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

Frédéric Pascal, Y. Chitour, Jean-Philippe Ovarlez, Philippe Forster, Pascal Larzabal

▶ **To cite this version:**

Frédéric Pascal, Y. Chitour, Jean-Philippe Ovarlez, Philippe Forster, Pascal Larzabal. Covariance Structure Maximum-Likelihood Estimates in Compound Gaussian Noise: Existence and Algorithm Analysis. IEEE Transactions on Signal Processing, 2008, 56 (1), pp.34-48. 10.1109/TSP.2007.901652 . hal-01816367

**HAL Id: hal-01816367**
**https://hal.science/hal-01816367**

Submitted on 29 Feb 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

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

**Abstract**

Recently, a new adaptive scheme [1], [2] has been introduced for covariance structure matrix estimation in the context of adaptive radar detection under non Gaussian noise. This latter has been modelled by compound-Gaussian noise, which is the product $\mathbf{c}$ of the square root of a positive unknown variable $\tau$ (deterministic or random) and an independent Gaussian vector $\mathbf{x}$, $\mathbf{c} = \sqrt{\tau}\,\mathbf{x}$.

Because of the implicit algebraic structure of the equation to solve, we called the corresponding solution, the Fixed Point (FP) estimate. When $\tau$ is assumed deterministic and unknown, the FP is the exact Maximum Likelihood (ML) estimate of the noise covariance structure, while when $\tau$ is a positive random variable, the FP is an Approximate Maximum Likelihood (AML).

This estimate has been already used for its excellent statistical properties without proofs of its existence and uniqueness. The major contribution of this paper is to fill these gaps. Our derivation is based on some Likelihood functions general properties like homogeneity and can be easily adapted to other recursive contexts. Moreover, the corresponding iterative algorithm used for the FP estimate practical determination is also analyzed and we show the convergence of this recursive scheme, ensured whatever the initialization.

F. Pascal is with the Office National d'Etudes et de Recherches Aérospatiales, DEMR/TSI, BP 72, 92322 Chatillon Cedex, France (e-mail: frederic.pascal@onera.fr).

Y. Chitour is with the Laboratoire des Signaux et Systèmes, Supélec, 3 rue Joliot-Curie, 91190 Gif-sur-Yvette, France (e-mail: yacine.chitour@lss.supelec.fr)

J. P. Ovarlez is with the Office National d'Etudes et de Recherches Aérospatiales, DEMR/TSI, BP 72, 92322 Chatillon Cedex, France (e-mail: jean-philippe.ovarlez@onera.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).

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).

**Index Terms**

Compound-Gaussian, SIRV, Maximum likelihood estimate, adaptive detection, CFAR detector.

## I. INTRODUCTION

The basic problem of detecting a complex signal embedded in an additive Gaussian noise has been extensively studied these last decades. In these contexts, adaptive detection schemes required an estimate of the noise covariance matrix generally obtained from signal free data traditionally called secondary data or reference data. The resulting adaptive detectors, as those proposed by [7] and [8], are all based on the Gaussian assumption for which the Maximum Likelihood (ML) estimate of the covariance matrix is given by the sample covariance matrix. However, these detectors may exhibit poor performance when the additive noise is no more Gaussian [6].

This is the case in radar detection problems where the additive noise is due to the superposition of unwanted echoes reflected by the environment and traditionally called the clutter. Indeed, experimental radar clutter measurements showed that these data are non-Gaussian. This fact arises for example when the illuminated area is non-homogeneous or when the number of scatterers is small. This kind of non-Gaussian noises is usually described by distributions such as K-distribution, Weibull, ... Therefore, this non-Gaussian noise characterization has gained a lot of interest in the radar detection community.

One of the most general and elegant non-Gaussian noise model is the compound-Gaussian process which includes the so-called *Spherically Invariant Random Vectors* (SIRV). These processes encompass a large number of non-Gaussian distributions mentioned above and include of course Gaussian processes. They have been recently introduced, in radar detection, to model clutter for solving the basic problem of detecting a known signal. This approach resulted in the adaptive detectors development such as the Generalized Likelihood Ratio Test-Linear Quadratic (GLRT-LQ) in [1], [2] or the Bayesian Optimum Radar Detector (BORD) in [3], [4]. These detectors require an estimate of the covariance matrix of the noise Gaussian component. In this context, ML estimates based on secondary data have been introduced in [11], [12], together with a numerical procedure supposed to obtain them. However, as noticed in [12] p.1852, "*existence of the ML estimate and convergence of iteration [...] is still an open problem*".

To the best of our knowledge, the proofs of existence, uniqueness of the ML estimate and convergence of the algorithm proposed in [1] have never been established. The main purpose of this paper is to fill these gaps.

The paper is organized as follows. In the Section II, we present the two main models of interest in our ML estimation framework. Both models lead to ML estimates which are solution of a transcendental

equation. Section IV presents the main results of this paper while a proofs outline is given in Section V: for presentation clarity, full demonstrations are provided in Appendices. Finally, Section VI gives some simulations results which confirm the theoretical analysis.

## II. State of the art and problem formulation

A compound-Gaussian process $\mathbf{c}$ is the product of the square root of a positive scalar quantity $\tau$ called the texture and a $m$-dimensional zero mean complex Gaussian vector $\mathbf{x}$ with covariance matrix $\mathbf{M} = \mathbb{E}(\mathbf{x}\mathbf{x}^H)$ usually normalized according to $\mathrm{Tr}(\mathbf{M}) = m$, where $H$ denotes the conjugate transpose operator and $\mathrm{Tr}(.)$ stands for the trace operator:

$$\mathbf{c} = \sqrt{\tau}\,\mathbf{x}\,. \tag{1}$$

This general model leads to two distinct approaches: the well-known SIRV modeling where the texture is considered random and the case where the texture is treated as an unknown nuisance parameter.

Generally, the covariance matrix $\mathbf{M}$ is not known and an estimate $\widehat{\mathbf{M}}$ is required for the Likelihood Ratio (LR) computation. Classically, such an estimate $\widehat{\mathbf{M}}$ is obtained from Maximum Likelihood (ML) theory, well known for its good statistical properties. In this problem, estimation of $\mathbf{M}$ must respect the previous $\mathbf{M}$-normalization, $\mathrm{Tr}(\widehat{\mathbf{M}}) = m$. This estimate $\widehat{\mathbf{M}}$ will be built using $N$ independent realizations of $\mathbf{c}$ denoted $\mathbf{c}_i = \sqrt{\tau_i}\,\mathbf{x}_i$ for $i = 1, \ldots, N$.

It straightforwardly appears that the Likelihood will depend on the assumption relative to texture. The two most often met cases are presented in the two following subsections.

### A. SIRV case

Let us recap that a SIRV [5] is the product of the square root of a positive random variable $\tau$ (*texture*) and a $m$-dimensional independent complex Gaussian vector $\mathbf{x}$ (*speckle*) with zero mean normalized covariance matrix $\mathbf{M}$. This model led to many investigations [1], [2], [3], [4].

To obtain the ML estimate of $\mathbf{M}$, with no proofs of existence and uniqueness, Gini *et al.* derived in [12] an Approximate Maximum Likelihood (AML) estimate $\widehat{\mathbf{M}}$ as the solution of the following equation

$$\widehat{\mathbf{M}} = f(\widehat{\mathbf{M}})\,, \tag{2}$$

where $f$ is given by

$$f(\widehat{\mathbf{M}}) = \frac{m}{N}\sum_{i=1}^{N}\frac{\mathbf{c}_i\,\mathbf{c}_i^H}{\mathbf{c}_i^H\,\widehat{\mathbf{M}}^{-1}\,\mathbf{c}_i}. \tag{3}$$

### B. Unknown deterministic $\tau$ case

This approach has been developed in [13] where the $\tau_i$'s are assumed to be unknown deterministic quantities. The corresponding Likelihood function to maximize with respect to $\mathbf{M}$ and $\tau_i$'s, is given by

$$p_C(\mathbf{c}_1,\ldots,\mathbf{c}_N;\mathbf{M},\tau_1,\ldots,\tau_N) = \frac{1}{(\pi)^{mN}\,|\mathbf{M}|^N}\prod_{i=1}^{N}\frac{1}{\tau_i^m}\exp\left(-\frac{\mathbf{c}_i^H\,\mathbf{M}^{-1}\,\mathbf{c}_i}{\tau_i}\right), \tag{4}$$

where $|\mathbf{M}|$ denotes the determinant of matrix $\mathbf{M}$.

Maximization with respect to $\tau_i$'s, for a given $\mathbf{M}$, leads to $\hat{\tau}_i = \dfrac{\mathbf{c}_i^H\,\mathbf{M}^{-1}\,\mathbf{c}_i}{m}$, and then by replacing the $\tau_i$'s in (4) by their ML estimates $\hat{\tau}_i$'s , we obtain the reduced likelihood function

$$\hat{p}_C(\mathbf{c}_1,\ldots,\mathbf{c}_N;\mathbf{M}) = \frac{1}{(\pi)^{mN}\,|\mathbf{M}|^N}\prod_{i=1}^{N}\frac{m^m\,\exp(-m)}{(\mathbf{c}_i^H\,\mathbf{M}^{-1}\,\mathbf{c}_i)^m}.$$

Finally, maximizing $\hat{p}_C(\mathbf{c}_1,\ldots,\mathbf{c}_N;\mathbf{M})$ with respect to $\mathbf{M}$ is equivalent to maximize the following function $F$, written in terms of $\mathbf{x}_i$'s and $\tau_i$'s thanks to (1)

$$F(\mathbf{M}) = \frac{1}{|\mathbf{M}|^N}\prod_{i=1}^{N}\frac{1}{\tau_i^m(\mathbf{x}_i^H\,\mathbf{M}^{-1}\,\mathbf{x}_i)^m}. \tag{5}$$

By cancelling the gradient of $F$ with respect to $\mathbf{M}$, we obtain the following equation

$$\widehat{\mathbf{M}} = f(\widehat{\mathbf{M}}), \tag{6}$$

where $f$ is given again by (3) and whose solution is the Maximum Likelihood Estimator in the deterministic texture framework.

Note that $f$ can be rewritten from (1) as

$$f(\widehat{\mathbf{M}}) = \frac{m}{N}\sum_{i=1}^{N}\frac{\mathbf{x}_i\,\mathbf{x}_i^H}{\mathbf{x}_i^H\,\widehat{\mathbf{M}}^{-1}\,\mathbf{x}_i}. \tag{7}$$

Equation (7) shows that $f(\widehat{\mathbf{M}})$ does not depend on the texture $\tau$ but only on the Gaussian vectors $\mathbf{x}_i$'s.

### C. Problem Formulation

It has been shown in [12], [13] that estimation schemes developed under both the stochastic case (Section II-A) and the deterministic case (Section II-B) lead to the analysis of the same equation ((2) and (6)), whose solution is a fixed point of $f$ (7). A first contribution of this paper is to establish the existence

and the uniqueness, up to a scalar factor, of this fixed point $\widehat{\mathbf{M}}_{FP}$ which is is the Approximate Maximum Likelihood (AML) estimate under the stochastic assumption and the exact ML under the deterministic assumption.

Moreover, a second contribution is to analyze an algorithm based on the key equation (6), which defines $\widehat{\mathbf{M}}_{FP}$. The convergence of this algorithm will be established. Then, numerical results of Section VI will illustrate the computational efficiency of the algorithm for obtaining the FP estimate.

Finally, the complete statistical properties investigation of the corresponding ML estimate will be addressed in a forthcoming paper.

## III. Statement of the main result

We first provide some notations. Let $m$ and $N$ be positive integers such that $m < N$. We use $\mathbb{R}^{+*}$ to denote the set of strictly positive real scalars, $M_m(\mathbb{C})$ to denote the set of $m \times m$ complex matrices, and $\mathcal{G}$, the subset of $M_m(\mathbb{C})$ defined by the positive definite Hermitian matrices. For $\mathbf{M} \in M_m(\mathbb{C})$ , $||\mathbf{M}|| := \text{Tr}\left(\mathbf{M}^H\mathbf{M}\right)^{1/2}$ the Frobenius norm of $\mathbf{M}$ which is the norm associated to an inner product on $M_m(\mathbb{C})$. Moreover, from the statistical independence hypothesis of the $N$ complex $m$-vectors $\mathbf{x}_i$, it is natural to assume the following

$(H)$: Let us set $\mathbf{x}_i = \mathbf{x}_i^{(1)} + j\mathbf{x}_i^{(2)}$ . Any $2m$ distinct vectors taken in

$$\left\{ \begin{pmatrix} \mathbf{x}_1^{(1)} \\ \mathbf{x}_1^{(2)} \end{pmatrix}, \ldots, \begin{pmatrix} \mathbf{x}_N^{(1)} \\ \mathbf{x}_N^{(2)} \end{pmatrix}, \begin{pmatrix} -\mathbf{x}_1^{(2)} \\ \mathbf{x}_1^{(1)} \end{pmatrix}, \ldots, \begin{pmatrix} -\mathbf{x}_N^{(2)} \\ \mathbf{x}_N^{(1)} \end{pmatrix} \right\}$$

are linearly independent.

From (5) and (7), one has

$$\begin{array}{rcccl} F & : & \mathcal{G} & \longrightarrow & \mathbb{R}^{+*} \\ & & \mathbf{M} & \longrightarrow & F(M) = \dfrac{1}{|\mathbf{M}|^N} \displaystyle\prod_{i=1}^{N} \dfrac{1}{\tau_i^m \left(\mathbf{x}_i^H \mathbf{M}^{-1} \mathbf{x}_i\right)^m}, \end{array}$$

and

$$\begin{array}{rcccl} f & : & \mathcal{G} & \longrightarrow & \mathcal{G} \\ & & \mathbf{M} & \longrightarrow & f(M) = \dfrac{m}{N} \displaystyle\sum_{i=1}^{N} \dfrac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \mathbf{M}^{-1} \mathbf{x}_i}. \end{array}$$

**Theorem III.1**

$(i)$ *There exists* $\widehat{\mathbf{M}}_{FP} \in \mathcal{G}$ *with unit norm such that, for every* $\alpha > 0$, $f$ *admits a unique fixed point of norm* $\alpha > 0$ *equal to* $\alpha \widehat{\mathbf{M}}_{FP}$. *Moreover,* $F$ *reaches its maximum over* $\mathcal{G}$ *only on* $\mathcal{L}_{\widehat{\mathbf{M}}_{FP}}$, *the open half-line spanned by* $\widehat{\mathbf{M}}_{FP}$.

