A GENERALIZATION OF THE FIXED POINT ESTIMATE

Abigael Taylor[1], Philippe Forster[2], Franck Daout[2], Hélène Oriot[1], Laurent Savy[1]

[1]ONERA – The French Aerospace Lab, F-91761 – Palaiseau, France
[2]SATIE, ENS Cachan, 61 av du Président Wilson 94235 Cachan CEDEX

ABSTRACT

In this paper, the problem of proportional covariance matrices estimation for random Gaussian complex vectors is investigated. The maximum likelihood estimates of the matrix and the scale factors are derived, and their statistical performances are studied, through bias, consistency and asymptotic distribution. It is also shown that the problem treated here generalizes the covariance estimation problem for Spherically Invariant Random Vector (SIRV). An iterative estimation algorithm is proposed. A simulation based on a detection problem is presented. The results suggest that the asymptotic distribution obtained is a really good approximation, even for a small number of data.

Index Terms— Maximum likelihood estimate, covariance estimate, proportional covariance matrices, spherically invariant random vector (SIRV)

1. INTRODUCTION

The problem of estimating M proportional covariance matrices for random Gaussian complex vectors is considered. This problem finds applications in array processing and the authors were faced to it in this context. However, due to the lack of space, potential applications in array processing (such as covariance estimation in radar) will not be developped. The first papers to calculate the maximum likelihood (ML) estimates of the covariance matrix R and the scale factors λ_i, $i = 1..M$, and to propose a way of how to actually obtain them, are [1] and [2]. Their works are carried out for the case where $\lambda_1 = 1$, and where the data used are real-valued. In this paper, complex data are considered, and no constraint is imposed on λ_1. This latter point does not have a strong impact on the calculation, but enables a different interpretation of the results.

A formulation of the problem is first presented in Section 2. In Section 3, the ML estimates are derived, and based on these estimates, an algorithm is proposed. Their statistical properties are studied, through the derivation of the bias, consistency and asymptotic distribution. Section 4 presents experimental results based on a detection problem, which corroborate the theoretical results presented.

Notations: The notation $\mathbf{x} \sim N_c(\mathbf{0}, \mathbf{R})$ means that $\mathbf{x}$ is a zero mean complex Gaussian vector, with covariance $\mathbf{R}$. $\mathbf{W} \sim W(K, \mathbf{R})$ means that $\mathbf{W}$ follows a Wishart law of parameter K and expectation $\mathbf{R}$.
Operator $E(.)$ and $Tr(.)$ stand for expectation and trace. $.^H$ denotes the transpose conjugate operator. $\overset{P}{\rightarrow}$ denotes convergence in probability.

2. PROBLEM FORMULATION

Let $\{\mathbf{x}_{1,1}, \dots, \mathbf{x}_{1,K_1}\}$, …, $\{\mathbf{x}_{\mathrm{M},1}, \dots, \mathbf{x}_{\mathrm{M},\mathrm{K_M}}\}$ be M sets of independent N-dimensional zero-mean random gaussian complex vectors with equal unknown covariance matrices, up to unknown scale factors:

$$\mathbf{x}_{i,j} \sim N_c(\mathbf{0}, \lambda_i \mathbf{R}) \quad (1)$$

where λ_i , $\mathbf{R}$ are unknown, and must be estimated. For that purpose, a maximum likelihood procedure will be followed and the estimates statistical properties will be investigated. It can be noticed that a constraint must be imposed on $\mathbf{R}$ and its estimate $\widehat{\mathbf{R}}$ to avoid over-parameterization, such as, for instance, a constraint on their traces. In this paper, the constraint used will be in the form of:

$$Tr(\mathbf{AR}) = Tr(\mathbf{A}\widehat{\mathbf{R}}) = N \quad (2)$$

where $\mathbf{A}$ will be either $\mathbf{I}$ (in practice) or $\mathbf{R}^{-1}$ (for theoretical purpose only, since it is not available in real life).

