Journal of The Franklin Institute, Vol. 341, Issue 7, pp. 595-614, October 2004

Performance Analysis of a Spreading Sequence Estimator for Spread Spectrum Transmissions

C. Bouder, S. Azou and G. Burel

Laboratoire d'Electronique et Systèmes de Télécommunications (UMR CNRS 6165)

6, avenue Le Gorgeu, CS 93837, 29238 BREST cedex 3, FRANCE

June 17, 2004

Abstract

Direct sequence spread spectrum transmissions (DS-SS) are now widely used for secure communications, as well as for multiple access. They have many interesting properties, including low probability of interception. Indeed, DS-SS transmitters use a periodical pseudo-random sequence to modulate the baseband signal before transmission. A receiver which does not know the sequence cannot demodulate the signal.

In this paper, we propose a new method which can estimate the spreading sequence in a non cooperative context. The method is based on eigenanalysis techniques. The received signal is divided into windows, from which a covariance matrix is computed. We show that the sequence can be reconstructed from the two first eigenvectors of this matrix, and that useful information, such as desynchronisation time, can be extracted from the eigenvalues.

The main achievement of the present paper is a performance analysis of the proposed spreading sequence estimation procedure. An analytical approach is first considered owing to matrix perturbation theory and Wishart matrix properties. Then, complementary Monte Carlo simulations are performed to show the effectiveness of the proposed method.

Key Words : Digital communications, Direct Sequence Spread Spectrum, Blind Estimation, Performance analysis, Wishart matrix, Gold sequence.

1. Introduction

Spread spectrum transmissions have been in practical use since the 1950's. They found many applications in military systems due to their suitability for covert message transmission and resistance to jamming [1]. In the early 1980's, spread spectrum technology was proposed for private and commercial use, especially in Code Division Multiple Access (CDMA) transmissions. CDMA system has been adopted for use in commercially available wireless local area networks (WLAN's) [2]. Another useful feature of CDMA for indoor systems is its low power spectral density. This allows a CDMA system to coexist with licensed communications systems in environments where low levels of electromagnetic interference are desirable, such as hospital. In such environments, CDMA is ideally suited to high data rates being transmitted over hostile fading channels with the minimum of interference to sensitive equipments [3].

DS-SS is a transmission technique in which a pseudo-random sequence or pseudo-noise (PN) code [4], independent of the information data, is employed as a modulation waveform to spread the signal energy over a bandwidth much greater than the information signal bandwidth [5]. In practical systems, the bandwidth expansion factor, which is the ratio between the chip rate F_C and the data symbol rate F_S, is usually an integer. The amplitude, and thus the power in the spread spectrum signal, is the same as in the information signal. Due to the increased bandwidth of the spread spectrum signal, the power spectral density must be lower and can then be below the noise level [6]. Furthermore, the autocorrelation of a PN code has properties similar to those of white noise, so the spread spectrum signal looks like a white noise, hence it is very difficult to intercept. At the receiver, the signal is despread using a synchronized replica of the pseudo-noise sequence used at the transmitter. The optimum multiple user CDMA receiver is based on a correlator, or a bank of sequence matched filters, each followed by maximum likelihood sequence estimation detectors (MLSE). The objective of the MLSE is to find the input sequence that maximizes the conditional probability, or likelihood of the given output sequence [7].

In the context of spectrum surveillance, the pseudo-random sequence used by the transmitter is unknown, as well as other transmitter parameters such as duration of the sequence, symbol frequency and carrier frequency. Moreover, the longer the period of the pseudo-noise code is, the closer the transmitted signal will be to a truly random binary wave, and the harder it is to detect [8]. In this context, only Tsatsanis *et al.* [9] have proposed a reliable method to recover the convolution of the spreading sequence and the channel response in multipath environment. Their approach uses a multichannel identification technique due to Moulines *et al.* [10], where the orthogonality property between the signal and noise subspaces is exploited to yield the desired estimate.

In this paper, we propose a new method for estimating the pseudo-random sequence without prior knowledge about the transmitter. Our procedure does not rely on a multichannel framework and is computationally less expensive than that of Tsatsanis *et al.* The received signal is sampled and divided into temporal windows, the size of which is the pseudo-random sequence period, which is assumed to have been estimated. We prove that the spreading waveform can be recovered from the first and second eigenvectors of the sample covariance matrix. This property provides a simple way to estimate the pseudo-random sequence used at the transmitter. Furthermore, useful information about desynchronisation time can be extracted from the eigenvalues. In order to prove the reliability of the proposed technique, we analytically explore the statistics of the eigenparameters estimators, under a small perturbation assumption on the received signal covariance matrix. Numerical simulations are also proposed to complete this performance analysis and to show that the technique performs well in a multipath environment, even at low signal to noise ratio (SNR).

The paper is organized as follows. In Section 2, we give the notations, the hypotheses and we describe our spreading sequence estimation technique. Then, numerical experiments are given in section 3 to illustrate the method in multipath environment. Section 4 is devoted to a performance analysis of our algorithm. Finally, conclusions are drawn in section 5.

2. Problem formulation

In a direct sequence spread spectrum transmission, a pseudo-noise code generated at the modulator is used in conjunction with an M-ary Phase Shift Keying (PSK) modulation to shift the phase of the PSK signal randomly at the chip rate F_C, a rate that is an integer multiple of the symbol rate F_S. The bandwidth of the transmitted signal is determined by the chip rate.

Journal of The Franklin Institute, Vol. 341, Issue 7, pp. 595-614, October 2004

2.1. Notations and hypotheses

Let a_k be a QPSK or BPSK symbol transmitted at time kT_S, where T_S is the symbol duration. This symbol is multiplied by a pseudo-random sequence of chip duration T_C, which spreads the signal bandwidth. This discrete signal is then filtered, sent through the communication channel and filtered again at the receiver side. The resulting baseband signal is given by :

$$y(t) = \sum_{k=-\infty}^{+\infty} a_k h(t - kT_S) + n(t) \tag{2.1}$$

where $n(t)$ is the noise at the output of the receiver filter, and $h(t)$ encompasses the effects of the transmission filter, reception filter, channel response and the pseudo-random sequence :

$$h(t) = \sum_{k=0}^{P-1} c_k p(t - kT_C) \tag{2.2}$$

where $p(t)$ denotes the convolution of all the filters of the transmission chain and $\{c_k\}_{k=0\cdots P-1}$, is the pseudo-noise sequence of length P.

