ROBUSTNESS ANALYSIS OF COVARIANCE MATRIX ESTIMATES

M. Mahot[123], *P. Forster*[3], *J. P. Ovarlez*[12] *and F. Pascal*[1],

[1] SONDRA, Supelec
3 rue Joliot-Curie
91190 Gif-sur-Yvette, France
phone: + 33 1 6985 1817,
melanie.mahot@supelec.fr

[2] ONERA, DEMR/TSI
Chemin de la Hunière,
91761 Palaiseau Cedex, France

[3] SATIE, ENS Cachan, CNRS, UniverSud
61, Av. du Pdt Wilson,
F-94230 Cachan, France

ABSTRACT

Standard covariance matrix estimation procedures can be very affected by either the presence of outliers in the data or some mismatch in their statistical model. In the Spherically Invariant Random Vectors (SIRV) framework, this paper proposes the statistical analysis of the Normalized Sample Covariance Matrix (NSCM) and the Fixed Point (FP) estimates in disturbances context. The main contribution of this paper is to theoretically derive the bias of the NSCM and the FP arising from disturbances in the data used to build these estimates. The superiority of these two estimates is then highlighted in Gaussian or SIRV noise corrupted by strong deterministic disturbances. This robustness can be helpful for applications such as adaptive radar detection or sources localization methods.

1. INTRODUCTION

Many signal processing applications require the estimation of the data covariance matrix. This is the case for instance for source localization techniques such as conventional beamforming and high resolution methods (CAPON, MUSIC, ESPRIT,...) [1, 2, 3]. Adaptive radar and sonar detection methods also depend on the noise covariance matrix estimate [4]. In these cases, the estimation accuracy has a strong influence on the resulting performance. However, standard estimation process can be very affected by either the presence of outliers in the data or some mismatch on their statistical model.

In the conventional Gaussian framework, the well-known Sample Covariance Matrix (SCM) [5] is the Maximum Likelihood Estimate (MLE) and is therefore widely used for its good statistical properties : unbiasedness, efficiency, asymptotic Gaussianity,... Unfortunately, this estimate may perform poorly when the noise is not Gaussian anymore. One of the most general and elegant non-Gaussian noise model is provided by the so-called Spherically Invariant Random Vectors (SIRV). Indeed, these models encompass a large number of non-Gaussian distributions, including the Gaussian one. Within this modeling, it has been shown that the Normalized Sample Covariance Matrix (NSCM) and the Fixed Point (FP) are appropriate in terms of statistical performance [6, 7]. Moreover we will show in this paper that the NSCM and FP are also less sensitive to disturbances (outliers) than the SCM.

The authors would like to thank the Direction Générale de l'Armement (DGA) to fund this project.

More precisely, one of the contributions of this paper is to derive the theoretical bias of the NSCM and the FP arising from disturbances. The paper is organized as follows : section 2 formulates the problem while section 3 provides the main results. In section 4, simulations validate the theoretical analysis and illustrate the robustness of these estimates. Finally, section 5 concludes this work.

2. PROBLEM STATEMENT

A SIRV is a non-homogeneous Gaussian process with random power. More precisely, a SIRV [8] is the product of the square root of a positive random variable τ (*texture*), and an m-dimensional independent complex Gaussian vector $\mathbf{x}$ (*speckle*) with zero mean, covariance matrix $\mathbf{M} = E(\mathbf{x}\mathbf{x}^H)$ normalized according to $\mathrm{Tr}(\mathbf{M}) = m$:

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

Nowadays, SIRVs are increasingly used to model impulsive noise. In most applications, the speckle covariance matrix is of great importance (e.g adaptative detection in radar/sonar) and must be estimated if unknown. For that purpose, N independent snapshots $\mathbf{y}_1, ..., \mathbf{y}_N$ are usually available. Ideally, these N data should share the same distribution as $\mathbf{c}$ in (1).

However, in many situations, it may happen that some of these data, let us say the K first $\mathbf{y}_1, ..., \mathbf{y}_K$, are outliers with a different distribution than $\mathbf{c}$. Thus, $\mathbf{y}_1, ..., \mathbf{y}_N$ may be split into two sets :

$$\begin{cases} \mathbf{y}_k = \mathbf{p}_k & \text{for } 1 \leq k \leq K; \\ \mathbf{y}_n = \mathbf{c}_n = \sqrt{\tau_n}\,\mathbf{x}_n & \text{for } K < n \leq N; \end{cases} \tag{2}$$

where $\mathbf{c}_n$, τ_n and $\mathbf{x}_n$ share the same distribution as $\mathbf{c}$, τ and $\mathbf{x}$.

