Application of Minimum Coherence Criterion For Blind Source Separation

P. Ballatore and L. Bedini

Istituto di Scienza e Tecnologie dell'Informazione, C.N.R.,
Via Moruzzi, Pisa, 56124 – Italy

Abstract. The Blind Source Separation (BSS) is performed using multiple observations of mixtures of dependent or independent astrophysical sources of diffuse microwave radiation. Two separation methods are used and compared: (a) FastICA, which has already been used for similar astrophysical data, (b) the minimum coherence criterion. Results show that both methods give better results in the case of independence among the sources than in the case of dependence. In particular, in the case of the data sets considered, method (b) (newly applied to similar astrophysical data) generally gives better source estimates than method (a), especially in the case of correlated sources, where FastICA is not optimal because of its assumption of orthogonality among the sources.

1. Introduction

Under certain hypothesis, an unknown linear mixture of source signals can be separated into its components by Blind Source Separation (BSS) methods. These methods are devised to be applied to whatever source signal and they are based on the a-priori principle of redundancy reduction [Barlow, 1989]. This principle, as a form of neural coding, states that this coding is carried out so that the outputs are as independent as possible.

Based on this principle, the Independent Component Analysis performs BSS assuming that the components are mutually independent [Comon, 1994]. Intuitively, an equivalent approach to achieve independence is to maximize non-Gaussianity. In particular, Hyvärinen [1999] and Hyvärinen and Oja [2000]

have proposed the negative entropy ('neg-entropy') as a measure of non-Gaussianity and have performed separation by maximising it via the Newton algorithm. Apart from the independence of the sources, this procedure assumes that these sources have non-Gaussian distribution, save at most one. This algorithm is known as FastICA [see, e.g., Maino et al., 2002) and it is capable to solve component separation, even in presence of a noise whose covariance matrix is known.

If the component sources are not independent the redundancy reduction can be considered still valid, but the independence of the sources cannot be postulated. The so-called Dependent Component Analysis approach achieves the maximum source independence by minimizing the information common to the estimated components in the mixtures. In this context, the strategy of reducing redundancy by minimizing the output mutual spectral overlap has already been investigated, and the related algorithm is essentially based on the minimization of the averaged coherence function of the outputs [Barros, 2000]. In particular, this procedure takes also into account the case of independent sources as a special case in which the minimum coherence achievable is close to zero. In the following, we call this approach 'MCM' as short of 'Minimum Coherence Method'.

This paper deals with the comparisons among BSS results obtained using the FastICA and the DCA coherence method in cases of mixtures of dependent or independent components.

2. Data and Data Analysis

The data considered here are derived from measurements made by the Differential Microwave Radiometers (DMR), mounted on the COBE satellite [Bennett et al., 1992]. In particular, the DMR-COBE database contains images of all the sky for three galactic diffuse emissions (synchrotron, thermal dust and free-free) and for the extragalactic CMB (Cosmological Microwave Background) component. These sources were derived from observations of the whole sky (composition of all overlapping signals on the same line of sight),

made with a 7° beam and are presented in 2.8° pixels. The separation of the sky in its source components was computed using traditional mathematics methods and astrophysical modelling [e.g., Bennet et al., 1992]. In consequence, the galactic and CMB source values might be affected by the approximations applied in the estimation procedures. At the same time, these data represent the best measure of the considered source emissions for synchrotron, dust and CMB. In any case, this is not of interest in this context, where the image patches considered represent solely three image components, two of which are correlated and the third one has correlation coefficients with the other two close to zero.

From a theoretical point of view, statistically significant correlations are expected among the galactic components, while the CMB should be uncorrelated. Experimentally, the significance of the observed correlations is found to be dependent on the galactic coordinates [Ballatore et al., 2002]. Here, we use, specifically, the three patches for the astrophysical radiations synchrotron, dust and CMB. Each patch image has 32x32 pixels and is centred on the North Galactic Pole. The correlation coefficients among them are: (a) 0.67 between synchrotron and dust, (b) 0.08 between dust and CMB, (c) 0.09 between synchrotron and CMB.

