Short communication

Recursive linearly constrained minimum variance estimator in linear models with non-stationary constraints

François Vincent, Eric Chaumette*

Department of Electronics, Optronics and Signal, ISAE-SUPAERO, University of Toulouse, 10 Avenue Edouard Belin, Toulouse 31055, France

Article history:
Received 11 August 2017
Revised 18 March 2018
Accepted 21 March 2018
Available online 22 March 2018

Keywords:
Parameters estimation
Linearly constrained minimum variance estimator
Robust adaptive beamforming

A B S T R A C T

In parameter estimation, it is common place to design a linearly constrained minimum variance estimator (LCMVE) to tackle the problem of estimating an unknown parameter vector in a linear regression model. So far, the LCMVE has been mainly studied in the context of stationary constraints in stationary or non-stationary environments, giving rise to well-established recursive adaptive implementations when multiple observations are available. In this communication, provided that the additive noise sequence is temporally uncorrelated, we determine the family of non-stationary constraints leading to LCMVEs which can be computed according to a predictor/corrector recursion similar to the Kalman Filter. A particularly noteworthy feature of the recursive formulation introduced is to be fully adaptive in the context of sequential estimation as it allows at each new observation to incorporate or not new constraints.

1. Introduction

In the signal processing literature dealing with parameter estimation, one of the most studied estimation problem is that of identifying the components of a N-dimensional observation vector ($\mathbf{y}$) formed from a linear superposition of P individual signals ($\mathbf{x}$) to noisy data ($\mathbf{v}$): $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}$[1], a.k.a. the linear regression problem, where $\mathbf{H}$ is a N-by-P matrix and $\mathbf{v}$ is a N-dimensional vector. The importance of this problem stems from the fact that a wide range of problems in communications, array processing, and many other areas can be cast in this form [1,2]. As in [3, Section 5.1], we adopt a joint proper complex signals assumption for $\mathbf{x}$ and $\mathbf{v}$, which allows to resort to standard estimation in the mean squared error (MSE) sense defined on the Hilbert space of complex random variables with finite second-order moment. A proper complex random variable is uncorrelated with its complex conjugate. Any result derived with joint proper complex random vectors are valid for real random vectors provided that one substitutes the matrix/vector transpose conjugate for the matrix/vector transpose. Additionally, it is assumed that: (a) $\mathbf{v}$ is zero mean, (b) $\mathbf{x}$ is uncorrelated with $\mathbf{v}$, (c) the model matrix $\mathbf{H}$ and the noise covariance matrix $\mathbf{C_v}$ are either known or specified according to known parametric models. In this setting, the weighted least squares estimator of $\mathbf{x}$ [4]:[2]

$$\widehat{\mathbf{x}}^b = \arg\min_{\mathbf{x}} \left\{ (\mathbf{y} - \mathbf{H}\mathbf{x})^H \mathbf{C}_\mathbf{v}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) \right\} \tag{1a}$$

$$= \left(\mathbf{H}^H \mathbf{C}_\mathbf{v}^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^H \mathbf{C}_\mathbf{v}^{-1} \mathbf{y}, \tag{1b}$$

coincides with the maximum-likelihood estimator [5], if $\mathbf{x}$ is deterministic and $\mathbf{v}$ is Gaussian, and is known to minimize the MSE matrix among all linear unbiased estimators of $\mathbf{x}$, that is $\widehat{\mathbf{x}}^b = \mathbf{W}^{bH}\mathbf{y}$ where [6]:

$$\mathbf{W}^b = \arg\min_{\mathbf{W}} \left\{ E\left[\left(\mathbf{W}^H\mathbf{y} - \mathbf{x}\right)\left(\mathbf{W}^H\mathbf{y} - \mathbf{x}\right)^H \right] \right\} \text{ s.t. } \mathbf{W}^H\mathbf{H} = \mathbf{I} \tag{2a}$$

$$= \mathbf{C}_\mathbf{v}^{-1}\mathbf{H}\left(\mathbf{H}^H \mathbf{C}_\mathbf{v}^{-1} \mathbf{H}\right)^{-1}, \tag{2b}$$

whatever $\mathbf{x}$ is deterministic or random. Furthermore, since the matrix $\mathbf{W}^b$ is as well the solution of [2,6]:

$$\mathbf{W}^b = \arg\min_{\mathbf{W}} \left\{ \mathbf{W}^H \mathbf{C}_\mathbf{v} \mathbf{W} \right\} \text{ s.t. } \mathbf{W}^H\mathbf{H} = \mathbf{I}, \tag{2c}$$

$\widehat{\mathbf{x}}^b$ is also known as the minimum variance distortion less response estimator (MVDRE) [1,2,6]. However, it is well known that the performance achievable by the MVDRE strongly depends on the accurate knowledge on the parametric model of the observations, that is on $\mathbf{H}$ and $\mathbf{C_v}$, and are not particularly robust in

* Corresponding author.
E-mail addresses: francois.vincent@isae.fr (F. Vincent), eric.chaumette@isae.fr (E. Chaumette).

[1] Throughout the present communication, scalars, vectors and matrices are represented, respectively, by italic, bold lowercase and bold uppercase characters. The scalar/matrix/vector transpose conjugate is indicated by the superscript ^{H}. $[\mathbf{A}\,\mathbf{B}]$ and $\begin{bmatrix}\mathbf{A}\\ \mathbf{B}\end{bmatrix}$ denotes respectively the matrix resulting from the horizontal and the vertical concatenation of $\mathbf{A}$ and $\mathbf{B}$. $E[\ \cdot\ ]$ denotes the expectation operator.

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

https://doi.org/10.1016/j.sigpro.2018.03.016

the presence of various types of differences between the model and the actual environment [1, Section 6.6], [7, Section 1], [8]. Thus linearly constrained minimum variance estimators (LCMVEs) [6,9,10] have been developed in which additional linear constraints are imposed to make the MVDRE more robust [1, Section 6.7], [7, Section 1], [8]:

$$\mathbf{W}^b = \arg\min_{\mathbf{W}}\left\{\mathbf{W}^H\mathbf{C_v}\mathbf{W}\right\} \text{ s.t. } \mathbf{W}^H\boldsymbol{\Lambda} = \boldsymbol{\Gamma},\ \boldsymbol{\Lambda} = [\mathbf{H}\ \boldsymbol{\Omega}],\ \boldsymbol{\Gamma} = [\mathbf{I}\ \boldsymbol{\Upsilon}], \tag{3a}$$

$$= \mathbf{C_v}^{-1}\boldsymbol{\Lambda}\left(\boldsymbol{\Lambda}^H\mathbf{C_v}^{-1}\boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Gamma}^H, \tag{3b}$$

