

A Map-Making algorithm for the Planck Surveyor

P. Natoli¹, G. de Gasperis¹, C. Gheller², and N. Vittorio¹

¹ Dipartimento di Fisica, Università di Roma “Tor Vergata”, via della Ricerca Scientifica 1, 00133 Roma, Italy

² Cineca, Via Magnanelli 6/3, 40033 Caselecchio di Reno (BO), Italy

Received 16 January 2001 / Accepted 9 March 2001

Abstract. We present a parallel implementation of a map-making algorithm for CMB anisotropy experiments which is both fast and efficient. We show for the first time a Maximum Likelihood, minimum variance map obtained by processing the entire data stream expected from the PLANCK Surveyor, under the assumption of a symmetric beam profile. Here we restrict ourselves to the case of the 30 GHz channel of the PLANCK Low Frequency Instrument. The extension to PLANCK higher frequency channels is straightforward. If the satellite pointing periodicity is good enough to average data that belong to the same sky circle, then the code runs very efficiently on workstations. The serial version of our code also runs on very competitive time-scales the map-making pipeline for current and forthcoming balloon borne experiments.

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

1. Introduction

Making CMB maps out of balloon (e.g. QMAP, de Oliveira-Costa et al. 1998; MAXIMA-1, Hanany et al. 2000; BOOMERanG, de Bernardis et al. 2000) or space borne (MAP¹ and the PLANCK Surveyor²) experiments is an important step of the standard data analysis pipeline. It allows for: a major lossless compression of the raw data with a minimum number of assumptions; checks on systematics; tests for and/or estimation of foreground contamination; etc. Map-making is also important when simulating a CMB experiment. It allows to: look for possible systematics; optimize the focal plane design and/or the experiment scanning strategy; etc. It is then highly desirable to develop tools able to attack the map-making problem also for forthcoming, high-resolution CMB space experiments.

Map-making a la COBE (see e.g. Lineweaver 1994) can be extended to the differential, high resolution MAP experiment (Wright et al. 1996). Moreover Wright (1996) has discussed how to perform map-making in the case of one-horned experiments significantly affected by $1/f$ noise, which introduces spurious correlations in the data time stream (see e.g. Janssen et al. 1996). To our knowledge this algorithm has not yet been implemented for the PLANCK Surveyor mission. To date, maps for PLANCK have only been produced by means of *destriping*

algorithms (Delabrouille et al. 1998; Maino et al. 1999). We remind that PLANCK spins at 1 rpm, has a boresight angle of 85° and observes the same circle of the sky for 1 hour (every hour the spin axis is moved, say along the ecliptic, by $2.5'$). The destriping algorithms implemented for PLANCK assume that, after averaging data taken along a given circle (i.e. in one hour), the residual noise along the circle is well approximated by white noise plus an offset. These offsets are determined by minimizing the set of all possible differences between the antenna temperatures of the same sky pixel observed in two different scan circles. This method rests on a number of assumptions, the stronger being that the knee frequency of the $1/f$ noise is smaller than or at most comparable to the spin frequency. Wright's method does not suffer this limitation because it assumes to know the statistical properties of the noise, directly derived from the data themselves (see e.g. Ferreira & Jaffe 2000; Prunet et al. 2000).

The purpose of this paper is to present the first implementation of Wright's method to the PLANCK Surveyor, for the moment assuming a symmetric beam profile. In fact, we want to show, to our knowledge for the first time, the analysis of the entire time stream (14 months) of PLANCK simulated data to produce Maximum Likelihood, minimum variance CMB maps (we stress that maps obtained from destriping algorithms are not necessarily minimum variance). The parallel, Message Passing Interface (hereafter MPI) implementation of the algorithm has been tested on a SGI Origin 2000 (hereafter O2K) and runs on time scales that might render Monte Carlo simulations feasible. This opens up the possibility of evaluating the CMB angular power spectrum via Monte Carlo

Send offprint requests to: P. Natoli,
e-mail: Paolo.Natoli@roma2.infn.it

¹ <http://map.gsfc.nasa.gov/>

² <http://astro.estec.esa.nl/SA-general/Projects/Planck/>

techniques (Wandelt et al. 2000) rather than by maximizing a Likelihood (see e.g. Bond et al. 1998) or by directly evaluating the map angular correlation function (Szapudi et al. 1999). We want to stress that our implementation does not assume any exact periodicity of sky pointing within single circle observation (i.e. no average on the circle is performed, as instead required by destriping algorithms). If, on the contrary, it turns out that pointing periodicity is a good assumption (i.e. average on circles is performed) then the same code runs very efficiently on medium sized workstations (e.g. Pentium based PCs), again in quite short time scales.

Although this paper might seem rather technical, we think that it can be of interest to a large community of CMB data analysts, involved either in the PLANCK collaboration or in the forthcoming, new generation balloon experiments.

The plan of this paper is as follows. In Sects. 2 and 3 we will briefly discuss the method and its implementation. In Sect. 4 we will show tests and benchmarks of our implemented software. In Sect. 5 we will apply our tools to PLANCK and BOOMERanG simulated data. Finally, in Sect. 6 we will briefly review our conclusions.

2. Method

2.1. Statement of the problem

Let us for completeness outline here the map-making algorithm and its assumptions. The primary output of a CMB experiment are the Time Ordered Data (TOD), \mathbf{d} , which consist of \mathcal{N}_d sky observations made with a given scanning strategy and at a given sampling rate (three points per $FWHM$, say). A map, \mathbf{m} , can be thought as a vector containing \mathcal{N}_p temperature values, associated with sky pixels of dimension $\sim FWHM/3$. Following Wright (1996) we assume that the TOD depend linearly on the map:

$$\mathbf{d} = \mathbf{P}\mathbf{m} + \mathbf{n}, \quad (1)$$

where \mathbf{n} is a vector of random noise and \mathbf{P} is some known matrix³. The rectangular, $\mathcal{N}_d \times \mathcal{N}_p$ matrix \mathbf{P} is dubbed a *pointing matrix*. That is, applying \mathbf{P} on a map “unrolls” the latter on a TOD according to a given scanning strategy. Conversely, applying \mathbf{P}^T on the TOD “sums” them into a map⁴. The structure of \mathbf{P} depends on what we assume for \mathbf{m} . If \mathbf{m} contains a pixelized but unsmeared image of the sky then \mathbf{P} must account for beam smearing. This is the most general assumption and allows one to properly treat, for instance, an asymmetric beam profile. In this case, applying \mathbf{P} to \mathbf{m} implies both convolving the sky pattern with the detector beam response and unrolling

³ Note that the elements of \mathbf{m} need not to be sky temperature values. Other parameters may be fitted provided that they also depend linearly on the TOD. Sometimes these parameters are called “virtual” pixels.

⁴ The value of a pixel of this map is the sum of all the observations of that pixel made at different times according to a given scanning strategy.

\mathbf{m} into a “signal-only” time stream. If, on the other hand, the beam is – at least approximately – symmetric then it is possible – and certainly more convenient – to consider \mathbf{m} as the beam smeared pixelized sky. The structure of \mathbf{P} for a one-horned experiment would then be very simple. Only one element per row would be different from zero, the one connecting the observation of j th pixel to the i th element of the TOD. Hereafter we will restrict ourselves to this case and we will address the implementation of an asymmetric beam profile in a forthcoming paper.

2.2. Least squares approach to map-making

Many methods have been proposed to estimate \mathbf{m} in Eq. (1) (for a review see, e.g. Tegmark 1997). Since the problem is linear in \mathbf{m} , the use of a Generalized Least Squares (GLS) method appears well suited. This involves the minimization of the quantity

$$\chi^2 = \mathbf{n}^T \mathbf{V} \mathbf{n} = (\mathbf{d}^T - \mathbf{m}^T \mathbf{P}^T) \mathbf{V} (\mathbf{d} - \mathbf{P} \mathbf{m})$$

for some nonsingular, symmetric matrix \mathbf{V} . Deriving with respect to \mathbf{m} yields an estimator, say $\tilde{\mathbf{m}}$, for the map:

$$\tilde{\mathbf{m}} = (\mathbf{P}^T \mathbf{V} \mathbf{P})^{-1} \mathbf{P}^T \mathbf{V} \mathbf{d}. \quad (2)$$

The proof that this estimator is unbiased is straightforward. Just note that:

$$\tilde{\mathbf{m}} - \mathbf{m} = (\mathbf{P}^T \mathbf{V} \mathbf{P})^{-1} \mathbf{P}^T \mathbf{V} \mathbf{n}.$$