The purpose of this study is to test how FastICA and MCM can perform the component separation of correlated and uncorrelated source-image mixtures having a number of observations $\underline{x}_i$ equal to the number of sources $\underline{s}_i$. Sources and observations are simultaneous and time stationary.

In order to simulate mixtures of correlated and uncorrelated image signals, we have calculated the toy observations $\underline{x}_1$ and $\underline{x}_2$: (1) as linear mixtures of dust and synchrotron, (2) as linear mixtures of dust and CMB. Then, the purpose of FastICA and MCM is the estimation of the source images $\underline{s}_1$ and $\underline{s}_2$, given the mixtures $\underline{x}_1$ and $\underline{x}_2$

$$\begin{cases} \underline{x}_1 = a_{11} \cdot \underline{s}_1 + a_{12} \cdot \underline{s}_2 \\ \underline{x}_2 = a_{21} \cdot s_1 + a_{22} \cdot \underline{s}_2 \end{cases} \qquad [1]$$

or

$$\underline{x}=A\cdot\underline{s} \qquad [2]$$

Where a_{ij} are the elements of the matrix A, $\underline{x}_1$ and $\underline{x}_2$ are the elements of $\underline{x}$, $\underline{s}_1$ and $\underline{s}_2$ are the elements of $\underline{s}$. The elements, a_{ij}, of the mixing matrix, A, have been chosen randomly. Specifically, in the case shown in Figure 1, these are:

$$A=\begin{pmatrix}0.2 & 1.4\\ 1.7 & 0.5\end{pmatrix} \qquad [3]$$

For the mixture of synchrotron, $\underline{s}_1$, and dust, $\underline{s}_2$. In the case (shown in Figure 2) of dust, $\underline{s}_1$, and CMB, $\underline{s}_2$, the mixing matrix A has been chosen as:

$$A=\begin{pmatrix}0.7 & 1.0\\ 1.2 & 0.5\end{pmatrix} \qquad [4]$$

Results obtained by application of FastICA and MCM are independent from the specific channel simulations given in Figure 1 and 2, as verified by using different A matrices.

2a. FastICA algorithm

This algorithm is devised to find the W matrix in the general case of noise data:

$$\underline{\tilde{s}}=W\cdot(A\cdot\underline{s}+\underline{\varepsilon})=W\cdot\underline{x} \qquad [3]$$

where $\underline{\varepsilon}$ is the noise, $\underline{\tilde{s}}$ components are the source estimates and $\underline{x}$ components are the observations. The mean of the data in each observation $\underline{x}_i$ is removed and the covariance matrix C of the zero mean data is computed as the expectation over the set of all available pixels

$$C=E\{\underline{x}\underline{x}^T\} \qquad [4]$$

Given the covariance matrix for the noise Σ, the pre-processing consists in evaluating the matrix:

$$\hat{\Sigma}=(C-\Sigma)^{-0.5}\Sigma(C-\Sigma)^{-0.5} \qquad [5]$$

and the quasi-whitened data set:

$$\underline{\hat{x}}=(C-\Sigma)^{-0.5}\underline{x} \qquad [6]$$

Then the algorithm estimates the matrix W row by row as follows:

(a) choose an initial vector w (w^T is a row of W);

(b) update it through

$$w_{new} = E\{\underline{\hat{x}} g(w^T \underline{\hat{x}})\} - (I + \hat{\Sigma}) w E\{g'(w^T \underline{\hat{x}})\} \quad [7]$$

where E denotes expectation over all the available samples, g is a regular non quadratic function such as $g(u)=u^3$, or $g(u)=tanh(u)$, or $g(u)=exp(-u^2)$ and g' is its first derivative [Hyvärinen, 1999];

(c) $w_{new}=w_{new}/\|w_{new}\|$ [8]

(d) compare w_{new} with previous one: if not converged go back otherwise begin another process.

This procedure maximizes the non-Gaussianity of $w^T \underline{\hat{x}}$, so that, once estimated the K^{th} of W, the line $(K+1)^{th}$ is searched in the sub-space orthogonal to the first K rows, via orthogonalization between step (b) and (c).

