New Results on LMVDR Estimators for LDSS Models

Eric Chaumette, François Vincent, Benoit Priot, Gael Pages and Arnaud Dion
University of Toulouse - ISAE-Supaero, Toulouse, France
Emails: (eric.chaumette, francois.vincent, benoit.priot, gael.pages, arnaud.dion)@isae.fr

***Abstract*—In the context of linear discrete state-space (LDSS) models, we generalize a result lately introduced in the restricted case of invertible state matrices, namely that the linear minimum variance distortionless response (LMVDR) filter shares exactly the same recursion as the linear least mean squares (LLMS) filter, aka the Kalman filter (KF), except for the initialization. An immediate benefit is the introduction of LMVDR fixed-point and fixed-lag smoothers (and possibly other smoothers or predictors), which has not been possible so far. This result is particularly noteworthy given the fact that, although LMVDR estimators are sub-optimal in mean-squared error sense, they are infinite impulse response distortionless estimators which do not depend on the prior knowledge on the mean and covariance matrix of the initial state. Thus the LMVDR estimators may outperform the usual LLMS estimators in case of misspecification of the prior knowledge on the initial state. Seen from this perspective, we also show that the LMVDR filter can be regarded as a generalization of the information filter form of the KF. On another note, LMVDR estimators may also allow to derive unexpected results, as highlighted with the LMVDR fixed-point smoother.**

I. Introduction

We consider the general class of linear discrete state-space (LDSS) models represented with the state and measurement equations, respectively[1],

$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{w}_{k-1} \quad (1a)$$
$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \quad (1b)$$

where the time index $k \geq 1$, $\mathbf{x}_k$ is the P_k-dimensional state vector, $\mathbf{y}_k$ is the N_k-dimensional measurement vector and the model matrices $\mathbf{F}_k$ and $\mathbf{H}_k$ are known. Unless otherwise stated, the process noise sequence $\{\mathbf{w}_k\}$ and the measurement noise sequence $\{\mathbf{v}_k\}$, as well as the initial state $\mathbf{x}_0$ are random vectors with arbitrary distributions but at least known covariance and cross-covariance matrices. The process and the measurement noise sequences have zero-mean values and the initial state has a known mean value. The objective is to estimate $\mathbf{x}_k$ based on the measurements and our knowledge of the model dynamics. If the estimate of $\mathbf{x}_k$ is based on measurements up to and including time l, we denote the estimator as $\widehat{\mathbf{x}}_{k|l} \triangleq \widehat{\mathbf{x}}_{k|l}(\overline{\mathbf{y}}_l)$ where $\overline{\mathbf{y}}_l^T = (\mathbf{y}_1^T, \ldots, \mathbf{y}_l^T)$, and we use the term estimator to refer to the class of algorithms that includes filtering, prediction, and smoothing. A filter estimates $\mathbf{x}_k$ based on measurements up to and including time k. A predictor estimates $\mathbf{x}_k$ based on measurements prior to time k. A smoother estimates $\mathbf{x}_k$ based on measurements prior to time k, at time k, and later than time k.

Since the seminal paper of Kalman [1], it is known that, if $\{\mathbf{w}_k\}$, $\{\mathbf{v}_k\}$ and $\mathbf{x}_0$ verify the following uncorrelation conditions [1][2][3]:

$$\mathbf{C}_{\mathbf{x}_0,\mathbf{w}_k} = \mathbf{0},\ \mathbf{C}_{\mathbf{x}_0,\mathbf{v}_k} = \mathbf{0},\ \mathbf{C}_{\mathbf{w}_l,\mathbf{v}_k} = \mathbf{0},$$
$$\mathbf{C}_{\mathbf{v}_l,\mathbf{v}_k} = \mathbf{C}_{\mathbf{v}_k}\delta_k^l,\ \mathbf{C}_{\mathbf{w}_l,\mathbf{w}_k} = \mathbf{C}_{\mathbf{w}_k}\delta_k^l, \quad (2)$$

and are Gaussian, the minimum mean squared error (MSE) filter of $\mathbf{x}_k$ (1a) has a recursive predictor/corrector format:

$$\widehat{\mathbf{x}}_{k|k}^b = \mathbf{F}_{k-1}\widehat{\mathbf{x}}_{k-1|k-1}^b + \mathbf{K}_k^b\left(\mathbf{y}_k - \mathbf{H}_k\mathbf{F}_{k-1}\widehat{\mathbf{x}}_{k-1|k-1}^b\right), \quad (3)$$

$\forall k \geq 1$, so-called the Kalman filter (KF)[2]. Even if the noises are non-Gaussian, the KF is the linear least mean squares (LLMS) filter of $\mathbf{x}_k$. As the computation of the KF depends on prior information on $E[\mathbf{x}_0]$ and on $\mathbf{C}_{\mathbf{x}_0}$ [2][3], the KF can be looked upon as an "initial state first and second order statistics" matched filter [4]. However in numerous applications $E[\mathbf{x}_0]$ and/or $\mathbf{C}_{\mathbf{x}_0}$ is unknown. A commonly used solution to circumvent this lack of prior information is the Fisher initialization [5][6, §II]. The Fisher initialization consists in initializing the KF recursion at time $k = 1$ with the best linear unbiased estimator (BLUE) of $\mathbf{x}_1$ associated to the measurement model (1b), where $\mathbf{x}_1$ is regarded as a deterministic unknown parameter vector. In the deterministic framework, the BLUE of $\mathbf{x}_1$ is also known as the linear minimum variance distortionless response (LMVDR) estimator of $\mathbf{x}_1$ [7, §5.6][8]. If $\mathbf{H}_1$ and $\mathbf{C}_{\mathbf{v}_1}$ have full rank, the Fisher initialization yields:

$$\widehat{\mathbf{x}}_{1|1}^b = \mathbf{P}_{1|1}^b\mathbf{H}_1^H\mathbf{C}_{\mathbf{v}_1}^{-1}\mathbf{y}_1,\ \mathbf{P}_{1|1}^b = \left(\mathbf{H}_1^H\mathbf{C}_{\mathbf{v}_1}^{-1}\mathbf{H}_1\right)^{-1}. \quad (4)$$