$(ii)$ *Let* $(S)_{dis}$ *be the discrete dynamical system defined on* $\mathcal{D}$ *by*

$$(S)_{dis} \ : \mathbf{M}_{k+1} = f(\mathbf{M}_k). \tag{8}$$

*Then, for every initial condition* $\mathbf{M}_0 \in \mathcal{G}$, *the resulting sequence* $(\mathbf{M}_k)_{k\geq 0}$ *converges to a fixed point of* $f$, *i.e. to a point where* $F$ *reaches its maximum;*

$(iii)$ *Let* $(S)_{cont}$ *be the continuous dynamical system defined on* $\mathcal{G}$ *by*

$$(S)_{cont} \ : \dot{\mathbf{M}} = \quad \nabla F(\mathbf{M}). \tag{9}$$

*Then, for every initial condition* $\mathbf{M}(0) = \mathbf{M}_0 \in \mathcal{G}$, *the resulting trajectory* $\mathbf{M}(t)$, $t \geq 0$, *converges when* $t$ *tends to* $+\infty$, *to the point* $\|\mathbf{M}_0\| \widehat{\mathbf{M}}_{FP}$, *i.e. to a point where* $F$ *reaches its maximum.*

Consequently to $(i)$, $\widehat{\mathbf{M}}_{FP}$ is the unique positive definite $m \times m$ matrix of norm one satisfying

$$\widehat{\mathbf{M}}_{FP} = \frac{m}{N} \sum_{i=1}^{N} \frac{\mathbf{x}_i \, \mathbf{x}_i^H}{\mathbf{x}_i^H \, \widehat{\mathbf{M}}_{FP}^{-1} \, \mathbf{x}_i}. \tag{10}$$

*Proof:* The same problem and the same result can be formulated with real numbers instead of complex numbers and symmetric matrices instead of hermitian matrices, while hypothesis $(H)$ becomes hypothesis $(H2)$ stated below (just before Remark IV.1). The proof of Theorem III.1 breaks up into two stages. We first show in Appendix I how to derive Theorem III.1 from the corresponding real results. Then, the rest of the paper is devoted to the study of the real case. ■

## IV. Notations and statements of the results in the real case

### A. Notations

In this paragraph, we introduce the main notations of the paper for the real case. Notations already defined in the complex case are translated in the real one. Moreover, real results will be valid for every integer $m$. For every positive integer $n$, $[\![1, n]\!]$ denotes the set of integers $\{1, \ldots, n\}$. For vectors of $\mathbb{R}^m$, the norm used is the Euclidean one. Throughout the paper, we will use several basic results on square matrices, especially regarding diagonalization of real symmetric and orthogonal matrices. We refer to [14] for such standard results.

We use $M_m(\mathbb{R})$ to denote the set of $m \times m$ real matrices, $SO(m)$ to denote the set of $m \times m$ orthogonal matrices and $\mathbf{M}^\top$, the transpose of $\mathbf{M}$. We denote the identity matrix of $M_m(\mathbb{R})$ by $\mathbf{I}_m$.

We next define and list the several sets of matrices used in the sequel:

* $\mathcal{D}$, the subset of $M_m(\mathbb{R})$ defined by the symmetric positive definite matrices;
* $\overline{\mathcal{D}}$, the closure of $\mathcal{D}$ in $M_m(\mathbb{R})$, i.e. the subset of $M_m(\mathbb{R})$ defined by the symmetric non negative matrices;
* For every $\alpha > 0$, $\begin{cases} \mathcal{D}(\alpha) = \{\mathbf{M} \in \mathcal{D} |\ ||\mathbf{M}|| = \alpha\} \\ \overline{\mathcal{D}}(\alpha) = \left\{\mathbf{M} \in \overline{\mathcal{D}} |\ ||\mathbf{M}|| = \alpha\right\} \end{cases}$.

It is obvious that $\overline{\mathcal{D}}(\alpha)$ is compact in $M_m(\mathbb{R})$.

For $\mathbf{M} \in \mathcal{D}$, we use $\mathcal{L}_{\mathbf{M}}$ to denote the open-half line spanned by $\mathbf{M}$ in the cone $\mathcal{D}$, i.e. the set of points $\lambda\,\mathbf{M}$, with $\lambda > 0$. Recall that the order associated with the cone structure of $\mathcal{D}$ is called the Loewner order for symmetric matrices of $M_m(\mathbb{R})$ and is defined as follows. Let $\mathbf{A}, \mathbf{B}$ be two symmetric $m \times m$ real matrices. Then $\mathbf{A} \leq \mathbf{B}$ ($\mathbf{A} < \mathbf{B}$ respectively) means that the quadratic form defined by $\mathbf{B} - \mathbf{A}$ is non negative (positive definite respectively), i.e., for every non zero $\mathbf{x} \in \mathbb{R}^m$, $\mathbf{x}^\top (\mathbf{A} - \mathbf{B})\,\mathbf{x} \leq 0$, ($> 0$ respectively). Using that order, one has $\mathbf{M} \in \mathcal{D}$ ($\in \overline{\mathcal{D}}$ respectively) if and and only if $\mathbf{M} > \mathbf{0}$ ($\mathbf{M} \geq \mathbf{0}$ respectively).

As explained in Appendix I, we will study in this section the applications $F$ and $f$ (same notations as in the complex case) defined as follows:

$$\begin{array}{rcccl} F & : & \mathcal{D} & \longrightarrow & \mathbb{R}^{+*} \\ & & \mathbf{M} & \longrightarrow & \dfrac{1}{|\mathbf{M}|^N} \displaystyle\prod_{i=1}^{N} \dfrac{1}{\tau_i^m \left(\mathbf{x}_i^\top\, \mathbf{M}^{-1}\, \mathbf{x}_i\right)^m}, \end{array}$$

and

$$\begin{array}{rcccl} f & : & \mathcal{D} & \longrightarrow & \mathcal{D} \\ & & \mathbf{M} & \longrightarrow & \dfrac{m}{N} \displaystyle\sum_{i=1}^{N} \dfrac{\mathbf{x}_i\, \mathbf{x}_i^\top}{\mathbf{x}_i^\top\, \mathbf{M}^{-1}\, \mathbf{x}_i}. \end{array}$$

Henceforth, $F$ and $f$ stay for the real formulation. In the above, the vectors $(\mathbf{x}_i)$, $1 \leq i \leq N$, belong to $\mathbb{R}^m$ and verify the next two hypothesis:

- $(H1)$ : $\|\mathbf{x}_i\| = 1, 1 \leq i \leq N$;
- $(H2)$ : For any $m$ two by two distinct indices $i(1) < ... < i(m)$ chosen in $[\![1, N]\!]$, the vectors $\mathbf{x}_{i(1)}, \ldots, \mathbf{x}_{i(m)}$ are linearly independent.

Consequently, the vectors $\mathbf{c}_1, \ldots, \mathbf{c}_m$ verify $(H2)$.

Hypothesis $(H1)$ stems from the fact that function $f$ does not depend on $\mathbf{x}_i$'s norm.

Let us already emphasize that hypothesis $(H2)$ is the key assumption for getting all our subsequent results. Hypothesis $(H2)$ has the following trivial but fundamental consequence that we state as a remark.

**Remark IV.1**

*For every $n$ vectors $\mathbf{x}_{i(1)}, \ldots, \mathbf{x}_{i(n)}$ (respectively $\mathbf{c}_{i(1)}, \ldots, \mathbf{c}_{i(n)}$) with $1 \leq n \leq m$, $1 \leq i \leq N$, the vector space generated by $\mathbf{x}_{i(1)}, \ldots, \mathbf{x}_{i(n)}$ (respectively $\mathbf{c}_{i(1)}, \ldots, \mathbf{c}_{i(n)}$) has dimension $n$.*

In the sequel, we use $f^n$, $n \geq 1$, to denote the $n$-th iterate of $f$ i.e., $f^n := f \circ ... \circ f$, where $f$ is repeated $n$ times. We also adopt the following standard convention $f^0 := Id_{\mathcal{D}}$.

The two functions $F$ and $f$ are related by the following relation, which is obtained after an easy computation. For every $\mathbf{M} \in \mathcal{D}$, let $\nabla F(\mathbf{M})$ be the gradient of $F$ at $\mathbf{M} \in \mathcal{D}$ i.e. the unique symmetric matrix verifying, for every matrix $M \in \mathcal{S}$,

$$\nabla F(\mathbf{M}) = N\, F(\mathbf{M})\, \mathbf{M}^{-1} \left( f(\mathbf{M}) - \mathbf{M} \right) \mathbf{M}^{-1}.$$

Clearly $\mathbf{M}$ is a fixed point of $f$ if and only if $M$ is a critical point of the vector field defined by $\nabla F$ on $\mathcal{D}$.

### *B. Statements of the results*

The goal of this paper is to establish the following theorems whose proofs are outlined in the next Section.

**Theorem IV.1**

*There exists $\widehat{\mathbf{M}}_{FP} \in \mathcal{D}$ with unit norm such that, for every $\alpha > 0$, $f$ admits a unique fixed point of norm $\alpha > 0$ equal to $\alpha \widehat{\mathbf{M}}_{FP}$. Moreover, $F$ reaches its maximum over $\mathcal{D}$ only on $\mathcal{L}_{\widehat{\mathbf{M}}_{FP}}$, the open half-line spanned by $\widehat{\mathbf{M}}_{FP}$.*

Consequently, $\widehat{\mathbf{M}}_{FP}$ is the unique positive definite $m \times m$ matrix of norm one satisfying

$$\widehat{\mathbf{M}}_{FP} = \frac{m}{N} \sum_{i=1}^{N} \frac{\mathbf{x}_i\, \mathbf{x}_i^\top}{\mathbf{x}_i^\top\, \widehat{\mathbf{M}}_{FP}^{-1}\, \mathbf{x}_i}. \tag{11}$$

**Remark IV.2**

*Theorem IV.1 relies on the fact that $F$ reaches its maximum on $\mathcal{D}$. Roughly speaking, that issue is proved as follows. The function $F$ is continuously extended by the zero function on the boundary of $D$, excepted on the zero matrix. Since $F$ is positive and bounded on $\mathcal{D}$, we conclude. Complete argument is provided in Appendix II.*

As a consequence of Theorem IV.1, one obtains the next result.

**Theorem IV.2**

• *Let $(S)_{dis}$ be the discrete dynamical system defined on $\mathcal{D}$ by*

$$(S)_{dis} \ : \mathbf{M}_{k+1} = f(\mathbf{M}_k). \tag{12}$$

*Then, for every initial condition $\mathbf{M}_0 \in \mathcal{D}$, the resulting sequence $(\mathbf{M}_k)_{k\geq 0}$ converges to a fixed point of $f$, i.e. to a point where $F$ reaches its maximum;*

• *Let $(S)_{cont}$ be the continuous dynamical system defined on $\mathcal{D}$ by*

$$(S)_{cont} \ : \dot{\mathbf{M}} = \quad \nabla F(\mathbf{M}). \tag{13}$$

*Then, for every initial condition $\mathbf{M}(0) = \mathbf{M}_0 \in \mathcal{D}$, the resulting trajectory $\mathbf{M}(t)$, $t \geq 0$, converges, when $t$ tends to $+\infty$, to the point $\|\mathbf{M}_0\| \, \widehat{\mathbf{M}}_{FP}$, i.e. to a point where $F$ reaches its maximum.*

The last theorem can be used to characterize numerically the points where $F$ reaches its maximum and the value of that maximum.

Notice that algorithm defined by (12) does not allow the control of the FP norm. Therefore, for practical convenient, we propose a slightly modified algorithm in which the $\mathbf{M}$-normalization is applied at each iteration. This is summarized in the following corollary:

**Corollary IV.1**

*The following scheme*

$$\mathbf{M}'_{k+1} = \frac{f(\mathbf{M}'_k)}{Tr\left(f(\mathbf{M}'_k)\right)}. \tag{14}$$

*yields the matrices sequence $\{\mathbf{M}'_0, \ldots, \mathbf{M}'_k\}$ which is related to the matrices sequence $\{\mathbf{M}_0, \ldots, \mathbf{M}_k\}$, provided by (12), by, for $1 \leq i \leq k$ ,*

$$\mathbf{M}'_i = \frac{\mathbf{M}_i}{Tr(\mathbf{M}_i)}.$$

*This algorithm converges to $\widehat{\mathbf{M}}_{FP}$ up to a scaling factor which is: $\dfrac{1}{Tr(\widehat{\mathbf{M}}_{FP})}$.*

As a consequence of Theorem IV.1, we can prove a matrix inequality which is interesting on its own. It simply expresses that the Hessian computed at a critical point of $F$ is non positive. We also provide an example showing that, in general, the Hessian is not definite negative. Therefore, in general, the convergence rate to the critical points of $F$ for the dynamical systems $(S)_{dis}$ and $(S)_{cont}$ is not exponential.

**Proposition IV.1**

*Let $m, N$ be two positive integers with $m < N$ and $\mathbf{x}_1, \ldots, \mathbf{x}_N$ be unit vectors of $\mathbb{R}^m$ subject to $(H2)$ and such that*

$$\frac{m}{N} \sum_{i=1}^{N} \mathbf{x}_i \, \mathbf{x}_i^\top = \mathbf{I}_m. \tag{15}$$

*Then, for every matrix $\mathbf{M}$ of $M_m(\mathbb{R})$, we have*

$$\frac{m}{N} \sum_{i=1}^{N} (\mathbf{x}_i^\top \, \mathbf{M} \, \mathbf{x}_i)^2 \leq \|\mathbf{M}\|^2. \tag{16}$$

Assuming Theorem IV.1, the proof of the proposition is short enough to be provided next.