In this paper, the outliers $\mathbf{p}_k$ will be assumed to be random vectors with arbitrary distributions, and our purpose is to study the robustness of two speckle covariance matrix estimates : the NSCM and the FP. The NSCM, originally introduced in [9, 10], is defined by :

$$\widehat{\mathbf{M}}_{\mathrm{NSCM}} = \frac{m}{N}\sum_{n=1}^{N}\frac{\mathbf{y}_n\mathbf{y}_n^H}{\|\,\mathbf{y}_n\,\|^2}. \tag{3}$$

Its statistical properties have been derived in [6] in an ideal outlier-free context : (3) is a biased estimate of $\mathbf{M}$,

© EURASIP, 2010

unless $\mathbf{M}$ is the identity matrix.

The FP estimate [11, 12, 13], defined as the unique solution of the following equation

$$\widehat{\mathbf{M}}_{FP} = \frac{m}{N}\sum_{n=1}^{N}\frac{\mathbf{y}_n\mathbf{y}_n^H}{\mathbf{y}_n^H\widehat{\mathbf{M}}_{FP}^{-1}\mathbf{y}_n}, \tag{4}$$

is obtained in practice, by an appropriate convergent algorithm [11]. It exhibits good statistical performance (consistency, unbiasedness and asymptotic gaussianity) [6] in the outlier-free case.

3. MAIN RESULTS

Let $\widehat{\mathbf{M}}$ denote a speckle covariance matrix estimate (NSCM, FP, ...). The goal of this section is to derive its robustness to outliers within the framework (2). For that purpose, the additionnal bias $\mathbf{\Delta}$ due to the presence of outliers is defined as

$$\mathbf{\Delta} = E(\widehat{\mathbf{M}}|K\text{outliers}) - E(\widehat{\mathbf{M}}|\text{ no outlier}). \tag{5}$$

In the sequel, we will simply write $E(\widehat{\mathbf{M}})$ for $E(\widehat{\mathbf{M}}|K$ outliers). Two results will be provided for each estimate : a general expression of $\mathbf{\Delta}$, and a more specific one valid only for a particular type of outliers :

$$\mathbf{p}_k = \mathbf{m}_k + \mathbf{c}_k, \tag{6}$$

where $\mathbf{m}_k$ is deterministic and much stronger than the SIRV component : $\| \mathbf{m}_k \| >> \| \mathbf{c}_k \|$. Expression (6) accounts for the so-called "data contamination" case, often met in some adaptative detection problems. In the sequel, it will be referred to as the "data contamination model" for outliers.

Theorem 3.1
Let us denote $\mathbf{M}_{NSCM} = E(\widehat{\mathbf{M}}_{NSCM}|\textit{no outliers})$. The additional bias (5) due to outliers is given by :

$$\mathbf{\Delta}_{NSCM} = -\frac{K}{N}\mathbf{M}_{NSCM} + \frac{m}{N}\sum_{k=1}^{K}E\left[\frac{\mathbf{p}_k\,\mathbf{p}_k^H}{\|\mathbf{p}_k\|^2}\right]. \tag{7}$$

For the data contamination model (6), the term $E\left[\frac{\mathbf{p}_k\,\mathbf{p}_k^H}{\|\mathbf{p}_k\|^2}\right]$ in (7) is given by :

$$\begin{aligned} E\left[\frac{\mathbf{p}_k\,\mathbf{p}_k^H}{\|\mathbf{p}_k\|^2}\right] &= \\ \left(1 - m\frac{E[\tau]}{\|\mathbf{m}_k\|^2} + E[\tau]\frac{\mathbf{m}_k^H\mathbf{M}\,\mathbf{m}_k}{\|\mathbf{m}_k\|^4}\right)&\frac{\mathbf{m}_k\mathbf{m}_k^H}{\|\mathbf{m}_k\|^2} \\ + \frac{E[\tau]}{\|\mathbf{m}_k\|^2}\left(\mathbf{I} - \frac{\mathbf{m}_k\mathbf{m}_k^H}{\|\mathbf{m}_k\|^2}\right)\mathbf{M}&\left(\mathbf{I} - \frac{\mathbf{m}_k\mathbf{m}_k^H}{\|\mathbf{m}_k\|^2}\right). \end{aligned} \tag{8}$$

