.0	10.3 ± 0.2

Table 2. Reconstruction errors vs threshold, $S/N = 1$.

Threshold	Per cent map error
hard	23.5
$1.5\ \sigma_n$	21.7
$1.0\ \sigma_n$	20.7
$0.7\ \sigma_n$	20.5
$0.6\ \sigma_n$	20.6
$0.5\ \sigma_n$	20.7
$0.4\ \sigma_n$	21.0
$0.3\ \sigma_n$	21.3
signal + noise	100.0

0.3–1.5, although changes in this interval do not significantly affect the results). Soft thresholding consists of removing all coefficients with absolute values smaller than the threshold defined in terms of the noise dispersion (σ_n). Coefficients with absolute values above the defined threshold are rescaled by subtracting the threshold to the positive ones and adding it to the negative ones. To define these thresholds we use the so-called SURE thresholding technique introduced by Donoho & Johnstone (1995). This technique is based on finding an estimator of the signal that will minimize the expected loss or risk defined as the mean value of $(1/N_p)\sum_{i=1}^{N_p}(T_{d_i} - T_i)^2$, where T_i is the temperature at pixel i in the original signal map and T_{d_i} is the estimator at pixel i (temperature in the final denoised map). The minimization is finally achieved in the wavelet domain by choosing a threshold value that minimizes the risk at each wavelet level (see for instance Ogden 1997).

Results of the errors in the map reconstruction are shown in Table 1. The map error is defined as:

$$\left(\frac{\sum_{i=1}^{n_{\rm pixels}} (T_i - T_{d_i})^2}{\sum_{i=1}^{n_{\rm pixels}} T_i^2} \right)^{1/2} . \qquad (21)$$

Performing 20 simulations (proved to be enough as the results reached stable values) at each *S/N* level we have also calculated the 1σ error. The error improvement achieved with the denoising technique applied goes from factors of 3 to 5 for $S/N = 3$ to $S/N = 0.7$.

It is also interesting to see how well the denoising method performs at reconstructing the temperature power spectrum. Mean values and 1σ errors of the relative errors, $|\Delta C_\ell/C_\ell|$, are shown in Fig. 3 for three *S/N* ratios and the case of non-uniform noise considered in this work. The C_ℓs are reconstructed from the denoised maps with $|\Delta C_\ell/C_\ell| \leqslant 10$ per cent up to $l \sim 1000$ in cases $S/N \geqslant 1$. This error can only be achieved up to an $l \sim 700$ in the $S/N = 0.7$ case. Higher order multipoles ($\ell \leqslant 1500$) are reconstructed with $|\Delta C_\ell/C_\ell| \leqslant 20$ per cent. Absolute relative errors and reconstructed C_ℓs for a given map are presented in Figs 4 and 5, respectively.

In order to check the performance of the SURE thresholding technique, knowing the original maps we can find the optimal threshold to get a reconstructed map with a minimum error (as defined above). In $S/N = 1$ maps the optimal threshold is found to be $0.6–0.7\sigma_n$. Thresholds between $0.3–1\sigma_n$ do not make substantial changes in the reconstructed map (see Table 2). The hard case included in this table stands for a case where all the coefficients below a signal-to-noise dispersion ratio of 1.5 are removed, leaving the others untouched. For comparison, the error

