

Generalized conditionnal maximum likelihood estimators in the large sample regime

Eric Chaumette, François Vincent, Alexandre Renaux

► **To cite this version:**

Eric Chaumette, François Vincent, Alexandre Renaux. Generalized conditionnal maximum likelihood estimators in the large sample regime. 26th European Signal Processing Conference (EUSIPCO 2018), Sep 2018, Rome, Italy. 10.23919/eusipco.2018.8553249 . hal-01895290

HAL Id: hal-01895290

<https://hal-centralesupelec.archives-ouvertes.fr/hal-01895290>

Submitted on 15 Oct 2018

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.

GENERALIZED CONDITIONAL MAXIMUM LIKELIHOOD ESTIMATORS

Eric Chaumette⁽¹⁾, *François Vincent*⁽¹⁾ and *Alexandre Renaux*⁽²⁾

⁽¹⁾ Isae-Supaero, Université de Toulouse, 10 av. Edouard Belin, Toulouse, France ([eric.chaumette, francois.vincent]@isae.fr)

⁽²⁾ Université Paris-Sud/LSS, 3 Rue Joliot-Curie, Gif-sur-Yvette, France (renaux@lss.supelec.fr)

ABSTRACT

In modern array processing or spectral analysis, mostly two different signal models are considered: the conditional signal model (CSM) and the unconditional signal model. The discussed signal models are Gaussian and the parameters are connected either with the expectation value in the conditional case or with the covariance matrix in the unconditional one. We focus on the CSM where several independent observations of the same individual signals are available, which are allowed to perform a Gaussian random walk between observations. In the proposed generalized CSM, the parameters are connected with both the expectation value and the covariance matrix, which is a significant change in comparison with the usual CSM. Even if the batch form of the associated generalized conditional maximum likelihood estimators (GCMLEs) can be easily exhibited, it becomes uncomputable as the number of observations increases. As a main contribution, we introduce a recursive form of GCMLEs which allows to explore, by Monte-Carlo simulations, their asymptotic performance in terms of mean-squared error. We exhibit non consistent GMLEs when the number of observations tends to infinity, which highlights the consequence of combining (even slightly) dependent observations.

Index Terms— Deterministic parameter estimation, conditional maximum likelihood estimators, mean-squared error, consistency

1. INTRODUCTION

In many practical problems of interest (radar, sonar, communication, ...) dealing with deterministic parameters estimation, the observations consists of a complex circular vector [1][2]. In this instance, one of the most studied estimation problem is that of identifying the components of observations (\mathbf{y}_1) formed from a linear superposition of P individual signals (\mathbf{x}_1) to noisy data (\mathbf{v}_1) [3][4][5][6]: $\mathbf{y}_1 = \mathbf{H}_1(\boldsymbol{\theta}) \mathbf{x}_1 + \mathbf{v}_1$, $\mathbf{y}_1, \mathbf{v}_1 \in \mathcal{M}_{\mathbb{C}}(N_1, 1)$ ¹, where the mixing matrix depends on an unknown deterministic parameter vector $\boldsymbol{\theta}$. This problem has received considerable attention during the last fifty years, both for time series analysis [5] and array processing [6]. In the first case, one usually has to estimate the frequencies of complex sine waves from a single experiment data. In the second case, one looks for the directions of arrival (or spatial frequencies) of multiple plane waves impinging on a narrow-band array of sensors using

This work has been partially supported by the by the DGA/MRIS, and the iCODE institute, research project of the IDEX Paris-Saclay

¹Throughout the present paper, scalars, vectors and matrices are represented, respectively, by italic, bold lowercase and bold uppercase characters. \mathbf{I} is the identity matrix. $\mathcal{M}_{\mathbb{C}}(N, P)$ denotes the vector space of complex matrices with N rows and P columns. The matrix resulting from the vertical concatenation of k matrices $\mathbf{A}_1, \dots, \mathbf{A}_k$ of same column number is denoted $\overline{\mathbf{A}}_k$. The scalar/matrix/vector transpose conjugate is indicated by the superscript H .

multiple snapshots. These two problems have been merged into the framework of modern array processing [6][7] where mostly two different signal models are considered: the conditional signal model (CSM) and the unconditional signal model [4][8][9]. The discussed signal models are Gaussian and the angular/frequency dependency is given by parameters which are connected with the expectation value in the conditional case and with the covariance matrix in the unconditional one. In this paper, we focus on the CSM where k independent observations of \mathbf{x}_1 are available: $\mathbf{y}_l = \mathbf{H}_l(\boldsymbol{\theta}) \mathbf{x}_1 + \mathbf{v}_l$, and $\mathbf{y}_l, \mathbf{v}_l \in \mathcal{M}_{\mathbb{C}}(N_l, 1)$, $1 \leq l \leq k$, $2 \leq k$. In the standard CSM, $\mathbf{v}_l \sim \mathcal{CN}(0, \sigma_v^2 \mathbf{I})$ and the individual signals \mathbf{x}_1 are assumed to remain perfectly constant during the k observations. If one concatenates the observation vectors \mathbf{y}_l on a horizon of k observations from the first observation, i.e. $\overline{\mathbf{y}}_k^T = (\mathbf{y}_1^T, \dots, \mathbf{y}_k^T)$, then one obtains the following global CSM:

$$\overline{\mathbf{y}}_k = \overline{\mathbf{H}}_k(\boldsymbol{\theta}) \mathbf{x}_1 + \overline{\mathbf{v}}_k, \quad \overline{\mathbf{y}}_k \sim \mathcal{CN}(\overline{\mathbf{H}}_k(\boldsymbol{\theta}) \mathbf{x}_1, \sigma_v^2 \mathbf{I}), \quad (1a)$$

where $\overline{\mathbf{y}}_k, \overline{\mathbf{v}}_k \in \mathcal{M}_{\mathbb{C}}(\mathcal{N}_k, 1)$, $\overline{\mathbf{H}}_k(\boldsymbol{\theta}) \in \mathcal{M}_{\mathbb{C}}(\mathcal{N}_k, P)$, $\mathcal{N}_k = \sum_{l=1}^k N_l$. However, in a real-life experiment some experimental factors may prevent from observing perfectly constant individual signals \mathbf{x}_1 (see Section 4). In that perspective, we address the case where the P signals \mathbf{x}_1 are allowed to perform a Gaussian random walk between observations:

$$\mathbf{x}_l = \mathbf{F}_{l-1} \mathbf{x}_{l-1} + \mathbf{w}_{l-1}, \quad 2 \leq l \leq k, \quad (2a)$$

$$\mathbf{y}_l = \mathbf{H}_l(\boldsymbol{\theta}) \mathbf{x}_l + \mathbf{v}_l, \quad 1 \leq l \leq k, \quad (2b)$$

where $\mathbf{x}_l, \mathbf{w}_{l-1} \in \mathcal{M}_{\mathbb{C}}(P, 1)$, $\mathbf{F}_{l-1} \in \mathcal{M}_{\mathbb{C}}(P, P)$, and the Gaussian fluctuation noise sequence $\{\mathbf{w}_l\}_{l=1}^{k-1}$ is white and uncorrelated with the Gaussian measurement noise sequence $\{\mathbf{v}_l\}_{l=1}^k$. The Gaussian random walk (2a) of the individual signals \mathbf{x}_1 allows to define a more general class of CSM. The most noteworthy point introduced by the proposed generalized CSM, is that the parameters $\boldsymbol{\theta}$ are now connected with both the expectation value and the covariance matrix, which is a significant change in comparison with the usual CSM. Indeed, since:

$$\mathbf{x}_l = \mathbf{B}_{l,1} \mathbf{x}_1 + \sum_{q=1}^{l-1} \mathbf{B}_{l,q+1} \mathbf{w}_q, \quad \mathbf{B}_{l,q} = \begin{cases} \mathbf{F}_{l-1} \mathbf{F}_{l-2} \dots \mathbf{F}_q, & l > q \\ \mathbf{I} & , l = q \\ \mathbf{0} & , l < q \end{cases},$$

an equivalent form of (2b) is:

$$\mathbf{y}_l = \mathbf{A}_l(\boldsymbol{\theta}) \mathbf{x}_1 + \mathbf{n}_l, \quad \mathbf{A}_l(\boldsymbol{\theta}) = \mathbf{H}_l(\boldsymbol{\theta}) \mathbf{B}_{l,1},$$

$$\begin{cases} \mathbf{n}_1 = \mathbf{v}_1 \\ \mathbf{n}_l = \mathbf{v}_l + \mathbf{H}_l(\boldsymbol{\theta}) \mathbf{G}_l \overline{\mathbf{w}}_{l-1}, & 2 \leq l \leq k \end{cases}, \quad (3a)$$

where $\mathbf{G}_l \overline{\mathbf{w}}_{l-1} = \sum_{q=1}^{l-1} \mathbf{B}_{l,q+1} \mathbf{w}_q$, $\mathbf{G}_l \in \mathcal{M}_{\mathbb{C}}(P, (l-1)P)$, leading to:

$$\overline{\mathbf{y}}_k = \overline{\mathbf{A}}_k(\boldsymbol{\theta}) \mathbf{x}_1 + \overline{\mathbf{n}}_k, \quad \overline{\mathbf{y}}_k \sim \mathcal{CN}(\overline{\mathbf{A}}_k(\boldsymbol{\theta}) \mathbf{x}_1, \mathbf{C}_{\overline{\mathbf{n}}_k}(\boldsymbol{\theta})). \quad (3b)$$

As shown in Section 2, even in the simplest case where the set of unknown parameters is restricted to \mathbf{x}_1 and $\boldsymbol{\theta}$, the MLEs of $\boldsymbol{\theta}$ based on $\bar{\mathbf{y}}_k$ (3b), so-called in the following the generalized CMLE (GCMLE) of $\boldsymbol{\theta}$, is the solution of the maximization of a non-linear multidimensional optimization problem involving the computation of $\mathbf{C}_{\bar{\mathbf{n}}_k}^{-1}(\boldsymbol{\theta})$ and $|\mathbf{C}_{\bar{\mathbf{n}}_k(\boldsymbol{\theta})}|$, where $\mathbf{C}_{\bar{\mathbf{n}}_k(\boldsymbol{\theta})}$ is not block diagonal (except if $\mathbf{C}_{\mathbf{w}_l} = \mathbf{0}$, $1 \leq l \leq k-1$). Therefore, at first sight, the computation of the GCMLE of $\boldsymbol{\theta}$ seems to become computationally prohibitive as the number of observations k increases, which would limit the interest of the proposed model. Fortunately, the observation model of interest (2a-2b) belongs to the general class of linear discrete state-space (LDSS) models [10][11] represented with the state (2a) and measurement (2b) equations. By exploiting some new results on linear minimum variance distortionless response filters (LMVDRFs) for LDSS models [12], we show that the GCMLEs of \mathbf{x}_1 and $\boldsymbol{\theta}$ can be recursively computed from observation to observation without the need to compute at each new observation $\mathbf{C}_{\bar{\mathbf{n}}_k}^{-1}(\boldsymbol{\theta})$ nor $|\mathbf{C}_{\bar{\mathbf{n}}_k(\boldsymbol{\theta})}|$. The recursive form of the GCMLE allows to explore, by Monte-Carlo simulations, its asymptotic performance in terms of mean-squared error (MSE). For instance, the example given in Section 4 exemplifies the non negligible impact of an amplitude fluctuation which introduces a lower limit in the achievable MSE of GCMLEs. From a practical point of view, the existence of this lower limit shows that, when signal sources fluctuate, there exists an optimal number of observations that can be combined coherently in order to estimate their amplitudes and other unknown associated parameters with the minimum (or almost minimum) achievable MSE. From a theoretical point of view, we exhibit non consistent MLEs when the number of observations tends to infinity, which highlights the consequence of combining (even slightly) dependent observations. Last but not least, the recursive form of GCMLEs is also a key feature for real-world applications [10][11] where the observations become available sequentially and, immediately upon receipt of new observations, it is desirable to determine new estimates based upon all previous observations (including the current ones).

