A&A 417, 1189–1194 (2004)
DOI: 10.1051/0004-6361:20034066

Astronomy & Astrophysics

Noise power spectrum estimation and fast map making for CMB experiments

A. Amblard[1,4] and J.-Ch. Hamilton[1,2,3]

[1] Physique Corpusculaire et Cosmologie, Collège de France, 11 pl. Marcelin Berthelot, 75231 Paris Cedex 05, France
[2] Laboratoire de Physique Nucléaire et de Hautes Energies, 4, Place Jussieu, Tour 33, 75252 Paris Cedex 05, France
[3] Laboratoire de Physique Subatomique et de Cosmologie, 53 Avenue des Martyrs, 38026 Grenoble Cedex, France
[4] University of California, Department of Astronomy, 601 Campbell Hall, Berkeley, CA 94720-3411, USA
e-mail: amblard@in2p3.fr

Received 10 July 2003 / Accepted 27 December 2003

Abstract. We present a method designed to estimate the noise power spectrum in the time domain for CMB experiments. The noise power spectrum is extracted from the time ordered data avoiding the contamination coming from sky signal and accounting the pixellisation of the signal and the projection of the noise when making intermediate sky projections. This method is simple to implement and relies on Monte-Carlo simulations, it runs on a simple desk computer. We also propose a trick for filtering data before making coadded maps in order to avoid ringing due to the presence of signal in the timelines. These algorithms were succesfully tested on Archeops data.

Key words. cosmic microwave background – cosmology: observations – large-scale structure of the Universe – methods: numerical

1. Introduction

Measuring the Cosmic Microwave Background anisotropies angular power spectrum has proved to be a powerful cosmological tool giving direct information on both the cosmological parameters and the primordial Universe through the origin of the initial perturbations.

The usual method used nowadays for measuring the CMB temperature fluctuations is to scan the sky with a detector (with a bolometer for instance) that measures the temperature variations in a given beam. This provides time data streams along with pointing directions that are subsequently compressed into a sky map. In the general case, the noise in the time streams is not white (due to atmospheric contamination, spurious noises, electronic and detector temperature fluctuations). The sky is also in general not regularly scanned so that at the end, the noise in the maps is neither homogeneous nor white.

This causes difficulties when trying to extract the signal power spectrum from the maps as it has to be separated from the noise. Fully optimal methods finding the maximum likelihood solution for both the map and the CMB power spectrum have been proposed such as MADCAP (Borrill 1999) but they require the inversion of large covariance matrices on parallel supercomputers and are hard to implement for the present generation of experiments that cover a large portion of the sky and have a large number of pixels (WMAP, Archeops and Planck in the near future). Alternative methods that replace the full inversion by a Monte-Carlo simulation of the noise angular power spectrum have been proposed (Hivon et al. 2002; Szapudi et al. 2001). The results obtained with such methods are satisfactory and they consume little time compared to the maximum likelihood methods. The analysis process is done in the following way in these Monte-Carlo methods:

1. **Noise power spectrum estimation in the data:** it is very important not to be contaminated by signal at this stage. This is particularly difficult when dealing with large portions of the sky where Galactic dust clouds can be found at high frequency and synchrotron and free-free emission at lower ones. We propose in Sect. 2 a method that provides an unbiased estimate of the noise Fourier power spectrum by correcting an initial guess via Monte-Carlo simulations.
2. **Fast map making:** can be a simple coaddition of filtered timelines or a optimal iterative mapmaking that converges to the maximum likelihood solution for instance with a conjugate gradient method. This map making has to be very fast as it is repeated for each Monte-Carlo realisation. We propose in Sect. 3 a method for making simply coadded

Send offprint requests to: J.-Ch. Hamilton,
e-mail: hamilton@in2p3.fr

maps of filtered data avoiding the ringing due to the effect of the filtering on bright sources such as the Galactic plane.

3. C_ℓ **estimation on the map:** this is done very easily with the Healpix[1] package (Gorski et al. 1999). The angular power spectrum obtained here is called *pseudo-*C_ℓ as it differs from the true signal C_ℓ due to various effects: presence of noise, beam and time domain filtering effects on the signal, incomplete sky coverage that degrades the resolution in harmonic space. All these effects can be accounted for and corrected with the MASTER (Hivon et al. 2002) or SpICE (Szapudi et al. 2001) algorithms.

