Separation of Anthropogenic Noise and Extremely Low Frequency Natural Magnetic Field Using Statistical Features

1st Jesús Rodríguez-Camacho
Department of Applied Physics
University of Granada
Granada, Spain
jesrodcam@ugr.es

2nd David Blanco-Navarro
Department of Applied Physics
University of Granada
Granada, Spain
dblanco@ugr.es

3rd Juan Francisco Gómez-Lopera
Department of Applied Physics
University of Granada
Granada, Spain
jfgomez@ugr.es

4th Jesús Fornieles-Callejón
Department of Electromagnetism
University of Granada
Granada, Spain
jforniel@ugr.es

5th M. Carmen Carrión
Department of Applied Physics
University of Granada
Granada, Spain
mcarrion@ugr.es

***Abstract*—This contribution is aimed at separating the anthropogenic noise from the natural magnetic field in the measurements recorded in a Extremely Low Frequency band station located in Sierra Nevada, Spain, to experimentally study Schumann resonances. First, we use a scheme based on Independent Component Analysis that provides unsatisfactory results. In order to achieve a better separation, we develop a new method by exploiting the information that statistical moments give us, under the assumption that the statistical distributions of the anthropogenic noise and of the natural magnetic field are different. This method consists of finding the rotation of the two original directions of the magnetometers (North-South and East-West) that maximizes or, equivalently, minimizes, the value of a certain statistical parameter. Our purpose is that this rotation is equal to that rotation which makes the anthropogenic noise completely disappear from one of the outputs, which we will show that always exists.**

***Index Terms*—Signal processing, Anthropogenic noise, Independent Component Analysis, Schumann resonances**

I. Introduction

Schumann resonances are a natural electromagnetic phenomenon that is attracting scientific interest because of its possible environmental applications. They are the resonances of the electromagnetic cavity delimited by the surface of the Earth and the lower ionosphere, such as when this cavity is excited by lightning activity the result is a modulated random electromagnetic field, appearing the resonances as maxima in the frequency spectra. The parameters of the resonances - central frequency, heigh and width- are related to the properties of this cavity, so its study gives information about atmospheric properties and evolution. These resonances belong to the Extremely Low Frequency (ELF) band of the electromagnetic spectrum, the three first resonances being within the range 3 to 30 Hz. A magnetic field measurement station was placed in Sierra Nevada, Granada (Spain) in 2013 to study these resonances. This station has two magnetometers, oriented in the North-South and the East-West directions respectively, which record the natural magnetic field with a sampling frecuency $f_s = 256$ Hz [1]. Although this station is placed in a natural park, the measurements it provides are affected by anthropogenic noise, what makes longer to analyse their frequency spectra and very difficult to study transient events in time domain. Examples of spectra recorded at the station are given in Figure 1.

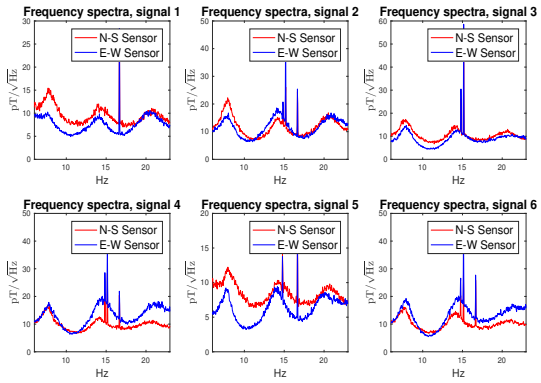

Fig. 1. Frequency spectra of six 1-hour length recordings at the ELF station in Sierra Nevada, for both North-South and East-West magnetometers. The y-axis of the plots represents the amplitud of the spectra, in units pT/$\sqrt{\text{Hz}}$ [1]. The three first Schumann resonances are seen at frequencies 7.5, 14.5 and 21.5 Hz approximately. Three narrow peaks often appear at frequencies 14.85, 15.25 and 16.67 Hz as a consequence of the anthropogenic noise in the area where the station is located.