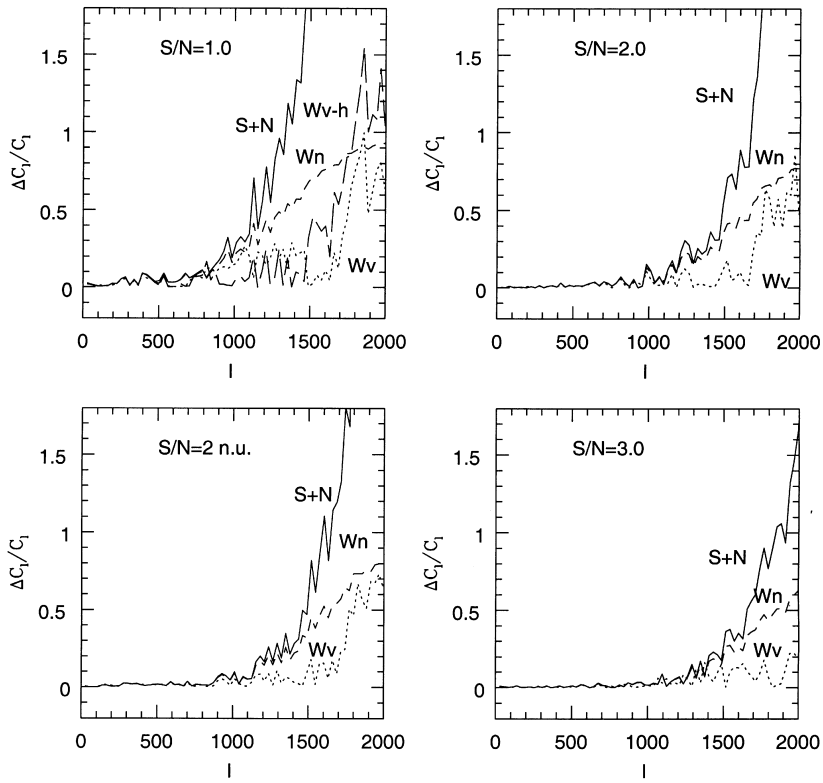

Figure 4. Absolute value of the relative errors, $\Delta C_\ell/C_\ell$, of the CMB power spectrum obtained from signal-plus-noise maps (solid lines), wavelet denoised maps (short dashed lines) and Wiener denoised maps (dashed lines). Top-left panel corresponds to $S/N = 1$ (wavelet denoised maps removing all coefficients at levels 1, 2, 3d and 3h is included as long dashed lines), top-right to $S/N = 2$, bottom-left to $S/N = 2$ with non-uniform noise and bottom-right to $S/N = 3$.

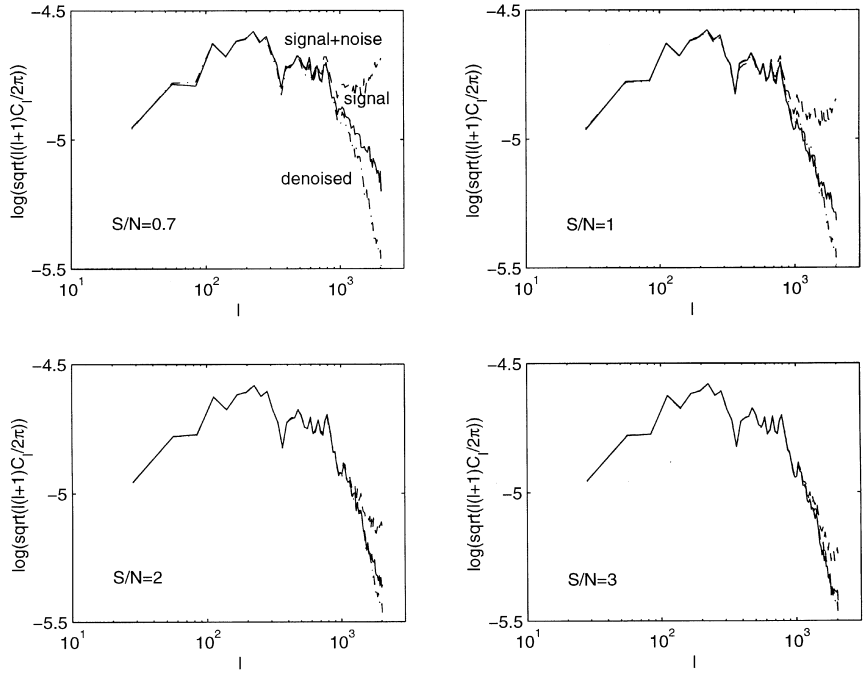

Figure 5. Power spectrum obtained from signal-plus-noise maps (dashed lines), signal maps (solid lines) and denoised maps (dashed-dotted lines). Top-left $S/N = 0.7$, top-right $S/N = 1$, bottom-left $S/N = 2$ and bottom-right $S/N = 3$.

obtained comparing the signal plus noise map with the original signal map is also presented in Table 2. We can see that the error reconstruction achieved with the SURE technique equals the one obtained with the optimal threshold.