In this article we propose methods for solving points 1 (in Sect. 2) and 2 (in Sect. 3), examples will refer to Archeops-like simulations. These methods have been sucessfully used for the Archeops data analysis (Benoît et al. 2003a,b). We used the MASTER algorithm for going from *pseudo-*C_ℓ to true C_ℓ and for estimating the error bars.

2. Noise spectrum estimation

We present in this section the method we developed for determining the noise Fourier power spectrum of a time data stream composed of CMB fluctuations, some strong astrophysical foreground (here the Galactic dust emission) and coloured Gaussian and stationary noise. The data is denoted as an N elements vector $\boldsymbol{d}$ (N is the number of time samples) and is composed of noise $\boldsymbol{n}$ and sky signal $\boldsymbol{s}$:

$$\boldsymbol{d} = \boldsymbol{s} + \boldsymbol{n}. \tag{1}$$

The noise power spectrum is given by $P(\nu) = \langle \tilde{\boldsymbol{n}}^* \tilde{\boldsymbol{n}} \rangle$ where $\tilde{}$ is the Fourier transform and $\langle\rangle$ the ensemble average over many realisations. The point here is to avoid contamination from signal (CMB and in a larger part Galactic dust) that would tend to overestimate the estimated noise power spectrum in time domain and therefore its equivalent in harmonic space leading to a biased C_ℓ spectrum at the end.

2.1. Raw estimation of the noise spectrum

The first raw estimation of the noise power can come from the data itself (that is $\boldsymbol{n} \simeq \boldsymbol{d}$) as we are dealing with low signal to noise experiments (at least in the time domain as the noise is reduced by redundancy in the map). This is a correct approximation only in the case where the observed region of the sky is free from any bright source. In most recent experiments such as WMAP and Archeops (as well as in Planck) the instrument covers a large portion of the sky by making large circles that often cross the Galactic plane twice per rotation. There is therefore a very bright source in all the portions of the time streams that prevent us from using the naive approximation $\boldsymbol{n} \simeq \boldsymbol{d}$.

Prior to any treatment we apply a high pass filter at $f_{\rm min}$ (in our example 0.3 Hz), estimating the noise power spectrum only at frequencies higher than $f_{\rm min}$, from now $\boldsymbol{d}$ will therefore refer to the filtered data. This $f_{\rm min}$ corresponds to the frequency where the non-astrophysical foreground or systematic (such as atmospheric emission) become negligible and where the low frequency noise is not too strong (in the $1/f$ noise model that would roughly corresponds to $f_{\rm knee}$).

A simple method to get rid of the signal contribution in the noise power spectrum estimation is to project the data onto a map of the sky portion observed. In such a map, the noise level is reduced but the signal is almost unchanged. Reading back this map with the scanning strategy leads to a timeline where the signal to noise is greatly improved. This latter timeline can then be subtracted from the initial timeline giving a noise estimate that is, in principle, free from signal contamination. In the matrix notation that has become common in CMB analysis this corresponds to the following operation:

$$\boldsymbol{n}_{\rm raw} = \boldsymbol{d} - A\left(A^T A\right)^{-1} A^T \boldsymbol{d} \tag{2}$$

where A is the pointing matrix. The noise spectrum estimation is then just obtained from the Fourier transform of $\boldsymbol{n}_{\rm raw}$.

2.2. Biasing effects

If one rewrites Eq. (2) using the definition (1), one gets:

$$\begin{aligned} \boldsymbol{n}_{\rm raw} &= \boldsymbol{n} - \underbrace{A\left(A^T A\right)^{-1} A^T \boldsymbol{n}}_{\boldsymbol{r}_{\rm n}} + \underbrace{\left(\boldsymbol{s} - A\left(A^T A\right)^{-1} A^T \boldsymbol{s}\right)}_{\boldsymbol{r}_{\rm s}} \\ &= \boldsymbol{n} - \boldsymbol{r}_{\rm n} + \boldsymbol{r}_{\rm s} \end{aligned} \tag{3}$$

