



# An example of non-standard behavior of a maximum likelihood estimator in the large sample regime

Lucien Bacharach, Jérôme Galy, Eric Chaumette, François Vincent, Alexandre Renaux, Mohammed Nabil El Korso

## ► To cite this version:

Lucien Bacharach, Jérôme Galy, Eric Chaumette, François Vincent, Alexandre Renaux, et al.. An example of non-standard behavior of a maximum likelihood estimator in the large sample regime. 8th IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2019), Dec 2019, Le Gosier, Guadeloupe, France. pp.525-529, 10.1109/CAMSAP45676.2019.9022505 . hal-02333941

**HAL Id: hal-02333941**

**<https://hal.parisnanterre.fr/hal-02333941>**

Submitted on 25 Oct 2019

**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.

# AN EXAMPLE OF NON-STANDARD BEHAVIOR OF A MAXIMUM LIKELIHOOD ESTIMATOR IN THE LARGE SAMPLE REGIME

Lucien Bacharach<sup>\*,◊</sup>, Jérôme Galy<sup>‡</sup>, Éric Chaumette<sup>\*</sup>, François Vincent<sup>\*</sup>, Alexandre Renaux<sup>†</sup>, Mohammed Nabil El Korso<sup>\*</sup>

<sup>\*</sup>University of Toulouse/ISAE-Supaéro, Toulouse, France (eric.chaumette@isae-supaero.fr, francois.vincent@isae-supaero.fr)

<sup>◊</sup>Université Paris-Sud/SATIE, Cachan, France (lucien.bacharach@u-psud.fr)

<sup>‡</sup>University of Montpellier/LIRMM, Montpellier, France (jerome.galy@lirmm.fr)

<sup>†</sup>Université Paris Nanterre/LEME, Ville d'Avray, France (m.elkorso@parisnanterre.fr)

<sup>\*</sup>Université Paris-Sud/L2S, Gif-sur-Yvette, France (renaux@lss.supelec.fr)

## ABSTRACT

In this paper, the performance of a maximum likelihood estimator (MLE) for a signal model accounting for possible coherence of the signal sources is studied. It is done by combining a dynamical evolution model of the source amplitude, namely a Gaussian random walk, with the usual observation model, assumed to be Gaussian. This implies an underlying relation between the signal source parameters and both the mean and the covariance matrix of the observations, and hence leads to a more general model than those usually used in array processing, such as the conditional signal model or the unconditional one. The resulting so-called “generalized conditional MLE” can be expressed both in batch form and in recursive form, which is updated since a term was missing in a previous paper. In addition, we derive the Cramér-Rao bounds associated with a model of this type, accounting for the fluctuations of a radar target’s backscattering coefficient. Simulation results highlight a non-standard behavior of the estimators: consistent but not efficient for frequencies, efficient but non-consistent for amplitudes.

**Index Terms**— parameter estimation, conditional maximum likelihood estimators, consistency, efficiency, Cramér-Rao bounds

## 1. INTRODUCTION

Since its introduction by R. A. Fisher in deterministic parameter estimation [1, 2], the method of maximum likelihood (ML) estimation has become one of the most widespread methods of estimation. The ongoing success of ML estimators (MLEs) originates from the fact that, under reasonably general conditions on the probabilistic observation model [2, pp.500-504], the MLEs are, as the number of independent observations  $N$  approaches infinity (a.k.a. the large sample regime), consistent, efficient and Gaussian distributed, that is MLEs converge in probability to the correct values as  $N \rightarrow \infty$  and MLEs are asymptotically unbiased and their covariance matrix asymptotically converges to the Cramér-Rao bound [2]. Last, still in the large sample regime, Fisher [1] showed that if the MLEs are consistent then they are also asymptotically efficient. Thus an attractive result would be the generalization of the aforementioned properties to the situation where the observations are slightly dependent, where the affordable degree of dependency would be still to be assessed. Unfortunately, if one considers the generalized conditional MLE

(GCMLE) lately introduced in [3], a further study introduced in the present communication shows that the known asymptotic properties of the MLEs in the large sample regime does not hold if the observations are not strictly independent, even in the case of a Gaussian observation model. Indeed, let us consider the following complex Gaussian observation model in the single signal source case:

$$x_l = f_{l-1} x_{l-1} + w_{l-1}, \quad 2 \leq l \leq k, \quad (1a)$$

$$y_l = \mathbf{h}_l(\theta) x_l + \mathbf{v}_l, \quad 1 \leq l \leq k, \quad (1b)$$