where $\boldsymbol{\Omega}$ and $\boldsymbol{\Upsilon}$ are known matrices of the appropriate dimensions, at the expense of an increase of the minimum MSE achieved, since additional degrees of freedom are used by the LCMVE in order to satisfy these constraints. However, firstly, the closed-form solution of the LCMVE (3b) requires the inversion of $\mathbf{C_v}$, which can be too computationally complex for numerous real-world applications. Secondly, $\mathbf{C_v}$ may be unknown and must be learned by an adaptive technique. Interestingly enough, if $\mathbf{x}$ and $\mathbf{v}$ are uncorrelated, $\mathbf{C_v}$ can be replaced by $\mathbf{C_y}$ in (1b), (2b) and (3b), which means that either $\mathbf{C_v}$ can be learned from auxiliary data containing noise only, if available, or $\mathbf{C_y}$ can be used instead and learned from the observations. Therefore, when several observations $\{\mathbf{y}_1, \ldots, \mathbf{y}_k\}$ are available, adaptive implementations of the LCMVE have been developed resorting to constrained stochastic gradient [6,11], constrained recursive least squares [12,13] and constrained Kalman-type [14,15] algorithms. The known equivalence between the LCMVE and the generalized side lobe canceller processor [9,10,16] allows to resort as well to standard (unconstrained) stochastic gradient or recursive least squares [2]. These recursive algorithms belongs to the set of sequential estimation algorithms compatible with applications 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). It is an attractive formulation for embedded systems in which computational time and memory are at a premium, since it does not require that all observations are available for simultaneous ("batch") processing. Last, this can be computationally beneficial in cases in which the number of observations is much larger than the number of signals [17].

However, the aforementioned recursive algorithms can only update sequentially the LCMVE (3b) in non-stationary environments, i.e. when the observation model changes over time ($\mathbf{y}_l = \mathbf{H}_l\mathbf{x} + \mathbf{v}_l$, $1 \le l \le k$), for a given set of linear constraints $\mathbf{W}^H\boldsymbol{\Lambda} = \boldsymbol{\Gamma}$ [2,6,11–15], which defines the set of recursive LCMVEs for stationary constraints. An example of a recursive LCMVE for non-stationary constraints in non-stationary environments is given by the MVDRE $\widehat{\mathbf{x}}_k^b$ of $\mathbf{x}$, based on observations up to and including time k. Indeed, provided that the additive noise sequence $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is temporally uncorrelated, $\widehat{\mathbf{x}}_k^b$ follows a predictor/corrector recursion similar to the Kalman Filter [17, Section 1] [18]:

$$\widehat{\mathbf{x}}_k^b = \widehat{\mathbf{x}}_{k-1}^b + \mathbf{W}_k^{bH}\left(\mathbf{y}_k - \mathbf{H}_k\widehat{\mathbf{x}}_{k-1}^b\right),\ \widehat{\mathbf{x}}_1^b = \left(\mathbf{H}_1^H\mathbf{C}_{\mathbf{v}_1}^{-1}\mathbf{H}_1\right)^{-1}\mathbf{H}_1^H\mathbf{C}_{\mathbf{v}_1}^{-1}\mathbf{y}_1, \tag{4}$$

where $\mathbf{W}_k^b$ is analogous to a Kalman gain at time k. In this case, the set of constraints (2c) is non-stationary since it is defined as $\overline{\mathbf{W}}^H\overline{\mathbf{H}}_k = \mathbf{I}$, where $\overline{\mathbf{H}}_k$ is the matrix resulting from the vertical concatenation of k matrices $\mathbf{H}_1, \ldots, \mathbf{H}_k$, and $\overline{\mathbf{W}}$ is an unknown matrix of the appropriate dimensions. Off course, from a theoretical point of view, if all the observations $\{\mathbf{y}_1, \ldots, \mathbf{y}_k\}$ are stacked into a single vector $\overline{\mathbf{y}}_k^T = \left(\mathbf{y}_1^T, \ldots, \mathbf{y}_k^T\right)$, the "batch form" (3b) obtained from $\overline{\mathbf{y}}_k$ allows to implement LCMVEs with non-stationnary constraints, which are, unfortunately, hardly likely to be computable as the size of $\overline{\mathbf{y}}_k$ increases. Therefore the novel contribution of the present communication is to introduce, provided that the additive noise sequence $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is temporally uncorrelated, the family of linear constraints yielding a LCMVE which can be computed recursively in the form of (4) in place of the "batch form" (3b). It appears that this family only contains non-stationary constraints, including the aforementioned MVDRE. A particularly noteworthy feature of the recursive formulation introduced is to be fully adaptive in the context of sequential estimation as it allows at each new observation to incorporate or not new constraints. The relevance of the proposed recursive formulation of the LCMVE is exemplified in Section 3 in the context of array processing.

2. Recursive linearly constrained minimum variance estimators

In the following: a) the vector space of complex matrices with N rows and P columns is denoted $\mathcal{M}_{\mathbb{C}}(N, P)$, b) the matrix resulting from the vertical concatenation of k matrices $\mathbf{A}_1, \ldots, \mathbf{A}_k$ of same column number is denoted $\overline{\mathbf{A}}_k$. We consider the linear measurement/observation model:

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x} + \mathbf{v}_k, \quad k \ge 1, \tag{5a}$$

where $\mathbf{x}$ is a P-dimensional complex unknown vector, $\mathbf{y}_k$ is a N_k-dimensional complex measurement/observation vector, $\mathbf{H}_k \in \mathcal{M}_{\mathbb{C}}(N_k, P)$ and the complex noise sequence $\{\mathbf{v}_k\}_{k\ge 1}$ is zero-mean and temporally uncorrelated. Then (5a) can be extended on a horizon of k points from the first observation as:

$$\overline{\mathbf{y}}_k = \begin{pmatrix}\mathbf{y}_1\\ \vdots\\ \mathbf{y}_k\end{pmatrix} = \begin{bmatrix}\mathbf{H}_1\\ \vdots\\ \mathbf{H}_k\end{bmatrix}\mathbf{x} + \begin{pmatrix}\mathbf{v}_1\\ \vdots\\ \mathbf{v}_k\end{pmatrix}$$

$$= \overline{\mathbf{H}}_k\mathbf{x} + \overline{\mathbf{v}}_k, \quad \left|\begin{array}{l}\overline{\mathbf{y}}_k, \overline{\mathbf{v}}_k \in \mathcal{M}_{\mathbb{C}}(\mathcal{N}_k, 1)\\ \overline{\mathbf{H}}_k \in \mathcal{M}_{\mathbb{C}}(\mathcal{N}_k, P)\\ \mathcal{N}_k = \sum_{l=1}^k N_l\end{array}\right. . \tag{5b}$$