A particularly noteworthy feature of this alternative initialization of the KF (4) is that it may yield the stochastic LMVDR filter, which shares the same recursion as the KF, except at

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

[1]Hereinafter, scalars, vectors and matrices are represented, respectively, by italic, bold lowercase and bold uppercase characters. $[\mathbf{A}\ \mathbf{B}]$ and $\left[\begin{smallmatrix}\mathbf{A}\\ \mathbf{B}\end{smallmatrix}\right]$ denote the matrix resulting from the horizontal and the vertical concatenation of $\mathbf{A}$ and $\mathbf{B}$, respectively. The matrix resulting from the vertical concatenation of k matrices $\mathbf{A}_1$, ... , $\mathbf{A}_k$ of same column number is denoted $\overline{\mathbf{A}}_k$. $E[\cdot]$ denotes the expectation operator. If $\mathbf{x}$ and $\mathbf{y}$ are two complex random vectors: a) $\mathbf{C}_\mathbf{x}$, $\mathbf{C}_\mathbf{y}$ and $\mathbf{C}_{\mathbf{x},\mathbf{y}}$ are respectively the covariance matrices of $\mathbf{x}$, of $\mathbf{y}$ and the cross-covariance matrix of $\mathbf{x}$ and $\mathbf{y}$, b) $\mathbf{C}_{\mathbf{x}|\mathbf{y}} \triangleq \mathbf{C}_\mathbf{x} - \mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_\mathbf{y}^{-1}\mathbf{C}_{\mathbf{x},\mathbf{y}}^H$, provided that $\mathbf{C}_\mathbf{y}$ is invertible.

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

time $k = 1$. Indeed, authors in [4] have lately shown that, under mild regularity conditions on the noises covariance matrices, this property holds for the restricted subset of LDSS models for which $\mathbf{F}_k$, $k \geq 1$, are invertible. Unfortunately, this restricted subset of LDSS models does not include fixed-point and fixed-lag smoothers (and possibly other smoothers or predictors) which are obtained by running the KF on modified LDSS models [2, §9] incorporating at least one non invertible state matrix.

As shown in the present paper, the invertibility of $\mathbf{F}_k$, $k \geq 1$, is actually not required, provided that $\mathbf{H}_1$ has full rank, which is a substantial extension of [4]. The proposed extension is based on a different approach than the one previously used in [4] which allows to derive the key results in an easily comprehensible manner. An immediate benefit is the introduction of LMVDR fixed-point and fixed-lag smoothers (see §III.C, and possibly other smoothers or predictors, which is left for future research). This result is particularly noteworthy given the fact that, although LMVDR estimators are sub-optimal in MSE sense, they are infinite impulse response distortionless estimators which do not depend on the prior knowledge on $E[\mathbf{x}_0]$ and $\mathbf{C}_{\mathbf{x}_0}$. Thus the LMVDR estimators may outperform the usual LLMS estimators in case of misspecification of the prior knowledge on $\mathbf{x}_0$ (see §III-A). Seen from this perspective, we also show that the LMVDR filter can be regarded as a generalization of the information filter form of the KF (see §III.B), since the LMVDR filter exists under more general assumptions. Last but not least, the interest of the LMVDR estimators can go beyond the property of being robust to uncertainties on the initial state as highlighted in §III.D.

A. Signal model

As in [4][7, §5.1], we adopt a joint proper (proper and cross-proper) complex signals assumption for the set of vectors $(\mathbf{x}_0, \{\mathbf{w}_k\}, \{\mathbf{v}_k\})$ which allows to resort to standard estimation in the 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 [7]. Moreover, 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 [7, §5.4.1]. First, as (1a) can be rewritten as, $\forall k \geq 2$:

$$\mathbf{x}_k = \mathbf{B}_{k,1}\mathbf{x}_1 + \mathbf{G}_k\overline{\mathbf{w}}_{k-1},\ \mathbf{G}_k\overline{\mathbf{w}}_{k-1} = \sum_{l=1}^{k-1} \mathbf{B}_{k,l+1}\mathbf{w}_l, \quad (5)$$

$$\mathbf{G}_k \in \mathcal{M}_{\mathbb{C}}(P_k, \mathcal{P}_{k-1}),\ \mathbf{B}_{k,l} = \left| \begin{array}{cl} \mathbf{F}_{k-1}\mathbf{F}_{k-2}...\mathbf{F}_l, & k > l \\ \mathbf{I} & ,k = l \\ \mathbf{0} & ,k < l \end{array} \right. ,$$

where $\mathcal{P}_k = \sum_{l=1}^{k} P_l$, an equivalent form of (1b) is:

$$\mathbf{y}_k = \mathbf{A}_k\mathbf{x}_1 + \mathbf{n}_k,\ \mathbf{A}_k = \mathbf{H}_k\mathbf{B}_{k,1}, \quad \left| \begin{array}{l} \mathbf{n}_1 = \mathbf{v}_1 \\ \mathbf{n}_k = \mathbf{v}_k + \mathbf{H}_k\mathbf{G}_k\overline{\mathbf{w}}_{k-1},\ k \geq 2 \end{array} \right. . \quad (6a)$$

Second, (1b) can be extended on a horizon of k points from the first observation as:

$$\overline{\mathbf{y}}_k = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_k \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_k \end{bmatrix} \mathbf{x}_1 + \begin{bmatrix} \mathbf{n}_1 \\ \vdots \\ \mathbf{n}_k \end{bmatrix} = \overline{\mathbf{A}}_k\mathbf{x}_1 + \overline{\mathbf{n}}_k, \quad (6b)$$

$\overline{\mathbf{y}}_k, \overline{\mathbf{n}}_k \in \mathbb{C}^{\mathcal{N}_k}$, $\overline{\mathbf{A}}_k \in \mathbb{C}^{\mathcal{N}_k \times P_k}$, $\mathcal{N}_k = \sum_{l=1}^{k} N_l$.

II. LMVDR filter for LDSS Models

In this section, we consider a completely different approach than the one previously used in [4]. Indeed, we provide a general definition of a distortionless filter in the context of LDSS models (7), which encompasses the definition used in [4]. And, it is the combination of this general definition with the Joseph stabilized version of the covariance measurement update equation that allows to prove that, whenever it exists, the LMVDR filter shares the same recursion as the KF except at initialization.