where $\boldsymbol{r}_{\rm n}$ and $\boldsymbol{r}_{\rm s}$ can be interpreted in the following way:

residual noise ($\boldsymbol{r}_{\rm n}$): the number of time samples per pixel is finite (in Archeops and for a resolution of 15 arcmin we have in average 25 samples per pixel with a very asymmetric distribution) and therefore there remains some noise in the projected map. This noise in the map, once re-read with the scanning stategy is strongly correlated to the initial noise in the timeline. We therefore slightly underestimate the noise spectrum. Finally $\boldsymbol{r}_{\rm n}$ increases if the pixel size is reduced. This residual noise is also discussed in Stompor et al. (2002).

residual signal ($\boldsymbol{r}_{\rm s}$): because of the pixellisation of the projected map, the high temporal frequencies of the signal (mainly dust at low galactic latitude or point sources) are badly subtracted from the initial time stream. This leads to an overestimation of the noise power spectrum. Finally $\boldsymbol{r}_{\rm s}$ decreases if the pixel size is reduced as it only originates from the large size of the pixels compared to the frequencies present in the data.

When computing the noise estimate power spectrum, one gets:

$$P_{\rm raw}(\nu) = \langle (\tilde{\boldsymbol{n}} - \tilde{\boldsymbol{r}}_{\rm n} + \tilde{\boldsymbol{r}}_{\rm s})^* (\tilde{\boldsymbol{n}} - \tilde{\boldsymbol{r}}_{\rm n} + \tilde{\boldsymbol{r}}_{\rm s}) \rangle \tag{4}$$

$$= \langle (\tilde{\boldsymbol{n}} - \tilde{\boldsymbol{r}}_{\rm n})^* (\tilde{\boldsymbol{n}} - \tilde{\boldsymbol{r}}_{\rm n}) \rangle + \langle \tilde{\boldsymbol{r}}_{\rm s}^* \tilde{\boldsymbol{r}}_{\rm s} \rangle \tag{5}$$

because noise and signal are supposed uncorrelated. The first term can be rewritten as $(1 - \epsilon(\nu))\, P(\nu)$. Simulations show that $\epsilon(\nu)$ does not depend significantly on the noise power spectrum in the frequency range considered here ($\nu \geq 0.3$ Hz), but

[1] http://www.eso.org/science/healpix

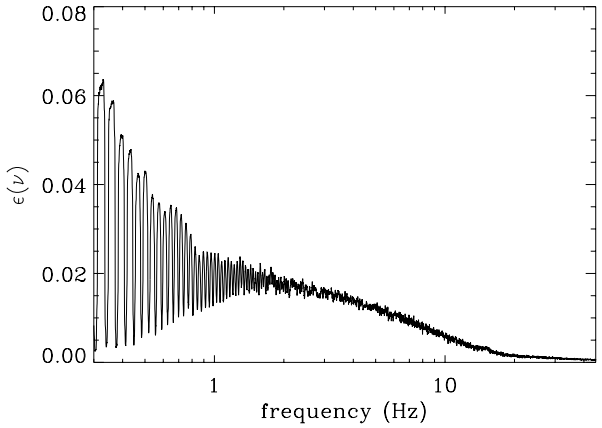

Fig. 1. Noise bias $\epsilon(\nu)$ (multiplicative).

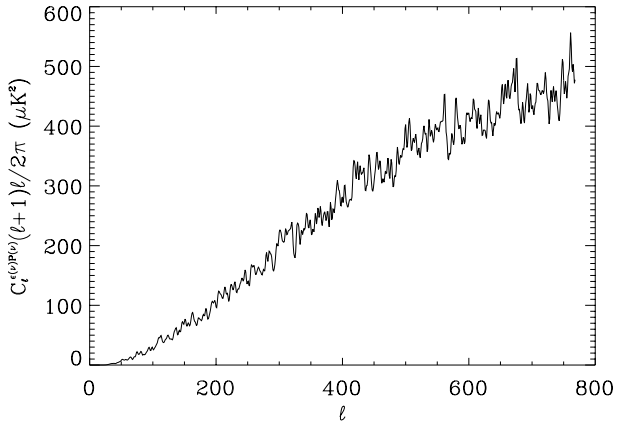

Fig. 2. Power spectrum of the noise bias $\epsilon(\nu) \times P(\nu)$ in harmonic space in Archeops-like case.