Let $\overline{\mathbf{W}}_k = \begin{bmatrix}\overline{\mathbf{D}}_{k-1}\\ \mathbf{W}_k\end{bmatrix}$ where $\overline{\mathbf{D}}_{k-1} \in \mathcal{M}_{\mathbb{C}}\left(\mathcal{N}_{k-1}, P\right)$ and $\mathbf{W}_k \in \mathcal{M}_{\mathbb{C}}(N_k, P)$. The aim is to look for the family of linear constraints:

$$\overline{\mathbf{W}}_k^H\overline{\boldsymbol{\Lambda}}_k = \boldsymbol{\Gamma}_k,\ \overline{\boldsymbol{\Lambda}}_k = \left[\overline{\mathbf{H}}_k\ \overline{\boldsymbol{\Omega}}_k\right],\ \boldsymbol{\Gamma}_k = [\mathbf{I}\ \boldsymbol{\Upsilon}_k], \tag{6}$$

where $\overline{\boldsymbol{\Omega}}_k$ and $\boldsymbol{\Upsilon}_k$ are known matrices of the appropriate dimensions, yielding a LCMVE $\widehat{\mathbf{x}}_k^b = \overline{\mathbf{W}}_k^{bH}\overline{\mathbf{y}}_k$ where (3a) and (3b):

$$\overline{\mathbf{W}}_k^b = \arg\min_{\overline{\mathbf{W}}_k}\left\{\overline{\mathbf{W}}_k^H\mathbf{C}_{\overline{\mathbf{v}}_k}\overline{\mathbf{W}}_k\right\} \text{ s.t. } \overline{\mathbf{W}}_k^H\overline{\boldsymbol{\Lambda}}_k = \boldsymbol{\Gamma}_k \tag{7a}$$

$$= \mathbf{C}_{\overline{\mathbf{v}}_k}^{-1}\overline{\boldsymbol{\Lambda}}_k\left(\overline{\boldsymbol{\Lambda}}_k^H\mathbf{C}_{\overline{\mathbf{v}}_k}^{-1}\overline{\boldsymbol{\Lambda}}_k\right)^{-1}\boldsymbol{\Gamma}_k^H, \tag{7b}$$

which can be computed according to a predictor/corrector recursion of the form, $\forall k \ge 2$:

$$\widehat{\mathbf{x}}_k^b = \widehat{\mathbf{x}}_{k-1}^b + \mathbf{W}_k^{bH}\left(\mathbf{y}_k - \mathbf{H}_k\widehat{\mathbf{x}}_{k-1}^b\right) = \left(\mathbf{I} - \mathbf{W}_k^{bH}\mathbf{H}_k\right)\widehat{\mathbf{x}}_{k-1}^b + \mathbf{W}_k^{bH}\mathbf{y}_k. \tag{8}$$

A key point to solve the problem at hand is to notice that, since $\mathbf{C}_{\mathbf{v}_l,\mathbf{v}_k} = \mathbf{C}_{\mathbf{v}_k}\delta_k^l$, then for any $\overline{\mathbf{W}}_k$ satisfying (6):

$$\mathbf{P}_k\left(\overline{\mathbf{W}}_k\right) = \overline{\mathbf{W}}_k^H\mathbf{C}_{\overline{\mathbf{v}}_k}\overline{\mathbf{W}}_k = \overline{\mathbf{D}}_{k-1}^H\mathbf{C}_{\overline{\mathbf{v}}_{k-1}}\overline{\mathbf{D}}_{k-1} + \mathbf{W}_k^H\mathbf{C}_{\mathbf{v}_k}\mathbf{W}_k = \mathbf{P}_k\left(\overline{\mathbf{D}}_{k-1},\ \mathbf{W}_k\right), \tag{9}$$

which suggests that some ad hoc linear constraints (6) could yield separable solutions for $\overline{\mathbf{D}}_{k-1}$ and $\mathbf{W}_k$, which is investigated in a first step.

- **First step**

If we recast $\overline{\boldsymbol{\Lambda}}_k = \left[\overline{\mathbf{H}}_k\ \overline{\boldsymbol{\Omega}}_k\right]$ as $\overline{\boldsymbol{\Lambda}}_k = \begin{bmatrix}\overline{\boldsymbol{\Phi}}_{k-1}\\ \boldsymbol{\Phi}_k\end{bmatrix}$ where $\overline{\boldsymbol{\Phi}}_{k-1} = \left[\overline{\mathbf{H}}_{k-1}\ \overline{\boldsymbol{\Omega}}_{k-1}\right]$ and $\boldsymbol{\Phi}_k = [\mathbf{H}_k\ \boldsymbol{\Omega}_k]$, then an equivalent form of (6) is:

$$\overline{\mathbf{W}}_k^H\overline{\boldsymbol{\Lambda}}_k = \boldsymbol{\Gamma}_k \ \Leftrightarrow\ \overline{\mathbf{D}}_{k-1}^H\overline{\boldsymbol{\Phi}}_{k-1} = \boldsymbol{\Gamma}_k - \mathbf{W}_k^H\boldsymbol{\Phi}_k. \tag{10}$$

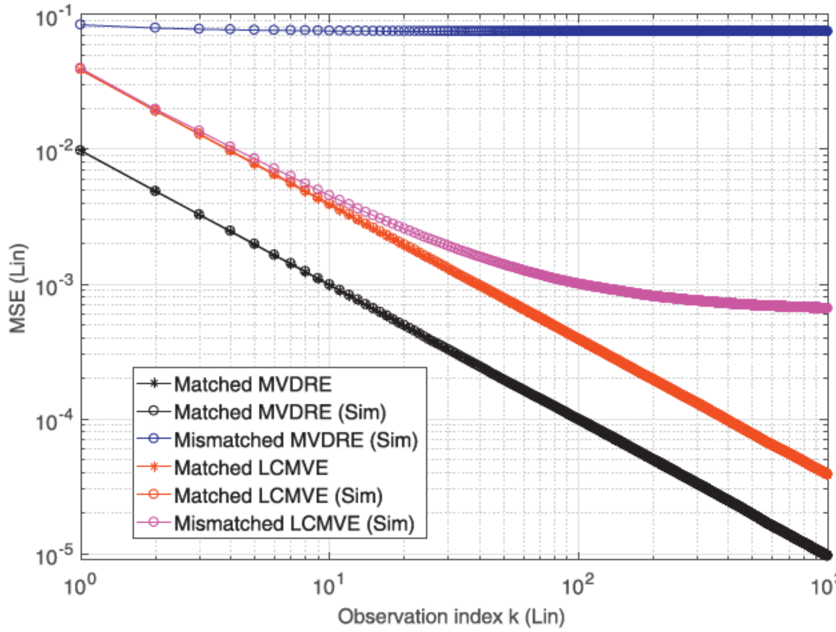

Fig. 1. MSE of recursive MVDREs and LCMVEs versus k.

Therefore, according to (9):

$$\overline{\mathbf{D}}_{k-1}^{b} = \arg\min_{\overline{\mathbf{D}}_{k-1}} \left\{ \overline{\mathbf{D}}_{k-1}^{H} \mathbf{C}_{\overline{\mathbf{v}}_{k-1}} \overline{\mathbf{D}}_{k-1} \right\} \text{ s.t. } \overline{\mathbf{D}}_{k-1}^{H} \overline{\boldsymbol{\Phi}}_{k-1} = \boldsymbol{\Gamma}_{k} - \mathbf{W}_{k}^{H} \boldsymbol{\Phi}_{k}, \tag{11a}$$

that is, provided that $\overline{\boldsymbol{\Phi}}_{k-1}$ and $\mathbf{C}_{\overline{\mathbf{v}}_{k-1}}$ are full rank, (3a) and (3b):

$$\overline{\mathbf{D}}_{k-1}^{b} = \mathbf{C}_{\overline{\mathbf{v}}_{k-1}}^{-1} \overline{\boldsymbol{\Phi}}_{k-1} \left(\overline{\boldsymbol{\Phi}}_{k-1}^{H} \mathbf{C}_{\overline{\mathbf{v}}_{k-1}}^{-1} \overline{\boldsymbol{\Phi}}_{k-1} \right)^{-1} \left(\boldsymbol{\Gamma}_{k} - \mathbf{W}_{k}^{H} \boldsymbol{\Phi}_{k} \right)^{H}. \tag{11b}$$

