On LMVDR Estimators for LDSS Models: Conditions for Existence and Further Applications

Eric Chaumette, François Vincent, Benoit Priot, Gael Pages, and Arnaud Dion

***Abstract*—For linear discrete state-space models, under certain conditions, the linear least mean squares (LLMS) filter estimate has a recursive format, a.k.a. the Kalman filter (KF). Interestingly, the linear minimum variance distortionless response (LMVDR) filter, when it exists, shares exactly the same recursion as the KF, except for the initialization. If LMVDR estimators are suboptimal in mean-squared error sense, they do not depend on the prior knowledge on 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. In this perspective, we establish the general conditions under which existence of the LMVDRF is guaranteed. 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. Indeed, the LMVDR fixed-point smoother can be used to compute recursively the solution of a generalization of the deterministic least-squares problem.**

***Index Terms*—Filtering, minimum mean-squared error (MSE) upper bound, smoothing, state estimation, unbiased filter.**

I. Introduction

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

$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{w}_{k-1} \tag{1a}$$

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \tag{1b}$$

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. 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 known covariance and cross-covariance matrices. The process and the measurement noise sequences have zero-mean values. 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 $\widehat{\mathbf{x}}_{k|l} \triangleq \widehat{\mathbf{x}}_{k|l}(\mathbf{y}_1, \ldots, \mathbf{y}_l)$ 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 certain uncorrelation conditions (lately extended in [2]) and are Gaussian, the minimum variance or minimum mean squared error (MSE) filter estimate for LDSS models has a convenient recursive predictor/corrector format, $\forall k \geq 1$

$$\widehat{\mathbf{x}}^b_{k|k} = \mathbf{F}_{k-1}\widehat{\mathbf{x}}^b_{k-1|k-1} + \mathbf{K}^b_k\left(\mathbf{y}_k - \mathbf{H}_k\mathbf{F}_{k-1}\widehat{\mathbf{x}}^b_{k-1|k-1}\right) \tag{2}$$

where $\widehat{\mathbf{x}}^b_{0|0} = E[\mathbf{x}_0]$, so-called the Kalman filter (KF).[1] Even if the noises are nonGaussian, the KF is the linear least mean squares (LLMS) filter (LLMSF) estimate. As the computation of the KF depends on prior information on the mean ($E[\mathbf{x}_0]$) and on the covariance matrix ($\mathbf{C}_{\mathbf{x}_0}$) of $\mathbf{x}_0$ [3]–[5], the KF can be looked upon as an "initial state first and second order statistics" matched filter [2]. 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 [6], [7, Sec. 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 $\mathbf{x}_1$ [8, Sec. 6], [9, Sec. 5.6], [10] and coincides with the weighted least squares estimator (WLSE) of $\mathbf{x}_1$ [11]. If $\mathbf{H}_1$ and the covariance matrix of $\mathbf{v}_1$ ($\mathbf{C}_{\mathbf{v}_1}$) are full rank, the Fisher initialization yields

$$\widehat{\mathbf{x}}^b_{1|1} = \mathbf{P}^b_{1|1}\mathbf{H}_1^H\mathbf{C}_{\mathbf{v}_1}^{-1}\mathbf{y}_1,\ \mathbf{P}^b_{1|1} = \left(\mathbf{H}_1^H\mathbf{C}_{\mathbf{v}_1}^{-1}\mathbf{H}_1\right)^{-1}. \tag{3}$$

A particularly noteworthy feature of this alternative initialization of the KF (3) is that it may yield the stochastic LMVDR filter (LMVDRF), which shares the same recursion as the KF, except at time $k = 1$. Indeed, Chaumette *et al.* in [2] have lately shown that this property holds for the restricted subset of LDSS models for which the state matrices $\mathbf{F}_k$, $k \geq 1$, are invertible.

II. Problem Statement and Novelty

Unfortunately, this restricted subset of LDSS models does not include fixed-point and fixed-lag smoothers, which are obtained by running the KF on augmented LDSS models [4, Sec. 9] incorporating at least one noninvertible state matrix. To solve this issue, we show that the invertibility of $\mathbf{F}_k$, $k \geq 1$, is actually not required. More specifically, by resorting to a different approach than the one previously used in [2], we establish the general conditions, in terms of noises covariance matrices and model matrices, under which existence of the LMVDRF is guaranteed. In a nutshell, it is shown that provided that $\mathbf{H}_1$ is full rank, the LMVDRF exists whenever the KF exists. If $\mathbf{H}_1$ is not full rank, the LMVDRF may not exist, and if it exists, then its numerical

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

The authors are with the ISAE-SUPAERO, Université de Toulouse, Toulouse 31055, France (e-mail: eric.chaumette@isae.fr; francois.vincent@isae.fr; benoit.priot@isae.fr; gael.pages@isae.fr; arnaud.dion@isae.fr).

computation may be untractable. First, these results allows for a comparison between the LMVDRF and the information filter (IF) form of the KF [4, Sec. 6.2], another well established solution to cope with a lack of prior information on $\mathbf{x}_0$. Second, LMVDR fixed-point and fixed-lag smoothers are introduced (and possibly other smoothers or predictors, which is left for future research), whose existence could not be proven from [2].

On another note, LMVDR estimators may allow to derive unexpected results, as highlighted with the LMVDR fixed-point smoother. Indeed, it is shown that the LMVDR fixed-point smoother can be used to compute recursively the solution of a generalization of the deterministic least-squares problem, that is a generalized WLSE (possibly regularized).

III. Notations and Signal Model

The notational convention adopted is as follows: scalars, vectors, and matrices are represented, respectively, by italic, bold lowercase, and bold uppercase characters. $\mathcal{M}_{\mathbb{C}}(N,P)$ denotes the vector space of complex matrices with N rows and P columns. The scalar/matrix/vector transpose conjugate is indicated by the superscript ^{H}. $\mathbf{1}_N$ denotes a N-dimensional vector with components equal to 1. $\mathbf{I}$ is the identity matrix. $[\mathbf{A}\,\mathbf{B}]$ 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 $\mathbf{A}_1,\ldots,\mathbf{A}_k$ of same column number is denoted $\overline{\mathbf{A}}_k$. $E[\cdot]$ denotes the expectation operator. If $\mathbf{x}$ and $\mathbf{y}$ are following two complex random vectors:

1) $\mathbf{C}_{\mathbf{x}}$, $\mathbf{C}_{\mathbf{y}}$, 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 $\mathbf{y}$;
2) if $\mathbf{C}_{\mathbf{y}}$ is invertible, then $\mathbf{C}_{\mathbf{x}|\mathbf{y}} \triangleq \mathbf{C}_{\mathbf{x}} - \mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1}\mathbf{C}_{\mathbf{x},\mathbf{y}}^{H}$.

As in [2], [3, Sec. 3], [9, Sec. 5.4], we adopt a joint proper[2] (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. 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 [3, Sec. 3.2.5], [9, Sec. 5.4.1].

First, as (1a) can be rewritten as, for $k \geq 2$

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

$$\mathbf{G}_k \in \mathcal{M}_{\mathbb{C}}(P_k,\mathcal{P}_{k-1}),\ \mathbf{B}_{k,l} = \left|\begin{array}{ll}\mathbf{F}_{k-1}\mathbf{F}_{k-2}\cdots\mathbf{F}_l, & k>l\\ \mathbf{I} & ,k=l\\ \mathbf{0} & ,k<l\end{array}\right. \tag{4}$$

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

$$\mathbf{y}_k = \mathbf{A}_k\mathbf{x}_1 + \mathbf{n}_k,\ \mathbf{A}_k = \mathbf{H}_k\mathbf{B}_{k,1}$$
$$\mathbf{n}_k = \mathbf{v}_k + \mathbf{H}_k\mathbf{G}_k\overline{\mathbf{w}}_{k-1}. \tag{5a}$$

Second, let $\mathbf{A}_1 = \mathbf{H}_1$ and $\mathbf{n}_1 = \mathbf{v}_1$; then (1b) can be extended on a horizon of k points from the first measurement as follows:

$$\overline{\mathbf{y}}_k = \overline{\mathbf{A}}_k\mathbf{x}_1 + \overline{\mathbf{n}}_k$$

$$\overline{\mathbf{y}}_k, \overline{\mathbf{n}}_k \in \mathcal{M}_{\mathbb{C}}(\mathcal{N}_k,1), \overline{\mathbf{A}}_k \in \mathcal{M}_{\mathbb{C}}(\mathcal{N}_k,P_k), \mathcal{N}_k = \sum_{l=1}^{k} N_l. \tag{5b}$$

[2]A proper complex random variable is uncorrelated with its complex conjugate [9].

IV. LLMSF for LDSS Models

In this section, first, the general assumptions required on LDSS models to obtain a LLMSF satisfying the same predictor/corrector format as the KF (2) (without extension of the state and measurement equations) are introduced in a more comprehensible manner than in [2], [12, Sec. II]. Second, under these general assumptions, we feature an insightful breakdown of the MSE of linear filters, which allows not only to derive easily the general form of the KF recursion released in [2, (19a–c)], but also to prove that, whenever it exits, the LMVDRF shares the same recursion as the KF except at initialization.