So, provided that $\langle \mathbf{n} \rangle = 0$, we have $\langle \tilde{\mathbf{m}} \rangle = \mathbf{m}$ (the symbol $\langle \cdot \rangle$ indicates, as usual, an average over the ensemble). The map covariance matrix is, then:

$$\begin{aligned} \Sigma^{-1} &\equiv \langle (\mathbf{m} - \tilde{\mathbf{m}})(\mathbf{m}^T - \tilde{\mathbf{m}}^T) \rangle \\ &= (\mathbf{P}^T \mathbf{V} \mathbf{P})^{-1} \mathbf{P}^T \mathbf{V} \langle \mathbf{n} \mathbf{n}^T \rangle \mathbf{V} \mathbf{P} (\mathbf{P}^T \mathbf{V} \mathbf{P})^{-1}. \end{aligned}$$

In order to have a “low noise” estimator, one has to find \mathbf{V} that minimizes the variance of $\tilde{\mathbf{m}}$. This is attained if we take \mathbf{V} to be the noise inverse covariance matrix, i.e. $\mathbf{V}^{-1} = \mathbf{N} \equiv \langle \mathbf{n} \mathbf{n}^T \rangle$. Then $\tilde{\mathbf{m}}$ has the very nice property of being, amongst all linear and unbiased estimators, the one of minimum variance⁵. The GLS solution to the map-making problem is, then:

$$\tilde{\mathbf{m}} = \Sigma^{-1} \mathbf{P}^T \mathbf{N}^{-1} \mathbf{d} \quad (3)$$

where

$$\Sigma = \mathbf{P}^T \mathbf{N}^{-1} \mathbf{P}. \quad (4)$$

⁵ To see why this is the case, remember that any linear estimator can be written in the form $\hat{\mathbf{m}} = \mathbf{A} \mathbf{d} = \mathbf{A} \mathbf{P} \mathbf{m} + \mathbf{A} \mathbf{n}$. So, the condition $\mathbf{A} \mathbf{P} = \mathbf{I}$ must hold for it to be unbiased. To prove that the variance of $\hat{\mathbf{m}}$ can not be smaller than the variance of $\tilde{\mathbf{m}}$, it helps (see, e.g. Lupton 1993) to write $\hat{\mathbf{m}} = \tilde{\mathbf{m}} + (\hat{\mathbf{m}} - \tilde{\mathbf{m}})$. It follows that

$$\text{Var}(\hat{\mathbf{m}}) = \text{Var}(\tilde{\mathbf{m}}) + \text{Var}(\hat{\mathbf{m}} - \tilde{\mathbf{m}}) + 2\text{Cov}(\tilde{\mathbf{m}}, \hat{\mathbf{m}} - \tilde{\mathbf{m}}).$$

By direct substitution, one finds that $\text{Cov}(\tilde{\mathbf{m}}, \hat{\mathbf{m}} - \tilde{\mathbf{m}}) = (\mathbf{P}^T - \mathbf{I})(\mathbf{P}^T \mathbf{N}^{-1} \mathbf{P})^{-1} = 0$. Thus, the variance of $\hat{\mathbf{m}}$ is equal to the variance of $\tilde{\mathbf{m}}$ plus a nonnegative quantity, which is all we need to show.

2.3. ML solutions

The statistical properties of detector noise are, as usual, described by a multivariate Gaussian distribution. This fact has two consequences. The first one is rather obvious: if the noise distribution is Gaussian, $\tilde{\mathbf{m}}$ is indeed the ML estimator. In fact, in the Gaussian case, the Likelihood of the data time stream given the (true) map is

$$\mathcal{L}(\mathbf{d}|\mathbf{m}) = (2\pi)^{-\mathcal{N}_p/2} \times \exp \left\{ -\frac{1}{2} [(\mathbf{d}^T - \mathbf{m}^T \mathbf{P}^T) \mathbf{N}^{-1} (\mathbf{d} - \mathbf{P}\mathbf{m}) + \text{Tr}(\ln \mathbf{N})] \right\}. \quad (5)$$

Solving for the maximum would clearly give Eqs. (3) and (4) again. The second remark has to do with the notion of a *lossless map*, that is, a map which contains all the relevant cosmological information contained in the TOD. A good way to prove that a method is lossless is to show that the TOD and the related map have the same Fisher information matrix. The estimator defined in Eq. (3) leads to a lossless map (Tegmark 1997).

3. Implementation

The method outlined in Sect. 2 has the advantage of being simple, linear and, hence, computationally appealing for PLANCK. In this section we discuss our implementation of the map-making algorithm to the 30 GHz channel of the PLANCK Low Frequency Instrument (LFI), with a nominal resolution $\sim 30'$ FWHM. This implies $\mathcal{N}_d \sim 10^9$ (14 months of observation, i.e. about two sky coverages) and $\mathcal{N}_p \sim 10^6$. Note, however, that the implementation we present here is rather general since the algorithm does not take specific advantage of the details of PLANCK's scanning strategy and instrumental performances. In fact, we want to stress that our implementation of Wright's algorithm is, in principle, well suited for any one-horned, balloon- or space-borne CMB experiment.

3.1. The noise covariance matrix

As already mentioned in the introduction, the GLS method assumes to know the statistical properties of the noise. They are fully described by \mathbf{N} , the noise covariance matrix in the time domain. It is of course highly desirable to estimate \mathbf{N} directly from the data since there is no a priori guarantee that the noise statistical properties match with what is measured during ground testing. The measured TOD is a combination of signal and noise. Thus, one has to estimate the signal, subtract it from the TOD and only then estimate \mathbf{N} . Ferreira & Jaffe (2000) and Prunet et al. (2000) discuss iterative methods to attack this issue.

The problem of evaluating the noise properties directly from the TOD is a bit far from the main point we want to make here: the possibility, given \mathbf{N} , to analyze the entire PLANCK simulated data set to produce a ML, minimum variance map. In any case, we will show in Sect. 5 that for the PLANCK case, to *zerth* order, we can estimate the

signal by projecting the TOD onto the sky. This implies applying \mathbf{P}^T (summing the TOD into a map), $(\mathbf{P}^T \mathbf{P})^{-1}$ (dividing the pixel values by the number of hits) and \mathbf{P} (unrolling a map into a time stream) to \mathbf{d} . The noise estimator, $\tilde{\mathbf{n}}$, can then be written as follows:

$$\tilde{\mathbf{n}} = \mathbf{d} - \mathbf{P}(\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \mathbf{d} = [\mathbf{I} - \mathbf{P}(\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T] \mathbf{n} \quad (6)$$

and does not include any contribution from the signal. In a forthcoming paper (Natoli et al. 2001) we will address in more detail, and specifically for the PLANCK Surveyor, the problem of estimating the noise statistical properties directly from flight data. Here, in Sect. 5, we will only show that $\tilde{\mathbf{n}}$ allows a reasonably good estimate of the "in-flight" noise behavior.

3.2. Stationary noise

It is customary to assume that the statistical properties of detector noise do not change over the mission life time. The formal way to state this property is to write the elements of the noise covariance matrix as follows:

$$N_{ij} = \xi(|i - j|) \quad (7)$$

or, equivalently, to say that N is a Toeplitz matrix. If ξ vanishes for $|i - j| > \mathcal{N}_\xi \ll \mathcal{N}_d$, then N is band diagonal and it can be well approximated by a circulant matrix:

$$N_{i+1, j+1} = N_{ij} \quad \text{mod } \mathcal{N}_d. \quad (8)$$

Obviously, circularity does not hold exactly for N : there is no physical reason for the noise correlation function to wrap around itself past the edges. However, the number of elements to modify in order to turn N into a circulant matrix is a mere $\mathcal{N}_\xi(\mathcal{N}_\xi + 1) \ll \mathcal{N}_d^2$. A circulant matrix is diagonal in Fourier space: $\mathbf{N} = \mathbf{F}^\dagger \mathbf{\Xi} \mathbf{F}$, where \mathbf{F} (\mathbf{F}^\dagger) is the discrete Fourier transform (antitransform) operator (DFT), and $\mathbf{\Xi}$ is diagonal. It is trivial to show that the inverse of a circulant matrix is still circulant and, in particular, that $\mathbf{N}^{-1} = \mathbf{F}^\dagger \mathbf{\Xi}^{-1} \mathbf{F}$. Circularity ensures that we will not have edge problems: this is exactly the boundary condition imposed by the DFT.