We may assume $\mathbf{M}$ to be symmetric since it is enough to prove the result for $\left(\mathbf{M} + \mathbf{M}^T\right)/2$, the symmetric part of $\mathbf{M}$. Applying Theorem IV.1, it is clear that the function $F$ associated to the $\mathbf{x}_i$'s reaches its maximum over $\mathcal{D}$ at $\mathbf{I}_m$. The expression of $H_{\mathbf{I}_m}$, the Hessian of $F$ at $\mathbf{I}_m$ is the following. For every symmetric matrix $\mathbf{M}$, we have

$$H_{\mathbf{I}_m}(\mathbf{M}, \mathbf{M}) = N \, F(\mathbf{I}_m) \Big( \frac{m}{N} \sum_{i=1}^{N} (\mathbf{x}_i^\top \, \mathbf{M} \, \mathbf{x}_i)^2 - \|\mathbf{M}\|^2 \Big).$$

Since $H_{\mathbf{I}_m}$ is non positive, (16) follows. Note that a similar formula can be given if, instead of (15), the $\mathbf{x}_i$'s verify the more general equation (11).

Because of the homogeneity properties of $F$ and $f$ and in order to prove that the rates of convergence of both $(S)_{dis}$ and $(S)_{cont}$ are not exponential, one must prove that the Hessian $H_{\mathbf{I}_m}$ is not negative definite on the orthogonal to $\mathbf{I}_m$ in the set of all symmetric matrices. The latter is simply the set of symmetric matrices with null trace. We next provide a numerical example describing that situation. Here, $m = 3$, $N = 4$ and

$$\mathbf{x}_1 = \begin{pmatrix} \frac{2\sqrt{2}}{3} \\ 0 \\ \frac{1}{3} \end{pmatrix}, \; \mathbf{x}_2 = \begin{pmatrix} -\frac{\sqrt{2}}{3} \\ \frac{\sqrt{2}}{\sqrt{3}} \\ \frac{1}{3} \end{pmatrix}, \; \mathbf{x}_3 = \begin{pmatrix} -\frac{\sqrt{2}}{3} \\ -\frac{\sqrt{2}}{\sqrt{3}} \\ \frac{1}{3} \end{pmatrix}, \; \mathbf{x}_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Then, hypotheses $(H1)$, $(H2)$ and (15) are satisfied. Moreover, it is easy to see that, for every diagonal matrix $\mathbf{D}$, we have equality in (16).

## V. Proofs outline

In that Section, we give Theorem IV.1 proof and Theorem IV.2 one. Each proof is decomposed in a sequence of lemmas and propositions whose arguments are postponed in the Appendices.

### A. *Proof of Theorem IV.1*

Theorem conclusions are the consequences of several propositions whose statements are listed below.

First of all, it is clear that $F$ is homogeneous of degree zero and $f$ is homogeneous of degree one, i.e., for every $\lambda > 0$ and $\mathbf{M} \in \mathcal{D}$, one has

$$F(\lambda\,\mathbf{M}) = F(\mathbf{M}), \quad f(\lambda\mathbf{M}) = \lambda f(\mathbf{M}).$$

The first proposition is the following.

**Proposition V.1**

*The supremum of $F$ over $\mathcal{D}$ is finite and is reached at a point $\widehat{\mathbf{M}}_{FP} \in \mathcal{D}$ with $\|\widehat{\mathbf{M}}_{FP}\| = 1$. Therefore, $f$ admits the open-half line $\mathcal{L}_{\widehat{\mathbf{M}}_{FP}}$ as fixed points.*

*Proof:* See Appendix II ■

It remains to show that there are no other fixed points of $f$ except $\mathcal{L}_{\widehat{\mathbf{M}}_{FP}}$. For that purpose, one must study the function $f$. We first establish the following result.

**Proposition V.2**

*The function $f$ verifies the following properties.*

- *(P1) : For every $\mathbf{M}, \mathbf{Q} \in \mathcal{D}$, if $\mathbf{M} \leq \mathbf{Q}$, then $f(\mathbf{M}) \leq f(\mathbf{Q})$ (also true with strict inequalities);*
- *(P2) : for every $\mathbf{M}, \mathbf{Q} \in \mathcal{D}$, then*

$$f(\mathbf{M} + \mathbf{Q}) \geq f(\mathbf{M}) + f(\mathbf{Q}), \tag{17}$$

*and equality occurs if and only if $\mathbf{M}$ and $\mathbf{Q}$ are colinear.*

*Proof:* See Appendix III ■

The property of $f$ described in the next proposition turns out to be basic for the proofs of both theorems.

**Proposition V.3**

*The function f is eventually strictly increasing, i.e. for every $\mathbf{Q}, \mathbf{P} \in \mathcal{D}$ such that $\mathbf{Q} \geq \mathbf{P}$ and $\mathbf{Q} \neq \mathbf{P}$, then $f^m(\mathbf{Q}) > f^m(\mathbf{P})$.*

*Proof:* See Appendix IV ■

We next proceed by establishing another property of $f$, which can be seen as an intermediary step towards the conclusion.

Recall that the *orbit* of $f$ associated to $\mathbf{M} \in \mathcal{D}$ is the trajectory of $(S)_{dis}$ (12) starting at $\mathbf{M}$.

**Proposition V.4**

*The following statements are equivalent.*

(A) $f$ *admits a fixed point;*

(B) $f$ *has one bounded orbit in* $\mathcal{D}$*;*

(C) *every orbit of* $f$ *is bounded in* $\mathcal{D}$*.*

*Proof:* See Appendix V ■

From proposition V.1, $f$ admits a fixed point. Thus, proposition V.4 ensures that every orbit of $f$ is bounded in $\mathcal{D}$.

Finally, using Proposition V.3, we get the following corollary, which concludes the proof of Theorem IV.1.

**Corollary V.1**

*Assume that every orbit of* $f$ *is bounded in* $\mathcal{D}$*. The following holds true.*

- *(C1) : Let* $\mathbf{P} \in \mathcal{D}$ *and* $n \geq 1$ *such that* $\mathbf{P}$ *can be compared with* $f^n(\mathbf{P})$*, i.e.* $\mathbf{P} \geq f^n(\mathbf{P})$ *or* $\mathbf{P} \leq f^n(\mathbf{P})$*. Then,* $\mathbf{P} = f^n(\mathbf{P})$*. In particular, if* $\mathbf{P} \geq f(\mathbf{P})$ *or* $\mathbf{P} \leq f(\mathbf{P})$*, then* $\mathbf{P}$ *is a fixed point of* $f$*;*
- *(C2) : All the fixed points of* $f$ *are colinear.*

*Proof:* See Appendix VI ■

To summarize, proposition V.1 establishes the existence of a fixed point while corollary V.1 ensures the uniqueness of the unit norm fixed point.

### *B. Proof of Theorem IV.2*

*1) Convergence results for* $(S)_{dis}$*:* In the previous Section, we already proved several important facts relative to the trajectories of $(S)_{dis}$ defined by (12), i.e. the orbits of $f$. Indeed, since $f$ has fixed points, then all the orbits of $f$ are bounded in $\mathcal{D}$. It remains to show now that each of them is convergent to a fixed point of $f$.

For that purpose, we consider, for every $\mathbf{M} \in \mathcal{D}$, the positive limit set $\omega(\mathbf{M})$ associated to $\mathbf{M}$, i.e., the set made of the cluster points of the sequence $(\mathbf{M}_k)_{k\geq 0}$, where $\mathbf{M}_{k+1} = f(\mathbf{M}_k)$ with $\mathbf{M}_0 = \mathbf{M}$. Since the orbit of $f$ associated to $\mathbf{M}$ is bounded in $\mathcal{D}$, the set $\omega(\mathbf{M})$ is a compact of $\mathcal{D}$ and is invariant by $f$: for every $\mathbf{P} \in \omega(\mathbf{M})$, $f(\mathbf{P}) \in \omega(\mathbf{M})$. It is clear that the sequence $(\mathbf{M}_k)_{k\geq 0}$ converges if and only if $\omega(\mathbf{M})$ reduces to a single point.

The last part of the proof is divided into two lemmas, whose statements are given below.

**Lemma V.1**

*For every $\mathbf{M} \in \mathcal{D}$, $\omega(\mathbf{M})$ contains a periodic orbit of $f$ (i.e. contain a finite number of points).*

*Proof:* See Appendix VII ■

**Lemma V.2**

*Let $\mathbf{M}_1$ and $\mathbf{M}_2 \in \mathcal{D}$ be such that their respective orbits are periodic. Then $\mathbf{M}_1$ and $\mathbf{M}_2$ are colinear and are both fixed points of $f$.*

*Proof:* See Appendix VIII ■

We now complete the proof of theorem IV.2 in the discrete case.

Let $\mathbf{M} \in \mathcal{D}$. Using both lemmas, it is easy to deduce that $\omega(\mathbf{M})$ contains a fixed point of $f$, which will be denoted by $\mathbf{Q}$. Notice that there exists a compact $\mathcal{K}$ containing both the orbit of $f$ associated to $\mathbf{M}$ and $\omega(\mathbf{M})$. We next prove that, for every $\varepsilon > 0$, there exists a positive integer $n_\varepsilon > 0$ such that

$$\Big(1 - \varepsilon\Big)\,\mathbf{Q} \leq f^{n_\varepsilon}(\mathbf{M}) \leq \Big(1 + \varepsilon\Big)\,\mathbf{Q}. \tag{18}$$

Indeed, since $\mathbf{Q} \in \omega(\mathbf{M})$, for every $\varepsilon > 0$, there exists a positive integer $n_\varepsilon > 0$ such that

$$\|f^{n_\varepsilon}(\mathbf{M}) - \mathbf{Q}\| \leq \varepsilon.$$

After standard computations, one can see that there exists a constant $K > 0$, only depending on the compact $\mathcal{K}$, such that, for $\varepsilon > 0$ small enough,

$$(1 - K\varepsilon)\,\mathbf{Q} \leq f^{n_\varepsilon}(\mathbf{M}) \leq (1 + K\varepsilon)\,\mathbf{Q}.$$

The previous inequality implies at once (18).

Applying $f^l$, $l \geq 0$, to (18), and taking into account that $\mathbf{Q}$ is a fixed point of $f$, one deduces that

$$\Big(1 - \varepsilon\Big)\,\mathbf{Q} \leq f^{l+n_\varepsilon}(\mathbf{M}) \leq \Big(1 + \varepsilon\Big)\,\mathbf{Q}.$$

This is nothing else but the definition of the convergence of the sequence $(f^l(\mathbf{M}))_{l\geq 0}$ to $\mathbf{Q}$. ■

*2) Convergence results for $(S)_{cont}$:* Let $t \to \mathbf{M}(t)$, $t \geq 0$, be a trajectory of $(S)_{cont}$ with initial condition $\mathbf{M}_0 \in \mathcal{D}$.

Thanks to equation (II.27) which appears in the proof of proposition V.1 in Appendix II, we have for every trajectory $\mathbf{M}(t)$ of $(S)_{cont}$

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{M}\|^2 = 2\,\mathrm{Tr}(\mathbf{M}\dot{\mathbf{M}}) = 2\,\mathrm{Tr}\left(\nabla F(\mathbf{M}).\mathbf{M}\right) = 0.$$

Then, for every $t \geq 0$, $\mathbf{M}(t)$ keeps a constant norm equal to $\|\mathbf{M}_0\|$. Moreover, one has for every $t \geq 0$

$$F(\mathbf{M}(t)) - F(\mathbf{M}(0)) = \int_0^t \frac{\mathrm{d}}{\mathrm{d}t} F(\mathbf{M}) = \int_0^t \|\nabla F(\mathbf{M})\|^2 > 0.$$

Since $F$ is bounded over $\mathcal{D}(\|\mathbf{M}_0\|)$, we deduce that

$$\int_0^{+\infty} \|\nabla F(\mathbf{M})\|^2 < +\infty. \tag{19}$$

In addition, since $t \to F(\mathbf{M}(t))$ is an increasing function, then $\mathbf{M}(t)$ remains in a compact subset $\mathcal{K}$ of $\mathcal{D}(\|\mathbf{M}_0\|)$ which is independent of the time $t$. As $\mathcal{D}(\|\mathbf{M}_0\|)$ contains a unique equilibrium point of $(S)_{cont}$, we proceed by proving theorem IV.2 in the continuous case

$$\forall \mathbf{M}_0 \in \mathcal{D}, \mathbf{M}(t) \xrightarrow[t \to +\infty]{} \|\mathbf{M}_0\| \, \widehat{\mathbf{M}}_{FP}. \tag{20}$$

Without loss of generality, we assume that $\|\mathbf{M}_0\| = 1$. Let $F_0$ be the limit of $F(\mathbf{M}(t))$ as $t$ tends to $+\infty$. Thanks to Theorem IV.1 and the fact that $\|\mathbf{M}(t)\|$ is constant, it is easy to see that (20) follows if one can show that $F_0 = F(\widehat{\mathbf{M}}_{FP})$. We assume the contrary and will reach a contradiction.

Indeed, if we assume that $F_0 < F(\widehat{\mathbf{M}}_{FP})$ then there exists $\varepsilon_0$ such that $\|\mathbf{M}(t) - \widehat{\mathbf{M}}_{FP}\| \geq \varepsilon_0$, for every $t \geq 0$. This implies together with the fact that $\widehat{\mathbf{M}}_{FP}$ is the unique fixed point of $f$ in $\mathcal{D}(1)$ and $\|\nabla F(\mathbf{M})\|$ is continuous, that there exists $C_0$ such that $\|\nabla F(\mathbf{M})\| \geq C_0$, for every $t \geq 0$. Then, $\int_{t_0}^{+\infty} \|\nabla F(\mathbf{M})\|^2 = +\infty$, which contradicts (19). Therefore, (20) holds true. ■

## VI. Simulations

The main purpose of this section is to give some tools for computing of the FP estimate regardless of its statistical properties; in particular, we investigate the numerical accuracy and the algorithm convergence in different contexts for the complex case.

