Performance Analysis of Covariance Matrix Estimates in Impulsive Noise

— Source link ↗

Frédéric Pascal, Philippe Forster, J.-P. Ovarlez, Pascal Larzabal

Institutions: Centre national de la recherche scientifique, University of Paris-Sud

Related papers:

- Covariance Structure Maximum-Likelihood Estimates in Compound Gaussian Noise: Existence and Algorithm Analysis

- Covariance matrix estimation for CFAR Detection in correlated heavy tailed clutter

- Recursive estimation of the covariance matrix of a compound-Gaussian process and its application to adaptive CFAR detection

- A Distribution-Free M-Estimator of Multivariate Scatter

- Complex Elliptically Symmetric Distributions: Survey, New Results and Applications

Performance analysis of covariance matrix estimates in impulsive noise

Frédéric Pascal, Philippe Forster, Jean-Philippe Ovarlez, Pascal Larzabal

HAL Id: hal-00353591

https://hal-supelec.archives-ouvertes.fr/hal-00353591

Submitted on 4 Mar 2020

Performance Analysis of Covariance Matrix Estimates in Impulsive Noise

Frédéric Pascal, Philippe Forster , Jean-Philippe Ovarlez and Pascal Larzabal

Abstract

This paper deals with covariance matrix estimates in impulsive noise environments. Physical models based on compound noise modeling (SIRV, Compound Gaussian Processes) allow to correctly describe reality (*e.g.*, range power variations or clutter transitions areas in radar problems). However, these models depend on several unknown parameters (covariance matrix, statistical distribution of the texture, disturbance parameters) which have to be estimated. Based on these noise models, this paper presents a complete analysis of the main covariance matrix estimates used in the literature: four estimates are studied, the well-known Sample Covariance Matrix $\widehat{\mathbf{M}}_{SCM}$ and a normalized version $\widehat{\mathbf{M}}_N$, the Fixed Point (FP) estimate $\widehat{\mathbf{M}}_{FP}$, and a theoretical benchmark $\widehat{\mathbf{M}}_{TFP}$. Among these estimates, the only one of practical interest in impulsive noise is the FP. The three others, which could be used in a Gaussian context, are, in this paper, only of academic interest, *i.e.*, for comparison with the FP. A statistical study of these estimates is performed through bias analysis, consistency, and asymptotic distribution. This study allows to compare the performance of the estimates and to establish simple relationships between them. Finally, theoretical results are emphasized by several simulations corresponding to real situations.

Index Terms

SIRV, covariance matrix estimates, statistical performance analysis, bias, consistency, asymptotic distribution, non-Gaussian noise.

F. Pascal is with SATIE, ENS Cachan, UMR CNRS 8029, 94235 Cachan Cedex, France (e-mail: pascal@satie.ens-cachan.fr).

P. Forster is with the Groupe d'Electromagnétisme Appliqué (GEA), Institut Universitaire de Technologie de Ville d'Avray, 92410 Ville d'Avray, France (e-mail: philippe.forster@cva.u-paris10.fr).

J. P. Ovarlez is with the French Aerospace Lab, ONERA, DEMR/TSI, BP 72, 92322 Chatillon Cedex, France (e-mail: jean-philippe.ovarlez@onera.fr).

P. Larzabal is with the IUT de Cachan, C.R.I.I.P, Université Paris Sud, 94234 Cachan Cedex, France, and also with the SATIE, ENS Cachan, UMR CNRS 8029, 94235 Cachan Cedex, France (e-mail: larzabal@satie.ens-cachan.fr).

I. INTRODUCTION

It is often assumed that signals, interferences or noises are Gaussian stochastic processes. Indeed, this assumption makes sense in many applications. Among them, we can cite: sources localization in passive sonar where signals and noises are generally assumed to be Gaussian, radar detection where thermal noise and clutter are often modeled as Gaussian processes and digital communications where the Gaussian hypothesis is widely used for interferences and noises.

In these contexts, Gaussian models have been thoroughly investigated in the framework of Statistical Estimation and Detection Theory [1], [2], [3]. They have led to attractive algorithms. For instance, we can cite the stochastic Maximum Likelihood method for sources localization in array processing [4], [5], and the matched filter in radar detection [6], [7] and in digital communications [8].

However, such widespread techniques are sub-optimal when the noise process is a non-Gaussian stochastic process [9]. Therefore, non-Gaussian noise modeling has gained many interest in the last decades and currently leads to active researches in the literature. High order moment methods [10] have initiated this research activity and particle filtering methods [11] are now intensively investigated. In radar applications, experimental clutter measurements, performed by MIT [12], showed that these data are not correctly described by Gaussian statistical models. More generally, numerous non-Gaussian models have been developed in several engineering fields. For example, we can cite the K-distribution already used in the area of radar detection [13], [14]. Moreover, let us note that the Weibull distribution is a widely spread model in biostatistics and in radar detection [15].

One of the most general and elegant impulsive noise model is provided by the so-called *Spherically Invariant Random Vectors* (SIRV). Indeed, these processes encompass a large number of non-Gaussian distributions, included, of course, Gaussian processes and also, the aforementioned distributions. SIRV and their variants have been used in various problems such as: bandlimited speech signals [16], radar clutter echoes [17], [18], and wireless radio fading propagation problems [19], [20]. Moreover, SIRVs are also connected to other interesting processes such as the "heavy-tailed" processes, which have been used to model impulse radio noises as well as processes used in financial engineering models [21].

A SIRV is a compound process. It is the product of a Gaussian random process with the square root of a non-negative random scalar variable (called the *texture* in the radar context). Thus, the SIRV is fully

characterized by the texture (representing an unknown power) and the unknown covariance matrix of the zero-mean Gaussian vector. One of the major challenging difficulties in SIRV modeling, is to estimate these two unknown quantities [22], [23], [24]. These problems have been investigated in [25] for the texture estimation while [26] and [27] have proposed different estimates for the covariance matrix. The knowledge of the estimates statistical properties is essential to use them in different contexts.

This paper deals with three covariance matrix estimates: the well-known Sample Covariance Matrix (SCM) [28], the Theoretical Fixed Point (TFP), both studied for academic purposes, and the Fixed Point (FP) which may easily be implemented in practice [29]. These three estimates arise as the solutions of Maximum Likelihood (ML) or Approximate Maximum Likelihood (AML) problems. The main contribution of this paper is to derive and to compare their statistical properties: bias, consistency, second order moment and asymptotical distribution.

The paper is organized as follows. In Section II, a background on the SIRV covariance matrix estimates is given. Sections III, IV and V present the main results of this paper,*i.e.*, performance analysis of the estimates in terms of bias, consistency, covariance matrices, and asymptotic distribution. For clarity, long proofs are reported in Appendices. Finally, Section VI gives some simulation results confirming the theoretical analysis.

II. PROBLEM FORMULATION

In this section, we introduce the SIRV noise model under study and the associated covariance matrix estimates. In the following, H denotes the conjugate transpose operator, $\top$ denotes the transpose operator, E stands for the statistical mean of a random variable, and $\text{Tr}(\mathbf{M})$ is the trace of the matrix $\mathbf{M}$.

A. Statistical Framework

Let us recap some SIRV theory results. A SIRV is a complex compound Gaussian process with random power. More precisely, a SIRV $\mathbf{c}$ [30], [31] is the product between the square root of a positive random variable τ and a m-dimensional independent zero-mean complex Gaussian vector $\mathbf{x}$

$$\mathbf{c} = \sqrt{\tau}\,\mathbf{x}.$$

For identifiability considerations, the covariance matrix $\mathbf{M} = E\left[\mathbf{x}\mathbf{x}^H\right]$ is normalized according to $\text{Tr}(\mathbf{M}) = m$ (see [26]) and called, in the sequel, normalization #1.

The SIRV Probability Density Function (PDF) expression is

$$p(\mathbf{c}) = \frac{1}{\pi^m \, |\mathbf{M}|} \int_0^{+\infty} \frac{1}{\tau^m} \exp\left(-\frac{\mathbf{c}^H \, \mathbf{M}^{-1} \, \mathbf{c}}{\tau}\right) q(\tau) \, \mathrm{d}\tau \, ,$$

where q is the texture PDF.

Notice that, when τ is Dirac distributed, *i.e.*, $q(\tau) = \delta(\tau - \tau_0)$, the resulting SIRV is a zero-mean Gaussian vector with covariance matrix $\tau_0 \, \mathbf{M}$ while when τ is Gamma distributed, the resulting SIRV is the well known K-distribution. However, the closed-form expression of the texture PDF q is not always available (*e.g.*, Weibull SIRV). Thus, in problems where $\mathbf{M}$ is unknown and has to be estimated from SIRV realizations, it would be of interest to have a covariance matrix structure estimate independent of the texture.

B. Covariance Matrix Structure Estimation

The covariance matrix has to be normalized to identify the SIRV noise model. Consequently, it is reasonable to think that the same normalization has to be applied to its estimate. However, as it will be shown later, the appropriate normalization for performance analysis is $\mathrm{Tr}(\mathbf{M}^{-1}\widehat{\mathbf{M}}) = m$ and will be called normalization #2 for any estimate $\widehat{\mathbf{M}}$ of $\mathbf{M}$:

$$\begin{cases} \text{Normalization \#1:} & \mathrm{Tr}(\mathbf{M}) = m \, , \\ \text{Normalization \#2:} & \mathrm{Tr}(\mathbf{M}^{-1}\widehat{\mathbf{M}}) = m \quad \text{or, equivalently,} \quad \mathrm{Tr}(\mathbf{M}^{-1/2}\widehat{\mathbf{M}}\mathbf{M}^{-1/2}) = m \, . \end{cases} \quad (1)$$

When $\mathbf{M}$ is unknown, it could be objected that normalization #2 is only of theoretical interest while only normalization #1 can be applied in practice. In most applications, however, $\widehat{\mathbf{M}}$ is exploited in such a way that any scale factor on $\widehat{\mathbf{M}}$ has no influence on the final result. (*e.g.*, in radar detection, the detector could be a likelihood ratio, homogeneous in terms of $\widehat{\mathbf{M}}$ [26]). This would also be the case in general estimation problems where estimated parameters only depend on the structure of the covariance matrix. Hence, the normalization, chosen for any studied case, is of little importance.

In this framework, the three estimates will be built from N independent realizations of $\mathbf{c}$ denoted $\mathbf{c}_1 = \sqrt{\tau_1} \, \mathbf{x}_1, \dots, \mathbf{c}_N = \sqrt{\tau_N} \, \mathbf{x}_N$ and called secondary data.

First, if we had access to the N independent realizations $\mathbf{x}_1, \dots, \mathbf{x}_N$ of the underlying Gaussian vector $\mathbf{x}$, the ML estimate would lead to the Sample Covariance Matrix (SCM) which is Wishart distributed

[33] and defined by

$$\widehat{\mathbf{M}}_{SCM} = \frac{1}{N} \sum_{k=1}^{N} \mathbf{x}_k \mathbf{x}_k^H . \tag{2}$$

But, in practice, we only have access to the N independent realizations of a SIRV, $\mathbf{c}_1, \ldots, \mathbf{c}_N$ and it is impossible to isolate the Gaussian process $\mathbf{x}$. However, it is used as a benchmark for comparison with the other estimates. Moreover, $\widehat{\mathbf{M}}_{SCM}$ performance analysis will lead to an interpretation of the theoretical expressions obtained for the other estimates.

To fulfill normalization #2, $\widehat{\mathbf{M}}_N$ will be defined by

$$\widehat{\mathbf{M}}_N = \frac{m}{\text{Tr}(\mathbf{M}^{-1}\widehat{\mathbf{M}}_{SCM})} \widehat{\mathbf{M}}_{SCM} , \tag{3}$$

Eq. (3) can also be written

$$\widehat{\mathbf{M}}_N = \frac{m}{\displaystyle\sum_{i=1}^{N} \mathbf{x}_i^H \mathbf{M}^{-1} \mathbf{x}_i} \sum_{k=1}^{N} \mathbf{x}_k \mathbf{x}_k^H . \tag{4}$$

Estimates (2) and (3) have only a theoretical interest since the $\mathbf{x}_k$'s are not available. Practical estimates are functions of the $\mathbf{c}_k$'s and "good" ones should not depend on the τ_k's.