2. BATCH FORM OF GCMLES

For the sake of simplicity, we assume that σ_v^2 , $\{\mathbf{F}_l\}_{l=1}^{k-1}$, $\{\mathbf{C}_{\mathbf{w}_l}\}_{l=1}^{k-1}$ are known. Thus the set of unknown parameters is restricted to \mathbf{x}_1 and $\boldsymbol{\theta}$. Since $\bar{\mathbf{y}}_k \sim \mathcal{CN}(\bar{\mathbf{A}}_k(\boldsymbol{\theta})\mathbf{x}_1, \mathbf{C}_{\bar{\mathbf{n}}_k}(\boldsymbol{\theta}))$ (3a-3b), the log likelihood function, is up to a constant value, defined as [5][6][7]:

$$L(\bar{\mathbf{y}}_k; \boldsymbol{\theta}, \mathbf{x}_1) = -\ln |\mathbf{C}_{\bar{\mathbf{n}}_k(\boldsymbol{\theta})}| - (\bar{\mathbf{y}}_k - \bar{\mathbf{A}}_k(\boldsymbol{\theta})\mathbf{x}_1)^H \mathbf{C}_{\bar{\mathbf{n}}_k(\boldsymbol{\theta})}^{-1} (\bar{\mathbf{y}}_k - \bar{\mathbf{A}}_k(\boldsymbol{\theta})\mathbf{x}_1), \quad (4)$$

leading to the following definition of the GCMLEs of \mathbf{x}_1 and $\boldsymbol{\theta}$:

$$(\hat{\mathbf{x}}_{1|k}, \hat{\boldsymbol{\theta}}_k) = \arg \max_{\mathbf{x}_1, \boldsymbol{\theta}} \{L(\bar{\mathbf{y}}_k; \boldsymbol{\theta}, \mathbf{x}_1)\}. \quad (5)$$

It is then well known [5][6][7] that $\hat{\mathbf{x}}_{1|k} = \mathbf{x}_{1|k}(\hat{\boldsymbol{\theta}}_k)$ where:

$$\mathbf{x}_{1|k}(\boldsymbol{\theta}) = \left(\bar{\mathbf{A}}_k^H(\boldsymbol{\theta}) \mathbf{C}_{\bar{\mathbf{n}}_k(\boldsymbol{\theta})}^{-1} \bar{\mathbf{A}}_k(\boldsymbol{\theta}) \right)^{-1} \bar{\mathbf{A}}_k^H(\boldsymbol{\theta}) \mathbf{C}_{\bar{\mathbf{n}}_k(\boldsymbol{\theta})}^{-1} \bar{\mathbf{y}}_k, \quad (6)$$

$$\hat{\boldsymbol{\theta}}_k = \arg \max_{\boldsymbol{\theta}} \{L(\bar{\mathbf{y}}_k; \boldsymbol{\theta}, \mathbf{x}_{1|k}(\boldsymbol{\theta}))\}, \quad (7)$$

or equivalently:

$$\hat{\boldsymbol{\theta}}_k = \arg \max_{\boldsymbol{\theta}} \{\mathcal{I}_k(\boldsymbol{\theta}) - \mathcal{J}_k(\boldsymbol{\theta})\}, \quad (8a)$$

$$\mathcal{I}_k(\boldsymbol{\theta}) = \left\| \frac{\mathbf{C}_{\bar{\mathbf{n}}_k(\boldsymbol{\theta})}^{-1} \bar{\mathbf{y}}_k}{\Pi_{\bar{\mathbf{A}}_k(\boldsymbol{\theta})} \mathbf{C}_{\bar{\mathbf{n}}_k(\boldsymbol{\theta})}^{-1}} \right\|^2, \quad \mathcal{J}_k(\boldsymbol{\theta}) = \ln |\mathbf{C}_{\bar{\mathbf{n}}_k(\boldsymbol{\theta})}|, \quad (8b)$$

where $\Pi_{\bar{\mathbf{A}}}^{\mathbf{C}} = \mathbf{A}(\mathbf{A}^H \mathbf{C} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{C}$ and $\|\mathbf{u}\|_{\mathbf{C}}$ denote, respectively, the orthogonal projection matrix on $\text{span}\{\mathbf{A}\}$ and the norm of vector \mathbf{u} for the Hermitian inner product defined by the Hermitian positive-definite matrix \mathbf{C} . According to (8a-b), the GCMLE of $\boldsymbol{\theta}$ is the solution of the maximization of a non-linear multidimensional optimization problem involving the computation of $\mathbf{C}_{\bar{\mathbf{n}}_k}^{-1}(\boldsymbol{\theta})$ and $|\mathbf{C}_{\bar{\mathbf{n}}_k(\boldsymbol{\theta})}|$, where $\mathbf{C}_{\bar{\mathbf{n}}_k(\boldsymbol{\theta})}$ is not block diagonal (except if $\mathbf{C}_{\mathbf{w}_l} = \mathbf{0}$, $1 \leq l \leq k-1$). As a consequence, if one resorts to a grid search approach to solve the maximization problem, for each selected value $\boldsymbol{\theta}^i$ of the grid, the evaluation of $\mathcal{I}_k(\boldsymbol{\theta}^i)$ and $\mathcal{J}_k(\boldsymbol{\theta}^i)$ request $O(\mathcal{N}_k^3)$ multiplications, where $\mathcal{N}_k = \sum_{l=1}^k N_l$, which becomes rapidly computationally prohibitive as the number of observations k increases.

3. RECURSIVE FORM OF GCMLES

In this section, we consider the computation of $\mathbf{x}_{1|k}(\boldsymbol{\theta})$ (6) and $\{\mathcal{I}_k(\boldsymbol{\theta}), \mathcal{J}_k(\boldsymbol{\theta})\}$ (8b) for a selected value $\boldsymbol{\theta}$ of the parameter space. We show that $\{\mathbf{x}_{1|k}(\boldsymbol{\theta}), \mathcal{I}_k(\boldsymbol{\theta})\}$ and $\mathcal{J}_k(\boldsymbol{\theta})$ can be computed recursively by means of two distinct recursions; a first one associated with a LMVDRF and a second one associated with a KF. For legibility, we omit the dependency of \mathbf{H}_k on $\boldsymbol{\theta}$ and $\mathbf{H}_k(\boldsymbol{\theta})$ is simply denoted \mathbf{H}_k ; the same applies to $\mathbf{C}_{\bar{\mathbf{n}}_k}(\boldsymbol{\theta})$, $\mathbf{x}_{1|k}(\boldsymbol{\theta})$, $\mathcal{I}_k(\boldsymbol{\theta})$ and $\mathcal{J}_k(\boldsymbol{\theta})$ simply denoted $\mathbf{C}_{\bar{\mathbf{n}}_k}$, $\mathbf{x}_{1|k}$, \mathcal{I}_k and \mathcal{J}_k .