A. LLMSF for LDSS Models

It has been known for ages [9, Sec. 5.4.1], [13] that, if $\mathbf{x}$ and $\mathbf{y}$ are two zero mean proper complex random vectors, then, provided that $\mathbf{C}_{\mathbf{y}}$ is invertible, the linear estimator of $\mathbf{x}$ which minimizes the error covariance matrix w.r.t. the Löwner ordering [14, Sec. 7.7], so called the LLMS estimator, is given by the following:

$$\widehat{\mathbf{x}}^b \triangleq \mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1}\mathbf{y},\quad E\left[\left(\widehat{\mathbf{x}}^b-\mathbf{x}\right)\left(\widehat{\mathbf{x}}^b-\mathbf{x}\right)^H\right] = \mathbf{C}_{\mathbf{x}|\mathbf{y}}. \tag{6}$$

Therefore, if $E[\mathbf{x}_0]=\mathbf{0}$, the LLMSF of $\mathbf{x}_k$ based on measurements up to and including time k, $k\geq 2$, is simply

$$\widehat{\mathbf{x}}^b_{k|k} = \mathbb{K}^b_k\overline{\mathbf{y}}_k = \left[\mathbb{J}^b_{k-1}\ \mathbf{K}^b_k\right]\overline{\mathbf{y}}_k \mid \left[\mathbb{J}^b_{k-1}\ \mathbf{K}^b_k\right]\mathbf{C}_{\overline{\mathbf{y}}_k} = \mathbf{C}_{\mathbf{x}_k,\overline{\mathbf{y}}_k} \tag{7}$$

provided that $\mathbf{C}_{\overline{\mathbf{y}}_k}$ is invertible, where $\mathbb{J}^b_{k-1}\in\mathcal{M}_{\mathbb{C}}(P_k,\mathcal{N}_{k-1})$ and $\mathbf{K}^b_k\in\mathcal{M}_{\mathbb{C}}(P_k,N_k)$. Thus, $\mathbb{J}^b_{k-1}$ and $\mathbf{K}^b_k$ are the solution of the system of linear equations

$$\begin{cases}\mathbb{J}^b_{k-1}\mathbf{C}_{\overline{\mathbf{y}}_{k-1}} + \mathbf{K}^b_k\mathbf{C}_{\mathbf{y}_k,\overline{\mathbf{y}}_{k-1}} = \mathbf{C}_{\mathbf{x}_k,\overline{\mathbf{y}}_{k-1}}\\ \mathbb{J}^b_{k-1}\mathbf{C}_{\overline{\mathbf{y}}_{k-1},\mathbf{y}_k} + \mathbf{K}^b_k\mathbf{C}_{\mathbf{y}_k} = \mathbf{C}_{\mathbf{x}_k,\mathbf{y}_k}\end{cases}$$

which yields

$$\mathbb{J}^b_{k-1} = \mathbf{C}_{\mathbf{x}_k,\overline{\mathbf{y}}_{k-1}}\mathbf{C}^{-1}_{\overline{\mathbf{y}}_{k-1}} - \mathbf{K}^b_k\mathbf{C}_{\mathbf{y}_k,\overline{\mathbf{y}}_{k-1}}\mathbf{C}^{-1}_{\overline{\mathbf{y}}_{k-1}}.$$

Consequently, from (6), (7) can be rewritten as follows:

$$\widehat{\mathbf{x}}^b_{k|k} = \mathbb{J}^b_{k-1}\overline{\mathbf{y}}_{k-1} + \mathbf{K}^b_k\mathbf{y}_k = \widehat{\mathbf{x}}^b_{k|k-1} + \mathbf{K}^b_k\left(\mathbf{y}_k - \widehat{\mathbf{y}}^b_{k|k-1}\right) \tag{8}$$

which is the general form of the so-called predictor/corrector format of the LLMSF. Moreover, as $\mathbf{C}_{\mathbf{y}_k,\overline{\mathbf{y}}_{k-1}} = \mathbf{H}_k\mathbf{C}_{\mathbf{x}_k,\overline{\mathbf{y}}_{k-1}} + \mathbf{C}_{\mathbf{v}_k,\overline{\mathbf{y}}_{k-1}}$ and $\mathbf{C}_{\mathbf{x}_k,\overline{\mathbf{y}}_{k-1}} = \mathbf{F}_{k-1}\mathbf{C}_{\mathbf{x}_{k-1},\overline{\mathbf{y}}_{k-1}} + \mathbf{C}_{\mathbf{w}_{k-1},\overline{\mathbf{y}}_{k-1}}$, then according to (6)

$$\widehat{\mathbf{y}}^b_{k|k-1} = \mathbf{H}_k\widehat{\mathbf{x}}^b_{k|k-1} + \widehat{\mathbf{v}}^b_{k|k-1}$$
$$\widehat{\mathbf{x}}^b_{k|k-1} = \mathbf{F}_{k-1}\widehat{\mathbf{x}}^b_{k-1|k-1} + \widehat{\mathbf{w}}^b_{k-1|k-1}$$

and (8) can be recasted as follows:

$$\begin{aligned}\widehat{\mathbf{x}}^b_{k|k} &= \left(\mathbf{I}-\mathbf{K}^b_k\mathbf{H}_k\right)\mathbf{F}_{k-1}\widehat{\mathbf{x}}^b_{k-1|k-1} + \mathbf{K}^b_k\mathbf{y}_k\\ &\quad + \left(\mathbf{I}-\mathbf{K}^b_k\mathbf{H}_k\right)\widehat{\mathbf{w}}^b_{k-1|k-1} - \mathbf{K}^b_k\widehat{\mathbf{v}}^b_{k|k-1},\ k\geq 2\end{aligned} \tag{9}$$

a general form already released in [2, (16)] but at the expense of a more complex derivation [12, Sec. II]. It is noteworthy that (9) has two additional terms in comparison with the recursive predictor/corrector form (2) introduced by Kalman [1]. Therefore, the general assumptions required to obtain the Kalman form (2) of the LLMSF for LDSS (9) are as follows:

$$\mathbf{C}_{\overline{\mathbf{y}}_k}\ \text{invertible}, \forall\overline{\mathbf{y}}_{k-1}: \begin{cases}\widehat{\mathbf{w}}^b_{k-1|k-1} = \mathbf{0}\\ \widehat{\mathbf{v}}^b_{k|k-1} = \mathbf{0}\end{cases},\ k\geq 2 \tag{10a}$$

that is

$$\mathbf{C}_{\overline{\mathbf{y}}_k}\ \text{invertible}, \mathbf{C}_{\mathbf{w}_{k-1},\overline{\mathbf{y}}_{k-1}} = \mathbf{0},\ \mathbf{C}_{\mathbf{v}_k,\overline{\mathbf{y}}_{k-1}} = \mathbf{0}, k\geq 2. \tag{10b}$$

B. Insightful Breakdown of the MSE of Linear Filters

Another noteworthy point is that under the general assumptions (10b), the MSE of any linear filter $\widehat{\mathbf{x}}_{k|k} = \mathbb{K}_k \overline{\mathbf{y}}_k$, $\mathbb{K}_k = [\mathbb{J}_{k-1}\ \mathbf{K}_k]$, where $\mathbb{J}_{k-1} \in \mathcal{M}_{\mathbb{C}}(P_k, \mathcal{N}_{k-1})$ and $\mathbf{K}_k \in \mathcal{M}_{\mathbb{C}}(P_k, N_k)$, that is,

$$\begin{aligned}\mathbf{P}_{k|k}(\mathbb{K}_k) &= E\left[\left(\widehat{\mathbf{x}}_{k|k} - \mathbf{x}_k\right)\left(\widehat{\mathbf{x}}_{k|k} - \mathbf{x}_k\right)^H\right] \\ &= \mathbf{P}_{k|k}(\mathbb{J}_{k-1}, \mathbf{K}_k) \qquad (11)\end{aligned}$$

breaks down into

$$\begin{aligned}&\mathbf{P}_{k|k}(\mathbb{J}_{k-1}, \mathbf{K}_k) = \mathbf{Q}_{k-1}(\mathbb{J}_{k-1}, \mathbf{K}_k) \\ &+ (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\begin{pmatrix}\mathbf{C}_{\mathbf{w}_{k-1}} + \mathbf{F}_{k-1}\mathbf{C}^H_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}} \\ +\mathbf{C}_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}}\mathbf{F}^H_{k-1}\end{pmatrix}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^H \\ &- (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{C}_{\mathbf{x}_k,\mathbf{v}_k}\mathbf{K}^H_k - \mathbf{K}_k\mathbf{C}^H_{\mathbf{x}_k,\mathbf{v}_k}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^H \\ &+ \mathbf{K}_k\mathbf{C}_{\mathbf{v}_k}\mathbf{K}^H_k \qquad (12a)\end{aligned}$$

where

$$\begin{aligned}\mathbf{Q}_{k-1}(\mathbb{J}_{k-1}, \mathbf{K}_k) &= E\left[\widehat{\mathbf{q}}_{k-1}\widehat{\mathbf{q}}^H_{k-1}\right] \\ \widehat{\mathbf{q}}_{k-1} &= \mathbb{J}_{k-1}\overline{\mathbf{y}}_{k-1} - (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{F}_{k-1}\mathbf{x}_{k-1} \qquad (12b)\end{aligned}$$

which is a key result in order to derive straightforwardly (in comparison with [2], [12, Sec. II]) the general form of the KF recursion (without extension of the state and measurement equations). The MSE breakdown (12a) is easily obtained from the combination of (10b) and the next breakdown of the error

$$\widehat{\mathbf{x}}_{k|k} - \mathbf{x}_k = \widehat{\mathbf{q}}_{k-1} + (\mathbf{K}_k\mathbf{v}_k - (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{w}_{k-1}).$$

C. General Form of the KF Recursion

From (12a), it is obvious that

$$\mathbb{J}^b_{k-1} = \arg\min_{\mathbb{K}_{k-1}}\{\mathbf{Q}_{k-1}(\mathbb{J}_{k-1}, \mathbf{K}_k)\} \qquad (13a)$$

that is (6)

$$\begin{aligned}\mathbb{J}^b_{k-1}\overline{\mathbf{y}}_{k-1} &= \mathbf{C}_{(\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)\mathbf{F}_{k-1}\mathbf{x}_{k-1},\overline{\mathbf{y}}_{k-1}}\mathbf{C}^{-1}_{\overline{\mathbf{y}}_{k-1}}\overline{\mathbf{y}}_{k-1} \\ &= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{F}_{k-1}\widehat{\mathbf{x}}^b_{k-1|k-1} \qquad (13b)\end{aligned}$$

leading to the general form of the Joseph stabilized version of the covariance measurement update equation [2]

$$\begin{aligned}&\mathbf{P}_{k|k}\left(\mathbb{J}^b_{k-1}, \mathbf{K}_k\right) = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}^b_{k|k-1}(\mathbf{I} - \mathbf{K}_k\mathbf{W}_k)^H \\ &- (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{C}_{\mathbf{x}_k,\mathbf{v}_k}\mathbf{K}^H_k - \mathbf{K}_k\mathbf{C}^H_{\mathbf{x}_k,\mathbf{v}_k}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^H \\ &+ \mathbf{K}_k\mathbf{C}_{\mathbf{v}_k}\mathbf{K}^H_k \qquad (14)\end{aligned}$$

where

$$\begin{aligned}\mathbf{P}^b_{k|k-1} &= \mathbf{F}_{k-1}\mathbf{P}^b_{k-1|k-1}\mathbf{F}^H_{k-1} + \mathbf{C}_{\mathbf{w}_{k-1}} \\ &\quad + \mathbf{F}_{k-1}\mathbf{C}^H_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}} + \mathbf{C}_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}}\mathbf{F}^H_{k-1} \\ &= E\left[\left(\widehat{\mathbf{x}}^b_{k|k-1} - \mathbf{x}_k\right)\left(\widehat{\mathbf{x}}^b_{k|k-1} - \mathbf{x}_k\right)^H\right]\end{aligned}$$

since under (10a–b)

$$\widehat{\mathbf{x}}^b_{k|k-1} = \mathbf{F}_{k-1}\widehat{\mathbf{x}}^b_{k-1|k-1} + \widehat{\mathbf{w}}^b_{k-1|k-1} = \mathbf{F}_{k-1}\widehat{\mathbf{x}}^b_{k-1|k-1}.$$