where the Gaussian measurement noise sequence  $\{\mathbf{v}_l\}_{l=1}^k$  is temporally and spatially white ( $\mathbf{v}_l \sim \mathcal{CN}(\mathbf{0}, \sigma_v^2 \mathbf{I})$ ), and the fluctuation noise sequence  $\{w_l\}_{l=1}^{k-1}$  is Gaussian, temporally white and uncorrelated with the noise sequence  $\{\mathbf{v}_l\}_{l=1}^k$ . Then the MLEs of  $x_1$  and  $\theta$  based on observation vectors  $\mathbf{y}_l$  on a horizon of  $k$  observations yields the so-called GCMLEs  $\hat{x}_{1|k}$  and  $\hat{\theta}_k$ . Indeed, (1a–1b) is a generalization of the usual conditional signal model (CSM) obtained where  $\{w_l = 0\}_{l=1}^{k-1}$  [4] and yielding the usual conditional MLEs (CMLEs). Due to the Gaussian random walk between observations of the signal source amplitudes, the parameter vector  $\theta$  is connected with both the expectation value and the covariance matrix [3], which is a significant change in comparison with the CSM where  $\theta$  is connected with the expectation only [4]. Interestingly enough, the amplitudes Gaussian random walk (1a) introduces as well a correlation from one observation to another (since  $\mathbf{C}_{y_l, y_{l-1}} = f_{l-1} \sigma_{x_{l-1}}^2 \mathbf{h}_l(\theta) \mathbf{h}_{l-1}^H(\theta)$ ), and thus, a degree of statistical dependency from one observation to another. As a contribution to the characterization of the MLEs in the large sample regime, we show that the combination of (even slightly) dependent observations, may lead to exhibit, on the one hand, a MLE which is asymptotically not consistent but efficient ( $\hat{x}_{1|k}$ ), and on the other hand a MLE ( $\hat{\theta}_k$ ) which is asymptotically consistent but not efficient, both for the same observation model. To do this, we compute the Cramér-Rao bounds for the parameter vector  $\boldsymbol{\xi} = (x_1, \theta)^T$  and we also correct the GCMLE for the parameter of interest  $\theta$  previously released [3, (7a-7b)], which appears to be incorrect since a term is missing in  $\mathcal{J}_k(\theta)$  (7c), term for which we propose a recursive expression as well. The correctness of the noticed behavior is also checked by highlighting the perfect match between the recursive and the batch forms of the estimator.

The rest of this paper is organized as follows. In Section 2, we provide some results about GCMLEs, including their batch as well as their recursive formulation. A correction is brought to the recursive form presented in [3]. In Section 3, we focus on a particular

This work was supported in part by DGA/AID agency (2018.60.0072.00.470.75.01), and in part by ANR-ASTRID MARGARITA (ANR-17-ASTR-0015).

model covered by (1a-1b), that may account for the fluctuations of a radar target's backscattering coefficient. The Cramér-Rao bounds for the parameter vector  $\xi$  are derived, and simulation results are provided, showing the non-standard behavior of the estimators described above.

## 2. BACKGROUND ON GCMLES FOR A SINGLE SOURCE

### 2.1. Analytical expressions (batch form)

First, we can notice that the dynamics equation (1a) can be recast, for  $2 \leq l \leq k$ , as

$$x_l = b_{l,1} x_1 + \mathbf{g}_l^T \bar{\mathbf{w}}_{l-1}, \quad (2a)$$

where, for  $i \geq 1$ , we have defined

$$b_{l,i} \triangleq \begin{cases} f_{l-1} f_{l-2} \dots f_i, & \text{if } l > i, \\ 1, & \text{if } l = i, \\ 0, & \text{if } l < i, \end{cases}$$

as well as, for  $l \geq 2$ ,  $\mathbf{g}_l \triangleq (b_{l,2}, \dots, b_{l,l})^T \in \mathbb{C}^{l-1}$ , and  $\bar{\mathbf{w}}_{l-1} \triangleq (w_1, \dots, w_{l-1})^T \in \mathbb{C}^{l-1}$ . In turn, the measurement equation (1b) can be rewritten, for  $l = 1, \dots, k$ , as

$$\mathbf{y}_l = \mathbf{a}_l(\theta) x_1 + \mathbf{n}_l(\theta), \quad (2b)$$

after defining  $\mathbf{a}_l(\theta) \triangleq \mathbf{h}_l(\theta) b_{l,1}$ , and

$$\begin{cases} \mathbf{n}_1 \triangleq \mathbf{v}_1, \\ \mathbf{n}_l(\theta) \triangleq \mathbf{v}_l + \mathbf{h}_l(\theta) (\mathbf{g}_l^T \bar{\mathbf{w}}_{l-1}), \end{cases} \text{ for } 2 \leq l \leq k.$$

