MADAM – a map-making method for CMB experiments

E. Keihänen,[1,2][⋆] H. Kurki-Suonio[1] and T. Poutanen[2]

[1]*University of Helsinki, Department of Physical Sciences, PO Box 64, FIN-00014, Helsinki, Finland*
[2]*Helsinki Institute of Physics, PO Box 64, FIN-00014, Helsinki, Finland*

ABSTRACT

We present a new map-making method for cosmic microwave background (CMB) measurements. The method is based on the destriping technique, but it also utilizes information about the noise spectrum. The low-frequency component of the instrument noise stream is modelled as a superposition of a set of simple base functions, whose amplitudes are determined by means of maximum-likelihood analysis, involving the covariance matrix of the amplitudes. We present simulation results with $1/f$ noise and show a reduction in the residual noise with respect to ordinary destriping. This study is related to *Planck* Low Frequency Instrument (LFI) activities.

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

1 INTRODUCTION

Construction of the cosmic microwave background (CMB) map from time-ordered data (TOD) is an important part of the data analysis of CMB experiments. Future space missions like *Planck*[1] present new challenges for the data analysis. The amount of data *Planck* produces is far larger than that of any earlier experiments.

The destriping technique (Burigana et al. 1997; Delabrouille 1998; Maino et al. 1999, 2002; Keihänen et al. 2004) provides an efficient map-making method for large data sets. The method is non-optimal in accuracy but fast and stable. Other methods have been developed which aim at finding the optimal minimum-variance map for *Planck*-like data (e.g. Borrill 1999; Doré et al. 2001; Natoli et al. 2001).

In this paper we present a new map-making method for CMB measurements, called MADAM (Map-making through Destriping for Anisotropy Measurements). The method is built on the destriping technique but unlike ordinary destriping, it also utilizes information about the noise spectrum. The aim is to improve the accuracy of the output map as compared to destriping, while still keeping, at least partly, the speed and stability of the destriping method.

The basic idea of the method is the following. The low-frequency component of the instrument noise in the TOD is modelled as a superposition of simple base functions, whose amplitudes are determined by means of maximum-likelihood analysis, involving the covariance matrix of the amplitudes. The covariance matrix is computed from the noise spectrum, assumed to be known.

This paper is organized as follows. In Section 2 we go through the maximum-likelihood analysis that forms the basis of our map-making method. We describe the map-making algorithm in Section 3. In Section 4 we consider the covariance matrix of compo-

nent amplitudes. Some technical calculations related to this section are presented in Appendix A. We give results from numerical simulations in Section 5 and present our conclusions in Section 6.

2 MAP-MAKING PROBLEM

2.1 Maximum-likelihood analysis of the destriping problem

In the following we present a maximum-likelihood analysis on which our map-making method is based. The analysis is similar to that presented in Keihänen et al. (2004), the main difference being that here we include the covariance of the correlated noise component.

We write the time-ordered data (TOD) stream as

$$y = Pm + n'. \tag{1}$$

Here the first term presents the CMB signal and the second term presents noise. The vector m presents the pixelized CMB map and pointing matrix P spreads it into TOD.

We divide the noise contribution into a correlated noise component and white noise, and model the correlated part as a linear combination of some orthogonal base functions,

$$n' = Fa + n. \tag{2}$$

The vector a contains the unknown amplitudes of the base functions and the matrix F spreads them into TOD. Each column of the matrix F contains the values of the corresponding base function along the TOD. Assuming that the white noise component and the correlated noise component are independent, the total noise covariance is given by

$$C_t = \langle n'(n')^T \rangle = FC_aF^T + C_n \tag{3}$$

where $C_n = \langle nn^T \rangle$ is the white noise covariance, $C_a = \langle aa^T \rangle$ is the covariance matrix for the component amplitudes a, and $\langle x \rangle$ denotes the expectation value of quantity x.

[⋆]E-mail: elina.keihanen@helsinki.fi
[1] http://astro-estec.esa.nl/SA-general/Projects/Planck

© 2005 RAS

Our aim is to find the maximum-likelihood estimate of m and a simultaneously, for given data y. We consider the likelihood of the data, which by the chain rule of probability theory can be written as

$$P(y) = P(y|m, a)P(a|m)P(m). \tag{4}$$

Here $P(a|b)$ denotes the conditional probability of a under condition b. Now $P(m)$ is constant, since we are treating the underlying CMB sky as deterministic (we associate no probability distribution to it). The probability distribution of a is independent of the map so that $P(a|m) = P(a)$. We assume Gaussianity and write

$$P(a) = (2\pi \det \mathbf{C}_a)^{-1/2} \exp\left(-\frac{1}{2}a^{\mathrm{T}}\mathbf{C}_a^{-1}a\right). \tag{5}$$

With m and a fixed, the likelihood of the data is given by the white noise distribution

$$P(y|m, a) = (2\pi \det \mathbf{C}_n)^{-1/2} \exp\left(-\frac{1}{2}n^{\mathrm{T}}\mathbf{C}_n^{-1}n\right) \tag{6}$$

where now $n = y - \mathbf{F}a - \mathbf{P}m$. The white noise covariance $\mathbf{C}_n$ is assumed to be diagonal, but not necessarily uniform.

Maximizing the likelihood (4) is equivalent to minimizing the inverse of its logarithm. We obtain the chi-square minimization function

$$\begin{aligned}
\chi^2 &= -2\ln P(y) = -2\ln[P(y|m, a)P(a)] \\
&= (y - \mathbf{F}a - \mathbf{P}m)^{\mathrm{T}}\mathbf{C}_n^{-1}(y - \mathbf{F}a - \mathbf{P}m) \\
&\quad + a^{\mathrm{T}}\mathbf{C}_a^{-1}a + \text{constant.}
\end{aligned} \tag{7}$$

We want to minimize (7) with respect to both a and m. Minimization with respect to m gives

$$m = \left(\mathbf{P}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{P}\right)^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{C}_n^{-1}(y - \mathbf{F}a). \tag{8}$$

Substituting equation (8) back into equation (7) we get the chi-square into the form

$$\chi^2 = (y - \mathbf{F}a)^{\mathrm{T}}\mathbf{Z}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}(y - \mathbf{F}a) + a^{\mathrm{T}}\mathbf{C}_a^{-1}a, \tag{9}$$

where

$$\mathbf{Z} = \mathbf{I} - \mathbf{P}\left(\mathbf{P}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{P}\right)^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{C}_n^{-1}. \tag{10}$$

Here $\mathbf{I}$ denotes the unit matrix.

We minimize χ^2 with respect to a, to obtain an estimate for the amplitude vector a. The solution is given by

$$\left(\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F} + \mathbf{C}_a^{-1}\right)a = \mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}y. \tag{11}$$

Here we have used the property $\mathbf{Z}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z} = \mathbf{C}_n^{-1}\mathbf{Z}$.

The MADAM algorithm uses the conjugate gradient technique to solve vector a from (11). Note that the matrix on the left-hand side is symmetric. An estimate for the CMB map can then be computed using equation (8).

2.2 Comparison to the minimum-variance solution

If the underlying CMB map is treated as deterministic, noise is Gaussian, and its statistical properties are known, the optimal minimum-variance map is given by

$$m = \left(\mathbf{P}^{\mathrm{T}}\mathbf{C}_t^{-1}\mathbf{P}\right)^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{C}_t^{-1}y \tag{12}$$

where $\mathbf{C}_t$ is the covariance of the noise TOD.

In the following we show that if the total noise covariance is of the form (3), the map estimate given by equations (8) and (11) equals the minimum-variance map (12).