The next section derives the ML estimates of λ_i and $\mathbf{R}$, as well as an iterative algorithm to obtain them. Their statistical properties are also investigated. In the paper, the Sample Covariance Matrices (SCM) of the different data-sets will be noted $\widehat{\mathrm{R}}_{\mathrm{i}}$,

$$\widehat{\mathbf{R}}_i = \frac{1}{K_i} \sum_{k=1}^{K_i} \mathbf{x}_{i,k} \mathbf{x}_{i,k}{}^H \quad (3)$$

and we will set

$$K = \sum_{i=1}^{M} K_i \quad (4)$$

$\widehat{\mathbf{R}}_i$ follows a complex Wishart distribution with K_i degrees of freedom and expectation $\lambda_i \mathbf{R}$.

It can be noticed that our problem is a generalization of a problem well-investigated in the recent signal processing literature. Indeed, the case $K_1 = K_2 = \dots = K_M = 1$ leads to a SIRV (Spherically Invariant Random Vector) model, for which the ML estimates have been derived and theoretically analyzed [3][4][5]. Examples of application can be found in [6] and [7], where the authors consider sets of compound-Gaussian (included in SIRV model) data with proportional covariance matrices, which they have to estimate for constant false alarm rate detection purpose.

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

3. MAXIMUM LIKELIHOOD ESTIMATES

Let L be the negated log-likelihood. We have:

$$L = \sum_{i=1}^{M}(NK_i ln(\pi) + K_i N \ln(\lambda_i) + K_i \ln(|\mathbf{R}|) + \frac{K_i}{\lambda_i} Tr(\widehat{\mathbf{R}}_i \mathbf{R}^{-1})). \quad (5)$$

ML estimates minimize L. Optimization over λ_i leads to:

$$\lambda_i = \frac{Tr(\widehat{\mathbf{R}}_i \mathbf{R}^{-1})}{N}. \quad (6)$$

By substituting λ_i into L, we get:

$$L = \sum_{i=1}^{M}(NK_i ln(\pi) + K_i N \ln\left(\frac{Tr(\widehat{\mathbf{R}}_i \mathbf{R}^{-1})}{N}\right) + K_i \ln(|\mathbf{R}|) + K_i N). \quad (7)$$

To optimize L with respect to $\mathbf{R}$, let us differentiate the above expression:

$$\mathrm{d}L = Tr\left[(K\mathbf{R}^{-1} - \sum_{i=1}^{M} \frac{K_i N}{Tr(\widehat{\mathbf{R}}_i \mathbf{R}^{-1})} \mathbf{R}^{-1}\widehat{\mathbf{R}}_i \mathbf{R}^{-1}) d\mathbf{R}\right]. \quad (8)$$

Cancelling the differential for all $d\mathbf{R}$ leads to the estimate $\widehat{\mathbf{R}}$ of $\mathbf{R}$:

$$\widehat{\mathbf{R}} = \frac{N}{K} \sum_{i=1}^{M} \frac{K_i}{Tr(\widehat{\mathbf{R}}_i \widehat{\mathbf{R}}^{-1})} \widehat{\mathbf{R}}_i \quad (9)$$

A few remarks can be made about the above expression:

- *Remark 1:* (9) doesn't give an explicit solution for $\widehat{\mathbf{R}}$, and one has to resort to an iterative procedure which will be detailed in section 3.1.
- *Remark 2:* As already mentioned in the introduction, if $\widehat{\mathbf{R}}$ is solution, then $\alpha\widehat{\mathbf{R}}$ is also solution: a constraint must be imposed, such as (2).
- *Remark 3:* (9) also gives:
$$\mathbf{R}^{-1/2}\widehat{\mathbf{R}}\mathbf{R}^{-1/2} = \frac{N}{K} \sum_{i=1}^{M} K_i \mathbf{R}^{-1/2}\widehat{\mathbf{R}}_i\mathbf{R}^{-1/2} / Tr\left(\mathbf{R}^{-1/2}\widehat{\mathbf{R}}_i\mathbf{R}^{-1/2}\left(\mathbf{R}^{-1/2}\widehat{\mathbf{R}}\mathbf{R}^{-1/2}\right)^{-1}\right) \quad (10)$$
which means that if $\widehat{\mathbf{R}}$ is solution, $\mathbf{R}^{-1/2}\widehat{\mathbf{R}}\mathbf{R}^{-1/2}$ is solution of a similar expression, where $\widehat{\mathbf{R}}_i$ is replaced with $\mathbf{R}^{-1/2}\widehat{\mathbf{R}}_i\mathbf{R}^{-1/2}$ which follows a complex Wishart law, with K_i degrees of freedom, and expectation $\lambda_i \mathbf{I}$.
- *Remark 4:* Notice also that in the SIRV case ($K_i = 1$) (9) provides the Fixed Point Estimate (see for example [4]).
- *Remark 5:* $\widehat{\mathbf{R}}$ is unchanged in (9) if we replace $\widehat{\mathbf{R}}_i$ by $\frac{1}{\lambda_i}\widehat{\mathbf{R}}_i$ which is $W(K_i, \mathbf{R})$ distributed. Thus, we can assume without loss of generality that $\widehat{\mathbf{R}}_i \sim W(K_i, \mathbf{R})$ in (9): this will be assumed when studying the statistical properties of $\widehat{\mathbf{R}}$.