This work has been supported by the Ministry of Education, Science and Sport of Spain through the FPU grants for PhD studentship (reference: FPU15/04291). This work has also been supported by the Ministry of Economy and Competitiveness of Spain under the project with reference FIS2017-90102-R, co-financed with FEDER funds of the EU.

ISBN 978-90-827970-1-5 © EURASIP 2018

Our purpose is to remove the anthropogenic noise in time domain. The detection and removal in time domain is faster than in frequency domain (detecting the peaks and mask them in the frequency spectra by an empirical searching algorithm), and it allows to study the response of the Earth-ionosphere cavity to transient events. One of the most known techniques that achieves it is Independent Component Analysis (ICA), which requires a series of hypothesis, as we will explain in the next section, that we cannot be sure to be fully satisfied in our measurements. The main contribution of this work is an alternative method that improves the performance of ICA in the separation of natural electromagnetic signals and noise.

In Section II we present an overview of ICA and a separation scheme based on it. In Section III we implement a separation scheme based on statistical features, and in Section IV we compare the results obtained using both methods.

II. Anthropogenic Noise Removal based on ICA

Independent Component Analysis (ICA) is a powerful tool to solve the blind source separation problem, i.e. to separate the different independent sources that contribute to a signal. One of the main uses of ICA has been for denoising, as in image [2] or EGG [3] analysis. For ICA to be able to solve the problem, the following assumptions are needed [4]:

1) The sources that contribute to the measured signal must be statistically independent.
2) The sources must have non-gaussian distributions. If just one of the independent components is gaussian, the ICA model can still be estimated.
3) The number of observed mixtures (independent measurements of the mixed signals) must be at least equal to the number of independent components.

In this work we use the FastICA algorithm [5] to perform ICA on the data recorded by the N-S and the E-W magnetometers, since it is a fast well stablished method that has been adapted to work in different situations [6] (and references therein) and applied successfully used in many real applications. It consists of a whitening process, which decorrelates both series of measurements, followed by the search of a unitary matrix that maximizes the non-gaussivity of the independent components. In our case, we have two measures in each instant, and, if $x_1(t)$ and $x_2(t)$ are the outputs after the whitening process, the unitary matrix is a 2x2 rotation (of a certain angle θ) matrix (except permutation and sign, which are intrinsic undeterminations in ICA), and the independent components $y_1(t)$ and $y_2(t)$ obtained by ICA can be written as:

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{1}$$

$$\begin{cases} y_1(t) = x_1(t)\cos\theta - x_2\sin\theta \\ y_2(t) = x_1(t)\sin\theta + x_2\cos\theta \end{cases} \tag{2}$$

Thus, the purpose of the application of ICA is to separate the natural electromagnetic field, on the one hand, and the anthropogenic noise, on the other hand, that make up our recorded signal. Due to the very different behavior of the anthropogenic noise (quasi-armonic) and the natural magnetic field, we expect that these contributions to verify statistical independence and would be identified as the independent components. In this work, the FastICA algorithm has been used. It allows to choose between four different nonlinearities [7] to find the 2x2 matrix that maximizes the nongaussivity of the independent components:

- Non-linearity 'pow3'

$$g_1(y) = y^3 \tag{3}$$

- Non-linearity 'skew'

$$g_2(y) = y^2 \tag{4}$$

- Non-linearity 'tanh'

$$g_3(y) = \tanh(a_1 * y) \tag{5}$$

- Non-linearity 'gauss'

$$g_4(y) = \exp(-a_2 * y^2/2) \tag{6}$$

The parameters a_1 and a_2 are fixed to 1 in this work. The approach selected is 'deflation', the stopping criterion is 10^{-6}, and a whitening process is performed before the maximization of the non-linearity in order to decorrelate x_1 and x_2. The FastICA algorithm will find the independent component y_1 which maximizes the selected non-linearity and, after that, the other independent y_2 will be found by orthogonality.

