



The next generation radiotelescopes / Les radiotélescopes du futur

## Cyclostationary approaches for spatial RFI mitigation in radio astronomy

### *Approches spatiales cyclostationnaires pour le traitement des interférences en radioastronomie*

Grégory Hellbourg<sup>a,b</sup>, Rodolphe Weber<sup>a,c,\*</sup>, Cécile Capdessus<sup>c</sup>, Albert-Jan Boonstra<sup>b</sup>

<sup>a</sup> Observatoire de Paris, Station de radioastronomie de Nançay, 18330 Nançay, France

<sup>b</sup> ASTRON, P.O. Box 2, 7990 AA Dwingeloo, the Netherlands

<sup>c</sup> Laboratoire PRISME – université d'Orléans, site Galilée, 12 rue de Blois, 45067 Orléans cedex 2, France

#### ARTICLE INFO

##### Article history:

Available online 11 December 2011

##### Keywords:

Cyclostationarity  
RFI mitigation  
Phased array antenna

##### Mots-clés:

Cyclostationnarité  
Traitement d'interférences  
Réseaux d'antennes phasées

#### ABSTRACT

Radio astronomical observations are increasingly corrupted by radio frequency interferences (RFIs), and real time filtering algorithms are becoming essential. In this article, it is shown how spatial processing techniques can limit the impact of the incoming RFIs for phased array radio telescopes. The proposed approaches are based on estimation of the RFI spatial signature. It requires the diagonalization of either the classic correlation matrix or the cyclic correlation matrix of the array. Different diagonalization techniques are compared. Then, RFI detection and RFI filtering techniques are illustrated through simulations on data acquired with the Low Frequency Array Radio telescope, LOFAR. The originality of the study is the use of the cyclostationarity property, in order to improve the spatial separation between cosmic sources and RFIs.

© 2011 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

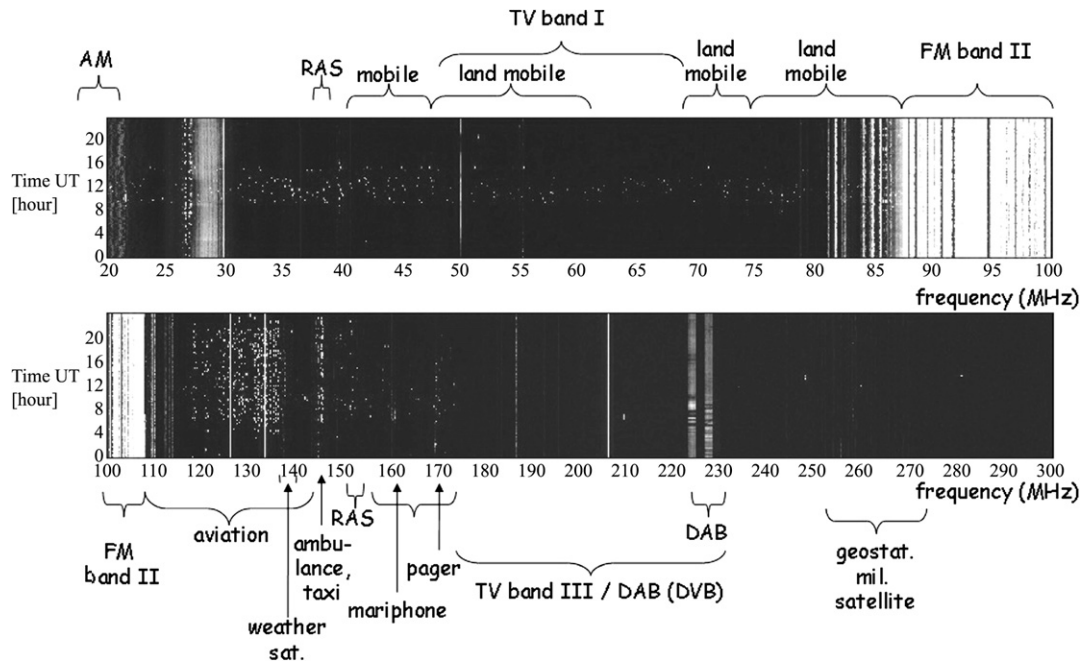
#### RÉSUMÉ

L'environnement électromagnétique est de plus en plus exploité par les activités de télécommunications, limitant ainsi la sensibilité et les performances des radiotélescopes. Il devient donc nécessaire de développer des traitements efficaces afin de limiter l'impact de ces interférences radioélectriques (RFI). Les approches proposées reposent sur l'estimation de la signature spatiale de l'interférence grâce à la diagonalisation de la matrice de corrélation classique ou cyclique. Dans cet article, différentes techniques de diagonalisation de ces matrices sont comparées. Ensuite, des techniques de détection ou de filtrage des interférences sont illustrées sur des observations acquises avec le radiotélescope LOFAR. L'originalité de l'étude repose sur l'utilisation des caractéristiques cyclostationnaires intrinsèques aux RFIs.

© 2011 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

\* Corresponding author at: Laboratoire PRISME – université d'Orléans, site Galilée, 12 rue de Blois, 45067 Orléans cedex 2, France.

E-mail addresses: [gregory.hellbourg@obs-nancay.fr](mailto:gregory.hellbourg@obs-nancay.fr) (G. Hellbourg), [rodolphe.weber@univ-orleans.fr](mailto:rodolphe.weber@univ-orleans.fr) (R. Weber), [cecile.capdessus@univ-orleans.fr](mailto:cecile.capdessus@univ-orleans.fr) (C. Capdessus), [boonstra@astron.nl](mailto:boonstra@astron.nl) (A.-J. Boonstra).