The unicity of the solution of (9) has been proved for the real data case in [2], and could be extended in the complex case. This solution will be called the Generalized Fixed Point Estimate (GPFE), to distinguish it from the Fixed Point Estimate corresponding to the case $K_i = 1$.

3.1. The Generalized Fixed Point Estimate

The proposed iterative procedure is an alternate maximization algorithm. At iteration k, let $\hat{\lambda}_i^{(k)}$ be the estimate of λ_i and $\widehat{\mathbf{R}}^{(k)}$ the estimate of $\mathbf{R}$. Maximizing the likelihood with respect to λ_i for fixed $\mathbf{R} = \widehat{\mathbf{R}}^{(k)}$ leads to:

$$\hat{\lambda}_i^{(k+1)} = Tr\left(\widehat{\mathbf{R}}_i {\mathbf{R}^{(k)}}^{-1}\right)/N \quad (11)$$

and maximizing the likelihood with respect to $\mathbf{R}$ for fixed $\lambda_i = \hat{\lambda}_i^{(k+1)}$ leads to

$$\widehat{\mathbf{R}}^{(k+1)} = \frac{1}{K} \sum_{i=1}^{M} \frac{K_i}{\hat{\lambda}_i^{(k+1)}} \widehat{\mathbf{R}}_i \quad (12)$$

which can be rewritten into:

$$\widehat{\mathbf{R}}^{(k+1)} = \frac{N}{K} \sum_{i=1}^{M} \frac{K_i}{Tr\left(\widehat{\mathbf{R}}_i {\mathbf{R}^{(k)}}^{-1}\right)} \widehat{\mathbf{R}}_i \ . \quad (13)$$

It should be noted that each iteration increases the likelihood. A proof of convergence in the real data case can be found in [2].

3.2. Statistical properties

From remarks 2 and 3 following equation (9), studying $\widehat{\mathbf{R}}$ statistical properties when $\widehat{\mathbf{R}}_i$ is replaced by $\mathbf{R}^{-1/2}\widehat{\mathbf{R}}_i\mathbf{R}^{-1/2}$ and $\mathbf{R} = \mathbf{I}$ is sufficient. Therefore, whithout loss of generality, we will assume that $\widehat{\mathbf{R}}_i$ in (9) is complex Wishart distributed, with K_i degrees of freedom, and expectation $\lambda_i \mathbf{I}$. In the following subsection, consistency as well as asymptotical distribution are derived, under the following conditions:

$$K_i \to \infty \quad \text{with} \quad \frac{K_i}{K} \to \alpha_i \quad (14)$$

and with the normalization constraint $Tr(\widehat{\mathbf{R}}) = N$.
First, the bias of the estimate is derived.

3.2.1. Bias

Unbiasedness: $\widehat{\mathbf{R}}$ is an unbiased estimate of $\mathbf{R}$.