3.3. Non-stationary noise

The ML solution of the map-making problem (see Eq. (3) and Eq. (4)) can be easily generalized to the case of piecewise stationary noise. In this case, the whole TOD can be split into a set $\{\mathbf{d}_{(\kappa)}\}$ of smaller pieces $\kappa = 1, \mathcal{N}_c$. Clearly,

$$\mathbf{d}_{(\kappa)} = \mathbf{P}_{(\kappa)} \mathbf{m} + \mathbf{n}_{(\kappa)}, \quad (9)$$

where $\mathbf{P}_{(\kappa)}$ is a pointing matrix and $\mathbf{n}_{(\kappa)}$ is a vector of random, stationary noise with covariance matrix $\mathbf{N}_{(\kappa)}$. Neglecting cross-correlations between different $\mathbf{n}_{(\kappa)}$'s (a reasonable assumption if $\mathcal{N}_\xi \ll \mathcal{N}_d$) reduces the Likelihood given in Eq. (5) to the product of the Likelihood of the different pieces, which of course do not

need to share the same noise covariance matrix. Thus, in this case, the ML map is given by

$$\tilde{\mathbf{m}} = \left[\sum_{\kappa=1}^{\mathcal{N}_c} \mathbf{P}_{(\kappa)}^T \mathbf{N}_{(\kappa)}^{-1} \mathbf{P}_{(\kappa)} \right]^{-1} \sum_{\kappa=i}^{\mathcal{N}_c} \mathbf{P}_{(\kappa)}^T \mathbf{N}_{(\kappa)}^{-1} \mathbf{d}_{(\kappa)}. \quad (10)$$

Since neglecting the correlation between different $\mathbf{n}_{(\kappa)}$'s is de facto equivalent to treating different sections of the same TOD as independent experiment outcomes, we can obviously think of Eq. (10) as a simple recipe to produce maps from \mathcal{N}_c different horns whose sky coverage is at least partially overlapped.

3.4. The noise inverse covariance matrix

In what follows we do not need to explicitly use \mathbf{N}^{-1} , although we have in principle all the relevant information to do so. In fact, the first row of \mathbf{N}^{-1} , usually called a noise filter in the literature, can be computed by taking a DFT of Ξ^{-1} , the noise inverse spectral density. Then, by exploiting the circularity of \mathbf{N}^{-1} we could in principle reconstruct the whole matrix. Needless to say, handling a $\mathcal{N}_d \times \mathcal{N}_d$ matrix is, for the PLANCK case, unconceivable even for the 30 GHz case.

Fortunately, under the assumption of circularity the application of the noise inverse covariance matrix to the TOD is simply a convolution,

$$\mathbf{N}^{-1} \mathbf{d} = \mathbf{F}^\dagger \Xi^{-1} \mathbf{F} \mathbf{d}, \quad (11)$$

and we do not need to store a huge matrix.

Asking how long a noise filter should be is a very important question. In fact, from the one hand, it is desirable to convolve the TOD with a filter properly sampled in \mathcal{N}_d points (something already numerically feasible even for a PLANCK-like TOD). On the other hand, Eq. (11) works for real data in the limit $\mathcal{N}_\xi \ll \mathcal{N}_d$ (so that \mathbf{N}^{-1} can be considered circulant). Also note that an obvious bound on \mathcal{N}_ξ comes from the instrument: it is difficult to imagine that the PLANCK detectors will remain “coherent” for the whole mission life-time. A further benefit in having $\mathcal{N}_\xi \ll \mathcal{N}_d$ is the possibility of performing a piecewise convolution in Eq. (11). In any case, in Sect. 5 we will show the sensitivity of our results to \mathcal{N}_ξ .

3.5. The conjugate gradient solver

Evaluating $\tilde{\mathbf{m}}$ requires solving Eq. (3), that for convenience we rewrite as

$$\mathbf{H}^{-1} \mathbf{P}^T \mathbf{N}^{-1} (\mathbf{P} \tilde{\mathbf{m}}) = \mathbf{H}^{-1} \mathbf{P}^T \mathbf{N}^{-1} \mathbf{d} \quad (12)$$

where \mathbf{H}^{-1} is a preconditioning matrix (or preconditioner). This linear system can be solved by means of a Conjugate Gradient (CG) minimization technique (see e.g. Axelsson & Barker 1984). CG methods are known to converge efficiently (i.e. in less than the $\mathcal{O}(\mathcal{N}_p^3)$ operations required by standard matrix inversion) if the eigenvalues of Σ spawn a few orders of magnitudes. If this is not the

case, the method should be preconditioned. This implies finding a matrix \mathbf{H} such that its inverse approximates⁶ Σ^{-1} . The method then solves the equivalent system given in Eq. (12). Other than being a good approximation to the original matrix, the preconditioner should also be fast to compute and invert. These often are conflicting requirements: finding a good preconditioner may be hard. In our case, we find more than satisfactory to precondition our system with the diagonal part of the symmetric and positive definite matrix Σ . This choice is motivated by the fact that Σ is diagonally dominant. Furthermore, the diagonal part of Σ is very close to the number of hits per pixel $\mathbf{P}^T \mathbf{P}$ so in practice we are normalizing each row of the inverse covariance matrix to the “redundancy” of the corresponding pixel⁷.

A CG algorithm does not need to explicitly invert Σ . We want to stress that earlier methods implemented to solve the map-making problem for balloon-borne experiments (Borrill 1999) require the inversion of the same matrix, surely an unpleasant task when considering large maps. Such implementations usually rely on the fact that Σ is (analytically speaking) positive definite and, as such, a candidate to be Cholesky decomposed in $\mathcal{O}(\mathcal{N}_p^3)$ operations. This fact makes the procedure prohibitively expensive – even for present day supercomputers – when \mathcal{N}_p becomes large ($\gtrsim 10^5$). Also, Σ may not be exactly positive definite due to small inaccuracies when estimating the noise correlation properties, making Cholesky decomposition critical. Note that using an iterative solver could potentially reduce the above operation count by a factor $\mathcal{N}_p/\mathcal{N}_{\text{iter}}$. However, storage requirement for Σ would still be $\mathcal{O}(\mathcal{N}_p^2)$, prohibitive for PLANCK.

To conclude this Section, let us stress that the map we obtain, $\tilde{\mathbf{m}}$, has the correct noise covariance matrix, Σ , even if we never evaluate nor store it directly.

3.6. Parallelization

The scheme outlined above naturally lends itself to a multi-processor environment. In fact, it is straightforward to divide the entire TOD in a set of smaller, partially overlapped, time streams and force each of the Processor Elements (PE) to perform map-making with its own time stream. This is accomplished by explicitly assigning work

⁶ To be quantitative, an optimal preconditioner would give $\mathbf{H}^{-1} \Sigma = \mathbf{I} + \mathbf{R}$ with all the eigenvalues of \mathbf{R} less than 1 (Axelsson & Barker 1984). It is then clear that multiplying Σ by the preconditioner changes its spectrum to give an equivalent but better conditioned system.

⁷ One other problem should be mentioned: Σ is remarkably ill conditioned. In fact, any monopole term belongs to its kernel since applying \mathbf{N}^{-1} kills the corresponding TOD offset. Our preconditioner does not cure this problem which is instead fixed by specifically imposing that $\tilde{\mathbf{m}}$ has a given (usually zero) average. This prescription, critical if one attempts a matrix inversion, is not needed in our case: we can, of course, always subtract out the monopole.

loads to each PE by means of MPI calls embedded in the Fortran 95 code.

As stated above, the working assumption that validates Eq. (11) is that $\mathcal{N}_\xi \ll \mathcal{N}_d$. Since the number of PE, \mathcal{N}_{PE} , is at most $\mathcal{O}(100)$ the previous condition is even fulfilled for the single PE. Thus, each PE performs the convolution of Eq. (11) by executing its own serial FFT as discussed in Sect. 3.4. This is more efficient than spreading over the PE's an intrinsically parallel FFT, the advantage being that we limit inter-processor communications. However, as a drawback, each PE has to handle a partial map of size \mathcal{N}_p and cross talk among PE's is necessary to merge \mathcal{N}_{PE} of these maps at the end of each CG iteration. Note that increasing the number of processors \mathcal{N}_{PE} while keeping \mathcal{N}_d constant may result in the operation count being dominated by this merging (see however Sect. 4.2 below).

4. Performances

In this Section we benchmark an implementation of the map-making algorithm described above. The benchmarks are produced within the framework of the PLANCK's simulation pipeline. However, since the algorithm does not take specific advantage of PLANCK's scanning strategy and instrumental performance, they are indeed very general. To show that this is the case we will apply our code to a simulation of the BOOMERanG experiment (see Sect. 5.3 below).

