Comment on the Application of the Independent Component Analysis for the Blind Source Separation of Microwave Astrophysical Backgrounds

P.Ballatore

Istituto di Scienza e Tecnologie dell'Informazione, National Research Council, Pisa, Italy

ABSTRACT

In order to separate the microwave astrophysical components superposed in maps of the sky, Maino et al. (2002) apply the Fastica algorithm, which is able to recover emissions from statistically independent processes with the only a priori assumption that all the components, except possibly one, must have non-Gaussian distributions. This comment highlights that Fastica is programmed in order to find orthogonal components, while the microwave astrophysical diffuse emissions cannot be considered, a priori, as orthogonal. Therefore, in this case, whatever blind source component separation technique that is not assuming orthogonality among the sources is expected to be better than Fastica. Merely for purposes of example, one of these possible algorithms is presented and compared with Fastica in a very simple toy model.

Keywords: data analysis methods, image processing techniques.

1. INTRODUCTION

Maino et al. (2002) ('M02') consider that the microwave sky maps are originated by the superposition of the cosmic microwave background (CMB) and other Galactic and Extragalactic components. Among these sources of radiation overlapping the CMB, they take into specific account the Galactic free-free, synchrotron and thermal dust emissions, which can be considered as the major sources of Galactic diffuse microwave radiation. After introducing the astrophysical necessity to separating the different components with high accuracy and reliability, M02 proposes the FASTICA as an extremely promising technique for analyzing the maps that will be obtained by the forthcoming high-resolution CMB experiments.

Considering the sources as elements of the vector $\underline{s}$ and the observations at different frequency channels as elements of the vector $\underline{x}$

$$\underline{x}=A\cdot\underline{s}+\underline{\varepsilon} \qquad (1)$$

A is the mixing matrix and the elements of the vector $\underline{\varepsilon}$ are the noises at the corresponding observational channels. Fastica derives the matrix W that gives the maximum neg-entropy among the elements of $\underline{y}$:

$$\underline{y}=W\cdot\underline{x} \qquad (2)$$

In fact, under certain hypothesis, W is equal to the inverse of the mixing matrix A and $\underline{y}$ is the best estimate of $\underline{s}$ (M02). This equivalence, as in other blind source separation methods, is essentially based on the principle of redundancy reduction, which refers to the neural coding and states that this coding is carried out so that the outputs are as independent as possible [Barlow, 1989].

In particular, Fastica is based on the assumption that the CMB and the Galactic emissions are independent. In fact each source component is assumed to be orthogonal to each other: if a certain number, k, of rows of W have been evaluated at one specific step, FASTICA algorithm estimates the (k+1) row by searching in the sub-space orthogonal to the first k rows (as reported in M02).

From Eq. 2, the perpendicularity among the rows of the matrix W implies the perpendicularity among the elements of $\underline{y}$, which are the estimated sources. Regarding this point, it is important that this orthogonality among the estimated sources is not confused with some effect related to one step of the pre-processing that is defined the 'whitening'. The whitening is the decorrelation of the observations (the elements of $\underline{x}$, not the estimated source elements of $\underline{y}$) at the different channels among themselves. At first, these observations are obviously correlated because these are different combinations of the same sources. If the sources are orthogonal, the observations are still correlated. The whitening, reducing to the minimum the correlations among observations, eliminates the common information repeated in each channel.

This paper highlights that significant correlations can be observed between Galactic emissions. Requiring orthogonality of the estimated sources, Fastica implies that the astrophysical emissions do absolutely not cross-correlate. In fact the orthogonality between $\underline{y}_1$ and $\underline{y}_2$ (the estimated sources, elements of $\underline{y}$) means that their scalar product is null, or, geometrically, that $\underline{y}_2$ has not components along the direction of $\underline{y}_1$. On the contrary, a linear cross-correlation between $\underline{y}_1$ and $\underline{y}_2$ implies that $\underline{y}_2$ can be expressed as a function of and $\underline{y}_1$, which implies that $\underline{y}_1$ and $\underline{y}_2$ have a common component along the same direction, geomatrically. Many different demonstrations can be found showing that the perpendicularity excludes the possibility of correlations. Therefore, FASTICA cannot be the best component separation technique for astrophysical data sets that might eventually be correlated. Specifically, even in the case of separating the CMB from the rest of Galactic component (considered all together and un-correlated with the CMB), the reconstruction of the CMB and its spectrum is not as good as any other algorithm similar to the Fastica, but which does not rely on the orthogonality assumption.