A. A general definition of the LMVDR filter

We adopt the notation used in the deterministic framework for the LMVDR filter [7, §5.6][8] to stress the fact that the LMVDR filter is different from the LLMS filter, aka the KF. Indeed, for LDSS models one can define a "state-former" in the same way as a beamformer in array processing or a frequency-bin former in spectral analysis [7, §5.6][8], that is $\overline{\mathbf{W}}_k \in \mathbb{C}^{\mathcal{N}_k \times P_k}$ yielding the state vector $\overline{\mathbf{W}}_k^H \overline{\mathbf{y}}_k$, which can be recasted as (6b):

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

Therefore, according to (5), a filter $\overline{\mathbf{W}}_k$ is distortionless iff:

$$\overline{\mathbf{W}}_k^H \overline{\mathbf{y}}_k = \mathbf{x}_k + \overline{\mathbf{W}}_k^H \overline{\mathbf{n}}_k - \mathbf{G}_k\overline{\mathbf{w}}_{k-1} \Leftrightarrow \overline{\mathbf{W}}_k^H \overline{\mathbf{A}}_k = \mathbf{B}_{k,1}. \quad (7)$$

If $\mathbf{H}_1$ is full rank, the set of distortionless state-formers is non empty. Indeed, if $\mathbf{W}_1 = \mathbf{H}_1 \left(\mathbf{H}_1^H \mathbf{H}_1 \right)^{-1}$, then:

$$\left(\mathbf{W}_1 \mathbf{B}_{k,1}^H \right)^H \mathbf{y}_1 = \mathbf{x}_k + \left(\mathbf{W}_1 \mathbf{B}_{k,1}^H \right)^H \mathbf{v}_1 - \mathbf{G}_k\overline{\mathbf{w}}_{k-1},$$

is a distortionless state-former. Therefore, there exists a best distortionless state-former in the MSE sense, aka the LMVDR filter, which minimizes the error covariance matrix

$$\mathbf{P}_{k|k}\left(\overline{\mathbf{W}}_k \right) = E\left[\left(\overline{\mathbf{W}}_k^H \overline{\mathbf{y}}_k - \mathbf{x}_k \right) \left(\overline{\mathbf{W}}_k^H \overline{\mathbf{y}}_k - \mathbf{x}_k \right)^H \right] \quad (8)$$

w.r.t. the Löwner ordering [9] under linear constraints (7):

$$\overline{\mathbf{W}}_k^b = \arg\min_{\overline{\mathbf{W}}_k} \left\{ \mathbf{P}_{k|k}\left(\overline{\mathbf{W}}_k \right) \right\} \text{ s.t. } \overline{\mathbf{W}}_k^H \overline{\mathbf{A}}_k = \mathbf{B}_{k,1}. \quad (9)$$

Note that (9) is equivalent to:

$$\overline{\mathbf{W}}_k^b = \arg\min_{\overline{\mathbf{W}}_k} \left\{ E\left[\widehat{\mathbf{r}}_k \left(\widehat{\mathbf{r}}_k \right)^H \right] \right\} \text{ s.t. } \overline{\mathbf{W}}_k^H \overline{\mathbf{A}}_k = \mathbf{B}_{k,1},$$

$$\widehat{\mathbf{r}}_k = \overline{\mathbf{W}}_k^H \overline{\mathbf{n}}_k - \mathbf{G}_k\overline{\mathbf{w}}_{k-1}. \quad (10)$$

Since the KF is the solution of the following unconstrained minimization problem [1][2][3]:

$$\overline{\mathbb{K}}_k^b = \arg\min_{\overline{\mathbf{W}}_k} \left\{ \mathbf{P}_{k|k}\left(\overline{\mathbf{W}}_k\right) \right\}, \tag{11}$$

it follows that the LMVDR filter (9) is sub-optimal in MSE sense in comparison with the KF (11). Let us recall that, if one considers the following breakdown of $\overline{\mathbf{W}}_k$:

$$\overline{\mathbf{W}}_k = \begin{bmatrix} \overline{\mathbf{D}}_{k-1} \\ \mathbf{W}_k \end{bmatrix}, \quad \begin{cases} \overline{\mathbf{D}}_{k-1} \in \mathbb{C}^{\mathcal{N}_{k-1} \times P_k} \\ \mathbf{W}_k \in \mathbb{C}^{N_k \times P_k} \end{cases}, \tag{12}$$

then the KF $\widehat{\mathbf{x}}_{k|k}^b = \overline{\mathbb{K}}_k^{bH} \overline{\mathbf{y}}_k$, $\overline{\mathbb{K}}_k^b = \left[\begin{smallmatrix} \overline{\mathbb{D}}_{k-1}^b \\ \mathbb{K}_k^b \end{smallmatrix}\right]$ (11), can be recasted in the convenient recursive predictor/corrector form (3) where $\mathbf{K}_k^b = \left(\mathbb{K}_k^b\right)^H$.

B. LMVDR filter for LDSS Models

It is known that, under assumptions (2), the error covariance matrix (8) of any linear filter $\widehat{\mathbf{x}}_{k|k} = \overline{\mathbf{W}}_k \overline{\mathbf{y}}_k$ satisfies the Joseph stabilized version of the covariance measurement update equation [1][2][3]:

$$\begin{aligned} \mathbf{P}_{k|k}\left(\overline{\mathbf{W}}_k\right) &= \mathbf{Q}_{k-1}\left(\overline{\mathbf{D}}_{k-1}, \mathbf{W}_k\right) + \mathbf{W}_k^H \mathbf{C}_{\mathbf{v}_k} \mathbf{W}_k \\ &+ \left(\mathbf{I} - \mathbf{W}_k^H \mathbf{H}_k\right) \mathbf{C}_{\mathbf{w}_{k-1}} \left(\mathbf{I} - \mathbf{H}_k^H \mathbf{W}_k\right), \end{aligned} \tag{13a}$$

where:

$$\begin{aligned} \mathbf{Q}_{k-1}\left(\overline{\mathbf{D}}_{k-1}, \mathbf{W}_k\right) &= E\left[\widehat{\mathbf{q}}_{k-1} \widehat{\mathbf{q}}_{k-1}^H\right], \\ \widehat{\mathbf{q}}_{k-1} &= \overline{\mathbf{D}}_{k-1}^H \overline{\mathbf{y}}_{k-1} - \left(\mathbf{I} - \mathbf{W}_k^H \mathbf{H}_k\right) \mathbf{F}_{k-1} \mathbf{x}_{k-1}. \end{aligned} \tag{13b}$$

