Mem. S.A.It. Suppl. Vol. 13, 84

Memorie della
Supplementi

The GLS approach to CMB map making

P. Natoli[1,2], M. Botti[3], G. de Gasperis[1], G. De Troia[1], and F. Massaioli[3]

[1] Dipartimento di Fisica, Università di Roma "Tor Vergata", Via della Ricerca Scientifica, 1 00133 Roma (Italy), e-mail: paolo.natoli@roma2.infn.it
[2] I.N.F.N., sezione Tor Vergata Via della Ricerca Scientifica, 1 00133 Roma (Italy)
[3] CASPUR, Via dei Tizii, 6, 00185 Roma (Italy)

Abstract. The field of Cosmic Microwave Background (CMB) data analysis is a smashing source of computational problems. Here we discuss the ROMA map making code. Originally designed to generate a map of the CMB anisotropy and polarization fields using optimal statistical methods, ROMA can be applied also to other microwave oriented surveys. The computational details of the subject will be set forth.

1. Introduction

The Cosmic Microwave Background anisotropy is a unique window on the early Universe. It is widely regarded as one of the most important sources of cosmological information. Several experiments, either ground based, balloon or space borne have sought after this weak signal (the CMB is anisotropic only for about one part out of 10^5). Such an experimental effort will be crowned by Planck[1], an ESA satellite due to launched in late 2008 that will measure the CMB anisotropy to a level only limited by residual astrophysical foreground emission. Together with experimental know how, CMB datasets also have grown in size: modern experiment can host up to several thousand detectors, that are typically read out in the range 10 to 1000 Hz. As a consequence, the sheer size of the dataset is an issue. At the same time, the weakness of the target signal calls for the implementation of optimal data reduction techniques.

Map Making is a key part of the data analysis pipeline of an experiment aiming at extracting scientific information from the CMB. It allows for a radical compression of the dataset, reducing its size by several orders of magnitudes while preserving (when properly carried out) all of the relevant cosmological information. In addition, the frequencies probed by CMB experiments are broadly relevant for Galactical and extra Galactical studies. For these, it is often critical to rely on accurate spatial characterization of the target source. Last, but by no means least, map making is often beneficial to the dataset, since many timeline specific systematic effects are reduced in magnitude when projected from time to pixel space. When this is not possible, careful analysis of a CMB map might flag instrumental effects in the data that are not evident at the timeline level.

Several map making algorithms have been proposed for CMB experiments. The most efficient from the computational point of view

Send offprint requests to: P. Natoli

[1] www.rssd.esa.int/index.php?project=planck

take advantage from peculiarities of the scanning strategy of the experiment to deliver approximate but quick solutions. This is the case of so called "destriping" algorithms widely used in Planck (see (Maino 1999) and (Revenu *et al.* 2000)). While some of these algorithms achieve excellent performances and accuracy their use is usually restricted to particular scanning strategies and noise characteristics. Hereafter we describe a particular implementation of a map making algorithm (the so called GLS algorithm) that can be applied to virtually every CMB experiment. A Finnish group (Keihanen *et al.* 2005) has recently proposed an hybrid algorithm (dubbed madam) that can perform as a destriper when this is desirable or appropriate, but capable to reach the accuracy of a GLS when a higher computational cost can be afforded. We do not discuss this algorithm here, the interested reader is refereed to the reference above and to (Ashdown *et al.* 2007, 2008) for a thorough comparison of several mapmaking algorithms of interest for Planck.

1.1. The data model

State of the art CMB detectors do not possess the sensitivity to image directly the target field. Instead, the beams are scanned across by means of a dedicated movement of the telescope. The observations hence form a time series (or timeline) and signal to noise ratio is gained by observing repeatedly the target region. The classical model of a data set in time space is:

$$d_t = s_t + n_t \tag{1}$$

where d_t, the data vector as a function of time, is modelled as a linear combination of the sky signal s_t and detectors noise n_t. The two can be assumed to be statistically independent, i.e. $\langle s_t n_t \rangle = 0$, where $\langle \cdot \rangle$ denotes the ensemble expectation value. On the other hand, noise from a state-of-the art CMB detector almost always exhibits a time correlation:

$$\langle n_t n_t' \rangle = \xi(|t - t'|), \tag{2}$$

where ξ is some correlation function, whose explicit dependence on the time arguments holds in the case of stationary noise. As the telescope perform its scan, the sky signal is projected onto the timeline. We can think of this operation as if carried out by means of a linear operator A:

$$d_t = \sum_p A_{tp} m_p + n_t \tag{3}$$

where m_p is the (unknown) underlying sky map, labeled by its pixel element p. Note that so doing we are implicitly subdiving the sky into finite resolution elements: the pointing matrix A then couples the time and pixel domain. The sky signal is convolved with an optical transfer function: in general, the information about beam smearing could be included in A. We prefer however, to assume here that A is pointwise (i.e. having a single nonzero entry per row, occurring when a pixel falls into the line of sight) and hide the beam map (i.e. solve for a pixelized, beam smeared image of the sky). This is only possible if the beam can be considered axially symmetric with respect to boresight because in the case the ordering of the "pointing" and "beam smearing" operations is irrelevant: convolving with a symmetric beam produces the same result regardless of the scan direction (see however (Armitage & Wandelt 2004) for an algorithm that deals with the effects of an asymmetric beam).

The CMB is known to be linearly polarized due to the dependence of the Thomson cross section on photon polarization and the anisotropic character of the radiation field (Kaiser, 1983). It has long been realized that measurements of this polarization pattern can provide useful cosmological information. Hence, there is a neeed to provide polarization maps as well. In the formalism of the Stokes parameters (Chandrasekhar 1960) one needs also to estimate Q and U maps (the V parameter is not relevant here since the CMB does not exhibit circular polarization). Thus, the above data model can be modified as:

$$\mathcal{D}_t = \frac{1}{2} A_{tp} \left(I_p + Q_p \cos 2\phi_t + U_p \sin 2\phi_t \right) + n_t \tag{4}$$

where I, Q and U are the Stokes parameters for total intensity and linear polarisation, and the angle ϕ_t identifies the polarimeter orientation with respect to the chosen celestial frame. Note that we do not include the contribution of circular polarisation, associated to the Stokes parameter V. As a consequence, a V signal is seen by our polarimeter as a contribution to I. This fact is not a problem in CMB science because circular polarisation is not present. All three relevant Stokes parameters can be extracted by either combining the output of many detectors with different mutual orientation, or by allowing enough focal plane rotation.

We can now safely insert the sines and cosines in the A matrix, and then consider a set of k polarimeters, recasting Eq. (4) into a more compact formalism:

$$\mathcal{D}_t = \mathbf{A}_{tp}\mathbf{S}_p + \mathbf{n} \tag{5}$$

where we assume that the timelines of each detector are combined:

$$\mathcal{D}_t \equiv \begin{pmatrix} \mathcal{D}_t{}^1 \\ \vdots \\ \mathcal{D}_t{}^k \end{pmatrix},$$

and introduced a generalised pointing matrix that takes the form:

$$\mathbf{A}_{tp} \equiv \frac{1}{2}\begin{pmatrix} A^1_{tp} & \cos 2\phi_t A^1_{tp} & \sin 2\phi_t A^1_{tp} \\ \vdots & \vdots & \vdots \\ A^k_{tp} & \cos 2\phi_t A^k_{tp} & \sin 2\phi_t A^k_{tp} \end{pmatrix}.$$

We can also express the sky signal as:

$$\mathbf{S}_p \equiv \begin{pmatrix} I_p \\ Q_p \\ U_p \end{pmatrix},$$

and the noise timeline as:

$$\mathbf{n}_t \equiv \begin{pmatrix} n^1_t \\ \vdots \\ n^k_t \end{pmatrix}.$$

The job of map making is to estimate the three (I, Q, U) maps. Since eq. (5) describes a linear problem, whose unknown parameters (the map pixel values) can be constrained using a the generalized least squared (GLS) solution (e.g. Lupton, (1993)):

$$\widetilde{\mathbf{S}}_\mathbf{p} = \left(\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^t\mathbf{N}^{-1}\mathcal{D}, \tag{6}$$

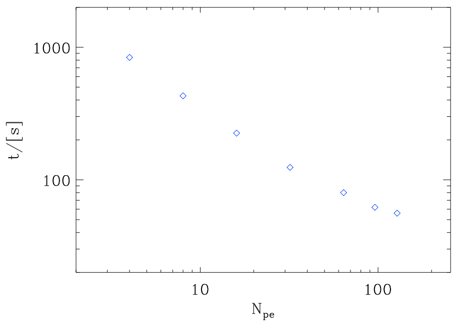

Fig. 1. Scaling of ROMA with the number of processor elements on an IBM Power 5 machines. The code scales "optimally" until $N_{pe} \sim 80$. The dataset was composed of about 100 million samples.