We develop the solution (11) into a Taylor series as

$$a = \left(\mathbf{C}_a - \mathbf{C}_a\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F}\mathbf{C}_a + \cdots\right)\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}y. \tag{13}$$

We write $y - \mathbf{F}a$ out and recollapse the resulting expansion, to get

$$y - \mathbf{F}a = \left(\mathbf{I} + \mathbf{F}\mathbf{C}_a\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{Z}\right)^{-1}y. \tag{14}$$

We now use equation (3) and write $\mathbf{F}\mathbf{C}_a\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1} = \mathbf{C}_t\mathbf{C}_n^{-1} - \mathbf{I}$. Using this and writing $\mathbf{Z}$ out we arrive at

$$\begin{aligned}
&\mathbf{P}^{\mathrm{T}}\mathbf{C}_n^{-1}(y - \mathbf{F}a) \\
&= \mathbf{P}^{\mathrm{T}}\left[\mathbf{C}_t - \left(\mathbf{C}_t\mathbf{C}_n^{-1} - \mathbf{I}\right)\mathbf{P}\left(\mathbf{P}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{P}\right)^{-1}\mathbf{P}^{\mathrm{T}}\right]^{-1}y \\
&= \left[\mathbf{I} - \mathbf{P}^{\mathrm{T}}\left(\mathbf{C}_n^{-1} - \mathbf{C}_t^{-1}\right)\mathbf{P}\left(\mathbf{P}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{P}\right)^{-1}\right]^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{C}_t^{-1}y.
\end{aligned} \tag{15}$$

In the last equality we have taken $\mathbf{C}_t$ out from the left and used the identity $\mathbf{A}(\mathbf{I}+\mathbf{B}\mathbf{A})^{-1} = (\mathbf{I}+\mathbf{A}\mathbf{B})^{-1}\mathbf{A}$, which is easily verified by expanding both sides as Taylor series. The MADAM solution for the map (equation 8) then becomes

$$m = \left[\mathbf{P}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{P} - \mathbf{P}^{\mathrm{T}}\left(\mathbf{C}_n^{-1} - \mathbf{C}_t^{-1}\right)\mathbf{P}\right]^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{C}_t^{-1}y \tag{16}$$

which readily simplifies into (12).

If the chosen base function set accurately models the correlated noise component, the CMB map estimate given by MADAM equals the minimum-variance solution. This is necessarily true at the limit where the number of base functions L per ring approaches the number of samples n, since the base functions then form a complete orthogonal basis. In practice, however, it is not possible to use that many base functions, since both the required memory and CPU time increase with increasing L so that the method finally becomes unfeasible.

3 IMPLEMENTATION

3.1 Map-making algorithm

Equations (8) and (11) form the basis of the MADAM map-making method. In this section we consider the implementation of the method.

Our starting point is a *Planck*-like scanning strategy, where the detector scans the CMB sky in circles which fall on top of each other on consecutive rotations of the instrument. In order to reduce the amount of data, the data from consecutive scan circles is averaged, a process called 'co-adding'. We call one co-added circle a 'ring'. In the nominal *Planck* scanning strategy, the same circle is scanned 60 times before repointing. We denote by M the number of co-added circles. On each ring, we model the correlated noise component as a linear combination of some simple orthogonal arithmetic functions, such as sine and cosine functions or Legendre polynomials. In the simplest case we fit only uniform baselines.

The MADAM algorithm can easily be generalized to a scanning strategy with no co-adding, by setting the co-adding factor M equal to one. We consider both types of scanning strategy in the simulation section. If no co-adding is done, one can choose any length of data to represent a ring. The concept of ring then loses its connection to the spin period and its length becomes a freely chosen parameter.

The most frequently used parameters and symbols of this paper have been collected in Table 1.

The MADAM map-making algorithm consists of the following steps.

(i) Choose a set of base functions to model the correlated noise component and compute the covariance matrix $\mathbf{C}_a$ of their amplitudes. The computation of the covariance matrix is discussed in Section 4.

Table 1. Main parameters and symbols used in this paper and values used in simulation.

Symbol	Parameter	Value
n	number of samples/ring	4608
N	number of rings	8640
M	co-adding factor	60
L	number of base functions	1–65
f_s	sampling frequency	76.8 Hz
f_{sp}	spin frequency of the spacecraft	1/60 Hz
σ	white noise standard deviation (std)	2700 μK
f_{min}	minimum freq. (noise spectrum)	10^{-5} Hz
f_{kn}	knee frequency (noise spectrum)	0.1 Hz
γ	slope of the noise spectrum	−1.0
$\boldsymbol{m}$	pixelized CMB map	
$\mathbf{P}$	pointing matrix	
$\mathbf{C}_n$	white noise covariance	
$\mathbf{F}$	base function matrix	
$\boldsymbol{a}$	base function amplitudes	
$\tilde{\boldsymbol{a}}$	reference values for $\boldsymbol{a}$	
$\mathbf{C}_a$	covariance of amplitudes $\boldsymbol{a}$	
$\mathbf{C}^k$	kth component covariance	
b_k	. . . and its coefficient	
g_k	kth characteristic frequency	

(ii) Using the conjugate gradient technique, solve $\boldsymbol{a}$ from equation (11). The tricky part here is the evaluation of the term $\mathbf{C}_a^{-1}\boldsymbol{a}$, since matrix $\mathbf{C}_a$ is very large. For *Planck*-like data its dimension typically varies from thousands to hundreds of thousands. Fortunately, the matrix has symmetries, which allow us to evaluate $\mathbf{C}_a^{-1}\boldsymbol{a}$ in a quite efficient manner, as we show in Section 3.2.

(iii) Solve the CMB map according to equation (8). This means simple binning of the destriped TOD into pixels, weighting by the white noise covariance. Here we have used HEALPix[2] (Górski, Hivon & Wandelt 1999; Górski et al. 2005) pixelization.

3.2 Evaluation of $\mathbf{C}_a^{-1}\boldsymbol{a}$

The conjugate gradient solution of equation (11) requires that we evaluate

$$\boldsymbol{x} = \mathbf{C}_a^{-1}\boldsymbol{a} \tag{17}$$

several times for different $\boldsymbol{a}$. Here $\boldsymbol{a}$ and $\boldsymbol{x}$ are vectors of (NL) elements, where N is the number of rings and L is the number of base functions per ring. Matrix $\mathbf{C}_a$ has dimension (NL, NL) and is thus expensive to invert. However, $\mathbf{C}_a$ has symmetries which allow us to perform the inversion quite efficiently.

We use index notation in the following. Evaluation of equation (17) is equivalent to solving x_{il} from

$$a_{il} = \sum_{i'l'} C_{a,ii'll'} x_{i'l'}. \tag{18}$$

Here i, i' label rings and l, l' label base functions. The matrix has the symmetry property $C_{a,ii'll'} = C_{a,i'il'l}$.

We assume that the properties of the correlated noise component do not change with time. Matrix $\mathbf{C}_a$ then depends on indices i, i' only through their difference, being thus approximately circulant in indices i, i'. The matrix can be stored as a table of L^2N elements.

A general symmetric matrix equation of moderate size can be solved by Cholesky decomposition. Crout's algorithm to find the

decomposition is given, e.g. in Press et al. (1992). On the other hand, circulant matrix equations can be solved by the Fourier technique.

We solve equation (18) using a combined technique, where we apply Cholesky decomposition to the indices l, l', and a Fourier technique to indices i, i'.

We drop indices i, i' for a while and introduce the following notation. We denote by $\hat{C}_{ll'}$ an (N, N) submatrix of $\mathbf{C}_a$. Now $\hat{C}$ can be understood as an (L, L) matrix whose elements are themselves (N, N) matrices. Similarly, we denote by $\hat{a}_l$ an N-element subvector of $\boldsymbol{a}$.

We now have

$$\hat{a}_l = \sum_{l'} \hat{C}_{ll'}\hat{x}_{l'} \tag{19}$$

where it must be understood that each term $\hat{C}_{ll'}\hat{x}_{l'}$ involves a matrix multiplication of order N. Matrix $\hat{C}$ has the symmetry property $\hat{C}_{ll'} = (\hat{C}_{l'l})^{\mathrm{T}}$. In particular, the diagonal elements $\hat{C}_{ll}$ are symmetric.

We apply Crout's algorithm to equation (19). We follow the procedure presented in Press et al. (1992), keeping in mind that instead of scalar elements we are now operating with matrices.

We want to decompose $\hat{C}$ into

$$\hat{C}_{lk} = \sum_j \hat{L}_{lj}\hat{L}_{kj}^{\mathrm{T}} \tag{20}$$

where $\hat{L}_{lj} = 0$ for $j > l$. Here $\hat{L}$ is a lower triangular matrix, whose elements are again (N, N) matrices. Note that the transpose sign refers to the element $\hat{L}_{kj}$, not $\hat{L}$ itself.