As shown hereinafter, the covariance update equation (13a) allows to breakdown the initial constrained minimization problem (9) into two separable minimization problems: a first constrained minimization problem w.r.t. $\overline{\mathbf{D}}_{k-1}$, namely:

$$\overline{\mathbf{D}}_{k-1}^b = \arg\min_{\overline{\mathbf{D}}_{k-1}} \left\{ \mathbf{Q}_{k-1}\left(\overline{\mathbf{D}}_{k-1}, \mathbf{W}_k\right) \right\} \text{ s.t. } \overline{\mathbf{W}}_k^H \overline{\mathbf{A}}_k = \mathbf{B}_{k,1}, \tag{14a}$$

where $\overline{\mathbf{D}}_{k-1}^b \triangleq \overline{\mathbf{D}}_{k-1}^b\left(\mathbf{W}_k\right)$, followed by a second unconstrained minimization problem w.r.t. $\mathbf{W}_k$, namely:

$$\mathbf{W}_k^b = \arg\min_{\mathbf{W}_k} \left\{ \mathbf{P}_{k|k}\left(\overline{\mathbf{D}}_{k-1}^b, \mathbf{W}_k\right) \right\}, \tag{14b}$$

where $\mathbf{P}_{k|k}\left(\overline{\mathbf{D}}_{k-1}, \mathbf{W}_k\right) \triangleq \mathbf{P}_{k|k}\left(\overline{\mathbf{W}}_k\right)$.

• Solution of (14a)

Firstly, let us notice that $\widehat{\mathbf{q}}_{k-1}$ in (13b) can be recasted as:

$$\begin{aligned} \widehat{\mathbf{q}}_{k-1} &= \overline{\mathbf{D}}_{k-1}^H \overline{\mathbf{A}}_{k-1} \mathbf{x}_1 - \left(\mathbf{I} - \mathbf{W}_k^H \mathbf{H}_k\right) \mathbf{F}_{k-1} \mathbf{x}_{k-1} \\ &+ \overline{\mathbf{D}}_{k-1}^H \overline{\mathbf{n}}_{k-1}. \end{aligned} \tag{15}$$

Secondly, since $\overline{\mathbf{W}}_k^H \overline{\mathbf{A}}_k = \overline{\mathbf{D}}_{k-1}^H \overline{\mathbf{A}}_{k-1} + \mathbf{W}_k^H \mathbf{A}_k$ where $\mathbf{A}_k = \mathbf{H}_k \mathbf{B}_{k,1} = \mathbf{H}_k \mathbf{F}_{k-1} \mathbf{B}_{k-1,1}$, then if $\overline{\mathbf{W}}_k$ is a distortionless filter, it verifies (7):

$$\begin{aligned} &\overline{\mathbf{W}}_k^H \overline{\mathbf{A}}_k = \mathbf{B}_{k,1} \Leftrightarrow \\ &\quad \overline{\mathbf{D}}_{k-1}^H \overline{\mathbf{A}}_{k-1} = \left(\mathbf{I} - \mathbf{W}_k^H \mathbf{H}_k\right) \mathbf{F}_{k-1} \mathbf{B}_{k-1,1}, \end{aligned}$$

leading to:

$$\begin{aligned} \widehat{\mathbf{q}}_{k-1} &= \left(\mathbf{I} - \mathbf{W}_k^H \mathbf{H}_k\right) \mathbf{F}_{k-1} \left(\mathbf{B}_{k-1,1} \mathbf{x}_1 - \mathbf{x}_{k-1}\right) \\ &+ \overline{\mathbf{D}}_{k-1}^H \overline{\mathbf{n}}_{k-1}. \end{aligned} \tag{16}$$

Thirdly:

$$\begin{aligned} &\mathbf{x}_{k-1} = \mathbf{B}_{k-1,1} \mathbf{x}_1 + \mathbf{G}_{k-1} \overline{\mathbf{w}}_{k-2} \Leftrightarrow \\ &\quad \mathbf{B}_{k-1,1} \mathbf{x}_1 - \mathbf{x}_{k-1} = -\mathbf{G}_{k-1} \overline{\mathbf{w}}_{k-2}. \end{aligned}$$

Finally, for any distortionless state-former $\overline{\mathbf{W}}_k$:

$$\widehat{\mathbf{q}}_{k-1} = \overline{\mathbf{D}}_{k-1}^H \overline{\mathbf{n}}_{k-1} - \left(\mathbf{I} - \mathbf{W}_k^H \mathbf{H}_k\right) \mathbf{F}_{k-1} \mathbf{G}_{k-1} \overline{\mathbf{w}}_{k-2}, \tag{17a}$$

leading to the following equivalent form of (14a):

$$\begin{aligned} \overline{\mathbf{D}}_{k-1}^b &= \arg\min_{\overline{\mathbf{D}}_{k-1}} \left\{ E\left[\widehat{\mathbf{q}}_{k-1} \widehat{\mathbf{q}}_{k-1}^H\right] \right\} \\ &\text{s.t. } \overline{\mathbf{D}}_{k-1}^H \overline{\mathbf{A}}_{k-1} = \left(\mathbf{I} - \mathbf{W}_k^H \mathbf{H}_k\right) \mathbf{F}_{k-1} \mathbf{B}_{k-1,1}. \end{aligned} \tag{17b}$$