3.1. Background on LMVDRFs

In Bayesian estimation, if the estimate of \mathbf{x}_k is based on measurements up to and including time l , it is denoted as $\hat{\mathbf{x}}_{k|l} \triangleq \hat{\mathbf{x}}_{k|l}(\mathbf{y}_1, \dots, \mathbf{y}_l)$. A filter estimates \mathbf{x}_k based on measurements up to and including time k . Let $\bar{\mathbf{W}}_k = \begin{bmatrix} \bar{\mathbf{D}}_{k-1} \\ \bar{\mathbf{w}}_k \end{bmatrix}$ where $\bar{\mathbf{D}}_{k-1} \in \mathcal{M}_{\mathbb{C}}(\mathcal{N}_{k-1}, P)$ and $\bar{\mathbf{w}}_k \in \mathcal{M}_{\mathbb{C}}(N_k, P)$. From (3a-3b), $\bar{\mathbf{W}}_k$ defines the following filtering:

$$\bar{\mathbf{W}}_k^H \bar{\mathbf{y}}_k = \left(\left(\bar{\mathbf{W}}_k^H \bar{\mathbf{A}}_k \right) \mathbf{x}_1 + \mathbf{G}_k \bar{\mathbf{w}}_{k-1} \right) + \bar{\mathbf{W}}_k^H \bar{\mathbf{n}}_k - \mathbf{G}_k \bar{\mathbf{w}}_{k-1}.$$

Therefore, a filter $\bar{\mathbf{W}}_k$ is distortionless iff:

$$\bar{\mathbf{W}}_k^H \bar{\mathbf{y}}_k = \mathbf{x}_k + \bar{\mathbf{W}}_k^H \bar{\mathbf{n}}_k - \mathbf{G}_k \bar{\mathbf{w}}_{k-1} \Leftrightarrow \bar{\mathbf{W}}_k^H \bar{\mathbf{A}}_k = \mathbf{B}_{k,1}. \quad (10)$$

If \mathbf{H}_1 and $\mathbf{C}_{\mathbf{v}_1}$ are full rank, there exists a best distortionless filter in the MSE sense, aka the LMVDRF, defined by² [12]:

$$\bar{\mathbf{W}}_k^b = \arg \min_{\bar{\mathbf{W}}_k} \{ \mathbf{P}_{k|k}(\bar{\mathbf{W}}_k) \} \text{ s.t. } \bar{\mathbf{W}}_k^H \bar{\mathbf{A}}_k = \mathbf{B}_{k,1}, \quad (11a)$$

where $\mathbf{P}_{k|k}(\bar{\mathbf{W}}_k) = E \left[\left(\bar{\mathbf{W}}_k^H \bar{\mathbf{y}}_k - \mathbf{x}_k \right) \left(\bar{\mathbf{W}}_k^H \bar{\mathbf{y}}_k - \mathbf{x}_k \right)^H \right]$, which is equivalent to [12]:

$$\bar{\mathbf{W}}_k^b = \arg \min_{\bar{\mathbf{W}}_k} \left\{ E \left[\hat{\mathbf{r}}_k(\hat{\mathbf{r}}_k)^H \right] \right\} \text{ s.t. } \bar{\mathbf{W}}_k^H \bar{\mathbf{A}}_k = \mathbf{B}_{k,1}, \quad \hat{\mathbf{r}}_k = \bar{\mathbf{W}}_k^H \bar{\mathbf{n}}_k - \mathbf{G}_k \bar{\mathbf{w}}_{k-1}. \quad (11b)$$

Then, provided that:

$$\mathbf{C}_{\mathbf{w}_{l-1}, \bar{\mathbf{y}}_{l-1}} = \mathbf{0}, \quad \mathbf{C}_{\mathbf{v}_l, \bar{\mathbf{y}}_{l-1}} = \mathbf{0}, \quad 2 \leq l \leq k, \quad (12)$$

and $\mathbf{C}_{\bar{\mathbf{n}}_l}$, $2 \leq l \leq k$, are invertible, the solution of (11a-11b) shares the same recursion as the Kalman Filter (KF)[12][13], except at time $k=1$ where $\mathbf{W}_1^b = \mathbf{C}_{\mathbf{v}_1}^{-1} \mathbf{H}_1 \mathbf{P}_{1|1}^b$, $\mathbf{P}_{1|1}^b = (\mathbf{H}_1^H \mathbf{C}_{\mathbf{v}_1}^{-1} \mathbf{H}_1)^{-1}$.

²The superscript b is used to remind the reader that the value under consideration is the "best" one according to a given criterion.

3.2. Recursive form of $\mathbf{x}_{1|k}$ and \mathcal{I}_k

By noticing that:

$$\Pi_{\bar{\mathbf{A}}_k}^{-1} \bar{\mathbf{y}}_k = \bar{\mathbf{A}}_k \left(\bar{\mathbf{A}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{A}}_k \right)^{-1} \bar{\mathbf{A}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{y}}_k = \bar{\mathbf{A}}_k \mathbf{x}_{1|k},$$

\mathcal{I}_k can be rewritten as:

$$\mathcal{I}_k = \mathbf{x}_{1|k}^H \mathbf{P}_{1|k}^{-1} \mathbf{x}_{1|k}, \quad \mathbf{P}_{1|k} = \left(\bar{\mathbf{A}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{A}}_k \right)^{-1}. \quad (13)$$