essentially on the scanning strategy. The second term is additive and depends on the signal contained in the map and can be defined as $\epsilon_s(\nu)$. One can therefore rewrite Eq. (5) as:

$$P_{\rm raw}(\nu) = (1 - \epsilon(\nu))\, P(\nu) + \epsilon_s(\nu). \quad (6)$$

In the case of the Archeops experiment these two effects are not negligible for the pixel sizes we are interested in. As they cannot be reduced simultaneously by increasing or decreasing the pixel size, both have to be accounted for in order to correctly estimate the noise power spectrum. In the following, we have chosen to use relatively small pixel size (15 arcmin) in order to reduce the contribution of ϵ_s to the bias as its determination, as will be seen later, relies on calibration and is model-dependent.

If one can determine the shape of both $\epsilon(\nu)$ and $\epsilon_s(\nu)$, then debiasing the noise power is straighforward by inverting Eq. (6).

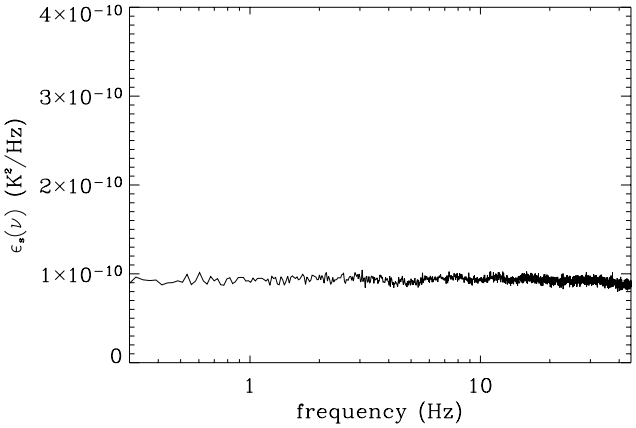

Fig. 3. Signal bias $\epsilon_s(\nu)$ (additive) in Archeops-like case.

2.3. Estimating the noise effect

The simplest way to estimate $\epsilon(\nu)$ is to make noise only Monte-Carlo realizations of data streams with spectrum $P_{\rm raw}$ (this assumes that $P_{\rm raw}(\nu)$ is close enough to P so that ϵ is computed with enough accuracy although the algorithm could be iterated) and perform the operation of Eq. (2) on each realisation. The average power spectrum of all realizations is biased by $1-\epsilon(\nu)$ that can be calculated by taking the ratio of the input power spectrum to the output. The result obtained is shown in Fig. 1. Large oscillations can be seen at lower frequencies with peaks at harmonics of the spinning frequency (0.035 Hz in our Archeops-like case). This is not surprising as we expect the noise to project onto the sky at these frequencies. It is therefore at these particular frequencies that the noise bias is the largest, leading to large values of $\epsilon(\nu)$. The noise bias transformed into harmonic power spectrum using the Archeops pointing strategy is shown in Fig. 2, showing that this effect should not be neglected as it amounts to a significant fraction of the error bars (shown in black in Fig. 6).

2.4. Estimating the signal effect

The signal bias estimation does not require averaging over realizations as it depends only on the sky signal. Here, we consider Galactic dust emission as an example of foreground signal contamination. We therefore implicitely consider observations at high frequency where it dominates. The same procedure could be applied to lower frequency observations just replacing the dust map (Schlegel et al. 1998) by either a synchrotron tracer maps (Haslam et al. 1982) map and/or a free-free tracer map (Finkbeiner 2003). The estimation is simply done by adding realistic Galactic dust signal on a noise only timeline (with spectrum $P_{\rm raw}(\nu)$). The realistic simulated Galactic dust signal is obtained from Schlegel et al. (1998) dust maps filtered with the corresponding time constants, beam and electronic filter. The noise+signal timeline is then deglitched and filtered the same way as for the real data. We then remove the initial noise before applying the operation of Eq. (2). The galactic bias $\epsilon_s(\nu)$ is the Fourier power spectrum of the residual signal. The signal bias $\epsilon_s(\nu)$ is shown in Fig. 3 in Fourier space and in Fig. 4 in harmonic space. Again, the power spectrum bias amounts to a significant fraction of the error bars, showing that the signal bias should not be neglected. The level of galactic bias highly depends on the level of Galactic signal in the data and therefore the calibration. The biggest uncertainty in the estimation of $\epsilon_s(\nu)$ therefore arises from the uncertainty in the instrumental calibration typically around 10% (in temperature). The residual error is finally relatively small compared to the signal as it only