In the form of (17b), $\overline{\mathbf{D}}_{k-1}^b$ is analogous to a linearly constrained Wiener filter [10, §2.5].
If $\mathbf{H}_1$ has full rank, since $\mathbf{A}_1 = \mathbf{H}_1$ in (6b), then $\overline{\mathbf{A}}_k$, $k \geq 1$, has full rank as well. Therefore, if $\mathbf{C}_{\overline{\mathbf{n}}_{k-1}}$ is invertible, then $\overline{\mathbf{D}}_{k-1}^b$ can be computed as [10, (2.113)]:

$$\overline{\mathbf{D}}_{k-1}^b = \overline{\mathbf{W}}_{k-1}^b \left(\left(\mathbf{I} - \mathbf{W}_k^H \mathbf{H}_k\right) \mathbf{F}_{k-1}\right)^H, \tag{18a}$$

$$\begin{aligned} \overline{\mathbf{W}}_{k-1}^b &= \arg\min_{\overline{\mathbf{W}}_{k-1}} \left\{ E\left[\widehat{\mathbf{r}}_{k-1} \widehat{\mathbf{r}}_{k-1}^H\right] \right\} \\ &\text{s.t. } \overline{\mathbf{W}}_{k-1}^H \overline{\mathbf{A}}_{k-1} = \mathbf{B}_{k-1,1}, \end{aligned} \tag{18b}$$

where $\overline{\mathbf{W}}_{k-1}^b$ is the LMVDR filter at time $k-1$, since, from (9-10), (18b) is equivalent to:

$$\begin{aligned} \overline{\mathbf{W}}_{k-1}^b &= \arg\min_{\overline{\mathbf{W}}_{k-1}} \left\{ \mathbf{P}_{k-1|k-1}\left(\overline{\mathbf{W}}_{k-1}\right) \right\} \\ &\text{s.t. } \overline{\mathbf{W}}_{k-1}^H \overline{\mathbf{A}}_{k-1} = \mathbf{B}_{k-1,1}. \end{aligned}$$

Finally, $\forall k \geq 2$:

$$\begin{aligned} \mathbf{Q}_{k-1}\left(\overline{\mathbf{D}}_{k-1}^b, \mathbf{W}_k\right) &= \left(\mathbf{I} - \mathbf{W}_k^H \mathbf{H}_k\right) \mathbf{F}_{k-1} \times \\ &\mathbf{P}_{k-1|k-1}\left(\overline{\mathbf{W}}_{k-1}^b\right) \mathbf{F}_{k-1}^H \left(\mathbf{I} - \mathbf{H}_k^H \mathbf{W}_k\right). \end{aligned} \tag{19}$$

• Solution of (14b)

According to (19), the solution $\overline{\mathbf{D}}_{k-1}^b \triangleq \overline{\mathbf{D}}_{k-1}^b\left(\mathbf{W}_k\right)$ (18a) of the first constrained minimization problem (14a) leads to:

$$\begin{aligned} \mathbf{P}_{k|k}\left(\overline{\mathbf{D}}_{k-1}^b, \mathbf{W}_k\right) &= \left(\mathbf{I} - \mathbf{W}_k^H \mathbf{H}_k\right) \mathbf{P}_{k|k-1}^b \left(\mathbf{I} - \mathbf{H}_k^H \mathbf{W}_k\right) \\ &+ \mathbf{W}_k^H \mathbf{C}_{\mathbf{v}_k} \mathbf{W}_k, \end{aligned} \tag{20a}$$

$$\mathbf{P}_{k|k-1}^b = \mathbf{F}_{k-1} \mathbf{P}_{k-1|k-1}^b \mathbf{F}_{k-1}^H + \mathbf{C}_{\mathbf{w}_{k-1}} \tag{20b}$$

Thus, the solution $\mathbf{W}_k^b$ of the minimization of (20a), that is:

$$\mathbf{W}_k^b = \arg\min_{\mathbf{W}_k} \left\{ \mathbf{P}_{k|k}\left(\overline{\mathbf{D}}_{k-1}^b\left(\mathbf{W}_k\right), \mathbf{W}_k\right) \right\}, \tag{21a}$$

can be computed according to the following recursion for $k \geq 2$ [1][2][3]:

$$\mathbf{P}^{b}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{P}^{b}_{k-1|k-1}\mathbf{F}^{H}_{k-1} + \mathbf{C}_{\mathbf{w}_{k-1}} \tag{21b}$$

$$\mathbf{S}^{b}_{k|k-1} = \mathbf{H}_{k}\mathbf{P}^{b}_{k|k-1}\mathbf{H}^{H}_{k} + \mathbf{C}_{\mathbf{v}_{k}} \tag{21c}$$

$$\mathbf{W}^{b}_{k} = \left(\mathbf{S}^{b}_{k|k-1}\right)^{-1}\mathbf{H}_{k}\mathbf{P}^{b}_{k|k-1}, \tag{21d}$$

$$\mathbf{P}^{b}_{k|k} = \left(\mathbf{I} - \mathbf{W}^{bH}_{k}\mathbf{H}_{k}\right)\mathbf{P}^{b}_{k|k-1} \tag{21e}$$

- **Summary**

For $k \geq 2$, according to (18a) and (21a), the LMVDR filter (9) yields the state-former:

$$\begin{aligned}
\widehat{\mathbf{x}}^{b}_{k|k} &= \overline{\mathbf{W}}^{bH}_{k}\overline{\mathbf{y}}_{k} = \overline{\mathbf{D}}^{bH}_{k-1}\overline{\mathbf{y}}_{k-1} + \mathbf{W}^{bH}_{k}\mathbf{y}_{k} \\
&= \left(\mathbf{I} - \mathbf{W}^{bH}_{k}\mathbf{H}_{k}\right)\mathbf{F}_{k-1}\left(\overline{\mathbf{W}}^{bH}_{k-1}\overline{\mathbf{y}}_{k-1}\right) + \mathbf{W}^{bH}_{k}\mathbf{y}_{k} \\
\widehat{\mathbf{x}}^{b}_{k|k} &= \mathbf{F}_{k-1}\widehat{\mathbf{x}}^{b}_{k-1|k-1} + \mathbf{W}^{bH}_{k}\left(\mathbf{y}_{k} - \mathbf{H}_{k}\mathbf{F}_{k-1}\widehat{\mathbf{x}}^{b}_{k-1|k-1}\right)
\end{aligned}$$

where $\mathbf{W}^{b}_{k}$ is given by the recursion (21b-21e), similar to the KF recursion [1][2][3]. At time $k = 1$:

$$\mathbf{W}^{b}_{1} = \arg\min_{\mathbf{W}_{1}}\left\{\mathbf{P}_{1|1}\left(\mathbf{W}_{1}\right)\right\} \text{ s.t. } \mathbf{W}^{H}_{1}\mathbf{H}_{1} = \mathbf{I},$$

leading to (4):

$$\widehat{\mathbf{x}}^{b}_{1|1} = \mathbf{P}^{b}_{1|1}\mathbf{H}^{H}_{1}\mathbf{C}^{-1}_{\mathbf{v}_{1}}\mathbf{y}_{1},\ \mathbf{P}^{b}_{1|1} = \left(\mathbf{H}^{H}_{1}\mathbf{C}^{-1}_{\mathbf{v}_{1}}\mathbf{H}_{1}\right)^{-1}. \tag{22}$$