A comparison of wavelet techniques with a Wiener filter (see, for instance, Press et al. 1994) has also been performed. In relation to map reconstruction the error affecting the Wiener reconstructed maps is comparable to the error for the wavelet reconstructed

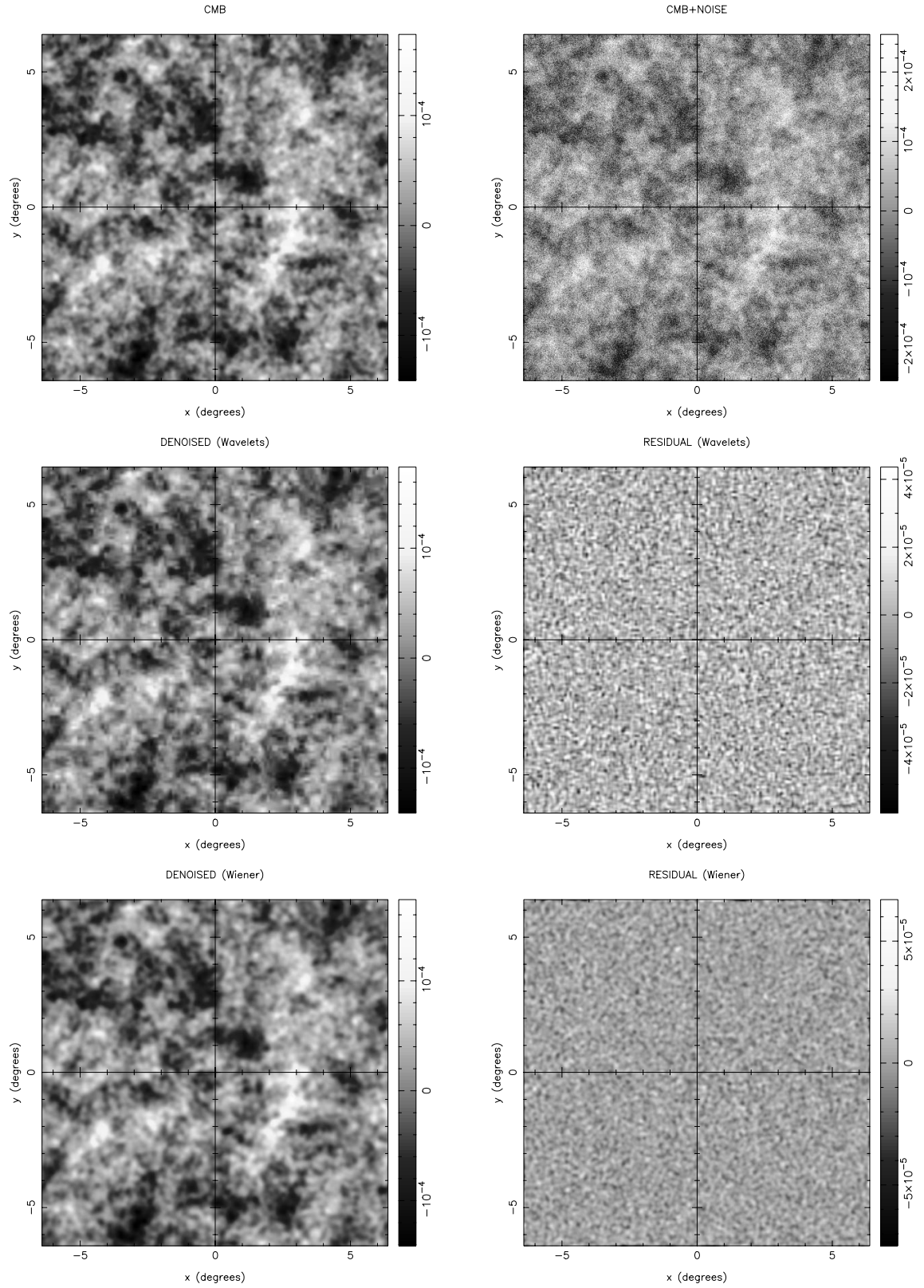

Figure 6. $12.8 \times 12.8\,\mathrm{deg}^2$ maps of the cosmological signal (top left), signal plus noise with $S/N = 1$ (top right), denoised map using a soft thresholding as explained in the text (middle left) and residual map obtained from the CMB signal map minus the denoised one (middle right). For comparison a denoised map using Wiener filter is presented in the bottom-left panel together with the residuals in the bottom-right panel.

maps, in all cases. However, in order to apply the Wiener filter, previous knowledge of the signal power spectrum is required. Reconstructed and residual maps using both, wavelets and the Wiener filter, are shown in Fig. 6. Regarding the C_ℓs, the performance of the Wiener filter is clearly worse than wavelets for $\ell = 1000$–1500, as can be seen in Fig. 4. For example, for $S/N = 1$, the C_ℓs are recovered using a Wiener filter with an error between 20 and 70 per cent for ℓs between 1000 and 1500, this error being smaller by a factor of 2–4 for the wavelet reconstruction. The error is also clearly larger for Wiener reconstruction than for wavelet reconstruction, up to $\ell \sim 2000$ in cases with $S/N > 1$.