The solution $\mathbf{K}^b_k$ of the minimization of (14) can be computed according to the following recursion [2] for $k \geq 2$:

$$\begin{aligned}\mathbf{P}^b_{k|k-1} &= \mathbf{F}_{k-1}\mathbf{P}^b_{k-1|k-1}\mathbf{F}^H_{k-1} + \mathbf{C}_{\mathbf{w}_{k-1}} \\ &\quad + \mathbf{F}_{k-1}\mathbf{C}^H_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}} + \mathbf{C}_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}}\mathbf{F}^H_{k-1} \qquad (15a) \\ \mathbf{S}^b_{k|k} &= \mathbf{H}_k\mathbf{P}^b_{k|k-1}\mathbf{H}^H_k + \mathbf{C}_{\mathbf{v}_k} + \mathbf{H}_k\mathbf{C}^H_{\mathbf{v}_k,\mathbf{x}_k} + \mathbf{C}_{\mathbf{v}_k,\mathbf{x}_k}\mathbf{H}^H_k \\ \mathbf{K}^b_k &= \left(\mathbf{P}^b_{k|k-1}\mathbf{H}^H_k + \mathbf{C}^H_{\mathbf{v}_k,\mathbf{x}_k}\right)\left(\mathbf{S}^b_{k|k}\right)^{-1} \qquad (15b) \\ \mathbf{P}^b_{k|k} &= \left(\mathbf{I} - \mathbf{K}^b_k\mathbf{H}_k\right)\mathbf{P}^b_{k|k-1} - \mathbf{K}^b_k\mathbf{C}_{\mathbf{v}_k,\mathbf{x}_k}. \qquad (15c)\end{aligned}$$

The above recursion is also valid for $k = 1$ provided that $\mathbf{P}^b_{0|0} = \mathbf{C}_{\mathbf{x}_0}$ and $\widehat{\mathbf{x}}^b_{0|0} = \mathbf{0}$ [2]. The case of a nonzero mean initial state $\mathbf{x}_0$ is addressed by simply setting $\widehat{\mathbf{x}}^b_{0|0} = E[\mathbf{x}_0]$. Last, let us remind that $\widehat{\mathbf{x}}^b_{k|k-1}$ is also known as the *a priori* estimate of $\mathbf{x}_k$, and $\mathbf{S}^b_{k|k} = E[\varepsilon^b_k(\varepsilon^b_k)^H]$, where $\varepsilon^b_k = \mathbf{y}_k - \mathbf{H}_k\widehat{\mathbf{x}}^b_{k|k-1} = \mathbf{y}_k - \widehat{\mathbf{y}}^b_{k|k-1}$ is the innovations vector.

V. LMVDRF FOR LDSS MODELS

In this section, we consider a completely different approach than the one previously used in [2]. Indeed, we provide a general definition of a distortionless filter in the context of LDSS models (16), which encompasses the definition used in [2] . And, it is the combination of this general definition with the MSE breakdown (12a) that allows to prove that, whenever it exits, the LMVDRF shares the same recursion as the KF except at initialization.

As in [2], we adopt the notation used in the deterministic framework for the LMVDRF [8, Sec. 6], [9, Sec. 5.6] to stress the fact that the LMVDRF is different from the LLMSF, a.k.a. 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 [8, Sec. 6], [9, Sec. 5.6], that is, $\overline{\mathbf{W}}_k \in \mathcal{M}_{\mathbb{C}}(\mathcal{N}_k, P_k)$ yielding the state vector $\overline{\mathbf{W}}^H_k\overline{\mathbf{y}}_k$, which can be recast as (5b)

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

Therefore, according to (4), a filter $\overline{\mathbf{W}}_k$, $k \geq 2$, is distortionless iff

$$\overline{\mathbf{W}}^H_k\overline{\mathbf{y}}_k = \mathbf{x}_k + \overline{\mathbf{W}}^H_k\overline{\mathbf{n}}_k - \mathbf{G}_k\overline{\mathbf{w}}_{k-1} \Leftrightarrow \overline{\mathbf{W}}^H_k\overline{\mathbf{A}}_k = \mathbf{B}_{k,1} \qquad (16)$$

which leads to the following definition of the best distortionless state-former in the MSE sense, a.k.a. the LMVDRF:

$$\overline{\mathbf{W}}^b_k = \arg\min_{\overline{\mathbf{W}}_k}\left\{\mathbf{P}_{k|k}\left(\overline{\mathbf{W}}_k\right)\right\} \text{ s.t. } \overline{\mathbf{W}}^H_k\overline{\mathbf{A}}_k = \mathbf{B}_{k,1} \qquad (17)$$

$$\mathbf{P}_{k|k}\left(\overline{\mathbf{W}}_k\right) = E\left[\left(\overline{\mathbf{W}}^H_k\overline{\mathbf{y}}_k - \mathbf{x}_k\right)\left(\overline{\mathbf{W}}^H_k\overline{\mathbf{y}}_k - \mathbf{x}_k\right)^H\right]. \qquad (18)$$

Since $\mathbf{C}_{\overline{\mathbf{y}}_k}$ is invertible (10b), then provided that $\overline{\mathbf{A}}_k$ is full rank, $\overline{\mathbf{W}}^b_k$ is analogous to a linearly constrained Wiener filter (LCWF) [15, Sec. 2.5] whose batch form expression is given by [15, (2.113)]

$$\begin{aligned}\overline{\mathbf{W}}^b_k &= \mathbf{C}^{-1}_{\overline{\mathbf{y}}_k}\mathbf{C}_{\overline{\mathbf{y}}_k,\mathbf{x}_k} + \mathbf{C}^{-1}_{\overline{\mathbf{y}}_k}\overline{\mathbf{A}}_k\left(\overline{\mathbf{A}}^H_k\mathbf{C}^{-1}_{\overline{\mathbf{y}}_k}\overline{\mathbf{A}}_k\right)^{-1}\mathbf{B}^H_{k,1} \\ &\quad - \mathbf{C}^{-1}_{\overline{\mathbf{y}}_k}\overline{\mathbf{A}}_k\left(\overline{\mathbf{A}}^H_k\mathbf{C}^{-1}_{\overline{\mathbf{y}}_k}\overline{\mathbf{A}}_k\right)^{-1}\overline{\mathbf{A}}^H_k\mathbf{C}^{-1}_{\overline{\mathbf{y}}_k}\mathbf{C}_{\overline{\mathbf{y}}_k,\mathbf{x}_k}. \qquad (19)\end{aligned}$$

Since the KF is the solution of the following unconstrained minimization problem (6–7) [1], [3]–[5]:

$$\mathbb{K}_k^{bH} = \arg\min_{\overline{\mathbf{W}}_k} \left\{\mathbf{P}_{k|k}\left(\overline{\mathbf{W}}_k\right)\right\} \tag{20}$$

it follows that the LMVDRF (17) is suboptimal in MSE sense in comparison with the KF (20).

A. $\mathbf{H}_1$ is Full Rank

If $\mathbf{H}_1$ is full rank, then $\overline{\mathbf{A}}_k$ (5b), $k \geq 2$, is full rank as well and $\overline{\mathbf{W}}_k^b$ (19) exists. Let $\overline{\mathbf{W}}_k = \left[\begin{smallmatrix}\overline{\mathbf{D}}_{k-1}\\ \mathbf{W}_k\end{smallmatrix}\right]$, $\overline{\mathbf{D}}_{k-1} \in \mathcal{M}_{\mathbb{C}}\left(\mathcal{N}_{k-1}, P_k\right)$ and $\mathbf{W}_k \in \mathcal{M}_{\mathbb{C}}\left(N_k, P_k\right)$. The MSE breakdown (12a–b) is also valid for any distortionless state-former, provided that one substitutes $\overline{\mathbf{D}}_{k-1}^H$ for $\mathbb{J}_{k-1}$ and $\mathbf{W}_k^H$ for $\mathbf{K}_k$, yielding

$$\begin{aligned}\mathbf{Q}_{k-1}\left(\overline{\mathbf{D}}_{k-1}, \mathbf{W}_k\right) &= E\left[\widehat{\mathbf{q}}_{k-1}\widehat{\mathbf{q}}_{k-1}^H\right]\\ \widehat{\mathbf{q}}_{k-1} &= \overline{\mathbf{D}}_{k-1}^H\overline{\mathbf{y}}_{k-1} - \left(\mathbf{I} - \mathbf{W}_k^H\mathbf{H}_k\right)\mathbf{F}_{k-1}\mathbf{x}_{k-1}.\end{aligned} \tag{21}$$

It is also a key result in order to derive the recursive form of the LMVDRF (17), (19). Indeed, as shown in the following, the MSE breakdown (12a) allows to breakdown the initial constrained minimization problem (17) into two separable minimization problems: a first constrained minimization problem w.r.t. $\overline{\mathbf{D}}_{k-1}$, namely

$$\overline{\mathbf{D}}_{k-1}^b = \arg\min_{\overline{\mathbf{D}}_{k-1}} \left\{\mathbf{Q}_{k-1}\left(\overline{\mathbf{D}}_{k-1}, \mathbf{W}_k\right)\right\} \text{ s.t. } \overline{\mathbf{W}}_k^H\overline{\mathbf{A}}_k = \mathbf{B}_{k,1} \tag{22a}$$

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

$$\mathbf{W}_k^b = \arg\min_{\mathbf{W}_k}\left\{\mathbf{P}_{k|k}\left(\overline{\mathbf{D}}_{k-1}^b, \mathbf{W}_k\right)\right\}. \tag{22b}$$

• Solution of (22a)

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

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

and (22a) is equivalent to

$$\begin{aligned}\overline{\mathbf{D}}_{k-1}^b &= \arg\min_{\overline{\mathbf{D}}_{k-1}}\left\{E\left[\widehat{\mathbf{q}}_{k-1}\widehat{\mathbf{q}}_{k-1}^H\right]\right\}\\ \text{s.t. } \overline{\mathbf{D}}_{k-1}^H\overline{\mathbf{A}}_{k-1} &= \left(\mathbf{I} - \mathbf{W}_k^H\mathbf{H}_k\right)\mathbf{F}_{k-1}\mathbf{B}_{k-1,1}.\end{aligned} \tag{23}$$