We write

$$\hat{C}_{lk} = \hat{L}_{lk}\hat{L}_{kk}^{\mathrm{T}} + \sum_{j<k} \hat{L}_{lj}\hat{L}_{kj}^{\mathrm{T}} \quad (l > k) \tag{21}$$

$$\hat{C}_{ll} = \hat{L}_{ll}\hat{L}_{ll}^{\mathrm{T}} + \sum_{j<l} \hat{L}_{lj}\hat{L}_{lj}^{\mathrm{T}}. \tag{22}$$

From this we can solve for the elements of $\hat{L}$,

$$\hat{L}_{lk} = \left[\hat{L}_{kk}^{-1}(\hat{C}_{lk}^{\mathrm{T}} - \sum_{j<k} \hat{L}_{kj}\hat{L}_{lj}^{\mathrm{T}})\right]^{\mathrm{T}} \quad (l > k) \tag{23}$$

$$\hat{L}_{ll} = \left[\hat{C}_{ll} - \sum_{j<l} \hat{L}_{lj}\hat{L}_{lj}^{\mathrm{T}}\right]^{1/2}. \tag{24}$$

For each l, we first use equation (23) to solve $\hat{L}_{lk}$ for $k = 1, \ldots, l-1$ and then equation (24) to solve $\hat{L}_{ll}$. Once $\hat{L}$ is known, elements $\hat{x}_k$ can be solved by back-substitution as

$$\hat{z}_k = \hat{L}_{kk}^{-1}\left(\hat{a}_k - \sum_{j<k} \hat{L}_{kj}\hat{z}_j\right) \qquad k = 1, \ldots, L \tag{25}$$

$$\hat{x}_k = \hat{L}_{kk}^{-1}\left(\hat{z}_k - \sum_{j>k} \hat{L}_{jk}^{\mathrm{T}}\hat{x}_j\right) \qquad k = L, \ldots, 1. \tag{26}$$

The procedure presented above contains operations between (N, N) matrices and vectors of length N. These (N, N) matrices are nearly circulant, except for the corners, where they do not 'wrap around' like circulant matrices do. However, the deviation is small, and we may treat the matrices as circulant.

Circulant matrix operations are most conveniently carried out in Fourier space. We pad each vector with zeros up to the next power

of 2, and use the fast Fourier transform (FFT) technique to perform the Fourier transforms.

To each elementary operation involving a circulant matrix there corresponds an operation in Fourier space. The corresponding Fourier operations are the following:

(i) matrix multiplication – element by element multiplication of Fourier transforms;
(ii) matrix transpose – complex conjugate of the Fourier transform;
(iii) square root of a matrix – square root of the Fourier transform;
(iv) matrix inversion – inverse of the Fourier transform.

The procedure for determining $x = \mathbf{C}_a^{-1}a$ can be summarized as follows. First perform a Fourier transform to the circulant ring index $i - i'$ of $\mathbf{C}_a$. Then carry out Cholesky decomposition in index l as described above, and store the resulting $\hat{L}$ matrix.

For each vector a, carry out a Fourier transform along the ring index i, perform back-substitution as given by (25)–(26), and do an inverse Fourier transform to obtain x.

The total operation count of the above procedure is proportional to $L^3 N \ln N$, as contrasted to $L^3 N^3$ of normal matrix inversion. Furthermore, the decomposition can be done 'in place' in the space of $L^2 N$ elements, instead of $L^2 N^2$.

3.3 Covariance of the output map

The covariance of the output map of MADAM is given by

$$\mathbf{C}_m = \left[\mathbf{P}^{\mathrm{T}}\left(\mathbf{C}_n + \mathbf{F}\mathbf{C}_a\mathbf{F}^{\mathrm{T}}\right)^{-1}\mathbf{P}\right]^{-1}, \tag{27}$$

assuming that the noise is well modelled by the noise model (2). The inverse of (27) can be put into the form

$$\mathbf{C}_m^{-1} = \mathbf{P}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{P} - \mathbf{P}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{F}\left(\mathbf{C}_a^{-1} + \mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{F}\right)^{-1}\mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{P}. \tag{28}$$

We presented in Section 3.2 a procedure for evaluating $\mathbf{C}_a^{-1}a$ for arbitrary a. By running this procedure L times one can compute the inverse of matrix $\mathbf{C}_a$. The matrix $\mathbf{C}_a^{-1} + \mathbf{F}^{\mathrm{T}}\mathbf{C}_n^{-1}\mathbf{F}$ can then again be decomposed and stored using the same procedure. When this is done, one can then easily compute any element of $\mathbf{C}_m^{-1}$ using formula (28).

4 COVARIANCE OF THE COMPONENT AMPLITUDES

4.1 General

In this section we consider the computation of the covariance matrix $\mathbf{C}_a$.

First we define reference values for the amplitude vector a. Suppose the actual co-added noise TOD, denoted by u, is known. We consider here the correlated noise component only. We fit the noise model $\mathbf{F}a$ to the noise stream. A least-squares fit gives

$$\tilde{a} = (\mathbf{F}^{\mathrm{T}}\mathbf{F})^{-1}\mathbf{F}^{\mathrm{T}}u. \tag{29}$$

Equation (29) gives the reference values or best estimates for the amplitude vector. We compute the covariance matrix as

$$\mathbf{C}_a = \langle\tilde{a}\tilde{a}^{\mathrm{T}}\rangle. \tag{30}$$

Now let y_p be the original, unco-added noise stream. We assume that the noise is stationary and its autocorrelation $c_{p-p'} = \langle y_p y_{p'}\rangle$ is known.

The co-added noise stream is

$$u_{ij} = \frac{1}{M}\sum_{m=0}^{M-1} y_{[iMn+mn+j]}. \tag{31}$$

Here n is the number of samples per ring and M is the number of co-added circles ($M = 60$ for the nominal *Planck* scanning strategy). The index $i = 0, \ldots, N - 1$ labels rings, $j = 0, \ldots, n - 1$ labels samples on a ring, and $m = 0, \ldots, M - 1$ labels circles co-added into a ring.

Let F_{lj} be the values of the base functions $l = 1, \ldots, L$ on a ring. We assume orthogonality and write

$$\bar{F}_{lj} = \left(\sum_{j'} F_{lj'}^2\right)^{-1} F_{lj}. \tag{32}$$

The reference values for component amplitudes are, according to (29), given by

$$\tilde{a}_{il} = \sum_{j=0}^{n-1} \bar{F}_{lj}u_{ij} = \sum_{j=0}^{n-1} \bar{F}_{lj}\frac{1}{M}\sum_{m=0}^{M-1} y_{[iMn+mn+j]}. \tag{33}$$

For uniform baselines, for instance, the reference value is simply the average of the noise over the ring.

Next we calculate the theoretical covariance of the reference values (33). The elements of the covariance matrix are given by

$$C_{a,ii'll'} = \langle\tilde{a}_{il}\tilde{a}_{i'l'}\rangle$$
$$= \sum_{j,j'=0}^{n-1} \bar{F}_{lj}\bar{F}_{l'j'}\frac{1}{M^2}\sum_{m,m'=0}^{M-1} \langle y_{[inM+mn+j]}y_{[i'nM+m'n+j']}\rangle$$
$$= \sum_{j,j'=0}^{n-1} \bar{F}_{lj}\bar{F}_{l'j'}\frac{1}{M^2}\sum_{m,m'=0}^{M-1} c_{[(i-i')nM+(m-m')n+j-j']}. \tag{34}$$

The sum over m, m' can be combined into one sum over $m'' = m - m'$,

$$C_{a,ii'll'} = \sum_{j,j'=0}^{n-1} \bar{F}_{lj}\bar{F}_{l'j'}\sum_{m''=-M}^{M} \frac{M - |m''|}{M^2}c_{[(i-i')nM+m''n+j-j']}. \tag{35}$$

This is a general formula for the elements of the covariance matrix $\mathbf{C}_a$.

4.2 Exponential expansion of the autocorrelation function

Covariance (35) is rather heavy to evaluate computationally, since it contains a three-dimensional sum. However, if the autocorrelation function can be expanded in exponential functions, the covariance can be computed in a very efficient way. This holds, e.g. for $1/f$ and $1/f^2$ noise.

Suppose the autocorrelation function can be expanded as

$$c(t) = \sum_k b_k c_k(t) \tag{36}$$

where

$$c_k(t) = \exp(-g_k|t|) \tag{37}$$

where g_k is a selected set of characteristic frequencies and and b_k are coefficients to be determined.