The results after performing FastICA with the four nonlinearities above are shown in Figure 2, where the spectra of the independent components obtained by FastICA are presented. The algorithm has been performed on 1-hour long trial signals. The plots show that none of the nonlinearities achieves a satisfactory denoising in all the signals. The reason can be that the natural electromagnetic field cannot be considered as a single source or single contribution to the measurements.

This field is generated by lightning discharges in different locations and different times. There are three main lightning nucleus in the world (South-East Asia, Central America and Central Africa) [8] whose distances and relative positions to the ELF station are different. Although the electromagnetic field near a lightning event would present magnetic measures in the S-N and E-W directions that are correlated in time, the contribution of many of this lightning events in the magnetometers, after their electromagnetic field has travelled through the atmosphere taking different paths and interacting with the earth surface and ionosphere, may produce time independent measures in the S-N and E-W sensors. In this case, the number of independent sources would be three: two natural electromagnetic contributions and the anthropogenic noise. After realising that ICA is not able to separate the noise and the natural magnetic field, in the next section we will try to transform the original measurements $x_1(t)$ and $x_2(t)$ into two channels $y_1(t)$ and $y_2(t)$, our purpose being that one of these two channels is free of noise. Then, one of

ISBN 978-90-827970-1-5 © EURASIP 2018

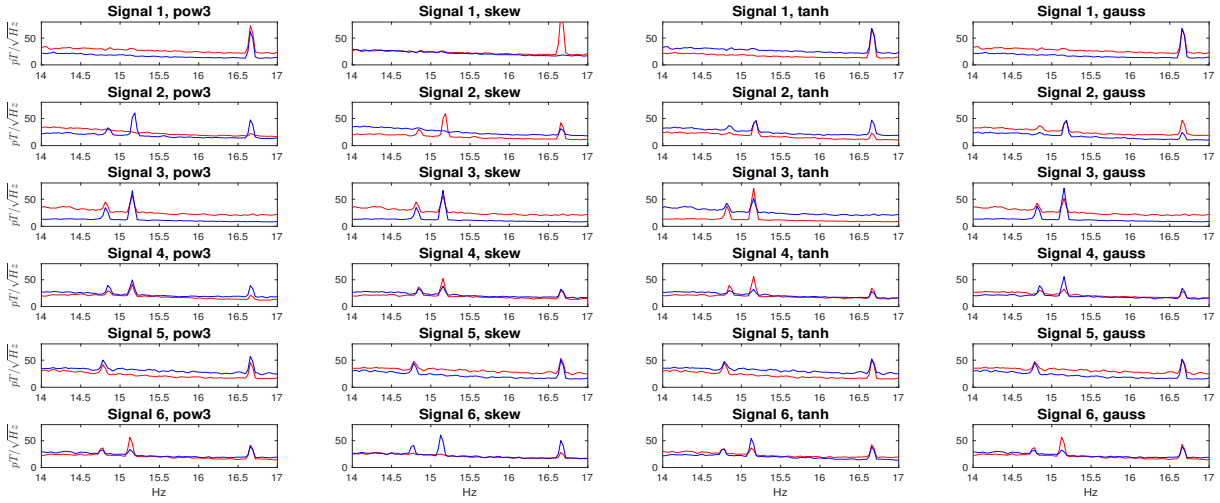

Fig. 2. Spectra of the independent components obtained with the different non-linearities of FastICA, for 6 measurements recorded at the ELF station in Sierra Nevada.

this new channels will collect both natural magnetic field and anthropogenic noise, but the other one will only collect natural magnetic field. This is what we try in the next section.

III. Anthropogenic Noise Removal Based on Statistical Features

In this section, we propose a method to isolate the anthropogenic noise in one channel based on finding the orthogonal transformation that maximizes (or equivalently minimizes) the difference between the value of a certain statistical parameter for the two transformed signals. A hypothesis that justifies this approach is that the anthropogenic noise and the natural magnetic field may have different statistical distributions, so if we maximize (or equivalently minimize) a certain parameter whose value is different for the noise and for the magnetic field, then by maximizing its value we will have performed the orthogonal transformation that points one of our axes in the direction in which anthropogenic noise field is vibrating. We can experimentally check that this direction exists for all the analyzed recordings, as we show in Figure 3.