If $\mathbf{H}_1$ is full rank, then $\overline{\mathbf{A}}_{k-1}$, $k \geq 2$, is full rank as well. Moreover, since $\mathbf{C}_{\overline{\mathbf{y}}_k}$ is invertible, $\mathbf{C}_{\overline{\mathbf{y}}_{k-1}}$ is invertible as well. Therefore, $\overline{\mathbf{D}}_{k-1}^b$ (23) is a LCWF [15, Sec. 2.5] whose batch form expression can be computed as [15, eq. (2.113)]

$$\overline{\mathbf{D}}_{k-1}^b = \overline{\mathbf{W}}_{k-1}^b\left(\left(\mathbf{I} - \mathbf{W}_k^H\mathbf{H}_k\right)\mathbf{F}_{k-1}\right)^H \tag{24a}$$

$$\begin{aligned}\overline{\mathbf{W}}_{k-1}^b &= \mathbf{C}_{\overline{\mathbf{y}}_{k-1}}^{-1}\mathbf{C}_{\overline{\mathbf{y}}_{k-1},\mathbf{x}_{k-1}}\\ &+ \mathbf{C}_{\overline{\mathbf{y}}_{k-1}}^{-1}\overline{\mathbf{A}}_{k-1}\left(\overline{\mathbf{A}}_{k-1}^H\mathbf{C}_{\overline{\mathbf{y}}_{k-1}}^{-1}\overline{\mathbf{A}}_{k-1}\right)^{-1}\mathbf{B}_{k-1,1}^H\\ &- \mathbf{C}_{\overline{\mathbf{y}}_{k-1}}^{-1}\overline{\mathbf{A}}_{k-1}\left(\overline{\mathbf{A}}_{k-1}^H\mathbf{C}_{\overline{\mathbf{y}}_{k-1}}^{-1}\overline{\mathbf{A}}_{k-1}\right)^{-1}\overline{\mathbf{A}}_{k-1}^H\mathbf{C}_{\overline{\mathbf{y}}_{k-1}}^{-1}\mathbf{C}_{\overline{\mathbf{y}}_{k-1},\mathbf{x}_{k-1}}\end{aligned} \tag{24b}$$

where $\overline{\mathbf{W}}_{k-1}^b$ coincides with the LMVDRF at time $k-1$. Indeed, (24b) is the solution of (19)

$$\begin{aligned}\overline{\mathbf{W}}_{k-1}^b &= \arg\min_{\overline{\mathbf{W}}_{k-1}}\left\{\mathbf{P}_{k-1|k-1}\left(\overline{\mathbf{W}}_{k-1}\right)\right\}\\ \text{s.t. } \overline{\mathbf{W}}_{k-1}^H\overline{\mathbf{A}}_{k-1} &= \mathbf{B}_{k-1,1}.\end{aligned} \tag{24c}$$

Finally, $\forall k \geq 2$

$$\begin{aligned}\mathbf{Q}_{k-1}\left(\overline{\mathbf{D}}_{k-1}^b, \mathbf{W}_k\right) &= \left(\mathbf{I} - \mathbf{W}_k^H\mathbf{H}_k\right)\mathbf{F}_{k-1}\\ &\times \mathbf{P}_{k-1|k-1}\left(\overline{\mathbf{W}}_{k-1}^b\right)\mathbf{F}_{k-1}^H\left(\mathbf{I} - \mathbf{H}_k^H\mathbf{W}_k\right).\end{aligned} \tag{25}$$

• Solution of (22b)

According to (25), the solution $\overline{\mathbf{D}}_{k-1}^b \triangleq \overline{\mathbf{D}}_{k-1}^b\left(\mathbf{W}_k\right)$ (24a) of the first constrained minimization problem (22a) leads to the following:

$$\begin{aligned}\mathbf{P}_{k|k}\left(\overline{\mathbf{D}}_{k-1}^b, \mathbf{W}_k\right) &= \left(\mathbf{I} - \mathbf{W}_k^H\mathbf{H}_k\right)\mathbf{P}_{k|k-1}^b\left(\mathbf{I} - \mathbf{H}_k^H\mathbf{W}_k\right)\\ &- \left(\mathbf{I} - \mathbf{W}_k^H\mathbf{H}_k\right)\mathbf{C}_{\mathbf{x}_k,\mathbf{v}_k}\mathbf{W}_k - \mathbf{W}_k^H\mathbf{C}_{\mathbf{x}_k,\mathbf{v}_k}^H\\ &\times \left(\mathbf{I} - \mathbf{H}_k^H\mathbf{W}_k\right) + \mathbf{W}_k^H\mathbf{C}_{\mathbf{v}_k}\mathbf{W}_k\end{aligned} \tag{26a}$$

$$\begin{aligned}\mathbf{P}_{k|k-1}^b &= \mathbf{F}_{k-1}\mathbf{P}_{k-1|k-1}^b\mathbf{F}_{k-1}^H + \mathbf{C}_{\mathbf{w}_{k-1}}\\ &+ \mathbf{F}_{k-1}\mathbf{C}_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}}^H + \mathbf{C}_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}}\mathbf{F}_{k-1}^H\end{aligned} \tag{26b}$$

which is the general form of the Joseph stabilized version of the covariance measurement update (14), provided that one substitutes $\mathbf{W}_k^H$ for $\mathbf{K}_k$. Therefore, the solution $\mathbf{W}_k^b$ of the minimization of (26a), that is,

$$\mathbf{W}_k^b = \arg\min_{\mathbf{W}_k}\left\{\mathbf{P}_{k|k}\left(\overline{\mathbf{D}}_{k-1}^b\left(\mathbf{W}_k\right), \mathbf{W}_k\right)\right\} \tag{27a}$$

can be computed according to (15a–c) provided that one substitutes $\left(\mathbf{W}_k^b\right)^H$ for $\mathbf{K}_k^b$, i.e.,

$$\begin{aligned}\mathbf{P}_{k|k-1}^b &= \mathbf{F}_{k-1}\mathbf{P}_{k-1|k-1}^b\mathbf{F}_{k-1}^H + \mathbf{C}_{\mathbf{w}_{k-1}}\\ &+ \mathbf{F}_{k-1}\mathbf{C}_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}}^H + \mathbf{C}_{\mathbf{w}_{k-1},\mathbf{x}_{k-1}}\mathbf{F}_{k-1}^H\end{aligned} \tag{27b}$$

$$\mathbf{S}_{k|k}^b = \mathbf{H}_k\mathbf{P}_{k|k-1}^b\mathbf{H}_k^H + \mathbf{C}_{\mathbf{v}_k} + \mathbf{H}_k\mathbf{C}_{\mathbf{v}_k,\mathbf{x}_k}^H + \mathbf{C}_{\mathbf{v}_k,\mathbf{x}_k}\mathbf{H}_k^H$$

$$\mathbf{W}_k^b = \left(\mathbf{S}_{k|k}^b\right)^{-1}\left(\mathbf{H}_k\mathbf{P}_{k|k-1}^b + \mathbf{C}_{\mathbf{v}_k,\mathbf{x}_k}\right) \tag{27c}$$

$$\mathbf{P}_{k|k}^b = \left(\mathbf{I} - \mathbf{W}_k^{bH}\mathbf{H}_k\right)\mathbf{P}_{k|k-1}^b - \mathbf{W}_k^{bH}\mathbf{C}_{\mathbf{v}_k,\mathbf{x}_k}. \tag{27d}$$

• **Summary**

For $k \geq 2$, according to (24a) and (27a), the LMVDRF (17) yields the state-former

$$\widehat{\mathbf{x}}_{k|k}^{b} = \left(\mathbf{I} - \mathbf{W}_{k}^{bH}\mathbf{H}_{k}\right)\mathbf{F}_{k-1}\left(\overline{\mathbf{W}}_{k-1}^{bH}\overline{\mathbf{y}}_{k-1}\right) + \mathbf{W}_{k}^{bH}\mathbf{y}_{k}$$
$$= \left(\mathbf{I} - \mathbf{W}_{k}^{bH}\mathbf{H}_{k}\right)\mathbf{F}_{k-1}\widehat{\mathbf{x}}_{k-1|k-1}^{b} + \mathbf{W}_{k}^{bH}\mathbf{y}_{k} \tag{27e}$$

where $\mathbf{W}_{k}^{b}$ is given by the recursion (27b–d), similar to the general form of the KF recursion (15a–c). At time $k=1$: $\mathbf{W}_{1}^{b} = \arg\min_{\mathbf{W}_1}\left\{\mathbf{P}_{1|1}\left(\mathbf{W}_{1}\right)\right\}$ s.t. $\mathbf{W}_{1}^{H}\mathbf{H}_{1} = \mathbf{I}$, leading to $\widehat{\mathbf{x}}_{1|1}^{b} = \mathbf{P}_{1|1}^{b}\mathbf{H}_{1}^{H}\mathbf{C}_{\mathbf{v}_1}^{-1}\mathbf{y}_{1}$, $\mathbf{P}_{1|1}^{b} = \left(\mathbf{H}_{1}^{H}\mathbf{C}_{\mathbf{v}_1}^{-1}\mathbf{H}_{1}\right)^{-1}$, which is the Fisher estimate of $\mathbf{x}_1$ (3).

To summarize the derivation above, provided that $\mathbf{H}_1$ is full rank and that $\mathbf{C}_{\overline{\mathbf{y}}_k}$ is invertible, then the Fisher initialization (3) of the KF does not yield a LLMSF any longer but a LMVDRF. Since $\mathbf{P}_{k|k}^{b} \triangleq \mathbf{P}_{k|k}(\overline{\mathbf{W}}_{k}^{b})$ depends on neither $E\left[\mathbf{x}_0\right]$ nor $\mathbf{C}_{\mathbf{x}_0}$, the LMVDRF is suboptimal 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 LMVDRF is an upper bound on the performance of the KF whatever the initial conditions $E\left[\mathbf{x}_0\right]$ and $\mathbf{C}_{\mathbf{x}_0}$.