We denote by $\mathbf{C}^k$ the covariance matrix that corresponds to an exponential autocorrelation function of the form (37). Once the

component covariances $\mathbf{C}^k$ and coefficients b_k have been determined, the total covariance $\mathbf{C}_a$ can be computed as

$$\mathbf{C}_a = \sum_k b_k \mathbf{C}^k. \tag{38}$$

The component covariances $\mathbf{C}^k$ can be computed in a very efficient way making use of the basic properties of the exponent function.

Expanding the autocorrelation as (36) is equivalent to expanding the power spectrum as

$$P(f) = \sum_k b_k \frac{2g_k}{g_k^2 + (2\pi f)^2}. \tag{39}$$

The expansion does not exist for arbitrary noise spectra. In Appendix A we calculate the expansion explicitly for a power-law spectrum of the form

$$P(f) = \left(\frac{f_{kn}}{f}\right)^\gamma \frac{\sigma^2}{f_s} \quad (f > f_{min}) \tag{40}$$

for $-2 \leqslant \gamma < 0$. Here f_{kn} is the knee frequency, at which the spectral power equals the white noise power σ^2/f_s, σ is the white noise std, and f_s is the sampling frequency.

For $1/f$ noise ($\gamma = -1$) the expansion is particularly simple. If the characteristic frequencies g_k are chosen uniformly in logarithmic scale, the correct spectrum is obtained with

$$b_k = 2\sigma^2 \frac{f_{kn}}{f_s} \Delta, \tag{41}$$

where Δ is the logarithmic interval in g_k.

The $1/f$ spectrum, as given by the expansion (39) with coefficients (41), is shown in Fig. 1. The $1/f$ form holds inside the frequency range $f_{min} < f < f_{max}$ spanned by the characteristic frequencies g_k. Below f_{min} the spectrum levels out, as can be seen from the figure.

Another simple example is the $1/f^2$ spectrum ($\gamma = -2$). In that case the desired spectrum is given by one single g component with coefficient

$$b = \frac{2\pi^2 f_{kn}^2}{g} \frac{\sigma^2}{f_s} \tag{42}$$
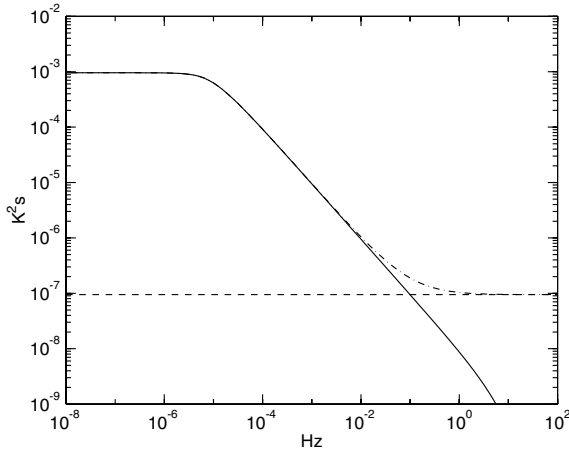
and $g = 2\pi f_{min}$.

Figure 1. $1/f$ noise power spectrum, as given by expansion (39) and equation (41). The solid line shows the pure $1/f$ noise spectrum. The dashed line shows the white noise level ($P_{wn} = \sigma^2/f_s$). The dash-dotted line shows the total noise spectrum, including both components. The parameters were $\sigma = 2700$ μK, $f_s = 76.8$ Hz, $f_{kn} = 0.1$ Hz, $\Delta = 1$, $f_{min} = 10^{-5}$ Hz, $f_{max} = 10$ Hz.

4.3 Component covariance matrices

In this section we give explicit formulae for the elements of the component covariance matrix $\mathbf{C}^k$ that corresponds to the autocorrelation function (37). The derivation is given in Appendix A. Here we just quote the results.

We use again the index notation, where indices i, i' refer to rings and l, l' to base functions. The elements of the component covariance matrix are given by

$$C_{ii'll'}^k = G_k S_{kl}^+ S_{kl'}^- \exp\left[-\frac{g_k}{f_{sp}}(i - i' - 1)M\right] \quad (i > i') \tag{43}$$

$$C_{ii'll'}^k = G_k S_{kl}^- S_{kl'}^+ \exp\left[-\frac{g_k}{f_{sp}}(i' - i - 1)M\right] \quad (i < i') \tag{44}$$

$$\begin{aligned} C_{ii'll'}^k = \frac{1}{M} \sum_{j=0}^{n-1} (\bar{F}_{lj}\bar{F}_{l'j} + \bar{F}_{lj}A_{l'j} + \bar{F}_{l'j}A_{lj}) \\ + G_k^0 \left(S_{kl}^+ S_{kl'}^- + S_{kl}^- S_{kl'}^+\right). \quad (i = i') \end{aligned} \tag{45}$$

Here $f_{sp} = f_s/n$, where n is the number of samples on a ring, represents the spin frequency of the instrument. If no co-adding is applied, n can be chosen freely, and f_{sp} does not need to have any connection to the scanning pattern of the instrument. In that case f_{sp} represents simply the inverse of the chosen baseline length.

Factors S^+ and S^- are defined as

$$S_{kl}^+ = \sum_{j=0}^{n-1} \bar{F}_{lj} \exp\left(-\frac{g_k}{f_s} j\right) \tag{46}$$

$$S_{kl}^- = \sum_{j=0}^{n-1} \bar{F}_{lj} \exp\left[-\frac{g_k}{f_s}(n - j)\right]. \tag{47}$$

Co-adding brings in the factors

$$G_k = \frac{1}{M^2} \frac{[1 - \exp(-\frac{g_k}{f_{sp}}M)]^2}{[1 - \exp(-\frac{g_k}{f_{sp}})]^2} \tag{48}$$

and

$$G_k^0 = \frac{1}{M} \frac{1}{1 - \exp(-\frac{g_k}{f_{sp}})} \left[1 - \frac{1}{M} \frac{1 - \exp(-\frac{g_k}{f_{sp}}M)}{1 - \exp(-\frac{g_k}{f_{sp}})}\right]. \tag{49}$$

If no co-adding is done ($M = 1$), we have $G_k = 1$ and $G_k^0 = 0$.

Factor A_{lj} is defined by

$$A_{lj} = \sum_{j'=0}^{j-1} \bar{F}_{lj'} \exp\left[-\frac{g_k}{f_s}(j - j')\right] \tag{50}$$

and can be computed rapidly using the recurrence relation

$$A_{lj} = (A_{l,j-1} + \bar{F}_{l,j-1}) \exp\left(-\frac{g_k}{f_s}\right), \tag{51}$$

with the starting value $A_{l0} = 0$.

Formulae (43)–(45) are very fast to evaluate numerically, as compared to the general formula (35).

Fig. 2 presents the theoretical covariance, computed using expansion (38), between uniform baselines for $1/f$ noise with $f_{kn} = 0.1$ Hz. Other parameters used were $f_s = 76.8$ Hz, $f_{sp} = 1/60$ Hz, $n = 4608$ and $M = 60$. We show in the same figure the covariance as determined from simulated $1/f$ noise. We generated 10 realizations of noise TOD of one year length, and computed their
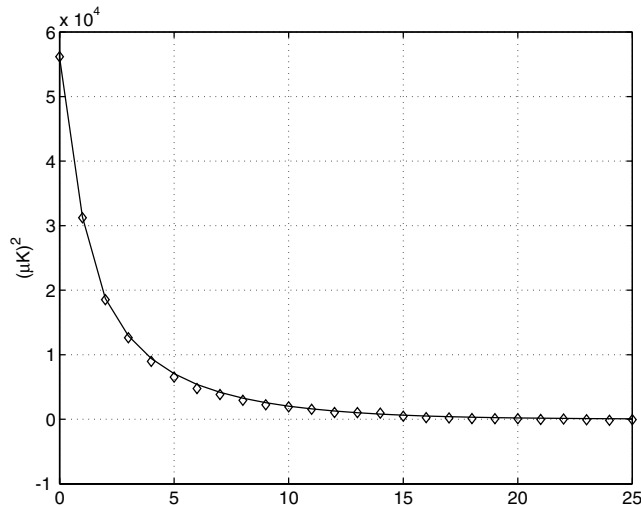
Figure 2. Covariance between uniform baselines, as a function of the distance between rings. The solid line shows the theoretical curve, computed using equations (43)–(45). Diamonds present the values determined from simulated noise. The noise parameters were $\sigma = 2700\ \mu K$, $f_s = 76.8$ Hz, $f_{kn} = 0.1$ Hz, $\Delta = 1$, $f_{min} = 10^{-5}$ Hz, $f_{max} = 10$ Hz.