It is noteworthy that (11b) can be recasted as:

$$\overline{\mathbf{D}}_{k-1}^{b} = \overline{\mathbf{W}}_{k-1}^{b} \left(\mathbf{I} - \mathbf{W}_{k}^{H} \mathbf{H}_{k} \right)^{H} + \mathbf{C}_{\overline{\mathbf{v}}_{k-1}}^{-1} \overline{\boldsymbol{\Phi}}_{k-1} \left(\overline{\boldsymbol{\Phi}}_{k-1}^{H} \mathbf{C}_{\overline{\mathbf{v}}_{k-1}}^{-1} \overline{\boldsymbol{\Phi}}_{k-1} \right)^{-1} \boldsymbol{\Theta}_{k-1}^{H}, \tag{12a}$$

where $\boldsymbol{\Theta}_{k-1} = \boldsymbol{\Gamma}_{k} - \mathbf{W}_{k}^{H} \boldsymbol{\Phi}_{k} - \left(\mathbf{I} - \mathbf{W}_{k}^{H} \mathbf{H}_{k} \right) \boldsymbol{\Gamma}_{k-1}$ and:

$$\overline{\mathbf{W}}_{k-1}^{b} = \arg\min_{\overline{\mathbf{W}}_{k-1}} \left\{ \overline{\mathbf{W}}_{k-1}^{H} \mathbf{C}_{\overline{\mathbf{v}}_{k-1}} \overline{\mathbf{W}}_{k-1} \right\} \text{ s.t. } \overline{\mathbf{W}}_{k-1}^{H} \overline{\boldsymbol{\Phi}}_{k-1} = \boldsymbol{\Gamma}_{k-1}. \tag{12b}$$

Thus, the LCMVE (7a) follows a predictor/corrector recursion (8) with separable solutions for $\overline{\mathbf{D}}_{k-1}$ and $\mathbf{W}_{k}$ iff, $\forall \mathbf{W}_{k}$:

$$\begin{aligned} \overline{\mathbf{D}}_{k-1}^{b} = \overline{\mathbf{W}}_{k-1}^{b} \left(\mathbf{I} - \mathbf{W}_{k}^{H} \mathbf{H}_{k} \right)^{H} &\Leftrightarrow \boldsymbol{\Theta}_{k-1} = \mathbf{0} \\ &= \boldsymbol{\Gamma}_{k} - \mathbf{W}_{k}^{H} \boldsymbol{\Phi}_{k} - \left(\mathbf{I} - \mathbf{W}_{k}^{H} \mathbf{H}_{k} \right) \boldsymbol{\Gamma}_{k-1}, \end{aligned}$$

that is iff, $\forall \mathbf{W}_{k}$: $\boldsymbol{\Gamma}_{k} - \mathbf{W}_{k}^{H} \boldsymbol{\Phi}_{k} = \left[\mathbf{I} - \mathbf{W}_{k}^{H} \mathbf{H}_{k} \ \ \boldsymbol{\Upsilon}_{k} - \mathbf{W}_{k}^{H} \boldsymbol{\Omega}_{k} \right] = \left(\mathbf{I} - \mathbf{W}_{k}^{H} \mathbf{H}_{k} \right) \boldsymbol{\Gamma}_{k-1}$, which requires that $\boldsymbol{\Gamma}_{k-1} = \left[\mathbf{I} \ \ \boldsymbol{\Upsilon}_{k-1} \right]$ where $\boldsymbol{\Upsilon}_{k-1} = \boldsymbol{\Upsilon}_{k}$, and $\boldsymbol{\Omega}_{k} = \mathbf{H}_{k} \boldsymbol{\Upsilon}_{k-1}$. Ergo, the LCMVE (7a) follows a predictor/corrector recursion (8) iff (11a):

$$\overline{\mathbf{D}}_{k-1}^{H} \overline{\boldsymbol{\Phi}}_{k-1} = \boldsymbol{\Gamma}_{k} - \mathbf{W}_{k}^{H} \boldsymbol{\Phi}_{k} = \left(\mathbf{I} - \mathbf{W}_{k}^{H} \mathbf{H}_{k} \right) [\mathbf{I} \ \ \boldsymbol{\Upsilon}_{k-1}],$$

or equivalently, iff:

$$\mathcal{C}_{k}^{1} : \mathbf{W}_{k}^{H} \begin{bmatrix} \overline{\mathbf{H}}_{k-1} & \overline{\boldsymbol{\Omega}}_{k-1} \\ \mathbf{H}_{k} & \mathbf{H}_{k} \boldsymbol{\Upsilon}_{k-1} \end{bmatrix} = [\mathbf{I} \quad \boldsymbol{\Upsilon}_{k-1}], \tag{13}$$

and (12b) amounts to:

$$\begin{aligned} \overline{\mathbf{W}}_{k-1}^{b} = \arg\min_{\overline{\mathbf{W}}_{k-1}} \left\{ \overline{\mathbf{W}}_{k-1}^{H} \mathbf{C}_{\overline{\mathbf{v}}_{k-1}} \overline{\mathbf{W}}_{k-1} \right\} \text{ s.t. } \overline{\mathbf{W}}_{k-1}^{H} \left[\overline{\mathbf{H}}_{k-1} \ \ \overline{\boldsymbol{\Omega}}_{k-1} \right] \\ = [\mathbf{I} \ \ \boldsymbol{\Upsilon}_{k-1}], \end{aligned} \tag{14}$$

which means that $\widehat{\mathbf{x}}_{k-1}^{b} = \overline{\mathbf{W}}_{k-1}^{bH} \overline{\mathbf{y}}_{k-1}$ is a LCMVE as well. The specific form of (13) reflects the fact that the linear constraints at time $k-1$ (14) propagates at time k via $\mathcal{C}_{k}^{1}$ (13). Interestingly enough, additional linear constraints on $\mathbf{W}_{k}$ can be introduced on-line as shown in a second step.