B. Conditions of Existence When $\mathbf{H}_1$ is Not Full Rank

If $\mathbf{H}_1$ is not full rank, then the set of distortionless state-formers may be empty. For instance, let us consider a time-invariant LDSS system of the form: $\mathbf{x}_k = \mathbf{x}_{k-1} + \mathbf{w}_{k-1}$, $\mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k$. Then, $P_k = P$, $N_k = N$, and (5b) becomes $\overline{\mathbf{y}}_k = \overline{\mathbf{A}}_k\mathbf{x}_1 + \overline{\mathbf{n}}_k$, $\overline{\mathbf{A}}_k = \mathbf{1}_k \otimes \mathbf{H}$, where $\otimes$ denotes the Kronecker product. Since $\operatorname{rank}\left(\mathbf{1}_k \otimes \mathbf{H}\right) = \operatorname{rank}\left(\mathbf{1}_k\right)\operatorname{rank}\left(\mathbf{H}\right)$ [16, p. 235], therefore, $\operatorname{rank}\left(\overline{\mathbf{A}}_k\right) = \operatorname{rank}\left(\mathbf{H}\right)$. Thus, if $\mathbf{H}$ is rank-deficient, so is $\overline{\mathbf{A}}_k$, and the distortionless constraints (16) $\overline{\mathbf{W}}_{k}^{H}\overline{\mathbf{A}}_k = \mathbf{B}_{k,1} = \mathbf{I}$ can not be satisfied. On the other hand, if there exists a time q for which $\overline{\mathbf{A}}_q$ is full rank, then the LMVDRF exists and its batch form is given by (19)

$$\overline{\mathbf{W}}_{q}^{b} = \mathbf{C}_{\overline{\mathbf{y}}_q}^{-1}\mathbf{C}_{\overline{\mathbf{y}}_q,\mathbf{x}_q} + \mathbf{C}_{\overline{\mathbf{y}}_q}^{-1}\overline{\mathbf{A}}_q\left(\overline{\mathbf{A}}_q^{H}\mathbf{C}_{\overline{\mathbf{y}}_q}^{-1}\overline{\mathbf{A}}_q\right)^{-1}\mathbf{B}_{q,1}^{H}$$
$$- \mathbf{C}_{\overline{\mathbf{y}}_q}^{-1}\overline{\mathbf{A}}_q\left(\overline{\mathbf{A}}_q^{H}\mathbf{C}_{\overline{\mathbf{y}}_q}^{-1}\overline{\mathbf{A}}_q\right)^{-1}\overline{\mathbf{A}}_q^{H}\mathbf{C}_{\overline{\mathbf{y}}_q}^{-1}\mathbf{C}_{\overline{\mathbf{y}}_q,\mathbf{x}_q}. \tag{28}$$

However, $\overline{\mathbf{W}}_{q}^{b}$ (28) can not be expressed in a recursive form, which does not prevent from computing it theoretically, but may make its numerical computation untractable if q is too large. If $\widehat{\mathbf{x}}_{q|q}^{b}$ and $\mathbf{P}_{q|q}^{b} = E[(\widehat{\mathbf{x}}_{q|q}^{b} - \mathbf{x}_q)(\widehat{\mathbf{x}}_{q|q}^{b} - \mathbf{x}_q)^H]$ are numerically computable, then $\widehat{\mathbf{x}}_{q+1|q+1}^{b}, \ldots$ are computed according to the standard LMVDRF recursion (27b–e).

C. Prior-Free Estimate of $\mathbf{x}_1$ via the IF Form of the KF

If one assumes that: $\mathbf{C}_{\mathbf{x}_0,\mathbf{w}_k} = \mathbf{0}$, $\mathbf{C}_{\mathbf{x}_0,\mathbf{v}_k} = \mathbf{0}$, $\mathbf{C}_{\mathbf{w}_l,\mathbf{w}_k} = \mathbf{C}_{\mathbf{w}_k}\delta_k^l$, $\mathbf{C}_{\mathbf{v}_l,\mathbf{v}_k} = \mathbf{C}_{\mathbf{v}_k}\delta_k^l$, and $\mathbf{C}_{\mathbf{w}_l,\mathbf{v}_k} = \mathbf{0}$, then (15a–c) become

$$\mathbf{P}_{k|k-1}^{b} = \mathbf{F}_{k-1}\mathbf{P}_{k-1|k-1}^{b}\mathbf{F}_{k-1}^{H} + \mathbf{C}_{\mathbf{w}_{k-1}} \tag{29a}$$

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

$$\mathbf{K}_{k}^{b} = \mathbf{P}_{k|k-1}^{b}\mathbf{H}_k^{H}\left(\mathbf{S}_{k|k}^{b}\right)^{-1} \tag{29b}$$

$$\mathbf{P}_{k|k}^{b} = \left(\mathbf{I} - \mathbf{K}_{k}^{b}\mathbf{H}_k\right)\mathbf{P}_{k|k-1}^{b}. \tag{29c}$$

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

$$\mathbf{I}_{k|k} = \left(\mathbf{P}_{k|k}^{b}\right)^{-1}, \quad \mathbf{I}_{k|k-1} = \left(\mathbf{P}_{k|k-1}^{b}\right)^{-1}. \tag{30}$$

Then, the usual form of the KF recursion (29a–c) can be rewritten in the following IF form [4, Sec. 6.2]:

$$\mathbf{I}_{k|k-1} = \mathbf{C}_{\mathbf{w}_{k-1}}^{-1} - \mathbf{C}_{\mathbf{w}_{k-1}}^{-1}\mathbf{F}_{k-1}\left(\mathbf{I}_{k-1|k-1} + \mathbf{F}_{k-1}^{H}\mathbf{C}_{\mathbf{w}_{k-1}}^{-1}\mathbf{F}_{k-1}\right)^{-1}\mathbf{F}_{k-1}^{H}\mathbf{C}_{\mathbf{w}_{k-1}}^{-1} \tag{31a}$$

$$\mathbf{I}_{k|k} = \mathbf{I}_{k|k-1} + \mathbf{H}_k^{H}\mathbf{C}_{\mathbf{v}_k}^{-1}\mathbf{H}_k \tag{31b}$$

$$\mathbf{K}_{k}^{b} = \mathbf{I}_{k|k}^{-1}\mathbf{H}_k^{H}\mathbf{C}_{\mathbf{v}_k}^{-1} \tag{31c}$$

where $\mathbf{I}_{0|0} = \mathbf{C}_{\mathbf{x}_0}^{-1}$ and $\widehat{\mathbf{x}}_{0|0}^{b} = E\left[\mathbf{x}_0\right]$. 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 $\mathbf{I}_{0|0} \rightarrow \mathbf{0}$, leading to

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

Moreover, if $\mathbf{F}_0$ is invertible, then $\mathbf{I}_{1|0} = \mathbf{0}$, $\mathbf{I}_{1|1} = \mathbf{H}_1^{H}\mathbf{C}_{\mathbf{v}_1}^{-1}\mathbf{H}_1$, $\mathbf{K}_1^{b} = \mathbf{I}_{1|1}^{-1}\mathbf{H}_1^{H}\mathbf{C}_{\mathbf{v}_1}^{-1}$, $\mathbf{K}_1^{b}\mathbf{H}_1 = \mathbf{I}$, and $\widehat{\mathbf{x}}_{1|1}^{b}$ does not depend on $\widehat{\mathbf{x}}_{0|0}^{b} = E\left[\mathbf{x}_0\right]$ any longer, since (2)

$$\widehat{\mathbf{x}}_{1|1}^{b} = \mathbf{F}_0\widehat{\mathbf{x}}_{0|0}^{b} + \mathbf{K}_1^{b}\left(\mathbf{y}_1 - \mathbf{H}_1\mathbf{F}_0\widehat{\mathbf{x}}_{0|0}^{b}\right) = \mathbf{K}_1^{b}\mathbf{y}_1.$$

Thus, if $\mathbf{F}_0$ is invertible, the use of a prior-free estimate of $\mathbf{x}_1$, obtained via the IF form (31a–c) coincides with the LMVDRF which does not depend neither on $E\left[\mathbf{x}_0\right]$ nor on $\mathbf{C}_{\mathbf{x}_0}$. However, if $\mathbf{F}_0$ if not full rank, provided that $\widehat{\mathbf{x}}_{0|0}^{b} = E\left[\mathbf{x}_0\right]$ is known and that the IF form (31a–c) exists, it should be used instead of the LMVDRF in absence of prior knowledge on $\mathbf{C}_{\mathbf{x}_0}$. Indeed, in this instance, $\mathbf{I}_{1|1}^{-1} \leq \left(\mathbf{H}_1^{H}\mathbf{C}_{\mathbf{v}_1}^{-1}\mathbf{H}_1\right)^{-1} = \mathbf{P}_{1|1}^{b}$ (w.r.t. the Löwner ordering [14, Sec. 7.7]), which implies that the MSE matrix of the IF will be less or equal than the MSE matrix of the LMVDRF at time $k \geq 2$ as well. Last, if $\mathbf{C}_{\mathbf{w}_0}$ is known and $E\left[\mathbf{x}_0\right]$ is unknown, then $\widehat{\mathbf{x}}_{1|1}^{b}$ can not be computed via the IF form, whereas the LMVDRF of $\mathbf{x}_1$ exists (if $\mathbf{H}_1$ is full rank). In a nutshell, in comparison with the IF form, the LMVDRF exists under more general assumptions (10b), which do not require $\mathbf{C}_{\mathbf{w}_{k-1}}$ and $\mathbf{C}_{\mathbf{v}_k}$ to be invertible. Moreover, the knowledge of $\mathbf{F}_0$, $\mathbf{C}_{\mathbf{w}_0}$, $E\left[\mathbf{x}_0\right]$, and $\mathbf{C}_{\mathbf{x}_0}$ is not required either.