A first candidate is the Normalized Sample Covariance Matrix (NSCM) [34] given by

$$\widehat{\mathbf{M}}_{NSCM} = \frac{m}{N} \sum_{k=1}^{N} \frac{\mathbf{c}_k \mathbf{c}_k^H}{\mathbf{c}_k^H \mathbf{c}_k},$$

which can be rewritten only in terms of $\mathbf{x}_k$'s

$$\widehat{\mathbf{M}}_{NSCM} = \frac{m}{N} \sum_{k=1}^{N} \frac{\mathbf{x}_k \mathbf{x}_k^H}{\mathbf{x}_k^H \mathbf{x}_k},$$

As the statistical performance of this estimate have been extensively studied in [35], only its statistical analysis results will be presented in order to compare all available estimates. Moreover, although this "heuristic" estimate respects normalization #1, it exhibits several severe drawbacks.

A second candidate, provided by the ML theory [26], [27], is the FP estimate $\widehat{\mathbf{M}}_{FP}$ of $\mathbf{M}$, defined as a fixed point of function $f_{N,\mathbf{M}}$

$$f_{N,\mathbf{M}} : \begin{cases} \mathcal{D} & \longrightarrow \quad \mathcal{D} \\[2mm] \mathbf{A} & \longrightarrow \quad \dfrac{m}{N} \displaystyle\sum_{k=1}^{N} \dfrac{\mathbf{c}_k \mathbf{c}_k^H}{\mathbf{c}_k^H \mathbf{A}^{-1} \mathbf{c}_k} \end{cases} \tag{5}$$

where $\mathcal{D} = \{\mathbf{A} \in M_m(\mathbb{C}) | \mathbf{A}^H = \mathbf{A}, \mathbf{A} \text{ positive definite}\}$ with $M_m(\mathbb{C}) = \{ m \times m \text{ matrices with elements in } \mathbb{C}\}$.

The notation $f_{N,\mathbf{M}}$ stresses the dependency on N and on the covariance matrix $\mathbf{M}$ involved in the $\mathbf{c}_k$'s.

As shown in [36], equation $\widehat{\mathbf{M}} = f_{N,\mathbf{M}}(\widehat{\mathbf{M}})$ has a solution of the form $\alpha \mathbf{M}$, where α is an arbitrary scaling factor. In this paper, the only solution $\widehat{\mathbf{M}}_{FP}$ satisfying normalization #2 is called the Fixed Point estimate. In other words, $\widehat{\mathbf{M}}_{FP}$ is the unique solution of

$$\widehat{\mathbf{M}}_{FP} = f_{N,\mathbf{M}}(\widehat{\mathbf{M}}_{FP}) = \frac{m}{N} \sum_{k=1}^{N} \frac{\mathbf{c}_k \, \mathbf{c}_k^H}{\mathbf{c}_k^H \, \widehat{\mathbf{M}}_{FP}^{-1} \, \mathbf{c}_k} = \frac{m}{N} \sum_{k=1}^{N} \frac{\mathbf{x}_k \, \mathbf{x}_k^H}{\mathbf{x}_k^H \, \widehat{\mathbf{M}}_{FP}^{-1} \, \mathbf{x}_k}, \tag{6}$$

such that $\widehat{\mathbf{M}}_{FP}$ respects normalization #2.

Notice that $\widehat{\mathbf{M}}_{FP}$, as $\widehat{\mathbf{M}}_{NSCM}$, does not depend on the texture τ as emphasized in Eq. (6).

Remark II.1

The FP is the ML estimate when the texture τ is assumed to be an unknown deterministic parameter, and is an Approximate Maximum Likelihood (AML) when τ is assumed to be a positive random variable [26], [27].

Finally, an analysis of another texture-independent estimate $\widehat{\mathbf{M}}_{TFP}$ of $\mathbf{M}$ is performed, where

$$\widehat{\mathbf{M}}_{TFP} = \frac{m}{N} \sum_{k=1}^{N} \frac{\mathbf{c}_k \, \mathbf{c}_k^H}{\mathbf{c}_k^H \, \mathbf{M}^{-1} \, \mathbf{c}_k} = \frac{m}{N} \sum_{k=1}^{N} \frac{\mathbf{x}_k \, \mathbf{x}_k^H}{\mathbf{x}_k^H \, \mathbf{M}^{-1} \, \mathbf{x}_k}. \tag{7}$$

In this paper, it is called the Theoretical Fixed Point (TFP) estimate. This estimate is only of theoretical interest since it depends on the unknown covariance matrix $\mathbf{M}$. It is closely related to the FP estimate (6) and it will be shown that its statistical performance have strong connections with those of $\widehat{\mathbf{M}}_{FP}$. Notice that $\widehat{\mathbf{M}}_{TFP}$ satisfies normalization #2: $\text{Tr}(\mathbf{M}^{-1}\widehat{\mathbf{M}}_{TFP}) = m$. So, the TFP estimate will be considered as the benchmark for $\widehat{\mathbf{M}}_{FP}$.

In this context of covariance matrix estimation in impulsive noise, the statistical properties of the three proposed estimates, $\widehat{\mathbf{M}}_N$, $\widehat{\mathbf{M}}_{TFP}$ and $\widehat{\mathbf{M}}_{FP}$ will be established in this paper, while existing results concerning $\widehat{\mathbf{M}}_{SCM}$ and $\widehat{\mathbf{M}}_{NSCM}$ will be reminded.

III. BIAS ANALYSIS

This section provides an analysis of the bias B denoted by

$$B(\widehat{\mathbf{M}}) = E\left[\widehat{\mathbf{M}}\right] - \mathbf{M},$$

for each estimate $\widehat{\mathbf{M}}$ introduced previously.

The SCM $\widehat{\mathbf{M}}_{SCM}$ has been studied in literature and is unbiased. Now, $\widehat{\mathbf{M}}_N$ bias will be analysed next. Its unbiasedness is presented in the following theorem

Theorem III.1 (Unbiasedness of $\widehat{\mathbf{M}}_N$)

$\widehat{\mathbf{M}}_N$ *is an unbiased estimate of* $\mathbf{M}$.

Proof: To prove that $E\left[\widehat{\mathbf{M}}_N\right] = \mathbf{M}$, the focus is put on

$$E\left[\widehat{\mathbf{M}}_N\right] = m \sum_{k=1}^{N} E\left[\frac{\mathbf{x}_k \mathbf{x}_k^H}{\displaystyle\sum_{i=1}^{N} \mathbf{x}_i^H \mathbf{M}^{-1} \mathbf{x}_i}\right]. \tag{8}$$

For this purpose, let us whiten the $\mathbf{x}_k$'s in Eq. (8), according to $\mathbf{y}_k = \mathbf{M}^{-1/2} \mathbf{x}_k$

$$E\left[\widehat{\mathbf{M}}_N\right] = m \, \mathbf{M}^{1/2} \left(\sum_{k=1}^{N} E\left[\frac{\mathbf{y}_k \mathbf{y}_k^H}{\displaystyle\sum_{i=1}^{N} \mathbf{y}_i^H \mathbf{y}_i}\right]\right) \mathbf{M}^{1/2}. \tag{9}$$

For $k = 1, \ldots, N$, $\mathbf{A}_k$ are given by

$$\mathbf{A}_k = \frac{\mathbf{y}_k \mathbf{y}_k^H}{\displaystyle\sum_{i=1}^{N} \mathbf{y}_i^H \mathbf{y}_i}. \tag{10}$$

Since $\mathbf{y}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ for any k, the jth element of $\mathbf{y}_k$ denoted by $y_k^{(j)}$ can be written as

$$y_k^{(j)} = \sqrt{(1/2) \, \chi_k^{2(j)}} \, \exp(i\theta_k^{(j)}),$$

where for $1 \le k \le N$ and $1 \le j \le m$, $\chi_k^{2(j)} \sim \chi^2(2)$ and $\theta_k^{(j)} \sim \mathcal{U}([0, 2\pi])$, where $\chi^2(2)$ denotes the Chi-squared distribution with 2 degrees of freedom and $\mathcal{U}([0, 2\pi])$, the uniform distribution on interval $[0, 2\pi]$, and $\chi_k^{2(j)}$ and $\theta_k^{(j)}$ are independent.

Thus, by replacing the $\mathbf{y}_k$'s in Eq. (10), element pq of matrix $\mathbf{A}_k$ is

$$A_k^{(pq)} = \frac{\sqrt{\chi_k^{2(p)} \chi_k^{2(q)}}}{\displaystyle\sum_{i=1}^{N} \sum_{r=1}^{m} \chi_i^{2(r)}} \exp(i(\theta_k^{(p)} - \theta_k^{(q)})) \,.$$

Since $E\left[\exp(i\theta_k^{(p)})\right] = 0$, for any (k, p), the non diagonal elements of $E\left[\mathbf{A}_k\right]$ are null. Then, diagonal element pp is

$$E\left[A_k^{(pp)}\right] = E\left[\frac{\chi_k^{2(p)}}{\displaystyle\sum_{i=1}^{N} \sum_{r=1}^{m} \chi_i^{2(r)}}\right] = E\left[\frac{\chi_k^{2(p)}}{\chi_k^{2(p)} + \displaystyle\sum_{\substack{i=1 \\ i \neq k}}^{N} \sum_{\substack{r=1 \\ r \neq p}}^{m} \chi_i^{2(r)}}\right],$$

where $\dfrac{\chi_k^{2(p)}}{\chi_k^{2(p)} + \displaystyle\sum_{\substack{i=1 \\ i \neq k}}^{N} \sum_{\substack{r=1 \\ r \neq p}}^{m} \chi_i^{2(r)}}$ is a Beta of the first kind distributed random variable with parameters 2

and $2Nm - 2$. Moreover, the statistical mean of a $\beta(2, 2Nm - 2)$ is $\dfrac{2}{2 + 2Nm - 2} = \dfrac{1}{Nm}$, thus

$E\left[A_k^{(pp)}\right] = \dfrac{1}{Nm}$ for $k = 1, \ldots, N$ and for $p = 1, \ldots, m$ and $E\left[\mathbf{A}_k\right] = (1/Nm)\,\mathbf{I}$. Using these results in Eq. (9) leads to

$$E\left[\widehat{\mathbf{M}}_N\right] = m\,\mathbf{M}^{1/2} \left(\sum_{k=1}^{N} (1/Nm)\,\mathbf{I}\right) \mathbf{M}^{1/2},$$

proving that $\widehat{\mathbf{M}}_N$ is unbiased

$$E\left[\widehat{\mathbf{M}}_N\right] = \mathbf{M},$$

∎

The analysis of $\widehat{\mathbf{M}}_{NSCM}$ bias is provided in two cases. In the most general case, theorem III.2 gives a closed form expression of $\widehat{\mathbf{M}}_{NSCM}$ bias. Then theorem III.3 proves the unbiasedness of $\widehat{\mathbf{M}}_{NSCM}$ when $\mathbf{M} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix.

Theorem III.2 ($\widehat{\mathbf{M}}_{NSCM}$ bias when M has distinct eigenvalues)

Assuming that $\mathbf{M}$ has distinct eigenvalues, $\widehat{\mathbf{M}}_{NSCM}$ bias is given by

$$B(\widehat{\mathbf{M}}_{NSCM}) = \mathbf{U}\left(diag\left(\mathbf{E}\,\mathbf{c}\right)\right)\mathbf{U}^{H} - \mathbf{M},$$

where

- *the operator diag reshapes a m-vector* $\mathbf{v} = (v_j)_{1 \leq j \leq m}$ *into a $m \times m$ diagonal matrix* $\mathbf{V} = (V_{ij})_{1 \leq i,j \leq m}$ *with elements $V_{ii} = v_i$,*

- $\mathbf{U}$ *is the orthogonal matrix of the m eigenvectors of* $\mathbf{M}$,

- $\mathbf{E} = (E_{ij})_{1 \leq i,j \leq m}$ *with* $E_{ij} = \dfrac{\log{(\lambda_j/\lambda_i)}}{\lambda_j/\lambda_i - 1} - \dfrac{\lambda_i}{\lambda_j}$ *if $i \neq j$ and $E_{ij} = 0$ if $i = j$, where λ_i is the ith eigenvalue of* $\mathbf{M}$,

- $\mathbf{c} = (c_j)_{1 \leq j \leq m}$ *with* $c_j = m \prod\limits_{k \neq j} \dfrac{1}{1 - \lambda_k/\lambda_j}$.