In this report, we do not take the noise into account.

2b. MCM algorithm

This algorithm consists in the minimization of the averaged coherence function defined as

$$\varsigma = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|P_{s_1 s_2}(\omega)|^2}{P_{s_1 s_1}(\omega) P_{s_2 s_2}(\omega)} d\omega \quad [9]$$

Where $P_{s_1 s_2}(\omega)$ is the cross-power spectrum between the estimates of the sources $\underline{\tilde{s}}_1$ and $\underline{\tilde{s}}_2$, and $P_{s_1 s_1}(\omega)$ [$P_{s_2 s_2}(\omega)$] is the power spectrum of the source $\underline{\tilde{s}}_1$ [$\underline{\tilde{s}}_2$]. The estimates of the sources are

$$\underline{\tilde{s}} = W \underline{x} = WA\underline{s} \quad [10]$$

Where Eq. 2 is used for the second equivalence. The best W matrix is the one for which WA is the closest to the identity matrix I.

The W coefficients, w_j^i, are derived as those coefficients that minimize the average coherence ς. Specifically, the minimization is performed using the simulating annealing method.

3. Comparisons between FastICA and MCM results

We have applied the FastICA and the MCM to the two-channel simulated observations reported in Figure 1 and Figure 2. For the dependent case (see Figure 1), the results are illustrated in Figure 3: the upper panels represent the sources estimated using FastICA, the middle panels represent the sources estimated using MCM and the bottom panels represent the original sources as reported in Figure 1. All the six images shown in Figure 3 are shown using a 255 color scale, so that zero corresponds to the minimum of specific image and 255 to its maximum value. In fact, the sources are derived independently from a possible multiplicative or additive constant. Figure 3 shows that MCM produces source estimates better than the FastICA.

In order to quantify the difference between the two BSS methods, we have reported in Figure 4 the histograms for the differences between the original sources and their estimates using FastICA or using the MCM. Figure 4 confirms and highlights results visible in Figure 3.

Apart from the histograms in Figure 3, another indicator of the precision of FastICA and MCM source estimates is the product between the derived estimates of W and the A reported in Eq. 3 and Eq. 4. The closest is this index to I, the best is the component separation reached. The matrix W given by the MCM for the dust and synchrotron mixtures is

$$W = \begin{pmatrix} -0.437 & 0.76 \\ 0.98 & -0.106 \end{pmatrix} \qquad [11]$$

Results associated with this W give a coherence value between the source estimates equal to $9.4 \cdot 10^{-4}$. The W derived by FastICA and associated to the results shown in Figure 3 and 4 is

$$W = \begin{pmatrix} 448.13 & -622.0 \\ 24.6 & -300.8 \end{pmatrix} \qquad [12]$$

This matrix W is the output of the FastICA and it can be normalized for convenience. It is easy to verify that the numbers out of the diagonal of the matrix WA obtained from MCM are closer to zero compared to those obtained from FastICA.

Similarly to the case of the mixtures between dust and synchrotron, we have applied FastICA and MCM algorithms to the two channel observations reported in Figure 2 for the independent sources and the results are shown in Figure 5. Again the upper panels show the sources estimated using FastICA, the middle panels those estimated using MCM and the bottom panels show the original sources as reported in Figure 2. In addition, similar to Figure 4, Figure 6 shows the histograms for the differences between the original sources and their estimates using FastICA or MCM. Figure 6 shows quite a surprising precision for the MCM separation obtained in this case: both sources are perfectly derived. The matrix W for this MCM case is

$$W = \begin{pmatrix} -0.34 & 0.68 \\ 1.61 & -1.18 \end{pmatrix} \qquad [13]$$

with a coherence value between the source estimates equal to $1.7 \cdot 10^{-3}$. Results reported in Figure 5 and 6 for the FastICA, correspond to

$$W = \begin{pmatrix} 476.3 & -953.8 \\ -19.48 & 27.22 \end{pmatrix} \qquad [14]$$

The comparison between Figure 3 and Figure 5 (or Figure 4 and Figure 6) demonstrates that both separation algorithms work better in the case of the independent sources and MCM is generally better than the FastICA, even in this latter case.