**Fig. 1.** Electromagnetic environment of the LOFAR radio telescope (20–300 MHz band). Some of the empty bands are progressively occupied by digital radio (DAB) and television (TV) transmissions (DAB stands for Digital Audio Broadcast and DVB stands for Digital Video Broadcast). RAS stands for Radio Astronomy protected band, AM and FM stand respectively for Amplitude Modulation and Frequency Modulation (radio transmission).

## 1. Introduction

Radio astronomy was born in the nineteen thirties and its growth followed that of the related technological fields, such as electromagnetism, electronics and computer science. Based on these same technological advances, telecommunications also developed extensively and became perturbation for sky observation. The radio observatories located close to populated areas are of course affected by many man-made Radio Frequency Interferences (RFIs). But even in the most remote locations, they still can be reached by satellite communications and positioning systems, so that the choice of an adequate location, though still a major issue, cannot ensure immunity to electromagnetic pollution (Fig. 1 shows an example of the potential electromagnetic activity that can be expected in the radio telescope LOFAR band). Furthermore, the sensitivity of radio telescopes has improved drastically, thanks to a wider reception surfaces and electronic devices that generate less noise, thus allowing faster sampling, increased data storage and averaging. For example, the Square Kilometre Array telescope (SKA), which is presently under study, should be 50 times more sensitive than any other radio instrument. The drawback of such sensitivity is that interferences that were once drowned in the noise now become prevailing. Some frequency bands have been dedicated to radio observation (see RAS indications in Fig. 1), based on the main wavelength produced by celestial bodies. However, the redshift effect alter these wavelengths and make the recorded signals of distant sources falling into an unprotected frequency band. Furthermore, the allocated frequency bands were chosen based on immediate knowledge of celestial bodies, and exploring the other frequency bands may lead to discoveries for radio astronomers.

So far, observation strategies have taken advantage of the sparse nature of telecommunication signals. First, their sparsity in the frequency domain allows sky observations in sub-bands of the allocated frequency bands. Second, most communication systems emit intermittently, so that corrupted parts of the signal can be eradicated and sky observation performed on the remaining ones. Unfortunately telecommunication companies also intend to make the most of that sparsity, for instance through Ultra Wide Band (UWB) communication systems or cognitive radio [1]. The development of automated RFI detection and mitigation techniques has thus become a necessity. The techniques developed so far have taken advantage of both the radio astronomy specificities and the RFI features [2]. However, new generation telescopes, such as LOFAR or SKA, are multisensor devices, or groups of such devices located in distant places, allowing interferometry and the construction of sky maps. They are associated to real-time correlators, and allow spatial processing, such as beamforming. RFI detection or mitigation algorithms can be based on either pre-correlation or post-correlation methods, and can take advantage of the fact that correlation matrices are estimated for the needs of radio astronomy. Indeed, it has been shown that under some assumptions on the noise features and assuming that the antenna is calibrated, diagonalization of the correlation matrix makes it possible to estimate the number of interferences and to eradicate them by spatial filtering [3,4]. These techniques can be improved by taking into account the cyclostationarity of RFI signals, since all communication signals exhibit cyclostationarity [5], tied either to their carrier frequency or to their baud rate. These characteristics can be a priori known from

lists of identified RFI or estimated from the recorded signals. Some cyclostationary detectors have already been proposed for single dish radio telescopes [6,7].

We propose here new detection and mitigation techniques based on the diagonalization of cyclic correlation matrices at given cyclic frequencies corresponding to the RFI to be eradicated. These new algorithms do not require prior antenna calibration nor do they assume any special features for the noise. Furthermore, slight modifications to the existing correlators would allow the estimation of cyclic correlation matrices for real-time implementation of the proposed algorithms. Preliminary work has already been published in [8]. In Section 2, the problem formulation will be addressed and the correlation and cyclic correlation matrices will be defined. In Section 3, several diagonalization techniques will be presented and evaluated. In Section 4, some RFI detection or mitigation algorithms will be described and applied to simulations as well as real-life signals. Section 5 presents the conclusion and some perspectives.

## 2. Data model

Consider a phased array radio telescope consisting of  $M$  antennas, each having a received signal  $z_k(t)$ ,  $k = 1, \dots, M$ . It is assumed that the narrowband condition holds and that the geometric delay for each antenna and each impinging source can be represented by a phase shift. In this case, the telescope output vector  $\mathbf{z}(t)$  can be modeled in complex baseband form as:

$$\mathbf{z}(t) = \mathbf{a}_r r(t) + \mathbf{A}_s \mathbf{s}(t) + \mathbf{n}(t) \quad (1)$$

where:

- $\mathbf{z}(t) = [z_1(t) \cdots z_M(t)]^T$  is the  $M \times 1$  vector of radio telescope observations at time  $t$ .  $(\cdot)^T$  is the transpose operator.
- $\mathbf{a}_r = [a_{1,r} \cdots a_{M,r}]^T$  is the spatial signature of the RFI signal,  $r(t)$ .
- $\mathbf{s}(t) = [s_1(t) \cdots s_K(t)]^T$  is the  $K \times 1$  vector of the  $K$  white Gaussian independent cosmic source signals.
- $\mathbf{A}_s = [\mathbf{a}_{s_1}, \mathbf{a}_{s_2}, \dots, \mathbf{a}_{s_K}]$  is an  $M \times K$  matrix where each  $\mathbf{a}_{s_k} = [a_{1,s_k} \cdots a_{M,s_k}]^T$  is the spatial signature of the corresponding  $k$ th cosmic source.
- $\mathbf{n}(t)$  is the  $M \times 1$  system noise vector with independent Gaussian entries. There is no specific assumption made concerning the whitening of the covariance matrix and calibration of the data (i.e.  $\langle \mathbf{n}(t) \mathbf{n}^H(t) \rangle_\infty \neq \sigma^2 \mathbf{I}$ , where  $\langle \cdot \rangle_\infty$  is the average operator,  $(\cdot)^H$  is the conjugate transpose operator,  $\mathbf{I}$  is the  $M \times M$  identity matrix and  $\sigma^2$  is the noise power). In particular, if there is no coupling between antennas, this noise covariance matrix becomes diagonal,  $\langle \mathbf{n}(t) \mathbf{n}^H(t) \rangle_\infty = \text{diag}(\sigma_k^2, k = 1, \dots, M)$  with  $\sigma_k^2$  the noise power on the  $k$ th antenna.

At this stage, the phased array is not calibrated. Thus, all spatial signatures are multiplied by unknown complex gain matrices which depend on ionosphere, antennas and system configuration [9]. The calibration procedure consists in estimating these gain matrices and the noise covariance matrix mentioned above. Since these estimations have to be based on RFI-free signals, RFI detection must be done prior to any calibration procedure. Consequently, our data model will consider RFI spatial signatures as random phase vectors without any predictable structure.

However, even with a calibrated array, there are two other reasons for considering RFI spatial signatures as random phase vectors. One reason is that RFI may enter through the side lobes which are not well characterized and the other is that the far field hypothesis may not always be valid for RFI signals.

### 2.1. Classical correlation matrix

In radio astronomy, sky images are obtained through the correlation matrices measured by phased array radio telescopes. All the spatial and statistical information on the cosmic sources are contained in these matrices.

Let us define  $\mathbf{R}_z$ , the classic correlation matrix of the phased array. With our previous signal model,  $\mathbf{R}_z$  is expressed by:

$$\mathbf{R}_z = \langle \mathbf{z}(t) \mathbf{z}^H(t) \rangle_\infty = \underbrace{\mathbf{a}_r \mathbf{a}_r^H}_{\mathbf{R}_{RFI}} \sigma_r^2 + \underbrace{\mathbf{A}_s \mathbf{R}_s \mathbf{A}_s^H}_{\mathbf{R}_{Sources}} + \underbrace{\mathbf{N}}_{\mathbf{R}_{Noise}} \quad (2)$$

where  $\sigma_r^2$  is the RFI power,  $\mathbf{R}_s$  is the correlation matrix ( $K \times K$ ) of the  $K$  cosmic sources, and  $\mathbf{N}$  the correlation matrix ( $M \times M$ ) of the system noise.

One classic simplification of this model is to consider that cosmic sources are negligible. This simplification is easily accepted since their power is usually very small compared to RFI power.

### 2.2. Cyclostationary approach

Most telecommunications signals present a hidden periodicity due to the periodic characteristics involved in the signal construction (carrier frequency, baud rate, coding scheme, ...). These parameters are usually scrambled and hidden by the randomness of the message to be transmitted. However, by using a cyclostationary approach, this hidden periodicity can be

recovered thus making identification of the telecommunication signal possible. An exhaustive overview of cyclostationarity theory and applications can be found in [10,11]. For array signal processing, cyclostationarity was introduced in [5]. In the present analysis, the correlation matrix,  $\mathbf{R}_z$ , is replaced by the cyclic correlation matrix:

$$\mathbf{R}_z^\alpha = \langle \mathbf{z}(t)\mathbf{z}^H(t) \exp(-j2\pi\alpha t) \rangle_\infty \quad (3)$$

where  $\alpha$  is the cyclic frequency. This parameter is linked to the above-mentioned periodic characteristics. Similarly, we can define the cyclic conjugated correlation matrix,  $\overline{\mathbf{R}}_z^\alpha$ , by replacing the operator  $(\cdot)^H$  by  $(\cdot)^T$  in Eq. (3). Then, another set of cyclic frequencies can be considered. For example, let us consider a binary phase shift keying (BPSK) modulation scheme with a baud rate,  $\frac{1}{T_0}$ , and a carrier frequency,  $f_0$ :

$$r(t) = \exp(j2\pi f_0 t + j\varphi_0) \cdot \sum_{k \in \mathbb{Z}} c_k \cdot h(t - kT_0 - t_0) \quad (4)$$

where  $c_k$  is the binary random sequence,  $t_0$  and  $\varphi_0$  are unknown synchronization parameters and  $h(t)$  is the pulse shaping function. The cyclic correlation of  $r(t)$  is non-zero for the cyclic frequencies  $\alpha = \frac{k}{T_0}$ ,  $k \in \mathbb{Z}^*$ , and the cyclic conjugated correlation is non-zero for the cyclic frequencies  $\alpha = 2f_0 + \frac{k}{T_0}$ ,  $k \in \mathbb{Z}$ .

The main idea behind these cyclic operators is to find some periodicities inside the second order statistics of the signal. Eq. (3) can be seen as the Fourier transform of the instantaneous signal power. One can also remark that by taking  $\alpha = 0$ , we retrieve the expression of the classical correlation matrix.

Any signal that is cyclostationary at frequency  $\alpha$  will generate non-zero cyclic or cyclic conjugated correlation matrices. Inversely, any stationary signal or cyclostationary signal with different cyclic frequencies will generate zero in Eq. (3).