4.1. Scaling with \mathcal{N}_ξ , \mathcal{N}_d and \mathcal{N}_p

Part of the l.h.s. of Eq. (12) can be thought as the application of the $P^T N^{-1}$ operator (the same acting on \mathbf{d} on the rhs) to $P \tilde{\mathbf{m}}^{(k)}$, the data stream obtained by unrolling the tentative solution produced by the CG at the k th step. While the application of $P^T N^{-1}$ to \mathbf{d} is done just once, the application of the same operator to $P \tilde{\mathbf{m}}^{(k)}$ must be done for every iteration of the CG algorithm. Therefore, speed of execution becomes critical.

If we use a radix-2 Fast Fourier Transform (FFT) to perform a piecewise convolution of the TOD, the operation count is expected $\propto \mathcal{N}_d \log_2 \mathcal{N}_\xi$. The lion share of CPU time is taken by the (real) DFT transform. The use of an efficient FFT library greatly speeds up the calculation. We used the publicly available FFTW library (Frigo & Johnson 1998). This code has the nice property of tailoring itself to the architecture over which is executed, therefore greatly enhancing cross-platform efficiency. In Fig. 1 (upper panel) we plot CPU time per CG iteration, τ , versus \mathcal{N}_ξ . The behavior shown in the figure is due to the non constant efficiency attained by the FFTW library when performing transforms of different lengths. FFT routines are usually more efficient when processing power-of-two transforms and FFTW is no exception. Thus, we fix \mathcal{N}_ξ to a power of two.

The application of P (P^T) to a map (TOD) is expected to scale linearly with \mathcal{N}_d . In Fig. 1 (middle panel) we give

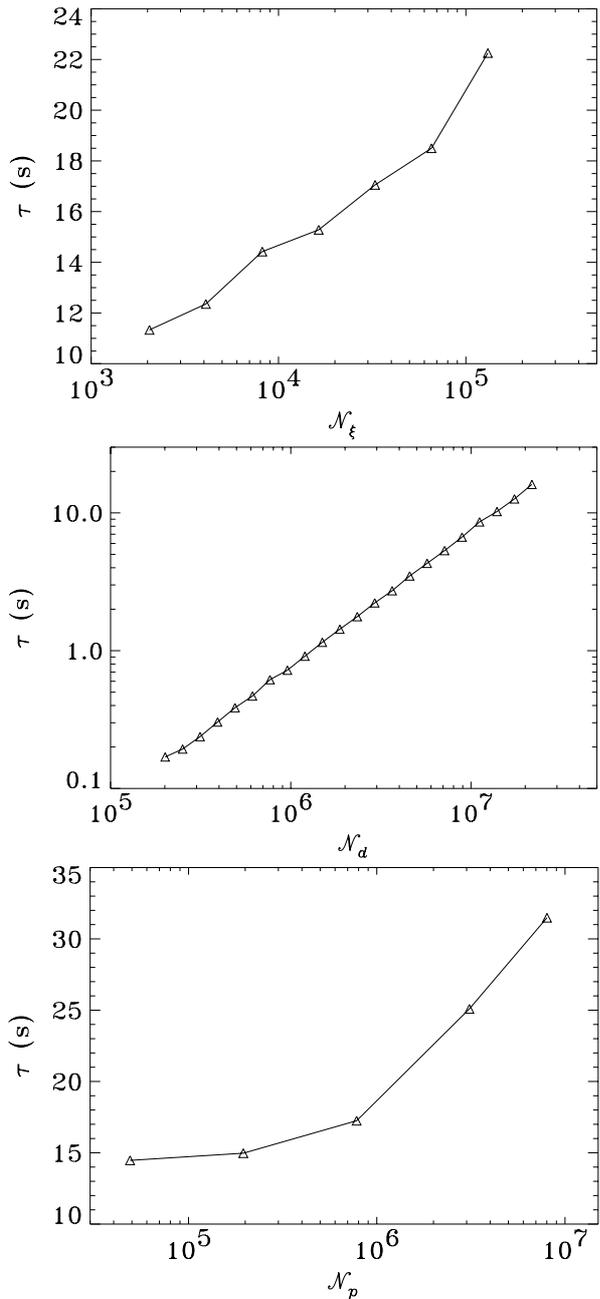


Fig. 1. The wall-clock time τ (s) on a single 500 MHz Pentium III CPU per CG iteration plotted vs. the noise filter length, \mathcal{N}_ξ , vs. the TOD length, \mathcal{N}_d and vs. the map pixel number, \mathcal{N}_p .

τ as a function of \mathcal{N}_d for a given map size ($\mathcal{N}_p = 786432$, as expected by choosing $\mathcal{N}_{side} = 256$ in the HEALPix pixelization (Górski et al. 1999)) and the PLANCK baseline scanning strategy. The scaling with \mathcal{N}_d is indeed linear: $\tau = ((\mathcal{N}_d/1400) + 50)$ ms.

Since we never build up the $\mathcal{N}_p \times \mathcal{N}_p$ inverse covariance matrix Σ , but rather perform the piecewise multiplication $P^T N^{-1}(P \mathbf{m})$ each time we need to do so, this operation is not expected to scale with the number of pixels. However, multiplying by the preconditioner, H^{-1} , does. When $\mathcal{N}_p \ll \mathcal{N}_d$ (a condition usually fulfilled for a CMB experiment)

the preconditioner should produce a negligible effect on the operation count. Conversely, we expect an increase in the operation count as, keeping \mathcal{N}_d fixed, \mathcal{N}_p increases. This is confirmed in Fig. 1 (lower panel). Note that, for the last two points in the plot, $\mathcal{N}_p \sim \mathcal{N}_d$. Another factor that should in principle contribute to the scaling of τ with \mathcal{N}_p is the performance overhead resulting from handling the arrays containing the maps. Note, however, that τ increases by only a factor ~ 2 when \mathcal{N}_p is boosted by a factor >200 .

All the scalings given in Fig. 1 have been obtained from a single processor job using a 500 MHz Pentium III CPU.

4.2. Scaling with \mathcal{N}_{PE}

In order to test our parallel software we use the following simulated data sets (relative to the LFI 30 GHz channel), for which \mathcal{N}_d changes by a couple of orders of magnitudes: $\mathcal{P}60$, the time stream obtained by averaging on circles the full PLANCK TOD ($\mathcal{N}_d = 20\,196\,000$); $\mathcal{P}3$, a fictitious time stream obtained assuming that each circle is observed only 20 times ($\mathcal{N}_d = 403\,920\,000$); \mathcal{P} , the full PLANCK TOD ($\mathcal{N}_d = 1\,211\,760\,000$). In the ideal case, the algorithm speed τ^{-1} should scale linearly with \mathcal{N}_{PE} and should be inversely proportional to \mathcal{N}_d . So, we can define an efficiency index

$$\mathcal{E} = \frac{1}{\tau} \frac{\mathcal{N}_d}{\mathcal{N}_{PE}} \quad (13)$$

that we measure on an O2K. For $\mathcal{P}60$, increasing \mathcal{N}_{PE} from 1 to 8 results in a loss of efficiency of $\sim 50\%$ due to inter-processor communications. If we keep $\mathcal{N}_{PE} = 8$ but change $\mathcal{P}60$ to $\mathcal{P}3$ and \mathcal{P} , \mathcal{E} increases. We find that \mathcal{E} is basically the same when comparing $\mathcal{P}60$ and $\mathcal{N}_{PE} = 1$ with \mathcal{P} and $\mathcal{N}_{PE} = 8$. Thus, for these cases, our code shows the ideal scaling given in Eq. (13).

4.3. Convergence criterion and precision

A natural criterion to stop the CG solver is to require that the fractional difference

$$\epsilon^{(k)} = \frac{\|P^T N^{-1} P \tilde{\mathbf{m}}^{(k)} - P^T N^{-1} \mathbf{d}\|}{\|P^T N^{-1} \mathbf{d}\|} \quad (14)$$

obtained at the k th iteration step is less than a supplied tolerance. The norm $\|\cdot\|$ in Eq. (14) represents the usual Euclidean norm⁸.

To test the CG convergence efficiency we plot in Fig. 2 $\epsilon^{(k)}$ for two datasets of different length: $\mathcal{N}_d = 20\,196\,000$ (corresponding to $\mathcal{P}60$) and $\mathcal{N}_d = 1\,211\,760\,000$ (corresponding to \mathcal{P}). Both datasets have been benchmarked

⁸ An alternative convergence criterion is the more stringent L_∞ norm, i.e. the absolute value of the largest element of a vector. We have tested that, for our applications, the CG solver returns almost identical results using either the Euclidean or L_∞ norm. However, in the latter case, convergence is significantly slower.