Remark 1
When $\|\mathbf{m}_k\| \to \infty$ (very strong data contamination), $\mathbf{\Delta}_{NSCM}$ simplifies to :

$$\mathbf{\Delta}_{NSCM} = -\frac{K}{N}\mathbf{M}_{NSCM} + \frac{m}{N}\sum_{k=1}^{K}\frac{\mathbf{m}_k\,\mathbf{m}_k^H}{\|\mathbf{m}_k\|^2}. \tag{9}$$

Proof 3.1
See Appendix A.

Now let us turn to the FP estimate.

Theorem 3.2
For $N >> K$, the additionnal bias (6) due to outliers is given by

$$\mathbf{\Delta}_{FP} = \frac{m+1}{N}\left(\sum_{k=1}^{K}E\left[\frac{\mathbf{p}_k\mathbf{p}_k^H}{\mathbf{p}_k^H\mathbf{M}^{-1}\mathbf{p}_k}\right] - \frac{K}{m}\mathbf{M}\right) \tag{10}$$

In the data contamination model (6), the term $E\left[\frac{\mathbf{p}_k\mathbf{p}_k^H}{\mathbf{p}_k^H\mathbf{M}^{-1}\mathbf{p}_k}\right]$ in (10) is given by

$$\begin{aligned} E\left[\frac{\mathbf{p}_k\mathbf{p}_k^H}{\mathbf{p}_k^H\mathbf{M}^{-1}\mathbf{p}_k}\right] &= \left(1 - m\frac{E[\tau]}{\mathbf{m}_k^H\mathbf{M}^{-1}\mathbf{m}_k} + \right. \\ E[\tau]\frac{\|\mathbf{m}_k\|}{(\mathbf{m}_k^H\mathbf{M}^{-1}\mathbf{m}_k)^2}\Bigg) &\frac{\mathbf{m}_k\mathbf{m}_k^H}{\mathbf{m}_k^H\mathbf{M}^{-1}\mathbf{m}_k} + \frac{E[\tau]}{\mathbf{m}_k^H\mathbf{M}^{-1}\mathbf{m}_k}\times \\ \left(\mathbf{M} - \frac{\mathbf{m}_k\mathbf{m}_k^H}{\mathbf{m}_k^H\mathbf{M}^{-1}\mathbf{m}_k}\right)&\left(\mathbf{M} - \frac{\mathbf{m}_k\mathbf{m}_k^H}{\mathbf{m}_k^H\mathbf{M}^{-1}\mathbf{m}_k}\right). \end{aligned} \tag{11}$$

Remark 2
When $\|\mathbf{m}_k\| \to \infty$ (very strong data contamination), $\mathbf{\Delta}_{FP}$ reduces to

$$\mathbf{\Delta}_{FP} = \frac{m+1}{N}\left(\sum_{n=1}^{K}\frac{\mathbf{m}_k\mathbf{m}_k^H}{\mathbf{m}_k^H\mathbf{M}^{-1}\mathbf{m}_k} - \frac{K}{m}\mathbf{M}\right). \tag{12}$$

Proof 3.2
See Appendix B.

Expressions (7) and (10) show that both the NSCM and the FP are robust estimates since their additional biases do not depend on the outliers norm but solely on the quantities $\frac{\mathbf{p}_k}{\|\mathbf{p}_k\|}$.

This is in contrast with the widely used SCM which is already known to be a poor estimate in impulsive noise. Furthermore, when outliers are present and in a SIRV context, the resulting bias is trivially shown to be equal to

$$\mathbf{\Delta}_{SCM} = \frac{1}{N}\sum_{n=1}^{K}E\left[\mathbf{p}_k\mathbf{p}_k^H\right] - \frac{K}{N}E[\tau]\mathbf{M}, \tag{13}$$

which is obviously very sensitive to strong outliers.