2. OBSERVATIONS

2.a. Correlations among microwave astrophysical sources of diffuse radiation

The major Galactic radiations contributing to the observations at the considered frequency range are from free-free, synchrotron and thermal dust emissions. If two directions of observation have a certain difference in physical parameters (dust abundance, temperature, electron density, etc.), this causes also a difference in the free-free, synchrotron and thermal dust emissions observed in these two directions. This difference is related to the level of differences in the physical parameters in each of the three cases. For this reason, there may be relationships among Galactic emissions from different processes. For example, considering the all-sky maps of the Analysed Science

Data Set (ASDS) for the Differential Microwave Radiometer (DMR) mounted on COBE satellite (Bennet et al., 1992), the following linear correlation coefficients are found: 0.9 between the synchrotron and thermal dust emissions, 0.85 between the free-free and the thermal dust emissions, 0.76 between the synchrotron and the free-free emissions. These correlation coefficients are found considering the maps of all the sky, with 6144 data points, and correspond to a statistical confidence level above 99.9%. This significance is not merely due to the enhancements in all emissions observed in the galctic plane. In fact, considering separately patches of the sky out of the Galactic plane significant correlations (with possibly smaller coefficients) can be observed, especially between the synchrotron and the thermal dust emissions. Details about the dependence of these correlations on the Galactic coordinates are reported by Ballatore et al. (2002). Here, for example, it can be worth mentioning that correlation coefficients of about 0.8 or about 0.55 are obtained between ASDS data of synchrotron and thermal dust for slices of sky at Galactic latitudes in the ranges, respectively, (50, 70)° or (-50, -30)°.

These large correlations can be in part attributed to the inclusion of the Galactic equatorial region (source of more intense emissions for all the processes of interest) and in part to the relatively low angular resolution of DMR-COBE (7° FWHM). In fact, it is reasonable that a decrease of correlations be associated with an increase of resolution, when smaller scale differences among emissions can be appreciated. However, based on previous observations, the Fastica orthogonality among the components cannot be postulated as an a priori characteristics of the astrophysical sources of microwave emissions.

Regarding the CMB component, it is worth mentioning that it is reasonably expected to be independent from the Galactic emissions. However, this cannot exclude accidental occurrence of relatively weak CMB/Galactic correlations, which would have no astrophysical meaning, but can affect the component separation computations, especially in some regions of the sky. In this context, the orthogonality of the CMB to the other Galactic emissions is a Fastica a-priori constrain that would be better to avoid, leaving the free possibility of eventual orthogonality appearing as a result of the component separation algorithm.

2.b. Average Coherence Minimization Algorithm

According to the principle of redundancy reduction, the reconstructed sources should be as independent as possible. As in the Fastica, this can be interpreted as the minimisation of the neg-entropy and, once one source is estimated, the orthogonality among the sources can be used for facilitating the estimation of the rest of the components. If it is not known about the source dependence, the principle of redundancy reduction still suggests the minimisation of whatever function expressing the mutual common information among the sources. In particular, one case, examined here for purposes of example, is the minimisation of the average spectral coherence (Barros, 2000). This Minimum Coherence Method ('MCM' in the following) estimates the elements of the matrix W of Eq. 2, by minimising the coherence (averaged over all the frequencies) among the sources estimated, i.e. among the elements of the vector $\underline{y}$. The case of independence among the sources can be solved by this technique, giving a minimum average coherence equal to zero. More specifically, in the simple case of only two observations and two sources, the MCM algorithm consists in the minimisation of the function

$$C = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{| S_{y_1 y_2}(\omega) |^2}{S_{y_1}(\omega) S_{y_2}(\omega)} d\omega$$

where $\underline{y}_1$ and $\underline{y}_2$ are the elements of the vector $\underline{y}$ given in the Eq. 2, $S_{y_1 y_2}$ is the cross-power spectrum between $\underline{y}_1$ and $\underline{y}_2$, S_{y1} and S_{y2} are the power spectra of $\underline{y}_1$ and $\underline{y}_2$, respectively. This algorithm has been developed with different minimisation methods. In particular, results described in the following have been derived using two minimisations: (a) the reticulum method: variation of parameters into ranges of discrete values; (b) the simulating annealing method (Kirkpatrick et al.; 1983), which is more convenient in the case of many sources and many channels. The difference between these two methods is not significant in the simple case considered in the following. Further studies are in course of development about the performance of the MCM algorithm and its possible applicapibility in more complicated cases of multiple unknown sources.