We have checked for non-Gaussian features possibly introduced by the non-linearity of the soft thresholding used in the wavelet methods applied for denoising. Distributions of skewness and kurtosis have been obtained for the original signal maps as well as for the denoised ones. No significant differences can be appreciated between both distributions. However, this method could not be good enough to detect non-Gaussian features. As recently claimed by Hobson et al. (1998a), the analysis of the distribution of wavelet coefficients is one of the most efficient methods to detect them. We have performed a similar analysis using the Daubechies 4 multiresolution wavelet coefficients. These coefficients are Gaussian-distributed in the case of a temperature Gaussian random field. The application of soft thresholding to the wavelet coefficients at a certain level clearly changes the Gaussian distribution by removing all coefficients of which the absolute values are below the imposed threshold and shifting the remaining ones by that threshold. As an example, in the previous case $S/N = 1$, the kurtosis of the diagonal level 3 distribution changes from 3.3 ± 0.1 to 34 ± 10! (notice that the change strongly depends on the threshold imposed). This result is not surprising as any non-linear method used for denoising or foreground separation will introduce non-Gaussinity at different levels in the reconstructed map. Fortunately there are two ways of overcoming the question of determining the Gaussianity of the CMB signal. One way would be to check the Gaussian character of the data before applying denoising to maps affected by Gaussian noise. We have checked this by looking at the multiresolution wavelet coefficient distributions in the case of $S/N = 1$. The addition of white noise did not change the mean value and error bar of the kurtosis. The second way would be to apply a linear denoising method. We have used a simple one consisting of removing all detail coefficients at levels with signal-to-noise dispersion ratio < 1.5 (notice that 1.5 corresponds to the upper value of the threshold interval where soft thresholding was applied). This method is equivalent to applying hard thresholding with a threshold above all the coefficients. The errors of the reconstructed map and its corresponding C_ls increase slightly compared with the SURE thresholding method (see Table 2 and top-left panel of Fig. 4). The same hard thresholding linear method will give even better results using 2D wavelets with two scales of dilation (Sanz et al., in preparation) instead of the one-scale multiresolution techniques, because the former works with many more resolution levels being therefore more selective in removing the coefficients.

5 DISCUSSION AND CONCLUSIONS

A wavelet multiresolution technique has been presented and used to analyse and denoise CMB maps. This method has been proved to be one of the best for reconstructing observed CMB maps as well as power spectra by removing a significant percentage of the noise. The analysis has been carried out assuming a uniform Gaussian noise as would be expected in a small sky patch, e.g. $12.8 \times 12.8\,\mathrm{deg}^2$, observed by satellite scans. Analysis of whole sky CMB maps using wavelets will be performed in a future work. As these data are affected by non-uniform noise, the use of wavelet techniques to localize map features is highly suitable.

A semi-analytical calculation of the variance of the wavelet coefficients has been presented. The behaviour of the variance of the *detail* coefficients is given for a standard CDM model in the case of Haar and Mexican Hat bases. The acoustic peaks can be noticed in the wavelet coefficient variance represented in Fig. 1. Moreover, these peaks are better defined for the Mexican Hat wavelet system because these wavelets are more localized than the Haar ones.

Denoising of CMB maps has been carried out by using a signal-independent prescription, the SURE thresholding method. The results are model-independent depending only on the observed data. However, a good knowledge of the noise affecting the observed CMB maps is required. For a typical case of $S/N \sim 1$ the high-order *detail* coefficients are dominated by the signal, whereas the lowest ones are noise-dominated. This behaviour is due to the expected dependence of the temperature power spectrum, $C_\ell \propto l^{-2}$. The applied wavelet method is able to reconstruct maps with an error improvement factor between 3 and 5, and the CMB power spectrum of the denoised maps carries relative errors below 20 per cent up to $l \sim 1500$ for $S/N \geqslant 1$. We have also checked that SURE thresholding methods are providing thresholds in agreement with the optimal ones.

For comparison a Wiener filter has also been applied to the simulations considered in this paper. This method reconstructs CMB maps after denoising with errors comparable to the wavelet method we propose, as shown in Fig. 6. However, the C_ℓs of the denoised maps obtained applying Wiener filter have relative errors larger than a factor of 2 than the relative errors of the C_ℓs obtained from the wavelet reconstructed maps in the range $\ell = 1000$–1700. In addition we have applied a Maximum Entropy Method (MEM) to the maps used in this work, with the definition of entropy given by Hobson & Lasenby (1998). This method provides reconstruction errors at the same level as multiresolution wavelet methods. However, the latter are easier (not requiring iteractive processes) and faster [$0(N)$] to apply than MEM.