Proof: see [35]. ∎

Theorem III.3 (Unbiasedness of $\widehat{\mathbf{M}}_{NSCM}$ when $\mathbf{M} = \mathbf{I}$)

$\widehat{\mathbf{M}}_{NSCM}$ *is an unbiased estimate of* $\mathbf{I}$

Proof: With the same reasoning as in theorem III.1's proof, it is shown that $E\left[\widehat{\mathbf{M}}_{NSCM}\right]$ is a diagonal matrix with elements

$$A^{(kk)} = mE\left[\chi^{2(k)} / \sum_{j=1}^{m} \chi^{2(j)}\right] = mE\left[\chi^{2(k)} / \left(\chi^{2(k)} + \sum_{\substack{j=1 \\ j \neq k}}^{m} \chi^{2(j)}\right)\right],$$

where $\chi^{2(k)} / \left(\chi^{2(k)} + \sum\limits_{\substack{j=1 \\ j \neq k}}^{m} \chi^{2(j)}\right)$ is a Beta of the first kind distributed random variable with parameters

2 and $2m - 2$. Since the statistical mean of a $\beta(2, 2m - 2)$ is $\dfrac{2}{2 + 2m - 2} = \dfrac{1}{m}$, $A^{(kk)} = 1$, which completes the proof. ∎

Theorem III.4 (Unbiasedness of $\widehat{\mathbf{M}}_{FP}$)

$\widehat{\mathbf{M}}_{FP}$ *is an unbiased estimate of* $\mathbf{M}$

Proof: For clarity, in this section $\widehat{\mathbf{M}}_{FP}$ will be denoted $\widehat{\mathbf{M}}$. The first part of the proof is the whitening of the data. By applying the following change of variable, $\mathbf{y}_k = \mathbf{M}^{-1/2}\mathbf{x}_k$ to Eq. (6), one has

$$\widehat{\mathbf{M}}_{FP} = \frac{m}{N}\sum_{k=1}^{N} \frac{\mathbf{M}^{1/2}\mathbf{y}_k\mathbf{y}_k^H\mathbf{M}^{1/2}}{\mathbf{y}_k^H\widehat{\mathbf{T}}^{-1}\mathbf{y}_k},$$

where

$$\widehat{\mathbf{T}} = \mathbf{M}^{-1/2}\widehat{\mathbf{M}}_{FP}\mathbf{M}^{-1/2}.$$

Therefore,

$$\widehat{\mathbf{T}} = \frac{m}{N}\sum_{k=1}^{N} \frac{\mathbf{y}_k\mathbf{y}_k^H}{\mathbf{y}_k^H\widehat{\mathbf{T}}^{-1}\mathbf{y}_k}.$$

$\widehat{\mathbf{T}}$ is thus the unique FP estimate (up to a scaling factor) of the identity matrix. Its statistics are clearly independent of $\mathbf{M}$ since the $\mathbf{y}_k$'s are $\mathcal{N}(\mathbf{0}, \mathbf{I})$.

Moreover, for any unitary matrix $\mathbf{U}$,

$$\mathbf{U}\,\widehat{\mathbf{T}}\,\mathbf{U}^H = \frac{m}{N}\sum_{k=1}^{N} \frac{\mathbf{z}_k\,\mathbf{z}_k^H}{\mathbf{z}_k^H\left(\mathbf{U}\,\widehat{\mathbf{T}}\,\mathbf{U}^H\right)^{-1}\mathbf{z}_k},$$

where $\mathbf{z}_k = \mathbf{U}\,\mathbf{y}_k$ are also i.i.d. $\mathcal{N}(\mathbf{0}, \mathbf{I})$. Therefore, $\mathbf{U}\,\widehat{\mathbf{T}}\,\mathbf{U}^H$ has the same distribution as $\widehat{\mathbf{T}}$, so

$$E\left[\widehat{\mathbf{T}}\right] = \mathbf{U}E\left[\widehat{\mathbf{T}}\right]\mathbf{U}^H,\ \text{for any unitary matrix}\ \mathbf{U}.$$

Since $E\left[\widehat{\mathbf{T}}\right]$ is different from $\mathbf{0}$, Lemma A.1, detailed in the Appendix, ensures that $E\left[\widehat{\mathbf{T}}\right] = \alpha\mathbf{I}$. Moreover, since $\widehat{\mathbf{T}} = \mathbf{M}^{-1/2}\widehat{\mathbf{M}}\mathbf{M}^{-1/2}$, thus $E\left[\widehat{\mathbf{M}}\right] = \alpha\mathbf{M}$ also since $\mathrm{Tr}(\mathbf{M}^{-1}\widehat{\mathbf{M}}) = m$, $\alpha = 1$.

In conclusion: $\widehat{\mathbf{M}}$ is an unbiased estimate of $\mathbf{M}$, for any integer N. ∎

Theorem III.5 (Unbiasedness of $\widehat{\mathbf{M}}_{TFP}$)

$\widehat{\mathbf{M}}_{TFP}$ *is an unbiased estimate of* $\mathbf{M}$

Proof: To prove Theorem III.5, changing variable $\mathbf{x} = \mathbf{M}^{1/2}\,\mathbf{y}$, where $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, leads to

$$E\left[\widehat{\mathbf{M}}_{TFP}\right] = \mathbf{M}^{1/2}\,E\left[m\,\frac{\mathbf{y}\,\mathbf{y}^H}{\mathbf{y}^H\,\mathbf{y}}\right]\mathbf{M}^{1/2},$$

The equality

$$E\left[m\,\frac{\mathbf{y}\,\mathbf{y}^H}{\mathbf{y}^H\,\mathbf{y}}\right] = \mathbf{I},$$

is proven by Theorem III.3.

Thus, $E\left[\widehat{\mathbf{M}}_{TFP}\right] = \mathbf{M}$, *i.e.*, $\widehat{\mathbf{M}}_{TFP}$ is an unbiased estimate of $\mathbf{M}$. ∎

IV. CONSISTENCY

An estimate $\widehat{\mathbf{M}}$ of $\mathbf{M}$ is consistent if

$$\forall \varepsilon > 0,\, Pr(\|\widehat{\mathbf{M}} - \mathbf{M}\| \geq \varepsilon)\xrightarrow[N\to+\infty]{} 0,$$

where N is the number of secondary data $\mathbf{c}_k$'s used to estimate $\mathbf{M}$ and $\|.\|$ stands for any matrix norm.

Remark IV.1

When $\mathbf{M}$ has distinct eigenvalues, Theorem III.2 shows that $\widehat{\mathbf{M}}_{NSCM}$ is a biased estimate of $\mathbf{M}$. Moreover, this bias does not depend on the number N of $\mathbf{x}_k$'s. Thus, $\widehat{\mathbf{M}}_{NSCM}$ is not a consistent estimate of $\mathbf{M}$. In the sequel, since $\widehat{\mathbf{M}}_{NSCM}$ suffers from the previous drawbacks (bias and non consistent), thus from here on,

this estimate will not be taken into account when $\mathbf{M}$ *has distinct eigenvalues. On the other hand, the NSCM estimate of* $\mathbf{I}$ *will be studied as a particular case of the TFP estimate.*

Under Gaussian assumptions, the SCM estimate is consistent. This result was established in [28] pp. 80-81.

Theorem IV.1 ($\widehat{\mathbf{M}}_N$ consistency)

$\widehat{\mathbf{M}}_N$ *is a consistent estimate of* $\mathbf{M}$ *.*

Proof: By whitening the $\mathbf{x}_k$'s in Eq. 4:

$$\widehat{\mathbf{M}}_N = m\,\mathbf{M}^{1/2} \left(\frac{\displaystyle\sum_{k=1}^{N} \mathbf{y}_k\mathbf{y}_k^H}{\displaystyle\sum_{i=1}^{N} \mathbf{y}_i^H\mathbf{y}_i} \right) \mathbf{M}^{1/2},$$

where $\mathbf{y}_k \sim \mathcal{CN}(\mathbf{0},\mathbf{I})$.

Then, the Weak Law of Large Numbers (WLLN) demonstrates that

$$\frac{1}{N}\sum_{k=1}^{N}\mathbf{y}_k\mathbf{y}_k^H \quad \xrightarrow[N\to+\infty]{Pr} \quad \mathbf{I},$$

$$\frac{1}{N}\sum_{i=1}^{N}\mathbf{y}_i^H\mathbf{y}_i \quad \xrightarrow[N\to+\infty]{Pr} \quad E\left[\mathbf{y}^H\mathbf{y}\right] = m.$$

Finally, basic theorems on stochastic convergence show that

$$\widehat{\mathbf{M}}_N \xrightarrow[N\to+\infty]{Pr} \mathbf{M},$$

which means that $\widehat{\mathbf{M}}_N$ is a consistent estimate of $\mathbf{M}$. ∎

Theorem IV.2 ($\widehat{\mathbf{M}}_{FP}$ consistency)

$\widehat{\mathbf{M}}_{FP}$ *is a consistent estimate of* $\mathbf{M}$ *.*

Proof: See Appendix B. ∎

Theorem IV.3 ($\widehat{\mathbf{M}}_{TFP}$ consistency)

$\widehat{\mathbf{M}}_{TFP}$ *is a consistent estimate of* $\mathbf{M}$ *.*

Proof: Theorem III.1 and the WLLN imply that $\widehat{\mathbf{M}}_{TFP}$ is a consistent estimate of $\mathbf{M}$. ∎

V. ASYMPTOTIC DISTRIBUTION

A. Notations

In this section, a perturbation analysis will be used to derive the asymptotic distribution of the three estimates, all denoted $\widehat{\mathbf{M}}$ for clarity. For this purpose, $\widehat{\mathbf{M}}$ is rewritten as

$$\widehat{\mathbf{M}} = \mathbf{M} + \delta\mathbf{M}.$$

The essential quantities for the analysis are defined as follows:

- $\boldsymbol{\Delta} = \mathbf{M}^{-1/2} \left(\delta\mathbf{M}\right) \mathbf{M}^{-1/2} = \mathbf{M}^{-1/2} \widehat{\mathbf{M}} \mathbf{M}^{-1/2} - \mathbf{I}$,

- $\boldsymbol{\delta} = \text{vec}(\boldsymbol{\Delta})$ where $\boldsymbol{\delta}$ is the vector containing all the elements of $\boldsymbol{\Delta}$ and vec denotes the operator which reshapes the $m \times n$ matrix elements into a mn column vector.

In the sequel, these quantities will be indexed according to the studied estimate: $\boldsymbol{\delta}_{SCM}, \boldsymbol{\delta}_N, \boldsymbol{\delta}_{TFP}$ and $\boldsymbol{\delta}_{FP}$.

The asymptotic distribution of $\widehat{\mathbf{M}}$ is obtained from the distribution of $\boldsymbol{\delta}$ with the following Proposition:

Proposition V.1

$$vec(\widehat{\mathbf{M}}) = \left(\mathbf{M}^{1/2} \otimes \mathbf{M}^{\top/2}\right) \boldsymbol{\delta} + vec(\mathbf{M}),$$

where $\otimes$ represents the Kronecker product.