4. SIMULATIONS

To illustrate previous theorems and analyze the robustness of studied estimates (SCM, NSCM, FP), we compare our theoretical values of $\mathbf{\Delta}$ with those obtained by simulations. In all cases $\mathbf{M} = \mathbf{I}$, $m = 3$, $N = 50$ and $K = 1$.

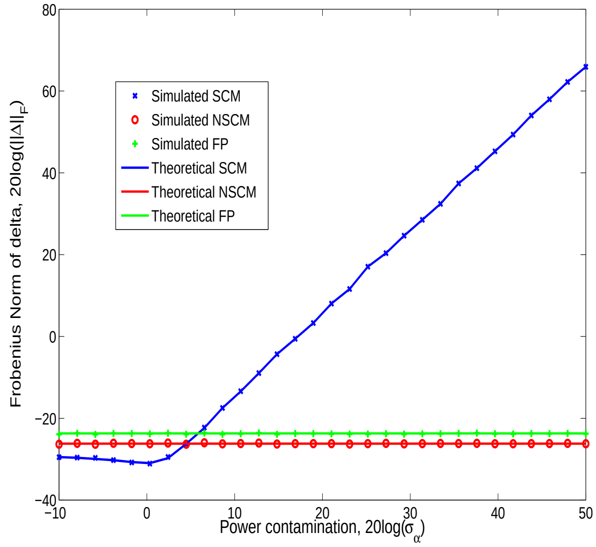

Figure 1: Frobenius norm of $\mathbf{\Delta}$ for the SCM, NSCM and FP versus disturbances power for a Gaussian noise.

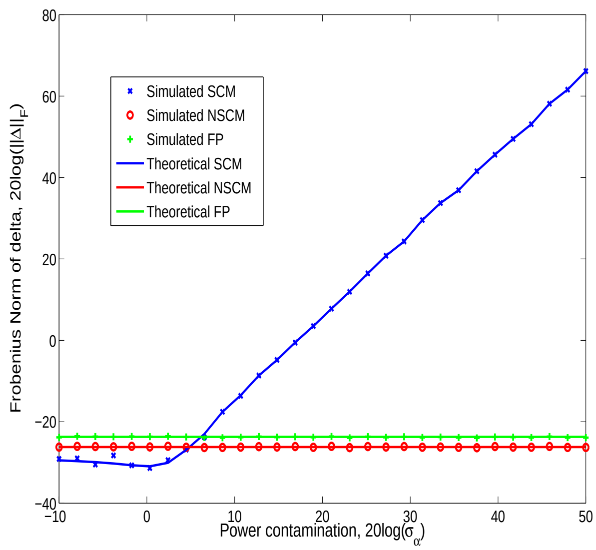

Figure 2: Frobenius norm of $\mathbf{\Delta}$ for the SCM, NSCM and FP versus disturbances power for a K-distributed noise.

In figures 1 and 2 we study the validity of the general expressions (7) and (10), for a single distubance of the form $\mathbf{p} = \alpha \mathbf{d}$ where α is a random Gaussian variable $\mathcal{N}(0, \sigma_\alpha^2)$ and $\mathbf{d}$ a fixed unit norm steering vector. Plots give the Frobenius norm $\|.\|_F$ (in dB) of $\mathbf{\Delta}_{SCM}$, $\mathbf{\Delta}_{NSCM}$ and $\mathbf{\Delta}_{FP}$ as a function of $20log(\sigma_\alpha)$. Figure 1 addresses the Gaussian noise case ($\tau_n = 1$ in equation (2)), while in figure 2 the noise is K-distributed (τ_n follows a Gamma distribution). The K-distribution parameter ν is equal to 0.1 which results in a highly impulsive noise.
These simulations prove the validity of the general expressions (7) and (10) of the bias. Furthermore, they show the insensitivity of the NSCM and the FP with respect to the outliers strength, while the SCM's performance is strongly degraded when the outliers power increases.