The two algorithms presented in section IV will be compared:

- the discrete case algorithm of theorem IV.2, called algorithm 1 in the sequel, defined by (12) and whose convergence to the FP estimate has been proved in Section V;
- the normalized algorithm, called algorithm 2 in the sequel, defined by (14).

The first purpose of simulations is to compare the two algorithms in order to choose the best one in terms of convergence speed.

Secondly, we study the parameters influence in the retained algorithm: the order $m$ of matrix $\mathbf{M}$, the number $N$ of reference data $(\mathbf{c}_1, \ldots, \mathbf{c}_N)$ and the algorithm starting point. Note that the distribution of the $\mathbf{c}_i$'s has no influence on the simulations because of the independence of equation (3) (which completely

defines the FP estimate) with respect to the distribution of the $\tau_i$'s. Thus, without loss of generality, the Gaussian distribution will be used in the sequel.

Convergence will be analyzed by evaluating the widely used criterion $C$

$$C(k) = \frac{\|\widehat{\mathbf{M}}_{k+1} - \widehat{\mathbf{M}}_k\|}{\|\widehat{\mathbf{M}}_k\|} \tag{21}$$

as a function of algorithm iteration $k$. The numerical limit of $C$ (when algorithm has converged) is called the floor level.

The first subsection compares algorithms 1 and 2 while the second subsection studies parameters influence.

### A. Comparison of the two Algorithms

This section is devoted to the comparison of Algorithm 1 and 2 for Toeplitz matrices which are met when the processes are stationary. We will use the set of Toeplitz matrices $\mathbf{M}$ defined by the following widely used structure:

$$M_{ij} = \rho^{|i-j|}, \tag{22}$$

for $1 \leq i, j \leq m$ and for $0 < \rho < 1$. Notice that the covariance matrix $\mathbf{M}$ is fully defined by the parameter $\rho$, which characterizes the correlation of the data.

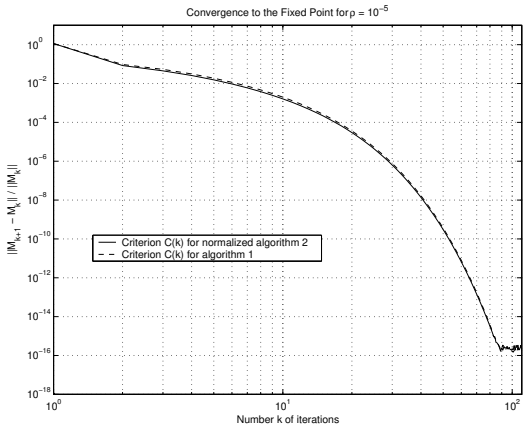
*1) Convergence behavior for different values of $\rho$:* Fig. 1 displays the criterion $C(k)$ versus the iterations number $k$ for the following set of parameters: $m = 10$, $N = 20$ and the starting point $\mathbf{M}_0 = \mathbf{I}_m$. Three typical cases are investigated: weak correlation ($\rho = 10^{-5}$, Fig. 1.a), medium correlation ($\rho = 0.9$, Fig. 1.b) and strong correlation ($\rho = 1 - 10^{-5}$, Fig. 1.c).

Fig. 1 leads to four main comments.

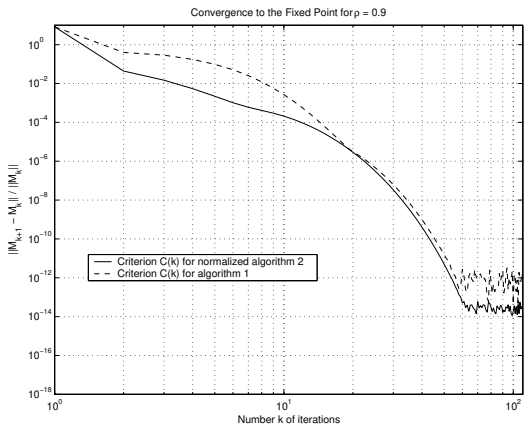
- For a given of $\rho$, both algorithms numerical convergence occurs for the same iteration number. Moreover, algorithm 2 always presents a better accuracy (in terms of floor level).
- Higher the $\rho$, faster the convergence is; for $\rho = 10^{-5}$, convergence is reached around 90 iterations, for $\rho = 0.9$, 60 iterations are enough and for $\rho = 1 - 10^{-5}$, only 20 iterations are required.
- Stronger the correlation, lower the limit accuracy is.
- The improvement of algorithm 2 in term of accuracy increases with $\rho$.

With this first analysis, we infer that algorithm 2 is better than algorithm 1.

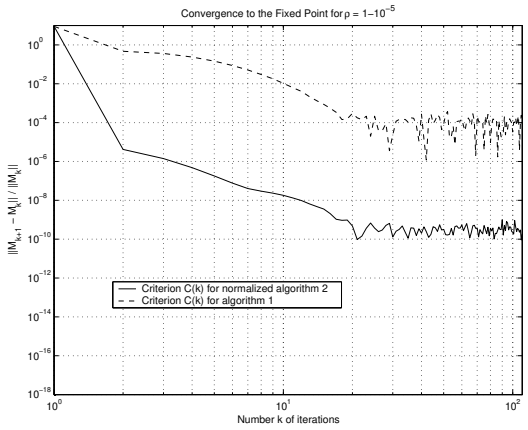
On Fig. 2, we have plotted the criterion $C$ versus $\rho$ when the convergence has occurred. Floor level is evaluated at the $150^{\text{th}}$ iteration. Both algorithms exhibit the same behavior: the floor level gets worth




(a) $\rho = 10^{-5}$




(b) $\rho = 0.9$




(c) $\rho = 1 - 10^{-5}$

Fig. 1. Convergence to the FP for three different $\rho$. a) $\rho = 10^{-5}$, b) $\rho = 0.9$, c) $\rho = 1 - 10^{-5}$

when correlation parameter $\rho$ increases. Floor level is always better for the normalized algorithm than for the algorithm 1. Moreover, the distance between the two curves increases with $\rho$.

Fig. 3 shows the required iteration number $k$ to achieve a relative error $C$ equal to $10^{-5}$. Plots are given as a function of correlation parameter $\rho$. Algorithm 1 is quite insensitive to the correlation parameter influence. The number of iteration $k$ is always close to 21. Conversely, for algorithm 2, the iteration number $k$ decreases with $\rho$, starting at $k = 20$ for small $\rho$ and ending at $k = 8$ for $\rho$ close to 1. Surprisingly, more the data are correlated, faster the convergence is (but according to Fig. 1.c, the floor level gets worse).

These results allow to conclude that algorithm 2 (normalized algorithm) is the best in all situations.

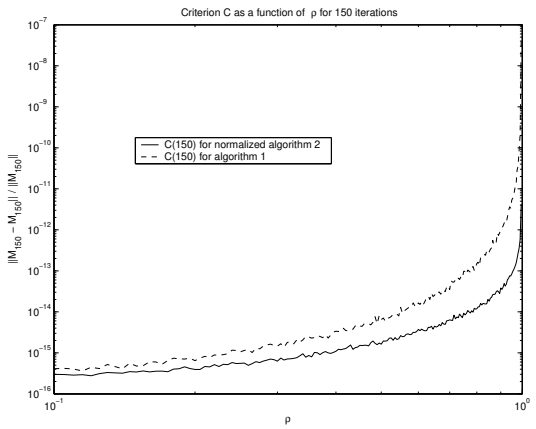



Fig. 2. Floor level, $C(150)$, against $\rho$

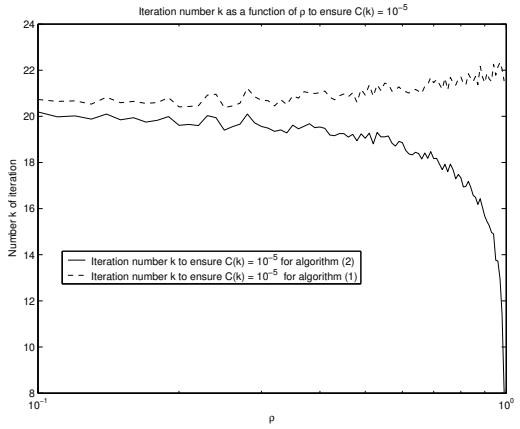


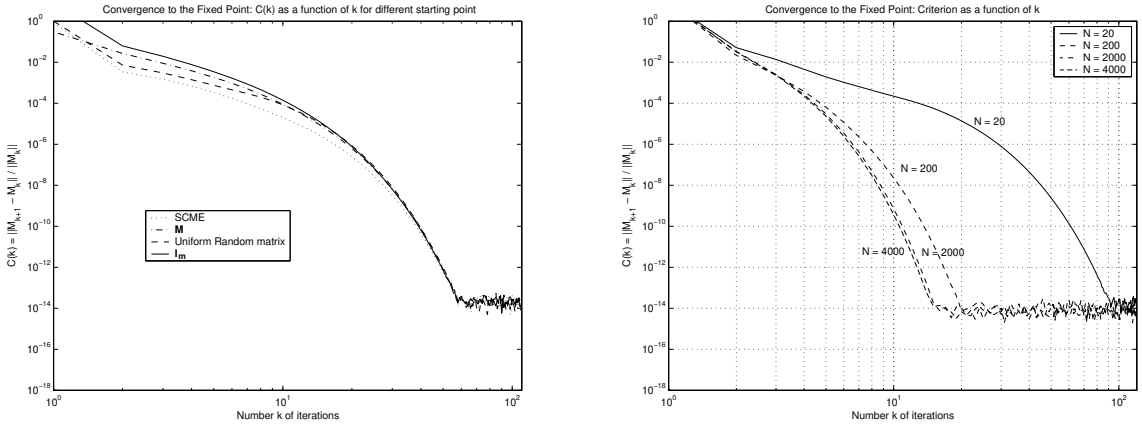

Fig. 3. Required iteration number $k$ to achieve the relative error $C = 10^{-5}$

That is why, in the sequel, we will study parameters influence on the normalized algorithm.

### *B. Parameters influence*

This section studies the influence on the normalized algorithm of the starting point $\mathbf{M}_0$ and the number $N$ of reference data.

Fig. 4.a shows the criterion $C(k)$ for four different initial conditions $\mathbf{M}_0$ and a medium correlation parameter $\rho = 0.9$: the well known Sample Covariance Matrix Estimate (SCME), the true covariance matrix $\mathbf{M}$, a random matrix whose elements are uniformly distributed and the identity matrix $\mathbf{I}_m$. Floor level and convergence speed are independent of the algorithm initialization, after 10 iterations, all the




(a) Error influence with starting points

(b) Error influence with the number $N$ of secondary data

Fig. 4. Convergence to the fixed point. a) $C(k)$ as a function of $k$ for different starting points $\mathbf{M}_0$. b) $C(k)$ as a function of $k$ for various values of $N$: 20, 200, 2000 and 4000

curves merge. Fig. 4.b represents $C(k)$ for various values of $N$: 20, 200, 2000 and 4000. Notice that convergence speed increases with $N$, while the floor level is almost independent of $N$.

## VII. Conclusion

In this work we have considered the problem of covariance matrix estimation for adaptive radar detection in compound-Gaussian clutter. The corresponding ML estimate of the covariance matrix built with secondary data is known to be the solution (if such a solution exists and is unique) of an equation for which no closed form solution is available. We have established in this paper a sound demonstration of the existence and uniqueness of this ML estimate, called FPE (Fixed Point Estimator). We have also derived two algorithms for obtaining the FPE. The convergence of each algorithm has been theoretically proved and emphasized by extensive simulations which have shown the superiority of one of them, the so-called normalized algorithm. The numerical behavior of the two algorithms in realistic scenario has been also investigated as a function of main parameters, correlation and number of reference data, highlighting their fast convergence and therefore their great practical interests. These important results will allow the use of the FPE in real radar detection scheme [15]. It remains now to analyze the statistical behavior of the FPE, preliminary results in that direction have been already obtained in [16].