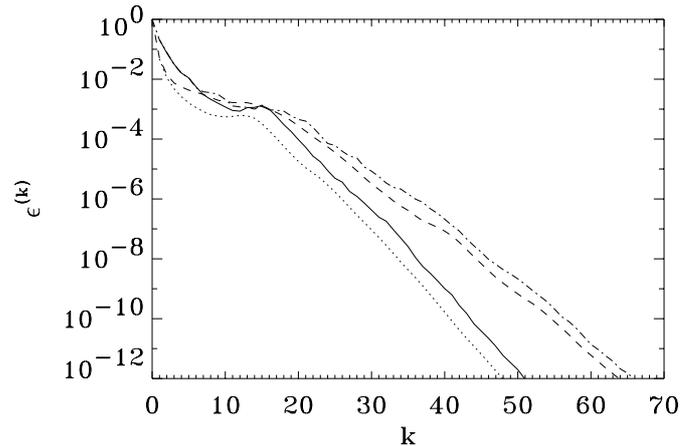


Fig. 2. Solver precision as a function of the CG iteration step. Solid and dotted curves refer to the parallel code running on the full \mathcal{P} TOD (no average on circles) for “signal-only” and “noise-only” TOD, respectively. Dot-dashed and dashed curves refer to the serial code running on $\mathcal{P}60$, again for “noise-only” and “signal-only”.

for the cases of “signal-only” and “noise-only” TOD. In the latter case, the noise properties are kept independent of the TOD length to facilitate comparisons. Note that, after a few steps, the precision $\epsilon^{(k)}$ decreases exponentially and drops to the 10^{-6} level in a few tens of iterations. Although always exponential, the rate of convergence of the CG solver depends on the TOD length. This is a consequence of having kept fixed \mathcal{N}_ξ and \mathcal{N}_p while increasing \mathcal{N}_d . In fact, the number of elements of Σ connected by the noise filter is lower for a longer TOD, a fact that makes this matrix more efficiently preconditioned by its diagonal part.

To quote a single number, the algorithm converges, to 10^{-6} precision, in about 90 min ($\mathcal{N}_{PE} = 8$) to produce a map out of the entire (*not* averaged over circles) 14 months time stream, expected from the PLANCK Surveyor @ 30 GHz.

Of course, solver precision and accuracy (that is, “distance” from the true solution) are not the same thing. In other words, being confident that the algorithm converges quickly is only half the story. We also want it to converge to the exact solution. This is a rather important issue because we would not like the algorithm to kill modes in the map. Discussion of this point is deferred until the next Section.

5. Numerical results

5.1. “Signal-only”

In the limiting case of a noiseless TOD (i.e. $\mathbf{d} = P\mathbf{m}$) Eq. (3) and Eq. (4) yield $\tilde{\mathbf{m}} = \mathbf{m}$. Thus, any good map-making code should return the “true”, pixelized sky when fed with a “signal-only” TOD. We tested our code with a pure CMB anisotropy pattern⁹ and the same CMB

⁹ We use the standard Cold Dark Model with $\Omega_{CDM} = 0.95$, $\Omega_B = 0.05$ and $h = 0.5$. No dipole was included.

pattern contaminated by the foregrounds most relevant at 30 GHz¹⁰. The following scheme was employed to prepare the TOD. First, a “signal map” is built, pixelized at $\sim FWHM/3$ using the HEALPix pixelization ($\mathcal{N}_{\text{side}} = 256$) and $FWHM$ -smeared. The latter is “read” by simulating a 30 GHz PLANCK/LFI time stream. This TOD is the input to the map-making code.

In order for the code to run a noise filter, N^{-1} , is needed. First note that for this noiseless case the filter should, in principle, be irrelevant: as stated above, the GLS solution trivially gives back the “true” map independently of N^{-1} . We must keep in mind, however, that we only consider \mathcal{N}_{ξ} elements of the filter. In Fourier space this translates into estimating the noise inverse spectral density precisely at \mathcal{N}_{ξ} frequencies, the lowest, f_{min} say, being \mathcal{N}_{ξ} times smaller than the sampling frequency. One obvious limitation on \mathcal{N}_{ξ} is that it must be fixed accordingly, so that the Fourier representation of N^{-1} well comprises the spin frequency. The amplitude of modes with $f < f_{\text{min}}$, if any, is not recovered unless the corresponding frequencies are considered in the analysis. However, there are two reasons why one should not worry about this effect: first, and more importantly, it is small and becomes utterly negligible for $\ell \gtrsim$ a few. Furthermore, as previously stated, this contribution is likely to get lost in any case since it falls in a region of the time Fourier spectrum strongly dominated by the noise and/or, more realistically, suppressed by hardware-induced decoherence (e.g. instrumental high passing). We show (for the typical case $\mathcal{N}_{\xi} = 4096$) the angular power spectra of the CMB only input and output maps together with their difference (lower panel of Fig. 3). It is clear that our map-making code recovers, as expected, the input power spectrum to an impressively high precision over the whole range of explored multipoles (although with a somewhat coarser accuracy at low ℓ 's). The same happens when considering the case of CMB plus foregrounds (see Fig. 4, upper panel): the level of the residuals between the input and output map is remarkably low, underlying the high accuracy of our code in recovering the input signal even in the presence of sharp features (see Fig. 4, lower panel).

As shown above, the code also recovers to very high precision even the broadest features in the map, i.e. the ones corresponding to low ℓ 's. We want to stress that the CG solver converges to the corresponding modes in a very reasonable number of iterations ($\lesssim 50$, see Fig. 2). Other map-making codes, implementing different solving algorithms, seem not to share this property (Prunet et al. 2000) requiring more sophisticated methods to speed up the convergence of low multipoles and/or a change of variable to solve for the featureless noise-only map (see Doré et al. 2001).

A final comment before concluding this Section. As mentioned at the end of Sect. 4.3, precision and accuracy

¹⁰ Specifically, synchrotron and free free emission were included as well as simulated emission from SZ clusters and milimetric bright sources.

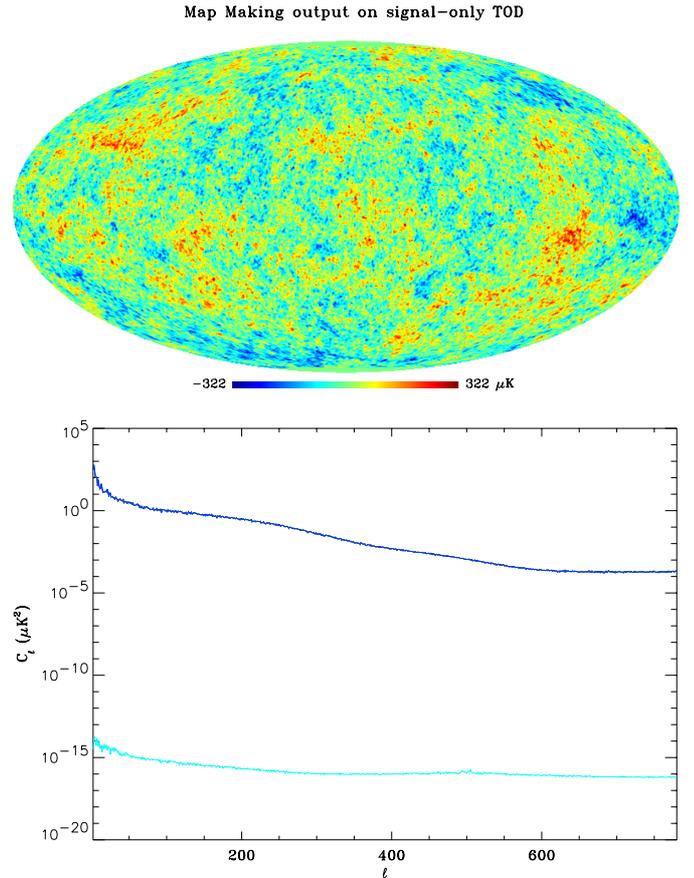


Fig. 3. Upper panel: map-making output on “signal-only” TOD (CMB); lower panel: angular power spectrum of the input map (dark blue line), together with the spectrum of the map given in upper panel. The two are virtually indistinguishable. The light blue line below shows the spectrum of the difference map.

are not the same thing. We want to be sure that the CG algorithm converges to the “true” solution. So, from a given “signal-only” TOD we generate different maps by changing the solver precision. For each of these maps we evaluate the maximum difference in pixel value between them and the input map. The results are shown in Fig. 5 where we plot this maximum difference, normalized to the L_{∞} norm of the input map, as a function of the required solver precision ϵ . There are two regimes separated by $\epsilon \sim 10^{-8}$: the map accuracy decreases when ϵ gets larger and saturates when ϵ gets smaller. So, $\epsilon \sim 10^{-8}$ is just about the right number to have a good compromise between speed of convergence and accuracy. In any case, any residual map error smaller than $\sim 1 \mu\text{K}$ is acceptable in most applications, given the PLANCK sensitivity goal.