5. Summary and Conclusion

Two BSS methods are applied to images representing mixtures of astrophysical sources of diffuse microwave radiations: (a) FastICA (e.g., Maino

et al., 2002), and (b) Minimum Coherence Method (Barros, 2000). The major difference between the two methods is that the FastICA is based on the orthogonality of the components, while MCM is not. In particular the FastICA has previously been applied to the component separation of microwave diffuse radiations similar to the present data sets, although these may present or not correlations among the sources (Ballatore et al., 2002). Differently, no previous applications of the minimum coherence method to similar astrophysical data are known in literature.

For the present data sets, the estimates of uncorrelated sources are more precise than the estimates of correlated sources for both the BSS algorithms considered. However, the precision of the minimum coherence method is generally higher than the precision of the FastICA, especially for the case of correlated components.

References

Ballatore P., et al., Correlation among COBE microwave astrophysical sources, *CNR Technical Report*, ISTI-2002-TR-16, ISTI/CNR, Pisa (Italy), 2002.

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

Barros, A.K., in Advamces in Independent Component Analysis, Ed. M. Girolami, Springer-Verlag, London, 2000.

Bennett, C.L., G.F. Smooth, M. Janssen, et al. 27 co-authors, COBE Differential Microwave radiometers: Calibration Techniques, ApJ, 391, 466, 1992.

Comon, P., Independent component analysis, a new concept ?, *Signal Processing*,24, 287, 1994.

Hyvärinen, A., *IEEE Signal Processing Lett.*, 6, 145, 1999.

Hyvärinen, A. and E. Oja, *Neural Networks*, 13, 411, 2000.

Maino, D., A. Farusi, C. Baccicalupi, F. Perrotta, et al., Mon.Not.R.astron.Soc., 334, 53, 2002.

Figure Captions

Figure 1. The two dependent sources Dust and Synchrotron are shown (upper panels), together with their combination into the simulated observed channels (bottom panels).

Figure 2. Similar to Figure 1, but for the independent sources Dust and CMB.

Figure 3. The original dependent sources are shown in the bottom panels, the sources estimated using MCM algorithm are shown in the middle panels and the sources estimated using FastICA are reported in the top panels.

Figure 4. Distribution of the absolute value of the differences between pixel value for the original dependent sources and for their estimates.

Figure 5. Similar to Figure 3, but for the independent sources.

Figure 6. Similar to Figure 4, but for the independent sources.

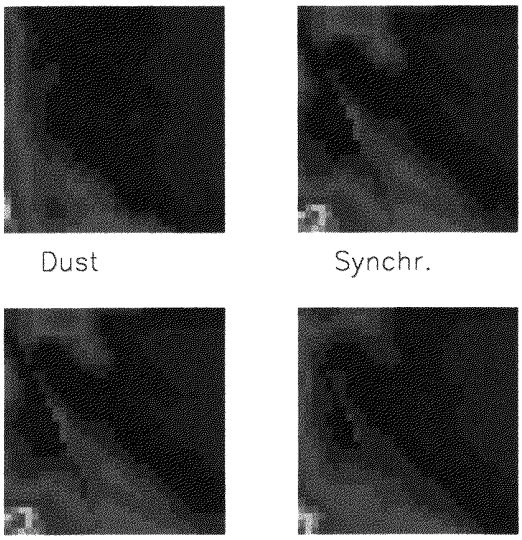
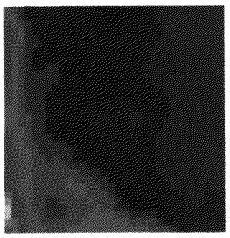
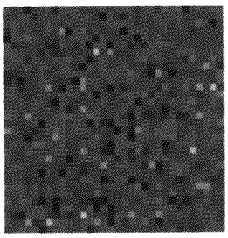


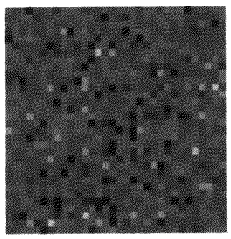
Figure 1.