Since $\mathbf{P}^{b}_{k|k} \triangleq \mathbf{P}_{k|k}\left(\overline{\mathbf{W}}^{b}_{k}\right)$ depends on neither $E\left[\mathbf{x}_{0}\right]$ nor $\mathbf{C}_{\mathbf{x}_{0}}$, the LMVDR filter is sub-optimal in MSE sense in comparison with the KF whatever the initial conditions $E\left[\mathbf{x}_{0}\right]$ and $\mathbf{C}_{\mathbf{x}_{0}}$. Thus, the LMVDR filter is an upper bound on the performance of the KF whatever the initial conditions $E\left[\mathbf{x}_{0}\right]$ and $\mathbf{C}_{\mathbf{x}_{0}}$.

III. On the Significance of LMVDR Estimators

A. A filter independent of a priori knowledge on $\mathbf{x}_0$

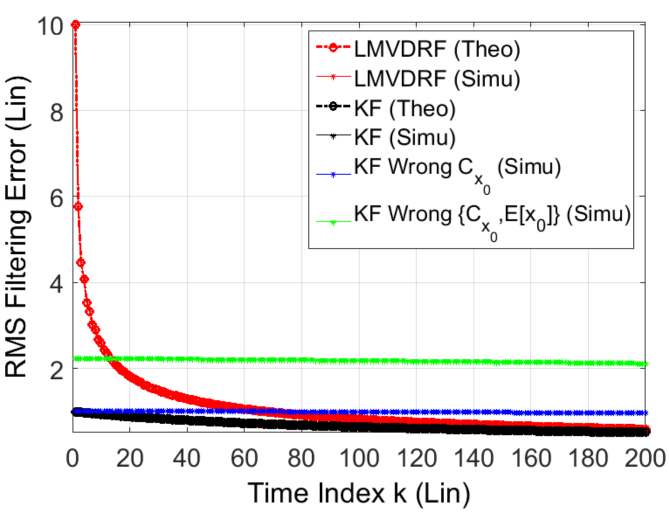

Fig. 1. Comparison of $\sqrt{P^{d}_{k|k}}$ and $\sqrt{P_{k|k}}$.

For the sake of illustration of the key properties of the LMVDR filter, in the general case where $\mathbf{F}_{k}$, $\forall k$, is not invertible, we consider the problem of bias estimation from noisy measurements based on the following simple time variant LDSS model:

$$\left|\begin{array}{l}\mathbf{x}_{2k+1} = \mathbf{F}_{2k}x_{2k} + \mathbf{w}_{2k} \\ \mathbf{y}_{2k+1} = \mathbf{x}_{2k+1} + \mathbf{v}_{2k+1}\end{array}\right., \left|\begin{array}{l}x_{2k+2} = \mathbf{F}_{2k+1}\mathbf{x}_{2k+1} + w_{2k+1} \\ y_{2k+2} = x_{2k+2} + v_{2k+2}\end{array}\right.$$

where $\mathbf{F}_{2k} = \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$ and $\mathbf{F}_{2k+1} = [0.5\ 0.5]$ are not invertible. The noise process, the measurement noise and the initial state x_0 are Gaussian and uncorrelated. Moreover, assume that $E\left[x_0\right] = -2$, $C_{x_0} = 1$, $\mathbf{C}_{\mathbf{w}_{2k}} = \sigma^{2}_{w}\mathbf{I}$, $C_{w_{2k+1}} = \sigma^{2}_{w}$, $\sigma^{2}_{w} = 4/3$, $\mathbf{C}_{\mathbf{v}_{2k+1}} = \sigma^{2}_{v}\mathbf{I}$, $C_{v_{2k}} = \sigma^{2}_{v}$, $\sigma^{2}_{v} = 100$. Fig. 1. highlights the consequence of a misspecification on C_{x_0} and on $\left\{C_{x_0}, E\left[x_0\right]\right\}$ when one initializes the KF with **wrong** assumed values $C_{x_0} = 10^{-2}$ and $\left\{C_{x_0} = 10^{-2}, E\left[x_0\right] = 0\right\}$. The empirical $\sqrt{P^{b}_{k|k}}$ ('... (Simu)') are assessed with 10^4 Monte-Carlo trials.

Fig. 1. clearly shows that, although the LMVDR filter is sub-optimal in terms of MSE when $\left\{\mathbf{C}_{\mathbf{x}_0}, E\left[\mathbf{x}_0\right]\right\}$ are perfectly known, in the presence of uncertainties on $\left\{\mathbf{C}_{\mathbf{x}_0}, E\left[\mathbf{x}_0\right]\right\}$, the LMVDR filter may offer better performance than a KF wrongly initialized.