Thus, the cyclic correlation matrix at  $\alpha_0$  for the model defined by Eq. (1) becomes:

$$\mathbf{R}_z^{\alpha_0} = \mathbf{a}_r \mathbf{a}_r^H R_r^{\alpha_0} + \underbrace{\mathbf{A}_s \mathbf{R}_s^{\alpha_0} \mathbf{A}_s^H}_{\rightarrow 0} + \underbrace{\mathbf{R}_n^{\alpha_0}}_{\rightarrow 0} \quad (5)$$

where  $\alpha_0$  is the expected RFI cyclic frequency and  $R_r^{\alpha_0}$  is the mono-dimensional version of Eq. (3) (i.e.  $\mathbf{z}$  becomes a scalar). To simplify the notations, RFIs with different cyclic frequencies are considered as source signals and are merged into  $R_s^{\alpha_0}$ . A similar asymptotic expression can be obtained for the cyclic conjugated correlation matrix.

Eq. (5) is expected to give better estimates of the RFI spatial signature than Eq. (2) due to the fact that all non-cyclostationary terms asymptotically vanish. In the following section, we will compare different approaches to obtain the right RFI spatial signature from the classic and cyclic correlation matrices of the observations.

### 3. Spatial signature vector estimation

The main point of any spatial signal processing is the estimation of source signature vectors. In this part, we will consider that we are in the presence of one RFI and  $K$  cosmic sources. We will present different methods allowing RFI spatial signature vector estimation (i.e. estimation of  $\mathbf{a}_r$ ). These methods are based on the use of the classic and cyclic correlation matrices.

#### 3.1. Classic eigenvalue decomposition

The classic correlation matrix expressed in Eq. (2) can be decomposed into eigenvalues:

$$\mathbf{R}_z = \mathbf{a}_r \mathbf{a}_r^H \sigma_r^2 + \mathbf{A}_s \mathbf{R}_s \mathbf{A}_s^H + \mathbf{N} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H \quad (6)$$

where  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_M)$  contains the eigenvalues of  $\mathbf{R}_z$  sorted in decreasing order and  $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_M]$  contains the corresponding eigenvectors. The eigenvalues of this decomposition give the power matched by a beamformer pointing in the direction of the related eigenvector. By assuming that the system noise is independent and identically distributed over the antennas (i.e.  $\mathbf{N} = \sigma^2 \mathbf{I}$ ), we can split the orthonormal basis made by the eigenvectors into two sub-bases generating two different subspaces: a signal subspace ( $(K+1)$  dominant eigenvalues) and a noise subspace [12,13]. If cosmic sources are neglected, the signal subspace (reduced to a one-dimensional vector space) generates exactly the same vector space as  $\mathbf{a}_r$  (i.e.  $\mathbf{u}_{\max} \propto \mathbf{a}_r$ , with  $\mathbf{u}_{\max}$  the eigenvector associated to the maximum eigenvalue).

#### 3.2. Cyclic singular value decomposition

As shown in Eq. (5), cyclic correlation matrices calculated for a cyclic frequency  $\alpha_0$  are asymptotically independent of cosmic sources and RFIs with a cyclic frequency different from  $\alpha_0$ . However, an eigenvalue decomposition of this type of matrix is not possible, since they are non-hermitian. The following singular value decomposition (SVD) is defined:

$$\mathbf{R}_z^{\alpha_0} = \mathbf{U}_{\text{cyclic}} \mathbf{\Lambda}_{\text{cyclic}} \mathbf{V}_{\text{cyclic}}^H \quad (7)$$

with  $\mathbf{A}_{cyclic}$  the singular value matrix of  $\mathbf{R}_z^{\alpha_0}$ ,  $\mathbf{U}_{cyclic}$  the input basis vector matrix and  $\mathbf{V}_{cyclic}$  the output basis vector matrix. By identification with Eq. (5), we find  $\lambda_1 = R_r^{\alpha_0}$ , with  $\lambda_1$  the only non-null singular value of  $\mathbf{A}_{cyclic}$ , and  $\mathbf{u}_1 = \mathbf{a}_r$ , with  $\mathbf{u}_1$  the singular vector associated with  $\lambda_1$ . This technique thus makes it possible to estimate the RFI spatial signature vector. Furthermore, no assumption has been made on the noise correlation matrix, nor on the presence of other sources. The cyclostationary approach is therefore an interesting technique to mitigate RFIs prior to antenna array calibration.

### 3.3. Joint diagonalization of multiple cyclic correlation matrices

The basic idea of joint diagonalization is to find a common transformation matrix that allows the diagonalization of a set of matrices. Most cyclostationary interferences present more than one cyclic frequency. In order to provide a better estimate of the RFI signature vector, we suggest applying a joint diagonalization to cyclic matrices calculated at different cyclic frequencies, since all these matrices contain the same spatial information. Finding a joint diagonalizer for these matrices should therefore decrease the error made by a single matrix diagonalization due to the finite length correlation estimation.