In order to exhibit a recursive formulation of $\mathbf{x}_{1|k}$ (6) and \mathcal{I}_k (8b), firstly, one builds from (2a-b) an auxiliary LDSS model consisting of the same observations associated with an augmented state for $k \geq 2$:

$$\begin{aligned} l=1: & \quad \mathbf{y}_1 = \mathbf{H}_1 \mathbf{x}_1 + \mathbf{v}_1, \\ l=2: & \quad \begin{cases} \begin{pmatrix} \mathbf{x}_2 \\ \varkappa_2 \end{pmatrix} = \begin{bmatrix} \mathbf{F}_1 \\ \mathbf{I} \end{bmatrix} \mathbf{x}_1 + \begin{pmatrix} \mathbf{w}_1 \\ \mathbf{0} \end{pmatrix} \\ \mathbf{y}_2 = [\mathbf{H}_2 \quad \mathbf{0}] \begin{pmatrix} \mathbf{x}_2 \\ \varkappa_2 \end{pmatrix} + \mathbf{v}_2 \end{cases} \\ l \geq 3: & \quad \begin{cases} \begin{pmatrix} \mathbf{x}_l \\ \varkappa_l \end{pmatrix} = \begin{bmatrix} \mathbf{F}_{l-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{pmatrix} \mathbf{x}_{l-1} \\ \varkappa_{l-1} \end{pmatrix} + \begin{pmatrix} \mathbf{w}_{l-1} \\ \mathbf{0} \end{pmatrix} \\ \mathbf{y}_l = [\mathbf{H}_l \quad \mathbf{0}] \begin{pmatrix} \mathbf{x}_l \\ \varkappa_l \end{pmatrix} + \mathbf{v}_l \end{cases} \end{aligned}$$

that is, in short:

$$\begin{cases} \mathbf{x}'_1 = \mathbf{x}_1 \\ \mathbf{y}_1 = \mathbf{H}'_1 \mathbf{x}'_1 + \mathbf{v}_1 \end{cases}, \quad \begin{cases} \mathbf{x}'_l = \mathbf{F}'_{l-1} \mathbf{x}'_{l-1} + \mathbf{w}'_{l-1} \\ \mathbf{y}_l = \mathbf{H}'_l \mathbf{x}'_l + \mathbf{v}_l \end{cases}, \quad (14)$$

where $\mathbf{H}'_1 = \mathbf{H}_1$, and (3a) becomes:

$$\mathbf{y}_l = \mathbf{A}'_l \mathbf{x}_l + \mathbf{n}'_l, \quad \mathbf{A}'_l = \mathbf{H}'_l \mathbf{B}'_{l,1}, \quad \begin{cases} \mathbf{n}'_1 = \mathbf{v}_1 \\ \mathbf{n}'_l = \mathbf{v}_l + \mathbf{H}'_l \mathbf{G}'_l \bar{\mathbf{w}}'_{l-1} \end{cases}.$$

By definition:

$$\begin{aligned} \mathbf{B}'_{l,1} &= \mathbf{F}'_{l-1} \dots \mathbf{F}'_2 \mathbf{F}'_1 = \begin{bmatrix} \mathbf{B}_{l,2} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{F}_1 \\ \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{B}_{l,1} \\ \mathbf{I} \end{bmatrix}, \\ \mathbf{A}'_l &= \mathbf{H}'_l \mathbf{B}'_{l,1} = [\mathbf{H}_l \quad \mathbf{0}] \begin{bmatrix} \mathbf{B}_{l,1} \\ \mathbf{I} \end{bmatrix} = \mathbf{H}_l \mathbf{B}_{l,1} = \mathbf{A}_l. \end{aligned}$$

Moreover since $\mathbf{G}'_l \bar{\mathbf{w}}'_{l-1} = \sum_{q=1}^{l-1} \mathbf{B}'_{l,q+1} \mathbf{w}'_q = (\mathbf{G}'_l \bar{\mathbf{w}}'_{l-1})$, then:

$$\mathbf{n}'_l = \mathbf{v}_l + \mathbf{H}'_l \mathbf{G}'_l \bar{\mathbf{w}}'_{l-1} = \mathbf{v}_l + \mathbf{H}_l \mathbf{G}_l \bar{\mathbf{w}}_{l-1} = \mathbf{n}_l.$$

Secondly, since \mathbf{H}'_1 and $\mathbf{C}_{\mathbf{v}_1}$ are full rank, if we consider the LDSS model (14), the LMVDRF of \mathbf{x}'_k exists and is defined by (11a):

$$\bar{\mathbf{W}}_k^b = \arg \min_{\bar{\mathbf{W}}_k} \{ \mathbf{P}_{k|k}(\bar{\mathbf{W}}_k) \} \text{ s.t. } \bar{\mathbf{W}}_k^H \bar{\mathbf{A}}_k = \mathbf{B}'_{k,1}, \quad (15a)$$

where $\mathbf{P}_{k|k}(\bar{\mathbf{W}}_k) = E \left[\left(\bar{\mathbf{W}}_k^H \bar{\mathbf{y}}_k - \mathbf{x}'_k \right) \left(\bar{\mathbf{W}}_k^H \bar{\mathbf{y}}_k - \mathbf{x}'_k \right)^H \right]$, which is equivalent to (11b):

$$\begin{aligned} \bar{\mathbf{W}}_k^b &= \arg \min_{\bar{\mathbf{W}}_k} \left\{ E \left[\hat{\mathbf{r}}_k^H \hat{\mathbf{r}}_k \right] \right\} \text{ s.t. } \bar{\mathbf{W}}_k^H \bar{\mathbf{A}}_k = \mathbf{B}'_{k,1}, \\ \hat{\mathbf{r}}_k &= \bar{\mathbf{W}}_k^H \bar{\mathbf{n}}'_k - \mathbf{G}'_k \bar{\mathbf{w}}'_{k-1}, \quad \bar{\mathbf{W}}_k = \begin{bmatrix} \bar{\mathbf{W}}_k^x & \bar{\mathbf{W}}_k^z \end{bmatrix}. \end{aligned} \quad (15b)$$