Table 2. The first elements of the covariance matrix $\mathbf{C}_a$ [in $(\mu K)^2$]. The noise parameters were the same as in Fig. 2. We used Fourier components as base functions, normalized as $\sum_j F_{lj}^2 = n$. We show elements $l, l' = 1, \ldots,$ 5 and $i - i' = 0, \ldots, 3$. Index value $l = 1$ refers to the uniform baseline and values $l = 2$ and $l = 3$ ($l = 4$ and $l = 5$) to the sine and cosine of the first (second) Fourier mode, respectively. The noise parameters were the same as in Fig. 2.

l	$i - i'$	$l' = 1$	$l' = 2$	$l' = 3$	$l' = 4$	$l' = 5$
1	0	56049	0	−2.41	0	−0.668
	1	31324	−89.6	1.10	−44.9	0.307
	2	18707	−28.9	0.0531	−14.5	0.0133
	3	12928	−16.3	0.0210	−8.17	5.25e-3
2	0	0	161	0	1.64	0
	1	89.6	−1.42	0.209	−0.745	0.0696
	2	28.9	−0.0751	2.35e-4	−0.0376	5.88e-5
	3	16.3	−0.0297	5.90e-5	−0.0148	1.48e-5
3	0	−2.41	0	158	0	−0.123
	1	1.10	−0.209	0.133	−0.139	0.0617
	2	0.0531	−2.35e-4	1.14e-6	−1.18e-4	2.85e-7
	3	0.0210	−5.90e-5	1.73e-7	−2.95e-5	4.32e-8
4	0	0	1.64	0	79.9	0
	1	44.9	−0.745	0.139	−0.401	0.0524
	2	14.5	−0.0376	1.18e-4	−0.0188	2.94e-5
	3	8.17	−0.0148	2.95e-5	−7.42e-3	7.38e-6
5	0	−0.668	0	−0.123	0	79.0
	1	0.307	−0.0696	0.0617	−0.0524	0.0334
	2	0.0133	−5.88e-5	2.85e-7	−2.94e-5	7.13e-8
	3	5.25e-3	−1.48e-5	4.32e-8	−7.38e-6	1.08e-8

autocorrelation using the Fourier technique. The agreement is very good.

As another example we show in Table 2 the first elements of the covariance matrix for Fourier components. Index $l = 1$ refers to the uniform baseline and indices $l = 2$ and $l = 3$ ($l = 4$ and $l = 5$) to the sine and cosine of the first (second) Fourier mode, respec-

tively. We have normalized all components to $\sum_j F_{lj}^2 = n$. Elements of matrix $\mathbf{F}$ are thus $F_{1j} = 1$, $F_{2j} = \sqrt{2}\sin(2\pi j/n)$, $F_{3j} = \sqrt{2}\cos(2\pi j/n)$, $F_{4j} = \sqrt{2}\sin(4\pi j/n)$, $F_{5j} = \sqrt{2}\cos(4\pi j/n)$, and $\bar{\mathbf{F}}$ is given by $\bar{F}_{lj} = F_{lj}/n$. We see that the dominant elements are those corresponding to uniform baselines.

5 SIMULATION RESULTS

5.1 Data sets

We have produced two sets of simulated TOD. We refer to them as 'co-added' and 'unco-added' data sets.

The co-added data set mimicks the one year TOD from one *Planck* Low Frequency Instrument (LFI) 70-GHz detector. The scanning pattern was the following. The spin axis remained in the equatorial plane and was turned 2.4 arcmin every hour, so that after 8640 h the spin axis had turned 360 degrees. The detector turned around the spin axis with an opening angle of 85 deg and spin frequency $f_{sp} = 1/60$ Hz. The sampling frequency was $f_s = 76.8$ Hz. We co-added data of 60 consecutive spin circles to form a ring. Our total data set consisted of 8640 rings, with 4608 samples in each. The sky coverage was 0.9964.

The unco-added data set was generated with a quite similar scanning pattern. The main difference was that instead of moving in steps, the spin axis turned continuously at the rate of 360 degrees in 8640 min. The sampling and spin frequencies as well as the opening angle were the same as in the first data set. Because the spin axis moved continuously, consecutive circles did not fall on top of each other, and no co-adding was done. The total length of the TOD was the same as in the co-added data set, i.e. 8640 × 4608 samples. The amount of data was equivalent to 6 d of one detector *Planck* data, spread over the whole sky. This scanning pattern was rather artificial, but our purpose was only to demonstrate the use of MADAM in the case of unco-added data. Full-scale simulations with realistic unco-added one-year *Planck* data are beyond the scope of this paper.

The underlying CMB map was created by the SYNFAST code of the HEALPIX package (Górski et al. 1999, 2005), starting from the CMB anisotropy angular power spectrum computed with the CMBFAST[3] code (Seljak & Zaldarriaga 1996) using the cosmological parameters $\Omega_{tot} = 1.00$, $\Omega_\Lambda = 0.7$, $\Omega_b h^2 = 0.02$, $h = 0.7$, $n = 1.00$, and $\tau = 0.0$. We created the input map with HEALPIX resolution $N_{side} = 2048$ and with a symmetric Gaussian beam with full width at half maximum (FWHM) of 14 arcmin. We then formed the signal TOD by picking temperatures from this map. Our output maps have resolution parameter $N_{side} = 512$, corresponding to an angular resolution of 7 arcmin.

We used the stochastic differential equation (SDE) technique to create the instrument noise stream, which we added to the signal TOD. We generated noise with power spectrum

$$P(f) = \left(1 + \frac{f_{kn}}{f}\right)\frac{\sigma^2}{f_s}, \quad (f > f_{min}) \tag{52}$$

with parameters $\sigma = 2700\ \mu K$, knee frequency $f_{kn} = 0.1$ Hz, and $f_{min} = 10^{-5}$ Hz. The white noise level 2700 μK (CMB temperature scale) corresponds to the estimated white noise level of one 70-GHz LFI detector. We used the same noise spectrum for both data sets.

We run our code on one processor of an IBM eServer Cluster 1600 supercomputer.

---

[3] http://www.cmbfast.org

Table 3. Estimated rms of the residual $1/f$ noise map (see text for explanation) for different numbers of base functions, for co-added data. We fit Fourier components and Legendre polynomials. We show also the number of iteration steps and CPU time in the Fourier case.

L	Legendre rms/µK	Fourier rms/µK	iter	CPU/s
1	19.232	19.232	16	59
2	18.801			
3	18.591	18.351	20	65
4	18.331			
5	18.147	17.970	28	106
7	17.914	17.749	28	123
9	17.752	17.604	32	164
11	17.634	17.503	32	191
15	17.482	17.370	36	267
25	17.284	17.213	40	530
35	17.194	17.137	44	758
45	17.141	17.096	52	1154
65		17.057	56	1959

5.2 Results for co-added data

We show our results for the co-added data set in Tables 3–5 and in Figs 3–5.

We created a reference map by co-adding the pure signal TOD into a map and computed the total residual noise map by subtracting the reference map from the output map. We computed the rms of the residual map (rms_tot) and averaged it over 10 noise realizations. The rms values show little variation, because they are dominated by white noise. Therefore we have used as a figure of merit the estimated rms of the residual $1/f$ noise, which was computed as follows. Using the white noise sigma (2700 µK) and the known distribution of measurements in the sky we obtained rms_wn = 108.95 µK for the expected contribution from white noise to the residual noise map. We computed an estimate for the rms of the residual $1/f$ noise as rms = sqrt((rms_tot)^2-(rms_wn)^2). Tables 3 and 4 quote these rms values.

The results for different numbers of base functions are given in Table 3. We tried two sets of base functions: Fourier components and Legendre polynomials. Because a Fourier fit always includes a uniform baseline plus an equal number of sine and cosine functions, the total number L of base functions is always an odd number. The rms values continue improving when we increase the number of base functions. Fourier components give lower rms values than Legendre polynomials for the same number of components. In the rest of the simulations in this section we fitted only Fourier components.

We show also the number of iterations and total CPU time taken by one run in the case of Fourier components. Since the CPU time naturally depends on the computer used and may vary from run to run, the values quoted should not be taken too seriously. They merely give an idea how the run time increases with increasing number of base functions.

In Fig. 3 we have plotted the rms values against the inverse of the number of base functions. At the limit $1/L \to 0$ the rms seems to be approaching the value 16.97 µK. We expect that to be the std of the minimum-variance map (Section 2.2).