In [14], Belouchrani et al. introduce the following *off* function for an  $n \times n$  matrix  $\mathbf{M}$  with entries  $M_{ij}$ :

$$off(\mathbf{M}) = \sum_{1 \leq i \neq j \leq n} |M_{ij}|^2 \quad (8)$$

By using this function, diagonalizing a matrix consists in finding a matrix  $\mathbf{P}$  such that:

$$off(\mathbf{P}^H \mathbf{M} \mathbf{P}) = 0 \quad (9)$$

The aim of joint diagonalization of a set of  $N$  matrices  $\mathcal{M} = \{\mathbf{M}_1, \mathbf{M}_2, \dots, \mathbf{M}_N\}$  is thus to find a matrix  $\mathbf{P}$  that minimizes the following cost function:

$$c(\mathcal{M}, \mathbf{P}) = \sum_{1, \dots, N} off(\mathbf{P}^H \mathbf{M}_k \mathbf{P}) \quad (10)$$

Let  $\mathcal{M}$  be a set of cyclic correlation matrices  $\mathbf{R}_z^\alpha$  calculated over several cyclic and conjugated cyclic frequencies. As seen in Section 3.2, the  $\mathbf{R}_z^\alpha$  matrices do not admit an eigenvalue decomposition (EVD). In order to fit to the model of Eq. (10), the following property can be used:

$$\mathbf{R}_z^\alpha \cdot \mathbf{R}_z^{\alpha H} = (\mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V}^H) \cdot (\mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V}^H)^H = \mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V}^H \cdot \mathbf{V} \cdot \mathbf{S}^H \cdot \mathbf{U}^H = \mathbf{U} \cdot \mathbf{S} \cdot \mathbf{S}^H \cdot \mathbf{U}^H = \mathbf{P} \cdot \mathbf{D} \cdot \mathbf{P}^H \quad (11)$$

with  $\mathbf{R}_z^\alpha = (\mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V}^H)$  (SVD of  $\mathbf{R}_z^\alpha$ ),  $\mathbf{P} = \mathbf{U}$  and  $\mathbf{D} = \mathbf{S} \cdot \mathbf{S}^H$ . By applying this property to all the matrices stored in  $\mathcal{M}$ , the RFI spatial signature vector can be estimated by finding a joint diagonalizer of  $\mathcal{M}$ .

The main advantage of this method is the use of multiple correlation matrices, thereby improving the estimation of the RFI signature vector. Although this advantage is important for the estimation accuracy, the implementation cost of this method is very high. The next section will present a more simple multiple matrix technique.

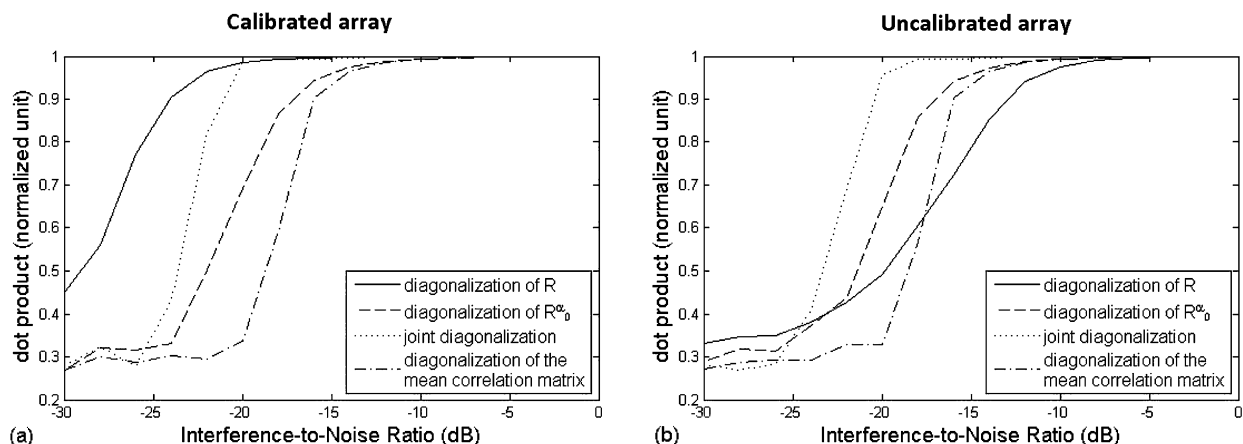
### 3.4. Mean correlation matrix diagonalization

The principle of this technique is to calculate a mean matrix over all the cyclic and conjugated cyclic correlation matrices. As all these correlation matrices are biased by the added system noise and by the signal finite length, computing a mean matrix will improve correlation matrix estimation.

Since all these matrices are defined on the same vector space (observation space), diagonalization of the mean matrix should provide a better estimate of the RFI spatial signature vector. Two options can be considered: applying an SVD of a mean cyclic correlation matrix or an EVD of a mean matrix calculated over cyclic correlation matrices after applying the transformation presented in Eq. (11) on each of them. It is however important to normalize matrices before any average calculation step, since cyclic power is not uniformly distributed over all the cyclic frequencies of a cyclostationary signal. We therefore divide each cyclic correlation matrix (or transformed cyclic correlation matrix) by its squared Frobenius Norm defined by the sum of the squared moduli of its elements. This matrix norm is linked to singular values of the matrix on which it is calculated by the following expression:

$$\|\mathbf{M}\|_F^2 = \sum_{k=1}^n \lambda_k^2 \quad (12)$$

with  $\|\cdot\|_F$  the Frobenius Norm operator,  $\mathbf{M}$  an  $(n \times n)$  matrix and  $\lambda_k$  the  $k$ th singular value of  $\mathbf{M}$ .



**Fig. 2.** Performance comparison of signature vector estimation techniques (dot product between estimated and generated signature vectors for different techniques and different INR, its value is 1 for an exact estimation). (a) 10-antenna calibrated array. (b) 10-antenna uncalibrated array (20% noise power fluctuation over antennas).