The chip duration T_C can be chosen such as .

Throughout the sequel, the hypotheses below will be assumed :

- The symbols are centered and uncorrelated;
- The noise is AWG (Averaged White Gaussian) and uncorrelated with the signal;
- The signal to noise ratio (SNR, in dB) at the output of the receiver filter is negative, i.e. the signal is hidden in the noise;
- The symbol period T_S is equal to the spreading code period, that is $T_C = T_S/P$; Its estimation is achieved in a pre-processing step, owing to the method proposed in [11] for example.

All other parameters are unknown.

2.2. Overview of the presence detection of direct-sequence spread spectrum signals (pre-processing step)

The basic principle of any intercept receiver is to take profit of the fact that the transmitted signal statistical properties are not the same as the noise statistical properties. For instance, in some simple applications, the signal and noise frequencies are not the same, hence filters are sufficient to detect the presence of a signal. Here, the application is much more complex, because a spread spectrum signal is specially built to be similar to a noise, in order to have a low probability of intercept (remind that spread spectrum was initially developed for military applications). For instance, the autocorrelation of a spread spectrum signal is close to a Dirac function, as well as the autocorrelation of a white noise (this is due to the pseudo-random sequence).

Our intercept receiver acts in two major steps : first the presence of any DS-SS transmission is detected and then we proceed to blind spreading sequence estimation before demodulation. The present paper focus on performance analysis of the second step. The pre-processing step, which is detailed in a recent paper [11] and briefly overviewed in this paragraph, also leads to an accurate spreading code period estimate which is then used in the rest of the algorithm.

The pre-processing step relies on the fluctuations of autocorrelation estimators, instead of on the autocorrelation itself. Although the autocorrelation of a DS-SS signal is similar to the autocorrelation of a noise, it is proved in [11] that the fluctuations of estimators are totally different.

Journal of The Franklin Institute, Vol. 341, Issue 7, pp. 595-614, October 2004

In order to compute the fluctuations, the received signal $y(t)$ is divided into M non-overlapping temporal windows, each of duration T large enough to contain a few symbols. Then an autocorrelation estimator is applied to each window :

$$\widehat{R_{yy}^{m}}(\tau) = \frac{1}{T}\int_{t_m}^{t_m+T} y(t)\, y^*(t-\tau)\, dt \quad , \quad m = 1, ..., M \tag{2.3}$$

Using the whole set of windows, we can then estimate the second order moment of this autocorrelation estimator :

$$\rho(\tau) = \frac{1}{M}\sum_{m=1}^{M} \left|\widehat{R_{yy}^{m}}(\tau)\right|^2 \tag{2.4}$$

By comparing the theoretical fluctuations of $\rho(\tau)$ in case of a "noise-only" hypothesis to that obtained by computation from the incoming signal, we can detect the presence of any DS-SS transmission.

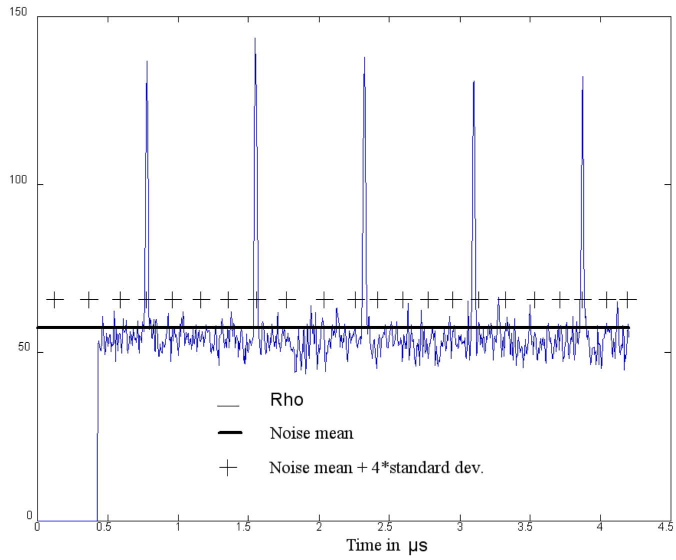

Fig. 1 - Example of DS-SS signals detector output (Gold code of length 31 at a chip rate Fc of 40 MHz in a Gaussian channel with SNR of -5 dB)

This result is illustrated by figure 1 where the fluctuations of $\rho(\tau)$ are plotted. The horizontal line show the theoretical average fluctuations $m_\rho^{(n)}$ and the crosses correspond to the theoretical average fluctuations plus 4 times the theoretical standard deviation $\sigma_\rho^{(n)}$ on the fluctuations ("noise-only" hypothesis). Large peaks that are observed, far above $m_\rho^{(n)} + 4\sigma_\rho^{(n)}$, indicate the presence of a DS-SS signal. These peaks are obtained for values of τ which are multiples of the spreading code period (0.775 μs for this example).

2.3. Blind estimation of the spreading sequence

Eigenanalysis techniques are exploited to recover the spreading sequence used at the transmitter. A first version of this approach, based on principal components analysis, was described in [12]. A second one, employing neural networks, was explained in [13]. In this paper we will focus on principal components analysis method.

The received signal is sampled and divided into non overlapping windows, the duration of which is T_S. Let us note $\mathbf{y}$ the content of a window. We can define the covariance matrix :

$$\mathbf{R} = E\left\{\mathbf{y}.\mathbf{y}^H\right\} \tag{2.5}$$

where ^{H} denotes the Hermitian transpose.

Since the window duration is equal to the symbol duration, a window always contains the end of a symbol for a duration $T_S - t_0$, followed by the beginning of the next symbol for a duration t_0, where t_0, the desynchronisation between a window and a symbol is unknown.

Hence we can write :

$$\mathbf{y} = a_m \mathbf{h}_0 + a_{m+1} \mathbf{h}_{-1} + \mathbf{n} \tag{2.6}$$

where $\mathbf{n}$ stands for the noise; $\mathbf{h}_0$ is a vector containing the end of the spreading waveform for a duration $T_S - t_0$, followed by zeroes for a duration t_0; $\mathbf{h}_{-1}$ is a vector containing zeroes for a duration $T_S - t_0$, followed by the beginning of the spreading waveform.