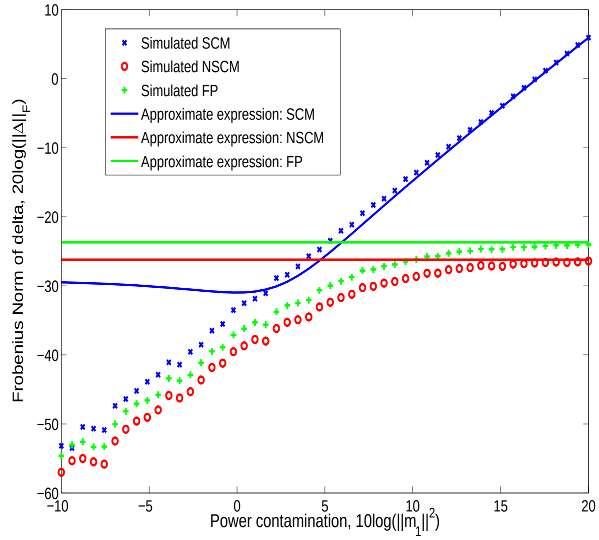

Figure 3: Frobenius norm of $\mathbf{\Delta}$ for the SCM, NSCM and FP versus disturbances power for a Gaussian noise: data contamination model.

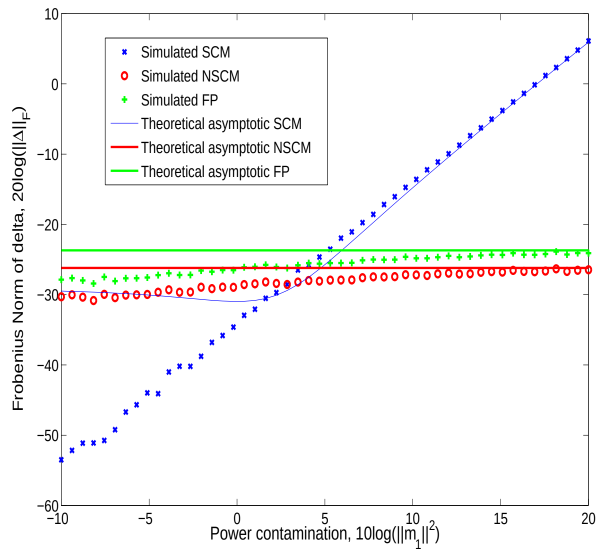

Figure 4: Frobenius norm of $\mathbf{\Delta}$ for the SCM, NSCM and FP versus disturbances power for a K-distributed noise: data contamination model.

In the next two simulations (figures 3 and 4), we investigate the domain of validity of the approximate expressions (9) and (12) in the data contamination case (6). In figure 3 the noise is Gaussian while in figure 4, it follows a

K-distribution. As expected, for large $\|\mathbf{m}_1\|$, experimental results are close to the approximate expressions. Indeed, theoretical results have been derived under the assumption that $\|\mathbf{m}_1\| >> \|\mathbf{c}_1\|$ (cf Remarks 1 and 2).
On the other hand, in figure 3, for low disturbance power, the NSCM and the FP are slightly better than the SCM. However, in figure 4 it is not the case. We also notice a better overall adequacy between experimental and approximate curves for the FP and for the NSCM estimates. We can roughly explain that behavior in the case of highly impulsive noise. Indeed, in this context, one has $\|\mathbf{m}_1\| >> \|\mathbf{c}_1\|$ with high probability, even for weak disturbances. Thus the approximate expressions remain valid in a wider domain for $\|\mathbf{m}_1\|$ than in the Gaussian case.

5. CONCLUSION

In this paper we have investigated the robustness of two covariance matrix estimates, the NSCM and the FP, when part of the data are outliers. In this context, we have derived theoretical formulas of the bias for an arbitrary distribution of the disturbances. In the data contamination case which is met in some applications, we have established simple approximate bias expressions.
These theoretical investigations have been validated by simulations results, and they demonstrate the superiority of the NSCM and the FP over the standard SCM, in terms of robustness. The results are of particular interest in applications such as adaptive radar and source localization methods.

A. PROOF OF THEOREM 3.1

(7) is trivial, therefore we provide only the proof of theorem 3.1 in the data contamination case. Let us rewrite $E\left[\frac{\mathbf{p}_k\,\mathbf{p}_k^H}{\|\mathbf{p}_k\|^2}\right]$ as

$$E\left[\frac{\mathbf{p}_k\,\mathbf{p}_k^H}{\|\mathbf{p}_k\|^2}\right] = E\left[\frac{(\mathbf{c}_k+\mathbf{m}_k)\,(\mathbf{c}_k+\mathbf{m}_k)^H}{\|\mathbf{c}_k+\mathbf{m}_k\|^2}\right],$$