Proof: $\widehat{\mathbf{M}} = \mathbf{M}^{1/2} \boldsymbol{\Delta} \mathbf{M}^{1/2} + \mathbf{M}$. This is proven by the property, $\text{vec}(\mathbf{EFG}) = \left(\mathbf{E} \otimes \mathbf{G}^{\top}\right) \text{vec}(\mathbf{F})$ for any matrix $\mathbf{E}, \mathbf{F}, \mathbf{G}$, (see [33] p.9). ∎

The aim of this section is to derive the asymptotic distribution of $\boldsymbol{\delta}$, *i.e.*, the asymptotic distribution of $\begin{pmatrix} \text{Re}(\boldsymbol{\delta}) \\ \text{Im}(\boldsymbol{\delta}) \end{pmatrix}$, where $\text{Re}(\boldsymbol{\delta})$ denotes the real part of the complex vector $\boldsymbol{\delta}$ and $\text{Im}(\boldsymbol{\delta})$ its imaginary part. It will be shown later that this distribution is Gaussian and therefore is fully characterized by its asymptotic covariance matrix. This matrix may be derived from the two quantities $E\left[\boldsymbol{\delta}\boldsymbol{\delta}^{\top}\right]$ and $E\left[\boldsymbol{\delta}\boldsymbol{\delta}^{H}\right]$. However, in this specific case, $\boldsymbol{\delta}$ is the vec of an Hermitian matrix so $E\left[\boldsymbol{\delta}\boldsymbol{\delta}^{\top}\right]$ can be obtained from $E\left[\boldsymbol{\delta}\boldsymbol{\delta}^{H}\right]$.

B. Results

The following original results use the notations $\mathbf{C}_1$ and $\mathbf{C}_2$, defined by

$$\mathbf{C}_1 \;=\; \frac{m}{m+1}\left(\mathbf{P} - \frac{1}{m}\,\text{vec}(\mathbf{I})\text{vec}(\mathbf{I})^\top\right), \tag{11}$$

$$\mathbf{C}_2 \;=\; \frac{m}{m+1}\left(\mathbf{I} - \frac{1}{m}\,\text{vec}(\mathbf{I})\text{vec}(\mathbf{I})^\top\right), \tag{12}$$

where $\mathbf{P}$ is defined by, for $1 \le p, p' \le m$,

$$\begin{cases} P_{kl} = 1 & \text{for} \quad k = l = p + m(p-1), \\ P_{kl} = 1 & \text{for} \quad k = p + m(p'-1) \text{ and } l = p' + m(p-1), \\ P_{kl} = 0 & \text{else}. \end{cases} \tag{13}$$

The covariance matrix of $\text{vec}(\widehat{\mathbf{M}}_{SCM})$ has been established in [28] and is reminded here below.

Theorem V.1 ($\text{vec}(\widehat{\mathbf{M}}_{SCM})$ Asymptotic Distribution)

1) $\sqrt{N}\begin{pmatrix} Re(\boldsymbol{\delta}_{SCM}) \\ Im(\boldsymbol{\delta}_{SCM}) \end{pmatrix} \xrightarrow[N\to+\infty]{dist.} \mathcal{N}(\mathbf{0}, \mathbf{C}_{SCM})$, where $\xrightarrow{dist.}$ stands for the convergence in distribu-

tion and $\mathbf{C}_{SCM}$ denotes the covariance matrix of $\begin{pmatrix} Re(\boldsymbol{\delta}_{SCM}) \\ Im(\boldsymbol{\delta}_{SCM}) \end{pmatrix}$ which can be straightforwardly

obtained from:

2) $N\,E\left[\boldsymbol{\delta}_{SCM}\boldsymbol{\delta}_{SCM}^\top\right] \xrightarrow[N\to+\infty]{} \mathbf{P}$ where $\mathbf{P}$ is defined above;

3) $N\,E\left[\boldsymbol{\delta}_{SCM}\boldsymbol{\delta}_{SCM}^H\right] \xrightarrow[N\to+\infty]{} \mathbf{I}$.

   Proof: See [33].

   ■

Theorem V.2 ($\text{vec}(\widehat{\mathbf{M}}_N)$ Asymptotic Distribution)

1) $\sqrt{N}\begin{pmatrix} Re(\boldsymbol{\delta}_N) \\ Im(\boldsymbol{\delta}_N) \end{pmatrix} \xrightarrow[N\to+\infty]{dist.} \mathcal{N}(\mathbf{0}, \mathbf{C}_N)$, where $\mathbf{C}_N$ denotes the covariance matrix and can be straight-
forwardly obtained from:

2) $N\,E\left[\boldsymbol{\delta}_N\boldsymbol{\delta}_N^\top\right] \xrightarrow[N\to+\infty]{} \left(\frac{m+1}{m}\right)\mathbf{C}_1$;

3) $N E \left[\boldsymbol{\delta}_N \boldsymbol{\delta}_N^H \right] \xrightarrow[N \to +\infty]{} \left(\dfrac{m+1}{m} \right) \mathbf{C}_2 .$

The two matrices $\mathbf{C}_1$ and $\mathbf{C}_2$ are defined by Eqs. (11) and (12).

Proof: See Appendix C.

∎

Theorem V.3 ($\text{vec}(\widehat{\mathbf{M}}_{TFP})$ Asymptotic Distribution)

1) $\sqrt{N} \begin{pmatrix} Re(\boldsymbol{\delta}_{TFP}) \\ Im(\boldsymbol{\delta}_{TFP}) \end{pmatrix} \xrightarrow[N \to +\infty]{dist.} \mathcal{N}(\mathbf{0}, \mathbf{C}_{TFP}), $ *where $\mathbf{C}_{TFP}$ denotes the covariance matrix and can be straightforwardly obtained from:*

2) $N E \left[\boldsymbol{\delta}_{TFP} \boldsymbol{\delta}_{TFP}^\top \right] \xrightarrow[N \to +\infty]{} \mathbf{C}_1 ;$

3) $N E \left[\boldsymbol{\delta}_{TFP} \boldsymbol{\delta}_{TFP}^H \right] \xrightarrow[N \to +\infty]{} \mathbf{C}_2 .$

Proof: See Appendix D.

∎

In [37], the following original results on $\widehat{\mathbf{M}}_{FP}$ with the above notations has been partially established.

Theorem V.4 ($\text{vec}(\widehat{\mathbf{M}}_{FP})$ Asymptotic Distribution)

1) $\sqrt{N} \begin{pmatrix} Re(\boldsymbol{\delta}_{FP}) \\ Im(\boldsymbol{\delta}_{FP}) \end{pmatrix} \xrightarrow[N \to +\infty]{dist.} \mathcal{N}(\mathbf{0}, \mathbf{C}_{FP}), $ *where $\mathbf{C}_{FP}$ denotes the covariance matrix and can be straightforwardly obtained from:*

2) $N E \left[\boldsymbol{\delta}_{FP} \boldsymbol{\delta}_{FP}^\top \right] \xrightarrow[N \to +\infty]{} \left(\dfrac{m+1}{m} \right)^2 \mathbf{C}_1 ;$

3) $N E \left[\boldsymbol{\delta}_{FP} \boldsymbol{\delta}_{FP}^H \right] \xrightarrow[N \to +\infty]{} \left(\dfrac{m+1}{m} \right)^2 \mathbf{C}_2 .$

Proof: Proof of theorem V.4 is fully given in Appendix E.

∎

C. Synthesis

All the results on the asymptotic second order moment of $\widehat{\mathbf{M}}$ are recapped in table I:

- The three estimates $\widehat{\mathbf{M}}_{FP}$, $\widehat{\mathbf{M}}_{TFP}$ and $\widehat{\mathbf{M}}_N$, *share the same asymptotic covariance matrix up to scaling factors.*

TABLE I

ASYMPTOTIC SECOND ORDER MOMENT

	$\widehat{\mathbf{M}}_{FP}$	$\widehat{\mathbf{M}}_{TFP}$	$\widehat{\mathbf{M}}_{SCM}$	$\widehat{\mathbf{M}}_N$
Limit of $NE\left[\boldsymbol{\delta\delta}^{\top}\right]$	$\left(\frac{m+1}{m}\right)^2\mathbf{C}_1$	$\mathbf{C}_1$	$\mathbf{P}$	$\frac{m+1}{m}\,\mathbf{C}_1$
Limit of $NE\left[\boldsymbol{\delta\delta}^{H}\right]$	$\left(\frac{m+1}{m}\right)^2\mathbf{C}_2$	$\mathbf{C}_2$	$\mathbf{I}$	$\frac{m+1}{m}\,\mathbf{C}_2$

- More precisely, $\sqrt{\left(\frac{m}{m+1}\right)^2 N}\,\boldsymbol{\delta}_{FP}$, $\sqrt{\frac{m}{m+1}N}\,\boldsymbol{\delta}_N$ and $\sqrt{N}\,\boldsymbol{\delta}_{TFP}$ have the same asymptotic distribution. Therefore $\widehat{\mathbf{M}}_{FP}$, with N secondary data, has the same asymptotic behavior as $\widehat{\mathbf{M}}_N$, with $\frac{m}{m+1}\,N$ secondary data. Since $\widehat{\mathbf{M}}_N$ is the SCM up to a scale factor, we may conclude that, in problems invariant with respect to a scale factor on the covariance matrix, the FP estimate is asymptotically equivalent to the SCM with a little less secondary data: $\frac{m+1}{m}\,N$ data.

VI. SIMULATIONS

In order to enlighten results provided in sections III, IV and V, some simulation results are presented. Since $\widehat{\mathbf{M}}_{FP}$ and $\widehat{\mathbf{M}}_{TFP}$ are texture independent and $\widehat{\mathbf{M}}_{SCM}$ and $\widehat{\mathbf{M}}_N$ are only valid under Gaussian assumption, simulations are performed with Gaussian noise.

Operator $\overline{A}$ is defined as the empirical mean of the quantities $A(i)$ obtained from I Monte Carlo runs. For each iteration i, a new set of N secondary data $\mathbf{x}_1,\ldots,\mathbf{x}_N$ is generated to compute $A(i)$.

Thus, for example, $\overline{\widehat{\mathbf{M}}}$ is defined by

$$\overline{\widehat{\mathbf{M}}} = \frac{1}{I}\sum_{i=1}^{I}\widehat{\mathbf{M}}(i)\,.$$

A. Bias analysis

The results presented in this section are obtained for complex Gaussian zero-mean data with covariance matrix $\mathbf{M}$ defined by

$$M_{ij} = \rho^{|i-j|}\,,\text{ for } 1 \le i,j \le m\,.$$

Figure 1 shows the bias of each estimate for different values of ρ: $\rho = 0.1$, 0.5 and 0.9. The length of each vector $\mathbf{x}_k$ is $m = 3$.

For that purpose, a plot of $C(\widehat{\mathbf{M}}, N) = \left\| \overline{\widehat{\mathbf{M}}} - \mathbf{M} \right\|$ versus the number N of $\mathbf{x}_k$'s and for any matrix norm $\|.\|$ is presented for each estimate.

It can be noticed that, as enlightened by theoretical analysis, the bias of $\widehat{\mathbf{M}}_N$, $\widehat{\mathbf{M}}_{TFP}$ and $\widehat{\mathbf{M}}_{FP}$ is very close to $\mathbf{0}$ for each value of N, while the bias of $\widehat{\mathbf{M}}_{NSCM}$ (when the covariance matrix $\mathbf{M}$ is different from $\mathbf{I}$) does not tend to zero with the number N of $\mathbf{x}_k$'s. Moreover, this simulation underlines the fact that $\widehat{\mathbf{M}}_{NSCM}$ bias does not depend on N, it is constant for all number N of $\mathbf{x}_k$'s.

Furthermore, the correlation coefficient ρ has of course no influence on the unbiased estimates. On the other hand, for each value of ρ, the bias of the NSCM estimate is different and it tends to 0 when ρ tends to 0 (case $\mathbf{M}$ tends to $\mathbf{I}$).

B. Consistency analysis

Figure 2 presents results of estimates consistency. For that purpose, a plot of $D(\widehat{\mathbf{M}}, N) = \left\| \widehat{\mathbf{M}} - \mathbf{M} \right\|$ versus the number N of $\mathbf{x}_k$'s is presented for each estimate.

It can be noticed that the above criterion $D(\widehat{\mathbf{M}}, N)$ tends to 0 when N tends to $+\infty$, for each estimate and each value of ρ. However, there are more fluctuations when data are strongly correlated (*i.e.*, $\rho \simeq 1$).

C. Second order moment analysis

Simulations relating to second order moment have been represented on two different graphics: one on transpose operator $\top$ results and the other for the transpose conjugate operator H results.