D. Illustration of LMVDRF Properties

For the sake of illustration of the key properties of the LMVDRF, in the general case where $\mathbf{F}_k$, $\forall k$, is not invertible, we consider the following simple time varying LDSS model:

$$\left| \begin{array}{l} \mathbf{x}_{2k+1} = \mathbf{F}_{2k}x_{2k} + \mathbf{w}_{2k} \\ \mathbf{y}_{2k+1} = \mathbf{x}_{2k+1} + \mathbf{v}_{2k+1} \end{array} \right. \tag{32a}$$

$$\left| \begin{array}{l} x_{2k+2} = \mathbf{F}_{2k+1}\mathbf{x}_{2k+1} + w_{2k+1} \\ y_{2k+2} = x_{2k+2} + v_{2k+2} \end{array} \right. \tag{32b}$$

where $\mathbf{F}_{2k} = \mathbf{1}_{10}$ and $\mathbf{F}_{2k+1} = (0.01, 0.22, -0.62, \; -0.08, -0.33 - 0.35, 0.29, 0.08, 0.22, -0.44)$ are not invertible. The noise process, the measurement noise, and the initial state x_0 are uncorrelated. Moreover, assume that $E\left[x_0\right] = -1$, $C_{x_0} = 1$, $\mathbf{C}_{\mathbf{w}_{2k}} = \sigma_w^2\mathbf{I}$, $C_{w_{2k+1}} = \sigma_w^2$, $\sigma_w^2 = 10^{-4}$, $\mathbf{C}_{\mathbf{v}_{2k+1}} = \sigma_v^2\mathbf{I}$, $C_{v_{2k}} = \sigma_v^2$, $\sigma_v^2 = 100$. Fig. 1 highlights the consequence of a misspecification on $\{C_{x_0}, E\left[x_0\right]\}$ when one initializes the KF with wrong assumed values $\{C_{x_0} = 510^{-2}, E\left[x_0\right] = 0\}$. The empirical MSE of the various filters considered, namely the LMVDRF, the IF form of the KF and the KF, are assessed with 5×10^4 Monte-Carlo trials and are denoted "... (Simu)." The theoretical MSE of the LMVDRF and the IF are computed, respectively, from (27b–d) and (31a–c), and are denoted "... (Theo)." A very broad prior distribution on $\mathbf{x}_0$ is assumed for the

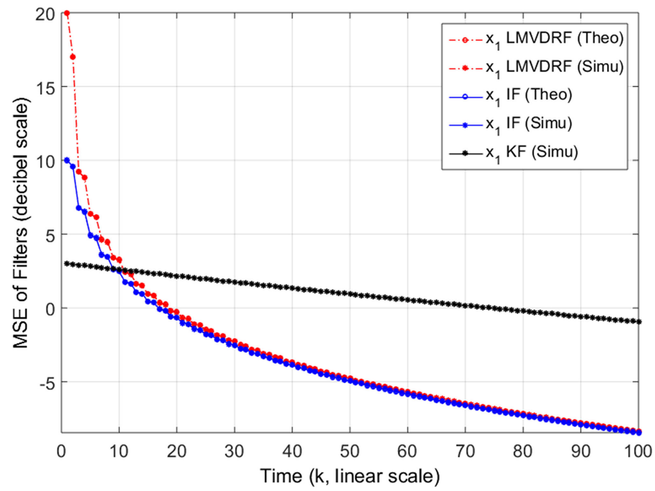

Fig. 1. Comparison of MSE of filters of $x_1 = x_{2k+2}$ and $x_1 = (\mathbf{x}_{2k+1})_1$.

IF ($\mathbf{I}_{0|0} = \mathbf{0}$). The software used to produce Fig. 1 is MATLAB. Fig. 1 clearly shows that, although the LMVDRF is suboptimal in terms of MSE when $\{\mathbf{C}_{\mathbf{x}_0}, E[\mathbf{x}_0]\}$ are perfectly known, in the presence of uncertainties on $\{\mathbf{C}_{\mathbf{x}_0}, E[\mathbf{x}_0]\}$, the LMVDRF may offer better performance than a KF wrongly initialized. Moreover, if $\mathbf{F}_0$ is not full rank ($\mathbf{F}_0 = \mathbf{1}_{10}$), provided that the IF exists and $E[x_0]$ is known, its use instead of the LMVDRF yields a lower (or equal) MSE.

VI. LMVDR Fixed-Point and Fixed-Lag Smoothers

Let us recall that the standard fixed-point smoother $\widehat{\mathbf{x}}^b_{l|k}$ is obtained by running the KF on the following augmented LDSS model [4, Sec. 9.2]:

$$
\begin{aligned}
k \le l &\quad \left|\begin{array}{l} \mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{w}_{k-1} \\ \mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \end{array}\right. \\
k = l+1 &\quad \left|\begin{array}{l} \begin{pmatrix}\mathbf{x}_{l+1}\\ \varkappa_{l+1}\end{pmatrix} = \begin{bmatrix}\mathbf{F}_l\\ \mathbf{I}\end{bmatrix}\mathbf{x}_l + \begin{pmatrix}\mathbf{w}_l\\ \mathbf{0}\end{pmatrix} \\ \mathbf{y}_l = \begin{bmatrix}\mathbf{H}_l & \mathbf{0}\end{bmatrix}\begin{pmatrix}\mathbf{x}_l\\ \varkappa_l\end{pmatrix} + \mathbf{v}_l \end{array}\right. \\
k \ge l+2 &\quad \left|\begin{array}{l} \begin{pmatrix}\mathbf{x}_k\\ \varkappa_k\end{pmatrix} = \begin{bmatrix}\mathbf{F}_{k-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\begin{pmatrix}\mathbf{x}_{k-1}\\ \varkappa_{k-1}\end{pmatrix} + \begin{pmatrix}\mathbf{w}_{k-1}\\ \mathbf{0}\end{pmatrix} \\ \mathbf{y}_k = \begin{bmatrix}\mathbf{H}_k & \mathbf{0}\end{bmatrix}\begin{pmatrix}\mathbf{x}_k\\ \varkappa_k\end{pmatrix} + \mathbf{v}_k \end{array}\right.
\end{aligned} \tag{33a}
$$

leading to $\widehat{\mathbf{x}}^b_{l|k} = \widehat{\varkappa}^b_{k|k}$ for $k \ge l+1$. Obviously, at time $l+1$, the state matrix of the augmented state is always non invertible, whatever $\mathbf{F}_l$ is invertible or not. Likewise, the standard fixed-lag smoother $\widehat{\mathbf{x}}^b_{k-N|k}$ [4, Sec. 9.3] is obtained by running the KF on an augmented system which state matrix is always noninvertible [4, eq. (9.41)]

$$
\begin{aligned}
k \le N &\quad \left|\begin{array}{l} \mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{w}_{k-1} \\ \mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \end{array}\right. \\
k > N &\quad \left|\begin{array}{l} \mathbf{z}_{k+1} = \begin{bmatrix}\mathbf{F}_k & \mathbf{0} & \dots & \mathbf{0}\\ \mathbf{I} & \mathbf{0} & \dots & \mathbf{0}\\ \vdots & \ddots & \ddots & \vdots\\ \mathbf{0} & \dots & \mathbf{I} & \mathbf{0}\end{bmatrix}\mathbf{z}_k + \begin{pmatrix}\mathbf{w}_k\\ \mathbf{0}\\ \vdots\\ \mathbf{0}\end{pmatrix} \\ \mathbf{y}_k = \begin{bmatrix}\mathbf{H}_k & \mathbf{0} & \dots & \mathbf{0}\end{bmatrix}\mathbf{z}_k + \mathbf{v}_k \end{array}\right.
\end{aligned} \tag{33b}
$$

where $\mathbf{z}_k^T = \left(\mathbf{x}_{k+1}^T, \mathbf{x}_{k+1,1}^T, \dots, \mathbf{x}_{k+1,N+1}^T\right)$, and $\widehat{\mathbf{x}}^b_{k+1,1|k} = \widehat{\mathbf{x}}^b_{k|k}$, $\widehat{\mathbf{x}}^b_{k+1,2|k} = \widehat{\mathbf{x}}^b_{k-1|k}, \dots, \widehat{\mathbf{x}}^b_{k+1,N+1|k} = \widehat{\mathbf{x}}^b_{k-N|k}$. Note that the conditions (10b) are satisfied for both the augmented LDSS models (33a–b) once they are satisfied for the initial LDSS model (1a–b). As a consequence, a major benefit of the relaxation on the conditions of existence of LMVDRF introduced here is the proof of the existence of the LMVDR fixed-point and fixed-lag smoothers [obtained by initializing the KF associated with (33a–b) with (3)], which cannot be proved with the result previously introduced in [2].

A. Generalization of the Deterministic Least-Squares Problem

In deterministic parameters estimation, one of the most studied estimation problem is that of identifying the components of measurements ($\mathbf{y}_1$) formed from a linear superposition of individual signals ($\mathbf{x}_1$) to noisy data ($\mathbf{v}_1$): $\mathbf{y}_1 = \mathbf{H}_1\mathbf{x}_1 + \mathbf{v}_1$, where the model matrix $\mathbf{H}_1$ and the noise covariance matrix $\mathbf{C}_{\mathbf{v}_1}$ are known, a.k.a. the linear regression problem. As mentioned in Section I, in this setting, the WLSE of $\mathbf{x}_1$ [11] coincides with the BLUE (a.k.a. the LMVDRE) of $\mathbf{x}_1$ [17]. These results still hold if k measurements of $\mathbf{x}_1$ are available: $\mathbf{y}_l = \mathbf{H}_l\mathbf{x}_1 + \mathbf{v}_l$, $1 \le l \le k$, $2 \le k$, and the measurement noise sequence $\{\mathbf{v}_l\}_{l=1}^k$ is temporally white. Indeed, the equivalent measurement model (5b) then becomes simply

$$
\overline{\mathbf{y}}_k = \overline{\mathbf{H}}_k\mathbf{x}_1 + \overline{\mathbf{v}}_k \tag{34}
$$

leading to the WLSE of $\mathbf{x}_1$ [3, Sec. 2.2.2]

$$
\begin{aligned}
\widehat{\mathbf{x}}^b_1(k) &= \arg\min_{\mathbf{x}_1}\left\{\left(\overline{\mathbf{y}}_k - \overline{\mathbf{H}}_k\mathbf{x}_1\right)^H \mathbf{C}^{-1}_{\overline{\mathbf{v}}_k}\left(\overline{\mathbf{y}}_k - \overline{\mathbf{H}}_k\mathbf{x}_1\right)\right\} \\
&= \arg\min_{\mathbf{x}_1}\left\{\sum_{l=1}^{k}\left(\mathbf{y}_l - \mathbf{H}_l\mathbf{x}_1\right)^H \mathbf{C}^{-1}_{\mathbf{v}_l}\left(\mathbf{y}_l - \mathbf{H}_l\mathbf{x}_1\right)\right\}
\end{aligned} \tag{35}
$$

which batch form solution is given by the following:

$$
\widehat{\mathbf{x}}^b_1(k) = \overline{\mathbf{H}}_k\left(\overline{\mathbf{H}}_k^H\mathbf{C}^{-1}_{\overline{\mathbf{v}}_k}\overline{\mathbf{H}}_k\right)^{-1}\overline{\mathbf{H}}_k^H\mathbf{C}^{-1}_{\overline{\mathbf{v}}_k}\overline{\mathbf{y}}_k. \tag{36}
$$