Proof:

Let $\mathbf{U}$ be a unitary matrix. From (9),

$$\mathbf{U}\widehat{\mathbf{R}}\mathbf{U}^H = \frac{N}{K} \sum \frac{K_i}{Tr(\mathbf{U}\widehat{\mathbf{R}}_i\mathbf{U}^H(\mathbf{U}\widehat{\mathbf{R}}\mathbf{U}^H)^{-1})} \mathbf{U}\widehat{\mathbf{R}}_i\mathbf{U}^H. \quad (15)$$

Since $\mathbf{U}\widehat{\mathbf{R}}_i\mathbf{U}^H$ follows the same Wishart distribution as $\widehat{\mathrm{R}}_i$, $\mathbf{U}\widehat{\mathbf{R}}\mathbf{U}^H$ follows the same distribution as $\widehat{\mathbf{R}}$. Therefore, we have, whatever the unitary matrix is:

$$\begin{aligned} &\forall \mathbf{U} \text{ unitary}, \mathrm{E}(\mathbf{U}\widehat{\mathbf{R}}\mathbf{U}^H) = \mathrm{E}(\widehat{\mathbf{R}}) \\ \Rightarrow\ &\forall \mathbf{U} \text{ unitary}, \mathbf{U}\mathrm{E}(\widehat{\mathbf{R}})\mathbf{U}^H = \mathrm{E}(\widehat{\mathbf{R}}) \end{aligned} \quad (16)$$

Let $\mathrm{E}(\widehat{\mathbf{R}}) = \sum \mu_i \mathbf{u}_i \mathbf{u}_i^H$ be the eigendecomposition of $E(\widehat{\mathbf{R}})$. Suppose that there exist 2 eigenvalues μ_1 and μ_2 such that

$$\mu_1 \neq \mu_2. \quad (17)$$

Let $\mathbf{V}$ be the unitary matrix such that

$$\begin{cases} \mathbf{V}\mathbf{u}_1 = \mathbf{u}_2 \\ \mathbf{V}\mathbf{u}_2 = \mathbf{u}_1 \\ \forall i \neq 1,2, \mathbf{V}\mathbf{u}_i = \mathbf{u}_i \end{cases}. \quad (18)$$

Then, from (16):

$$\begin{aligned} &\mathbf{V}E(\widehat{\mathbf{R}})\mathbf{V}^H = E(\widehat{\mathbf{R}}) \\ \Rightarrow\ &\mu_2\mathbf{u}_1\mathbf{u}_1^H + \mu_1\mathbf{u}_2\mathbf{u}_2^H = \mu_1\mathbf{u}_1\mathbf{u}_1^H + \mu_2\mathbf{u}_2\mathbf{u}_2^H \\ \Rightarrow\ &\mu_2 = \mu_1 \end{aligned} \quad (19)$$

which is in contradiction with (17). Thus:

$$\forall i,j\ \mu_i = \mu_j = \mu \quad (20)$$

And since $Tr(\widehat{\mathbf{R}}) = N$,

$$\mu = 1. \quad (21)$$

Finally, we have:

$$E(\widehat{\mathbf{R}}) = \mathbf{I}. \quad (22)$$

3.2.2. Consistency of the GFPE

Consistency: $\widehat{\mathbf{R}}$ is a consistent estimate of $\mathbf{R}$.

Proof:
To check the consistency of the estimate, we study the limit of (9) in probability when $K_i \to \infty$, with $\frac{K_i}{K} \to \alpha_i$.Let $\widehat{\mathbf{R}}_\infty$be the limit of $\widehat{\mathbf{R}}$ under these conditions.
Since $\mathbf{R}_i \sim W(K_i, \mathbf{R})$ (see remark 5 in Section 3), it follows that:

$$\widehat{\mathbf{R}}_i \xrightarrow{P} \mathbf{I} \quad \text{when} \quad K_i \to \infty. \quad (23)$$

Taking the limit of (9), we obtain:

$$\widehat{\mathbf{R}}_\infty = N \sum_{i=1}^{2} \alpha_i \frac{\mathbf{I}}{Tr\left(\widehat{\mathbf{R}}_\infty{}^{-1}\right)}. \quad (24)$$

Hence, $\widehat{\mathbf{R}}_\infty$ is proportional to $\mathbf{I}$, and the normalization constraint on $\widehat{\mathbf{R}}$ leads to:

$$\widehat{\mathbf{R}}_\infty = \mathbf{I}. \quad (25)$$

3.2.3. Normalized Wishart matrices and their properties

The asymptotic distribution of the GFPE is related to the complex Wishart distribution, as will be seen in the next section. For that purpose, some useful properties of Wishart matrices are presented in this section.