and set $\varepsilon_k = \frac{\mathbf{c}_k}{\|\mathbf{m}_k\|}$. For large $\|\mathbf{m}_k\|$, a second order expansion with respect to the ε_ks leads to :

$$\begin{aligned} &E\left[\frac{\mathbf{p}_k\,\mathbf{p}_k^H}{\|\mathbf{p}_k\|^2}\right] = \\ &\left(1 - m\frac{E[\tau]}{\|\mathbf{m}_k\|^2} + E[\tau]\frac{\mathbf{m}_k^H\mathbf{M}\,\mathbf{m}_k}{\|\mathbf{m}_k\|^4}\right)\frac{\mathbf{m}_k\mathbf{m}_k^H}{\|\mathbf{m}_k\|^2} \\ &+\frac{E[\tau]}{\|\mathbf{m}_k\|^2}\left(\mathbf{I}-\frac{\mathbf{m}_k\mathbf{m}_k^H}{\|\mathbf{m}_k\|^2}\right)\mathbf{M}\left(\mathbf{I}-\frac{\mathbf{m}_k\mathbf{m}_k^H}{\|\mathbf{m}_k\|^2}\right). \end{aligned} \tag{14}$$

This concludes the proof.

B. PROOF OF THEOREM 3.2

Within the framework (2), the FP estimate (4) can be written

$$\begin{aligned} \widehat{\mathbf{M}}_{FP} &= \frac{m}{N}\sum_{n=K+1}^{N}\frac{\mathbf{c}_n\,\mathbf{c}_n^H}{\mathbf{c}_n^H\,\widehat{\mathbf{M}}_{FP}^{-1}\mathbf{c}_n} + \frac{m}{N}\sum_{k=1}^{K}\frac{\mathbf{p}_k\,\mathbf{p}_k^H}{\mathbf{p}_k^H\,\widehat{\mathbf{M}}_{FP}^{-1}\,\mathbf{p}_k} \\ &= \frac{m}{N}\sum_{n=K+1}^{N}\frac{\mathbf{x}_n\,\mathbf{c}_n^H}{\mathbf{x}_n^H\,\widehat{\mathbf{M}}_{FP}^{-1}\mathbf{x}_n} + \frac{m}{N}\sum_{k=1}^{K}\frac{\mathbf{p}_k\,\mathbf{p}_k^H}{\mathbf{p}_k^H\,\widehat{\mathbf{M}}_{FP}^{-1}\,\mathbf{p}_k} \end{aligned} \tag{15}$$

By setting $\widehat{\mathbf{R}} = \mathbf{M}^{-1/2}\widehat{\mathbf{M}}_{FP}\mathbf{M}^{-1/2}$ one obtains:

$$\widehat{\mathbf{R}} = \frac{m}{N}\sum_{n=K+1}^{N}\frac{\mathbf{z}_n\,\mathbf{z}_n^H}{\mathbf{z}_n^H\,\widehat{\mathbf{R}}^{-1}\mathbf{z}_n} + \frac{m}{N}\sum_{k=1}^{K}\frac{\mathbf{q}_k\,\mathbf{q}_k^H}{\mathbf{q}_k^H\,\widehat{\mathbf{R}}^{-1}\,\mathbf{q}_k}, \tag{16}$$

where

- $\mathbf{z}_n = \mathbf{M}^{-1/2}\,\mathbf{x}_n$ is complex gaussian distributed with zero mean and covariance matrix $\mathbf{I}$
- $\mathbf{q}_k = \mathbf{M}^{-1/2}\,\mathbf{p}_k$.

Previous equation (16) admits a unique solution such that $\mathrm{Tr}(\widehat{\mathbf{R}}) = m$. When N tends to $+\infty$ and for fixed K, equation (16) tends to

$$\mathbf{A} = mE\left[\frac{\mathbf{z}\,\mathbf{z}^H}{\mathbf{z}^H\,\mathbf{A}^{-1}\,\mathbf{z}}\right]. \tag{17}$$

where $\mathbf{z} \sim \mathscr{CN}(\mathbf{0},\mathbf{I})$.
It has been proved in [6] that the unique solution of (17) which satisfies $\mathrm{Tr}(\mathbf{A}) = m$ is $\mathbf{A} = \mathbf{I}$.