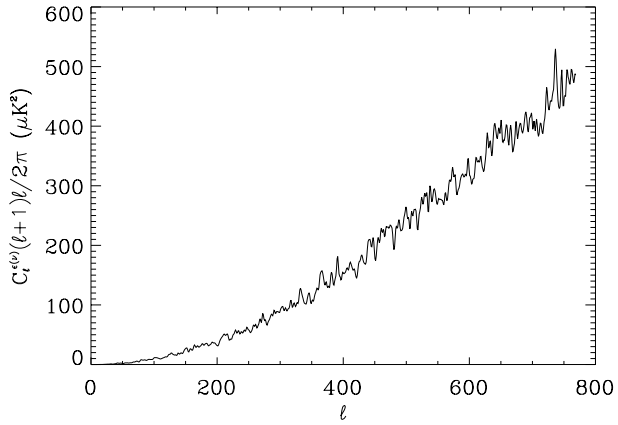

Fig. 4. Power spectrum of the signal bias $\epsilon_s(\nu)$ in harmonic space in Archeops-like case.

applies to the signal residual that is already a few percent of the whole signal (in Archeops-like case).

2.5. Accuracy on the reconstructed noise spectrum

After debiasing of both effects, we obtain a noise spectrum that should, in principle reflect the initial noise spectrum very accurately. In order to estimate this effect, we performed 60 simulations of the debiasing procedure based on the same inital noise spectrum (a realistic Archeops noise spectrum). The result is shown in Fig. 5. The noise reconstruction is better than two percent at all frequencies which is sufficient for accurate noise spectrum subtraction with MASTER in the case studied here.

3. Making maps with filtered data: Filtering without ringing

Making coadded maps using heavily filtered data (in the Archeops case only frequencies between 0.3 and 45 Hz are kept) is a simple and accurate alternative to optimal mapmaking when the knee frequency of the noise is low enough to prevent the filtering from removing too much signal in the low frequency part. Even when this is the case, one has to make sure that the filtering (obtained by setting low Fourier modes to zero) has the expected effect and does not pollute the data.

Spurious effects from filtering do arise if there is strong signal in the data. A good example of such a strong signal is the Galactic plane crossings that occur twice per rotation with Archeops. As the galactic peak is essentially non stationary the filtering creates ringing signal around the galactic plane in the maps. The structures created are not only located just around the galactic plane (where they are the strongest) but all over the map at a level comparable to the CMB anisotropies that are searched for, particularly at low multipoles as shown in blue in Fig. 6. We therefore need a trick to remove these ringing effects, that is, filtering without ringing. A similar effort is described in (Yvon et al. 2003).

The usual trick used to reduce ringing effects on filtered data is not to use too sharp filters. We used a filter with a sine shape ($\sin\nu$ between 0.24 and 0.36 Hz) going from 0 to 1 in 0.12 Hz. We obtained this way good reduction of the ringing effect but we wanted to improve further the mapmaking.