In figure 3, the quantity $S_1(\widehat{\mathbf{M}}, N)$

$$S_1(\widehat{\mathbf{M}}, N) = \left\| N \, \overline{\boldsymbol{\delta}\boldsymbol{\delta}^\top} - \mathbf{Q} \right\| ,$$

is plotted for the four studied estimates versus the number N of $\mathbf{x}_k$'s, where $\boldsymbol{\delta}$ is defined in the notations used in section V and the matrix $\mathbf{Q}$ represents the closed form expression of $N \, E\left[\boldsymbol{\delta}\boldsymbol{\delta}^\top \right]$ for the different estimates.

Let us recall these results:

- for $\widehat{\mathbf{M}}_{FP}$, $\mathbf{Q} = \left(\dfrac{m+1}{m}\right)^2 \mathbf{C}_1$,
- for $\widehat{\mathbf{M}}_{NSCM}$, $\mathbf{Q} = \mathbf{C}_1$,
- for $\widehat{\mathbf{M}}_{SCM}$, $\mathbf{Q} = \mathbf{P}$,
- for $\widehat{\mathbf{M}}_N$, $\mathbf{Q} = \left(\dfrac{m+1}{m}\right) \mathbf{C}_1$,

where $\mathbf{C}_1$ is defined by (11) and $\mathbf{P}$ by (13).

Figure 3.a validates results on second order moment estimates obtained in section V because the quantity $S_1(\widehat{\mathbf{M}}, N)$ tends to zero when N tends to infinity for all estimates.

From Figure 3.b, the same conclusion as Figure 3.a is drawn but with the transpose conjugate operator H, *i.e.*,

$$S_2(\widehat{\mathbf{M}}, N) = \left\| N\, \overline{\boldsymbol{\delta}\boldsymbol{\delta}^H} - \mathbf{R} \right\|,$$

where the matrix $\mathbf{R}$ represents the closed form expression of $N\, E\left[\boldsymbol{\delta}\boldsymbol{\delta}^H\right]$ for the different estimates.

VII. Conclusion

In this paper, the problem of covariance matrix estimation in impulsive noise modeled by Spherically Invariant Random Vectors was considered. Four estimates, $\widehat{\mathbf{M}}_{SCM}$, $\widehat{\mathbf{M}}_N$, $\widehat{\mathbf{M}}_{FP}$ and $\widehat{\mathbf{M}}_{TFP}$ have been studied through a theoretical statistical analysis : bias, consistency, asymptotic distribution and second order moments. All original results have been illustrated by simulations.

In this impulsive noise context, the Sample Covariance Matrix cannot be used in practice since this estimate of the Gaussian kernel is based on unavailable data. The same remark holds for the theoretical fixed point estimate as it depends on the unknown covariance matrix that needs to be estimated. Finally the well known Normalized Sample Covariance Matrix is biased and not consistent. Therefore the only appropriate estimate is the so called Fixed Point estimate, which is unbiased, consistent and has, up to a scale factor, the same second order moments as the sample covariance matrix of the Gaussian kernel.

Finally, this statistical study will allow a performance analysis of signal processing algorithms based on these estimates. For instance performance of radar detection algorithms using the Fixed Point estimate will be investigated in future work.

APPENDIX A

LEMMA A.1

Lemma A.1

Let **A** *denote a Hermitian matrix and let* **U** *stand for any unitary matrix, then*

$$\mathbf{A} = \mathbf{UAU}^H \,, \, \forall \mathbf{U} \Leftrightarrow \mathbf{A} = \alpha \mathbf{I}, \, \alpha \in \mathbb{R}$$

Proof:

- If $\mathbf{A} = \alpha \mathbf{I}$ then $\mathbf{A} = \mathbf{UAU}^H$.

- Now, assume that for a diagonalizable matrix **A** and for any unitary matrix **U**, $\mathbf{A} = \mathbf{UAU}^H$. Let **V** be the matrix of the eigenvectors of **A** and Λ the diagonal matrix of the eigenvalues of **A**, then $\mathbf{A} = \mathbf{V} \Lambda \mathbf{V}^H$.

  If $\mathbf{U} = \mathbf{V}^H$, one has $\mathbf{VDV}^H = \mathbf{D}$. This implies that **A** is a diagonal matrix. By taking for **U** the permutation matrix which reshapes elements the ith element of **A**, A_{ii} into the $(i+1)^{th}$ elements $A_{i+1,i+1}$, this leads to the conclusion.

∎

APPENDIX B

PROOF OF THEOREM IV.2 (CONSISTENCY OF $\widehat{\mathbf{M}}_{FP}$)

To show the consistency of $\widehat{\mathbf{M}}_{FP}$, denoted $\widehat{\mathbf{M}}(N)$ in the sequel and to show the dependence between $\widehat{\mathbf{M}}_{FP}$ and the number N of $\mathbf{x}'_k$s, several properties of the function $f_{N,\mathbf{M}}$ defined by (5) will be used. First, a new function g_N is defined by

$$g_N : \begin{cases} \mathcal{D} & \longrightarrow & \mathcal{D} \\ \mathbf{A} & \longrightarrow & g_N = A - f_{N,\mathbf{M}}(A) \end{cases}$$

where $\mathcal{D} = \{\mathbf{A} \in M_m(\mathbb{C}) | \, \mathbf{A}^H = \mathbf{A} \,, \mathbf{A} \text{ positive definite}\}$ with $M_m(\mathbb{C}) = \{ m \times m \text{ matrices with elements in } \mathbb{C}\}$, and $\mathbb{C}$, the set of complex scalar.

As $\widehat{\mathbf{M}}(N)$ is a fixed point of function $f_{N,\mathbf{M}}$, it is the unique zero, up to a scaling factor, of the random function g_N. To show the consistency of $\widehat{\mathbf{M}}(N)$, Theorem 5.9 pp. 46 of [38] will be used. Let us verify hypothesis of this theorem.

First, the Strong Law of Large Numbers (SLLN) gives

$$\forall \mathbf{A} \in \mathcal{D} \,, \, g_N(\mathbf{A}) \xrightarrow[N \to +\infty]{a.s.} g(\mathbf{A}) \,,$$

where $\xrightarrow[N \to +\infty]{a.s.}$ stands for the convergence almost surely and

$$\forall \mathbf{A} \in \mathcal{D}, \; g(\mathbf{A}) = \mathbf{A} - m \, E \left[\frac{\mathbf{x}\mathbf{x}^H}{\mathbf{x}^H \mathbf{A}^{-1} \mathbf{x}} \right],$$

for $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{M})$.

Then g defined by Eq.(**??**) is rewritten by applying an appropriate change of variable on $\mathbf{x}$. Let $\mathbf{y} = \mathbf{A}^{-1/2}\mathbf{x}$, where $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \mathbf{A}^{-1/2}\mathbf{M}\mathbf{A}^{-1/2})$ and thus

$$\forall \mathbf{A} \in \mathcal{D}, \; g(\mathbf{A}) = \mathbf{A}^{1/2} \left(\mathbf{I} - m \, E \left[\frac{\mathbf{y}\mathbf{y}^H}{\mathbf{y}^H\mathbf{y}} \right] \right) \mathbf{A}^{1/2}.$$

By using the same change of variable, g_N is rewritten

$$\forall \mathbf{A} \in \mathcal{D}, \; g_N(\mathbf{A}) = \mathbf{A}^{1/2} \left(\mathbf{I} - \frac{m}{N} \sum_{k=1}^{N} \frac{\mathbf{y}_k\mathbf{y}_k^H}{\mathbf{y}_k^H\mathbf{y}_i} \right) \mathbf{A}^{1/2}.$$

Now, it must be shown that for every $\varepsilon > 0$,

$$(H1): \quad \sup_{\mathbf{A} \in \mathcal{D}} \{\|g_N(\mathbf{A}) - g(\mathbf{A})\|\} \xrightarrow[N \to +\infty]{Pr} 0,$$

$$(H2): \quad \inf_{\mathbf{A}:\|\mathbf{A}-\mathbf{M}\| \geq \varepsilon} \{\|g(\mathbf{A})\|\} > 0 = g(\mathbf{M}).$$

Now, for every $\mathbf{A} \in \mathcal{D}$

$$\|g_N(\mathbf{A}) - g(\mathbf{A})\| = m \, \|A\| \, \left\| \frac{1}{N} \sum_{k=1}^{N} \left(\frac{\mathbf{y}_k\mathbf{y}_k^H}{\mathbf{y}_k^H\mathbf{y}_k} - E\left[\frac{\mathbf{y}\mathbf{y}^H}{\mathbf{y}^H\mathbf{y}} \right] \right) \right\|,$$

and thus, the SLLN, applied to the N i.i.d variables $\mathbf{Y}_k = \dfrac{\mathbf{y}_k\mathbf{y}_k^H}{\mathbf{y}_k^H\mathbf{y}_k}$, with same first order moment, ensures $(H1)$.

Now, to show $(H2)$, it suffices to use the bias of $\widehat{\mathbf{M}}_{NSCM}$ shown by Theorem III.2. Indeed, for every $\mathbf{A} \in \mathcal{D}$, with $\mathbf{A} \neq \mathbf{M}$,

$$\|g(\mathbf{A})\| > 0. \tag{14}$$

Eq.(14) is explained by

$$\|g(\mathbf{A})\| = \|B(\widehat{\mathbf{N}}_{NSCM,\mathbf{A}^{-1/2}\mathbf{M}\mathbf{A}^{-1/2}})\|,$$

where $\widehat{\mathbf{N}}_{NSCM,\mathbf{A}^{-1/2}\mathbf{M}\mathbf{A}^{-1/2}}$ is the NSCM estimate of $\mathbf{A}^{-1/2}\mathbf{M}\mathbf{A}^{-1/2}$ and $B(\widehat{\mathbf{N}}_{NSCM,\mathbf{A}^{-1/2}\mathbf{M}\mathbf{A}^{-1/2}})$ the bias of $\widehat{\mathbf{N}}_{NSCM,\mathbf{A}^{-1/2}\mathbf{M}\mathbf{A}^{-1/2}}$ defined by Theorem III.2.

Finally, Theorem 5.9 pp. 46 of [38] concludes the proof and $\widehat{\mathbf{M}}_{FP} \xrightarrow[N \to +\infty]{Pr} \mathbf{M}$, which is the definition of the consistency of $\widehat{\mathbf{M}}_{FP}$.

$$\boldsymbol{\Delta}_N = \mathbf{M}^{-1/2}\,\widehat{\mathbf{M}}_N\,\mathbf{M}^{-1/2} - \mathbf{I}.$$

By using the definition of $\widehat{\mathbf{M}}_N$, Eq. (3), one obtains

$$\boldsymbol{\Delta}_N = \frac{m}{\mathrm{Tr}(\boldsymbol{\Delta}_{SCM} + \mathbf{I})}\,(\boldsymbol{\Delta}_{SCM} + \mathbf{I}) - \mathbf{I} = \frac{1}{1 + \dfrac{\mathrm{Tr}(\boldsymbol{\Delta}_{SCM})}{m}}\,(\boldsymbol{\Delta}_{SCM} + \mathbf{I}) - \mathbf{I}.$$

It is supposed that N is large enough to ensure the validity of the first order expressions, with respect to $\boldsymbol{\Delta}_{SCM}$, and thus

$$\boldsymbol{\Delta}_N \simeq \left(1 - \frac{\mathrm{Tr}(\boldsymbol{\Delta}_{SCM})}{m}\right)\,(\boldsymbol{\Delta}_{SCM} + \mathbf{I}) - \mathbf{I},$$

and by omitting the second order term, *i.e.*, $\boldsymbol{\Delta}_{SCM}\,\dfrac{\mathrm{Tr}(\boldsymbol{\Delta}_{SCM})}{m}$, previous equation becomes

$$\boldsymbol{\Delta}_N \simeq \boldsymbol{\Delta}_{SCM} - \frac{\mathrm{Tr}(\boldsymbol{\Delta}_{SCM})}{m}\mathbf{I}.$$

Now, with the notations presented at the beginning of section V, it comes

$$\boldsymbol{\delta}_N \simeq \boldsymbol{\delta}_{SCM} - \frac{\mathrm{Tr}(\boldsymbol{\Delta}_{SCM})}{m}\,\mathrm{vec}(\mathbf{I}),$$

Moreover, from the expression of $\mathbf{C}_2$ in Eq. (12) and by noting that $\mathrm{Tr}(\mathbf{A}) = \mathrm{vec}(\mathbf{I})^{\top}\mathrm{vec}(\mathbf{A})$ for any matrix $\mathbf{A}$, one has

$$\boldsymbol{\delta}_N \simeq \frac{m+1}{m}\mathbf{C}_2\,\boldsymbol{\delta}_{SCM}. \tag{15}$$

Since $\sqrt{N}\,\boldsymbol{\delta}_{SCM}$ is asymptotically Gaussian, Eq. (15) ensures that the same result holds for $\sqrt{N}\,\boldsymbol{\delta}_N$.