## APPENDIX I
## REDUCTION OF THE COMPLEX CASE TO THE REAL CASE

Let $\mathcal{G}$ be the set of $m \times m$ definite positive Hermitian matrices and $\mathcal{S}$ the set of $2m \times 2m$ symmetric matrices. Let us define the function $g$ by

$$\begin{array}{rcccl} g & : & \mathcal{G} & \longrightarrow & \widetilde{\mathcal{D}} = g(\mathcal{G}) \subset \mathcal{S} \\ & & \mathbf{M} & \longrightarrow & g(\mathbf{M}) = \begin{pmatrix} \mathbf{M}^{(1)} & -\mathbf{M}^{(2)} \\ \mathbf{M}^{(2)} & \mathbf{M}^{(1)} \end{pmatrix}, \end{array}$$

where $\mathbf{M} = \mathbf{M}^{(1)} + j\,\mathbf{M}^{(2)}$ with $\mathbf{M}^{(1)}$, symmetric matrix, the real part of $\mathbf{M}$ and $\mathbf{M}^{(2)}$, antisymmetric matrix, the imaginary part. It is obvious that $g$ is a bijection between $\mathcal{G}$ and $\widetilde{\mathcal{D}}$. Moreover, we have the following proposition

**Proposition I.1**

$$\forall\, \mathbf{M} \in \mathcal{G}\,,\; g(f(\mathbf{M})) = f_R(g(\mathbf{M}))\,,$$

*where $f$ is given by (7) and $f_R$ by*

$$f_R(\mathbf{M}_r) = \frac{m}{N} \sum_{i=1}^{2N} \frac{\mathbf{w}_i\, \mathbf{w}_i^\top}{\mathbf{w}_i^\top\, \mathbf{M}_r^{-1}\, \mathbf{w}_i},$$

*with $\mathbf{M}_r \in \widetilde{\mathcal{D}}$, and the $2m$-vectors $\mathbf{w}_1\,, \ldots\,, \mathbf{w}_{2N}$ are defined by*

- *for the $N$ first vectors $\mathbf{w}_1\,, \ldots\,, \mathbf{w}_N$ (called $\mathbf{u}_i$ for clarity), $\mathbf{w}_i = \mathbf{u}_i = \begin{pmatrix} \mathbf{x}_i^{(1)} \\ \mathbf{x}_i^{(2)} \end{pmatrix}$,*
- *for the $N$ last vectors $\mathbf{w}_{N+1}\,, \ldots\,, \mathbf{w}_{2N}$ (called $\mathbf{v}_i$), $\mathbf{w}_{N+i} = \mathbf{v}_i = \begin{pmatrix} -\mathbf{x}_i^{(2)} \\ \mathbf{x}_i^{(1)} \end{pmatrix}$.*

*Proof:* We have

$$g(f(\mathbf{M})) = \frac{m}{N} \sum_{i=1}^{N} \frac{g(\mathbf{x}_i\, \mathbf{x}_i^H)}{\mathbf{x}_i^H\, \mathbf{M}^{-1}\, \mathbf{x}_i}.$$

Thanks to the following results: $g(\mathbf{M}^{-1}) = \Big(g(\mathbf{M})\Big)^{-1}$, $g(\mathbf{x}_i\, \mathbf{x}_i^H) = \mathbf{u}_i\, \mathbf{u}_i^\top + \mathbf{v}_i\, \mathbf{v}_i^\top$ and $\mathbf{x}_i^H \mathbf{M}^{-1} \mathbf{x}_i = \mathbf{u}_i^\top\, g(\mathbf{M}^{-1})\, \mathbf{u}_i = \mathbf{v}_i^\top\, g(\mathbf{M}^{-1})\, \mathbf{v}_i$, Proposition I.1 follows straightforwardly. ■

Hypothesis $(H)$ of Section III implies hypothesis $(H2)$ (just before Remark IV.1) of linear independence for the real problem just defined in $\mathbb{R}^{2m}$. Thanks to Theorem IV.1, there exists a unique fixed point $\mathbf{M}_R^{FP}$ (up to scalar factor) in $\mathcal{S}$. Thus, it remains to show that $\mathbf{M}_R^{FP}$ belongs to $\widetilde{D}$. Thanks to Proposition I.1, if initialization of algorithm defined in Theorem IV.2, Eqn. (12), belongs to $\widetilde{D}$, the resulting sequence

$\{\mathbf{M}_0, \ldots, \mathbf{M}_k\}$ obviously belongs to $\widetilde{D}$. Since this sequence converges in $\mathcal{S}$, by elementary topological considerations, the limit belongs to $\widetilde{D}$.

Now, since $f_R$ admits a unique fixed point $\mathbf{M}_R^{FP}$ (up to a scalar factor) in $\widetilde{D}$, the proof of Theorem III.1 is completed. Indeed, there exists a unique matrix $\mathbf{M}^{FP}$ (up to a scalar factor) which verifies

$$\mathbf{M}^{FP} = g^{-1}\left(\mathbf{M}_R^{FP}\right) = g^{-1}\left(f_R\left(\mathbf{M}_R^{FP}\right)\right) = g^{-1}\left(f_R\left(g(\mathbf{M}^{FP})\right)\right) = g^{-1}\left(g\left(f(\mathbf{M}^{FP})\right)\right) = f(\mathbf{M}^{FP}).$$

## APPENDIX II
## PROOF OF PROPOSITION V.1

If such a $\widehat{\mathbf{M}}_{FP}$ exists, then for every $\lambda > 0$, $\lambda\widehat{\mathbf{M}}_{FP}$ is also a fixed point of $f$, since $f$ is homogeneous of degree one. We start by demonstrating the following lemma.

**Lemma II.1**

*The function $F$ can be extended as a continuous function of $\overline{\mathcal{D}}\backslash\{\mathbf{0}\}$ so that, for every non invertible $\mathbf{M} \in \overline{\mathcal{D}}\backslash\{\mathbf{0}\}$, $F(\mathbf{M}) = \mathbf{0}$.*

*Proof:* It is enough to show that, for every non invertible $\mathbf{M} \in \overline{\mathcal{D}}\backslash\{\mathbf{0}\}$, and every sequence $(\mathbf{Q}^{(k)})_{k\geq 0}$ in $\overline{\mathcal{D}}$ converging to zero and so that $\mathbf{M} + \mathbf{Q}^{(k)}$ is invertible, we have

$$\forall \mathbf{Q} \in \mathcal{D}, \lim_{k\to\infty} F(\mathbf{M} + \mathbf{Q}^{(k)}) = 0.$$

Since $F$ is smooth, we may assume that $\mathbf{Q}^{(k)} \in \mathcal{D}$ for every $k \geq 0$. We introduce the notation $F^{\mathbf{c}}$ for the function $F$ in order to emphasize the dependence of $F$ with respect to the $N$-tuple $\mathbf{c} = (\mathbf{c}_1, \ldots, \mathbf{c}_N)$. If $\mathbf{R}$ is an invertible matrix, let $\mathbf{R}.\mathbf{c}$ be the $N$-tuple $\mathbf{R}.\mathbf{c} := (\mathbf{R}\mathbf{c}_1, \ldots, \mathbf{R}\mathbf{c}_N)$. Clearly one has for every $\mathbf{M} \in \mathcal{D}$,

$$F^{\mathbf{c}}(\mathbf{M}) = |\mathbf{R}|^{2N} F^{\mathbf{R}.\mathbf{c}}(\mathbf{R}\,\mathbf{M}\,\mathbf{R}^T).$$

Fix now a symmetric matrix $\mathbf{M}$ such that $\mathbf{M} \geq 0$ and the rank of $\mathbf{M}$, rk$(\mathbf{M})$, is equal to $l$, with $0 < l < m$. Thanks to the previous equation, we may assume that $\mathbf{M} = \mathbf{J}_l$, with $\mathbf{J}_l := \text{diag}(\mathbf{I}_l\ \mathbf{0}\cdots\mathbf{0})$, where $\mathbf{0}$ is repeated $m - l$ times. For $i \in [1, N]$, we write $\mathbf{c}_i$ as

$$\mathbf{c}_i = \begin{pmatrix} \mathbf{c}_i^1 \\ \mathbf{c}_i^2 \end{pmatrix}, \text{ with } \mathbf{c}_i^1 \in Im(\mathbf{J}_l) \text{ and } \mathbf{c}_i^2 \in Ker(\mathbf{J}_l).$$

According to that orthogonal decomposition, we write $\mathbf{Q}^{(k)}$ by blocks,

$$\mathbf{Q}^{(k)} = \begin{pmatrix} \mathbf{Q}_1^{(k)} & (\mathbf{Q}_2^{(k)})^T \\ \mathbf{Q}_2^{(k)} & \mathbf{Q}_3^{(k)} \end{pmatrix}.$$

Then,

$$\mathbf{M}+\mathbf{Q}^{(k)}=\begin{pmatrix}\mathbf{I}_l+\mathbf{Q}_1^{(k)} & (\mathbf{Q}_2^{(k)})^T\\ \mathbf{Q}_2^{(k)} & \mathbf{Q}_3^{(k)}\end{pmatrix}.$$

For every $k\geq 0$, set $\mathbf{P}_k:=(\mathbf{I}_l+\mathbf{Q}_1^{(k)})^{-1}$, and $\mathbf{R}_k:=\mathbf{Q}_3^{(k)}-\mathbf{Q}_2^{(k)}\mathbf{P}_k(\mathbf{Q}_2^{(k)})^T$. Then, for every $k\geq 0$, one has, after standard computations using the Schur complement formula (cf. [14] for instance), that

$$(\mathbf{M}+\mathbf{Q}^{(k)})^{-1}=\begin{pmatrix}\mathbf{P}_k+\mathbf{P}_k(\mathbf{Q}_2^{(k)})^T\mathbf{R}_k^{-1}\mathbf{Q}_2^{(k)}\mathbf{P}_k & -\mathbf{P}_k(\mathbf{Q}_2^{(k)})^T\mathbf{R}_k^{-1}\\ -\mathbf{R}_k^{-1}\mathbf{Q}_2^{(k)}\mathbf{P}_k & \mathbf{R}_k^{-1}\end{pmatrix},$$

and $\left|\mathbf{M}+\mathbf{Q}^{(k)}\right|=\left|\mathbf{I}_l+\mathbf{Q}_1^{(k)}\right||\mathbf{R}_k|$. We next compute $\mathbf{c}_i^T(\mathbf{M}+\mathbf{Q}^{(k)})^{-1}\mathbf{c}_i$ for $i\in[1,N]$ and $k\geq 0$. We get

$$\mathbf{c}_i^T(\mathbf{M}+\mathbf{Q}^{(k)})^{-1}\mathbf{c}_i=(\mathbf{c}_i^1)^T(\mathbf{P}_k+\mathbf{P}_k(\mathbf{Q}_2^{(k)})^T\mathbf{R}_k^{-1}\mathbf{Q}_2^{(k)}\mathbf{P}_k)\mathbf{c}_i^1-2(\mathbf{c}_i^1)^T\mathbf{P}_k(\mathbf{Q}_2^{(k)})^T\mathbf{R}_k^{-1}\mathbf{c}_i^2+(\mathbf{c}_i^2)^T\mathbf{R}_k^{-1}\mathbf{c}_i^2. \tag{II.23}$$

**Lemma II.2**

*With the above notations, we have*

$$\lim_{k\to\infty}\mathbf{P}_k+\mathbf{P}_k(\mathbf{Q}_2^{(k)})^T\mathbf{R}_k^{-1}\mathbf{Q}_2^{(k)}\mathbf{P}_k=\mathbf{I}_l,$$

*and, if* $\mathbf{c}_i^2\neq\mathbf{0}$*, then,*

$$\lim_{k\to\infty}\frac{\mathbf{c}_i^T(\mathbf{M}+\mathbf{Q}^{(k)})^{-1}\mathbf{c}_i}{\mathbf{c}_i^2\mathbf{R}_k^{-1}\mathbf{c}_i^2}=1. \tag{II.24}$$

*Proof:* Both results are a consequence of the following fact,

$$\lim_{k\to\infty}\mathbf{P}_k(\mathbf{Q}_2^{(k)})^T\mathbf{R}_k^{-1}\mathbf{Q}_2^{(k)}\mathbf{P}_k=0. \tag{II.25}$$

To see that, first recall that $\mathbf{S}_k:=\mathbf{Q}_3^{(k)}-\mathbf{Q}_2^{(k)}(\mathbf{Q}_1^{(k)})^{-1}(\mathbf{Q}_2^{(k)})^T$ is definite positive since $\mathbf{Q}^{(k)}$ is positive definite. Next, we write

$$\mathbf{R}_k=\mathbf{S}_k+\mathbf{Q}_2^{(k)}(\mathbf{Q}_1^{(k)})^{-1}(\mathbf{Q}_2^{(k)})^T-\mathbf{Q}_2^{(k)}\mathbf{P}_k(\mathbf{Q}_2^{(k)})^T=\mathbf{Q}_2^{(k)}(\mathbf{Q}_1^{(k)})^{-1}\mathbf{P}_k(\mathbf{Q}_2^{(k)})^T,$$

and we then have and we then have

$$\mathbf{P}_k(\mathbf{Q}_2^{(k)})^T\mathbf{R}_k^{-1}\mathbf{Q}_2^{(k)}\mathbf{P}_k=\mathbf{P}_k^{1/2}(\mathbf{Q}_1^{(k)})^{1/2}\mathbf{B}_k^T(\mathbf{I}_l+\mathbf{B}_k\mathbf{B}_k^T)^{-1}\mathbf{B}_k(\mathbf{Q}_1^{(k)})^{1/2}\mathbf{P}_k^{1/2},$$

where $\mathbf{B}_k:=\mathbf{S}_k^{-1/2}\mathbf{Q}_2^{(k)}(\mathbf{Q}_1^{(k)})^{-1/2}\mathbf{P}_k^{1/2}$.

It is now clear that (II.25) holds true if the $l \times l$ symmetric non negative matrix $\mathbf{B}_k^T(\mathbf{I}_l + \mathbf{B}_k\mathbf{B}_k^T)^{-1}\mathbf{B}_k$ is bounded. Computing the norm, we end up with

$$\|\mathbf{B}_k^T(\mathbf{I}_l + \mathbf{B}_k\mathbf{B}_k^T)^{-1}\mathbf{B}_k\|^2 = \|(\mathbf{I}_l + \mathbf{T}_k)^{-1}\mathbf{T}_k\|^2,$$

where $\mathbf{T}_k := \mathbf{B}_k\mathbf{B}_k^T \in \overline{\mathcal{D}}$. Since $(\mathbf{I}_l + \mathbf{T}_k)^{-1}\mathbf{T}_k \leq \mathbf{I}_l$, we conclude the proof of Lemma II.2. ■

We next consider the diagonalization of $\mathbf{R}_k$ in an orthonormal basis, given by

$$\mathbf{R}_k = \mathbf{U}_k^T\mathbf{D}_k\mathbf{U}_k, \text{ for } k \geq 0,$$

with $\mathbf{U}_k \in SO(m-l)$ and $\mathbf{D}_k = \text{ diag } (\varepsilon_k^{(l+1)}, \cdots, \varepsilon_k^{(m)})$. By definition, $\lim_{k\to\infty} \varepsilon_k^{(j)} = 0^+$, for every $j \in [l+1, m]$, and, with no loss of generality, we will assume that $\varepsilon_k^{(m)} = \min_{l+1\leq j\leq m} \varepsilon_k^{(j)}$ and $\lim_{k\to\infty} \mathbf{U}_k = \mathbf{U} \in SO(m-l)$.

We next establish the following lemma.