Consequently, when N tends to $+\infty$, $\widehat{\mathbf{R}}$ solution of equation (16) tends to $\mathbf{I}$, solution of equation (17). This establishes the consistency of $\widehat{\mathbf{R}}$. Thus, for large N, one has

$$\widehat{\mathbf{R}} = \mathbf{I} + \Delta\mathbf{R} \text{ where } \|\Delta\mathbf{R}\| << \|\mathbf{I}\|.$$

Assuming that $\Delta\mathbf{R}$ is small enough to ensure the validity of a first-order expansion, (16) can be written

$$\begin{aligned} \mathbf{I}+\Delta\mathbf{R} &= \frac{m}{N}\sum_{n=K+1}^{N}\frac{\mathbf{z}_n\,\mathbf{z}_n^H}{\mathbf{z}_n^H\,(\mathbf{I}-\Delta\mathbf{R})\,\mathbf{z}_n} \\ &+\frac{m}{N}\sum_{k=1}^{K}\frac{\mathbf{q}_k\,\mathbf{q}_k^H}{\mathbf{q}_k^H\,(\mathbf{I}-\Delta\mathbf{R})\,\mathbf{q}_k} \\ &\approx \mathbf{P} + \frac{m}{N}\sum_{n=K+1}^{N}\frac{\mathbf{z}_n\,\mathbf{z}_n^H}{\|\mathbf{z}_n\|^2}\left(1+\frac{\mathbf{z}_n^H\,\Delta\mathbf{R}\,\mathbf{z}_n}{\|\mathbf{z}_n\|^2}\right) \\ &+\frac{m}{N}\sum_{k=1}^{K}\frac{\mathbf{q}_k^H\,\Delta\mathbf{R}\mathbf{q}_k}{\|\mathbf{q}_k\|^4}\mathbf{q}_k\mathbf{q}_k^H \end{aligned} \tag{18}$$

where $\mathbf{P} = \frac{m}{N}\sum_{k=1}^{K}\frac{\mathbf{q}_k\,\mathbf{q}_k^H}{\|\mathbf{q}_k\|^2}$.

Since $K << N$ and $\|\Delta\mathbf{R}\| << 1$, the last term in the above equation may be neglected leading to :

$$\mathbf{I}+\Delta\mathbf{R} = \mathbf{P} + \frac{N-K}{N}\widehat{\mathbf{R}}_{NSCM} + \frac{m}{N}\sum_{n=K+1}^{N}\frac{\mathbf{z}_n\mathbf{z}_n^H}{\|\mathbf{z}_n\|^2}\frac{\mathbf{z}_n^H\Delta\mathbf{R}\mathbf{z}_n}{\|\mathbf{z}_n\|^2}$$

where $\widehat{\mathbf{R}}_{NSCM} = \frac{m}{N-K}\sum_{n=K+1}^{N}\frac{\mathbf{z}_n\,\mathbf{z}_n^H}{\|\mathbf{z}_n\|^2}$.

Now, let us define

- $\delta = \mathrm{vec}(\Delta\widehat{\mathbf{R}}), \mathbf{i} = \mathrm{vec}(\mathbf{I}), \mathbf{t}_n = \frac{\mathbf{z}_n}{\|\mathbf{z}_n\|}$,

- $\delta_{NSCM} = \text{vec}(\widehat{\mathbf{R}}_{NSCM} - \mathbf{I}) = \text{vec}(\Delta\mathbf{R}_{NSCM})$, where vec(.) denotes the operator which reshapes the $m \times n$ matrix elements into a mn column vector.

  Then, one obtains

$$\mathbf{i} + \delta = \mathbf{p} + \frac{N-K}{N}\left(\mathbf{i} + \delta_{NSCM} + \frac{m}{N-K}\sum_{n=K+1}^{N}(\mathbf{t}_n^* \otimes \mathbf{t}_n)(\mathbf{t}_n^* \otimes \mathbf{t}_n)^H \delta\right) \qquad (19)$$

where * denotes the conjugate operator and $\otimes$ the Kronecker product.