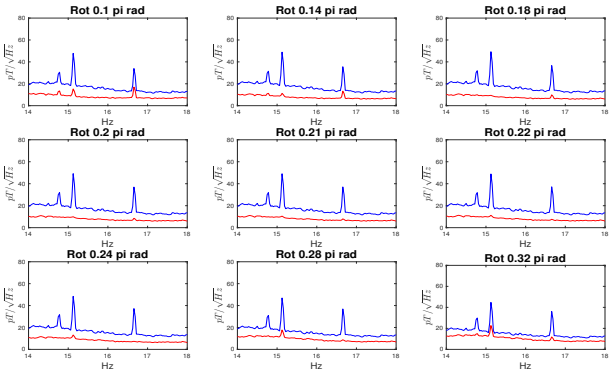

Fig. 3. Frequency spectra for the transformed signal, when orthogonal transformations corresponding to θ values within the range 0.1π to 0.32π rad are performed on signal 6. The best result is obtained for $\theta = 0.21\pi$ rad, when no peaks are observed in the spectrum of one of the components.

To try to find automatically the angle θ that gives the right orthogonal transformation to separate te noise, we are going to maximize the following parameters:

- $\zeta_1(\theta) = \xi\{y_1^3\} - \xi\{y_2^3\}$
- $\zeta_2(\theta) = \xi\{y_1^3\}^2 - \xi\{y_2^3\}^2$
- $\zeta_3(\theta) = \xi\{y_1^4\} - \xi\{y_2^4\}$

In order to improve the efficiency of the algorithm, it is possible to write the parameters ζ_i as function of θ, $x_1(t)$ and $x_2(t)$ taking into account (2). For example, for $\zeta_1(\theta)$:

$$\begin{aligned}\xi\{y_1^3\} &= S_{30}\cos^3\theta - 3S_{21}\cos^2\theta\sin\theta \\ &\quad +3S_{12}\cos\theta\sin^2\theta - S_{03}\sin^3\theta \\ \xi\{y_2^3\} &= S_{30}\sin^3\theta + 3S_{21}\sin^2\theta\cos\theta \\ &\quad +3S_{12}\sin\theta\cos^2\theta + S_{03}\cos^3\theta\end{aligned} \tag{7}$$

$$\begin{aligned}\zeta_1 = (S_{30} - S_{03})\cos^3\theta - 3(S_{21} + S_{12})\cos^2\theta\sin\theta \\ +3(S_{12} - S_{21})\cos\theta\sin^2\theta - (S_{03} + S_{30})\sin^3\theta,\end{aligned} \tag{8}$$

where

$$S_{ij} = \frac{1}{N}\sum_{k=0}^{N}(x_1(t_k))^i\,(x_2(t_k))^j, \tag{9}$$

with $x_1(t_k)$ and $x_2(t_k)$ the k-th samples recorded the magnetometers 1 (N-S) and 2 (E-W), and N the total number of samples of each series. As we work with 1-hour long signals, $N = 3600 \cdot f_s = 921600$. The methodology is exposed in the next items:

1) The measurements $x_1(t_k)$ and $x_2(t_k)$ are loaded and the sums S_{ij} are calculated, so the target function -the dependence of the parameter ζ_n we want to maximize on θ- is defined.
2) The maximum value of $\zeta_n(\theta)$ and its corresponding angle θ_0 are found.