**Lemma II.3**

*Let* $\mathbf{E}_m = (0 \cdots 0\ 1)^T$ *with* $0$ *repeated* $m-l-1$ *times. With the previous notations, there exist* $C > 0$ *and* $i^* \in [1, N]$ *such that, for* $k \geq 0$ *large enough, we have*

$$|\mathbf{E}_m^T\mathbf{U}_k\mathbf{c}_{i^*}^2| \geq C. \tag{II.26}$$

*Proof:* By a continuity argument, it is enough to show the existence of an index $i^*$ so that $\mathbf{E}_m^T\mathbf{U}\mathbf{c}_{i^*}^2 \neq 0$. Moreover, according to hypothesis $(H2)$, it is not possible to find $m$ vectors $\mathbf{c}_{i(1)}, \cdots, \mathbf{c}_{i(m)}$ linearly independent such that

$$\mathbf{e}_m^T\overline{\mathbf{U}}\mathbf{c}_i = \mathbf{E}_m^T\mathbf{U}\mathbf{c}_i^2 = 0,$$

where $\mathbf{e}_m = (0 \cdots 0\ 1)^T \in \mathbb{R}^m$ and $\overline{\mathbf{U}} = \text{diag}(\mathbf{I}_l, \mathbf{U})$. (Otherwise, there exist $m$ vectors $\mathbf{c}_{i(1)}, \cdots, \mathbf{c}_{i(m)}$ linearly independent belonging to the orthogonal of $\overline{\mathbf{U}}\mathbf{e}_m$, which has dimension $m-1$.)

By a simple counting argument, the index $i^*$ therefore exists. Indeed, otherwise the $N$ vectors $\mathbf{c}_i$'s, with $i \notin S$, verify $\mathbf{e}_m^T\overline{\mathbf{U}}\mathbf{c}_i = 0$, meaning that all the vectors $\mathbf{c}_i$, $1 \leq i \leq N$, are orthogonal to $\overline{\mathbf{U}}^T\mathbf{e}_m$, which is impossible. The proof of Lemma II.3 is complete. ■

We can now finish the proof of Lemma II.1. Let $\mathbf{c}^*$ be the $(N-1)$-tuple made of the $\mathbf{c}_i$'s for $i \in [1, N]/\{i^*\}$. For every $k \geq 0$, we have

$$F^{\mathbf{c}}(\mathbf{M} + \mathbf{Q}^{(k)}) = \frac{1}{\left|\mathbf{M} + \mathbf{Q}^{(k)}\right|}\left(\frac{1}{\mathbf{c}_{i^*}^T(\mathbf{M} + \mathbf{Q}^{(k)})^{-1}\mathbf{c}_{i^*}}\right)^m F^{\mathbf{c}^*}(\mathbf{M} + \mathbf{Q}^{(k)}).$$

Since $N-1 \geq m$, we apply the result of [13] which states that the supremum of $F^{\mathbf{c}^*}$ over $\mathcal{D}$ is finite, i.e., there exists a positive constant $C^*$ such that, for every $\mathbf{R} \in \mathcal{D}$, $F^{\mathbf{c}^*}(\mathbf{R}) \leq C^*$. Therefore, the conclusion holds true if

$$\lim_{k\to\infty} \frac{1}{\left|\mathbf{M}+\mathbf{Q}^{(k)}\right|} \left( \frac{1}{\mathbf{c}_{i^*}^T(\mathbf{M}+\mathbf{Q}^{(k)})^{-1}\mathbf{c}_{i^*}} \right)^m = 0.$$

Thanks to (II.24), that amounts to show that

$$\lim_{k\to\infty} \frac{1}{|\mathbf{D}_k|} \left( \frac{1}{(\mathbf{c}_{i^*}^2)^T(\mathbf{R}_k)^{-1}\mathbf{c}_{i^*}^2} \right)^m = 0.$$

It is clear that $|\mathbf{D}_k| \geq (\varepsilon_k^{(m)})^{m-l}$. In addition, by using Lemma II.3, we can write

$$(\mathbf{c}_{i^*}^2)^T(\mathbf{R}_k)^{-1}\mathbf{c}_{i^*}^2 = (\mathbf{U}_k\mathbf{c}_{i^*}^2)^T(\mathbf{D}_k)^{-1}\mathbf{U}_k\mathbf{c}_{i^*}^2 = \xi_k \frac{(\mathbf{E}_m^T\mathbf{U}_k\mathbf{c}_{i^*}^2)^2}{\varepsilon_k^{(m)}},$$

where $\xi_k$ is bounded below and above by positive constants independent on $k$. We finally get that

$$\frac{1}{|\mathbf{D}_k|} \left( \frac{1}{(\mathbf{c}_{i^*}^2)^T(\mathbf{R}_k)^{-1}\mathbf{c}_{i^*}^2} \right)^m \leq C(\varepsilon_k^{(m)})^l,$$

with a positive constant $C$ independent of $k$. By letting $k$ go to infinity, we conclude the proof of Lemma II.1. ■

End of the proof of Proposition V.1:

Recall that $\overline{\mathcal{D}}(1)$ is a compact subset of $\overline{\mathcal{D}}\backslash\{\mathbf{0}\}$. Then $F$ is well-defined on $\overline{\mathcal{D}}(1)$ and is continuous. The application $F$ reaches its maximum over $\overline{\mathcal{D}}(1)$ at a point $\widehat{\mathbf{M}}_{FP}$. Since $F$ is strictly positive on $\mathcal{D}(1)$ and equal to zero on $\overline{\mathcal{D}}(1)\backslash\mathcal{D}(1)$, then $F(\widehat{\mathbf{M}}_{FP}) > 0$, implying that $\widehat{\mathbf{M}}_{FP} \in \mathcal{D}(1)$. We complete the proof of Proposition V.1 by establishing the next lemma.

**Lemma II.4**

*Let $\widehat{\mathbf{M}}_{FP} \in \mathcal{D}(1)$ be defined as previously. Then, $\nabla F(\widehat{\mathbf{M}}_{FP}) = \mathbf{0}$, which implies that $\widehat{\mathbf{M}}_{FP}$ is a fixed point of $f$.*

*Proof:* By definition of $\widehat{\mathbf{M}}_{FP}$, one has $F(\widehat{\mathbf{M}}_{FP}) = \max_{\mathbf{M}\in\mathcal{D}(1)} F(\mathbf{M})$. By standard calculus, it results that $\nabla F(\widehat{\mathbf{M}}_{FP})$ and $\nabla\mathcal{N}(\widehat{\mathbf{M}}_{FP})$ are colinear, where $\mathcal{N}(\mathbf{M}) = \|\mathbf{M}\|^2$ for every $\mathbf{M} \in \mathcal{D}$. Since $\widehat{\mathbf{M}}_{FP} \in \mathcal{D}$, there exists a real number $\mu$ such that $\nabla F(\widehat{\mathbf{M}}_{FP}) = \mu\,\widehat{\mathbf{M}}_{FP}$. Recall that, since $F$ is homogeneous of degree zero, then,

$$\forall \mathbf{M} \in \mathcal{D}, \quad \nabla(\mathbf{M}).\mathbf{M} = 0. \tag{II.27}$$

One deduces that $\mu = \mu\ \|\widehat{\mathbf{M}}_{FP}\|^2 = \nabla F(\widehat{\mathbf{M}}_{FP}).\widehat{\mathbf{M}}_{FP} = 0$. The proof of Lemma II.4 is complete. ■

## APPENDIX III
## PROOF OF PROPOSITION V.2

We start by establishing $(P1)$. Let $\mathbf{M}, \mathbf{Q} \in \mathcal{D}$ with $\mathbf{M} \leq \mathbf{Q}$. Then, $\mathbf{M}^{-1} \geq \mathbf{Q}^{-1}$ and, for every $1 \leq i \leq N$, we have

$$\frac{1}{\mathbf{c}_i^\top \mathbf{M}^{-1} \mathbf{c}_i} \leq \frac{1}{\mathbf{c}_i^\top \mathbf{Q}^{-1} \mathbf{c}_i}.$$

The reasoning for the case with strict inequalities is identical. Then, clearly, $(P1)$ follows.

We next turn to the proof of $(P2)$. We first recall that, for every unit vector $\mathbf{c} \in \mathbb{R}^m$, $\|\mathbf{c}\| = 1$ and $\mathbf{M} \in \mathcal{D}$, then

$$\frac{1}{\mathbf{c}^\top \mathbf{M}^{-1} \mathbf{c}} = \inf_{\mathbf{z}^\top \mathbf{c} \neq 0} \frac{\mathbf{z}^\top \mathbf{M} \mathbf{z}}{(\mathbf{c}^\top \mathbf{z})^2}, \tag{III.28}$$

and the infimum is reached only on the line generated by $\mathbf{M}^{-1} \mathbf{c}$.

Let $\mathbf{M}, \mathbf{Q} \in \mathcal{D}$. Then, one has

$$f(\mathbf{M}+\mathbf{Q}) = \frac{m}{N} \sum_{i=1}^{N} \min_{\mathbf{z}^\top \mathbf{c}_i \neq 0} \frac{\mathbf{z}^\top (\mathbf{M}+\mathbf{Q}) \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2} = \frac{m}{N} \sum_{i=1}^{N} \min_{\mathbf{z}^\top \mathbf{c}_i \neq 0} \left( \frac{\mathbf{z}^\top \mathbf{M} \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2} + \frac{\mathbf{z}^\top \mathbf{Q} \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2} \right).$$

More generally, the following holds true,

$$\min_{\mathbf{z} \in \mathcal{A}} \Big( f_1(\mathbf{z}) + f_2(\mathbf{z}) \Big) \geq \min_{\mathbf{z} \in \mathcal{A}} f_1(\mathbf{z}) + \min_{\mathbf{z} \in \mathcal{A}} f_2(\mathbf{z}),$$

for every functions $f_1, f_2$ and set $\mathcal{A}$ giving a sense to the previous inequality. Then, $(P2)$ clearly holds true. It remains to study when equality occurs in $(P2)$. That happens if and only if, for every $1 \leq i \leq N$, one has

$$\min_{\mathbf{z}^\top \mathbf{c}_i \neq 0} \left( \frac{\mathbf{z}^\top \mathbf{M} \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2} + \frac{\mathbf{z}^\top \mathbf{Q} \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2} \right) = \min_{\mathbf{z}^\top \mathbf{c}_i \neq 0} \frac{\mathbf{z}^\top \mathbf{M} \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2} + \min_{\mathbf{z}^\top \mathbf{c}_i \neq 0} \frac{\mathbf{z}^\top \mathbf{Q} \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2}. \tag{III.29}$$

Let us first show that equality occurs in (III.29) if and only if there exists some $\mu_i > 0$ such that

$$\mathbf{M}^{-1} \mathbf{c}_i = \frac{1}{\mu_i} \mathbf{Q}^{-1} \mathbf{c}_i. \tag{III.30}$$

Indeed, for every vector $\mathbf{z} \in \mathbb{R}^m$ with $\mathbf{z}^\top \mathbf{c}_i \neq 0$, we have

$$\frac{\mathbf{z}^\top (\mathbf{M}+\mathbf{Q}) \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2} \geq \frac{1}{\mathbf{c}_i^\top \mathbf{M}^{-1} \mathbf{c}_i} + \frac{\mathbf{z}^\top \mathbf{Q} \mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2}.$$

Choosing $\mathbf{z} = (\mathbf{M}+\mathbf{Q})^{-1} \mathbf{c}_i$ yields

$$\frac{1}{\mathbf{c}_i^\top \mathbf{M}^{-1} \mathbf{c}_i} + \frac{1}{\mathbf{c}_i^\top \mathbf{Q}^{-1} \mathbf{c}_i} = \frac{1}{\mathbf{c}_i^\top (\mathbf{M}+\mathbf{Q})^{-1} \mathbf{c}_i} \geq \frac{1}{\mathbf{c}_i^\top \mathbf{M}^{-1} \mathbf{c}_i} + \frac{\mathbf{c}_i^\top (\mathbf{M}+\mathbf{Q})^{-1} \mathbf{Q} (\mathbf{M}+\mathbf{Q})^{-1} \mathbf{c}_i}{(\mathbf{c}_i^\top (\mathbf{M}+\mathbf{Q})^{-1} \mathbf{c}_i)^2}.$$

Therefore, the function of $z$ given by $\dfrac{\mathbf{z}^\top \mathbf{Q}\,\mathbf{z}}{(\mathbf{c}_i^\top \mathbf{z})^2}$ reaches its minimum value $\dfrac{1}{\mathbf{c}_i^\top \mathbf{M}^{-1}\,\mathbf{c}_i}$ at $\mathbf{z} = (\mathbf{M}+\mathbf{Q})^{-1}\,\mathbf{c}_i$. Using (III.28), we get that $(\mathbf{M}+\mathbf{Q})^{-1}\,\mathbf{c}_i$ is colinear to $\mathbf{Q}^{-1}\,\mathbf{c}_i$. Exchanging $\mathbf{M}$ and $\mathbf{Q}$ and proceeding as above yields that $(\mathbf{M}+\mathbf{Q})^{-1}\,\mathbf{c}_i$ is also colinear to $\mathbf{M}^{-1}\,\mathbf{c}_i$, which finally implies that $\mathbf{M}^{-1}\,\mathbf{c}_i$ and $\mathbf{Q}^{-1}\,\mathbf{c}_i$ are themselves colinear. (III.30) is proved.

To finish the proof, one must show that all the $(\mu_i)$'s, $1 \leq i \leq N$, as defined in (III.30), are equal.