A possible handicap of denoising methods based on soft thresholding of wavelet coefficients as well as other non-linear methods are the non-Gaussian features introduced in the reconstructed map. However, one can still detect the possible intrinsic non-Gaussianity of the CMB signal by studying it in the signal plus noise map using the wavelet coefficient distribution. Moreover, a valid reconstruction can be obtained by applying a 'hard' thresholding linear method as discussed in the text.

In a different work, we are studying the case of using a wavelet method based on two scales of dilation (Sanz et al., in preparation). Although this method has the advantage of keeping information on two different scales, for the purpose of denoising, both methods give comparable results. The linear hard thresholding method is expected to perform better for 2D wavelets than for multiresolution ones as the former works with many more resolution levels.

Summarizing, the main advantages of the wavelet method are: providing local information of the contribution from different scales, being computationally very fast $0(N)$, the absence of tuning

parameters, and most importantly, the good performance on denoising CMB maps. The best reconstruction is achieved using soft thresholding techniques. Concerning the Gaussianity of the signal one can apply the suggested linear method for denoising. Moreover, the soft thresholding technique will provide a good reference map and power spectrum for the signal, that can be used to check the quality of other reconstructions based on linear methods. Wavelets are also expected to be a very valuable tool for analysing future CMB maps such as those that will be provided by future missions like MAP and Planck.

ACKNOWLEDGMENTS

We acknowledge helpful discussions with B. Dugnol, C. Fernández, J. M. Gutíerrez, A. W. Jones and L. Tenorio. This work has been supported by the Spanish DGES Project no. PB95-1132-C02-02, Spanish CICYT Acción Especial no. ESP98-1545-E. JLS, LC, EMG and RBB acknowledge finantial support from the USA-Spain Science and Technology Program, ref. 98138. RBB has been supported by a Spanish MEC fellowship. LT acknowledges partial financial support from the Italian ASI and CNR.

REFERENCES

Bennett C. et al., 1996, MAP homepage http://map.gsfc.nasa.gov

Bouchet F. R., Gispert R., Puget J. L., 1996, in Dwek E., ed., AIP Conf. 348, Unvealing the Cosmic Infrared Background. AIP Press, New York, p. 255

Daubechies I., 1988, Comm. Pure Appl. Math., 41, 909

Donoho D. L., Johnstone I. M., 1995, J. Am. Statistical Assoc., 90, 1200

Hobson M. P., Lasenby A. N., 1998, MNRAS, 298, 905

Hobson M. P., Jones A. W., Lasenby A. N., 1998a, astro-ph/9810200

Hobson M.P., Jones A.W., Lasenby A.N., Bouchet F.R., 1998b, MNRAS, 1300, 1

Hobson M. P., Barreiro R. B., Toffolatti L., Lasenby A. N., Sanz J. L., Jones A. W., Bouchet F. R., 1999, MNRAS, 306, 232

Jawerth B., Sweldens W., 1994, SIAM review, 36, 377

Mandolesi N. et al., 1998, Proposal submitted to ESA for the Planck Low Frequency Instrument

Misiti M., Misiti Y., Oppenheim G., Poggi J.M., 1996, Matlab Wavelet Toolbox User's Guide. The MathWorks Inc., Natick, MA

Ogden, R. T., 1997, Essential Wavelets for Statistical Applications and Data Analysis, Birkhauser, Boston

Pando J., Valls-Gabaud D., Fang L., 1998, PRL, 81, 4568

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1994, Numerical Recipes in Fortran, 2539Cambridge University Press, Cambridge

Puget J. L. et al., 1998, Proposal submitted to ESA for the Planck High Frequency Instrument

Sanz J. L., Argüeso F., Cayón L., Martínez-González E., Barreiro R. B., Toffolatti L., 1998, in Banday A. J., Sheth R. K., Da Costa L. N., eds, Proc. MPA/ESO Cosmology Conf., Evolution of Large-Scale: From Recombination to Garching. Enschede, Print Partners Ipskamp, p. 53

Slezak E., de Lapparent V., Bijaoui A., 1993, ApJ, 409, 517

Tegmark M., Efstathiou G., 1996, MNRAS, 281, 129

Tenorio L., Jaffe A. H., Hanany S., Lineweaver C. H., 1999, astro-ph/9903206

This paper has been typeset from a TEX/LATEX file prepared by the author.