B. A generalization of the information filter form of the KF

If $\mathbf{C}_{\mathbf{w}_{k-1}}$ and $\mathbf{C}_{\mathbf{v}_{k}}$ are invertible, $k \geq 1$, thus $\mathbf{P}^{b}_{k|k}$ and $\mathbf{P}^{b}_{k|k-1}$ are invertible, which allows to define the information matrices:

$$\mathbf{I}_{k|k} = \left(\mathbf{P}^{b}_{k|k}\right)^{-1}, \quad \mathbf{I}_{k|k-1} = \left(\mathbf{P}^{b}_{k|k-1}\right)^{-1}. \tag{23}$$

Then the usual form of the KF recursion (21b-21e) can be rewritten in the following information filter (IF) form [2, §6.2]:

$$\begin{aligned}
\mathbf{I}_{k|k-1} &= \mathbf{C}^{-1}_{\mathbf{w}_{k-1}} - \\
&\mathbf{C}^{-1}_{\mathbf{w}_{k-1}}\mathbf{F}_{k-1}\left(\mathbf{I}_{k-1|k-1} + \mathbf{F}^{H}_{k-1}\mathbf{C}^{-1}_{\mathbf{w}_{k-1}}\mathbf{F}_{k-1}\right)^{-1}\mathbf{F}^{H}_{k-1}\mathbf{C}^{-1}_{\mathbf{w}_{k-1}},
\end{aligned} \tag{24a}$$

$$\mathbf{I}_{k|k} = \mathbf{I}_{k|k-1} + \mathbf{H}^{H}_{k}\mathbf{C}^{-1}_{\mathbf{v}_{k}}\mathbf{H}_{k}, \tag{24b}$$

$$\mathbf{K}^{b}_{k} = \mathbf{I}^{-1}_{k|k}\mathbf{H}^{H}_{k}\mathbf{C}^{-1}_{\mathbf{v}_{k}}. \tag{24c}$$

If a very broad prior distribution on $\mathbf{x}_0$ is assumed, i.e. in the limit case as $\mathbf{C}_{\mathbf{x}_0} \rightarrow \infty$, then $\mathbf{I}_{0|0} = \mathbf{C}^{-1}_{\mathbf{x}_0} \rightarrow \mathbf{0}$, leading to:

$$\mathbf{I}_{1|0} = \mathbf{C}^{-1}_{\mathbf{w}_0}\left(\mathbf{I} - \mathbf{F}_{0}\left(\mathbf{F}^{H}_{0}\mathbf{C}^{-1}_{\mathbf{w}_0}\mathbf{F}_{0}\right)^{-1}\mathbf{F}^{H}_{0}\mathbf{C}^{-1}_{\mathbf{w}_0}\right).$$

Moreover, if $\mathbf{F}_0$ is invertible then $\mathbf{I}_{1|0} = \mathbf{0}$ and:

$$\mathbf{I}_{1|1} = \mathbf{H}^{H}_{1}\mathbf{C}^{-1}_{\mathbf{v}_{1}}\mathbf{H}_{1},\ \mathbf{K}^{b}_{1} = \left(\mathbf{H}^{H}_{1}\mathbf{C}^{-1}_{\mathbf{v}_{1}}\mathbf{H}_{1}\right)^{-1}\mathbf{H}^{H}_{1}\mathbf{C}^{-1}_{\mathbf{v}_{1}}.$$

Thus, if $\mathbf{F}_0$ is invertible, the use of a prior-free estimate of $\mathbf{x}_1$, obtained via the IF form (24a-24c) coincides with the LMVDR filter. In comparison with the IF form, the LMVDR filter exists under more general assumptions since:

a) the invertibility of $\mathbf{C}_{\mathbf{w}_{k-1}}$ and $\mathbf{C}_{\mathbf{v}_{k}}$ is not required,

b) and the knowledge of $\mathbf{F}_0$ and $\mathbf{C}_{\mathbf{w}_0}$ is not required either.

C. LMVDR fixed-point and fixed-lag smoothers

The standard fixed-point smoother $\widehat{\mathbf{x}}_{l|k}^{b}$ is obtained by running the KF on the following LDSS models [2, §9.2]:

$$k \leq l \quad \begin{cases} \mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{w}_{k-1} \\ \mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \end{cases}$$

$$k = l+1 \begin{cases} \begin{pmatrix} \mathbf{x}_{l+1} \\ \varkappa_{l+1} \end{pmatrix} = \begin{bmatrix} \mathbf{F}_l \\ \mathbf{I} \end{bmatrix} \mathbf{x}_l + \begin{pmatrix} \mathbf{w}_l \\ \mathbf{0} \end{pmatrix} \\ \mathbf{y}_l = \begin{bmatrix} \mathbf{H}_l & \mathbf{0} \end{bmatrix} \begin{pmatrix} \mathbf{x}_l \\ \varkappa_l \end{pmatrix} + \mathbf{v}_l \end{cases}$$

$$k \geq l+2 \begin{cases} \begin{pmatrix} \mathbf{x}_k \\ \varkappa_k \end{pmatrix} = \begin{bmatrix} \mathbf{F}_{k-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{pmatrix} \mathbf{x}_{k-1} \\ \varkappa_{k-1} \end{pmatrix} + \begin{pmatrix} \mathbf{w}_{k-1} \\ \mathbf{0} \end{pmatrix} \\ \mathbf{y}_k = \begin{bmatrix} \mathbf{H}_k & \mathbf{0} \end{bmatrix} \begin{pmatrix} \mathbf{x}_k \\ \varkappa_k \end{pmatrix} + \mathbf{v}_k \end{cases}$$

leading to $\widehat{\mathbf{x}}_{l|k}^{b} = \widehat{\varkappa}_{k|k}^{b}$ for $k \geq l+1$. Obviously at time $l+1$ the state matrix of the augmented state, i.e. $\left[\begin{smallmatrix} \mathbf{F}_l \\ \mathbf{I} \end{smallmatrix}\right]$, is always non invertible, whatever $\mathbf{F}_l$ is invertible or not. Likewise, the standard fixed-lag smoother $\widehat{\mathbf{x}}_{k-N|k}^{b}$ [2, §9.3] is obtained by running the KF on an augmented system which state matrix is always non invertible [2, (9.41)]. As a consequence, a major benefit of the relaxation on the conditions of existence of LMVDR filter introduced here, is the proof of the existence of the LMVDR fixed-point and fixed-lag smoothers, which can not be proved with the result introduced in [4].