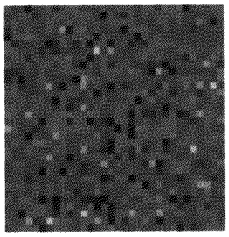
Dust



CMB



Channel 1



Channel 2

Figure 2.

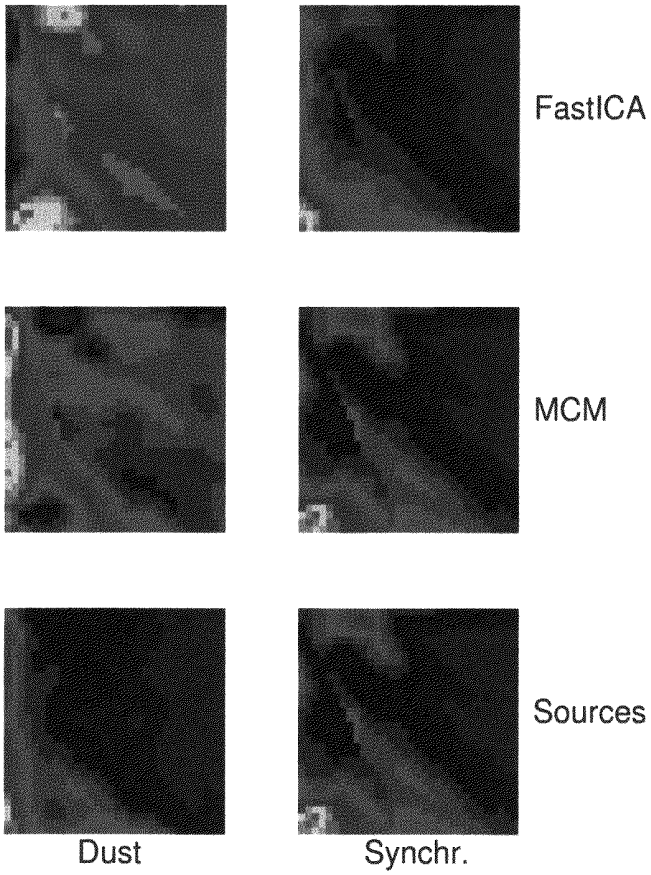


Figure 3.