The equation above then leads to the following expression for the covariance matrix (2.5) :

$$\mathbf{R} = E\left\{|a_m|^2\right\} \mathbf{h}_0.\mathbf{h}_0^H + E\left\{|a_{m+1}|^2\right\} \mathbf{h}_{-1}.\mathbf{h}_{-1}^H + \sigma_n^2 \mathbf{I} \tag{2.7}$$

From this equation it is clear that two eigenvalues will be larger than the others. The corresponding eigenvectors will be equal to $\mathbf{h}_0$ and $\mathbf{h}_{-1}$, up to multiplicative factors.

The eigenvalues can be expressed in decreasing order according to the period symbol T_S, the SNR ρ, the sampling period T_e and the noise variance σ_n^2 [12]:

$$\begin{cases} \lambda_1 = \left(1 + \rho \frac{T_S - t_0}{T_e}\right) \sigma_n^2 \\ \lambda_2 = \left(1 + \rho \frac{t_0}{T_e}\right) \sigma_n^2 \\ \lambda_i = \sigma_n^2 \quad , \forall\, i \geq 3 \end{cases} \tag{2.8}$$

The mathematical expressions of the two first eigenvalues must be exchanged if $t_0 > \frac{T_S}{2}$.

The spreading sequence can then be recovered, once the corresponding normalized eigenvectors $\mathbf{v}_0 = \frac{\mathbf{h}_0}{\|\mathbf{h}_0\|^2}$ and $\mathbf{v}_{-1} = \frac{\mathbf{h}_{-1}}{\|\mathbf{h}_{-1}\|^2}$ are concatenated.

Besides, it has been shown in [12] that the desynchronisation time t_0 and the SNR ρ can be estimated from (2.8) :

$$\begin{cases} \widehat{\rho} = \left(\frac{\lambda_1 + \lambda_2}{\sigma_n^2} - 2\right) \frac{T_e}{T_S} \\ \widehat{t_0} = \frac{T_e}{\widehat{\rho}} \left(\frac{\lambda_2}{\sigma_n^2} - 1\right) \end{cases} \tag{2.9}$$

In fact, the covariance matrix cannot be exactly defined, but only estimated by the sample covariance matrix :

$$\widehat{\mathbf{R}} = \frac{1}{N} \sum_{m=1}^{N} \mathbf{y}_m.\mathbf{y}_m^H \tag{2.10}$$

where N is the number of temporal windows of duration T_S .

The performance of the estimation of the spreading sequence will therefore depend on the number N besides the SNR. In the sequel we will suppose the sampling period is set to $T_e = T_C$, to make the interpretation of the results easier, but this is not a requirement. In this case the covariance matrix is a $P \times P$ complex matrix, where P is the length of the spreading sequence.

To illustrate our algorithm, computer simulations will be given in the next section, to show the method can provide a good estimation of the spreading sequence, even when the received signal is far below the noise level.

3. Numerical examples - Comparison

In this section, we provide representative comparisons of our approach with the method described in [9]. An example also illustrates the feasibility of our solution in multipath environment.

Example 1. .

The goal of this example is to show that our method is a good alternative to the algorithm proposed in [9], whose computational cost is higher due to its matrix dimensionality.

For illustration purposes, we consider a spread spectrum signal generated by spreading a QPSK modulation with a Gold sequence of length 31 (Fig. 2). For comparison the method of [9] is computed when multipath is absent. In this particular case, only the blind estimation of the spreading code used by the transmitter is performed. AWG noise is added to the received signal with $SNR = -5\ dB$. Both algorithms have been used with $N = 200$ analysis windows. We assume also, the length of the spreading code P and the chip period T_C are known as needed by subspace method with multichannel framework, although these hypotheses are not required by our algorithm.

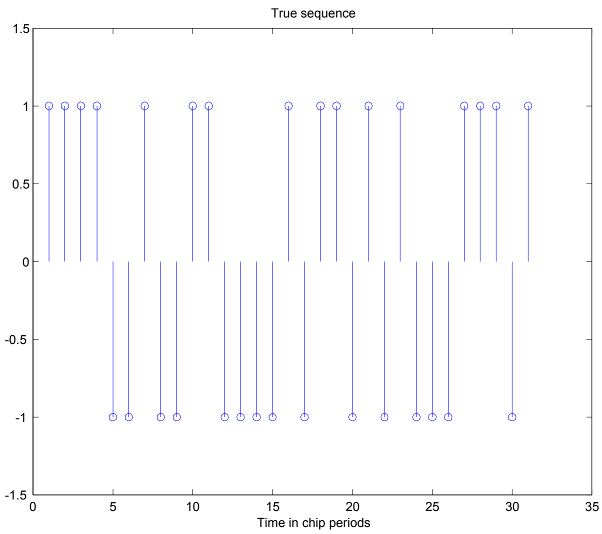

Fig. 2 : Gold sequence

The figures below show the estimated sequence by subspace method with multichannel framework (Fig. 3) and by our eigenanalysis technique (Fig. 4), when a desynchronisation time $t_0 = 8$ chips is present.