- **Second step**

Let us notice that $\mathbf{P}_{k}(\overline{\mathbf{D}}_{k-1}^{b}, \mathbf{W}_{k}) = \overline{\mathbf{D}}_{k-1}^{bH} \mathbf{C}_{\overline{\mathbf{v}}_{k-1}} \overline{\mathbf{D}}_{k-1}^{b} + \mathbf{W}_{k}^{H} \mathbf{C}_{\mathbf{v}_{k}} \mathbf{W}_{k}$ can be recasted as:

$$\mathbf{P}_{k}\left(\overline{\mathbf{D}}_{k-1}^{b}, \mathbf{W}_{k} \right) = \mathbf{W}_{k}^{H} \mathbf{C}_{\mathbf{v}_{k}} \mathbf{W}_{k} + \left(\mathbf{I} - \mathbf{W}_{k}^{H} \mathbf{H}_{k} \right) \mathbf{P}_{k-1}\left(\overline{\mathbf{W}}_{k-1}^{b} \right) \left(\mathbf{I} - \mathbf{W}_{k}^{H} \mathbf{H}_{k} \right)^{H},$$

that is as:

$$\mathbf{P}_{k}\left(\overline{\mathbf{D}}_{k-1}^{b}, \mathbf{W}_{k} \right) = E\left[\left(\overline{\mathbf{D}}_{k-1}^{bH} \overline{\mathbf{v}}_{k-1} - \mathbf{W}_{k}^{H} \mathbf{v}_{k} \right) \left(\overline{\mathbf{D}}_{k-1}^{bH} \overline{\mathbf{v}}_{k-1} - \mathbf{W}_{k}^{H} \mathbf{v}_{k} \right)^{H} \right], \tag{16}$$

where $\overline{\mathbf{D}}_{k-1}^{bH} \overline{\mathbf{v}}_{k-1} - \mathbf{W}_{k}^{H} \mathbf{v}_{k} = \overline{\mathbf{W}}_{k-1}^{bH} \overline{\mathbf{v}}_{k-1} - \mathbf{W}_{k}^{H} (\mathbf{H}_{k} \overline{\mathbf{W}}_{k-1}^{bH} \overline{\mathbf{v}}_{k-1} + \mathbf{v}_{k})$.

Then two cases are possible:

(1) no additional linear constraints on $\mathbf{W}_{k}$ are introduced. In that case, as shown at the first step, the LCMVE only propagates at time k the existing linear constraints at time $k-1$ (14) via $\mathcal{C}_{k}^{1}$ (13). Then the solution of: $\mathbf{W}_{k}^{b} = \arg\min_{\mathbf{W}_{k}} \{ \mathbf{P}_{k}(\overline{\mathbf{D}}_{k-1}^{b}, \mathbf{W}_{k}) \}$, is well known and given by [18], [17, Section 1]:

$$\widehat{\mathbf{x}}_{k}^{b} = \widehat{\mathbf{x}}_{k-1}^{b} + \mathbf{W}_{k}^{bH} \left(\mathbf{y}_{k} - \mathbf{H}_{k} \widehat{\mathbf{x}}_{k-1}^{b} \right), \tag{17a}$$

$$\mathbf{S}_{k} = \mathbf{H}_{k} \mathbf{P}_{k-1}^{b} \mathbf{H}_{k}^{H} + \mathbf{C}_{\mathbf{v}_{k}}, \quad \mathbf{W}_{k}^{b} = \mathbf{S}_{k}^{-1} \mathbf{H}_{k} \mathbf{P}_{k-1}^{b}, \tag{17b}$$

$$\mathbf{P}_{k}^{b} = \left(\mathbf{I} - \mathbf{W}_{k}^{bH} \mathbf{H}_{k} \right) \mathbf{P}_{k-1}^{b}. \tag{17c}$$

(2) additional linear constraints on $\mathbf{W}_{k}$, i.e. $\mathbf{W}_{k}^{H} \boldsymbol{\Delta}_{k} = \mathbf{T}_{k}$, are introduced on-line and (13) must be updated to take them into ac-

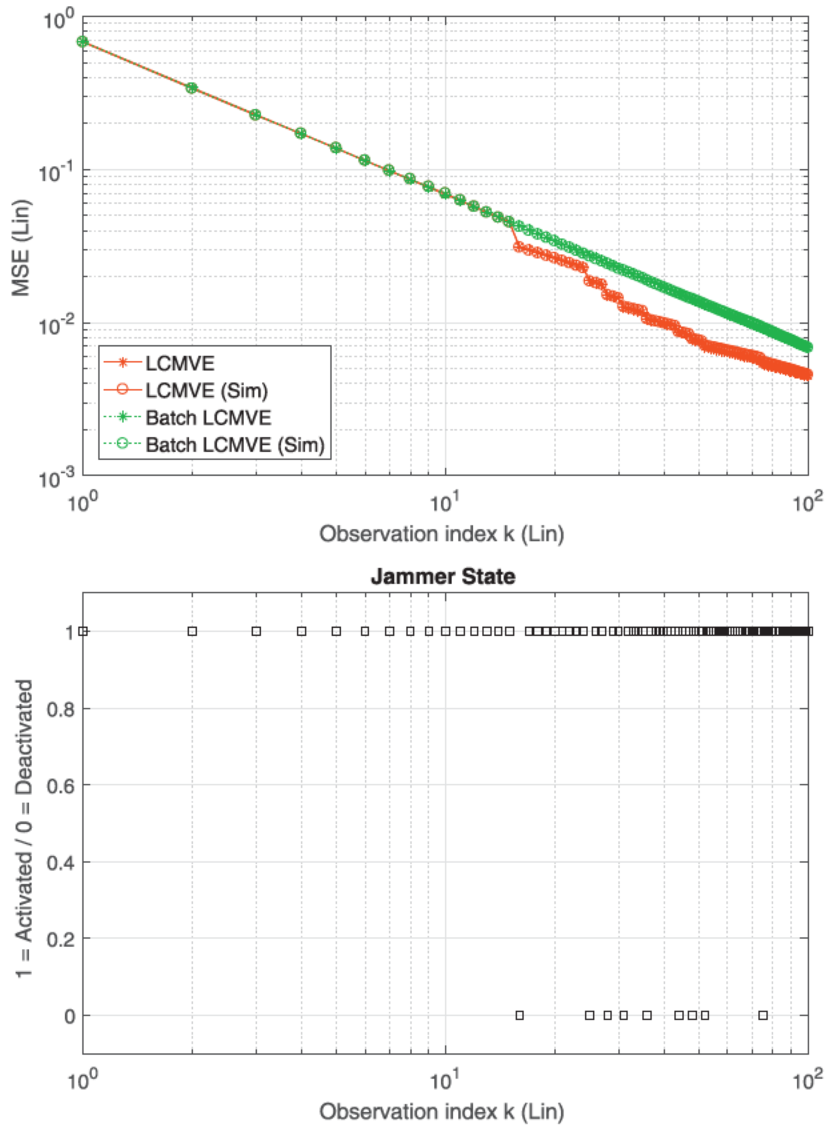

Fig. 2. MSE of recursive and batch form LCMVEs versus k, $\mathcal{P}_J = 0.9$.

count, leading to:

$$\mathcal{C}_k^2 : \overline{\mathbf{W}}_k^H \begin{bmatrix} \overline{\mathbf{H}}_{k-1} & \overline{\boldsymbol{\Omega}}_{k-1} & \mathbf{0} \\ \mathbf{H}_k & \mathbf{H}_k \boldsymbol{\Upsilon}_{k-1} & \boldsymbol{\Delta}_k \end{bmatrix} = [\mathbf{I}\ \boldsymbol{\Upsilon}_{k-1}\ \mathbf{T}_k]. \tag{18}$$