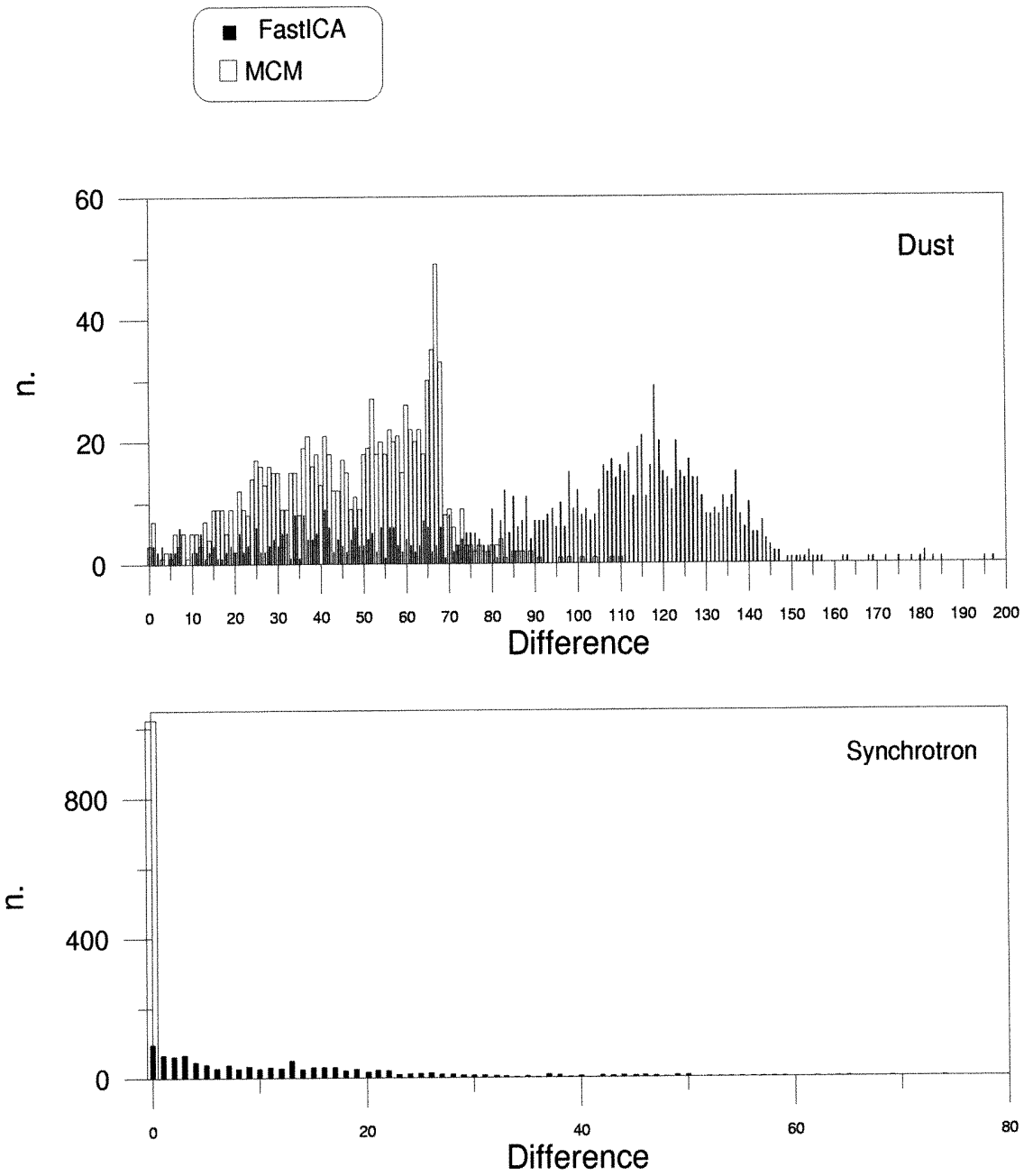


Figure 4.

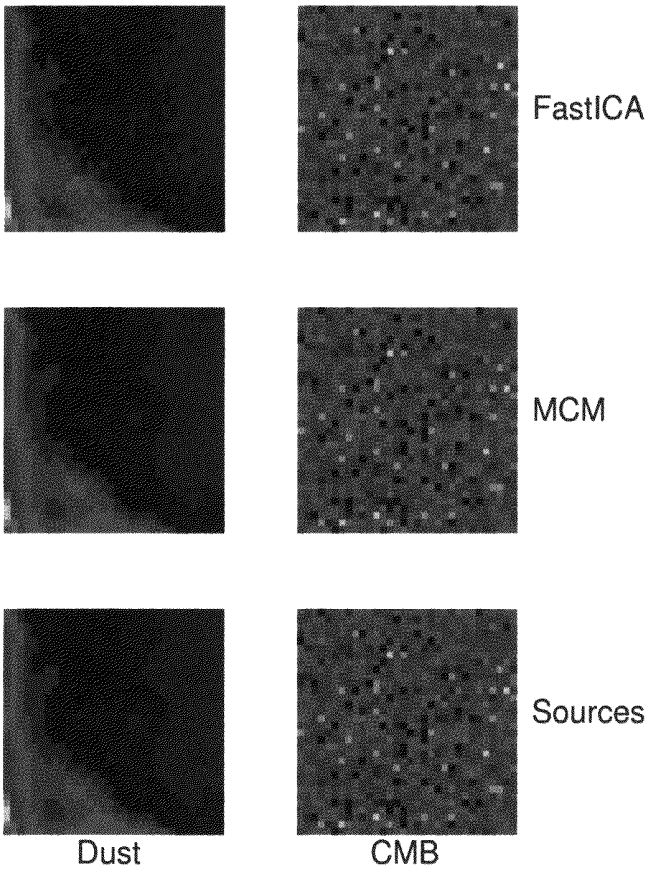


Figure 5.