ISBN 978-90-827970-1-5 © EURASIP 2018

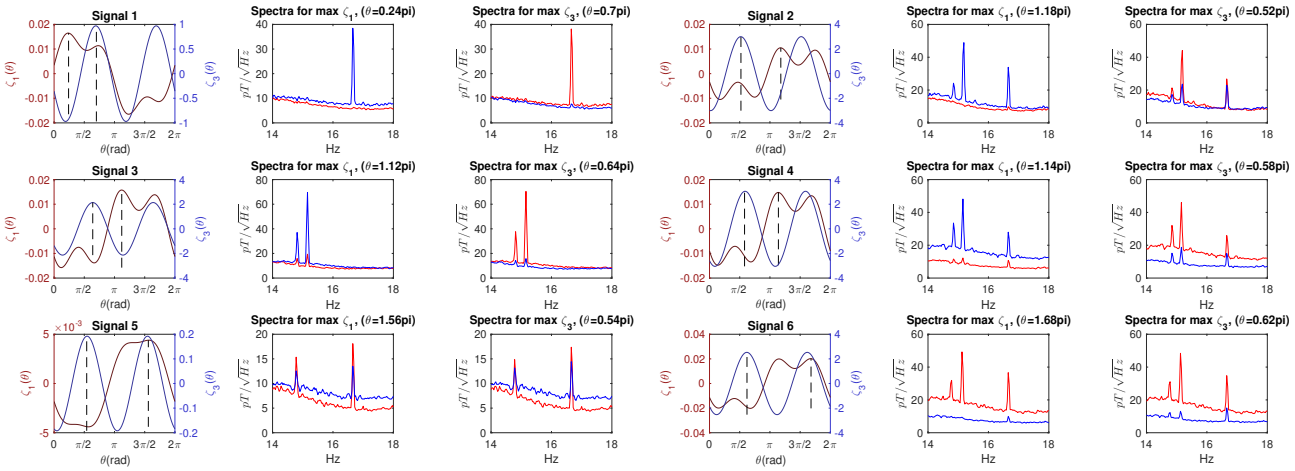

Fig. 4. Dependence of the statistical parameters ζ_1 and ζ_3 on θ with the maximum values marked, and frequency spectra after performing the orthogonal transformation that maximizes $\zeta_1\,(\theta)$ and $\zeta_3\,(\theta)$, for each one of the six signals analyzed in this contribution.

3) The transformed signals y_1 and y_2 are calculated using (2) for θ_0, and the amplitudes of their frequency spectra are plotted.

Finding the right orthogonal transformation by maximizing a 1-parameter function is by far more efficient and easier to implement than calculating a large set of spectra for different values of θ and identifying which of them show the best separation in frequency domain.

The spectra of the transformed signals, when the orthogonal transform is that which maximizes the value of the parameters ζ_1 and ζ_3, are shown in the Figure 4 together with the variation of the parameters in the range 0 to 2π rad. The spectra of the transformed signals when the orthogonal transform is that which maximizes the value of the parameter ζ_2 are shown in Figure 5 together with the variation of that parameters in the range 0 to π rad. It is equivalent to maximize and to minimize the parameters because they verify

$$\zeta_1(\theta+\pi) = -\zeta_1(\theta) \tag{10}$$

$$\zeta_2(\theta+\frac{\pi}{2}) = -\zeta_2(\theta) \qquad \zeta_3(\theta+\frac{\pi}{2}) = -\zeta_3(\theta) \tag{11}$$

and the physical meaning of a $\pi/2$ rotation is to permute y_1 and y_2 and inverse the sign of one of them.

The maximization of $\xi_2(\theta)$ clearly gives a better result than the maximization of the other two parameters. When maximizating ξ_2 one of the transformed series y_1, y_2 has a completely free of anthropogenic noise spectrum or at least an almost free of noise spectrum. The improvement with regard to ICA is also notorious.

For those measurements in which the anthropogenic noise is not separated, two reasons can cause it:

- The statistical parameters used have similar values both for the anthropogenic noise and for the magnetic field, so they are not useful to separate both contributions. Different statistical parameters could be tested to achieve separation except if the statistical distributions of both components are found to be identical.