In that case, the solution of: $\mathbf{W}_k^b = \arg\min_{\mathbf{W}_k} \{\mathbf{P}_k(\overline{\mathbf{D}}_{k-1}^b, \mathbf{W}_k)\}$ s.t. $\mathbf{W}_k^H \boldsymbol{\Delta}_k = \mathbf{T}_k$, is analogous to a linearly constrained Wiener filter [2, (2.113)]. Thus (7a) follows a predictor/corrector recursion given by:

$$\widehat{\mathbf{x}}_k^b = \widehat{\mathbf{x}}_{k-1}^b + \mathbf{W}_k^{bH}\left(\mathbf{y}_k - \mathbf{H}_k \widehat{\mathbf{x}}_{k-1}^b\right), \tag{19a}$$

$$\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_{k-1}^b \mathbf{H}_k^H + \mathbf{C}_{\mathbf{v}_k}, \ \mathbb{W}_k = \mathbf{S}_k^{-1} \mathbf{H}_k \mathbf{P}_{k-1}^b, \ \mathbb{T}_k = \mathbf{T}_k - \mathbb{W}_k^H \boldsymbol{\Delta}_k, \tag{19b}$$

$$\mathbf{W}_k^b = \mathbb{W}_k + \mathbf{S}_k^{-1} \boldsymbol{\Delta}_k \left(\boldsymbol{\Delta}_k^H \mathbf{S}_k^{-1} \boldsymbol{\Delta}_k\right)^{-1} \mathbb{T}_k^H, \tag{19c}$$

$$\mathbf{P}_k^b = \left(\mathbf{I} - \mathbb{W}_k^H \mathbf{H}_k\right) \mathbf{P}_{k-1}^b + \mathbb{T}_k \left(\boldsymbol{\Delta}_k^H \mathbf{S}_k^{-1} \boldsymbol{\Delta}_k\right)^{-1} \mathbb{T}_k^H. \tag{19d}$$

In both cases:

$$\mathbf{P}_{k-1}^b = \min_{\overline{\mathbf{W}}_{k-1}} \left\{ \overline{\mathbf{W}}_{k-1}^H \mathbf{C}_{\overline{\mathbf{v}}_{k-1}} \overline{\mathbf{W}}_{k-1} \right\} \text{ s.t. } \overline{\mathbf{W}}_{k-1}^H \left[\overline{\mathbf{H}}_{k-1}\ \overline{\boldsymbol{\Omega}}_{k-1}\right] = [\mathbf{I}\ \boldsymbol{\Upsilon}_{k-1}],$$

which means that the same rationale can be applied at time $k-1$ and so forth until time $k = 2$.

- **Summary**

The linear constraints (6) allowing the LCMVE to follow the predictor/corrector recursion:

$$\widehat{\mathbf{x}}_k^b = \widehat{\mathbf{x}}_{k-1}^b + \mathbf{W}_k^{bH}\left(\mathbf{y}_k - \mathbf{H}_k \widehat{\mathbf{x}}_{k-1}^b\right), \tag{20a}$$

are built as follows:

▹ at time $k = 1$, a set of linear constraints of the form:

$$\mathbf{W}_1^H \boldsymbol{\Lambda}_1 = \boldsymbol{\Gamma}_1, \ \{\boldsymbol{\Lambda}_1 = \mathbf{H}_1, \ \boldsymbol{\Gamma}_1 = \mathbf{I}\} \text{ or } \{\boldsymbol{\Lambda}_1 = [\mathbf{H}_1\ \boldsymbol{\Omega}_1], \ \boldsymbol{\Gamma}_1 = [\mathbf{I}\ \boldsymbol{\Upsilon}_1]\}, \tag{20b}$$

must be set, leading to:

$$\widehat{\mathbf{x}}_1^b = \mathbf{W}_1^{bH} \mathbf{y}_1, \ \mathbf{W}_1^b = \mathbf{C}_{\mathbf{v}_1}^{-1} \boldsymbol{\Lambda}_1 \left(\boldsymbol{\Lambda}_1^H \mathbf{C}_{\mathbf{v}_1}^{-1} \boldsymbol{\Lambda}_1\right)^{-1} \boldsymbol{\Gamma}_1^H,$$
$$\mathbf{P}_1^b = \boldsymbol{\Gamma}_1 \left(\boldsymbol{\Lambda}_1^H \mathbf{C}_{\mathbf{v}_1}^{-1} \boldsymbol{\Lambda}_1\right)^{-1} \boldsymbol{\Gamma}_1^H, \tag{20c}$$

▹ at time $k \geq 2$:

(a) either no additional linear constraints on $\mathbf{W}_k$ are introduced and $\widehat{\mathbf{x}}_k^b$ must be computed according to (17a)–(17c) in order to propagate the existing set of linear constraints,

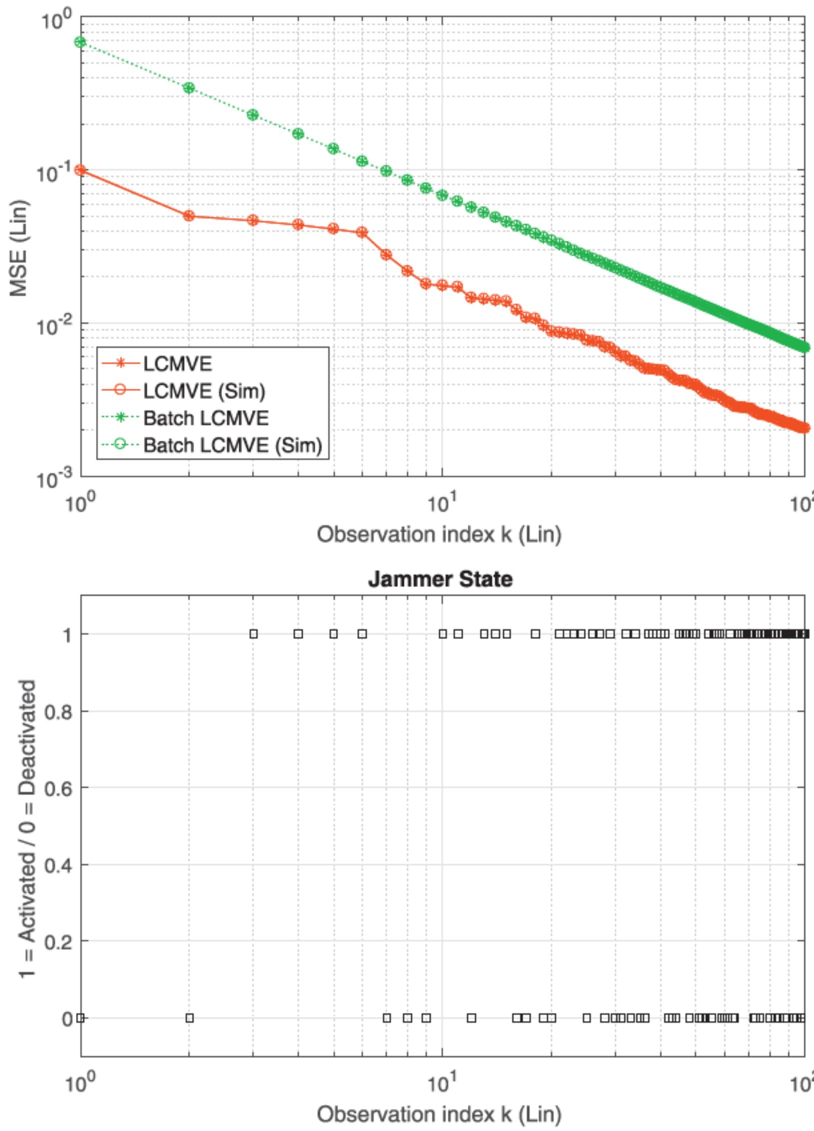

Fig. 3. MSE of recursive and batch form LCMVEs versus k, $\mathcal{P}_J = 0.5$.