By vertically concatenating the observation vectors in (2b), we obtain

$$\bar{\mathbf{y}}_k = \bar{\mathbf{a}}_k(\theta) x_1 + \bar{\mathbf{n}}_k(\theta) \quad (2c)$$

where  $\bar{\mathbf{a}}_k(\theta) \triangleq (\mathbf{a}_1^T(\theta), \dots, \mathbf{a}_k^T(\theta))^T$ , and  $\bar{\mathbf{n}}_k(\theta) \triangleq (\mathbf{n}_1^T, \mathbf{n}_2^T(\theta), \dots, \mathbf{n}_k^T(\theta))^T$ . Since we assume that  $\sigma_v^2$ ,  $\{f_l\}_{l=1}^{k-1}$  and  $\{\sigma_{w_l}^2\}_{l=1}^{k-1}$  are known, the set of unknown parameters reduces to the vector  $\xi = (\theta^T, x_1)^T$ , and we have  $\bar{\mathbf{y}}_k \sim \mathcal{CN}(\bar{\mathbf{a}}_k(\theta) x_1, \mathbf{C}_{\bar{\mathbf{n}}_k}(\theta))$ . Consequently, the log-likelihood function can be expressed, up to a constant value, as [5, 6]:

$$L(\bar{\mathbf{y}}_k; \theta, x_1) = -\ln |\mathbf{C}_{\bar{\mathbf{n}}_k}(\theta)| - (\bar{\mathbf{y}}_k - \bar{\mathbf{a}}_k(\theta) x_1)^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1}(\theta) (\bar{\mathbf{y}}_k - \bar{\mathbf{a}}_k(\theta) x_1), \quad (3)$$

and the GCMLEs of  $x_1$  and  $\theta$  based on  $\bar{\mathbf{y}}_k$  are given by:

$$(\hat{x}_{1|k}, \hat{\theta}_k) = \arg \max_{x_1, \theta} \{L(\bar{\mathbf{y}}_k; \theta, x_1)\}. \quad (4)$$

Then, it is well known [5, 6] that  $\hat{x}_{1|k} = x_{1|k}(\hat{\theta}_k)$ , where

$$x_{1|k}(\theta) = \left( \bar{\mathbf{a}}_k^H(\theta) \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1}(\theta) \bar{\mathbf{a}}_k(\theta) \right)^{-1} \bar{\mathbf{a}}_k^H(\theta) \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1}(\theta) \bar{\mathbf{y}}_k, \quad (5)$$

$$\hat{\theta}_k = \arg \max_{\theta} \{L(\bar{\mathbf{y}}_k; \theta, x_{1|k}(\theta))\}, \quad (6)$$

or equivalently, after some calculus:

$$\hat{\theta}_k = \arg \max_{\theta} \{\mathcal{I}_k(\theta) - \mathcal{J}_k(\theta)\}, \quad (7a)$$

in which we have defined

$$\mathcal{I}_k(\theta) = \|\Pi_{\bar{\mathbf{a}}_k(\theta)}^{\mathbf{C}_{\bar{\mathbf{n}}_k}^{-1}(\theta)} \bar{\mathbf{y}}_k\|_{\mathbf{C}_{\bar{\mathbf{n}}_k}^{-1}(\theta)}^2, \quad (7b)$$

$$\mathcal{J}_k(\theta) = \|\bar{\mathbf{y}}_k\|_{\mathbf{C}_{\bar{\mathbf{n}}_k}^{-1}(\theta)}^2 + \ln |\mathbf{C}_{\bar{\mathbf{n}}_k}(\theta)|, \quad (7c)$$

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 oblique 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  $\mathbf{C}$ , i.e.,  $\langle \mathbf{u}, \mathbf{v} \rangle \triangleq \mathbf{u}^H \mathbf{C} \mathbf{v}$ .

### 2.2. Recursive form of GCMLEs

It is worth noticing that the term  $\|\bar{\mathbf{y}}_k\|_{\mathbf{C}_{\bar{\mathbf{n}}_k}^{-1}(\theta)}^2$  is missing in [3], a mistake corrected in (7c). As mentioned in Section 1, even if the estimator of  $\theta$  and  $x_1$  can be computed from (5) and (7a)–(7c), this task rapidly becomes prohibitive as  $k$  increases, due to the necessity of inverting the covariance matrix  $\mathbf{C}_{\bar{\mathbf{n}}_k}(\theta)$  with size  $\mathcal{N}_k \times \mathcal{N}_k$ . This explains the necessity of seeking recursive forms of these estimators, that does not require such a computational power.

Recursive forms for  $x_{1|k}(\theta)$ ,  $\mathcal{I}_k(\theta)$  and  $\ln |\mathbf{C}_{\bar{\mathbf{n}}_k}(\theta)|$  have already been presented in [3], to which the reader is referred for the details on the way to obtain them. In the following, we focus on obtaining a recursive form for  $\|\bar{\mathbf{y}}_k\|_{\mathbf{C}_{\bar{\mathbf{n}}_k}^{-1}(\theta)}^2$ . By dropping the dependence on  $\theta$ , splitting the covariance matrix  $\mathbf{C}_{\bar{\mathbf{n}}_k}(\theta)$  into the four blocks  $\begin{bmatrix} \mathbf{C}_{\bar{\mathbf{n}}_{k-1}} & \mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \mathbf{n}_k} \\ \mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \mathbf{n}_k}^H & \mathbf{C}_{\mathbf{n}_k} \end{bmatrix}$ , and using the block inversion identity [7, (14.11), p. 293] we obtain  $\mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}^H & \mathbf{C}_{22} \end{bmatrix}$ , with