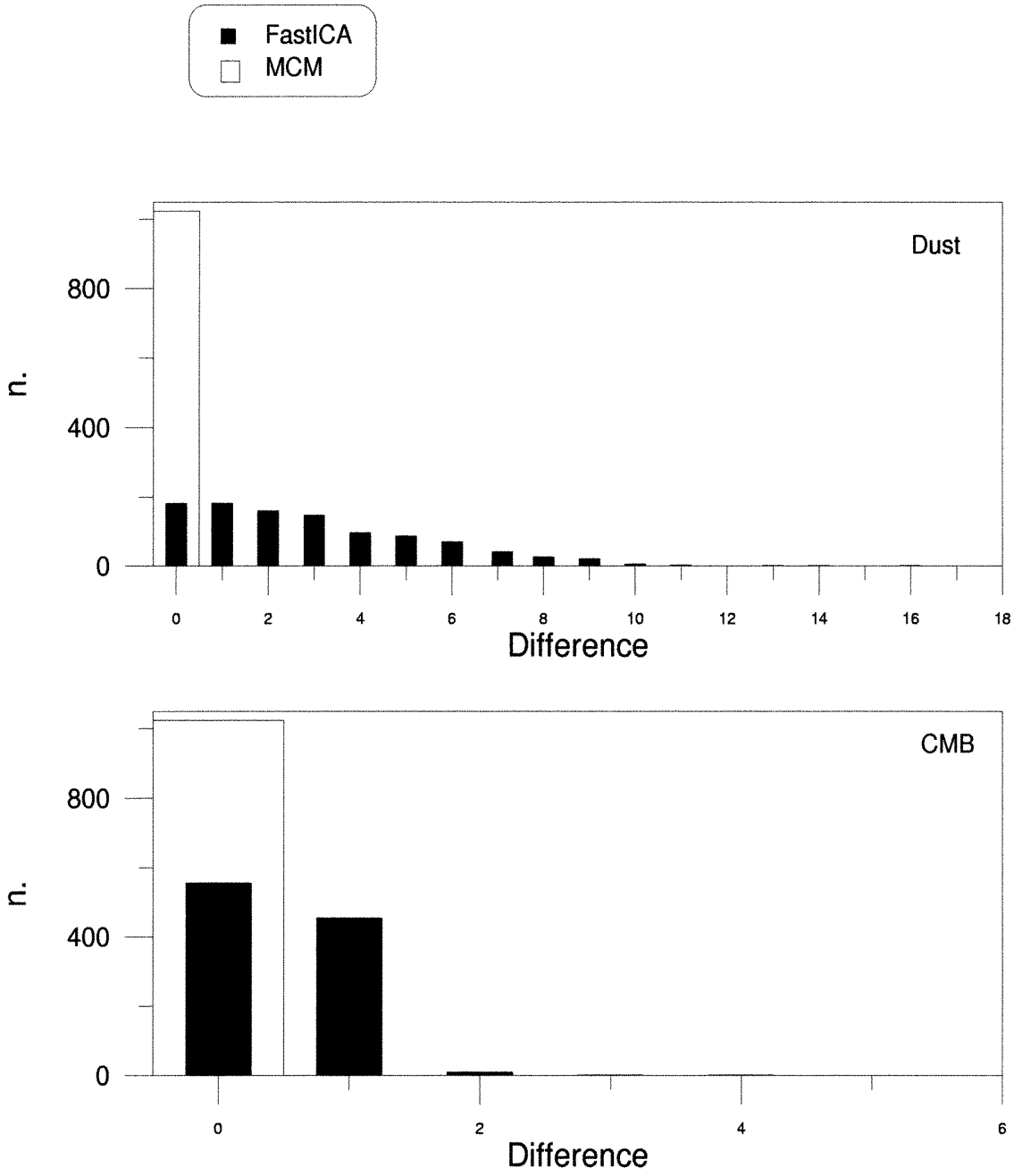

FastICA
MCM
Dust
800
400
0
n.
0
2
4
6
8
10
12
14
16
18
Difference
CMB
800
400
0
n.
0
2
4
6
Difference

Figure 6.