For purpose of comparison, the MCM and the Fastica have been applied to a toy case obtained by mixing the same CMB and synchrotron simulations as described by M02 in its paragraphs 4.1 and 5.1, respectively. Instead of all the sky, we consider only a patch image of 256 x 256 pixels centred around Galactic coordinates (30°, 30°), with pixel size 1.7 arcmin. These image arrays have been extrapolated to frequencies 100 GHz and 30 GHz and summed to simulate the two observations (elements of $\underline{x}$) at these frequency channels. In the case of mixtures of CMB and synchrotron, the Fastica is expected to work properly because of the theoretical independence of these radiations. Actually, the two image sources considered here are not orthogonal, having a correlation coefficient equal to about -0.13 and an average coherence of about $1.5e^{-2}$, both relatively small and possibly accidental.

After the matrix W has been derived, the CMB and synchrotron images have been reconstructed and illustrated in Figure 1 for the two methods, together with the sources considered. Results from MCM algorithm are better than those from Fastica, as further shown by the histogram in Figure 2. These histograms report the distributions of the differences (pixel by pixel) of the reconstructed CMB and synchrotron minus our CMB and synchrotron sources. These differences have been calculated after having scaled all the six images to values between 0 and 255, so that 0 corresponds to the specific minimum in each image and 255 to the maximum. This is in agreement with the fact that the source estimates are derived independently from possible normalizations.

In order to test how the spectral properties of the reconstructed sources correspond to the spectral properties of the sources, spectra have been calculated for all the six image arrays reported in Figure 1. In a plane approximation, the two-dimensional FFT subroutine 'fourn' (Press et al., 1992) has been used. Each output spectrum has been normalised according to the variance of the corresponding image, so that each spectrum can be compared with the others. Finally, the histograms of the distributions of the differences between the spectra of the sources and of their estimates by MCM and Fastica are plotted in Figure 3. These histograms show that the spectra reconstructed using MCM are more accurate than those derived using the Fastica.

Noise is not discussed in this comment, but it can be worth mentioning that preliminary tests demonstrate that a level of noise of 10\% in each channel is not appreciably degrading the MCM performance.

3. CONCLUSIONS

This brief paper highlights that the Fastica assumption of orthogonality among the sources is not appropriate to the microwave astrophysical diffuse emissions superposed in maps of the sky. In fact, this orthogonality is not always observed in the data of interest, especially in the case of separating the Galactic emissions among themselves. Therefore other possible blind source separation techniques can be considered more promising than the Fastica, if they do not require this assumption. For example, the results from Fastica and one possible different algorithm (MCM) are compared in an examplary simple toy model, showing that accuracy better than Fastica can be reached.

REFERENCES

Ballatore P., E. Salerno, E. Kuruoglu, L. Bedini, A. Tonazzini, *ISTI/CNR Technical Report*, TR-18-2002, 2002.

Barlow H.B., 1989, *Neural Computation*, 1, 295, 1989.

Barros, A.K., in *Advances in Independent Component Analysis*, Ed. M. Girolami, Springer-Verlag, pp. 63-71, 2000.

Bennett C.L. et al., *ApJ*, 396, L7, 1992.

Kirkpatrick S. et al., *Science*, 220, 671, 1983.

Maino D., A. Farusi, C. Baccicalupi, F. Perotta, et al., *MNRA*, 334, 53, 2002.

Press W.H. et al., *Numerical Recipes*, Cambridge University Press, New York, 1992.

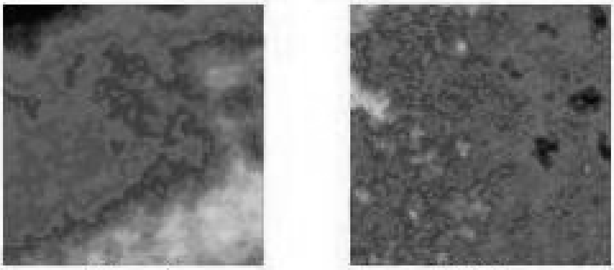
FIGURE CAPTIONS.

Figure 1. Simulated CMB and synchrotron signals (top panels), their re-constructions by MCM (middle panels) and by Fastica (bottom panels).

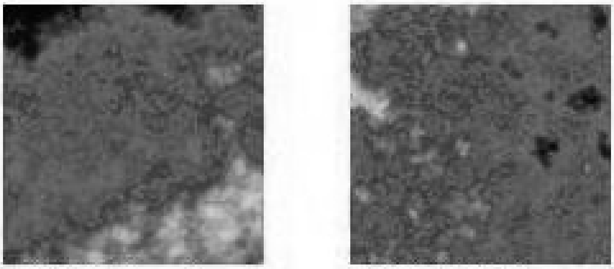
Figure 2. Histograms of the differences between the sources considered and their reconstructions by Fastica (white columns) and by MCM (gray columns).