(b) or additional linear constraints on $\mathbf{W}_k$, i.e. $\mathbf{W}_k^H \boldsymbol{\Delta}_k = \mathbf{T}_k$, are introduced on-line and $\widehat{\mathbf{x}}_k^b$ must be computed according to (19a)–(19d) in order to propagate the updated set of linear constraints.

3. Illustrative examples

To some extent, the LCMVE can robustify the MVDRE in the presence of parametric modelling errors in measurement matrices $\mathbf{H}_k \triangleq \mathbf{H}_k(\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a deterministic vector value. Indeed, if the value of $\boldsymbol{\theta}$ is known during the observations, then its value can be incorporated into any expression involving the parametric model, such as MLEs, MVDREs, LCMVEs, etc... Otherwise an estimated value $\widehat{\theta}$ must be provided. Experimental systems attempt to eliminate or minimize the parametric modelling errors $\widehat{\theta} - \theta$ by careful calibration of the system. However, system parameters $\boldsymbol{\theta}$ may change over time due to thermal effects, aging of components, changes in the location of the sensors, etc... Thus, for the batch form of MVDRE, it is common place to add derivative constraints in order to mitigate the effect on $\mathbf{H}_k(\boldsymbol{\theta})$ of a small change in system parameters $\boldsymbol{\theta}$ [1, Section 6.7.1], leading to a batch form of LCMVE. Thanks to the recursive form of the LCMVE released above, this mitigation technique can now be used in sequential estimation in the form of on-line linear constraints $\mathbf{W}_k^H \partial \mathbf{H}_k(\widehat{\theta}) / \partial \theta = \mathbf{0}$, if $\boldsymbol{\theta}$ reduces to a single value, for the sake of simplicity. As an example we consider a ULA of $N = 100$ sensors equally spaced at $\widehat{\theta} = \lambda/2$ (half-wavelength) and an impinging signal $x = (1 + j)/\sqrt{2}$ with broadside angle $\alpha = 10^\circ$, embedded in a spatially and temporally white noise: $\mathbf{y}_k = \mathbf{h}_k(\theta, \alpha)x + \mathbf{v}_k$, $\mathbf{h}_k^T(\theta, \alpha) = (1, \ldots, e^{j2\pi(N-1)\frac{\theta}{\lambda}\sin(\alpha)})$, $\mathbf{C}_{\mathbf{v}_l, \mathbf{v}_k} = \mathbf{I}\delta_k^l$. Due to a calibration error, or array deformation (thermal effects, aging, etc ...), the actual inter-sensor distance is $\theta = 0.99\widehat{\theta}$, i.e. $\widehat{\theta} - \theta = \lambda/200$. We superimpose on the Fig. 1 , the MSE (P_k^b) of the recursive MVDRE based on (17a)–(17c) and the MSE of the recursive LCMVE based on (19a)–(19d) where $T_k = 0$ and $\boldsymbol{\Delta}_k = \partial \mathbf{h}_k(\widehat{\theta}, \alpha)/\partial\theta$. The "Matched", respectively the "Mismatched", estimators are based on recursions computed with the true value θ, respectively with the assumed value $\widehat{\theta}$. To check the validity of (17a)–(17c) and (19a)–(19d) for the true value θ, the empirical MSEs of the MVDRE and of the LCMVE are assessed with 10^4 Monte-Carlo trials ("... (Sim)"). Fig. (1) exemplifies the strong dependency of MVDRE's MSE on the accurate knowledge on the parametric model, and the trade-off brought by the recursive LCMVE. If $\mathbf{H}_k(\boldsymbol{\theta})$ is perfectly known, then the addition of on-line linear constraints increases the minimum MSE achieved. However, if the on-line linear constraints are adequately chosen,