Journal of The Franklin Institute, Vol. 341, Issue 7, pp. 595-614, October 2004

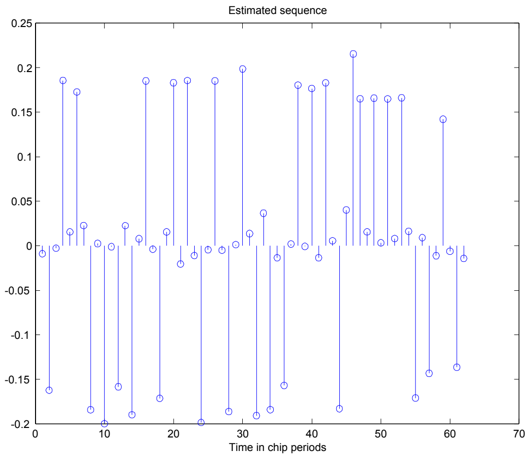

Fig. 3 : Subspace method with multichannel framework

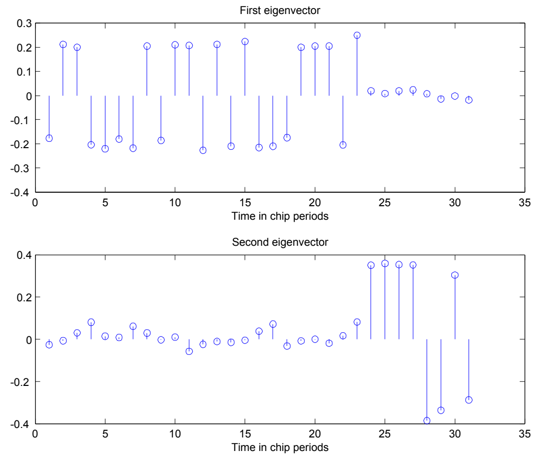

Fig. 4 : The proposed approach

The eigenanalysis of $\widehat{\mathbf{R}}$ shows there are two large eigenvalues as demonstrated in section 2. Once the desynchronisation time has been estimated according to the eigenvalues (2.9), it is possible to determine which eigenvector contains the beginning of the spreading sequence and which one the end. At the top of the Fig. 4, the first eigenvector is drawn, it contains the end of the spreading sequence during $(T_S - t_0)$ followed by zeroes during t_0, where t_0 is equal to 8 chips, while the second eigenvector at the bottom of Fig. 4 contains zeroes during $(T_S - t_0)$ followed by the beginning of the spreading code during t_0. The eigenvectors can then be concatenated to recover the spreading code used at the transmitter.

Example 2. *Results in multipath environment*

A Monte Carlo simulation is conducted to evaluate the performance of the proposed method in a multipath environment. A DS-SS signal is generated using a random sequence of length $P = 31$. The symbols belong to a BPSK constellation. The spreading sequence is shown in Fig. 5. The signal is then transmitted through a multipath channel, with the following impulse response :

$$h(t) = 1.2 - 0.8\ \delta(t - T_C) + 0.6\ \delta(t - 2T_C)$$

The channel impulse response is depicted in Fig. 6, and the overall response, i.e. the convolution of the spreading code and the channel impulse response, can be seen in Fig. 7. AWG noise is added to the received signal with $SNR = -10\ dB$, the top of Fig. 8 shows an example of the first transmitted BPSK chips and the bottom, the received chips corrupted by the noise. To verify the efficiency of our algorithm, the combined response of the spreading code with the channel impulse response has been estimated across a 100 runs Monte Carlo simulation, with $N = 200$ analysis windows.

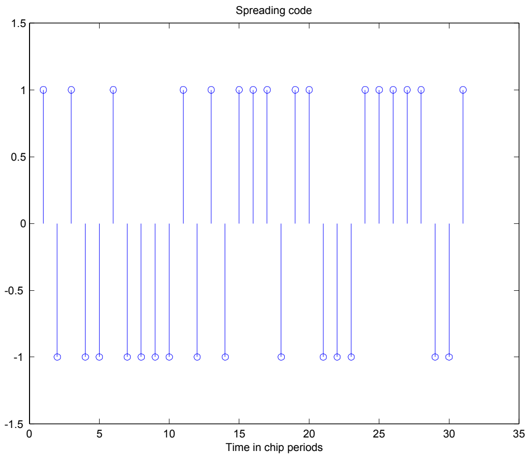

Fig. 5 : Spreading code

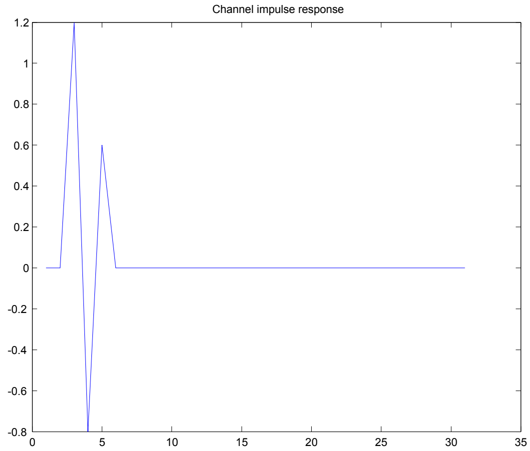

Fig. 6 : Channel impulse response

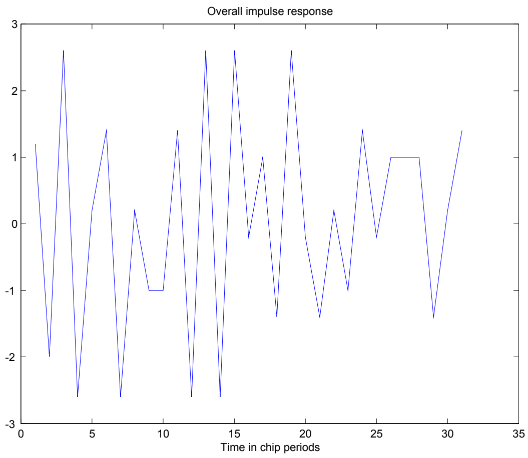

Fig. 7 : Overall impulse response

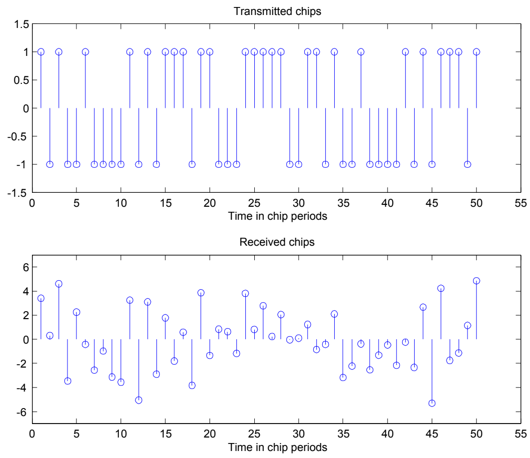

Fig. 8 : Transmitted and Received chips

Fig. 9 below shows the true overall impulse response ('o'), as well as the mean of the 100 runs ('*') and the standard deviation. The estimated response is clearly seen to be in accordance with the true one.

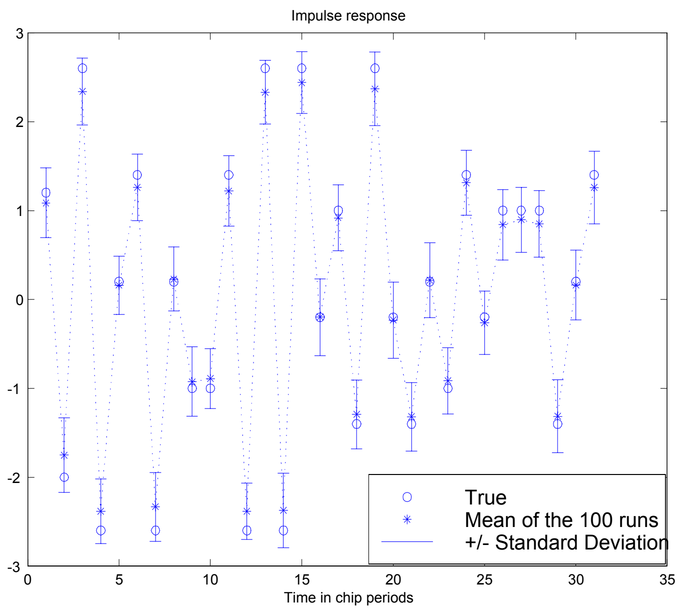

Fig. 9 : Monte Carlo results on estimated impulse response

Journal of The Franklin Institute, Vol. 341, Issue 7, pp. 595-614, October 2004

The estimated impulse response can then be used by a classical receiver in order to recover the transmitted information from the received data.