Moreover, if the measurement noise sequence $\{\mathbf{v}_l\}_{l=1}^k$ is Gaussian, the WLSE of $\mathbf{x}_1$ (35) coincides with the maximum-likelihood estimator [8] of $\mathbf{x}_1$ as well. As shown in [2], since usual assumptions on the recursive WLSE verify: $\mathbf{w}_k = \mathbf{0}$ and $\mathbf{C}_{\mathbf{v}_l,\mathbf{v}_k} = \mathbf{C}_{\mathbf{v}_k}\delta^l_k$, they verify (10b) as well, and (36) can also be computed recursively since it is a special case of the LMVDRF [2]. The importance of the linear regression 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 [3], [4], [8], [15]. In the standard WLSE, the individual signals $\mathbf{x}_1$ are assumed to remain perfectly constant during the k measurements. However, in a real-life experiment, some experimental factors may prevent from observing perfectly constant individual signals $\mathbf{x}_1$. For instance, in any problem dealing with signal transmission involving a transmitter device and a propagation medium, the transmitter noise may not be negligible and the fluctuation of the propagation medium are sometime unavoidable during the whole measurement time interval. These factors, and others, can be taken into account globally by introducing a random fluctuation from measurement to measurement

$$
\mathbf{x}_l = \mathbf{F}_{l-1}\mathbf{x}_{l-1} + \mathbf{w}_{l-1},\ 2 \le l \le k
$$

yielding a generalized form of (34), which consists of the class of LDSS models defined as follows:

$$
k \ge 1: \left|\begin{array}{l} \mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{w}_{k-1} \\ \mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \end{array}\right. \tag{37a}
$$

$$
\mathbf{C}_{\mathbf{x}_0} = \mathbf{0},\ \mathbf{C}_{\mathbf{w}_0} = \mathbf{0},\ \mathbf{F}_0 = \mathbf{I},\ \mathbf{x}_1 = \mathbf{x}_0 = E[\mathbf{x}_0]. \tag{37b}
$$

In this setting, (34) becomes (5b), that is, $\overline{\mathbf{y}}_k = \overline{\mathbf{A}}_k \mathbf{x}_1 + \overline{\mathbf{n}}_k$, which leads to the generalized WLSE of $\mathbf{x}_1$ (GWLSE) defined as [8, Sec. 6], [9, Sec. 5.6]

$$\widehat{\mathbf{x}}_1^b(k) = \arg\min_{\mathbf{x}_1} \left\{ \left(\overline{\mathbf{y}}_k - \overline{\mathbf{A}}_k \mathbf{x}_1\right)^H \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1} \left(\overline{\mathbf{y}}_k - \overline{\mathbf{A}}_k \mathbf{x}_1\right) \right\}$$

whose batch form solution is given by the following:

$$\widehat{\mathbf{x}}_1^b(k) = \mathbf{P}_1^b(k)\, \overline{\mathbf{A}}_k^H \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1} \overline{\mathbf{y}}_k,\ \mathbf{P}_1^b(k) = \left(\overline{\mathbf{A}}_k^H \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1} \overline{\mathbf{A}}_k\right)^{-1} \quad (38)$$

where $\mathbf{P}_1^b(k)$ denotes the MSE matrix of the GWLSE $\widehat{\mathbf{x}}_1^b(k)$. If the batch form (38) has the merit of offering a closed-form expression, it nevertheless suffers from two significant drawbacks. First, if $\overline{\mathbf{n}}_k$ is not block diagonal, then the determination of $\mathbf{C}_{\overline{\mathbf{n}}_k}^{-1}$ becomes computationally prohibitive as the number of observations k increases. Second, (38) is not compatible with real-world applications [4], [18] 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). Fortunately, a recursive form of $\widehat{\mathbf{x}}_1^b(k)$ (38) exists provided that (10b) are satisfied. Indeed, $\widehat{\mathbf{x}}_1^b(k)$ (38) coincides with the LMVDR fixed-point smoother of $\mathbf{x}_1$ as shown hereinafter.

First, let us consider the augmented LDSS model (33a), where $l = 1$, which can be recasted in a more compact form as follows:

$$\left| \begin{array}{l} \mathbf{x}'_1 = \mathbf{x}_1 \\ \mathbf{y}_1 = \mathbf{H}'_1 \mathbf{x}'_1 + \mathbf{v}_1 \end{array} \right., \quad \left| \begin{array}{l} \mathbf{x}'_l = \mathbf{F}'_{l-1} \mathbf{x}'_{l-1} + \mathbf{w}'_{l-1} \\ \mathbf{y}_l = \mathbf{H}'_l \mathbf{x}'_l + \mathbf{v}_l \end{array} \right. \quad (39)$$

where $\mathbf{H}'_1 = \mathbf{H}_1$, and (5a) becomes

$$\mathbf{y}_l = \mathbf{A}'_l \mathbf{x}_1 + \mathbf{n}'_l,\ \mathbf{A}'_l = \mathbf{H}'_l \mathbf{B}'_{l,1},\ \left| \begin{array}{l} \mathbf{n}'_1 = \mathbf{v}_1 \\ \mathbf{n}'_l = \mathbf{v}_l + \mathbf{H}'_l \mathbf{G}'_l \overline{\mathbf{w}}'_{l-1} \end{array} \right. .$$

By definition

$$\mathbf{B}'_{l,1} = \mathbf{F}'_{l-1} \cdots \mathbf{F}'_2 \mathbf{F}'_1 = \begin{bmatrix} \mathbf{B}_{l,2} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{F}_1 \\ \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{B}_{l,1} \\ \mathbf{I} \end{bmatrix}$$

$$\mathbf{A}'_l = \mathbf{H}'_l \mathbf{B}'_{l,1} = \begin{bmatrix} \mathbf{H}_l & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{l,1} \\ \mathbf{I} \end{bmatrix} = \mathbf{H}_l \mathbf{B}_{l,1} = \mathbf{A}_l.$$

Moreover, since $\mathbf{G}'_l \overline{\mathbf{w}'}_{l-1} = \sum_{q=1}^{l-1} \mathbf{B}'_{l,q+1} \mathbf{w}'_q = \binom{\mathbf{G}_l \overline{\mathbf{w}}_{l-1}}{\mathbf{0}}$, then

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

Second, since $\mathbf{H}'_1$ and $\mathbf{C}_{\overline{\mathbf{y}}_k}$ are full rank, if we consider the LDSS model (39), the LMVDRF of $\mathbf{x}'_k$ exists and is defined by (17)

$$\overline{\mathbf{W}}_k^b = \arg\min_{\overline{\mathbf{W}}_k} \left\{ \mathbf{P}_{k|k}\left(\overline{\mathbf{W}}_k\right) \right\} \text{ s.t. } \overline{\mathbf{W}}_k^H \overline{\mathbf{A}}'_k = \mathbf{B}'_{k,1} \quad (40a)$$

where $\mathbf{P}_{k|k}\left(\overline{\mathbf{W}}_k\right) = E[(\overline{\mathbf{W}}_k^H \overline{\mathbf{y}}_k - \mathbf{x}'_k)(\overline{\mathbf{W}}_k^H \overline{\mathbf{y}}_k - \mathbf{x}'_k)^H]$, which, according to (16), is equivalent to

$$\overline{\mathbf{W}}_k^b = \arg\min_{\overline{\mathbf{W}}_k} \left\{ E\left[\widehat{\mathbf{r}}_k \widehat{\mathbf{r}}_k^H\right] \right\} \text{ s.t. } \overline{\mathbf{W}}_k^H \overline{\mathbf{A}'}_k = \mathbf{B}'_{k,1}$$

$$\widehat{\mathbf{r}}_k = \overline{\mathbf{W}}_k^H \overline{\mathbf{n}}'_k - \mathbf{G}'_k \overline{\mathbf{w}'}_{k-1},\ \overline{\mathbf{W}}_k = \left[\overline{\mathbf{W}}_k^{\mathbf{x}}\ \overline{\mathbf{W}}_k^{\varkappa}\right]. \quad (40b)$$

Since $\overline{\mathbf{W}}_k^b$ is analogous to a LCWF [15, Sec. 2.5], its batch form is given by [15, Sec. 2]

$$\mathbf{C}_{\overline{\mathbf{n}}'_k} \overline{\mathbf{W}}_k^b = \overline{\mathbf{A}'}_k \left(\overline{\mathbf{A}'}_k^H \mathbf{C}_{\overline{\mathbf{n}}'_k}^{-1} \overline{\mathbf{A}'}_k\right)^{-1} \left(\mathbf{B}'_{k,1}\right)^H + \left(\mathbf{I} - \overline{\mathbf{A}'}_k \left(\overline{\mathbf{A}'}_k^H \mathbf{C}_{\overline{\mathbf{n}}'_k}^{-1} \overline{\mathbf{A}'}_k\right)^{-1} \overline{\mathbf{A}'}_k^H \mathbf{C}_{\overline{\mathbf{n}}'_k}^{-1}\right) \mathbf{C}_{\overline{\mathbf{n}}'_k, \mathbf{G}'_k \overline{\mathbf{w}}'_{k-1}} \quad (41a)$$

that is, as $\overline{\mathbf{n}}'_k = \overline{\mathbf{n}}_k$ and $\overline{\mathbf{A}}'_k = \overline{\mathbf{A}}_k$

$$\mathbf{C}_{\overline{\mathbf{n}}_k} \overline{\mathbf{W}}_k^b = \overline{\mathbf{A}}_k \left(\overline{\mathbf{A}}_k^H \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1} \overline{\mathbf{A}}_k\right)^{-1} \left[\mathbf{B}_{k,1}^H\ \mathbf{I}\right] + \left(\mathbf{I} - \overline{\mathbf{A}}_k \left(\overline{\mathbf{A}}_k^H \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1} \overline{\mathbf{A}}_k\right)^{-1} \overline{\mathbf{A}}_k^H \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1}\right) \left[\mathbf{C}_{\overline{\mathbf{n}}_k, \mathbf{G}_k \overline{\mathbf{w}}_{k-1}}\ \mathbf{0}\right]. \quad (41b)$$