$$\begin{aligned} \mathbf{C}_{\mathbf{n}_k | \bar{\mathbf{n}}_{k-1}} &\triangleq \mathbf{C}_{\mathbf{n}_k} - \mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \mathbf{n}_k}^H \mathbf{C}_{\bar{\mathbf{n}}_{k-1}}^{-1} \mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \mathbf{n}_k}, \\ \mathbf{C}_{11} &\triangleq \mathbf{C}_{\bar{\mathbf{n}}_{k-1}}^{-1} + \mathbf{C}_{\bar{\mathbf{n}}_{k-1}}^{-1} \mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \mathbf{n}_k} \mathbf{C}_{\bar{\mathbf{n}}_{k-1}}^{-1} \mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \mathbf{n}_k}^H \mathbf{C}_{\bar{\mathbf{n}}_{k-1}}^{-1}, \\ \mathbf{C}_{12} &\triangleq -\mathbf{C}_{\bar{\mathbf{n}}_{k-1}}^{-1} \mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \mathbf{n}_k} \mathbf{C}_{\mathbf{n}_k | \bar{\mathbf{n}}_{k-1}}^{-1}, \\ \mathbf{C}_{22} &\triangleq \mathbf{C}_{\mathbf{n}_k | \bar{\mathbf{n}}_{k-1}}^{-1}. \end{aligned}$$

Hence, by splitting  $\bar{\mathbf{y}}_k$  into  $(\bar{\mathbf{y}}_{k-1}^T, \mathbf{y}_k^T)^T$  and performing the block product in  $\|\bar{\mathbf{y}}_k\|_{\mathbf{C}_{\bar{\mathbf{n}}_k}^{-1}(\theta)}^2 = \bar{\mathbf{y}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{y}}_k$ , we obtain

$$\begin{aligned} \bar{\mathbf{y}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{y}}_k &= \bar{\mathbf{y}}_{k-1}^H \mathbf{C}_{\bar{\mathbf{n}}_{k-1}}^{-1} \bar{\mathbf{y}}_{k-1} \\ &\quad + (\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1})^H \mathbf{C}_{\mathbf{n}_k | \bar{\mathbf{n}}_{k-1}}^{-1} (\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}) \end{aligned}$$

where  $\hat{\mathbf{y}}_{k|k-1} \triangleq \mathbf{C}_{\mathbf{n}_k, \bar{\mathbf{n}}_{k-1}} \mathbf{C}_{\bar{\mathbf{n}}_{k-1}}^{-1} \bar{\mathbf{y}}_{k-1}$ . By the way,  $\mathbf{C}_{x_1} = 0$  implies (2c)  $\mathbf{C}_{\bar{\mathbf{y}}_k} = \mathbf{C}_{\bar{\mathbf{n}}_k}$ ,  $\mathbf{C}_{\mathbf{y}_k | \bar{\mathbf{y}}_{k-1}} = \mathbf{C}_{\mathbf{n}_k | \bar{\mathbf{n}}_{k-1}}$ . Thus  $\hat{\mathbf{y}}_{k|k-1} = \mathbf{C}_{\mathbf{y}_k, \bar{\mathbf{y}}_{k-1}} \mathbf{C}_{\bar{\mathbf{y}}_{k-1}}^{-1} \bar{\mathbf{y}}_{k-1}$  which can be connected to the Kalman filter estimate of  $\hat{x}_k$  whose batch form is given by  $\hat{x}_{k|k}^b = \mathbf{C}_{x_k, \bar{\mathbf{y}}_k} \mathbf{C}_{\bar{\mathbf{y}}_k}^{-1} \bar{\mathbf{y}}_k$ . Therefore  $\hat{\mathbf{y}}_{k|k-1}$  can be computed recursively as follows

$$\hat{\mathbf{y}}_{k|k-1}^b = \mathbf{h}_k f_{k-1} \hat{x}_{k-1|k-1}^b, \quad \hat{x}_{k|k}^b = \hat{x}_{k|k-1}^b + \mathbf{K}_k^b (\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}^b),$$

where  $\mathbf{K}_k^b$  is obtained from the following recursion:

$$C_{x_k | \bar{\mathbf{y}}_{k-1}} = f_{k-1} C_{x_{k-1} | \bar{\mathbf{y}}_{k-1}} f_{k-1}^* + C_{\mathbf{w}_{k-1}},$$

$$\mathbf{C}_{\mathbf{y}_k | \bar{\mathbf{y}}_{k-1}} = \mathbf{H}_k \mathbf{C}_{\mathbf{x}_k | \bar{\mathbf{y}}_{k-1}} \mathbf{H}_k^H + \mathbf{C}_{\mathbf{v}_k},$$

$$\mathbf{K}_k^b = C_{x_k | \bar{\mathbf{y}}_{k-1}} \mathbf{h}_k^H \mathbf{C}_{\mathbf{y}_k | \bar{\mathbf{y}}_{k-1}}^{-1},$$

$$C_{x_k | \bar{\mathbf{y}}_k} = (1 - \mathbf{K}_k^b \mathbf{h}_k) C_{x_k | \bar{\mathbf{y}}_{k-1}},$$

$$\text{and } \hat{x}_{1|1}^b = \mathbf{C}_{x_1, \mathbf{y}_1} \mathbf{C}_{\mathbf{y}_1}^{-1} \mathbf{y}_1 = 0.$$