Set $\mathbf{D} := \operatorname{diag}(\mu_1, \ldots, \mu_m)$ for the first $m$ indices of $[\![1, N]\!]$. Since $\left(\mathbf{c}_1, \ldots, \mathbf{c}_m\right)$ is a basis of $\mathbb{R}^n$ and $\mathbf{M}^{-1} - \mathbf{D}^{-1}\,\mathbf{Q}^{-1}$ is equal to $\mathbf{0}$ on that basis, we deduce that $\mathbf{M}^{-1} = \mathbf{D}^{-1}\mathbf{Q}^{-1}$. Consider now another basis of $\mathbb{R}^m$ defined by $\left(\mathbf{c}_2, \ldots, \mathbf{c}_{m+1}\right)$ and set $\widetilde{\mathbf{D}} = \operatorname{diag}\left(\mu_2, \ldots, \mu_{m+1}\right)$. Reasoning as previously, we obtain that $\mathbf{M}^{-1} = \widetilde{\mathbf{D}}^{-1}\mathbf{Q}^{-1}$, which firstly implies that $\widetilde{\mathbf{D}} = \mathbf{D}$ and, secondly, that $\mu_1 = \mu_2$, $\mu_2 = \mu_3$, ..., $\mu_m = \mu_{m+1}$. Repeating that reasoning for any pair of $m$-tuples of distinct indices $(i_1, \cdots, i_m)$ of $[\![1, N]\!]$, we get that, for every $i \in [\![1, N]\!]$, $\mu_i = \mu$, yielding $\mathbf{D} = \mu\,\mathbf{I}_m$. ∎

## APPENDIX IV
## PROOF OF PROPOSITION V.3

We first establish the following fact. For every $\mathbf{Q}, \mathbf{P} \in \mathcal{D}$, we have

$$\text{If } \mathbf{Q} \geq \mathbf{P} \text{ and } f(\mathbf{Q}) = f(\mathbf{P}), \text{ then } \mathbf{Q} = \mathbf{P}. \tag{IV.31}$$

Indeed, it is clear that $\mathbf{Q} \geq \mathbf{P}$ implies that $\mathbf{P}^{-1} - \mathbf{Q}^{-1} \geq \mathbf{0}$. Therefore, for every $1 \leq i \leq N$, we have

$$\frac{1}{\mathbf{c}_i^\top \mathbf{Q}^{-1}\,\mathbf{c}_i} \geq \frac{1}{\mathbf{c}_i^\top \mathbf{P}^{-1}\,\mathbf{c}_i}.$$

Assuming $f(\mathbf{Q}) = f(\mathbf{P})$ implies that, for every $1 \leq i \leq N$, we have $\mathbf{c}_i^\top \mathbf{Q}^{-1}\,\mathbf{c}_i = \mathbf{c}_i^\top \mathbf{P}^{-1}\,\mathbf{c}_i$ i.e.

$$\mathbf{c}_i^\top \left(\mathbf{P}^{-1} - \mathbf{Q}^{-1}\right) \mathbf{c}_i = 0.$$

Since $\mathbf{P}^{-1} - \mathbf{Q}^{-1} \geq \mathbf{0}$, the previous equality says that $\left(\mathbf{P}^{-1} - \mathbf{Q}^{-1}\right) \mathbf{c}_i = \mathbf{0}$, for every $1 \leq i \leq N$. By $(H2)$, the claim (IV.31) is proved.

We now turn to the proof of Proposition V.3. We consider $\mathbf{Q}, \mathbf{P} \in \mathcal{D}$ such that $\mathbf{Q} \geq \mathbf{P}$ and $\mathbf{Q} \neq \mathbf{P}$. From what precedes, we also have that $f(\mathbf{Q}) \geq f(\mathbf{P})$ and $f(\mathbf{Q}) \neq f(\mathbf{P})$. That implies the existence of an index $i_0 \in [\![1, N]\!]$ such that

$$\xi_{i_0} := \frac{m}{N}\left(\frac{1}{\mathbf{c}_{i_0}^\top \mathbf{Q}^{-1}\,\mathbf{c}_{i_0}} - \frac{1}{\mathbf{c}_{i_0}^\top \mathbf{P}^{-1}\,\mathbf{c}_{i_0}}\right) > 0.$$

Up to a relabel, we may assume that $i_0 = 1$. We then have

$$f(\mathbf{Q}) \geq f(\mathbf{P}) + \xi_1 \, \mathbf{c}_1 \, \mathbf{c}_1^\top. \tag{IV.32}$$

Next, we will show by induction on the index $l \leq m$ that there exist $l$ positive real numbers $\xi_k, 1 \leq k \leq l$, so that

$$f^l(\mathbf{Q}) \geq f^l(\mathbf{P}) + \sum_{k=1}^{l} \xi_k \, \mathbf{c}_k \, \mathbf{c}_k^\top \tag{IV.33}$$

In the previous equation, the vectors $(\mathbf{c}_k)_{1 \leq k \leq l}$ only need to be two by two distinct among all the vectors $(\mathbf{c}_i)_{1 \leq i \leq N}$. At each step of the induction, we will have the possibility to relabel the indices in $[\![l+1, N]\!]$ in such a way to get (IV.33). The induction starts for $l = 1$ and, in this case, (IV.33) reduces to (IV.32). Therefore the induction is initialized. We then assume that (IV.33) holds true for some index $l \leq m-1$ and proceed in showing the same for the index $l+1$. It is clear that it will be a consequence of the next lemma.

**Lemma IV.1**

*Let* $1 \leq l \leq m-1, \mathbf{Q}, \mathbf{P} \in \mathcal{D}$ *such that*

$$\mathbf{Q} \geq \mathbf{P} + \sum_{k=1}^{l} \xi_k \, \mathbf{c}_k \, \mathbf{c}_k^\top, \;\; \xi_k > 0. \tag{IV.34}$$

*Then, there exists a vector of* $\{\mathbf{c}_{l+1}, \ldots, \mathbf{c}_N\}$ *(to be set equal to* $\mathbf{c}_{l+1}$*, up to a relabelling of* $\{\mathbf{c}_{l+1}, \ldots, \mathbf{c}_N\}$*) and a positive real number* $\xi_{l+1} > 0$ *such that*

$$f(\mathbf{Q}) \geq f(\mathbf{P}) + \sum_{k=1}^{l+1} \xi_k \, \mathbf{c}_k \, \mathbf{c}_k^\top. \tag{IV.35}$$

*Proof:* Using (IV.34), we have for every $j \in [\![1, N]\!]$,

$$\frac{1}{\mathbf{c}_j^\top \, \mathbf{Q}^{-1} \, \mathbf{c}_j} = \min_{\mathbf{z}^\top \mathbf{c}_j \neq 0} \frac{\mathbf{z}^\top \, \mathbf{Q} \, \mathbf{z}}{(\mathbf{z}^\top \, \mathbf{c}_j)^2} \geq \min_{\mathbf{z}^\top \mathbf{c}_j \neq 0} \left( \frac{\mathbf{z}^\top \, \mathbf{P} \, \mathbf{z}}{(\mathbf{z}^\top \, \mathbf{c}_j)^2} + \sum_{k=1}^{l} \xi_k \, \frac{(\mathbf{z}^\top \, \mathbf{c}_k)^2}{(\mathbf{z}^\top \, \mathbf{c}_j)^2} \right). \tag{IV.36}$$

Using the induction hypothesis, we also have for every $1 \leq j \leq l$, that

$$\frac{1}{\mathbf{c}_j^\top \, \mathbf{Q}^{-1} \, \mathbf{c}_j} \geq \frac{1}{\mathbf{c}_j^\top \, \mathbf{P}^{-1} \, \mathbf{c}_j} + \xi_j.$$

We next show the following claim

$(C1)$ there exists two indices, one index $j_0 \in [\![l+1, N]\!]$ and another one $k_0 \in [\![1, l]\!]$ such that $\mathbf{c}_{k_0}^\top \mathbf{Q}^{-1} \, \mathbf{c}_{j_0} \neq 0$.

Claim $(C1)$ is proved reasoning by contradiction. Therefore, let us assume that $\mathbf{c}_k^\top \, \mathbf{Q}^{-1} \, \mathbf{c}_j = 0$, for every $1 \leq k \leq l$ and $l+1 \leq j \leq N$. Since $l < m$, the vectors $(\mathbf{Q}^{-1} \, \mathbf{c}_k)$, $1 \leq k \leq l$ generate a vector space $\mathbf{V}_l$

of dimension $l$, we deduce that, for every $j \in [\![l+1, N]\!]$, $\mathbf{c}_j$ is orthogonal to $\mathbf{V}_l$ and, therefore, belongs to an $m-l$-dimensional vector space of $\mathbb{R}^m$. But there are $N-l$ indices $j$ verifying the previous fact. According to $(H2)$, these vectors $(\mathbf{c}_j)_{l+1\leq j\leq N}$ generate a vector space of dimension $\min(N-l, m)$ in $\mathbb{R}^m$. We finally get that $\min(N-l, m) \leq m-l$ . This is impossible because $N > m$ and claim $(C1)$ is proved.

We now finish the proof of Lemma IV.1. Choosing in (IV.36) $\mathbf{z} = \mathbf{Q}^{-1}\,\mathbf{c}_{j_0}$, we get

$$\frac{1}{\mathbf{c}_{j_0}^\top\,\mathbf{Q}^{-1}\,\mathbf{c}_{j_0}} \geq \frac{(\mathbf{c}_{j_0}^\top\,\mathbf{Q}^{-1})\,\mathbf{P}\,(\mathbf{Q}^{-1}\,\mathbf{c}_{j_0})}{(\mathbf{c}_{j_0}^\top\,\mathbf{Q}^{-1}\,\mathbf{c}_{j_0})^2} + \xi_{k_0}\,\frac{(\mathbf{c}_{j_0}^\top\,\mathbf{Q}^{-1}\,\mathbf{c}_k)^2}{(\mathbf{c}_{j_0}^\top\,\mathbf{Q}^{-1}\,\mathbf{c}_{j_0})^2} \geq \frac{1}{\mathbf{c}_{j_0}^\top\,\mathbf{P}^{-1}\,\mathbf{c}_{j_0}} + \xi_{j_0}$$

with $\xi_{j_0} > 0$, thanks to claim $(C1)$. It is clear that $\mathbf{c}_{j_0}$ is the vector of $\{\mathbf{c}_{l+1}, \ldots, \mathbf{c}_N\}$ needed with $\xi_{j_0}$ so that, up to relabelling, yields (IV.35). Proofs of Lemma IV.1 and Proposition V.3 are now complete.

■

## APPENDIX V
## PROOF OF PROPOSITION V.4

We first need to make precise a definition. An orbit $(\mathbf{M}_k)_{k\geq 0}$ is bounded in $\mathcal{D}$ if it is contained in a compact subset of $\mathcal{D}$, i.e., there exists $\mathbf{M}, \mathbf{P} \in \mathcal{D}$ such that, for every $k \geq 0$, $\mathbf{M} \leq \mathbf{M}_k \leq \mathbf{P}$.

We will show the following chain of implications $(A) \Rightarrow (B) \Rightarrow (C) \Rightarrow (A)$.

$(A) \Rightarrow (B)$: Trivial (simply $\mathbf{M}_0 = \mathbf{P}$).

$(B) \Rightarrow (C)$: Assume that $f$ has a bounded orbit in $\mathcal{D}$, starting at $\mathbf{M}$. Then, there exists $\mu, \mu' > 0$ such that, for every $k \geq 0$, $\mu\,\mathbf{M} \leq \mathbf{M}_k \leq \mu'\,\mathbf{M}$, for every $k \geq 0$.

Let $\mathbf{Q}$ be an arbitrary matrix of $\mathcal{D}$. Then, there exists $\lambda, \lambda' > 0$ such that $\lambda\,\mathbf{M} \leq \mathbf{Q} \leq \lambda'\,\mathbf{M}$. Using the homogeneity of degree one of $f$, property $(P1)$ and the definition of an orbit of $f$, we get, after a trivial induction, that $\lambda\,\mu\,\mathbf{M} \leq \lambda\,\mathbf{M}_k \leq \mathbf{Q}_k \leq \lambda'\,\mathbf{M}_k \leq \lambda'\,\mu'\,\mathbf{M}$, for every $k \geq 0$. Then, the orbit associated to $\mathbf{Q}$ is bounded in $\mathcal{D}$.

$(C) \Rightarrow (A)$: Consider an orbit $(\mathbf{M}_k)_{k\geq 0}$ of $f$ starting at $\mathbf{M} \in \mathcal{D}$ and bounded in $\mathcal{D}$. It is then contained in a compact $\mathcal{K}$ of $\mathcal{D}$. For $l \geq 1$ , set

$$\mathbf{Q}_l := \frac{1}{l}\sum_{i=1}^{l} \mathbf{M}_i.$$

Then, the sequence $(\mathbf{Q}_l)_{l\geq 1}$ is bounded in $\mathcal{D}$ because every point $\mathbf{Q}_l$ belongs to the convex hull of $\mathcal{K}$, which is itself a compact subset of $\mathcal{D}$. For every $l \geq 1$, we have by using Proposition V.2 that

$$f(\mathbf{Q}_l) = \frac{1}{l}\,f\left(\sum_{i=1}^{l} \mathbf{M}_i\right) \geq \frac{1}{l}\sum_{i=2}^{l+1} \mathbf{M}_i = \frac{1}{l}\left(\sum_{i=1}^{l+1} \mathbf{M}_i - \mathbf{M}_1\right) = \mathbf{Q}_l + \frac{\mathbf{M}_{l+1} - \mathbf{M}_1}{l}.$$

Since $(\mathbf{Q}_l)_{l\geq 1}$ is bounded in $\mathcal{D}$, we have that, up to extracting a sub-sequence, that the sequence $(\mathbf{Q}_l)$ converges to $\overline{\mathbf{Q}}$, with $\overline{\mathbf{Q}} \in \mathcal{D}$, as $l$ tends to $+\infty$. From the last equation, it follows that $f(\overline{\mathbf{Q}}) \geq \overline{\mathbf{Q}}$.

We now consider the orbit of $f$ starting at $\overline{\mathbf{Q}}$. It defines an increasing, bounded in $\mathcal{D}$ sequence. It is therefore converging in $\mathcal{D}$ to a fixed point of $f$. ■

## APPENDIX VI
## PROOF OF COROLLARY V.1

The proof of $(C1)$ goes by contradiction. Let $\mathbf{P} \in \mathcal{D}$ with $f^l(\mathbf{P}) \geq \mathbf{P}$ and $f^l(\mathbf{P}) \neq \mathbf{P}$ for some positive integer $l \geq 1$. According to Proposition V.3, we have

$$f^n(f^l(\mathbf{P})) > f^n(\mathbf{P}) \Leftrightarrow f^l(f^n(\mathbf{P})) > f^n(\mathbf{P}).$$

Set $\mathbf{Q} := f^n(\mathbf{P})$ and $g := f^l$. It is clear that $g$ is a function from $\mathcal{D}$ to $\mathcal{D}$, homogeneous of degree one and verifies properties $(P1)$ and $(P2)$ of Proposition V.2. We will show that the orbit of $g$ associated to $\mathbf{Q}$ is not bounded, which will be the desired contradiction.