Since $\bar{\mathbf{W}}_k^b$ is analogous to a linearly constrained Wiener filter [14, §2.5], its batch form is given by [14, §2]:

$$\begin{aligned} \mathbf{C}_{\bar{\mathbf{n}}'_k} \bar{\mathbf{W}}_k^b &= \bar{\mathbf{A}}'_k \left(\bar{\mathbf{A}}'^H_k \mathbf{C}_{\bar{\mathbf{n}}'_k}^{-1} \bar{\mathbf{A}}'_k \right)^{-1} \left(\mathbf{B}'_{k,1} \right)^H + \\ & \left(\mathbf{I} - \bar{\mathbf{A}}'_k \left(\bar{\mathbf{A}}'^H_k \mathbf{C}_{\bar{\mathbf{n}}'_k}^{-1} \bar{\mathbf{A}}'_k \right)^{-1} \bar{\mathbf{A}}'^H_k \mathbf{C}_{\bar{\mathbf{n}}'_k}^{-1} \right) \mathbf{C}_{\bar{\mathbf{n}}'_k, \mathbf{G}'_k \bar{\mathbf{w}}'_{k-1}}, \end{aligned} \quad (16a)$$

that is, since $\bar{\mathbf{n}}'_k = \bar{\mathbf{n}}_k$ and $\bar{\mathbf{A}}'_k = \bar{\mathbf{A}}_k$:

$$\begin{aligned} \mathbf{C}_{\bar{\mathbf{n}}_k} \bar{\mathbf{W}}_k^b &= \bar{\mathbf{A}}_k \left(\bar{\mathbf{A}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{A}}_k \right)^{-1} \left[\mathbf{B}_{k,1}^H \mathbf{I} \right] + \\ & \left(\mathbf{I} - \bar{\mathbf{A}}_k \left(\bar{\mathbf{A}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{A}}_k \right)^{-1} \bar{\mathbf{A}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \right) \left[\mathbf{C}_{\bar{\mathbf{n}}_k, \mathbf{G}_k \bar{\mathbf{w}}_{k-1}} \mathbf{0} \right]. \end{aligned} \quad (16b)$$

Therefore, (15a-15b) yields the following separable solutions:

$$\begin{aligned} \left(\bar{\mathbf{W}}_k^x \right)^b &= \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{A}}_k \left(\bar{\mathbf{A}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{A}}_k \right)^{-1} \mathbf{B}_{k,1}^H + \\ \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} & \left(\mathbf{I} - \bar{\mathbf{A}}_k \left(\bar{\mathbf{A}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{A}}_k \right)^{-1} \bar{\mathbf{A}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \right) \mathbf{C}_{\bar{\mathbf{n}}_k, \mathbf{G}_k \bar{\mathbf{w}}_{k-1}}, \\ \left(\bar{\mathbf{W}}_k^z \right)^b &= \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{A}}_k \left(\bar{\mathbf{A}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{A}}_k \right)^{-1}, \end{aligned} \quad (17)$$

leading to $\hat{\mathbf{x}}_{k|k}^b = \left(\bar{\mathbf{A}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{A}}_k \right)^{-1} \bar{\mathbf{A}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{y}}_k = \mathbf{x}_{1|k}$, and

$E \left[\left(\hat{\mathbf{x}}_{k|k}^b - \hat{\mathbf{x}}_{k|k} \right) \left(\hat{\mathbf{x}}_{k|k}^b - \hat{\mathbf{x}}_{k|k} \right)^H \right] = \left(\bar{\mathbf{A}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{A}}_k \right)^{-1} \triangleq \mathbf{P}_{1|k}$.

Thirdly, the conditions (12) hold since: a) the noise sequences $\{\mathbf{w}_l\}$ and $\{\mathbf{v}_l\}$ are zero-mean, white, uncorrelated with known covariances $\mathbf{C}_{\mathbf{w}_l}$ and $\mathbf{C}_{\mathbf{v}_l}$, b) $\mathbf{x}'_1 = \mathbf{x}_1$ is uncorrelated with $\{\mathbf{w}_l, \mathbf{v}_l\}$. Based on these facts, the solution of (15a-15b) can also be computed recursively, since the LMVDRF shares then the same recursion as the KF [12][13]. Finally, $\mathbf{x}_{1|k}$ (6) and \mathcal{I}_k (8b) can be computed recursively as follows:

$$\mathcal{I}_k = \mathbf{x}_{1|k}^H \mathbf{P}_{1|k}^{-1} \mathbf{x}_{1|k}, \quad \begin{cases} \mathbf{x}_{1|k} = [\mathbf{0} \quad \mathbf{I}] \hat{\mathbf{x}}_{k|k}^b \\ \mathbf{P}_{1|k} = [\mathbf{0} \quad \mathbf{I}] \mathbf{P}_{k|k}^b \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \end{cases}, \quad (18)$$

where $\hat{\mathbf{x}}_{k|k}^b$ and $\mathbf{P}_{k|k}^b$ follow the recursion [12][13]:

$$\hat{\mathbf{x}}_{k|k}^b = \left(\mathbf{I} - \mathbf{W}_k^{bH} \mathbf{H}'_k \right) \mathbf{F}'_{k-1} \hat{\mathbf{x}}_{k-1|k-1}^b + \mathbf{W}_k^{bH} \mathbf{y}_k, \quad (19a)$$

$$\begin{aligned} \mathbf{P}_{k|k-1} &= \mathbf{F}'_{k-1} \mathbf{P}_{k-1|k-1}^b \mathbf{F}'_{k-1}^H + \mathbf{C}_{\mathbf{w}'_{k-1}} \\ \mathbf{W}_k^b &= \left(\mathbf{H}'_k \mathbf{P}_{k|k-1} \mathbf{H}'_k + \mathbf{C}_{\mathbf{v}_k} \right)^{-1} \mathbf{H}'_k \mathbf{P}_{k|k-1} \end{aligned} \quad (19b)$$

$$\mathbf{P}_{k|k}^b = \left(\mathbf{I} - \mathbf{W}_k^{bH} \mathbf{H}'_k \right) \mathbf{P}_{k|k-1}, \quad (19c)$$

except at time $k=1$ where: $\mathbf{x}_{1|1} = \mathbf{P}_{1|1}^b \mathbf{H}_1^H \mathbf{C}_{\mathbf{v}_1}^{-1} \mathbf{y}_1$, $\mathbf{P}_{1|1}^b = \left(\mathbf{H}_1^H \mathbf{C}_{\mathbf{v}_1}^{-1} \mathbf{H}_1 \right)^{-1}$.