Definition: Let $\widehat{\boldsymbol{\Gamma}} \sim W(K, \boldsymbol{\Gamma})$. The matrix $\widehat{\boldsymbol{\Gamma}}_n = \frac{N}{Tr(\boldsymbol{\Gamma}^{-1}\widehat{\boldsymbol{\Gamma}})}\widehat{\boldsymbol{\Gamma}}$ is said to follow a normalized complex Wishart distribution denoted by $\widehat{\boldsymbol{\Gamma}}_n \sim W_n(K, \boldsymbol{\Gamma})$.

Properties:
Let $\widehat{\boldsymbol{\Gamma}} \sim W(K, \mathbf{I})$, $\Delta\boldsymbol{\Gamma} = \widehat{\boldsymbol{\Gamma}} - \mathbf{I}$, $\Delta\boldsymbol{\Gamma}_n = \widehat{\boldsymbol{\Gamma}}_n - \mathbf{I}$. Then we have the following properties:

$$E(\widehat{\boldsymbol{\Gamma}}_n) = \mathbf{I} \quad (26)$$

$$\widehat{\boldsymbol{\Gamma}}_n \xrightarrow{P} \mathbf{I} \quad \text{when} \quad K \to \infty \quad (27)$$

$$\Delta\boldsymbol{\Gamma}_n = \Delta\boldsymbol{\Gamma} - \frac{1}{N} Tr(\Delta\boldsymbol{\Gamma})\mathbf{I} \quad (28)$$

when limiting to first order with respect to $\Delta\boldsymbol{\Gamma}$.

Proof:
Proof of (26) is similar to the proof of the GFPE unbiasedness.
Proof of (27) and (28) are obvious.

3.2.3. Asymptotic distribution of the GFPE

Asymptotic distribution: the asymptotic distribution of the GFPE $\widehat{\mathbf{R}}$ is $W_n(K, \mathbf{I})$.

Proof:
In this section, a perturbation analysis is performed to derive the asymptotic distribution of the estimate. Let us remind that $\widehat{\mathbf{R}}$ is the solution of (9): $\widehat{\mathbf{R}} = \frac{N}{K}\sum \frac{K_i}{Tr(\widehat{\mathbf{R}}_i\widehat{\mathbf{R}}^{-1})}\widehat{\mathbf{R}}_i$, where matrices $\widehat{\mathbf{R}}_i$ are independent complex Wishart distributed $W(K_i, \mathbf{I})$.

For large K_i's, we have:

$$\widehat{\mathbf{R}} = \mathbf{I} + \Delta\mathbf{R} \quad (29)$$

with $\Delta\mathbf{R}$ small since the estimate is consistent, and:

$$\widehat{\mathbf{R}}_i = \mathbf{I} + \Delta\mathbf{R}_i \quad (30)$$

with $\Delta\mathbf{R}_i$ small for these Wishart matrices. A first order expansion of (9) with respect to $\Delta\mathbf{R}$ and $\Delta\mathbf{R}_i$ can thus be derived:

$$\begin{aligned} \mathbf{I} + \Delta\mathbf{R} &\approx \frac{N}{K}\sum_{i=1}^{M} K_i \frac{\mathbf{I} + \Delta\mathbf{R}_i}{Tr((\mathbf{I} + \Delta\mathbf{R}_i)(\mathbf{I} - \Delta\mathbf{R}))} \\ &\approx \frac{1}{K}\sum_{i=1}^{M} K_i\left(1 - \frac{1}{N}Tr(\Delta\mathbf{R}_i) + \frac{1}{N}Tr(\Delta\mathbf{R})\right)(\mathbf{I} + \Delta\mathbf{R}_i) \ . \\ &\approx \frac{1}{K}\sum_{i=1}^{M} K_i\left(\mathbf{I} + \Delta\mathbf{R}_i - \frac{1}{N}Tr(\Delta\mathbf{R}_i)\mathbf{I} + \frac{1}{N}Tr(\Delta\mathbf{R})\mathbf{I}\right) \end{aligned} \quad (31)$$

Since $\sum_i K_i = K$, (31) can be simplified into:

$$\Delta\mathbf{R} - \frac{1}{N}Tr(\Delta\mathbf{R})\mathbf{I} \approx \frac{1}{K}\sum_{i=1}^{M} K_i\left(\Delta\mathbf{R}_i - \frac{1}{N}Tr(\Delta\mathbf{R}_i)\mathbf{I}\right). \quad (32)$$