In Fig. 4 we show the power spectrum of the residual noise map for $L = 1$, $L = 5$, and $L = 25$ (Fourier components), as an average over 10 noise realizations. Note that these spectra include the white noise contribution. The spectra look very similar. We therefore used the $L = 1$ spectrum of Fig. 4 as a reference spectrum, which we subtracted from other spectra. The differences are shown in Fig. 5, for

Table 4. Effect of misestimating the noise spectrum. We show the estimated $1/f$ residual map rms (in µK) for different numbers of Fourier components and for different assumed knee frequencies f_{kn} (upper panel), spectral slopes γ (middle), and minimum frequencies f_{min} (lower panel). The other two noise parameters were kept at their correct values ($f_{kn} = 0.1$ Hz, $\gamma = -1.0$, and $f_{min} = 10^{-5}$ Hz). The lowest rms value on each row is denoted by an asterisk. The correct parameter value is shown in boldface.

		f_{kn}		
L	0.03 Hz	0.05 Hz	**0.1 Hz**	0.2 Hz
1	19.232	19.232	19.232	19.232
3	18.396	18.345*	18.351	18.398
5	18.056	17.961*	17.970	18.083
7	17.888	17.754	17.749*	17.904
9	17.786	17.624	17.604*	17.788
11	17.721	17.538	17.503*	17.709
15	17.641	17.428	17.370*	17.603
25	17.556	17.326	17.213*	17.485
		γ		
L	-0.6	-0.8	$\mathbf{-1.0}$	-1.2
1	19.266	19.233	19.232*	19.232*
3	18.384	18.340*	18.351	18.373
5	18.008	17.951*	17.970	18.009
7	17.799	17.733*	17.749	17.788
9	17.666	17.592*	17.604	17.641
11	17.576	17.495*	17.503	17.536
15	17.461	17.369*	17.370	17.396
25	17.340	17.226	17.213*	17.230
35	17.291	17.159	17.137*	17.151
		f_{min}		
L	10^{-6} Hz	$\mathbf{10^{-5}}$ **Hz**	10^{-4} Hz	10^{-3} Hz
1	19.232*	19.232*	19.233	19.350
3	18.351*	18.351*	18.351*	18.501
5	17.970*	17.970*	17.971	18.220
7	17.749*	17.749*	17.750	18.039
9	17.604*	17.604*	17.605	17.926
11	17.503*	17.503*	17.505	17.849
15	17.370*	17.370*	17.371	17.748
25	17.213*	17.213*	17.215	17.626

Table 5. Pixelization noise. Rms of the residual map for noise-free TOD. The error is due to finite pixelization of the sky.

L	rms/µK
1	0.138
3	0.207
5	0.233
7	0.257
9	0.274
11	0.289
15	0.319
25	0.359
35	0.380
45	0.393

$L = 3, 5, 11, 25$. The spectra were binned into bins of 10 multipoles for clarity. Fig. 5 shows that increasing the number of Fourier components systematically reduces the residual noise at multipoles above $l = 60$.
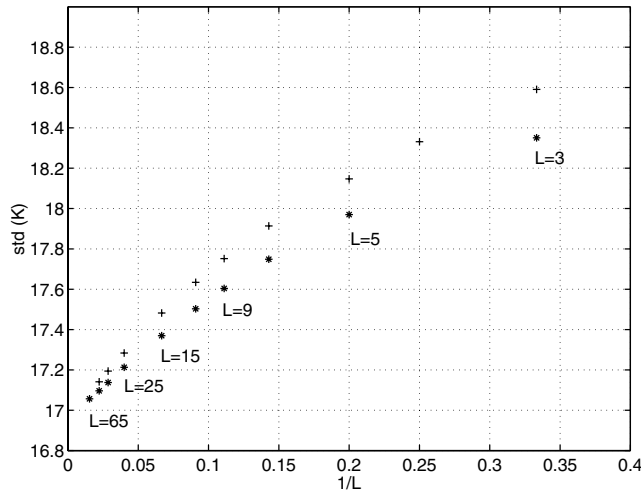
Figure 3. Rms of the residual $1/f$ noise map plotted against the inverse of the number of base functions, for Legendre polynomials (+) and Fourier components (*). The numerical values are given in Table 3. The rms seems to converge towards 16.97 μK at the limit $L \to n$.
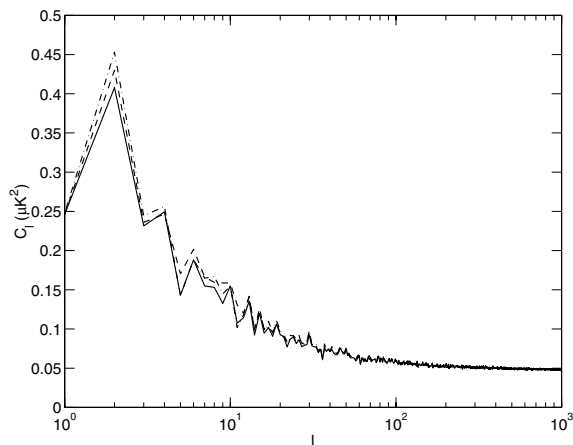

Figure 4. Power spectrum of the residual noise map (including white noise), for $L = 1$ (dot-dashed), $L = 5$ (dashed), and $L = 25$ (solid). We fit Fourier components.
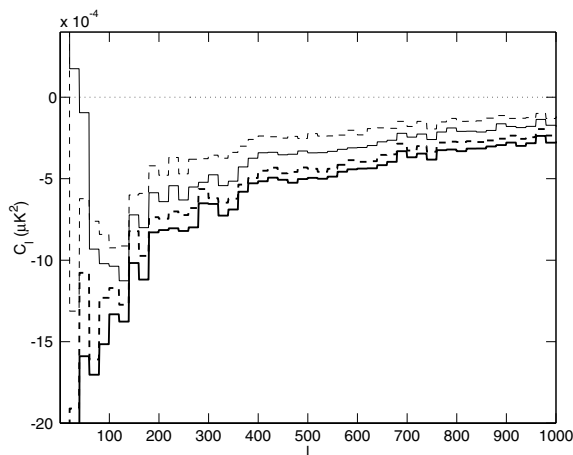

Figure 5. Difference of the power spectrum of the residual noise map, with respect to the $L = 1$ case, as bins of 10. From above: $L = 3$ (thin dashed), $L = 5$ (thin solid), $L = 11$ (thick dashed), and $L = 25$ (thick solid).

We have compared results of fitting uniform baselines using MADAM and ordinary destriping technique. We got the destriping results by running MADAM with $\mathbf{C}_a^{-1} = 0$. At this limit the method reduces to pure destriping. The destriping result was 19.232 18 μK (19.232 12 μK with MADAM). This indicates that when fitting uniform baselines only, the covariance plays a very small role, but the baselines can be determined from the data alone with good accuracy.

Keihänen et al. (2004) showed that fitting Fourier components beyond the uniform baseline with the ordinary destriping technique, without using the covariance matrix, did not improve the results. In this work we have found a clear improvement. This indicates that information about the noise spectrum is important when fitting base functions other than the uniform baseline.

We have also studied the effect of misestimating the noise spectrum. We varied in turn each of the three noise parameters (knee frequency, spectral slope, and minimum frequency) while keeping the other two at their correct values ($f_{kn} = 0.1$ Hz, $\gamma = -1.0$, $f_{min} = 10^{-5}$ Hz). We then recomputed the covariance matrix $\mathbf{C}_a$ for Fourier components with the new parameter values and rerun the map estimation. The results are shown in Table 4.

It is perhaps surprising that for small L underestimating the knee frequency or assuming a less deep slope seems to improve the results. This can be understood as follows. When L is small, the noise is not perfectly modelled by the base functions. There is an error involved, related to the higher Fourier components that have not been included in the analysis. This error disturbs the estimation of the lower components, leading to a too high variation in their amplitudes. The error in the covariance matrix, caused by misestimation of the noise spectrum, partly compensates for this error. We notice that the best results are obtained with a spectrum (a less deep slope or lower knee frequency) which leads to a smaller covariance for the low-frequency Fourier components. Smaller covariance tends to restrict the variation of the amplitudes, thus decreasing their error also. With larger L the phenomenon disappears, and the lowest rms is obtained with the correct noise spectrum, as expected.