### 3. CASE OF A PARTIALLY COHERENT SOURCE

#### 3.1. Model and assumptions

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  $N$  pulses with a pulse repetition interval  $T$ , backscattered by a “slow” moving target [8]. 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  $\rho$  along the trajectory. In order to increase the precision of the measurement of  $\rho$ ,  $k$  observations are made along the trajectory. For the sake of illustration, the time  $t_l$ ,  $1 \leq l \leq k$ , are set such that  $r_l^2 = r_1^2 / f^{l-1}$ . 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 the radar phase coherency may not be kept with a sufficient precision during all the observations; last, fluctuation of the propagation medium are sometimes 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 leading to following simplified observation model at the output of the range matched filter [8]:

$$2 \leq l \leq k : \begin{cases} \mathbf{y}_l = (\mathbf{h}(\theta) \frac{\beta}{r_1^2} f^{l-1}) \mathbf{x}_l + \mathbf{v}_l \\ \mathbf{x}_l = \kappa \mathbf{x}_{l-1} + \mathbf{w}_{l-1} \end{cases}, \quad (9)$$

with  $\mathbf{h}_l(\theta) = \mathbf{h}(\theta) = (1, e^{j2\pi\theta}, \dots, e^{j2\pi(N-1)\theta})^T \in \mathbb{C}^N$ ,  $\beta \in \mathbb{C}$ ,  $r_1 \in \mathbb{R}_+$ ,  $f \in \mathbb{R}_+$ , and  $\kappa \in \mathbb{C}$ ,  $0 \leq |\kappa| \leq 1$ . In (9),  $\theta = -2vT/\lambda_c$ ,  $-0.5 \leq \theta \leq 0.5$ , is the normalized Doppler frequency of the target,  $\lambda_c = c/f_c$  is the wavelength,  $\beta$  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  $\sigma_v^2$ . Since the backscattered energy of the target is given by  $|\rho|^2$  on the one hand, and by  $E[|x_k|^2]$  on the other hand, the assumption of a constant backscattering coefficient implies  $E[|x_k|^2] = |x_1|^2 = |\rho|^2$ , which leads to  $\sigma_{w_{l-1}}^2 = (1 - |\kappa|^2) |x_1|^2$ . By noticing that the model (9) belongs to the class described by (1a–1b), we can write it under the form (2a–2c), with  $b_{l,i} = \kappa^{l-i}$ ,  $\mathbf{a}_l(\theta) = \frac{\beta}{r_1^2} (f\kappa)^{l-1} \mathbf{h}(\theta)$ ,

$$\mathbf{n}_1 = \mathbf{v}_1, \quad \mathbf{n}_l(\theta) = \mathbf{v}_l + \frac{\beta}{r_1^2} f^{k-1} \mathbf{h}(\theta) \sum_{i=1}^{l-1} \kappa^{l-i-1} \mathbf{w}_i \quad (10)$$

for  $2 \leq l \leq k$ ,  $1 \leq i \leq l$ . A similar form as (2c) can thus be obtained, with  $\bar{\mathbf{a}}_k(\theta) \triangleq \frac{\beta}{r_1^2} (\mathbf{f}_k \otimes \mathbf{h}(\theta)) \in \mathbb{C}^{kN}$ , in which  $\mathbf{f}_k \triangleq (1, f\kappa, (f\kappa)^2, \dots, (f\kappa)^{k-1})^T \in \mathbb{C}^k$  and “ $\otimes$ ” denotes the Krönecker product. We consequently have  $\bar{\mathbf{y}}_k \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{C})$ , with  $\boldsymbol{\mu} = \bar{\mathbf{a}}_k(\theta) \mathbf{x}_1$ , and the covariance matrix  $\mathbf{C} \triangleq \mathbf{C}_{\bar{\mathbf{y}}_k}(\theta)$  can be written as the block matrix

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{\mathbf{n}_1} & \mathbf{C}_{\mathbf{n}_1 \mathbf{n}_2} & \dots & \mathbf{C}_{\mathbf{n}_1 \mathbf{n}_k} \\ \mathbf{C}_{\mathbf{n}_2 \mathbf{n}_1} & \mathbf{C}_{\mathbf{n}_2} & & \\ \vdots & & \ddots & \\ \mathbf{C}_{\mathbf{n}_k \mathbf{n}_1} & & \dots & \mathbf{C}_{\mathbf{n}_k} \end{bmatrix} \quad (11)$$

after dropping the dependency on  $\theta$  for convenience, and defining the blocks  $\mathbf{C}_{\mathbf{n}_l \mathbf{n}_{l'}} \triangleq E[\mathbf{n}_l \mathbf{n}_{l'}^H] = \mathbf{C}_{\mathbf{n}_{l'} \mathbf{n}_l}$  as the cross-covariance matrices between  $\mathbf{n}_l$  and  $\mathbf{n}_{l'}$ . In particular, we can notice that  $\mathbf{C}_{\mathbf{n}_1} = \sigma_v^2 \mathbf{I}$