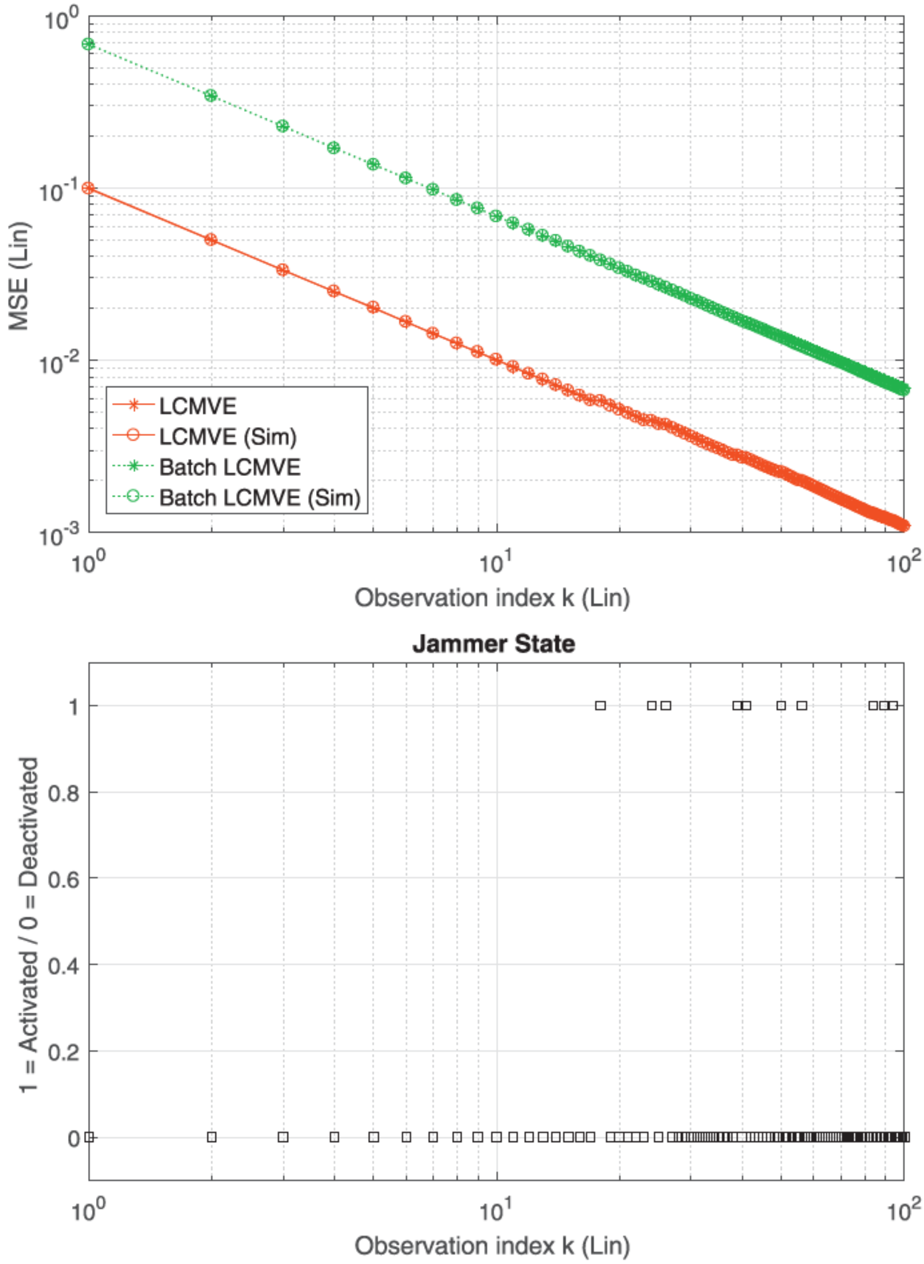

Fig. 4. MSE of recursive and batch form LCMVEs versus k, $\mathcal{P}_J = 0.1$.

the LCMVE obtained may become more robust to parametric modelling errors than the MVDRE. Please note that both the recursive MVDRE and the recursive LCMVE considered are associated to non-stationary constraints (13) and (18), respectively.

If calibration uncertainties must be mitigated for each new observations $\mathbf{y}_k$, in some sequential estimation problems it is more optimal to add on-line constraints that are triggered by a preprocessing of $\mathbf{y}_k$ or by external information on the environment. As an example we consider the same scenario as above but involving a smaller ULA of $N = 10$ sensors, which can be regarded as perfectly calibrated ($\theta = \widehat{\theta}$). An intermittent jammer is located at a known broadside angle α_J in the angular vicinity of the signal of interest, i.e. $\alpha = \alpha_J - \alpha_{3dB}/4$, where α_{3dB} denotes the beamwidth. The jammer to noise power (JNR) is 40 dB and its probability of activation at each observation is denoted by $\mathcal{P}_J$. We assume that the jammer is detected whenever it is activated. At each jammer detection, the null constraint [1, Section 6.7.1] $\mathbf{w}_k^H \mathbf{h}_k(\theta, \alpha_J) = 0$ is added to cancel the jammer signal, and the recursive LCMVE is updated according to (19a)–(19d) where $T_k = 0$ and $\boldsymbol{\Delta}_k = \mathbf{h}_k(\theta, \alpha_J)$. In the absence of jammer detection, the recursive LCMVE is updated without additional constraint, that is according to (17a)–(17c). We compare the proposed dynamic jammer cancellation with the standard procedure [1, Section 6.7.1] which consists in imposing a permanent null constraint $\overline{\mathbf{w}}_k^H \overline{\mathbf{h}}_k(\theta, \alpha_J) = 0$ in the batch form of the LCMVE ((17a)–(17c)), i.e. $\boldsymbol{\Gamma}_k = [0 \ 1]$ and $\overline{\boldsymbol{\Lambda}}_k = [\overline{\mathbf{h}}_k(\theta, \alpha) \ \overline{\mathbf{h}}_k(\theta, \alpha_J)]$. When the null constraint is set, the jammer signal is cancelled at the expense of an increase of the output noise power in comparison with a jammer free scenario, which increases the minimum MSE achieved. Therefore, to limit the increase of the MSE achieved, the null constraint must be set only when the jammer is activated, which is highlighted by Figs. 2–4 displaying the MSE of both solutions obtained for 3 values of $\mathcal{P}_J$: 0.9, 0.5, 0.1. As expected, the superiority of the recursive LCMVE over the batch form LCMVE increases as $\mathcal{P}_J$ decreases.

4. Conclusions

In this communication, for multiple noisy linear observations which noises are uncorrelated, we have introduced the family of linear constraints yielding a LCMVE computable via a predictor/corrector recursion similar to the Kalman Filter in place of the "batch form". It appears that this family only contains non-stationary constraints. A noteworthy feature of the recursive formulation introduced is to be fully adaptive in the context of sequential estimation as it allows optional constraints addition that can be triggered by a preprocessing of each new observation or external information on the environment.

Acknowledgment

This work has been partially supported by the DGA/MRIS (2015.60.0090.00.470.75.01).

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.sigpro.2018.03.016.

References

[1] H.L.V. Trees, Optimum Array Processing, Wiley-Interscience, 2002.
[2] P.S.R. Diniz, Adaptive Filtering: Algorithms and Practical Implementation (4Ed), in: Springer Science+Business Media, 2013.
[3] P.J. Schreier, L.L. Scharf, Statistical Signal Processing of Complex-Valued Data, Cambridge University Press, 2010.
[4] R. L. Plackett, Some theorems in least squares, Biometrika 37 (1950) 149–157.
[5] F. C. Schweppe, Sensor array data processing formultiple signal sources, IEEE Trans. IT 14 (1968) 294–305.
[6] O. L. Frost, An algorithm for linearly constrained adaptive array processing, Proc. IEEE 60 (8) (1972) 926–935.
[7] J. Li, P. Stoica, Robust Adaptive Beamforming, Wiley-Interscience, 2006.
[8] S. A. Vorobyov, Principles of minimum variance robust adaptive beamforming design, Elsevier Signal Process. 93 (2013) 3264–3277.
[9] L.J. Griffiths, C.W. Jim, An alternative approach to linearly constrained adaptive beamfonning, IEEE Trans. on AP 30 (1) (1982) 27–34.
[10] S. Werner, J.A. Apolinario Jr, M.L.R. De Campos, On the equivalence of RLS implementations of LCMV and GSC processors, IEEE SP Letters 10 (12) (2003) 356–359.
[11] X. Guo, L. Chu, B. Li, Robust adaptive LCMV beamformer based on an iterative suboptimal solution, Radioengineering 24 (2) (2015) 572–582.
[12] L.S. Resende, J.M.T. Romano, M.G. Bellanger, A fast least squares algorithm for linearly constrained adaptive filtering, IEEE Trans. SP 44 (5) (1996) 1168–1174.
[13] R.C. de Lamare, Adaptive reduced-rank LCMV beamforming algorithms based on joint iterative optimisation of filters, Electron. Lett. 44 (8) (2008) 565–566.
[14] Y.H. Chen, C.T. Chiang, Adaptive beamforming using the constrained kalman filter, IEEE Trans. AP 41 (11) (1993) 1576–1580.
[15] D. Cherkassky, S. Gannot, New insights into the kalman filter beamformer: applications to speech and robustness, IEEE SP Lett. 23 (3) (2016) 376–380.
[16] B.R. Breed, J. Strauss, A short proof of the equivalence of LCMV and GSC beamforming, IEEE SP Lett. 9 (6) (2002) 168–169.
[17] J.L. Crassidis, J.L. Junkins, Optimal Estimation of Dynamic Systems (2ed), CRC Press, Taylor & Francis Group, 2012.
[18] A.H. Sayed, T. Kailath, A state-space approach to adaptive RLS filtering, IEEE SP Mag. 11 (3) (1994) 18–60.