3.3. Recursive form of \mathcal{J}_k

Firstly, according to [15, 14.17]:

$$\begin{aligned} |\mathbf{C}_{\bar{\mathbf{n}}_k}| &= \left| \begin{bmatrix} \mathbf{C}_{\bar{\mathbf{n}}_{k-1}} & \mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \mathbf{n}_k} \\ \mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \mathbf{n}_k}^H & \mathbf{C}_{\mathbf{n}_k} \end{bmatrix} \right| = |\mathbf{C}_{\mathbf{n}_k | \bar{\mathbf{n}}_{k-1}}| |\mathbf{C}_{\bar{\mathbf{n}}_{k-1}}|, \\ \mathbf{C}_{\mathbf{n}_k | \bar{\mathbf{n}}_{k-1}} &= \mathbf{C}_{\mathbf{n}_k} - \mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \mathbf{n}_k}^H \mathbf{C}_{\bar{\mathbf{n}}_{k-1}}^{-1} \mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \mathbf{n}_k}. \end{aligned}$$

Secondly, according to (3b): $\mathbf{C}_{\bar{\mathbf{n}}_k} \triangleq \mathbf{C}_{\bar{\mathbf{y}}_k}$ and $\mathbf{C}_{\mathbf{n}_k|\bar{\mathbf{n}}_{k-1}} \triangleq \mathbf{C}_{\mathbf{y}_k|\bar{\mathbf{y}}_{k-1}}$. Therefore $\mathbf{C}_{\mathbf{n}_k|\bar{\mathbf{n}}_{k-1}} \triangleq \mathbf{C}_{\mathbf{y}_k|\bar{\mathbf{y}}_{k-1}}$ can be computed by the KF recursion associated to the LDSS model resulting from the addition to (2a-2b) of the following initial state equation:

$$\mathbf{x}_1 = \mathbf{F}_0 \mathbf{x}_0 + \mathbf{w}_0, \mathbf{C}_{\mathbf{x}_0} = \mathbf{0}, \mathbf{F}_0 = \mathbf{I}, \mathbf{C}_{\mathbf{w}_0} = \mathbf{0}. \quad (20)$$

Indeed then $\mathbf{C}_{\mathbf{n}_k|\bar{\mathbf{n}}_{k-1}} \triangleq \mathbf{C}_{\mathbf{y}_k|\bar{\mathbf{y}}_{k-1}} \triangleq \mathbf{S}_{k|k}$ [10][11] given by:

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k-1} \mathbf{P}_{k-1|k-1} \mathbf{F}_{k-1}^H + \mathbf{C}_{\mathbf{w}_{k-1}}, \quad (21a)$$

$$\mathbf{S}_{k|k} = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^H + \mathbf{C}_{\mathbf{v}_k}, \mathbf{K}_k^b = \mathbf{P}_{k|k-1} \mathbf{H}_k^H \mathbf{S}_{k|k}^{-1}, \quad (21b)$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k^b \mathbf{H}_k) \mathbf{P}_{k|k-1}. \quad (21c)$$

Finally \mathcal{J}_k can be computed recursively as: $\mathcal{J}_k = \ln |\mathbf{S}_{k|k}| + \mathcal{J}_{k-1}$.

4. MEASUREMENT OF THE BACKSCATTERING COEFFICIENT OF A TARGET

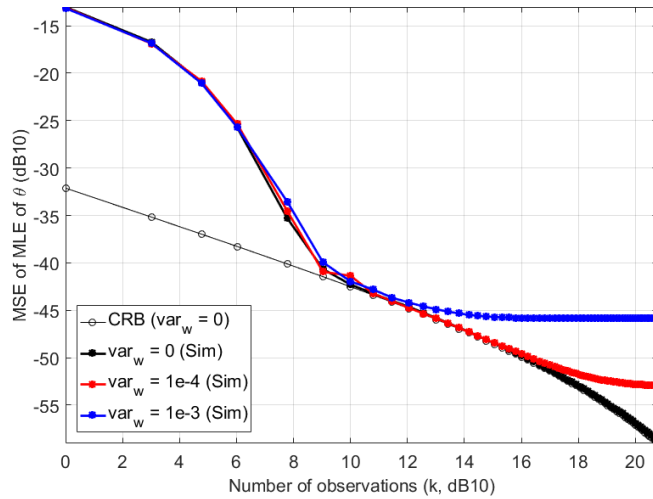


Fig. 1. MSE of the GCMLE of θ (22) versus k

Let us consider a radar system consisting of a 1-element antenna array receiving scaled, time-delayed, and Doppler-shifted echoes of a known complex bandpass signal $e_T(t) e^{j2\pi f_c t}$, where f_c is the carrier frequency and $e_T(t)$ is the envelope of the emitted signal. The antenna receives a pulse train (burst) of L pulses with a pulse repetition interval T , backscattered by a "slow" moving target [16] (no range migration during the burst and the Doppler effect on $e_T(t)$ is negligible). The target is assumed to have a radial motion towards the radar with an imposed constant radial speed v ($r(t) = r_0 + vt$) and a constant aspect angle, which leads to a constant complex backscattering coefficient ρ along the trajectory. At observation time t_l , a simplified observation model at the output of the range matched filter is given by [16]: $\mathbf{y}_l = \mathbf{h}_l(\nu) (\beta/r_l^2) \rho + \mathbf{v}_l$, where $\mathbf{h}_l^T(\theta) = (1, \dots, e^{j2\pi\theta(L-1)})$, $\theta = -2vT/\lambda_c$, $-0.5 \leq \theta \leq 0.5$, is the normalized Doppler frequency of the target, $\lambda_c = c/f_c$ is the wavelength, β represents the complex factor including transmission power, antenna gain and signal processing gains, and \mathbf{v}_l is a temporally white thermal noise with known power σ_v^2 . Indeed, in a radar system, the thermal noise power is accurately estimated from