Table 5 shows results from a run with noise-free data. The TOD contained only the contribution from the CMB signal, but no instrument noise. The error that still remains in the map is due to 'pixelization noise' (Doré et al. 2001) caused by the finite size of sky pixels. The pixelization error increases with increasing number of base functions, but is very small compared with the error due to instrument noise.

5.3 Results for unco-added data

If no co-adding is applied, the length n of a ring is not determined by the scanning pattern of the instrument, but is a free parameter to be chosen at will. We then have two parameters to select: the number L of base functions and their length n.

To keep things simple, we tried two schemes. First we kept the baseline length fixed at one minute and varied the number of Fourier components that we fitted. Secondly, we fitted uniform baselines only ($L = 1$) but varied their length. Results from the first case are shown in Table 6. The baseline length was fixed at $n = 4608$ samples (one minute). We show again the estimated rms of the residual $1/f$ noise map, computed in the same way as in the case of co-added data. The white noise level was higher than in the co-added case by a factor of $\sqrt{60}$. The expected white noise contribution to the rms of the residual noise map was 844.0 μK.

Table 7 shows results of fitting uniform baselines of different lengths. The first column gives the length n of the baseline, as the number of samples. The second column gives the baseline length

Table 6. Average residual noise rms for different numbers of Fourier components, for the unco-added data set. The ring length was $n = 4608$. Also shown are the number of iteration steps and the CPU time for one run.

L	rms/μK	iter	CPU s^{-1}
1	149.45	28	68
3	141.52	36	102
5	138.64	48	161
7	136.71	52	205
9	135.72	56	280
11	134.93	60	336
15	133.90	64	436
25	132.63	76	847

Table 7. Results of fitting uniform baselines of different lengths to unco-added data. The first three columns give the baseline length as the number of samples and in seconds, respectively, and the number of baselines per one minute of TOD. The next columns give the estimated residual $1/f$ noise rms, number of iteration steps and total CPU time taken by one run.

n	t s^{-1}	N min^{-1}	rms/μK	iter	CPU s^{-1}
4608	60.0	1	149.45	28	68
2304	30.0	2	144.14	32	80
1152	15.0	4	141.85	36	96
576	7.5	8	137.97	48	123
288	3.75	16	135.26	64	150
144	1.88	32	133.31	84	215
72	0.94	64	132.07	116	331
36	0.47	128	131.41	160	672
18	0.23	256	131.06	224	1284
9	0.12	512	130.91	320	2889

Table 8. Comparison between MADAM and destriping.

n	t s^{-1}	N min^{-1}	MADAM rms/μK	destriping rms/μK
4608	60.0	1	149.45	149.48
2304	30.0	2	144.14	144.26
1152	15.0	4	141.85	147.43
576	7.5	8	137.97	170.85
288	3.75	16	135.26	233.85

that with this extra information the CMB map can be estimated with better accuracy than with pure destriping.

We have tested the method with simulated co-added *Planck*-like data. As a figure of merit we have used the rms of the residual $1/f$ noise map. Our simulations show that fitting more base functions clearly improves the accuracy of the output map, with the cost of increasing requirements for CPU time and memory.

We have shown theoretically that the map estimate given by MADAM approaches the optimal minimum-variance map when the number of fitted base functions increases. In practice it is not possible to reach the exact minimum-variance map using MADAM, due to CPU time and memory limitations. Still, MADAM provides a fast and efficient map-making method. By varying the number of base functions the user may flexibly move from a very fast but less accurate map-making (small L) to a more accurate but more time-consuming map-making (large L), depending on what is desired.

We also demonstrated the use of MADAM for unco-added data. Although the data set used was quite artificial, in the sense that it does not mimick data from any existing CMB experiment, the method was shown to work well for unco-added data also.

The current implementation of the method is a serial one. With a parallelized version we expect to be able to process data sets equivalent to full-year unco-added *Planck* data.

in seconds. The third column shows the number of baselines per minute ($4608/n$). The shortest baseline we tried consisted of only nine samples.

The third column of Table 7 and the first column of Table 6 are comparable, since they give the total number of unknowns per one minute of TOD. We see that for a given number of unknowns, fitting Fourier components works better than fitting uniform baselines. However, when we compare CPU times, we see that fitting uniform baselines is more effective.

As in the case of co-added data, we compared results of fitting uniform baselines using MADAM and the ordinary destriping technique. The results are shown in Table 8. With one-minute baselines the difference between the methods is small, but increases with decreasing n, in favour of MADAM. The rms value obtained with destriping reaches a minimum around 0.5 min baseline length, while with MADAM the values continue improving. With small values of n MADAM is clearly superior to basic destriping.

6 CONCLUSIONS

We have presented a new map-making method for CMB experiments called MADAM. The method is based on the well-known destriping technique, but unlike basic destriping, it also uses information on the known statistical properties of the instrument noise in the form of the covariance matrix of the base function amplitudes. We have shown

REFERENCES

Borrill J., 1999, preprint (astro-ph/9911389)
Burigana C., Malaspina M., Mandolesi N., Danese L., Maino D., Bersanelli M., Maltoni M., 1997, Int. Rep. TeSRE/CNR, 198/1997 (astro-ph/9906360)
Delabrouille J., 1998, A&AS, 127, 555
Doré O., Teyssier R., Bouchet F. R., Vibert D., Prunet S., 2001, A&A, 374, 358
Górski K. M., Hivon E., Wandelt B. D., 1999, in Banday A. J., Sheth R. S., Da Costa L., eds, Proc. MPA/ESO Cosmology Conference Evolution of Large-Scale Structure. PrintPartners Ipskamp, the Netherlands, p. 37 (astro-ph/9812350)
Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelman M., 2005, ApJ, 622, 759
Keihänen E., Kurki-Suonio H., Poutanen T., Maino D., Burigana C., 2004, A&A, 428, 287

Maino D. et al., 1999, A&AS, 140, 383

Maino D., Burigana C., Górski K. M., Mandolesi N., Bersanelli M., 2002, A&A, 387, 356

Natoli P., de Gasperis G., Gheller C., Vittorio N., 2001, A&A, 372, 346

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical Recipes, 2nd edn. Cambridge Univ. Press, Cambridge

Seljak U., Zaldarriaga M., 1996, ApJ, 469, 437

APPENDIX A: COVARIANCE MATRIX FOR A POWER-LAW POWER SPECTRUM

We discussed the computation of the covariance matrix $\mathbf{C}_a$ in Section 4. In this appendix we present some technical calculations which were omitted there.

A1 Exponential expansion for a power-law spectrum

Assume that the autocorrelation function of the noise can be expanded as

$$c(t) = \sum_k b_k \exp(-g_k|t|) \tag{A1}$$

where g_k is a selected set of characteristic frequencies. We now derive the coefficients b_k in the case of a power-law power spectrum of the form

$$P(f) \propto f^\gamma \tag{A2}$$

for $-2 \leqslant \gamma < 0$.

The power spectrum $P(f)$ and the autocorrelation function $c(t)$ of stationary noise are related by the cosine transform

$$P(f) = \int_{-\infty}^{\infty} c(t) \cos(2\pi f t)\, \mathrm{d}t. \tag{A3}$$

The autocorrelation function $\exp(-g|t|)$ corresponds to the power spectrum

$$P(f, g) = \frac{2g}{g^2 + (2\pi f)^2}. \tag{A4}$$

The total power spectrum corresponding to the autocorrelation (A1) is given by

$$P(f) = \sum_k b_k P(f, g_k) = \sum_k \frac{2 b_k g_k}{g_k^2 + (2\pi f)^2}. \tag{A5}$$

We pick the frequencies g_k uniformly in a logarithmic scale inside some range $[f_{\min}, f_{\max}]$ and use the Ansatz

$$b_k = A g_k^{\gamma+1} \Delta \tag{A6}$$

where $\Delta = \ln(g_{k+1}/g_k)$ is the logarithmic step in g and A is a constant to be determined. The total power spectrum becomes

$$P(f) = A \sum_k \frac{2 g_k^{\gamma+2}}{g_k^2 + (2\pi f)^2} \Delta. \tag{A7}$$

We transform the sum into an integral

$$P(f) = A \int_{f_{\min}}^{f_{\max}} \frac{2 g^{\gamma+2}}{g^2 + (2\pi f)^2} \frac{\mathrm{d}g}{g}$$
$$\approx A \int_0^{\infty} \frac{2 g^{\gamma+1}}{g^2 + (2\pi f)^2}\, \mathrm{d}g \quad (f_{\min} \ll f \ll f_{\max}). \tag{A8}$$

The integral converges for $-2 < \gamma < 0$,

$$P(f) = A \frac{\pi (2\pi f)^\gamma}{\sin[(\gamma + 2)\pi/2]} \quad \propto f^\gamma. \tag{A9}$$

We choose

$$A = \frac{\sigma^2}{f_s} \frac{1}{\pi} (2\pi f_{\mathrm{kn}})^{-\gamma} \sin[(\gamma + 2)\pi/2] \tag{A10}$$

to obtain the desired power spectrum

$$P(f) = \frac{\sigma^2}{f_s} \left(\frac{f}{f_{\mathrm{kn}}} \right)^\gamma \quad (f_{\min} \ll f \ll f_{\max}). \tag{A11}$$

Here σ is the white noise std, f_s is the sampling frequency, and f_{kn} is the knee frequency, at which the power of the power-law noise component equals the white noise power σ^2/f_s. The maximum $f_{\max}$ should well exceed the knee frequency.