Furthermore, $Tr(\widehat{\mathbf{R}}) = N$ implies that $Tr(\Delta\mathbf{R}) = 0$, so that:

$$\Delta\mathbf{R} \approx \frac{1}{K}\sum_{i=1}^{M} K_i\left(\Delta\mathbf{R}_i - \frac{1}{N}Tr(\Delta\mathbf{R}_i)\mathbf{I}\right). \quad (33)$$

Based on Wishart matrices properties, we have:

$$\frac{1}{K}\sum_{i=1}^{M} K_i\widehat{\mathbf{R}}_i = \widehat{\boldsymbol{\Gamma}} \sim W(K, \mathbf{I}) \quad (34)$$

so that:

$$\frac{1}{K}\sum_{i=1}^{M} K_i\Delta\mathbf{R}_i = \Delta\boldsymbol{\Gamma} \quad (35)$$

and thus:

$$\Delta\mathbf{R} = \Delta\boldsymbol{\Gamma} - \frac{1}{N}Tr(\Delta\boldsymbol{\Gamma})\mathbf{I} \quad (36)$$

which is equal to $\Delta\boldsymbol{\Gamma}_n$ in (28). Therefore, $\widehat{\mathbf{R}}$ has the same asymptotical distribution as that of a complex normalized Wishart matrix.

4. SIMULATION

To illustrate the above theoretical results, we consider an application in the context of array processing which is met in radar detection. Let us consider the problem of detecting a known signal $\mathbf{p} \in \mathrm{C}^N$ corrupted by Gaussian clutter, which can be stated as the following binary hypothesis test:

$$\begin{cases} H_0: \quad \mathbf{y} = \mathbf{x}, \quad \mathbf{y}_{i,k} = \mathbf{x}_{i,k} \text{ for} 1 < k < K_i, i = 1,2 \\ H_1: \mathbf{y} = A\mathbf{p} + \mathbf{x}, \mathbf{y}_{i,k} = \mathbf{x}_{i,k} \text{ for} 1 < k < K_i, i = 1,2 \end{cases}$$

where $\mathbf{y}$ is the complex N-vector of the received signal, A is an unknown complex target amplitude, $\mathbf{p}$ stands for the known "steering vector", $\mathbf{x} \sim N(0, \mathbf{R})$ is the clutter in the primary data, and $\mathbf{x}_{i,k} \sim N(0, \lambda_i \mathbf{R})$ are target free secondary data available for clutter covariance estimation. In this model, we assume that the $\mathbf{x}_{i,k}$'s have been gathered in two subsets of size K_i which share the same unknown covariance matrix $\mathbf{R}$ up to an unknown scale factor λ_i.In this context, one can resort to the Adaptive Normalized Matched Filter (ANMF) [8] in which we use the secondary data to estimate the unknown covariance matrix $\mathbf{R}$ by means of our GFPE $\widehat{\mathbf{R}}$ (9) through the iterative procedure (12):

$$T(\widehat{\mathbf{R}}) = \frac{|\mathbf{p}^H \widehat{\mathbf{R}}^{-1} \mathbf{y}|^2}{(\mathbf{p}^H \widehat{\mathbf{R}}^{-1} \mathbf{p})(\mathbf{y}^H \widehat{\mathbf{R}}^{-1} \mathbf{y})} \underset{H_1}{\overset{H_0}{\gtrless}} \eta. \qquad (37)$$

Our results prove that the statistics of $T(\widehat{\mathbf{R}})$ are asymptotically (for large K_1 and K_2) the same as $T(\widehat{\mathbf{\Gamma}}_n)$ where $\widehat{\mathbf{\Gamma}}_n$ is $W_n(K_1 + K_2, \mathbf{R})$, which are also the same as the statistics of $T(\widehat{\mathbf{\Gamma}})$, where $\widehat{\mathbf{\Gamma}}$ is $W(K_1 + K_2, \mathbf{R})$. To verify this, the empirical cumulative distribution of $T(\widehat{\mathbf{R}})$ is compared to the exact distribution of $T(\widehat{\mathbf{\Gamma}})$ given in [8]. We recall that the exact distribution of $T(\widehat{\mathbf{\Gamma}})$ does not depend on $\mathbf{p}$ nor $\mathbf{R}$.