By noticing that $\text{Tr}(\Delta\mathbf{R}) = \mathbf{0}$ implies $\mathbf{i}^H\delta = 0$, the projection of equation (19) onto the orthogonal subspace of $\mathbf{i}$ gives :

$$\delta = \Pi_i^{\perp}\mathbf{p} + \frac{N-K}{N}\Pi_i^{\perp}\delta_{NSCM} + \left(\frac{m}{N}\Pi_i^{\perp}\sum_{n=K+1}^{N}(\mathbf{t}_n^* \otimes \mathbf{t}_n)(\mathbf{t}_n^* \otimes \mathbf{t}_n)^H \Pi_i^{\perp}\,\delta\right),$$

where $\Pi_i^{\perp} = \mathbf{I} - \dfrac{1}{m}\mathbf{i}\mathbf{i}^H$ and where the equality $\Pi_i^{\perp}\delta = \delta$ has been used.

This is equivalent to

$$\widehat{\alpha}\delta = \Pi_i^{\perp}\,\mathbf{p} + \frac{N-K}{N}\Pi_i^{\perp}\delta_{NSCM}. \qquad (20)$$

where $\widehat{\alpha} = \mathbf{I} - \left(\dfrac{m}{N}\Pi_i^{\perp}\displaystyle\sum_{n=K+1}^{N}(\mathbf{t}_n^* \otimes \mathbf{t}_n)(\mathbf{t}_n^* \otimes \mathbf{t}_n)^H \Pi_i^{\perp}\right)$.

It may be shown that

$$\widehat{\alpha} \xrightarrow[N\to\infty]{P} \alpha = (\mathbf{I} - \frac{1}{m+1}\Pi_i^{\perp}),$$

where $\xrightarrow{P}$ denotes the convergence in probability.

Therefore, $\widehat{\alpha}$ may be replaced by α in (20) without affecting the asymptotic distribution of δ. By noticing that

$$\alpha\delta = \frac{m}{m+1}\delta,$$

(20) leads to

$$\delta = \frac{m+1}{m}\Pi_i^{\perp}\mathbf{p} + \frac{m+1}{m}\frac{N-K}{N}\delta_{NSCM},$$

where the identity $\mathbf{i}^H\delta_{NSCM} = 0$ has been used.

Using the unvec operator (the vec inverse operator), one obtains

$$\Delta\mathbf{R} = \frac{m+1}{m}\left(\mathbf{P} - \frac{\text{Tr}(\mathbf{P})}{m}\mathbf{I}\right) + \frac{m+1}{m}\frac{N-K}{N}\Delta\mathbf{R}_{NSCM}.$$

Since $\boldsymbol{\Delta}_{FP} = E[\widehat{\mathbf{M}}_{FP} - \mathbf{M}] = E[\mathbf{M}^{1/2}\Delta\mathbf{R}\mathbf{M}^{1/2}]$ and $E[\widehat{\mathbf{R}}_{NSCM}] = \mathbf{I}$, the previous equation leads to theorem 3.2 which concludes the proof.

Starting from this general result, the additive bias obtained in the data contamination model can easily be calculated, using a similar method as in the NSCM case.

REFERENCES

[1] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, August 1969.

[2] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans.-ASSP*, vol. 34, no. 3, pp. 276–280, March 1986.

[3] R. Roy and T. Kailath, "ESPRIT-Estimation of signal parameters via rotational invariant techniques," *IEEE Trans.-ASSP*, vol. 37, no. 7, pp. 984–995, July 1989.

[4] S. M. Kay, *Fundamentals of Statistical Signal Processing - Detection Theory*, vol. 2, Prentice-Hall PTR, 1998.

[5] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons, New York, 1st edition, 1958.

[6] F. Pascal, P. Forster, J.-P. Ovarlez, and P. Larzabal, "Performance analysis of covariance matrix estimates in impulsive noise," *IEEE Trans.-SP*, vol. 56, no. 6, pp. 2206–2217, June 2008.

[7] 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.

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

[9] E. Conte, M. Lops, and G. Ricci, "Adaptive radar detection in compound-gaussian clutter," *Proceedings of Eusipco*, pp. 526–529, September 1994.

[10] F. Gini, M. V. Greco, and L. Verrazzani, "Detection problem in mixed clutter environment as a gaussian problem by adaptive pre-processing," *Electron. Lett. 31 (14)*, pp. 1189–1190, July 1995.

[11] 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*, vol. 56, no. 1, pp. 34–48, January 2008.

[12] 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, August 2002.

[13] 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, December 2002.