The coefficients b_k are given by

$$b_k = \frac{\sigma^2}{f_s} \frac{1}{\pi} (2\pi f_{\mathrm{kn}})^{-\gamma} \sin[(\gamma + 2)\pi/2] g_k^{\gamma+1} \Delta \tag{A12}$$

for $-2 < \gamma < 0$. In particular, for $1/f$ noise ($\gamma = -1$) we have the simple formula

$$b_k = 2\sigma^2 \frac{f_{\mathrm{kn}}}{f_s} \Delta. \tag{A13}$$

The case $\gamma = -2$ requires a separate treatment. We see directly from equation (A4) that the desired spectrum is given by one single g component with coefficient

$$b = \frac{2\pi^2 f_{\mathrm{kn}}^2}{g} \frac{\sigma^2}{f_s}. \tag{A14}$$

The spectrum then has the form

$$P(f) = \frac{f_{\mathrm{kn}}^2}{f_{\min}^2 + f^2} \frac{\sigma^2}{f_s} \tag{A15}$$

where $f_{\min} = g/(2\pi)$. The spectrum behaves as $\propto f^{-2}$ at $f \gg f_{\min}$ and levels out below $f_{\min}$.

A2 Component covariance matrices

Once the expansion (A1) has been found, the covariance matrix can be computed as

$$\mathbf{C}_a = \sum_k b_k \mathbf{C}^k. \tag{A16}$$

In the following we calculate the component covariance matrices $\mathbf{C}^k$.

In Section 4 we derived the general formula

$$C_{ii'll'} = \sum_{j,j'=0}^{n-1} \bar{F}_{lj} \bar{F}_{l'j'} \frac{1}{M^2} \sum_{m,m'=0}^{M-1} c_{[(i-i')nM+(m-m')n+j-j']}. \tag{A17}$$

Assume then that the autocorrelation function is of the exponential form

$$\langle y_p y_{p'} \rangle = c_{k,p-p'} = \exp(-g_k|t|) = \exp\left(-\frac{g_k}{f_s} |p - p'| \right) \tag{A18}$$

where f_s is the sampling frequency and indices p, p' label samples along the TOD. We substitute this into equation (A17). The covariance becomes

$$C_{ii'll'}^k = \sum_{jj'=0}^{n-1} \bar{F}_{lj} \bar{F}_{l'j'} \frac{1}{M^2} \sum_{m,m'=0}^{M-1}$$
$$\times \exp\left(-\frac{g_k}{f_s} |(i-i')nM + (m-m')n + j - j'| \right). \tag{A19}$$

We treat the cases $|i - i'| > 0$ and $i = i'$ separately.

(i) $|i - i'| > 0$. If $i - i' > 0$ the quantity inside the brackets is positive, and we can split the four-dimensional sum into a product of four sums as

$$
C^k_{ii'll'} = \exp\left[-\frac{g_k}{f_s}(i - i' - 1)nM\right]
$$
$$
\times \frac{1}{M^2}\sum_{m=0}^{M-1}\exp\left(-\frac{g_k}{f_s}mn\right)
$$
$$
\times \sum_{m=0}^{M-1}\exp\left[-\frac{g_k}{f_s}(M-1-m)n\right]
$$
$$
\times \sum_{j=0}^{n-1}\bar{F}_{lj}\exp\left(-\frac{g_k}{f_s}j\right)
$$
$$
\times \sum_{j'=0}^{n-1}\bar{F}_{l'j'}\exp\left[-\frac{g_k}{f_s}(n-j')\right]. \tag{A20}
$$

We have arranged the terms in such a way that the argument of an exponent function is always negative. This is helpful in numerical evaluation. The sum over m, m' can be carried out analytically, yielding

$$
G_k = \frac{1}{M^2}\sum_{m=0}^{M-1}\exp\left(-\frac{g_k}{f_s}mn\right)\sum_{m=0}^{M-1}\exp\left[-\frac{g_k}{f_s}(M-1-m)n\right]
$$
$$
= \frac{1}{M^2}\frac{\left[1 - \exp\left(-\frac{g_k}{f_{sp}}M\right)\right]^2}{\left[1 - \exp\left(-\frac{g_k}{f_{sp}}\right)\right]^2} \tag{A21}
$$

where $f_{sp} = f_s/n$.

The elements for which $i - i' < 0$ are obtained from the symmetry relation, $C^k_{ii'll'} = C^k_{i'il'l}$.

(ii) The case $i = i'$ is more complicated, since the quantity inside the brackets in equation (A19) takes both positive and negative values. We split the sum into three terms ($m = m'$, $m > m'$, and

$m < m'$) and the $m = m'$ term further into three terms ($j = j'$, $j > j'$, and $j < j'$),

$$
C^k_{iill'} = \sum_{jj'=0}^{n-1}\bar{F}_{lj}\bar{F}_{l'j'}\frac{1}{M^2}\sum_{m,m'=0}^{M-1}\exp\left[-\frac{g_k}{f_s}|(m-m')n + j - j'|\right]
$$
$$
= \frac{1}{M^2}\sum_{m=0}^{M-1}\sum_{m'<m}\exp\left[-\frac{g_k}{f_s}(m-m'-1)n\right]
$$
$$
\times \sum_{j=0}^{n-1}\bar{F}_{lj}\exp\left(-\frac{g_k}{f_s}j\right)\sum_{j'=0}^{n-1}\bar{F}_{l'j'}\exp\left[-\frac{g_k}{f_s}(n-j')\right]
$$
$$
+ \frac{1}{M^2}\sum_{m'=0}^{M-1}\sum_{m<m'}\exp\left[-\frac{g_k}{f_s}(m'-m-1)n\right]
$$
$$
\times \sum_{j=0}^{n-1}\bar{F}_{lj}\exp\left[-\frac{g_k}{f_s}(n-j)\right]\sum_{j'=0}^{n-1}\bar{F}_{l'j'}\exp\left(-\frac{g_k}{f_s}j\right)
$$
$$
+ \frac{1}{M}\sum_{j=0}^{n-1}\bar{F}_{lj}\bar{F}_{l'j} + \frac{1}{M}\sum_{j=0}^{n-1}\bar{F}_{lj}\sum_{j'<j}\bar{F}_{l'j'}\exp\left[-\frac{g_k}{f_s}(j-j')\right]
$$
$$
+ \frac{1}{M}\sum_{j'=0}^{n-1}\bar{F}_{l'j'}\sum_{j<j'}\bar{F}_{lj}\exp\left[-\frac{g_k}{f_s}(j'-j)\right]. \tag{A22}
$$

The sum over m, m' can again be calculated analytically,

$$
G_k^0 = \frac{1}{M^2}\sum_{m=0}^{M-1}\sum_{m'<m}\exp\left[-\frac{g_k}{f_{sp}}(m-m'-1)\right]
$$
$$
= \frac{1}{M}\frac{1}{1-\exp(-\frac{g_k}{f_{sp}})}\left[1 - \frac{1}{M}\frac{1-\exp(-\frac{g_k}{f_{sp}}M)}{1-\exp(-\frac{g_k}{f_{sp}})}\right]. \tag{A23}
$$

Formula (A22) may seem complicated, but is easy and fast to evaluate numerically.

Equations (A20) and (A22), when written in compact form, give the formulae (43)–(45) in Section 4.

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