5.2. “Noise-only”

In the opposite limiting case of a “noise-only” TOD (i.e. $\mathbf{m} \equiv 0$), the map-making solution (cf. Eqs. (3) and (4)) gives $\tilde{\mathbf{m}} = (\mathbf{P}^T \mathbf{N}^{-1} \mathbf{P})^{-1} \mathbf{P}^T \mathbf{N}^{-1} \mathbf{n}$. The efficiency of a map-making code can be tested, in this case, by examining the quality of the map obtained, a natural figure of merit

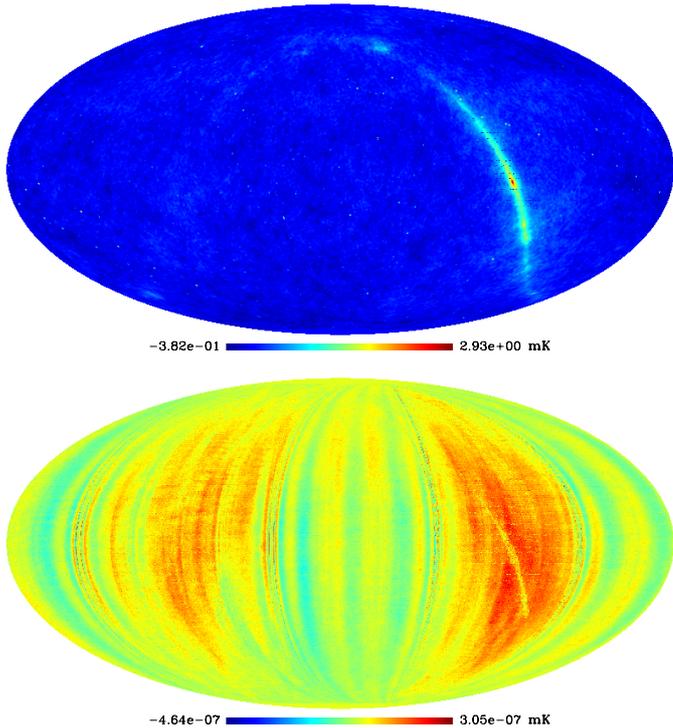


Fig. 4. Upper panel: map-making output on “signal-only” TOD (CMB and main foregrounds). Lower panel: Difference between the above and input maps.

being the variance of the reconstructed map or – better – its angular power spectrum.

We generate a noise time stream assuming the following form for the noise spectral density:

$$P(f) = A[1 + (|f|/f_k)^\alpha], \quad (15)$$

where f_k is the knee frequency. We choose $f_k = 0.1$ Hz (the PLANCK goal), $\alpha = -1$ and an amplitude A corresponding to the expected white noise level of the 30 GHz receivers. The minimum and maximum frequencies are set by the inverse of the mission life-time and the sampling frequency (33 Hz). We then generate a “noise-only” data stream that is the input to the map-making code.

Figure 6 shows the map obtained by just averaging different observations of the same pixel (upper panel), the ML solution (middle panel) and the angular power spectra of the two (lower panel). It is clear from the middle panel of this figure that the map-making algorithm strongly suppresses the stripes due to $1/f$ noise, very visible in the upper panel. This is confirmed by the angular power spectrum of the ML map (see lower panel of Fig. 6) which is basically flat (as expected in the case of white and isotropic noise) for $\ell \gtrsim 100$. The increasing power at $\ell \lesssim 100$ is due to two effects: residuals of spurious correlations in the noise and nonuniform sky coverage due to the PLANCK scanning strategy. In spite of that, the overall level of instrumental noise has been considerably lowered by the map-making algorithm.

The noise filter or, equivalently, the inverse of the noise spectral density, has been evaluated as explained in Sect. 3.1. In order to assess how good $\hat{\mathbf{n}}$ is as a noise

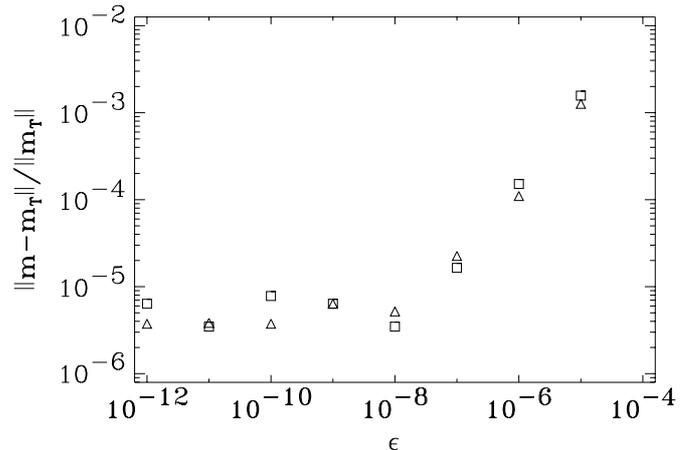


Fig. 5. Maximum deviation between input and output maps in the “signal-only” case, as a function of the solver precision. Here $\mathcal{N}_\xi = 2^{11}$ (triangles) and $\mathcal{N}_\xi = 2^{17}$ (squares).

estimator, we use the same “noise-only” TOD to produce two maps: the first, by using the “true” theoretical noise spectral density (cf. Eq. (15)); the second, by using the noise spectral density estimated with $\hat{\mathbf{n}}$. In both cases $\mathcal{N}_\xi = 4096$ as for the map in Fig. 6. In Fig. 7 we show, as a function of ℓ , the percentage difference between the spectra derived from the two maps. So, although not optimized to the specific case of the PLANCK Surveyor, the noise estimator $\hat{\mathbf{n}}$ reproduces to quite a good extent the in-flight noise properties. In any case, as also shown in Fig. 7, the uncertainties introduced in the estimate of the noise angular power spectrum are less than those due to cosmic variance.

As already discussed in the previous sections, the length of the noise filter is an important issue to carefully address. One should expect that in the “noise-only” case the dependence on \mathcal{N}_ξ could be stronger because of the $1/f$ tail in the noise power spectrum. To address this issue in detail and to show the sensitivity of our results to \mathcal{N}_ξ , the following scheme is employed: we generate the “true”, fully correlated noise time stream, accordingly to Eq. (15). Then, we do map-making using a noise filter with $\mathcal{N}_\xi = \mathcal{N}_d/2$, and evaluate the rms of this map, σ^{11} . Then, we generate a set of maps out of the same fully correlated noise time stream, but imposing a noise filter length, \mathcal{N}_ξ , shorter and shorter w.r.t. $\mathcal{N}_d/2$. In Fig. 8 we plot the fractional variation of the map rms w.r.t. σ as a function of \mathcal{N}_ξ . It is clear from the figure that, for the PLANCK Surveyor, we can lower \mathcal{N}_ξ by a factor of \sim thousand w.r.t. $\mathcal{N}_d/2$ and change the final rms of the map by much less than one percent. This shows that increasing \mathcal{N}_ξ above a given threshold does not induce any significant difference. This result holds even if we consider knee frequencies significantly higher than $f_k = 0.1$ Hz: in Fig. 8 we also plot the

¹¹ Note that we simulate the whole “noise-only”, \mathcal{N}_d long, time stream by using Eq. (15) and standard FFT techniques. This time stream is obviously periodic. Thus, N^{-1} is strictly circulant even for $\mathcal{N}_d = \mathcal{N}_\xi/2$. The same would not be true for real data.