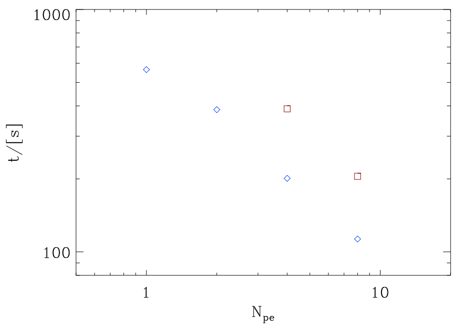

Fig. 2. As in figure 1 but for a 2.4 GHz Opteron cluster. The two different series of points refer to different dataset lengths.

where:

$$\mathbf{N} \equiv \langle \mathbf{n}_t \mathbf{n}_{t'} \rangle = \begin{pmatrix} \left\langle n^1_t n^1_{t'} \right\rangle & \cdots & \left\langle n^1_t n^k_{t'} \right\rangle \\ \vdots & \ddots & \vdots \\ \left\langle n^k_t n^1_{t'} \right\rangle & \cdots & \left\langle n^k_t n^k_{t'} \right\rangle \end{pmatrix}. \tag{7}$$

The noise matrix $\mathbf{N}$ reduces to block diagonal if the noise in different detectors is uncorrelated (that is, if there are no cross talks):

$$\left\langle n^i_t n^j_{t'} \right\rangle = \left\langle n^j_t n^i_{t'} \right\rangle = 0 \quad (i \neq j)\,. \tag{8}$$

1.2. Implementation

Eq. (6) is deceivingly simple. Its straightforward implementation for a modern CMB experiment would require the inversion of a matrix of order the number of pixels in the map, a number that can easily run in the million range. Such a task lies beyond the reach of present day supercomputers. Luckily, the problem can be recast in such a way that only matrix to vector products are required. Such a scheme was proposed by Natoli *et al.* (2001), has been demonstrated to work for intensity map-making even for a high resolution full sky survey such as PLANCK. The approach has been extended to polarization map making in (de Gasperis *et al.* 2000).

The idea (Wright (1996)) is to decompose the product $\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}$ ("unroll, convolve and rebin"), so that no large matrix needs to be stored, and make use of a preconditioned conjugate gradient (PCG) solver. The key assumptions are: (1) assume that the beam is axisymmetric, so to keep the structure of the pointing matrix $\mathbf{A}$ as simple as possible and (2) assume that the noise is (piecewise, at least) stationary and that its correlation function decays after a time lag much shorter than the length of the timeline piece being processed. Under the last assumption, the $\mathbf{N}$ matrix is approximately circulant, and as such it is diagonalized by a Fourier transform. This approach achieves linear scaling (with the number of time samples) per PCG iteration; if the system is preconditioned by the pixel hits counter (the number of times each pixel is seen), an accurate solution can be obtained in a few tens of iterations; see Natoli *et al.* (2001) for a discussion of algorithmic details.

To run, ROMA requires knowledge of the detector's noise properties, encoded in the noise covariance matrix $\mathbf{N}$. They must be measured directly from flight data which, however, are a combination of both signal and noise. This is accomplished by means of an iterative approach: a raw estimate of the signal is provided by performing a naive mapmaking, which is used to provide an estimate of the noise. The properties of the latter are then estimated by standard spectral techniques. The process is then repeated a few (typically 5) times to reach convergence.

The map making algorithm described above has been implemented into a code called ROMA. In the next section we will describe its performances.

1.3. Performances

ROMA is written in standard compliant fortran 95 and relies on the MPI libraries for inter processor communications. Since 2005, the original prototypical code has been completely rebuilt with the aid of personnel from the Caspur supercomputing center, with a significant boost in performance. It turns out that the lion share of CPU time is claimed by the discrete Fourier transform algorithm: ROMA uses the publicly available FFTW library (Frigo & Johnson 1998) to perform transforms. One of the nice characteristics of this library is that it can autonomously tailor itself to a given architecture. This greatly enhances cross platform efficiency. On an IBM power 5 machine, we have compared the performances of ROMA based on FFTW and on the vendor's optimized FFT library. The result was that very little (if anything) is to be gained by using the latter, while using FFTW enhances portability. Other than on IBM power 5 machines, ROMA has been tested on infiniband based Linux clusters. The scaling of the code with the number of processor elements is described in figure 1 and 2 for the two different architectures under consideration. Note how the scaling of ROMA is nearly optimal up to ~ 80 processors. Above this number, inter processor communications start to call their toll. This behaviour is due to the way ROMA is parallelized: each processor element only holds a portion of the timeline $\mathbf{A}$ and performs only its part of the product $\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}$, building a local submap. Only when the conjugate gradient solver is called, all the local maps are merged (`MPI_all_reduce`) to produce the full map. This means that most of the computation is performed without the necessity for inter processor communication and, when communications finally happen, the entire submap is transferred. ROMA runs very efficiently on a large number of processors, the highest number being tested being 1024. ROMA is expected to scale linearly with the number of samples in the timeline. This is confirmed by explicit tests