Nevertheless, when estimating the sample covariance matrix, an error $\mathbf{\Delta R}$ is made [15], which in turn induces an error in the estimated eigenvalues and eigenvectors. To complete our results, we have to pick out the effects of the perturbation on the eigenvalues and eigenvectors due to the estimation of the covariance matrix.

4. Performance evaluation

In order to demonstrate the performance of the proposed method, compact expressions for the variance and bias of the estimates of the eigenvalues and eigenvectors $\{\lambda_k, \mathbf{u}_k\}$ will be derived in the following. As a first step, each estimation error $\{\Delta\lambda_k, \Delta\mathbf{u}_k\}$ will be written as a Taylor expansion of the elements of $\Delta\mathbf{R}$ using well-known results from matrix perturbation theory. Then, as the sample covariance matrix follows a complex Wishart law, a statistical analysis of $\{\Delta\lambda_k, \Delta\mathbf{u}_k\}$ will be achieved owing to Wishart matrix properties [16], [17]. Similar studies have been previously carried out in the domain of high-resolution Direction-Of-Arrival (DOA) techniques [15], [18], [19].

Note that it is not the unique approach to solve our performance analysis problem. In the noise only case (AWGN), the joint eigenvalue distribution of a complex Wishart matrix is known [20], and use of zonal polynomials [21] even yields the distribution of each eigenvalue in a closed form. However, as noted by [22] in a different context, this gives intractable results for the moment. In the sequel the approach of [15] has been preferred, due to its simplicity and the standard matrix formalism required.

4.1. Theoretical analysis

As exposed in the previous section, our algorithm yields the spreading sequence through an eigenvalue decomposition of the sample covariance matrix (2.10).

As the random data set $\{\mathbf{y}_m\}_{m=1,\ldots,N}$ is composed from complex Gaussian, circular, independent vectors, $\widehat{\mathbf{R}}$ is distributed according to a complex Wishart law [17].

Denote $\Delta\mathbf{R}$ the small Hermitian perturbation on $\mathbf{R}$ due to finite number of temporal windows N :

$$\Delta\mathbf{R} = \widehat{\mathbf{R}} - \mathbf{R} \tag{4.1}$$

Then, the expected eigenparameters $\{\lambda_k, \mathbf{u}_k\}$ of $\mathbf{R}$ are changed into

$$\hat{\lambda}_k = \lambda_k + \Delta\lambda_k \tag{4.2}$$

$$\hat{\mathbf{u}}_k = \mathbf{u}_k + \Delta\mathbf{u}_k \tag{4.3}$$

with the following orthogonality constraint on eigenvectors :

$$\mathbf{u}_k^H \Delta\mathbf{u}_k = 0 \tag{4.4}$$

Remark 1. *The perturbation results derived throughout the sequel will be only valid for simple eigenvalue case of the covariance matrix* $\mathbf{R}$ *(and for associated eigenvectors). This is not a problem since the spreading sequence information is carried by the first two simple eigenvalues, in a decreasing order. Hence, the values for subscript* k *will be* 1 *or* 2 *if a desynchronisation between windows* $\{\mathbf{y}_m\}_{m=1,\ldots,N}$ *and symbols exists and* $k = 1$ *in the synchronized case. For a detailed discussion about matrix perturbation theory see [23] for example.*

Journal of The Franklin Institute, Vol. 341, Issue 7, pp. 595-614, October 2004

A statistical analysis of perturbations $\{\Delta\lambda_k, \Delta\mathbf{u}_k\}$ has now to be performed in order to check our algorithm efficiency. The following theorem, stated by Krim *et al.* [18], provides a relation between these perturbation terms and $\Delta\mathbf{R}$. Let $\{\delta^n\lambda_k,\ \delta^n\mathbf{u}_k\}$ be the n^{th} order terms of the Taylor expansion of perturbations $\{\Delta\lambda_k, \Delta\mathbf{u}_k\}$, respectively :

$$\Delta\lambda_k = \delta\lambda_k + \delta^2\lambda_k + ... + \delta^n\lambda_k + ... \tag{4.5}$$

$$\Delta\mathbf{u}_k = \delta\mathbf{u}_k + \delta^2\mathbf{u}_k + ... + \delta^n\mathbf{u}_k + ... \tag{4.6}$$

Theorem 4.1. *Then, by recurrence we have :*
First order terms,

$$\begin{aligned} \delta\lambda_k &= \mathbf{u}_k^H \Delta\mathbf{R}\mathbf{u}_k \\ \delta\mathbf{u}_k &= \mathbf{S}_k^\dagger \Delta\mathbf{R}\mathbf{u}_k \end{aligned} \tag{4.7}$$

Higher order terms ($n > 1$),

$$\begin{aligned} \delta^n\lambda_k &= \mathbf{u}_k^H \Delta\mathbf{R}\delta^{n-1}\mathbf{u}_k \\ \delta^n\mathbf{u}_k &= \mathbf{S}_k^\dagger \Delta\mathbf{R}\boldsymbol{\delta}^{\mathbf{n-1}}\mathbf{u}_k - \sum_{l=1}^{n-1} \delta^{n-l}\lambda_k \mathbf{S}_k^\dagger \delta^l \mathbf{u}_k \end{aligned} \tag{4.8}$$

where

$$\mathbf{S}_k = \lambda_k\mathbf{I} - \mathbf{R} \tag{4.9}$$

and $\mathbf{S}_k^\dagger$ is its pseudo-inverse :

$$\mathbf{S}_k^\dagger = \sum_{i=1\neq k}^{P} \frac{1}{(\lambda_k - \lambda_i)} \mathbf{u}_i\mathbf{u}_i^H \tag{4.10}$$

The next properties, due to Wishart distribution of $\widehat{\mathbf{R}}$, will be of central importance to derive the statistics (mean and variance) of the eigenparameters perturbations. Proofs of these properties can be found in multivariate statistics literature such as [16] or [17].