D. Recursive form of the generalized conditional maximum likelihood estimators (GCMLEs)

The interest of LMVDR estimators can go beyond the property of being robust to uncertainties on $\{\mathbf{C}_{\mathbf{x}_0}, E[\mathbf{x}_0]\}$. For instance, as shown in [12], the LMVDR fixed-point smoother $\widehat{\mathbf{x}}_{1|k}^{b}$ can be used to compute the GCMLEs of the unknown deterministic parameters $\mathbf{x}_1$ and $\boldsymbol{\theta}$ associated with the observation model:

$$\mathbf{x}_l = \mathbf{F}_{l-1}\mathbf{x}_{l-1} + \mathbf{w}_{l-1}, \quad 2 \leq l \leq k, \tag{26a}$$

$$\mathbf{y}_l = \mathbf{H}_l(\boldsymbol{\theta})\mathbf{x}_l + \mathbf{v}_l, \qquad 1 \leq l \leq k, \tag{26b}$$

where the Gaussian fluctuation noise sequence $\{\mathbf{w}_l\}_{l=1}^{k-1}$ is white and uncorrelated with the Gaussian white measurement noise sequence $\{\mathbf{v}_l\}_{l=1}^{k}$. The above generalized conditional signal model (GCSM) arises when k independent conditional signal models (CSMs) [11] involving $\mathbf{x}_1$ are available (26b) and where the signals $\mathbf{x}_1$ are allowed to perform a Gaussian random walk between observations (26a). In the GCSM, since (5-6b) lead to $\overline{\mathbf{y}}_k \sim \mathcal{CN}\left(\overline{\mathbf{A}}_k(\boldsymbol{\theta})\mathbf{x}_1, \mathbf{C}_{\overline{\mathbf{n}}_k}(\boldsymbol{\theta})\right)$, the parameters $\boldsymbol{\theta}$ are connected with both the expectation value and the covariance matrix, which is a significant change in comparison with the usual CSM where the parameters $\boldsymbol{\theta}$ are connected with expectation value only. Therefore, the computation of the GCMLE $\widehat{\boldsymbol{\theta}}(k)$ of $\boldsymbol{\theta}$ based on $\overline{\mathbf{y}}_k$ requires the computation of $\mathbf{C}_{\overline{\mathbf{n}}_k}^{-1}(\boldsymbol{\theta})$ and $\left|\mathbf{C}_{\overline{\mathbf{n}}_k(\boldsymbol{\theta})}\right|$, where $\mathbf{C}_{\overline{\mathbf{n}}_k(\boldsymbol{\theta})}$ is not block diagonal (except if $\mathbf{C}_{\mathbf{w}_l} = \mathbf{0}$, $1 \leq l \leq k-1$), which could become computationally prohibitive as the number of observations k increases. Fortunately, it can be shown [12] that the GCMLE $\widehat{\mathbf{x}}_1(k)$ of $\mathbf{x}_1$ coincides with the LMVDR smoother $\widehat{\mathbf{x}}_{1|k}^{b}$ of $\mathbf{x}_1$, which allows to compute both $\widehat{\mathbf{x}}_1(k)$ and $\widehat{\boldsymbol{\theta}}(k)$ recursively from observation to observation, without the need to compute at each new observation $\mathbf{C}_{\overline{\mathbf{n}}_k}^{-1}(\boldsymbol{\theta})$ nor $\left|\mathbf{C}_{\overline{\mathbf{n}}_k(\boldsymbol{\theta})}\right|$. Note that this result can not be obtained from the standard fixed-point smoother [2, §9.2] based on the LLMS filter, aka the KF.

IV. Conclusion

By relaxing the conditions of existence of LMVDR filters, the existence of the LMVDR fixed-point and fixed-lag smoothers has been proved (and possibly of other smoothers or predictors, which is left for future research). This result is quite interesting for filter/smoother/predictor performance analysis and design since it allows to synthesize IIR distortionless estimators which performance are robust to an unknown initial state $\mathbf{x}_0$. In a broader perspective, let us recall that the robustification of the KF to the presence of mismodeling has been reinvestigated lately by using unbiased finite impulse response (UFIR) [13], p-shift FIR [14] or minimum variance UFIR [15] filters. These algorithms have the same predictor/corrector format as the KF, often ignore initial estimations errors and the statistics of the noise, and become virtually optimal as the length of the FIR window increases. All in all, the LMVDR filter is not the best filter in terms of MSE, neither the more robust, but its performance can be assessed in advance and it can be pre-computed. On another note, LMVDR estimators may also allow to derive unexpected results, as highlighted with the LMVDR fixed-point smoother (see §III.D) [12].

References

[1] R. Kalman, "A new approach to linear filtering and prediction problems", ASME Journal of Basic Engineering, 82: 35-45, 1960

[2] D. Simon, *Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches*, Wiley InterScience, 2006

[3] J. L. Crassidis and J. L. Junkins, *Optimal Estimation of Dynamic Systems (2Ed)*, CRC Press, Taylor & Francis Group, 2012

[4] 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

[5] Catlin, D. E., "Estimation of random states in general linear models", IEEE Trans. on AC, 36(2): 248-252, 1991

[6] G. Chen, *Approximate Kalman Filtering*, World Scientific, Singapore, 1993

[7] P.J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data*, Cambridge University Press 2010

[8] S. A.Vorobyov, "Principles of minimum variance robust adaptive beamforming design, Elsevier Signal Processing", 93: 3264-3277, 2013

[9] R.A. Horn, C.R. Johnson, *Matrix Analysis (2Ed)*. Cambridge University Press, 2013

[10] Paulo S.R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation (4Ed)*, Springer Science+Business Media, 2013

[11] P. Stoica and A. Nehorai, "Performances study of conditional and unconditional direction of arrival estimation," IEEE Trans. on SP, 38(10): 1783-1795, 1990.

[12] E. Chaumette, F. Vincent, A. Renaux and J. Galy, "Generalized Conditional Maximum Likelihood Estimators in the Large Sample Regime", in Proc. EUSIPCO 2018

[13] Y. S. Shmaliy, "Linear optimal FIR estimation of discrete time-invariant state-space models", IEEE Trans. SP, 58(6): 3086-2010, 2010

[14] Y. S. Shmaliy, "An Iterative Kalman-Like Algorithm Ignoring Noise and Initial Conditions", IEEE Trans. SP, 59(6): 2465-2473, 2011

[15] S. Zhao, Y. S. Shmaliy, B. Huang, F. Liu, "Minimum variance unbiased FIR filter for discrete time-variant systems", Automatica, 53: 355-361, 2015