and  $\mathbf{C}_{\mathbf{n}_1 \mathbf{n}_l} = \mathbf{C}_{\mathbf{n}_l \mathbf{n}_1} = \mathbf{0}$ , from (10). The other blocks are given by

$$\mathbf{C}_{\mathbf{n}_l \mathbf{n}_{l'}} = \sigma_v^2 \delta_{l,l'} \mathbf{I}_N + \frac{|\beta|^2}{r_1^4} f^{l+l'-2} \mathbf{h}(\theta) \mathbf{h}^H(\theta) \times E \left[ \sum_{i=1}^{l-1} \kappa^{l-1-i} \mathbf{w}_i \sum_{i'=1}^{l'-1} (\kappa^*)^{l'-1-i'} \mathbf{w}_{i'}^* \right] \quad (12)$$

where  $\delta_{l,l'}$  denotes the Krönecker delta (equals 1 if  $l = l'$ , equals 0 otherwise). The expectation term in (12) can be further simplified as

$$E \left[ \left( \sum_{i=1}^{l-1} \kappa^{l-1-i} \mathbf{w}_i \right) \left( \sum_{i'=1}^{l'-1} (\kappa^*)^{l'-1-i'} \mathbf{w}_{i'}^* \right) \right] = \sigma_w^2 g_{l,l'}(\kappa)$$

after defining  $g_{l,l'}(\kappa) \triangleq \sum_{i=1}^{\min(l,l')-1} \kappa^{l-1-i} (\kappa^*)^{l'-1-i}$ , and according to the assumptions made on  $\{\mathbf{w}_l\}_{l=1}^{k-1}$ . In the case  $\kappa \in \mathbb{R}$ ,  $g_{l,l'}(\kappa)$  reduces to

$$g_{l,l'}(\kappa) = \kappa^{\max(l,l') - \min(l,l')} \sum_{i=0}^{\min(l,l')-2} \kappa^{2i} = \begin{cases} \kappa^{\max(l,l') - \min(l,l')} \frac{1 - \kappa^{2(\min(l,l')-1)}}{1 - \kappa^2} & \text{if } \kappa \neq 1 \\ \min(l,l') - 1 & \text{if } \kappa = 1 \end{cases}$$

Finally, (12) can be rewritten as

$$\mathbf{C}_{\mathbf{n}_l \mathbf{n}_{l'}} = \frac{|\beta|^2}{r_1^4} f^{l+l'-2} \mathbf{h}(\theta) \mathbf{h}^H(\theta) \sigma_w^2 g_{l,l'}(\kappa) + \sigma_v^2 \delta_{l,l'} \mathbf{I}_N. \quad (13)$$

These expressions are useful to obtain the explicit expressions of the MLEs for  $\theta$  and  $\mathbf{x}_1$ , as well as the associated CRBs.

#### 3.2. GCMLEs and associated CRBs

The batch form of the GCMLEs for  $\theta$  and  $\mathbf{x}_1$  can simply be obtained by plugging results of the previous section into (4)–(6). Even if the computation cost of the batch form is prohibitive, we have used it so as to check the correctness of the recursive forms presented in [3] and completed in Section 2.2.

Since  $\bar{\mathbf{y}}_k \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{C})$ , the CRBs on the variance of any unbiased estimator of  $\boldsymbol{\xi} = (\theta, \text{Re}(\mathbf{x}_1), \text{Im}(\mathbf{x}_1))^T$  can be obtained by inverting the Fisher information matrix (FIM)  $\mathbf{F}(\boldsymbol{\xi})$ , whose elements are given, for  $1 \leq m, n \leq 3$ , by the well-known Slepian-Bangs formulas [5]

$$F_{m,n}(\boldsymbol{\xi}) = 2\text{Re} \left( \frac{\partial \boldsymbol{\mu}^H}{\partial \xi_i} \mathbf{C}^{-1} \frac{\partial \boldsymbol{\mu}}{\partial \xi_j} \right) + \text{tr} \left( \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \xi_i} \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial \xi_j} \right). \quad (14)$$

In order to compute the elements of  $\mathbf{F}(\boldsymbol{\xi})$ , the following results are useful:

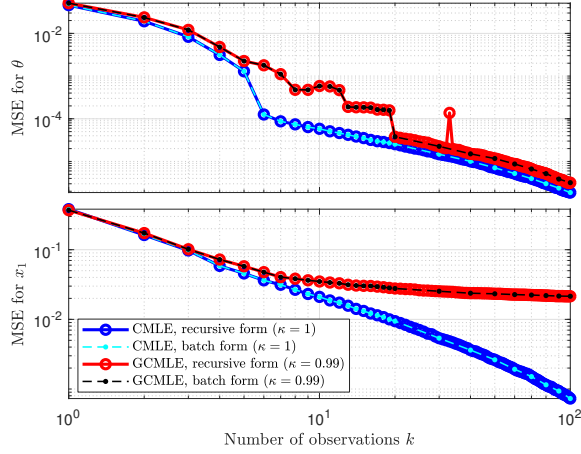
$$\frac{\partial \boldsymbol{\mu}}{\partial \theta} = \frac{\beta}{r_1^2} \mathbf{x}_1 \left( \mathbf{f}_k \otimes \frac{\partial \mathbf{h}(\theta)}{\partial \theta} \right), \quad (15a)$$

in which the elements of  $\partial \mathbf{h}(\theta) / \partial \theta$  are given, for  $1 \leq n \leq N$ , by  $[\partial \mathbf{h}(\theta) / \partial \theta]_n = j2\pi(n-1)e^{j2\pi(n-1)\theta}$ , and

$$\frac{\partial \boldsymbol{\mu}}{\partial \text{Re}(\mathbf{x}_1)} = \frac{\beta}{r_1^2} (\mathbf{f}_k \otimes \mathbf{h}(\theta)), \quad \frac{\partial \boldsymbol{\mu}}{\partial \text{Im}(\mathbf{x}_1)} = j \frac{\partial \boldsymbol{\mu}}{\partial \text{Re}(\mathbf{x}_1)}. \quad (15b)$$

The derivatives of the covariance matrix  $\mathbf{C}$  can be computed block-wise. We can first notice that, for any  $l = 1, \dots, k$  and any  $m = 1, 2, 3$ ,  $\frac{\partial \mathbf{C}_{\mathbf{n}_l \mathbf{n}_l}}{\partial \xi_m} = \frac{\partial \mathbf{C}_{\mathbf{n}_l \mathbf{n}_1}}{\partial \xi_m} = \mathbf{0}$ . Focusing on the other blocks' derivatives w.r.t.  $\theta$ , we have, for  $2 \leq l, l' \leq k$ ,

$$\frac{\partial \mathbf{C}_{\mathbf{n}_l \mathbf{n}_{l'}}}{\partial \theta} = \frac{|\beta|^2}{r_1^4} f^{l+l'-2} \frac{\partial \mathbf{h}(\theta) \mathbf{h}^H(\theta)}{\partial \theta} \sigma_w^2 g_{l,l'}(\kappa), \quad (15c)$$



**Fig. 1.** Comparison of recursive and batch forms of GCMLEs versus the number of observations  $k$ , for  $\kappa \in \{1, 0.99\}$ .

in which the elements of  $\partial \mathbf{h}(\theta) \mathbf{h}^H(\theta) / \partial \theta$  are given, for  $1 \leq n, n' \leq N$ , by  $[\frac{\partial \mathbf{h}(\theta) \mathbf{h}^H(\theta)}{\partial \theta}]_{n, n'} = j2\pi(n - n')e^{j2\pi(n - n')\theta}$ . Regarding the derivatives w.r.t.  $\text{Re}(x_1)$  and  $\text{Im}(x_1)$ , we notice that we simply have, for  $2 \leq l, l' \leq k$ ,

$$\frac{\partial \mathbf{C}_{\mathbf{n}_l \mathbf{n}_{l'}}}{\partial \text{Re}(x_1)} = \frac{\partial \mathbf{C}_{\mathbf{n}_l \mathbf{n}_{l'}}}{\partial \text{Im}(x_1)} = \frac{\partial \mathbf{C}_{\mathbf{n}_{l'} \mathbf{n}_l}}{\partial \text{Re}(x_1)} = \frac{\partial \mathbf{C}_{\mathbf{n}_{l'} \mathbf{n}_l}}{\partial \text{Im}(x_1)} = \mathbf{0}. \quad (15d)$$

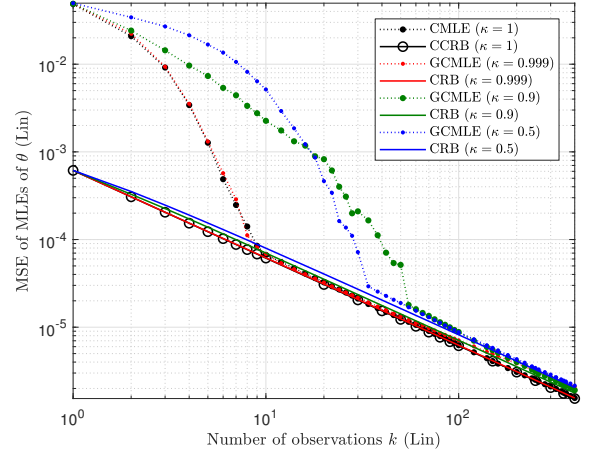
Hence, (15a)–(15d) can be plugged into (14) in order to compute the FIM.