### 3.5. Performance analysis

To compare the performance of the different signature vector estimation techniques, we ran a simulation involving a Binary Phase-Shift Keying (BPSK) RFI (see Eq. (4)) added to white Gaussian noise (carrier frequency = 0.1 (normalized frequency) and baud rate =  $1/12$  samples $^{-1}$ ). An array composed of  $M = 10$  antennas was used. The array noise was calibrated in a first scenario, and uncalibrated with 20% noise power fluctuations over antennas in a second scenario. During each run, we generated a new data set (32 768 samples  $\times$  10 antennas), estimated the signature vector of the impinging RFI and calculated the dot product between this estimate and the actual steering vector. In other words, we will use the dot product as a way to measure how our estimation matches the true signature vector. The Interference-to-Noise Ratio (INR) of the data was varied from  $-30$  to  $0$  dB. All our results were averaged over 100 runs for a fixed INR to provide statistics on signature vector estimation technique performances.

We calculate for each run a cyclic correlation matrix for  $\alpha_0 = 1/T_0$ , where  $T_0$  is the symbol duration of the interference, and three conjugated cyclic matrices at respectively  $\alpha_1 = 2 \cdot f_0 - 1/T_0$ ,  $\alpha_2 = 2 \cdot f_0$  and  $\alpha_3 = 2 \cdot f_0 + 1/T_0$ , where  $f_0$  is the carrier frequency of the interference.

Fig. 2 depicts the performance comparison of the RFI spatial signature vector estimation techniques: diagonalization of the correlation matrix  $\mathbf{R}$ , diagonalization of the cyclic correlation matrix  $\mathbf{R}^\alpha$  ( $\alpha_0 = 1/T_c$ ), joint diagonalization of cyclic and conjugated cyclic matrices ( $\alpha_0 = 1/T_c$ ,  $\alpha_1 = 2f_0 - 1/T_c$ ,  $\alpha_2 = 2f_0$ ,  $\alpha_3 = 2f_0 + 1/T_c$ ) and diagonalization of the mean cyclic matrix. It can be seen that these techniques react differently depending on whether the array noise is calibrated or not. Diagonalization of the classic correlation matrix gives quite good results, even for low INR, when the array is calibrated. However, its performance is severely impaired with an uncalibrated array noise. One very interesting point to notice with these plots is the steadiness of cyclic correlation matrices diagonalization, which remains completely independent of noise calibration (and of other sources, whether cyclostationary or not). The technique consisting in diagonalizing the mean correlation matrix and the joint diagonalization technique both present good performances with respect to the calibration hypothesis. This stems from the fact that they only involve cyclic and conjugated cyclic correlation matrices. Joint diagonalization gives the best result in all cases, but it suffers from a complex implementation (optimization algorithm).

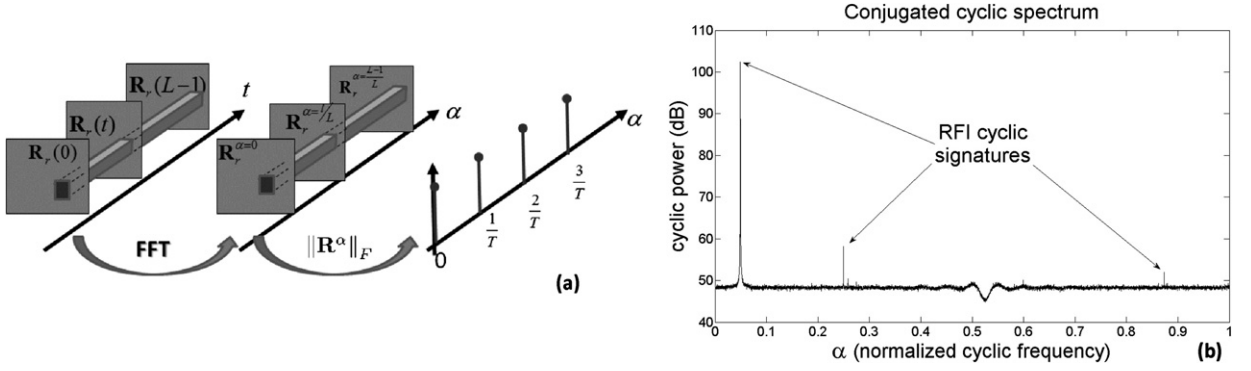
## 4. Application to LOFAR data

This section is dedicated to practical applications for the radio telescope LOFAR. After introducing this antenna array telescope, we will present the cyclostationary techniques used to detect and filter out RFIs from astronomical signals.

### 4.1. Description of LOFAR

LOFAR, LOw Frequency ARray [15], is a radio telescope built by the Netherlands Foundation for Research in Astronomy, ASTRON, and uses the principle of interferometry. Its main applications are low frequency astronomy ( $\approx 30$ – $240$  MHz) [16].

LOFAR is an antenna array distributed over 1000 km within the Netherlands, with extensions to other countries (Great Britain, France, Germany, Sweden). The antennas, which are low-cost and omnidirectional, are arranged in clusters, or stations. They are used as phased arrays, allowing electronic steering of the direction of observation by applying phase delays between them. All the data collected by LOFAR (several gigabits per second per station) are processed in the Netherlands, at the University of Groningen.



**Fig. 3.** Cyclic detection principle. (a) Cyclic features detection algorithm. (b) Cyclic spectrum calculated on a LOFAR data set ( $M = 10$  antennas and  $L = 2^{15}$  samples) corrupted by an aviation signal at 128.6 MHz. Spectral lines correspond to cyclostationary signatures.

#### 4.2. Detection approach

The cyclostationarity property of RFIs can be used in order to detect a known or unknown interference in a dataset. The detector principle is based on the estimation of the cyclic spectrum of an antenna array radio telescope data set.