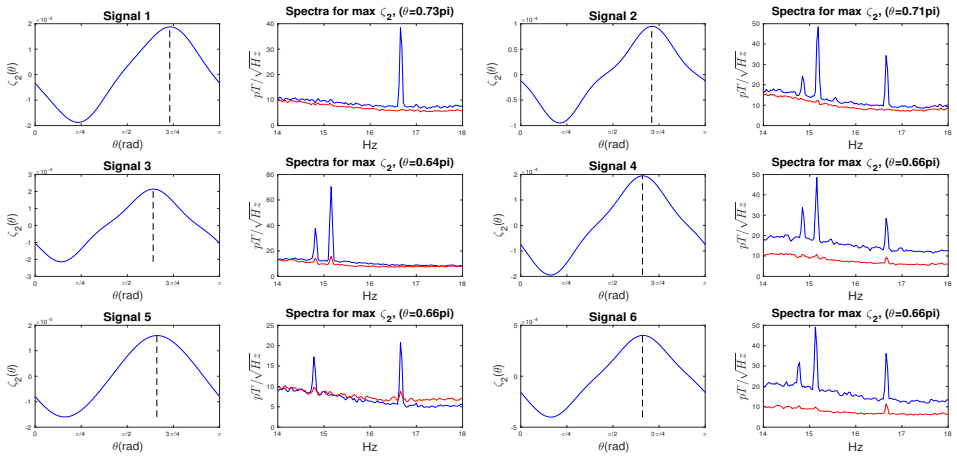

Fig. 5. Dependence of the statistical parameter ζ_2 on θ with the maximum values marked, and frequency spectra after performing the orthogonal transformation that maximizes $\zeta_2\,(\theta)$, for each one of the six signals analyzed in this contribution.

ISBN 978-90-827970-1-5

- The measurements of magnetic field recorded in the N-S and the E-W sensors have a difference in their parameters $\xi\{x_1^3\}$ and $\xi\{x_2^3\}$, what conditionates the dependence of ζ_2 on θ and masks the variation in this parameter due to the the anthropogenic noise.

IV. Comparison of the performance of ICA and the method based on statistical features

In this section we apply both ICA and the maximization of $\zeta_1(\theta)$, $\zeta_2(\theta)$ and $\zeta_3(\theta)$ to a set of 1-hour length measurements corresponding to the complete month of April, 2014. We have a total number of 720 1-hour length measurements recorded in each sensor for this month. The purpose in this section is to evaluate in how many cases ICA and the maximization of ζ_1, ζ_2 and ζ_3 succeed at providing the right orthogonal transformation that gives us one component free of noise. We follow these steps:

1) For each 1-hour length measurement, x_1 and x_2, we plot its N-S and E-W spectra and calculate the number of peaks of anthropogenic noise and its central frequencies, f_k. If there are no peaks, the measurement is classified as noisefree and the steps 2 and 3 are not performed.
2) We calculate the seven sets of transformed components y_1 and y_2 that respectively maximize the four non-linearities of ICA and the parameters ζ_1, ζ_2 and ζ_3.
3) For each of the seven sets of measuments, the next parameter is calculated

$$a = \min\left\{\sum_{k=1}^{N_l}\left(Y_1(f_k) - M_{1k}\right)^2, \sum_{k=1}^{N_l}\left(Y_2(f_k) - M_{2k}\right)^2\right\} \tag{12}$$

where N_l is the number of peaks of the frequency spectrum of y_l, $Y_l(f_k)$ is the amplitude of the frequency spectrum of y_l in the central peak frequency f_k, and M_{lk} is the mean of the frequency spectrum amplitude of y_l for two points inmediately out of the peak in f_k, one on the left side of the peak and the other one on the right side.

As an example, the values of a found for the 6 signals shown in Figures 1 to 5 are in Table 1.