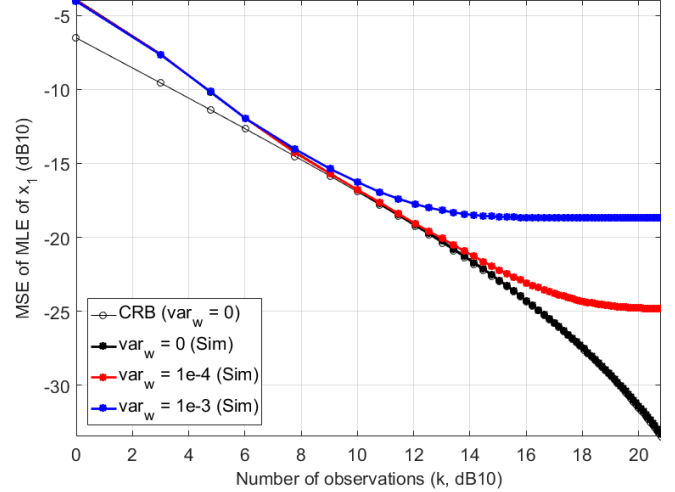


Fig. 2. MSE of the GCMLE of x_1 (22) versus k

snapshots obtained while the transmitter is turned off. In order to increase the precision of the measurement of ρ , k observations are made along the trajectory. For the sake of illustration, the time t_l , $1 \leq l \leq k$, are set such that $r_l^2 = r_0^2/f^l$. However, in a real-life experiment some experimental factors generally prevent from having a constant backscattering coefficient. For instance, it may be difficult for a target to keep a strictly constant radial trajectory, or fluctuation of the propagation medium are sometime unavoidable during the whole observation time interval. All these factors can be taken into account globally by introducing a random fluctuation from observation to observation:

$$\mathbf{y}_l = \mathbf{h}_l(\theta) \beta x_{l-1} + \mathbf{v}_l, x_l = f x_{l-1} + w_{l-1}, x_1 = \frac{\rho}{r_1^2}. \quad (22)$$

Secondly, due to adverse wind conditions, the true velocity of the target may differ from the desired one, and therefore the normalized Doppler frequency θ must be estimated as well. In this setting, the joint estimation of (x_1, θ) in the ML sense leads to GCMLEs $\hat{x}_{1|k}$ and $\hat{\theta}_k$ (6-8a), which MSEs are displayed respectively on figures (1) and (2), where $L = 10$, $\theta = 0.1$, $x_1 = (1 + j) / (2\sqrt{2})$, and $f = 1.01$, which means that the range of the target changes significantly as the number of observations increases ($1 \leq k \leq 120$). GCMLEs $\hat{x}_{1|k}$ and $\hat{\theta}_k$ are obtained via the recursive form of $x_{1|k}(\theta)$, $\mathcal{I}_k(\theta)$ and $\mathcal{J}_k(\theta)$ computed over a discretization of $]-0.5, 0.5[$ with a step of $1/2048$. The empirical MSEs are assessed with 210^4 Monte-Carlo trials. In order to highlight the impact of a target fluctuation on GCMLEs, we consider two cases with small fluctuations ($\sigma_{w_l}^2 = \sigma_w^2 \in \{10^{-4}, 10^{-3}\}$) and, for comparison, we also provide the ideal case with no fluctuation ($\sigma_{w_l}^2 = 0$) and the associated well known conditional Cramér-Rao bound (CRB) for θ and x_1 . Figure (1) and (2) exemplifies the non negligible impact of a target fluctuation on the MLEs asymptotic performance which introduces a lower limit in the achievable MSE. Practically speaking, this lower limit shows that, when a target fluctuates, there exists an optimal number of observations that can be combined coherently in order to estimate its parameters with a nearly minimum achievable MSE. Theoretically speaking, we exhibit non consistent MLEs when the number of observations tends to infinity, which highlights the consequence of combining (even slightly) dependent observations.

5. REFERENCES

- [1] H.L. Van Trees, *Detection, Estimation and Modulation Theory, Part I*, New York, Wiley, 1968
- [2] B. Picinbono, "On circularity," IEEE Trans. on SP, 42(12): 3473-3482, 1994
- [3] F.C.Schweppe, "Sensor array data processing formultiple signal sources", IEEE Trans. on IT, 14: 294-305, 1968
- [4] P. Stoica and A. Nehorai, "Performances study of conditional and unconditional direction of arrival estimation," IEEE Trans. on SP, 38(10): 1783-1795, 1990.
- [5] L. L. Scharf, *Statistical signal processing : detection, estimation, and time series analysis*, Addison-Wesley, 2002
- [6] H.L. Van Trees, *Optimum Array Processing*, Wiley-Interscience, 2002
- [7] P.J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data*, Cambridge University Press 2010
- [8] F. Vincent, O. Besson, E. Chaumette, "Approximate maximum likelihood estimation of two closely spaced sources", Signal Processing (Elsevier), 97:83-90, 2014
- [9] F. Vincent, O. Besson, E. Chaumette, "Approximate Unconditional Maximum Likelihood Direction of Arrival Estimation for Two Closely Spaced Targets", IEEE SP Letters, 22(1): 86-89, 2015
- [10] D. Simon, *Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches*, Wiley InterScience, 2006
- [11] J. L. Crassidis and J. L. Junkins, *Optimal Estimation of Dynamic Systems (2Ed)*, CRC Press, Taylor & Francis Group, 2012
- [12] E. Chaumette, F. Vincent, B. Priot, A. Dion and G. Pages, "New Results on LMVDR Estimators for LDSS Models", submitted to IEEE ICASSP Conference, 2018
- [13] E. Chaumette, B. Priot, F. Vincent, G. Pages and A. Dion, "Minimum Variance Distortionless Response Estimators For Linear Discrete State-Space Models", IEEE Trans. on AC, 62(4): 2048-2055, 2017
- [14] Paulo S.R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation (3Ed)*, Springer Science+Business Media, 2008
- [15] G.A.F. Seber, *Matrix Handbook forr Statisticians*, Wiley Series in Probability and Statistics, 2008
- [16] N Levanon, E. Mozeson, *Radar Signals*, Wiley-Interscience 2004