A data cube made up of instantaneous correlation matrices is calculated from the antenna network signals. If one or more RFIs are mixed with the data, these data become cyclostationary. This data cube will thus exhibit a periodicity. To bring it out, a Fourier transform is applied (see Fig. 3(a)). This produces a new data cube, made up of cyclic correlation matrices  $\mathbf{R}_z^\alpha$ , with  $\alpha = k/L$ ,  $k \in \mathbb{Z}^*$  and  $L$  the Fourier transform length. Asymptotically,  $\mathbf{R}_z^\alpha$  should have a rank of 1 according to our model, and  $\|\mathbf{R}_z^\alpha\|_F^2 = \mathbf{R}_z^\alpha$  (see Eq. (12)). Asymptotically, a single non-null value of  $\|\mathbf{R}_z^\alpha\|_F^2$  should then allow us to conclude on the presence of at least one  $\alpha$ -cyclostationary RFI.

By calculating the Frobenius norm of each of these matrices, we obtain a cyclic spectrum which highlights cyclic frequencies, and therefore a cyclostationary blind detection criterion.

However, since we are working on finite length signals,  $\mathbf{R}_z^\alpha$  is full rank (i.e.  $\lambda_k \neq 0, \forall k$ ). RFI detection then becomes less efficient.

Fig. 3 shows the detector principle, and an example of the method applied to LOFAR data corrupted by an aviation signal at 128.6 MHz.

#### 4.3. Spatial filtering approach

After having detected an RFI, it is possible to use its cyclostationary property to filter it. This filtering technique [3] is based on the use of a linear projection. As explained in Section 3.2, an SVD of the cyclic correlation matrix provides us with an estimation of the RFI spatial signature  $\mathbf{a}_r$ .

Let us define the following projector:

$$\mathbf{P}_{\text{pro}} = \mathbf{I} - \mathbf{a}_r \mathbf{a}_r^H \quad (13)$$

Assuming that  $\mathbf{a}_r^H \mathbf{a}_r = 1$ , it is straightforward to show that  $\mathbf{P}_{\text{pro}} \mathbf{a}_r = 0$ . Thus, by applying the projector  $\mathbf{P}_{\text{pro}}$  to the input signal vector  $\mathbf{z}(t)$  defined in Eq. (1), a clean version of the observation can be obtained:

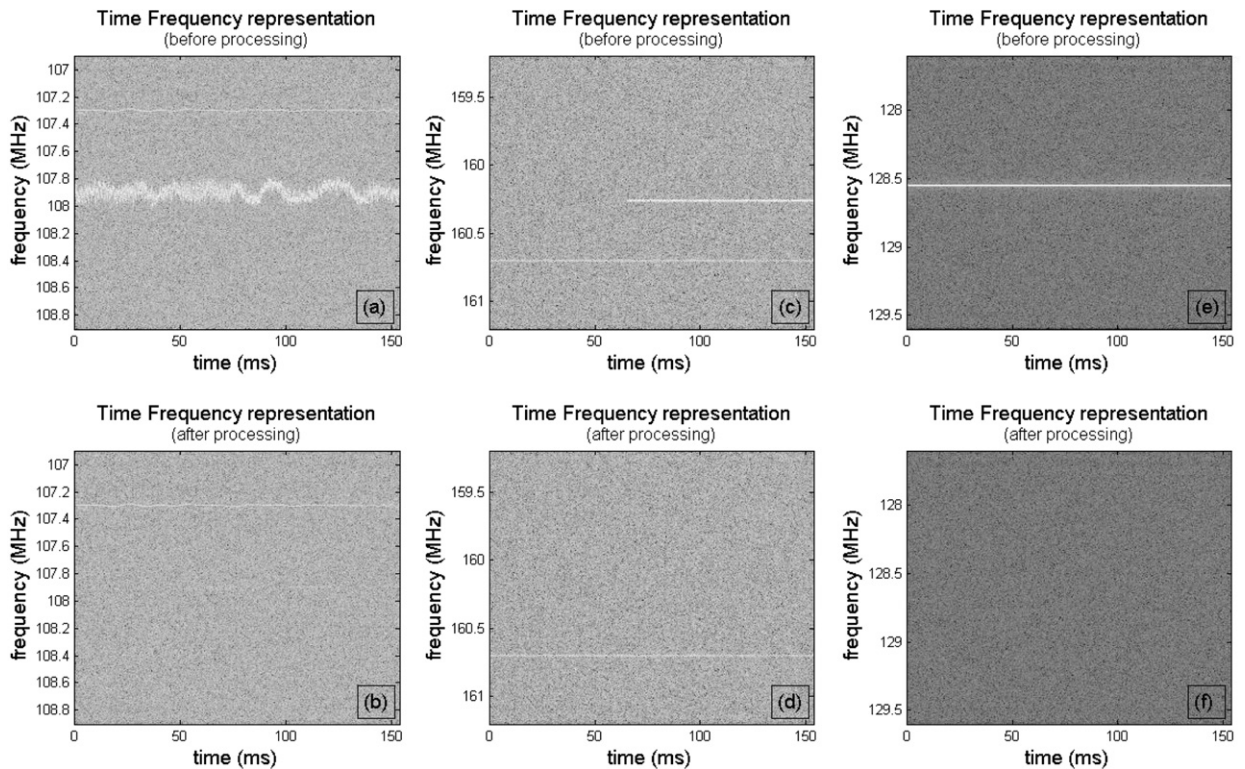
$$\mathbf{z}_{\text{clean}}(t) = \mathbf{P}_{\text{pro}} (\mathbf{A}_s \mathbf{s}(t) + \mathbf{n}(t)) \quad (14)$$

The bias induced by  $\mathbf{P}_{\text{pro}}$  in Eq. (14) can be removed as explained in [4]. Fig. 4 shows some results of this cyclic spatial filtering applied to LOFAR data. By using the cyclostationary approach, it is possible to focus the spatial filtering on one specific RFI.