Proposition 1.
1. The estimation error (4.1) satisfies :

$$E\{\Delta\mathbf{R}\} = \mathbf{0} \tag{4.11}$$

$$E\{\Delta\mathbf{R}(i,j)\Delta\mathbf{R}(i',j')\} = \frac{1}{N}\mathbf{R}(i,j')\mathbf{R}(i',j) \tag{4.12}$$

2. For any $P \times P$ complex matrix $\mathbf{A}$:

$$E\{\Delta\mathbf{R}\ \mathbf{A}\ \Delta\mathbf{R}\} = \frac{1}{N} tr(\mathbf{R}\mathbf{A})\mathbf{R} \tag{4.13}$$

Owing to these relations, approximate expressions for the moments of the estimators $\{\hat{\lambda}_k, \hat{\mathbf{u}}_k\}$ are easily derived, as stated in the next propositions. The mean is obtained by truncating the Taylor expansion (4.5) up to the second order. For the variance, a first order truncation is seen to be sufficiently accurate, as shown in the next subsection by Monte Carlo simulations.

Journal of The Franklin Institute, Vol. 341, Issue 7, pp. 595-614, October 2004

Proposition 2. *The expansions (4.5) and (4.6) used up to the second order yields the following approximate expressions for the first moment of the estimators (4.2) and (4.3) :*

$$E\{\hat{\lambda}_k\} = \lambda_k \left[1 + \frac{1}{N}\sum_{i=1\neq k}^{P} \frac{\lambda_i}{\lambda_k - \lambda_i}\right] \tag{4.14}$$

$$E\{\hat{\mathbf{u}}_k\} = \mathbf{u}_k \tag{4.15}$$

Proposition 3. *The expansions (4.5) and (4.6) used up to the first order yields the following approximate expressions for the second moment of the estimators (4.2) and (4.3) :*

$$var(\hat{\lambda}_k) = \frac{\lambda_k^2}{N} \tag{4.16}$$

$$cov(\hat{\mathbf{u}}_k) = \frac{\lambda_k}{N}\sum_{i=1\neq k}^{P} \frac{\lambda_i}{(\lambda_k - \lambda_i)^2}\mathbf{u}_i\mathbf{u}_i^H \tag{4.17}$$

It should be mentioned now that the spreading waveform is estimated up to an unknown multiplicative factor. This is not a problem because in any transmission system, the channel is modeled by at least an unknown multiplicative factor. Hence, this uncertainty is also included in the channel model. This is why transmission systems use either differential coding, or periodically occurring known symbol sequences, in order to remove this uncertainty.

Hence we have :

$$\hat{\mathbf{u}}_1 = z\ \mathbf{u}_1 + \Delta\mathbf{u}_1 \tag{4.18}$$

where z is the unknown multiplicative factor. Since this uncertainty is present in any transmission system, we can replace $z\mathbf{u}_1$ by $\mathbf{u}_1$, when our purpose is the evaluation of the performances of our method. Indeed, it is $\Delta\mathbf{u}_1$, and not z, which represents the quality of the estimator.

Since the demodulator uses a correlator to recover the data symbols, a good criterion to evaluate the quality of the estimated sequence is its normalized scalar product with the sequence :

$$\mathcal{C}(\hat{\mathbf{u}}_1, \mathbf{u}_1) = \frac{\hat{\mathbf{u}}_1^H\mathbf{u}_1}{\|\hat{\mathbf{u}}_1\|\,\|\mathbf{u}_1\|} \tag{4.19}$$

where $\mathbf{u}_1$ is the true sequence, $\hat{\mathbf{u}}_1$ is the eigenvector corresponding to the largest eigenvalue. Note that symbol synchronisation has been assumed above to simplify the presentation.

Proposition 4. *An approximate expression for the first moment of* $\mathcal{C}(\hat{\mathbf{u}}_1, \mathbf{u}_1)$ *is*

$$E\{\mathcal{C}\} = 1 - \frac{1}{2}\frac{\lambda_1}{N}\sum_{i=2}^{P}\frac{\lambda_i}{(\lambda_1 - \lambda_i)^2} \tag{4.20}$$

Proof. Because $\mathbf{u}_1$ is a normalized vector orthogonal to $\Delta\mathbf{u}_1$, it is clear that

$$\mathcal{C}(\hat{\mathbf{u}}_1, \mathbf{u}_1) = \frac{1}{\|\hat{\mathbf{u}}_1\|} = \frac{1}{[tr\{\hat{\mathbf{u}}_1\hat{\mathbf{u}}_1^H\}]^{1/2}}$$

Then, it is straightforward to see that

$$tr\{\hat{\mathbf{u}}_1\hat{\mathbf{u}}_1^H\} = \hat{\mathbf{u}}_1^H\hat{\mathbf{u}}_1 = 1 + \Delta\mathbf{u}_1^H\Delta\mathbf{u}_1$$

Assuming that $\Delta\mathbf{u}_1^H\Delta\mathbf{u}_1 \ll 1$, we obtain the following expression by truncating the series expansion up to the first order:

$$\mathcal{C}(\hat{\mathbf{u}}_1, \mathbf{u}_1) = 1 - \frac{1}{2}\Delta\mathbf{u}_1^H\Delta\mathbf{u}_1$$

Finally, replacing $\Delta\mathbf{u}_1$ by its first order approximation (4.7) yields the desired relation (4.20). ■

4.2. Verification by Monte Carlo simulations

Numerical simulations are now proposed in order to confirm the effectiveness of our algorithm in practical situations.

It should be noted that all the previous analytical approach is based on various approximations : The signal covariance error $\Delta\mathbf{R}$ has been assumed to be small in norm, which means that we have only considered the asymptotical case ($N \to \infty$). Also, other errors are caused by various Taylor series truncations when evaluating means or covariances.

Moreover, it is interesting to study the performance degradation in estimating the spreading sequence when only a short data record, i.e. a small number of windows N, is available.

In the following example, we show a comparison between the approximate analytical eigenvalue distribution (4.14 and 4.16) and the same distribution derived by Monte Carlo simulations, for various sizes of samples set N and a fixed SNR.