In the simulations, we have $N = 4$, $K_1 = K_2 = 2$. The figure display the threshold-false alarm rate relations for $1/\left(1 - T(\widehat{\mathbf{R}})\right)^N$ (black curve based on 100000 trials) and $1/\left(1 - T(\widehat{\mathbf{\Gamma}})\right)^N$ (dashed red curve) based on the theoretical distribution given in [8]. Despite the extremely small values for K_1 and K_2, we notice a surprising agreement between the two curves: this demonstrates the interest of our asymptotic result which turns out to be valid with only , $K_1 = K_2 = 2$ data.

5. CONCLUSION

In this paper, the problem of estimating proportional covariance matrices was considered. The maximum likelihood estimates of the covariance matrix and the scale factors were derived, as well as their statistical properties. Another interesting result was that the covariance estimation for SIRV model problem is a subcase of the problem treated here. A simulation based on a detection problem is presented, which suggests that the asymptotic distribution derived in this paper is valid, even for a very small data number.

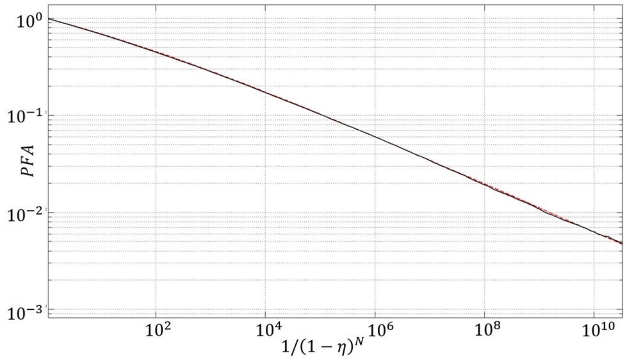

Fig. 1. PFA (vertical axis) evolution with $\mathbf{1}/(\mathbf{1} - \boldsymbol{T})^{\boldsymbol{N}}$ (horizontal axis). Dashed red curve represents the probability of false alarm (ordinate) against $\mathbf{1}/\left(\mathbf{1} - \boldsymbol{T}(\widehat{\mathbf{\Gamma}})\right)^{\boldsymbol{N}}$ (absciss). Black curve represents the probability of false alarm (ordinate) against $1/\left(\mathbf{1} - \boldsymbol{T}(\widehat{\mathbf{R}})\right)^{\boldsymbol{N}}$, based on 100000 trials.

REFERENCES

[1] B. K. Flury, "Proportionality of k covariance matrices", *Statistics & Probability Letters*, Vol. 4, pp. 29-33, January 1986.

[2] P. S. Eriksen, "Proportionality of covariance matrices", *The Annals of Statistics*, Vol. 15, No. 2, pp. 732-748, August 1986.

[3] Y. Chitour, and F. Pascal, "Exact maximum likelihood estimates for SIRV covariance matrix: existence and algorithm analysis", *IEEE Transactions on signal processing*, Vol. 56, No. 10, pp. 4563-4573, October 2008.

[4] L. Bombrun, F. Pascal, J-Y. Tourneret, and Y. Berthoumieu, "Performance of the maximum likelihood estimators for the parameters of multivariate generalized Gaussian distributions", *Proc. ICASSP*, Kyoto, 2012, pp. 3525-3528.

[5] F. Pascal, P. Forster, J-P. Ovarlez, P. Larzabal, "Performance analysis of covariance matrix estimates in impulsive noise", *IEEE Transactions on signal processing*, Vol. 56, No. 6, pp. 4563-4573, June 2008.

[6] F. Gini, M. Greco, "Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter", *Signal Processing*, Vol. 82, pp. 1847-1859, 2002

[7] E. Conte, A. De Maio and G. Ricci, " Recursive estimation of the covariance matrix of a compound-gaussian process and its application to adaptive CFAR Detection", *IEEE Transactions on signal processing*, vol. 50, No.2, pp. 1908-1915, August 2002.

[8] S. Kraut, L.L. Scharf, "Adaptive subspace detectors", *IEEE Transactions on signal processing*, vol. 49, No.1, pp. 1-16, January 2001.