TABLE I
Parameter a for each method and each signal

	pow3	**skew**	**tanh**	**gauss**	ζ_1	ζ_2	ζ_3
Signal 1	2500	1.5	2089	2163	0.08	0.08	2.4
Signal 2	7.7	91	317	37	1.4	1.4	125
Signal 3	595	573	410	405	61	27.0	27.0
Signal 4	205	185	166	97	11.3	4.4	65
Signal 5	557	505	472	460	27.3	3.5	35
Signal 6	186	40	251	168	3.7	8.9	32

It can be accepted that the value of a reflects the amount of anthropogenic noise collected in the less noisy component according to the visual information in Figures 2, 4 and 5. We resolve the next criterion based on the parameter a:

- If $a \leq 3$, then we can accept that one of the components is free of noise.
- If $3 < a \leq 10$, then we can accept that one of the components has very little noise.

The number of 1-hour long measurements free of noise for the month studied has been 37 out of 720. For the remaining 683 measurements, the number of measurements in which each method has succeded at completely or almost completely separating the anthropogenic noise is shown in Table 2.

TABLE II
Number of measurements denoised by each method

	pow3	**skew**	**tanh**	**gauss**	ζ_1	ζ_2	ζ_3
$a \leq 3$	35	139	4	3	285	255	110
$3 < a \leq 10$	27	75	2	2	115	110	99

In this case, the parameter which shows the best behavior is ζ_1, with a 42% of complete noise removal and a 17% of almost complete noise removal. The second best parameter is ζ_2, with a 37% of complete noise removal and a 16% of almost complete noise removal. Both perform much better than ICA, which when using the non-linearity 'skew' shows 20% and 11% success at completely and almost completely removing the noise from one component.

V. Conclusions

ICA does not achieve a separation of anthropogenic noise and the natural magnetic field in most of our measurements. The maximization of the statistical features exposed gives an acceptable separation in more cases than ICA. Among the statistical features tried, the parameters $\zeta_1 = \xi\{y_1^3\} - \xi\{y_2^3\}$ and $\zeta_2 = \xi\{y_1^3\}^2 - \xi\{y_2^3\}^2$ are those whose maximization achieves an acceptable separation in most of the cases.

This alternative separation scheme is valid when we dispose of signals coming from very different sources (from a statistical point of view), such as harmonic signals superposed to natural electromagnetic background.

References

[1] J. Fornieles-Callejon, A. Salinas, S. Toledo-Redondo, J. Porti, A. Mendez, E. A. Navarro, J. A. Morente-Molinera, C. Soto-Aranaz, and J. S. Ortega-Cayuela, "Extremely low frequency band station for natural electromagnetic noise measurement," *Radio Science*, vol. 50, no. 3, pp. 191–201, March 2015.

[2] K. Liang and J. Ye, "Ica-based image denoising: A comparative analysis of four classical algorithms," in *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)(*, March 2017, pp. 709–713.

[3] L. Albera, A. Kachenoura, P. Comon, A. Karfoul, F. Wendling, L. Senhadji, and I. Merlet, "Ica-based eeg denoising: a comparative analysis of fifteen methods," *Bulletin of the Polish Academy of Sciences: Technical Sciences*, vol. 60, no. 3, pp. 407–418, 2012.

[4] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent component analysis*. John Wiley & Sons, 2004, vol. 46.

[5] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural networks*, vol. 13, no. 4-5, pp. 411–430, 2000.

[6] S. Javidi, C. C. Took, and D. P. Mandic, "Fast independent component analysis algorithm for quaternion valued signals," *IEEE Transactions on Neural Networks*, vol. 22, no. 12, pp. 1967–1978, Dec 2011.

[7] A. Dermoune and T. Wei, "Fastica algorithm: Five criteria for the optimal choice of the nonlinearity function," *IEEE Transactions on Signal Processing*, vol. 61, no. 8, pp. 2078–2087, April 2013.

[8] R. L. Holle, "Some aspects of global lightning impacts," in *2014 International Conference on Lightning Protection (ICLP)*, Oct 2014, pp. 1390–1395.

ISBN 978-90-827970-1-5