Example 3. *We compute a QPSK signal spread by a Gold sequence of length* 31*, corrupted by a AWG noise with* $SNR = -5$ *dB. The symbol synchronisation is assumed here for the purpose of illustration, but this is not a requirement for the above analytical results to be valid. In this case, the signal information is exclusively carried by the first eigenvalue . Fig. 10 depicts the true eigenvalue ('o') and the average estimation ('+') predicted by equation (4.14)* $\pm$ *the standard deviation predicted by equation (4.16). In Fig. 11, the average eigenvalue ('*')* $\pm$ *the standard deviation is plotted across a Monte Carlo simulation consisting of 500 runs. It is clearly seen that our analytical results (Fig. 10) compare well with those obtained with Monte Carlo experiments (Fig. 11); we can note that the theoretical second order moment (4.16) tends to be overestimated but as will be seen on the next figure this will not affect the prediction of the spreading code estimate accuracy. Another observation is that to get a correct estimation of the eigenvalues/eigenvectors and as a consequence a correct estimation of the spreading code, about 150 analysis windows are required (for a SNR of -5 dB).*

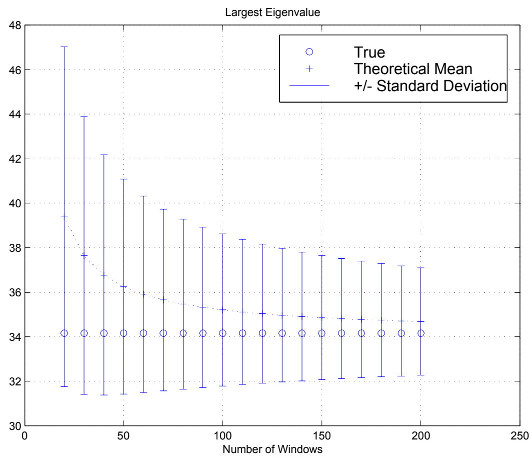

Fig. 10 : Analytical distribution

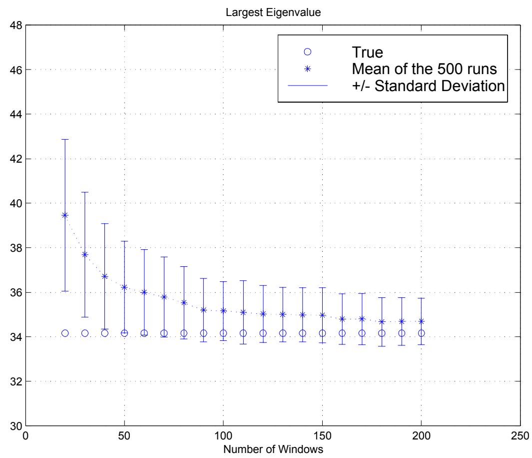

Fig. 11 : Monte Carlo simulations

Journal of The Franklin Institute, Vol. 341, Issue 7, pp. 595-614, October 2004

The next example is proposed to evaluate the average error in estimating the spreading sequence for different values of $\{N, SNR\}$.

Example 4. *We consider a QPSK modulation again, spread by a Gold sequence of length* 31*. Criterion (4.19) is chosen as the mean to reflect the efficiency of the spreading sequence estimation. Fig. 12 shows the mean of the normalized scalar product of the estimated sequence with the true one ('$*$') predicted by equation (4.20), as well as the average value ('o') obtained via 100 runs of a Monte Carlo simulation.*

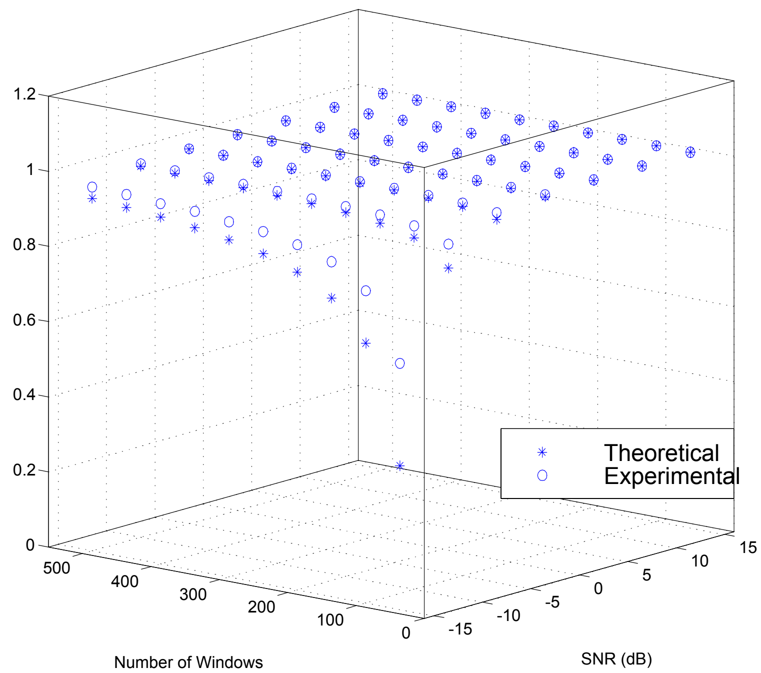

Fig. 12 : Quality of the estimated sequence evaluation owing to the criterion $\mathcal{C}$

From this figure, we can have an idea of the number of windows N and the SNR from which a good accordance between the estimated sequence and the true one can be reached.

Another way to bring out our results is to check the average number of sign errors between the true sequence and the estimated one. A 100 runs experiment has been done with a fixed number of temporal windows $N = 200$. The tables below show the percentage of sign errors versus the SNR in dB, when the signal is well synchronized (Table 1), as well as those obtained if a desynchronisation time $t_0 = 8$ chips is present (Table 2).

SNR (dB)	-16	-15	-14	-13	-12	-11	-10	-9	-8
Sign errors (%)	1.33	0.4	0.15	0.05	0.01	0	0	0	0

Table 1 : Percentage of sign errors in the spreading sequence versus SNR (synchronized case)

SNR (dB)	-16	-15	-14	-13	-12	-11	-10	-9	-8
Sign errors (%)	4.1	2.7	1.37	0.9	0.45	0.13	0.01	0	0

Table 2 : Percentage of sign errors in the spreading sequence versus SNR (desynchronized case)

We can draw a parallel between Fig. 12 and Table 1 when the number of windows is fixed to $N = 200$, to confirm the conclusions.