It just remains to derive the asymptotic behavior of the two quantities $NE\left[\boldsymbol{\delta}_N\boldsymbol{\delta}_N^{\top}\right]$ and $NE\left[\boldsymbol{\delta}_N\boldsymbol{\delta}_N^{H}\right]$. These limits follow from the results concerning $\boldsymbol{\delta}_{SCM}$ stated in table I:

$$\begin{cases} N\,E\left[\boldsymbol{\delta}_N\boldsymbol{\delta}_N^{\top}\right] & \xrightarrow[N\to+\infty]{} & \left(\frac{m+1}{m}\right)\mathbf{C}_1, \\ N\,E\left[\boldsymbol{\delta}_N\boldsymbol{\delta}_N^{H}\right] & \xrightarrow[N\to+\infty]{} & \left(\frac{m+1}{m}\right)\mathbf{C}_2, \end{cases}$$

where matrices $\mathbf{C}_1$ and $\mathbf{C}_2$ are defined by Eqs. (11-12).

APPENDIX D

PROOF OF THEOREM V.3 (VEC($\widehat{\mathbf{M}}_{TFP}$) ASYMPTOTIC DISTRIBUTION)

After the whitening of the $\mathbf{x}_k$'s, $\widehat{\mathbf{M}}_{TFP}$ defined by Eq. (7) becomes

$$\widehat{\mathbf{M}}_{TFP} = \mathbf{M}^{1/2} \left(\frac{m}{N} \sum_{k=1}^{N} \frac{\mathbf{y}_k \mathbf{y}_k^H}{\mathbf{y}_k^H \mathbf{y}_k} \right) \mathbf{M}^{1/2} .$$

where $\mathbf{y}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$. Now, in terms of $\boldsymbol{\delta}_{TFP}$, it comes

$$\boldsymbol{\delta}_{TFP} = \text{vec}\left(\mathbf{M}^{-1/2} \widehat{\mathbf{M}}_{TFP} \mathbf{M}^{-1/2} - \mathbf{I} \right) = \text{vec}\left(\frac{m}{N} \sum_{k=1}^{N} \left(\frac{\mathbf{y}_k \mathbf{y}_k^H}{\mathbf{y}_k^H \mathbf{y}_k} \right) - \mathbf{I} \right). \qquad (16)$$

The Central Limit Theorem (CLT) ensures the first point of Theorem V.3

$$\sqrt{N} \left[\begin{array}{c} \text{Re}(\boldsymbol{\delta}_{TFP}) \\ \text{Im}(\boldsymbol{\delta}_{TFP}) \end{array} \right] \xrightarrow[N \to +\infty]{dist.} \mathcal{N}\left(\mathbf{0}, \mathbf{C}_{TFP} \right),$$

where $\mathbf{C}_{TFP}$ is the covariance matrix of $\left[\begin{array}{c} \text{Re}(\boldsymbol{\delta}_{TFP}) \\ \text{Im}(\boldsymbol{\delta}_{TFP}) \end{array} \right]$.

Now, it just remains to derive the two quantities $E\left[\boldsymbol{\delta}_{TFP} \boldsymbol{\delta}_{TFP}^{\top} \right]$ and $E\left[\boldsymbol{\delta}_{TFP} \boldsymbol{\delta}_{TFP}^{H} \right]$.

One has, for a large N,

$$E\left[\boldsymbol{\delta}_{TFP} \boldsymbol{\delta}_{TFP}^{\top} \right] = \frac{m^2}{N} E\left[\text{vec}\left(\frac{\mathbf{y}\mathbf{y}^H}{\mathbf{y}^H\mathbf{y}} \right) \text{vec}\left(\frac{\mathbf{y}\mathbf{y}^H}{\mathbf{y}^H\mathbf{y}} \right)^{\top} \right] - \text{vec}(\mathbf{I}) \text{vec}(\mathbf{I})^{\top}, \qquad (17)$$

where $\mathbf{y} = \left(y^{(1)}, \ldots, y^{(m)} \right)^{\top} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$.

Then, focusing on the following variable

$$\mathbf{B} = E\left[\text{vec}\left(\frac{\mathbf{y}\mathbf{y}^H}{\mathbf{y}^H\mathbf{y}} \right) \text{vec}\left(\frac{\mathbf{y}\mathbf{y}^H}{\mathbf{y}^H\mathbf{y}} \right)^{\top} \right].$$

and rewriting the $y^{(j)}$'s as $y^{(j)} = \sqrt{\chi^{2(j)}/2} \exp(i\theta^{(j)})$ where for $j = 1, \ldots, m$, $\chi^{2(j)}$ and $\theta^{(j)}$ are independent variables, with $\chi^{2(j)} \sim \chi^2(2)$ and $\theta^{(j)} \sim \mathcal{U}([0, 2\pi])$, each element of matrix $\mathbf{B}$ becomes

$$B_{kl} = E\left[\frac{\sqrt{\chi^{2(p)} \chi^{2(q)} \chi^{2(p')} \chi^{2(q')}}}{\left(\sum_{j=1}^{m} \chi^{2(j)} \right)^2} \right] E\left[\exp(i(\theta^{(p)} - \theta^{(q)} + \theta^{(q')} - \theta^{(p')})) \right].$$

Now, it comes that $B_{kl} = 0$, except for the following indexes

1) $k = l = p + m(p - 1)$,

2) $k = p + m(p - 1), l = p' + m(p' - 1)$ and $p \neq p'$,

3) $k = p + m(p' - 1), l = p' + m(p - 1)$ and $p \neq p'$,

and in these cases, one has

1) $B_{p+m(p-1),p+m(p-1)} = \dfrac{2}{m(m+1)}$,

2) $B_{p+m(p-1),p'+m(p'-1)} = \dfrac{1}{m(m+1)}$,

3) $B_{p+m(p'-1),p'+m(p-1)} = \dfrac{1}{m(m+1)}$.

By replacing these results in Eq. (17), the following result is found

$$N E \left[\boldsymbol{\delta}_{TFP} \boldsymbol{\delta}_{TFP}^{\top} \right] \xrightarrow[N \to +\infty]{} \mathbf{C}_1 \,.$$

In the same way,

$$N E \left[\boldsymbol{\delta}_{TFP} \boldsymbol{\delta}_{TFP}^{H} \right] \xrightarrow[N \to +\infty]{} \mathbf{C}_2 \,.$$

APPENDIX E

PROOF OF THEOREM V.4 (VEC($\widehat{\mathbf{M}}_{FP}$) ASYMPTOTIC DISTRIBUTION)

First, $\widehat{\mathbf{M}}$ is written as $\widehat{\mathbf{M}} = \mathbf{M} + \delta\mathbf{M}$ where $\widehat{\mathbf{M}} = \widehat{\mathbf{M}}_{FP}$ in this subsection. For large N, $\delta\mathbf{M} \simeq \mathbf{0}$ because of the $\widehat{\mathbf{M}}$ consistency and N is assumed to be large enough to ensure the validity of the first order expressions, thus

$$\widehat{\mathbf{M}}^{-1} \simeq \left(\mathbf{M}^{-1} - \mathbf{M}^{-1}(\delta\mathbf{M})\mathbf{M}^{-1} \right) \,.$$

For N large enough, this implies that

$$\widehat{\mathbf{M}} \simeq \frac{m}{N} \sum_{k=1}^{N} \frac{\mathbf{x}_k \mathbf{x}_k^{H}}{\mathbf{x}_k^{H} \left(\mathbf{M}^{-1} - \mathbf{M}^{-1}(\delta\mathbf{M})\mathbf{M}^{-1} \right) \mathbf{x}_k},$$

and thus

$$\delta\mathbf{M} \simeq \frac{m}{N} \sum_{k=1}^{N} \left(\frac{\mathbf{x}_k \mathbf{x}_k^{H}}{\mathbf{x}_k^{H} \left(\mathbf{M}^{-1} - \mathbf{M}^{-1}(\delta\mathbf{M})\mathbf{M}^{-1} \right) \mathbf{x}_k} \right) - \mathbf{M} \,.$$

Let $\mathbf{y}_k = \mathbf{M}^{-1/2}\mathbf{x}_k$, then

$$\mathbf{M}^{-1/2}(\delta\mathbf{M})\mathbf{M}^{-1/2} \simeq \frac{m}{N} \sum_{k=1}^{N} \frac{\mathbf{y}_k \mathbf{y}_k^{H}}{\mathbf{y}_k^{H} \left(\mathbf{I} - \mathbf{M}^{-1/2}(\delta\mathbf{M})\mathbf{M}^{-1/2} \right) \mathbf{y}_k} - \mathbf{I},$$

or equivalently, by using expression $\boldsymbol{\Delta} = \mathbf{M}^{-1/2}(\delta\mathbf{M})\mathbf{M}^{-1/2}$,

$$\boldsymbol{\Delta} \simeq \frac{m}{N} \sum_{k=1}^{N} \frac{\mathbf{y}_k \mathbf{y}_k^H}{\mathbf{y}_k^H \mathbf{y}_k \left(1 - \dfrac{\mathbf{y}_k^H \boldsymbol{\Delta} \mathbf{y}_k}{\mathbf{y}_k^H \mathbf{y}_k}\right)} - \mathbf{I}.$$

At the first order, for large N and consequently small $\boldsymbol{\Delta}$, it results in

$$\boldsymbol{\Delta} \simeq \frac{m}{N} \sum_{k=1}^{N} \left(\frac{\mathbf{y}_k \mathbf{y}_k^H}{\mathbf{y}_k^H \mathbf{y}_k} \left(1 + \frac{\mathbf{y}_k^H \boldsymbol{\Delta} \mathbf{y}_k}{\mathbf{y}_k^H \mathbf{y}_k}\right)\right) - \mathbf{I}.$$

To find the explicit expression of $\boldsymbol{\Delta}$ in terms of data, the above expression can be reorganized as

$$\boldsymbol{\Delta} - \frac{m}{N} \sum_{k=1}^{N} \left(\frac{\mathbf{y}_k \mathbf{y}_k^H}{\mathbf{y}_k^H \mathbf{y}_k} \frac{\mathbf{y}_k^H \boldsymbol{\Delta} \mathbf{y}_k}{\mathbf{y}_k^H \mathbf{y}_k}\right) \simeq \frac{m}{N} \sum_{i=1}^{N} \left(\frac{\mathbf{y}_k \mathbf{y}_k^H}{\mathbf{y}_k^H \mathbf{y}_k}\right) - \mathbf{I}.$$

To solve this m^2-system, above equation is rewritten as

$$\mathbf{B}\,\boldsymbol{\delta}_{FP} \simeq \mathrm{vec}\left(\frac{m}{N} \sum_{i=1}^{N} \left(\frac{\mathbf{y}_k \mathbf{y}_k^H}{\mathbf{y}_k^H \mathbf{y}_k}\right) - \mathbf{I}\right), \tag{18}$$

where

- $\boldsymbol{\delta}_{FP} = \mathrm{vec}(\boldsymbol{\Delta})$,
- $\mathbf{D}_k$ is the $m^2 \times m^2$ matrix defined by $\mathbf{D}_k = (d_{ln})_{1 \leq l, n \leq m^2}^{(k)}$ with $d_{ln} = y^p \, \overline{y}^q \, \overline{y}^{p'} \, y^{q'}$ for $l = p + m(q-1)$ and $n = p' + m(q'-1)$, and

$$\mathbf{B} = \mathbf{I} - \frac{m}{N} \sum_{i=1}^{N} \frac{\mathbf{D}_k}{(\mathbf{y}_k^H \mathbf{y}_k)^2}. \tag{19}$$

From Eq. (16), the right-hand side member of Eq. (18) is seen to be equal to $\boldsymbol{\delta}_{TFP}$. Therefore,

$$\mathbf{B}\,\boldsymbol{\delta}_{FP} \simeq \boldsymbol{\delta}_{TFP}. \tag{20}$$

Now, normalization #2 for $\widehat{\mathbf{M}}$ ($\mathrm{Tr}(\mathbf{M}^{-1}\widehat{\mathbf{M}}) = m$) ensures that $\mathrm{Tr}(\boldsymbol{\Delta}) = 0$ which is equivalent to $\mathrm{vec}(\mathbf{I})^\top \boldsymbol{\delta}_{FP} = 0$.