This approach remains valid if  $K_r$   $\alpha_0$  cyclostationary RFIs are present. In that case, the vector  $\mathbf{a}_r$  will be replaced by the matrix  $\mathbf{U}_r$  formed by the singular vectors associated to the  $K_r$  largest singular values of the cyclic correlation matrix.

### 5. Conclusion

In this article, after a general introduction to the RFI mitigation problem in astrophysics, new techniques for RFI detection and mitigation have been proposed, based on the cyclostationary properties of the interferences. A model of the data recorded on an array radio telescope has been given and the cyclic statistics of the observation matrix have been derived and shown to be asymptotically independent of both the noise and the cosmic sources. Four different multisensor techniques for the estimation of the direction of arrival of an interference have then been presented. One of these techniques is based on the diagonalization of the correlation matrix of the sensors array. The other three techniques take advantage of some a priori knowledge about the cyclostationary features of RFIs and are based respectively on the diagonalization of a cyclic correlation



**Fig. 4.** Cyclic spatial filtering results on LOFAR data ( $M = 10$  antennas,  $\mathbf{R}_z^\alpha$  is estimated over  $L = 2^{15}$  samples). The spatial signature was estimated using cyclic singular value decomposition. (a) Data corrupted by FM signals. (b) Data shown in (a) after filtering of the 107.9 MHz RFI. (c) Data corrupted by land mobile signals. (d) Data shown in (c) after filtering of the starting RFI at 160.2 MHz. (e) Data corrupted by an aviation signal. (f) Data shown in (e) after filtering of the aviation RFI at 128.6 MHz.

matrix, the average of various cyclic correlation matrices related to the same interference and the joint diagonalization of these cyclic correlation matrices. These techniques have been compared through Monte Carlo simulations and the cyclic based techniques proved to be more robust to the array calibration. Lastly, we performed a two step application of the cyclic statistics to real-life data. The first one is the detection of cyclostationary interferences of unknown features by computing the Frobenius norms of a set of cyclic correlation matrices. The second one consists in building a projector from the estimated steering vector in order to remove the interference from the set of observations. The latter technique has been shown to be efficient on three different real-life situations, including an intermittent interference case. More generally, it has been shown that the cyclic statistics immunity to noise and poor calibration can be used to derive robust RFI detection and mitigation techniques.

## Acknowledgements

The authors would like to thank the European Commission Framework Program 7 (Project PrepSKA, contract no. 212243) and the French funding agency ANR (contract ANR-09-BLAN-0225-04) for funding part of this work.

## References

- [1] T. Yücek, H. Arslan, A survey of spectrum sensing algorithms for cognitive radio applications, *IEEE Communications Surveys & Tutorials* 11 (1) (2009), First Quarter.
- [2] A.J. Boonstra, R. Weber, RFI mitigation methods inventory, in: SKADS, FP6 European Project, DS4T3 Deliverable 1, 2009, 72 pp.
- [3] A.J. Boonstra, Radio frequency interference mitigation in radio astronomy, PhD thesis, University of Delft, Netherlands, June 2005.
- [4] S. van der Tol, A.J. van der Veen, Performance analysis of spatial filtering of RF interferences in radio astronomy, *IEEE Transactions on Signal Processing* 53 (3) (2005) 896–910.
- [5] W.A. Gardner, *Statistical Spectral Analysis – A Non-Probabilistic Theory*, Prentice-Hall, 1988.
- [6] S. Bretteil, R. Weber, Comparison between two cyclostationary detectors for RFI mitigation in radio astronomy, *Radio Science* 40 (2005).
- [7] R. Weber, et al., Data preprocessing for decametre wavelength exoplanet detection: an example of cyclostationary RFI detector, in: EUSIPCO, Ooznan, Poland, September 2007.
- [8] R. Feliachi, R. Weber, A.J. Boonstra, Cyclic spatial filtering in radio astronomy: Application to LOFAR data, in: EUSIPCO, Glasgow, UK, August 2009.
- [9] S.J. Wijnholds, et al., Calibration challenges for future radio telescopes, *IEEE Signal Processing Magazine* 27 (1) (2010) 32–42.
- [10] E. Serpedin, et al., Bibliography on cyclostationarity, *Signal Processing* 85 (12) (December 2005) 2233–2303.
- [11] W.A. Gardner, A. Napolitano, L. Paura, Cyclostationarity: Half a century of research, *Signal Processing* 86 (4) (April 2006) 639–697.



- [12] S.H. Friedberg, A.J. Insel, L.E. Spence, *Linear Algebra*, Prentice-Hall, July 1997.
- [13] H. Krim, M. Viberg, Two decades of array signal processing research, *IEEE Signal Processing Magazine* 13 (July 1997) 67–94.
- [14] A. Belouchrani, et al., A blind source separation technique using second order statistics, *IEEE Transactions on Signal Processing* 45 (February 1997) 434–444.
- [15] LOw Frequency ARray for radio astronomy, <http://www.lofar.org/>.
- [16] J.-M. Grießmeier, P. Zarka, M. Tagger, Radioastronomy with LOFAR, *C. R. Physique* 13 (1) (2012) 23–27, in this issue.