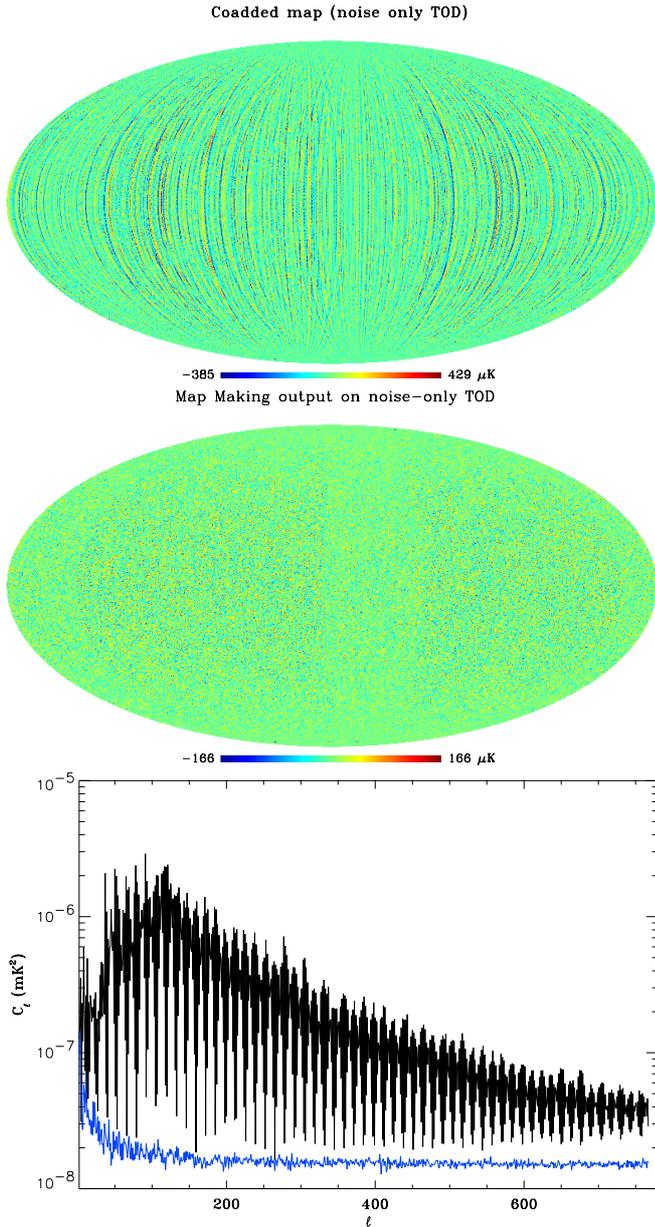


Fig. 6. Upper panel: coadded map made out of a “noise-only” TOD. Note the stripes. Middle panel: map-making solution on “noise-only” TOD. Lower panel: angular power spectrum of the maps given above (black and blue curves, respectively).

effect of having $f_k = 1$ Hz and $f_k = 10$ Hz, surely two very pessimistic assumptions in the PLANCK context.

In principle, the noise properties of the PLANCK detectors, as described by Eq. (15), suggest the existence of spurious correlations in the TOD over the whole mission life-time. In practice, the error introduced by neglecting them is negligible, at least when the rms is used as a figure of merit. To stress this point even further, we evaluate the angular power spectrum from two maps generated with the same input (i.e. the same TOD), but using noise filters of different lengths. Again, we plot the percentual difference between the spectrum obtained with $\mathcal{N}_\xi = \mathcal{N}_d/2$ and the spectra obtained with noise filters shorter by a

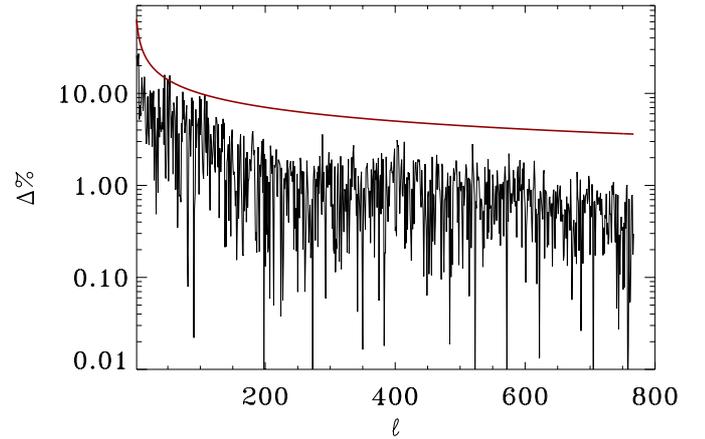


Fig. 7. Percentual difference, as a function of ℓ , of the power spectra of the maps obtained using the true noise properties and the noise estimator \tilde{n} discussed in the text.

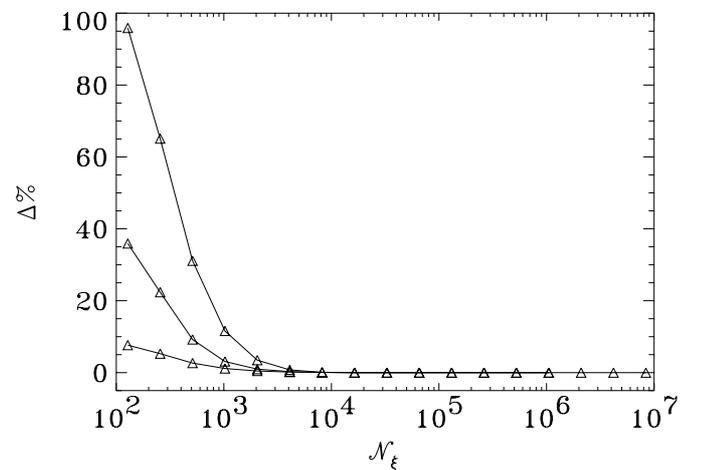


Fig. 8. Shown is the percentual difference between the rms of a “noise-only” map obtained with a filter of length \mathcal{N}_ξ and σ , the rms obtained for $\mathcal{N}_\xi = \mathcal{N}_d/2$. The three lines refer, from top to bottom, to $f_k = 10$ Hz, $f_k = 1$ Hz and $f_k = 0.1$ Hz (the PLANCK goal).

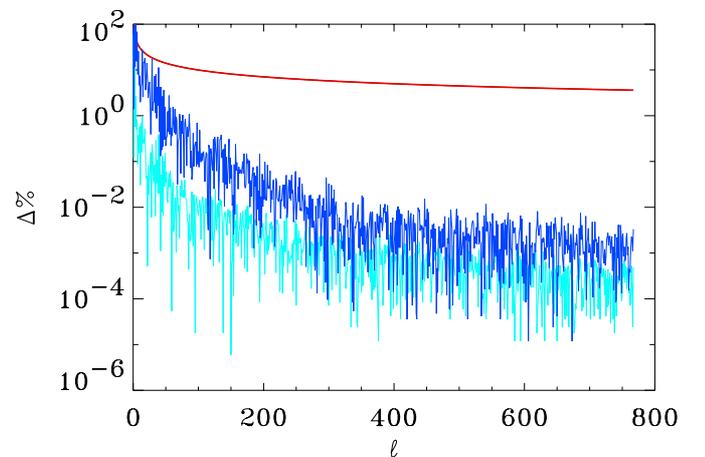


Fig. 9. Shown is the percentual difference between noise-only C_ℓ 's obtained for $\mathcal{N}_\xi = 2^{17}$ (light blue curve) and $\mathcal{N}_\xi = 2^{13}$ (dark blue curve) w.r.t. the power spectrum obtained with the longest possible filter ($\mathcal{N}_\xi = \mathcal{N}_d/2$). The red line is the uncertainty due to cosmic variance.

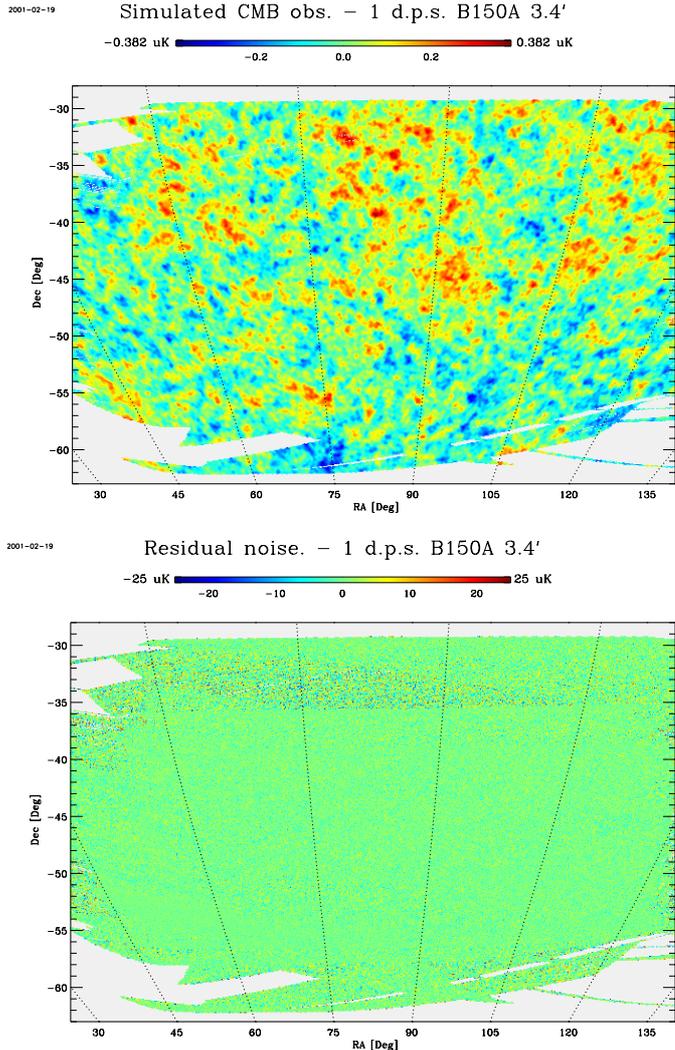


Fig. 10. Upper panel: map-making output on the simulated BOOMERanG TOD (signal plus noise). Lower panel: residual noise map obtained subtracting the input map from the one shown above. The maps have $\mathcal{N}_p \simeq 6 \times 10^5$.

factor of 64 and 1024, respectively. Figure 9 shows that this difference is below $\sim 1\%$ for $\ell \gtrsim 100$, well under cosmic variance in the region of interest for recovering the cosmological parameters.