As the ringing arises from the presence of the Galaxy crossing in the data, removing the Galactic signal should reduce drastically the effect. A first step is to remove simulated galactic signal from Schlegel et al. (1998) but the agreement between the simulated maps and the true Galactic signal is not perfect and there remain high peaks in the data. This method also introduces external data into the experimental data, a procedure which has to be avoided if possible. We therefore decided to remove an estimate of the Galactic signal done with the experimental data itself. This is done in the following way:

- We first fit the time domain data with a set of low frequency functions (in our case sinc functions multiplied by a Gaussian, but other functions could be used, such as splines) excluding data located in bright regions of the Galaxy (using a mask made from Schlegel et al. 1998 galactic dust maps). The low frequency template obtained is then free from Galactic contamination.
- We then remove this low frequency template from the initial data and obtain a timeline containing mostly high frequency noise and Galactic signal and no low frequency stripes.
- This timeline is then simply coadded to obtain a map. We then set to zero the pixels out of the Galactic mask used before so that the map contains Galactic signal in the region known to be contaminated and zero elsewhere (in particular, in the region where the CMB analysis is to be performed).
- This map is read with the scanning strategy and the resulting timeline is subtracted from the initial data. The resulting timeline is stricly identical to the initial one in the CMB region and differs in the Galactic region just by the estimate of the Galactic signal.
- This last timeline is then filtered using either a tophat bandpass filter (green points in Fig. 6) or the sine shape filter mentionned above (red points in Fig. 6) between 0.3 and 45 Hz and coadded to make the map that will be used for the CMB analysis.

This method allows a very efficient removal of the galactic signal in the contaminated region without changing anything in the clean region where one wants to perform the CMB analysis. The efficiency of the ringing effect removal is shown in Fig. 6. The ringing coming from the Galactic signal is efficiently reduced to a level much smaller than the CMB anisotropies and the error bars on the power spectrum estimation.

4. Conclusions

We have developed a method that allows an accurate estimate of the noise power spectrum for CMB experiments accounting for small but significant effects such as the noise reprojection and the pixellisation. This method relies on Monte-Carlo simulations and does not require specific platforms to be implemented. It is designed to serve as an input to CMB power spectrum estimation that generally requires a precise estimate of the noise properties. This method has been succesfully applied in

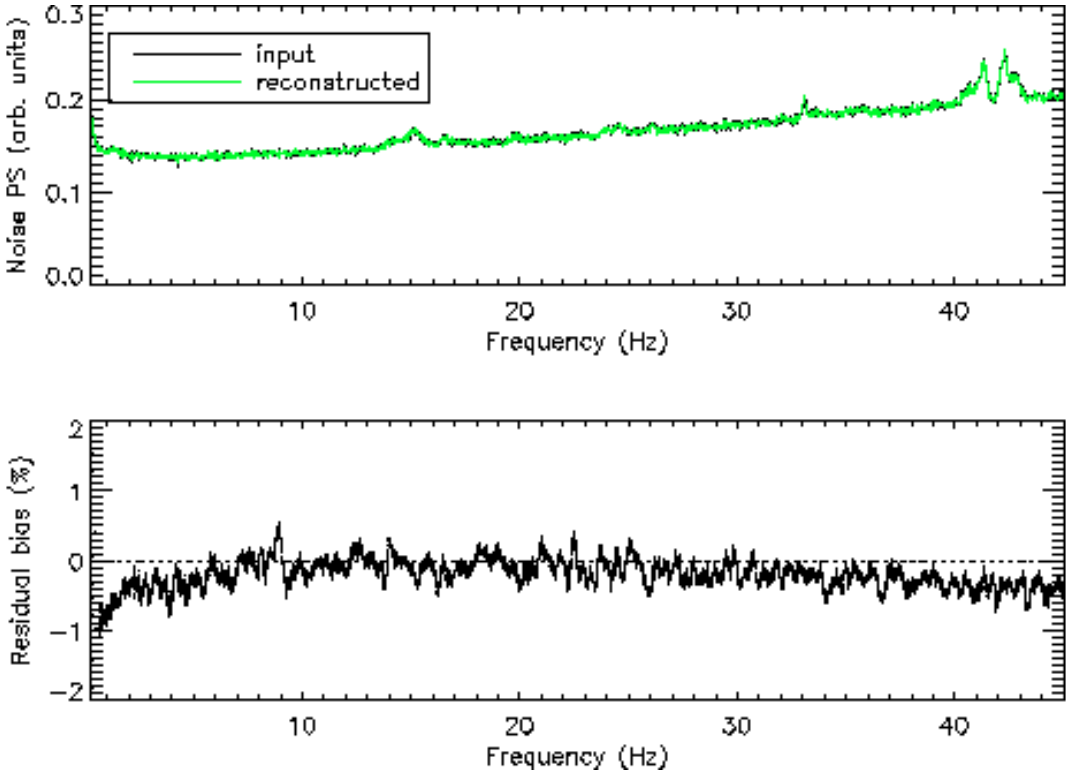

Fig. 5. Results of the noise spectrum reconstruction obtained with realistic realizations. The top panel shows the initial noise spectrum in green and the average reconstruction over the 60 realisations in black. The lower panel shows the relative residual bias which is much smaller than 2 percent at all frequencies.