### 3.3. Numerical results

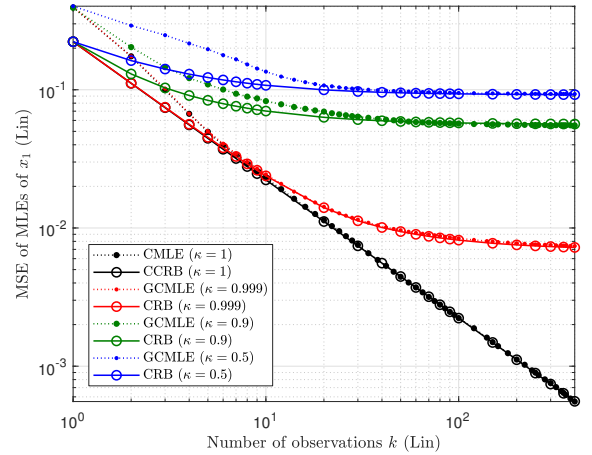
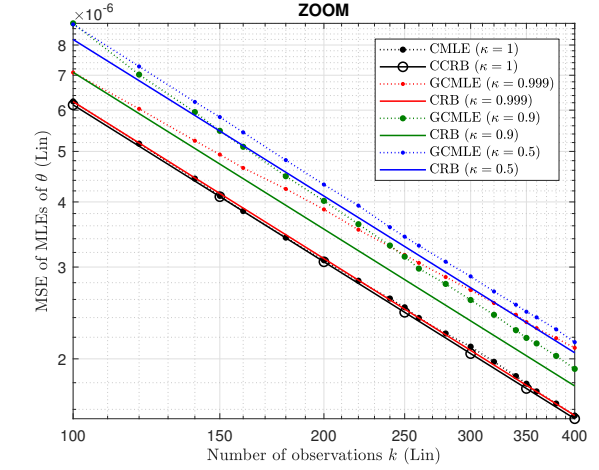
In this section, we provide results from simulation performed for model (9), for true values of the parameters  $\theta = 0.1$  and  $x_1 = (1 + j)/(2\sqrt{2})$ ,  $\beta = r_1 = 1$ ,  $\sigma_v^2 = 1$ ,  $N = 10$  and various values of  $\kappa$ . First, the correctness of the updated recursive forms of the GCMLES is checked by exhibiting the perfect match between the recursive and the batch forms in Fig. 1 for  $\kappa \in \{1, 0.99\}$ . Second, the non-standard asymptotic behavior of the GCMLEs is highlighted in Fig. 2–3. They present results regarding the estimation of  $\theta$  (including a zoom in the large-sample regime) and  $x_1$ , respectively, with values of  $\kappa \in \{1$  (in black), 0.999 (in red), 0.9 (in green), 0.5 (in magenta). The empirical mean-square errors (MSEs) have been assessed with  $10^4$  Monte-Carlo trials, and are displayed in solid line without marker, while the CRBs are displayed in solid line with circle marker. As soon as  $\kappa < 1$ , we notice in Fig. 2 that a gap appears between the MSE of  $\hat{\theta}$  and the associated CRB (non efficiency), even if it still decreases to zero (consistency). On the other hand, the MSE of  $\hat{x}_1$  in Fig. 3 attains the CRB (efficiency), but does not decrease below some lower limit (non consistency). Such a non-standard behavior has to be expected when fluctuations exist in the source amplitudes. This illustrates the fact that Cramér’s and Fisher’s results do not hold anymore when strict independence between the observations is not satisfied.

## 4. REFERENCES

- [1] R. A. Fisher, “Theory of statistical estimation,” *Proc. of the Cambridge Philos. Soc.*, vol. 22, no. 5, pp. 700–725, 1925.



**Fig. 2.** MSE of the GCMLE of  $\theta$  versus the number of observations  $k$ , for  $\kappa \in \{1, 0.999, 0.9, 0.5\}$ , and associated CRBs. Below, zoom on the large-sample regime.



**Fig. 3.** MSEs of the MLE of  $x_1$  versus the number of observations  $k$ , for  $\{1, 0.999, 0.9, 0.5\}$ , and associated CRBs.

- [2] H. Cramér, *Mathematical Methods of Statistics*, Princeton Univ. Press, Princeton, NJ, USA, 1946.
- [3] É. Chaumette, F. Vincent, A. Renaux, and J. Galy, “Generalized conditional maximum likelihood estimators in the large sample regime,” in *Proc. EUSIPCO*, Rome, Italy, 2018, pp. 271–275.
- [4] P. Stoica and A. Nehorai, “Performances study of conditional and unconditional direction of arrival estimation,” *IEEE Trans. on ASSP*, vol. 38, no. 10, pp. 1783–1795, 1990.
- [5] H. L. Van Trees, *Optimum Array Processing*, Wiley, New-York, NY, USA, 2002.
- [6] P. Schreier and L. Scharf, *Statistical Signal Processing of Complex-Valued Data*, Cambridge Univ. Press, UK, 2010.
- [7] G. A. F. Seber, *A Matrix Handbook for Statisticians*, John Wiley & Sons, Hoboken, NJ, USA, 2007.
- [8] N. Levanon and E. Mozeson, *Radar Signals*, John Wiley & Sons, Hoboken, NJ, USA, 2004.