Figure 3. Histograms of the differences between the spectra of the sources considered and the spectra of their reconstructions by Fastica (white columns) and by MCM (gray columns).

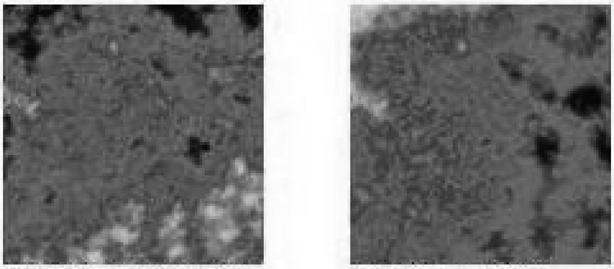
FIGURE 1.



Synchr. CMB



MCM Synchr. MCM CMB



Fastica Synchr. Fastica CMB

FIGURE 2.

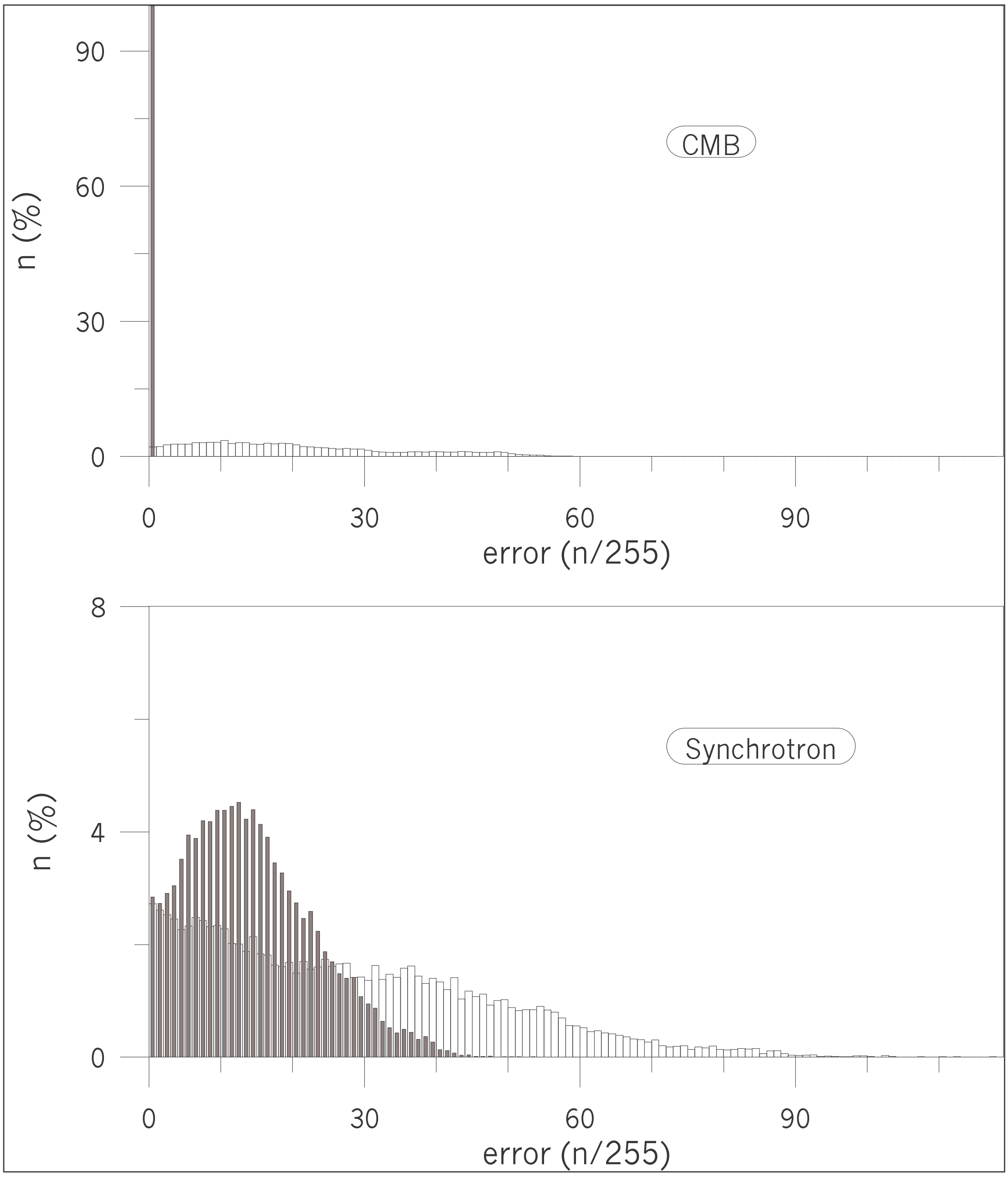
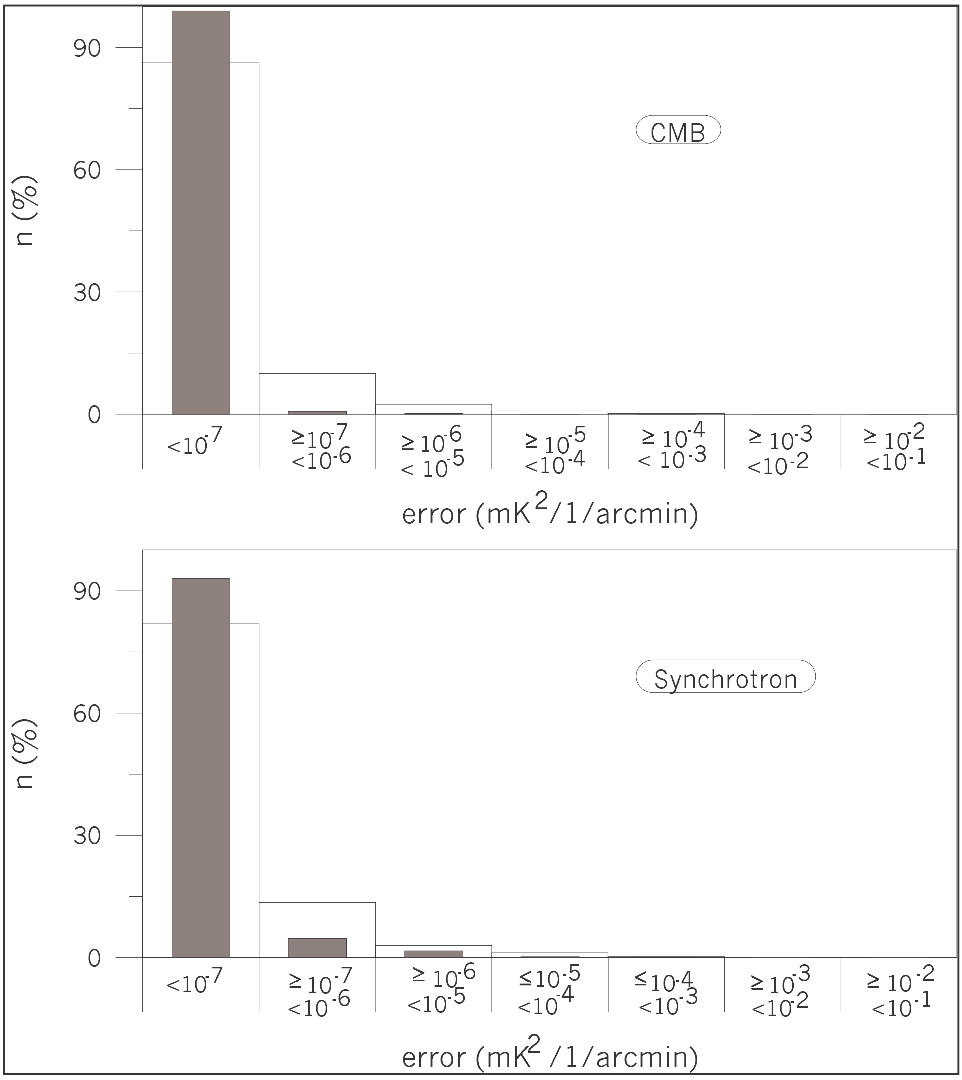

CMB
n (%)
90
60
30
0
0
30
60
90
error (n/255)
Synchrotron
n (%)
8
4
0
0
30
60
90
error (n/255)

FIGURE 3.


90
60
30
0
n (%)
CMB
<10-7
≥10-7 <10-6
≥ 10-6 < 10-5
≥ 10-5 <10-4
≥ 10-4 < 10-3
≥ 10-3 <10-2
≥ 10-2 <10-1
error (mK2/1/arcmin)
90
60
30
0
n (%)
Synchrotron
<10-7
≥ 10-7 <10-6
≥ 10-6 <10-5
≤10-5 <10-4
≤10-4 <10-3
≥ 10-3 <10-2
≥ 10-2 <10-1
error (mK2 /1/arcmin)