We have $g(\mathbf{Q}) > \mathbf{Q}$ which is equivalent to $g(\mathbf{Q}) - \mathbf{Q}$ being positive definite. By a simple continuity argument, there exists $\varepsilon_{\mathbf{Q}} > 0$ such that

$$\varepsilon_{\mathbf{Q}}\mathbf{Q} \leq f(\mathbf{Q}) - \mathbf{Q} \Leftrightarrow f(\mathbf{Q}) \geq (1+\varepsilon_{\mathbf{Q}})\,\mathbf{Q}.$$

By a trivial induction, we have $f^k(\mathbf{Q}) \geq (1+\varepsilon_{\mathbf{Q}})^k\,\mathbf{Q}$, for every $k \geq 0$, with the right-hand side of the above inequality tending to $+\infty$ as $k$ tends to $\infty$. Therefore, the orbit of $f$ associated to $\mathbf{M}$ is not bounded.

We now prove statement $(C2)$. Let $\widehat{\mathbf{M}}_{FP}$ and $\mathbf{P}_2$ be two fixed points of $f$. Applying $(P2)$, we have

$$f(\widehat{\mathbf{M}}_{FP} + \mathbf{P}_2) \geq f(\widehat{\mathbf{M}}_{FP}) + f(\mathbf{P}_2) = \widehat{\mathbf{M}}_{FP} + \mathbf{P}_2,$$

According to $(C1)$ above, we have that $\widehat{\mathbf{M}}_{FP} + \mathbf{P}_2$ is also a fixed point of $f$ and therefore, we have equality in $(P2)$. It implies that $\widehat{\mathbf{M}}_{FP}$ and $\mathbf{P}_2$ are colinear. The proofs of Corollary V.1 is complete and it concludes the argument of Theorem IV.1.

## APPENDIX VII
## PROOF OF LEMMA V.1

The argument goes by contradiction. We thus assume that $\omega(\mathbf{M})$ does not contain any periodic orbit. Let $\mathcal{K}$ be a compact subset of $\mathcal{D}$ containing both the orbit associated to $\mathbf{M}$ and $\omega(\mathbf{M})$.

Let $\mathbf{Q} \in \omega(\mathbf{M})$. Then, there exists a sequence $(f^{n_j}(\mathbf{M}))_{j\geq 0}$ converging to $\mathbf{Q}$, as $j$ tends to $+\infty$, with $(n_j)_{j\geq 0}$ a strictly increasing sequence of integers tending to $+\infty$.

Let $\varepsilon \in (0,1)$ small enough and $n_{j_0} \in \mathbb{N}$ such that $\|f^{n_{j_0}}(\mathbf{M}) - \mathbf{Q}\| \leq \varepsilon$. It is easy to see that there exists a constant $K$ only depending on $\mathcal{K}$ such that $(1-K\varepsilon)\mathbf{Q} \leq f^{n_{j_0}}(\mathbf{M}) \leq (1+K\varepsilon)\mathbf{Q}$. Using Proposition V.2, we have for every $p \geq 0$,

$$(1-K\varepsilon) f^p(\mathbf{Q}) \leq f^{n_0+p}(\mathbf{M}) \leq (1+K\varepsilon) f^p(\mathbf{Q}). \tag{VII.37}$$

Since $\mathbf{Q}$ is a cluster point for the orbit associated to $\mathbf{M}$, there exists $n_{j_1} \geq 0$ such that

$$\left(1 - \frac{K\varepsilon}{4}\right)\mathbf{Q} \leq f^{n_1}(\mathbf{M}) \leq \left(1 + \frac{K\varepsilon}{4}\right)\mathbf{Q}.$$

Using (VII.37) and the previous equation, there exists $p$ large enough such that

$$\left(1 - \frac{K\varepsilon}{2}\right)\mathbf{Q} \leq f^p(\mathbf{Q}) \leq \left(1 + \frac{K\varepsilon}{2}\right)\mathbf{Q}. \tag{VII.38}$$

We set $\mathbf{Q}_0 := \mathbf{Q}$ and $\varepsilon_0$ "maximal" with respect to (VII.38), i.e., $\varepsilon_0$ is the smallest positive real number so that $\Big(1-\varepsilon_0\Big)\mathbf{Q}_0 \leq f^p(\mathbf{Q}_0) \leq \Big(1+\varepsilon_0\Big)\mathbf{Q}_0$ holds true. Then, $\varepsilon_0 \leq \dfrac{K\varepsilon}{2}$ and one of the two previous inequalities is not strict, by maximality of $\varepsilon_0$. Moreover, $\varepsilon_0 > 0$. Indeed, if it were not the case, then $\mathbf{Q}_0$ and $f^p(\mathbf{Q}_0)$ would be comparable and, according to Corollary V.1, the orbit associated to $\mathbf{Q}_0$ would be periodic. We now consider the subset $V$ of $\omega(\mathbf{M})$, made of the matrices $\mathbf{P}$ such that there exists $\varepsilon(\mathbf{P}) > 0$ such that

$$\Big(1-\varepsilon(\mathbf{P})\Big)\mathbf{P} \leq f^p(\mathbf{P}) \leq \Big(1+\varepsilon(\mathbf{P})\Big)\mathbf{P}, \tag{VII.39}$$

and $\varepsilon(\mathbf{P})$ is "maximal" with respect to (VII.39).

We showed previously that $V$ is not empty since $\mathbf{Q} \in V$. We next show that $\overline{\varepsilon} = \inf_{\mathbf{P}\in V} \varepsilon(\mathbf{P}) = 0$.

By definition of $\overline{\varepsilon}$, there exists two sequences $\left(\mathbf{Q}^{(j)}\right)_{j\geq 0}$ and $\left(\varepsilon(\mathbf{Q}^{(j)})\right)_{j\geq 0}$ such that $\left(\varepsilon(\mathbf{Q}^{(j)})\right)_{j\geq 0}$ converges to $\overline{\varepsilon}$, as $j$ tends to $+\infty$. Up to considering a subsequence in the compact $\omega(\mathbf{M})$, we may assume that $\left(\mathbf{Q}^{(j)}\right)_{j\geq 0}$ converges to some $\overline{\mathbf{Q}} \in \omega(\mathbf{M})$. Passing to the limit in (VII.39), we get

$$\Big(1-\overline{\varepsilon}\Big)\overline{\mathbf{Q}} \leq f^p(\overline{\mathbf{Q}}) \leq \Big(1+\overline{\varepsilon}\Big)\overline{\mathbf{Q}}. \tag{VII.40}$$

If $\overline{\varepsilon} > 0$, then necessarily $\overline{\mathbf{Q}} \in V$ and $\overline{\varepsilon}$ is "maximal" with respect to (VII.40). Since $f$ is eventually strictly increasing, we get $\Big(1-\overline{\varepsilon}\Big) f^n(\overline{\mathbf{Q}}) < f^p\left(f^n(\overline{\mathbf{Q}})\right) < \Big(1+\overline{\varepsilon}\Big) f^n(\overline{\mathbf{Q}})$. Setting $\widetilde{\mathbf{Q}} := f^n(\overline{\mathbf{Q}})$, then $\widetilde{\mathbf{Q}}$ belongs to $\omega(\mathbf{M})$ since the latter is an invariant set with respect to $f$. Choosing $\widetilde{\varepsilon}$ "maximal" with respect to

$$\Big(1-\widetilde{\varepsilon}\Big)\widetilde{\mathbf{Q}} \leq f^p(\widetilde{\mathbf{Q}}) \leq \Big(1+\widetilde{\varepsilon}\Big)\widetilde{\mathbf{Q}},$$

we first have that $\widetilde{\varepsilon} > 0$ (otherwise we would have a periodic orbit) and $\widetilde{\varepsilon} < \overline{\varepsilon}$. We finally proved that $\widetilde{\mathbf{Q}} \in V$ with $0 < \widetilde{\varepsilon} = \varepsilon(\widetilde{\mathbf{Q}}) < \overline{\varepsilon}$. This is a contradiction with the minimality of $\overline{\varepsilon}$. Therefore, $\overline{\varepsilon} = 0$, which implies that $\overline{\mathbf{Q}} = f^p(\overline{\mathbf{Q}})$, i.e. $\omega(\mathbf{M})$ contains a periodic orbit. Lemma V.2 is proved. ■

## APPENDIX VIII
## PROOF OF LEMMA V.2

Let $\mathbf{M}_1, \mathbf{M}_2 \in \mathcal{D}$ whose associated orbits are periodic, with respective (positive) periods $l_1$ and $l_2$.

We first show that $\mathbf{M}_1$ and $\mathbf{M}_2$ are colinear, which will imply that $l_1 = l_2$.

For $i = 1, 2$, the orbit associated to $\mathbf{M}_i$ is the set $\left\{\mathbf{M}_i, f(\mathbf{M}_i), \ldots, f^{l_i - 1}(\mathbf{M}_i)\right\}$. Consider $\mathbf{M} := \mathbf{M}_1 + \mathbf{M}_2$ and $l := l_1 l_2$. Then, $f(\mathbf{M}) = f(\mathbf{M}_1 + \mathbf{M}_2) \geq f(\mathbf{M}_1) + f(\mathbf{M}_2)$ and, for every $k \geq 0$, we have

$$f^k(\mathbf{M}) \geq f^k(\mathbf{M}_1) + f^k(\mathbf{M}_2).$$

It implies that $f^l(\mathbf{M}) \geq f^l(\mathbf{M}_1) + f^l(\mathbf{M}_2) = \mathbf{M}_1 + \mathbf{M}_2 = \mathbf{M}$. By Corollary V.1, we get that $f^l(\mathbf{M}) = \mathbf{M}$. It implies that all the previous inequalities must be in fact equalities and, in particular, we have $f(\mathbf{M}) = f(\mathbf{M}_1) + f(\mathbf{M}_2)$. By $(P2)$, we deduce that $\mathbf{M}_1$ and $\mathbf{M}_2$ are colinear. It remains to show that a periodic orbit reduces to a single point.

Consider $\mathbf{M} \in \mathcal{D}$ such that

$$\begin{cases} l \geq 1, & f^l(\mathbf{M}) = \mathbf{M}, \\ (\text{if } \ l = 1, \text{no condition}) & f^{l-1}(\mathbf{M}) \neq \mathbf{M}. \end{cases}$$

We have to prove that $l = 1$.

Since the orbit associated to every $f^j(\mathbf{M})$, $0 \leq j \leq l$, is again $\omega(\mathbf{M})$ and thus finite, we deduce that $f^j(\mathbf{M})$ must be colinear to $\mathbf{M}$, according to what precedes. Then, for every $0 \leq j \leq l - 1$, we have $f^j(\mathbf{M}) = \lambda_j \mathbf{M}$, for some $\lambda_j > 0$. Obviously, $\lambda_0 = \lambda_l = 1$. In particular, we have $f(\mathbf{M}) = \lambda_1 \mathbf{M}$, implying that, either $f(\mathbf{M}) \leq \mathbf{M}$ or $f(\mathbf{M}) \geq \mathbf{M}$. By $(C1)$ of Corollary V.1, we get that $\mathbf{M}$ is a fixed point of $f$. The proof of Lemma V.1 is complete. ■

## REFERENCES

[1] E. Conte, M. Lops and G. Ricci, "Asymptotically optimum radar detection in compound-Gaussian clutter", *IEEE Trans. Aerosp. Electron. System*, vol. 31, no. 2, pp. 617-625, Apr. 1995.

[2] 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.

[3] 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.

[4] E. Jay, Détection en environnement non-Gaussien, *Ph.D. Thesis*, University of Cergy-Pontoise / ONERA, France, Jun. 2002.

[5] K. Yao, "A representation theorem and its applications to spherically invariant random processes", *IEEE Trans. Inform. Theory*, vol. 19, no. 5, pp. 600-608, Sep. 1973.

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

[7] E. J. Kelly "An adaptive detection algorithm", *IEEE Trans. Aerosp. Electron. System*, vol. 23, no. 1, pp. 115-127, Nov. 1986.

[8] F. C. Robey, D. R. Fuhrmann, E. J. Kelly and R. Nitzberg, "A CFAR adaptive matched filter detector", *Trans. Aerosp. Electron. System*, vol. 23, no. 1, pp. 208 - 216, Jan. 1992.

[9] E. Conte, M. Lops and G. Ricci, "Adaptive radar detection in compound-Gaussian clutter", *Proc. of the European Signal Processing Conf.*, Edinburgh, Scotland, Sep. 1994.

[10] F. Gini, M. V. Greco and L. Verrazzani, "Detection problem in mixed clutter environment as a Gaussian problem by adaptive pre-processing, *Electronics Letters*, vol. 31, no. 14, pp. 1189-1190, Jul. 1995.

[11] R. S. Raghavan and N. B. Pulsone, "A generalization of the adaptive matched filter receiver for array detection in a class of a non-Gaussian interference", *Proc. of the Adaptive Sensor Array Processing (ASAP) Workshop*, Lexinton, MA, pp. 499-517, Mar. 1996.

[12] F. Gini and M. V Greco, "Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter", *Signal Processing*, special section on Signal Processing with Heavy Tailed Distributions, vol. 82, no. 12, pp. 1847-1859, Dec. 2002.

[13] 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. Signal Process.*, vol. 50, no. 8, pp. 1908-1915, Aug. 2002.

[14] R. A. Horn and Ch. R. Johnson, *"Matrix analysis'"*, Cambridge University Press, Cambridge, U.K., 1985.

[15] 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, Austria, pp. 2143-2146, Sep. 2004.

[16] 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, Pennsylvania, USA, vol. IV, pp. 69-72, Mar. 2005.