Journal of The Franklin Institute, Vol. 341, Issue 7, pp. 595-614, October 2004

For comparison, the tables below present the performances of the subspace method with multichannel framework [9], in the same conditions as above. It is clearly seen that our algorithm is a good alternative, with a lower computational cost.

SNR (dB)	-16	-15	-14	-13	-12	-11	-10	-9	-8
Sign errors (%)	1.2	0.33	0.15	0	0	0	0	0	0

Table 3 : Percentage of sign errors in the spreading sequence versus SNR (synchronized case)

SNR (dB)	-16	-15	-14	-13	-12	-11	-10	-9	-8
Sign errors (%)	2.32	0.73	0.34	0.17	0.02	0	0	0	0

Table 4 : Percentage of sign errors in the spreading sequence versus SNR (desynchronized case)

5. Conclusion

The blind spreading sequence estimation problem in a DS-SS transmission system has been considered in this paper. The proposed algorithm is very easy to implement and requires a limited computational cost since it is based on a single eigenvalue decomposition of the received signal sample covariance matrix. This process also enables the estimation of parameters such as SNR or desynchronisation time. A performance analysis has been investigated. An analytical approach shows the efficiency of the proposed estimators asymptotically. By complementary Monte Carlo simulations it is seen that a good performance can even be achieved for a moderate length of the data record and a low SNR. No hypothesis was assumed on the nature of the spreading sequence : it can be a sequence generated by pseudo-random shift registers, such as Gold sequences, but this is not a requirement. Once estimated, the sequence can be used by a traditional spread spectrum receiver in order to retrieve the information symbols.

Journal of The Franklin Institute, Vol. 341, Issue 7, pp. 595-614, October 2004

References

[1] C. E. Cook, H. S. Marsh, An Introduction to Spread Spectrum, *IEEE Comm. Magazine*, March 1983, 8-16.

[2] P. G. Flikkema, Spread-Spectrum Techniques for Wireless Communication, *IEEE Signal Processing Magazine*, May 1997, 26-36.

[3] I Oppermann, P. Van Rooyen, R. Kohno, Guest Editorial Spread Spectrum for Global Communications II, *IEEE Journal on selected areas in Comm.*, Vol. 18, N°1, Jan. 2000, 1-5.

[4] D. V. Sarwate, M. B. Pursley, Crosscorrelation properties of pseudorandom and related sequences, *Proc. IEEE*, Vol. 68, N°5, May 1980, 593-619.

[5] J. G. Proakis, Digital Communications, Third Edition Mc Graw Hill International Edition, 1995, ISBN : 0-07-113814-5.

[6] R. L. Pickholtz, D. L. Schilling, L. B. Milstein, Theory of Spread Spectrum Communications - A Tutorial, *IEEE Trans. on Comm.*, Vol. COM 30, N°5, May 1982, 855-884.

[7] S. Glisic, B. Vucetic, Spread Spectrum CDMA Systems for Wireless Communications, Artech House Publishers, 1997, ISBN : 0-89006-858-5.

[8] R. Gold, Optimal Binary Sequences for Spread Spectrum multiplexing, *IEEE Trans. on Information Theory*, Oct. 1967, 619-621.

[9] N. K. Tsatsanis, G. B. Giannakis, Blind Estimation of direct sequence spread spectrum signals in multipath, *IEEE Trans on Signal Processing*, Vol. 45, N°12, May 1997, 1241-1251.

[10] E. Moulines, P. Duhamel, J-F. Cardoso, S. Mayrargue, Subspace Methods for the Blind Identification of Multichannel FIR Filters, *IEEE Trans. on Signal Processing*, Vol. 43, N°2, Feb. 1995,516-525.

[11] G. Burel, Detection of Spread Spectrum Transmissions using fluctuations of correlation estimators, *IEEE Int. Symp. on Intelligent Signal Processing and Communication Systems (ISPACS'2000),* Hawaï, Nov., 2000.

[12] G. Burel, C. Bouder, Blind Estimation of the Pseudo-random Sequence of a Direct Sequence Spread Spectrum Signal, *IEEE 21st Century Military Communications Conference (IEEE- MILCOM'2000)*, Oct., Los-Angeles.

[13] C. Bouder, G. Burel, Spread Spectrum Codes Identification by Neural Networks, *4th World Multiconference on Circuits, Systems, Communications & Computers (CSCC 2000)*, July 2000, Vougliameni, GREECE.

[14] C. Bouder, S. Azou, G. Burel, "A robust synchronisation procedure for blind estimation of the symbol period and the timing offset in spread spectrum transmissions", IEEE International Symposium on Spread Spectrum Techniques and Applications (ISSSTA), Prague, Czech Republic, September 2-5, 2002.

[15] H. Krim, J. G. Proakis, Smoothed Eigenspace-Based Parameter Estimation, *Automatica, Special Issue on Statistical Signal Processing and Control*, Jan. 1994.

Journal of The Franklin Institute, Vol. 341, Issue 7, pp. 595-614, October 2004

[16] Anderson, An Introduction to Multivariate Statistical Analysis, Wiley, 1958.

[17] R. J. Muirhead, Aspects of Multivariate statistical theory, J. Wiley & Sons, 1982.

[18] H. Krim, P. Forster, J. G. Proakis, Operator Approach to Performance Analysis of Root-MUSIC and Root-Min-Norm, *IEEE trans. on signal processing*, Vol. 40, N°7, July 1992, 1687-1696.

[19] S. Marcos, Les Méthodes à Haute Résolution - Traitement d'antennes et analyse spectrale, Hermès, 1998, ISBN : 2- 86601- 662- 9.

[20] A. Edelman, Eigenvalues and Condition Numbers of Random Matrices, Massachusetts Institute of Technology Doctoral Dissertation, Mathematics Department, May 1989.

[21] P. R. Krishnaiah, F. J. Schurmann, On the Evaluation of some Distribution that arise in Simultaneous Tests of the Equality of the Latents Roots of the Covariance Matrix, *Journal of multivariate analysis*, N°4, 1974, 265-282.

[22] J. Grouffaud, P. Larzabal, H. Clergeot, Some properties of ordered Eigenvalues of a Wishart Matrix : Application in detection test and model order selection, *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing Conference (ICASSP)*, May, 1996.

[23] G. H. Golub, C. F. Van Loan, Matrix Computations, Johns Hopkins, University Press, Baltimore, 1983.