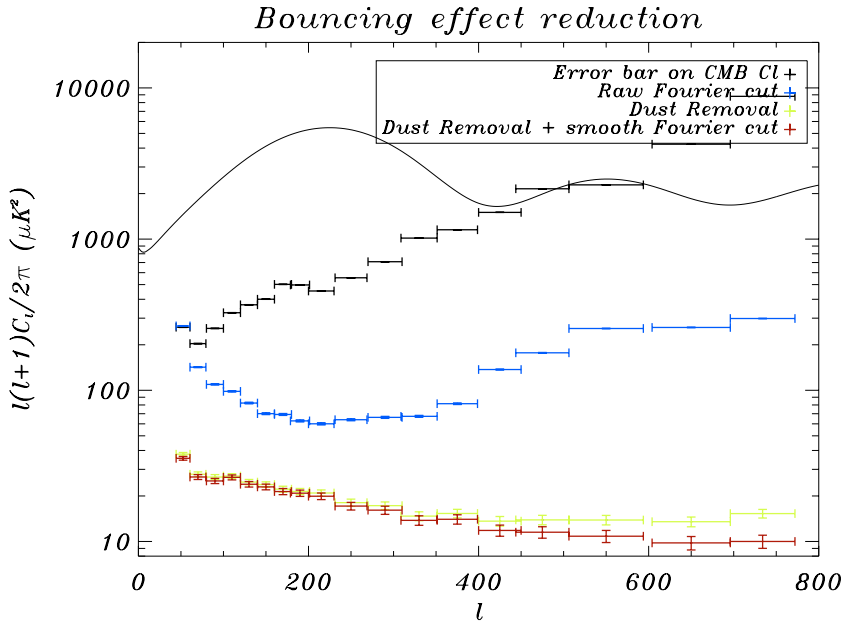

Fig. 6. The error bars on the power spectrum estimation are shown in black with a fiducial CMB power spectrum in solid black line. The excess power spectrum contribution coming from the ringing effect on the Galactic signal is shown in blue for a raw Fourier cut on the inital data. The effect of the dust removal by subtracting our estimate of the dust signal is shown in green. Adding the smooth sine shape filter instead of a raw Fourier cut leads to the red residuals.

the frame of the Archeops data analysis (Benoît et al. 2003a,b) but can be extended to any non-interferometric experiment.

Acknowledgements. The authors are grateful to the whole Archeops collaboration for uncountable stimulating discussions which have made this work possible. We wish to thank in particular M. Douspis, J.-F. Macías-Pérez and F.-X. Désert for their contributions.

References

Borrill, J. 1999 [astro-ph/9911389]
Hivon, E., Gorski, K. M., Netterfield, C. B., et al. 2002, ApJ, 567, 2
Szapudi, I., Prunet, S., & Colombi, S. 2001, ApJ, 561, L11
Benoît, A., Ade, P., Amblard, A., et al. 2003a, A&A, 399, L19
Benoît, A., Ade, P., Amblard, A., et al. 2003b, A&A, 399, L25
Stompor, R., Balbi, A., Borrill, J. D., et al. 2002, PhRvD, 65, 022003
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Gorski, K. M., et al. 1999, in Proceedings of the MPA/ESO Cosmology Conference "Evolution of Large-Scale Structure", 37 [astro-ph/9812350]
Haslam, C. G. T., Salter, C. J., Stoffel, H., & Wilson, W. E. 1982, A&AS, 47, 1
Finkbeiner, D. P. 2003, ApJS, 146, 407
Yvon, D., Mayet, F., & Magneville, Ch. 2003, to be published