Thus, Eq. (20) may be rewritten as

$$\left(\mathbf{B} + \frac{1}{m+1} \mathrm{vec}(\mathbf{I})\mathrm{vec}(\mathbf{I})^\top\right) \boldsymbol{\delta}_{FP} \simeq \boldsymbol{\delta}_{TFP},$$

and thus

$$\boldsymbol{\delta}_{FP} \simeq \mathbf{G}^{-1}\,\boldsymbol{\delta}_{TFP}, \tag{21}$$

where

$$\mathbf{G} = \mathbf{B} + \frac{1}{m+1} \text{vec}(\mathbf{I})\text{vec}(\mathbf{I})^\top \, , . \tag{22}$$

From the SLLN, $\mathbf{B}$ defined by Eq. (19) satisfies

$$\mathbf{B} \xrightarrow[N \to +\infty]{a.s.} \mathbf{C}_2 = \mathbf{I} - m \, E\left[\frac{\mathbf{D}}{(\mathbf{y}^H \mathbf{y})^2}\right] \, , \tag{23}$$

where $\mathbf{y} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and $\mathbf{D} = (d_{ln})_{1 \le l, n \le m^2}$, with $d_{ln} = y^p \, \overline{y}^q \, \overline{y}^{p'} \, y^{q'}$.

Thus, from standard probability convergence considerations, the first point of Theorem V.4 is obtained.

Now, from (23), one has

$$\mathbf{C}_2 = \mathbf{I} - m \, \mathbf{E} \, , \text{ where } \mathbf{E} = E\left[\frac{\mathbf{D}}{(\mathbf{y}^H \mathbf{y})^2}\right] \, .$$

In the same way as in the proof of theorem V.2, the non-zero elements of the matrix $\mathbf{E}$ are

1) $E_{p+m(p-1),p+m(p-1)} = \dfrac{2}{m(m+1)}$,

2) $E_{p+m(p-1),p'+m(p'-1)} = \dfrac{1}{m(m+1)}$,

3) $E_{p+m(q-1),p+m(q-1)} = \dfrac{1}{m(m+1)}$,

and thus

$$\mathbf{C}_2 = \frac{m}{m+1}\left(\mathbf{I} - \frac{1}{m} \, \text{vec}(\mathbf{I})\text{vec}(\mathbf{I})^\top\right) \, .$$

Therefore, $\mathbf{G}$ in Eq. (22) satisfies

$$\mathbf{G} \xrightarrow[N \to +\infty]{a.s.} \frac{m}{m+1}\mathbf{I}$$

And, it follows from Eq. (21) that $\sqrt{N}\,\boldsymbol{\delta}_{FP}$ has the same asymptotic distribution as $\dfrac{m+1}{m}\sqrt{N}\,\boldsymbol{\delta}_{TFP}$.

REFERENCES

[1] B. Widrow et al., "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter", *Proc. IEEE*, vol. 64, pp. 1151-1162, Aug. 1976.

[2] H.L. Van Trees, "Detection, Estimation and Modulation Theory", Part I, II and III, *John Wiley & Sons*, New York, 1971.

[3] L. Scharf and D.W. Lytle, "Signal Detection in Gaussian Noise of Unknown Level: an Invariance Application", *IEEE Trans.-IT*, Vol.17, pp.404-411, Jul. 1971.

[4] S. Haykin, ed., "Array Signal Processing", *Prentice-Hall Signal Processing Series*, Englewood Cliffs, New Jersey, 1985.

[5] H.L. Van Trees, "Optimum Array Processing", Part IV of "Detection, Estimation and Modulation Theory", *John Wiley & Sons*, New York, 2002.

[6] E.J. Kelly, "An Adaptive Detection Algorithm", *IEEE Trans.-AES*, vol. 23, no. 1, pp. 115-127, Nov. 1986.

[7] S. Kraut , L.L. Scharf and L.T. Mc Whorter, "Adaptive Subspace Detectors", *IEEE Trans.-SP*, vol. 49, no. 1, pp. 1-16, Jan. 2001.

[8] J.G. Proakis, "Digital Communications", *McGraw-Hill*, Third Ed., New York, 1995.

[9] M. Rangaswamy, J.H. Michels and D.D. Weiner, "Multichannel detection for correlated non-Gaussian random processes based on innovations", *IEEE Trans.-SP*, vol. 43, no. 8, pp. 1915-1922, Aug. 1995.

[10] J.-F. Cardoso, "Source Separation Using Higher Order Moments", *Proc. IEEE-ICASSP*, pp. 2109-2112, Glasgow, May 1989.

[11] P.M. Djuric et al., "Particle Filtering", *IEEE SP Magazine*, vol. 20, no. 5, pp. 19-38, Sep. 2003.

[12] J.B. Billingsley, "Ground Clutter Measurements for Surface-Sited Radar", *Technical Report 780, MIT*, Feb. 1993.

[13] S. Watts, "Radar detection prediction in sea clutter using the compound K-distribution model", *IEE Proceeding, Part. F*, vol. 132, no. 7, pp. 613-620, Dec. 1985.

[14] T. Nohara and S. Haykin, "Canada East Coast Trials and the K-Distribution", *IEE Proceeding, Part. F*, vol. 138, no. 2, pp. 82-88, 1991.

[15] A. Farina, A. Russo and F. Scannapieco, "Radar detection in coherent Weibull clutter", *IEEE Trans. on ASSP*, vol. 35, no. 6, pp. 893?895, June 1987.

[16] M. Rupp and R. Frenzel, "The Behavior of LMS and NLMS Algorithms with Delayed Coefficient Update in the Presence of Spherically Invariant Processes", *IEEE Trans.-SP*, vol. 42, no. 3, pp. 668-672, Mar. 1994.

[17] E. Conte and G. Ricci, "Performance Prediction in Compound-Gaussian Clutter", *IEEE Trans.-AES*, vol. 30, no. 2, pp. 611-616, Apr. 1994.

[18] F. Gini, "Sub-optimum Coherent Radar Detection in a Mixture of K-distributed and Gaussian Clutter", *IEE Proc. Radar, Sonar and Navigation*, vol. 144, no. 1, pp. 39-48, Feb. 1997.

[19] A. Abdi and S. Nader-Esfahani, "Expected Number of Maxima in the Envelope of a Spherically Invariant Random Process", *IEEE Trans.-IT*, vol. 49, no. 5, pp. 1369-1375, May 2003.

[20] K. Yao, M.K. Simon and E. Biglieri, "A Unified Theory on Wireless Communication Fading Statistics based on SIRV", *Fifth IEEE Workshop on SP Advances in Wireless Communications*, Lisboa, Portugal, July, 2004.

[21] L. Belkacem, J.L. Vhel and C. Walter, "Capm, Risk and Portfolio Selection in Alpha-stable Markets", *Fractals 8*, pp. 99-115, 2000.

[22] R. Little and D.B. Rubin, "Statistical Analysis with Missing Data", *John Wiley & Sons*, New York, 1987.

[23] C. Liu and D.B. Rubin, "ML Estimation of the t Distribution Using EM and its Extensions, ECM and ECME", *Statistica Sinica*, vol. 5, pp. 19-39, 1995.

[24] M. Rangaswamy, "Statistical Analysis of the nonhomogeneity detector for non-Gaussian interference backgrounds", *IEEE Trans.-SP*, vol. 53, no. 6, pp. 2101-2111, June 2005.

[25] E. Jay, J.-P. Ovarlez, D. Declercq and P. Duvaut, "BORD: Bayesian Optimum Radar Detector", *Signal Processing*, vol. 83, no. 6, pp. 1151-1162, Jun. 2003.

[26] F. Gini and M.V Greco, "Covariance Matrix Estimation for CFAR Detection in Correlated Heavy Tailed Clutter", *Signal Processing*, special section on SP with Heavy Tailed Distributions, vol. 82, no. 12, pp. 1847-1859, Dec. 2002.

[27] 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 Trans.-SP*, vol. 50, no. 8, pp. 1908-1915, Aug. 2002.

[28] T.W. Anderson, "An Introduction to Multivariate Statistical Analysis", *John Wiley & Sons*, New York, 1984.

[29] F. Pascal, J.-P. Ovarlez, P. Forster and P. Larzabal, "Constant False Alarm Rate Detection in Spherically Invariant Random Processes", *Proc. of the European Signal Processing Conf.*, Vienna, pp. 2143-2146, Sep. 2004.

[30] K. Yao, "A Representation Theorem and its Applications to Spherically Invariant Random Processes", *IEEE Trans.-IT*, vol. 19, no. 5, pp. 600-608, Sep. 1973.

[31] M. Rangaswamy, D.D. Weiner and A. Ozturk, "Non-Gaussian vector identification using spherically invariant random processes", *IEEE Trans.-AES*, vol. 29, no. 1, pp. 111-124, Jan. 1993.

[32] F. Pascal, J.-P. Ovarlez, P. Forster and P. Larzabal, "On a SIRV-CFAR Detector with radar experimentations in impulsive noise", *Proc. of the European Signal Processing Conf.*, Florence, Sep. 2006.

[33] A.K. Gupta and D.K. Nagar, "Matrix Variate Distributions", *Chapman & Hall/CRC*, 2000.

[34] E. Conte, M. Lops and G. Ricci, G., "Adaptive matched filter detection in spherically invariant noise", *IEEE SP Letters*, vol. 3, no. 8, pp. 248-250, Aug. 1996.

[35] S. Bausson, F. Pascal, P. Forster, J.-P. Ovarlez and P. Larzabal, "First and Second Order Moments of the Normalized Sample Covariance Matrix of Spherically Invariant Random Vectors", *IEEE SP Letters*, vol. 14, no. 6, pp. 425-428, June 2007.

[36] F. Pascal, Y. Chitour, J-P. Ovarlez, P. Forster and P. Larzabal, "Covariance Structure Maximum Likelihood Estimates in Compound Gaussian Noise : Existence and Algorithm Analysis", *IEEE Trans.-SP*. (To appear)

[37] F. Pascal, P. Forster, J.-P. Ovarlez and P. Larzabal, "Theoretical Analysis of an Improved Covariance Matrix Estimator in Non-Gaussian Noise", *Proc. IEEE-ICASSP*, Philadelphia, vol. IV, pp. 69-72, Mar. 2005.

[38] A.W. van der Vaart, "Asymptotic Statistics", *Cambridge University Press*, 1998.

Frédéric Pascal was born in Sallanches, France in 1979. He received the Master's degree ("Probabilities, Statistics and Applications : Signal, Image et Networks") with merit, in Applied Statistics from University Paris VII - Jussieu, Paris, France, in 2003. Then, he received the Ph.D. degree of Signal Processing, from university Paris X - Nanterre, advised by Pr. Philippe Forster : "Detection and Estimation in Compound Gaussian Noise" in 2006. This Ph.D thesis was in collaboration with the French Aerospace Lab (ONERA), Palaiseau, France. Since november 2006, he makes a post doctoral position in the Signal Processing and Information team of the laboratory SATIE (Système et Applications des Technologies de l'Information et de l'Energie), CNRS, Ecole Normale Supérieure, Cachan, France. His research interests are estimation in signal processing and radar detection.