Therefore, (40a–b) yields the following separable solutions:

$$\left(\overline{\mathbf{W}}_k^{\mathbf{x}}\right)^b = \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1} \overline{\mathbf{A}}_k \left(\overline{\mathbf{A}}_k^H \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1} \overline{\mathbf{A}}_k\right)^{-1} \mathbf{B}_{k,1}^H + \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1} \left(\mathbf{I} - \overline{\mathbf{A}}_k \left(\overline{\mathbf{A}}_k^H \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1} \overline{\mathbf{A}}_k\right)^{-1} \overline{\mathbf{A}}_k^H \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1}\right) \mathbf{C}_{\overline{\mathbf{n}}_k, \mathbf{G}_k \overline{\mathbf{w}}_{k-1}}$$

$$\left(\overline{\mathbf{W}}_k^{\varkappa}\right)^b = \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1} \overline{\mathbf{A}}_k \left(\overline{\mathbf{A}}_k^H \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1} \overline{\mathbf{A}}_k\right)^{-1} \quad (42)$$

leading to

$$\widehat{\varkappa}_{k|k}^b = \left(\overline{\mathbf{A}}_k^H \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1} \overline{\mathbf{A}}_k\right)^{-1} \overline{\mathbf{A}}_k^H \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1} \overline{\mathbf{y}}_k = \widehat{\mathbf{x}}_1^b(k) \quad (43a)$$

$$E\left[\left(\widehat{\varkappa}_{k|k}^b - \widehat{\varkappa}_{k|k}\right)\left(\widehat{\varkappa}_{k|k}^b - \widehat{\varkappa}_{k|k}\right)^H\right] = \left(\overline{\mathbf{A}}_k^H \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1} \overline{\mathbf{A}}_k\right)^{-1} \triangleq \mathbf{P}_1^b(k). \quad (43b)$$

Third, if the LDSS models associated to the GWLSE (37a–b) satisfies the conditions (10b), then the solution of (40a–b) can also be computed recursively as follows:

$$\widehat{\mathbf{x}}_1^b(k) = [\mathbf{0}\ \mathbf{I}]\, \widehat{\mathbf{x}'}_{k|k}^b,\ \mathbf{P}_1^b(k) = [\mathbf{0}\ \mathbf{I}]\, \mathbf{P}_{k|k}^b \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \quad (44)$$

where $\widehat{\mathbf{x}'}_{k|k}^b$ and $\mathbf{P}_{k|k}^b$ follow the LMVDRF's recursion (27b–d). Last, the extension of this result to a regularized GWLSE (RGWLSE)

$$\widehat{\mathbf{x}}_1^b(k) = \arg\min \left\{ (\mathbf{c} - \mathbf{x}_1)^H \mathbf{\Lambda}_0^{-1} (\mathbf{c} - \mathbf{x}_1) + \left(\overline{\mathbf{y}}_k - \overline{\mathbf{A}}_k \mathbf{x}_1\right)^H \mathbf{C}_{\overline{\mathbf{n}}_k}^{-1} \left(\overline{\mathbf{y}}_k - \overline{\mathbf{A}}_k \mathbf{x}_1\right) \right\}$$

where $\mathbf{\Lambda}_0$ is an Hermitian invertible matrix is simply obtained by adding a fictitious measurement at time $k = 0$

$$\mathbf{y}_0 = \mathbf{H}_0 \mathbf{x}_0 + \mathbf{v}_0,\ \mathbf{C}_{\mathbf{v}_0} = \mathbf{\Lambda}_0,\ \mathbf{y}_0 = \mathbf{c},\ \mathbf{H}_0 = \mathbf{I}$$
$$k \geq 1 : \mathbf{C}_{\mathbf{v}_0, \mathbf{v}_k} = \mathbf{0},\ \mathbf{C}_{\mathbf{v}_0, \mathbf{w}_{k-1}} = \mathbf{0}$$

and by starting the fixed-point smoother recursion (33a) at time $k = 0$

$$\widehat{\mathbf{x}}_{0|0}^b = \mathbf{P}_{0|0}^b \mathbf{H}_0^H \mathbf{C}_{\mathbf{v}_0}^{-1} \mathbf{y}_0 = \mathbf{c},\ \mathbf{P}_{0|0}^b = \left(\mathbf{H}_0^H \mathbf{C}_{\mathbf{v}_0}^{-1} \mathbf{H}_0\right)^{-1} = \mathbf{\Lambda}_0$$

instead of time $k = 1$.

1) Illustrative Example: We consider the reception by a uniform linear array of N sensors equally spaced at half-wavelength of an impinging signal of amplitude x_1 (which includes the transmission power, the propagation loss, and signal processing gains of the receivers) with broadside angle θ, embedded in a spatially and temporally white noise. In a stationary propagation medium, the N received signals can be modeled at the output of the Hilbert filter as [8]: $\mathbf{y}_l = \mathbf{h}(\theta)\, x_1 + \mathbf{v}_l$, $1 \leq l \leq k$, $\mathbf{h}^T(\theta) = \left(1, e^{j\pi \sin(\theta)}, \ldots, e^{j\pi(N-1)\sin(\theta)}\right)$. However, if x_1 is transmitted via a nonstationary propagation medium, fluctuating randomly from measurement to measurement, then the N received signals can be modeled (to first-order) as follows:

$$x_l = x_{l-1} + w_{l-1},\ \mathbf{y}_l = \mathbf{h}(\theta)\, x_l + \mathbf{v}_l,\ 1 \leq l \leq k \quad (45)$$

where the fluctuation noise sequence $\{w_l\}_{l=1}^{k-1}$ is white and uncorrelated with the measurement noise sequence $\{v_l\}_{l=1}^{k}$. In Fig. 2, we consider

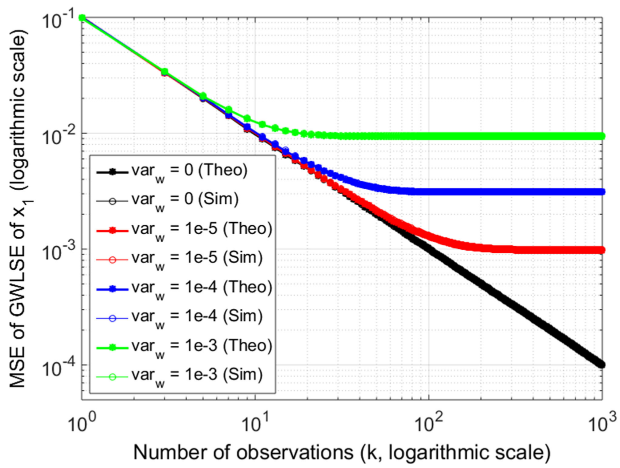

Fig. 2. MSE of the GWLSE of x_1 for $\sigma_w^2 \in \left\{10^{-5}, 10^{-4}, 10^{-3}\right\}$.

the scenario where $N = 10$, $x_1 = (1+j)/\sqrt{2}$, $\mathbf{C}_{\mathbf{v}_l} = \mathbf{I}$, $\theta = 1/4$ and $k = 1000$. We superimpose on Fig. 2 both the theoretical MSE of the GWLSE of x_1 denoted "...(Theo)" and the empirical MSE assessed with 5×10^4 Monte-Carlo trials and denoted "...(Sim)." The empirical GWLSE $\widehat{x}_1^b(k)$ of x_1 and its theoretical MSE $\mathbf{P}_1^b(k)$ are computed recursively according to (44) where the fluctuation noise sequence has a constant variance $\sigma_{w_l}^2 = \sigma_w^2 \in \{10^{-5}, 10^{-4}, 10^{-3}\}$ from observation to observation. The software used to produce Fig. 2 is MATLAB. Fig. 2 exemplifies the nonnegligible impact of a slight fluctuation of the unknown parameter on the WLSE asymptotic performance which introduces a lower limit in the achievable MSE. From a practical point of view, the existence of this lower limit shows that, if the unknown parameter is not perfectly constant during the k measurements, there exists an optimal number of observations that can be combined in order to estimate it with the minimum (or almost minimum) achievable MSE.

VII. Conclusion

By relaxing the conditions of existence of LMVDRFs, 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). From a general perspective, although the LMVDR estimators are suboptimal in terms of MSE, they have two merits, which are as follows:

1) They do not depend on the prior knowledge on $\mathbf{x}_0$.
2) They may outperform the usual LLMS estimators in case of misspecification of the prior knowledge on $\mathbf{x}_0$.

These features are quite interesting for filter/smoother/predictor performance analysis and design since they allow to synthesize infinite impulse response distortionless estimators which performance are robust to an unknown initial state. On another note, LMVDR estimators may allow to derive unexpected results, as highlighted with the link between the LMVDR fixed-point smoother and a generalized WLSE.

References

[1] R. Kalman, "A new approach to linear filtering and prediction problems," *ASME J. Basic Eng.*, vol. 82, pp. 35–45, 1960.

[2] E. Chaumette, B. Priot, F. Vincent, G. Pages, and A. Dion, "Minimum variance distortionless response estimators for linear discrete state-space models," *IEEE Trans. Autom. Control*, vol. 62, no. 4, pp. 2048–2055, Apr. 2017.

[3] T. Kailath, A. Sayed, and B. Hassibi, *Linear Estimation*. Upper Saddle River, NJ, USA: Prentice-Hall, 2000.

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

[5] B. P. Gibbs,*Advanced Kalman Filtering Least-Squares and Modeling*. Hoboken, NJ, USA: Wiley, 2011.

[6] D. E. Catlin, "Estimation of random states in general linear models," *IEEE Trans. Autom. Control*, vol. 36, no. 2, pp. 248–252, Feb. 1991.

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

[8] H. L. Van Trees, *Optimum Array Processing*. Hoboken, NJ, USA: Wiley-Interscience, 2002.

[9] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data*. Cambridge, U.K.: Cambridge Univ. Press, 2010.

[10] S. A. Vorobyov, "Principles of minimum variance robust adaptive beamforming design," *Elsevier Signal Process.*, vol. 93, pp. 3264–3277, 2013.

[11] R. L. Plackett, "Some theorems in least squares," *Biometrika*, vol. 37, pp. 149–157, 1950.

[12] E. Chaumette, B. Priot, F. Vincent, G. Pages, and A. Dion, "Minimum variance distortionless response estimators for linear discrete state-space models: Mathematical appendix," Institut supérieur de l'aéronautique et de l'espace, Toulouse, France, Res. memorandum, ISAE Supaéro, 2016.

[13] N. Wiener, *The Extrapolation, Interpolation and Smoothing of Stationary Time Series*. New York, NY, USA: Wiley, 1949.

[14] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2013.

[15] Paulo S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, 4th ed. Berlin, Germany: Springer, 2013.

[16] G. A. F. Seber, *Matrix Handbook Forr Statisticians (Wiley Series in Probability and Statistics)*. Hoboken, NJ, USA: Wiley, 2008.

[17] O. L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, no. 8, pp. 926–935, Aug. 1972.

[18] J. L. Crassidis and J. L. Junkins, *Optimal Estimation of Dynamic Systems*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2012.