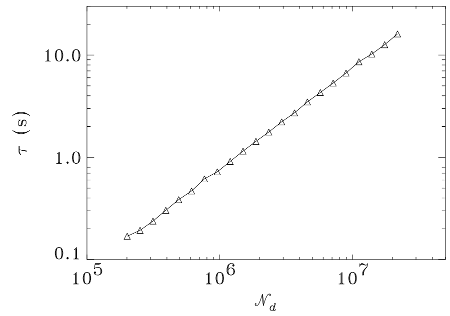

Fig. 3. Time per ROMA conjugate gradient iteration versus length of the timeline. Note linearity (reprinted from (Natoli *et al.* 2001)).

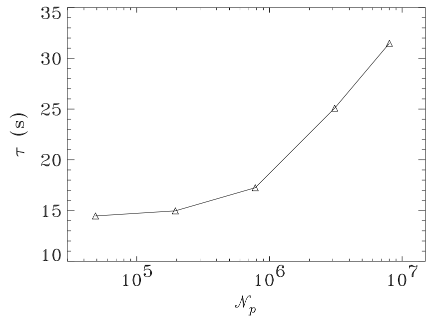

Fig. 4. Time per ROMA conjugate gradient iteration versus number of pixels in the map. Note how the time only doubles for a difference of two orders of magnitude in the number of pixels (reprinted from (Natoli *et al.* 2001)).

(figure 3). On the other hand, the scaling of the code with the number of pixels of the map is expected to be weak, though not totally insensitive because handling larger maps does have an impact on performances. In figure 4 we plot the measured pixel scaling. Note how it only doubles for a change of two orders of magnitude in pixel count.

1.4. Conclusions

We have descried the ROMA code for intensity of polarization map making. ROMA has been developed for data analysis of CMB experiment, but the approach is totally general and can be used for map making of any timeline exhibiting correlated noise. The code scales quasi optimally with the number of processor elements and can hence be employed for reducing a large dataset, such as the ones expected by forthcoming microwave oriented surveys.

Acknowledgements. We thank Caspur (http://www.caspur.it) for having provided the computational resources for part of the results presented here. We acknowledge the use of Healpix[2] (Górski *et al.* 2004).

References

Ashdown, M.A.J., *et al.*, 2007 A&A, 471 361
Ashdown, M.A.J., *et al.*, 2008 submitted to A&A(arXiv:0806.3167)
Armitage, C., & Wandelt, B. D. 2004, Phys. Rev. D, 70, 123007
Chandrasekhar S. 1960, "*Radiative transfer*", Dover, New York
de Gasperis G., *et al.*, 2005, A&A, 436, 1159
Frigo, M. & Johnson, S.G., 1998 ICASSP Conference, 3, 1381; also see http://www.fftw.org/
Górski, K.M., Hivon, E., Banday, A.J., Wandelt, B.D., Hansen, F.K., Reinecke, M., Bartelmann, M. 2004, Apj, 622, 729 [astro-ph/0409513]
Kaiser N., 1983, MNRAS, 202, 1169
Keihanen, *et al.*, 2005, MNRAS360, 390
Lupton, R., 1993 *Statistics in theory and practice*, Priceton University Press, Princeton
Maino, D, *et al.*, 1999 A&A140, 383
Montroy, T. *et al.*, 2003, New Astronomy Review, 47, 1057
Natoli P., de Gasperis, G., Gheller, C. & Vittorio, N., 2001, A&A, 372, 346
Natoli, P., Marinucci, D., Cabella, P., de Gasperis, G., & Vittorio, N. 2002, A&A, 383, 1100
Revenu B., Kim A., Ansari R., Couchot F., Delabrouille J., Kaplan J., 2000, A&AS, 142, 499
Wright, E.L., 1996, paper presented at the IAS CMB Data Analysis Workshop [astro-ph/9612006]

---
[2] http://healpix.jpl.nasa.gov