Philippe Forster (M'89) was born in Brest, France, in 1960. He received the Agrégation de physique appliquée degree in 1983 from the Ecole Normale Supérieure de Cachan, Cachan, France, and the Ph.D. degree in electrical engineering from the Université de Rennes, Rennes, France, in 1988. He is currently Professor of electrical engineering at the Institut Universitaire de Technologie de Ville d'Avray, Ville d'Avray, France, where he is a member of the GEA Laboratory. His research interests are in estimation and detection theory with applications to array processing, radar, and digital communications.

Jean-Philippe Ovarlez was born in Denain, France in 1963. He received jointly the engineering degree from Ecole Supérieure d'Electronique Automatique et Informatique (ESIEA), Paris, France and the Diplôme d'Etudes Approfondies degree in Signal Processing from University of Orsay (Paris XI), Orsay, France and the Ph.D. degree in Physics from the University of Paris VI, Paris, France, in 1987 and 1992, respectively. In 1992, he joined the Electromagnetic and Radar Division of the French Aerospace Lab (ONERA), Palaiseau, France, where he is is currently Chief Scientist and member of the Scientific Committee of the ONERA Physics Branch. His current activities of research are centered in the topic of Signal Processing for radar and SAR applications such as time-frequency, imaging, detection and parameters estimation.

Pascal Larzabal (M'93) was born in the Basque country in the south of France in 1962. He received the Agrégation degree in electrical engineering and the Dr.Sci. and Habilitation à diriger les recherches degrees in 1985, 1988, and 1998, respectively, all from the Ecole Normale Supérieure of Cachan, Cachan, France. He is now a Professor at the Institut Universitaire de Technologie of Cachan (University Paris Sud, Paris, France, where he teaches electronic and signal processing. He is at the head of the Signal Processing and Information team of the laboratory SATIE, CNRS, Ecole Normale Supérieure, Cachan. His research interests are estimation in array processing and spectral analysis.
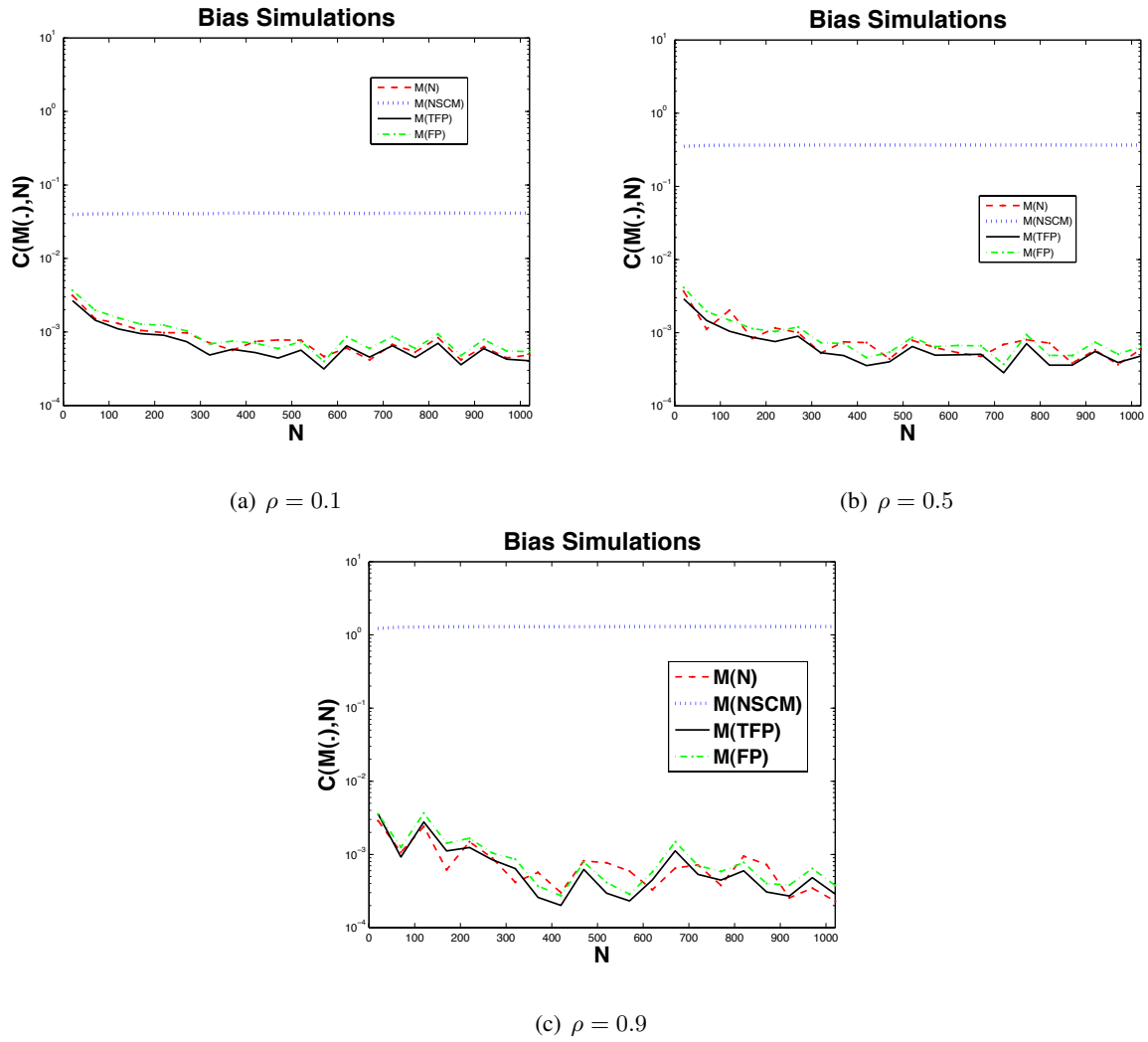
(a) $\rho = 0.1$

(b) $\rho = 0.5$

(c) $\rho = 0.9$

Fig. 1.   Estimates Bias for different values of ρ.

(a) $\rho = 0.1$

(b) $\rho = 0.5$

(c) $\rho = 0.9$
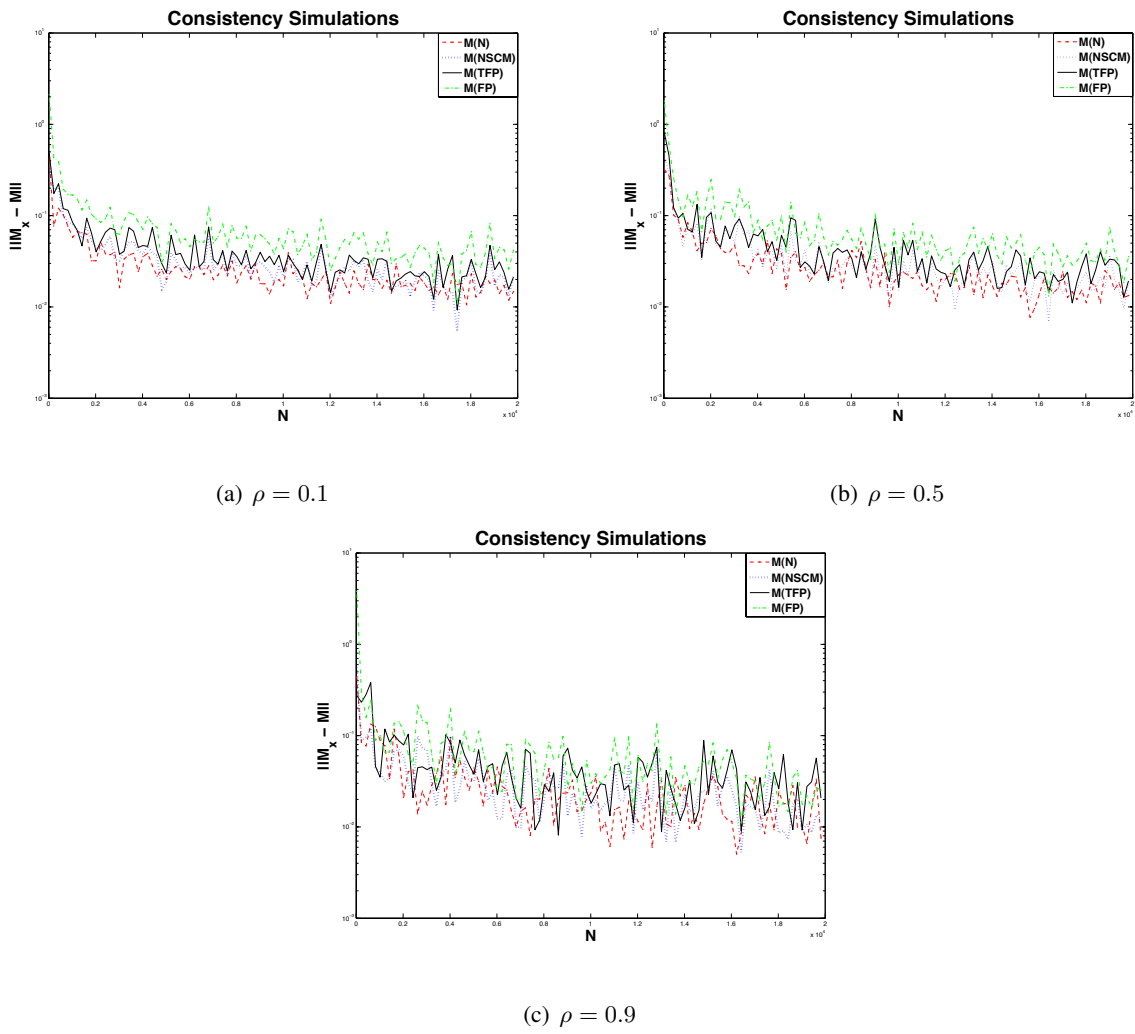
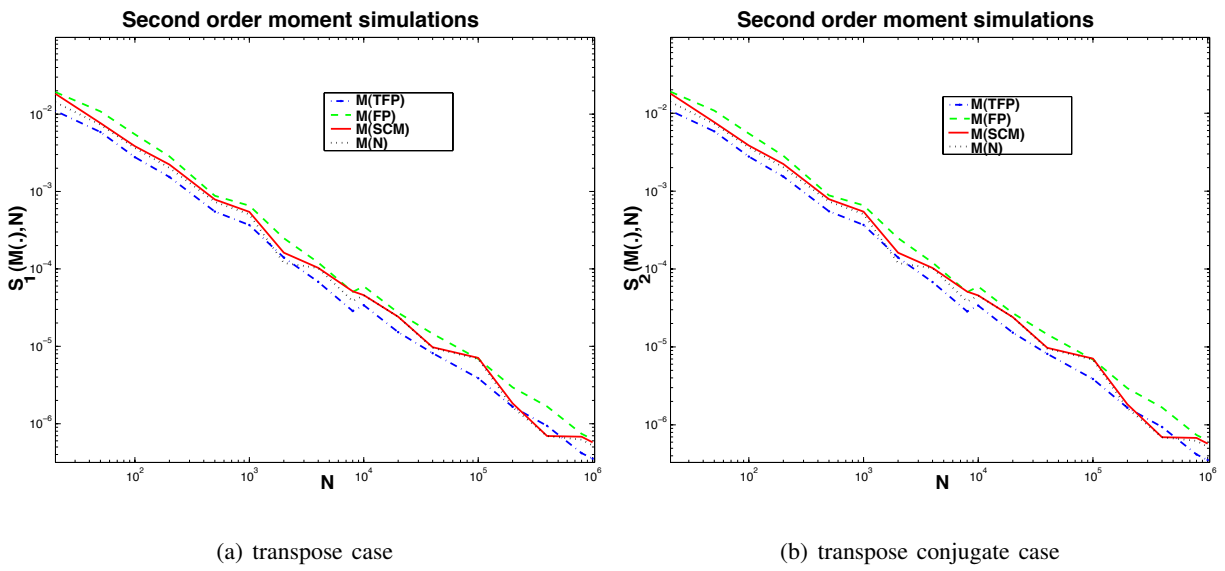Fig. 2.   Estimates Consistency for different values of ρ. $m = 3$.

(a) transpose case

(b) transpose conjugate case

Fig. 3.   Estimates second order moment