To conclude this section let us note that the algorithm discussed here is intrinsically linear (cf. Eq. (1)) since the map estimator can be written as the “true” map plus a noise term:

$$\tilde{\mathbf{m}} = \mathbf{m} + (\mathbf{P}^T \mathbf{N}^{-1} \mathbf{P})^{-1} \mathbf{P}^T \mathbf{N}^{-1} \mathbf{n}.$$

Our code shares this property. In fact, if we produce a map from a “signal-only” TOD and sum it to the map produced from a “noise-only” TOD the result is virtually indistinguishable (to machine precision) from the map obtained by summing the two (“signal-only” and “noise-only”) time streams.

5.3. Balloon borne experiments

As already mentioned, our code does not take any advantage from the specific PLANCK scanning strategy. It is then straightforwardly applicable to other one-horned experiments, such as BOOMERanG (de Bernardis et al. 2000). Here we do not want to enter in any data analysis issue. Rather, we want to show the flexibility and the efficiency of our code to process data coming from an experiment completely different from PLANCK.

As far as the signal is concerned, we simulate a theoretical input map⁹ pixelized at 1/3 of the BOOMERanG $FWHM$ ($\sim 10'$ @ 150 GHz) and properly smeared. Knowing its scanning strategy, we extract a typical BOOMERanG “signal-only” time stream. As far as the noise is concerned, we generate a “noise-only” time stream with the noise properties described in Eq. (15) and choosing $A = 150 \mu\text{K}\sqrt{s}$ and $f_k = 0.07$ Hz. In what follows we consider a noise filter of length $\mathcal{N}_\xi = 2^{17}$.

On the very same line of the previous subsections, we compare the theoretical input map with $\tilde{\mathbf{m}}$, which is reconstructed at a level of accuracy of 10^{-5} when the CG solver precision is chosen to be 10^{-6} . The length of the noise filter does not affect, even in this case, the final results.

In the above tests we only consider the 1 degree per second (d.p.s.) section of the BOOMERanG scan. However, in order to facilitate comparisons with PLANCK, we truncated the BOOMERanG simulated TOD at the length of $\mathcal{P}60$. Not surprisingly, our code runs the BOOMERanG map-making in ~ 15 min on a 500 MHz Pentium III workstation, while the whole 1 d.p.s. BOOMERanG scan is processed on the same machine in about 20 min.

Given these time scales, a parallel machine does not seem necessary for the analysis of a single channel. However, a parallel environment becomes quite appealing if one performs a combined analysis of more than one receiver. Also, having a parallel implementation may be important to perform extensive Monte Carlo simulations of the experiment. Specific applications of this code to BOOMERanG will be discussed elsewhere.

6. Summary

The purpose of this work was to present a parallel implementation of a map-making algorithm for CMB experiments. In particular, we have shown for the first time Maximum Likelihood, minimum variance maps obtained by processing the entire data stream expected from the PLANCK Surveyor, i.e. a TOD covering the full mission life span (14 months). Here we restrict ourselves to the simple case of the PLANCK/LFI 30 GHz channel. However, the extension of our implemented software to other, higher frequency, channels is straightforward. At the moment our implementation is limited to a symmetric antenna beam profile.

The code was shown to scale linearly with the number of time samples \mathcal{N}_d , logarithmically with the noise filter

length and very slowly with the number of pixels \mathcal{N}_p , the latter scaling being mostly influenced by the choice of the preconditioner. Furthermore, on a multiprocessor environment the code scales nearly optimally with the number of processor elements.

The code is extremely accurate and fast. We have shown that it recovers the true, input map to a remarkably high accuracy, over the whole range of considered multipoles. Moreover, we run the entire 30 GHz simulated TOD in ~ 90 min on an 8 processor job on a O2K machine, and the BOOMERanG case in only 20 min on a 500 MHz Pentium III workstation. On the latter machine, the PLANCK case runs in about 15 min if we perform the average on circles.

A key problem for a successful map-making pipeline is to obtain a reliable estimate of the noise properties directly from flight data. It is straightforward to further iterate the GLS solution implemented here to get more accurate estimates of the underlying noise. Here we have explicitly chosen not to do so. Rather, we stop to the *zereth* order solution, \tilde{n} , showing that this is more than enough for the purposes discussed here.

The possibility of running map-making algorithms on reasonably short time scales opens up the possibility of evaluating the CMB angular power spectrum via Monte Carlo simulations. We will address this point in a forthcoming paper, together with the extension to non-symmetric antenna beam profile.

Acknowledgements. We are indebted with C. Burigana and D. Maino for helping us with the generation of the TOD sets. We thank A. Balbi, P. Cabella, D. Marinucci and C. di Fiore for useful suggestions and comments. We also thank P. de Bernardis and the BOOMERanG collaboration for having provided us with the BOOMERanG scan and instrumental performances. PN acknowledges fruitful discussions with J. Staren and E. L. Wright. We acknowledge use of the HEALPix package and of the FFTW library. The supercomputing resources used for this work have been provided by Cineca.

References

Axelsson, O., & Barker, V. A. 1984, Finite elements solutions of boundary value problems – Theory and computation (Academy Press, Orlando)

- Bond, J. R., Jaffe, A. H., & Knox, L. 1998, Phys. Rev. D, 57, 2117
- Borrill, J. 1999, Proceedings of the Conference, 3 K Cosmology, AIP Conf. Proc., 476, 277
- de Bernardis, et al. 2000, Nature, 404, 955
- de Oliveira-Costa, et al. 1998, ApJ, 509, L77
- Delabrouille 1998, A&AS, 127, 255
- Doré, O., Teyssier, R., Bouchet, F. R., Vibert, D., & Prunet, S. 2001, submitted to A&A [[astro-ph/0101112](#)]
- Ferreira, P., & Jaffe, A. H. 2000, MNRAS, 312, 89
- Frigo, M., & Johnson, S. G. 1998 ICASSP Conf., 3, 1381; also see <http://www.fftw.org/>
- Górski, K. M., Hivon, E., & Wandelt, B., to appear in Proceedings of the MPA/ESO Conference on Large-Scale Structure: From Recombination to Garching, ed. Banday, et al. [[astro-ph/9812350](#)]
- Hanany, S., et al. 2000, ApJ, 545, L5
- Janssen, M. et al. 1996, internal report PSI-96-01 [[astro-ph/9602009](#)]
- Lineweaver, C. H., et al. 1994, ApJ, 436, 452
- Lupton, R. 1993, Statistics in theory and practice (Princeton University Press, Princeton)
- Maino, D., et al. 1999, A&A, 140, 383
- Natoli, P., et al. 2001, in preparation
- Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. 1992, Numerical Recipes in FORTRAN, The Art of Scientific Computing, 2nd edition (Cambridge University Press, Cambridge)
- Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. 1996, Numerical Recipes in Fortran 90, The Art of Scientific Computing (Cambridge University Press, Cambridge)
- Prunet, S., Netterfield, C. B., Hivon, E., & Crill, B. P., to be published in the proceedings of the XXXV Recontres de Moriond, Energy Densities in the Universe (Ed. Frontières) [[astro-ph/0006052](#)]
- Szapudi, I., Prunet, S., Pogosyan, D., Szalay, A. S., & Bond, R. 2000, submitted to ApJ [[astro-ph/0010256](#)]
- Tegmark, M. 1997, ApJ, 480, L87
- Wandelt, B., Hivon, E., & Górski, K. M., submitted to Phys. Rev. D [[astro-ph/0008111](#)]
- Wright, E. L. 1996, paper presented at the IAS CMB Data Analysis Workshop [[astro-ph/9612006](#)]
- Wright, E. L., Hinshaw, G., & Bennett, C. L. 1996, ApJ, 458, L53