



Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: <http://oatao.univ-toulouse.fr/>
Eprints ID: 16658

To link this article: <http://dx.doi.org/10.1109/TAC.2016.2594384>

To cite this version: Chaumette, Eric and Priot, Benoit and Vincent, François and Pages, Gael and Dion, Arnaud *Minimum Variance Distortionless Response Estimators For Linear Discrete State-Space Models*. (2016) IEEE Transactions on Automatic Control, vol. 11 (n° 4). ISSN 0018-9286

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

Minimum Variance Distortionless Response Estimators For Linear Discrete State-Space Models

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

Abstract—For linear discrete state-space models, under certain conditions, the linear least-mean-squares filter estimate has a convenient recursive predictor/corrector format, aka the Kalman filter. The purpose of this paper is to show that the linear minimum variance distortionless response (MVDR) filter shares exactly the same recursion, except for the initialization which is based on a weighted least-squares estimator. If the MVDR filter is sub-optimal in mean-squared error sense, it is an infinite impulse response distortionless filter (a deconvolver) which does not depend on the prior knowledge (first and second order statistics) on the initial state. In other words, the MVDR filter can be pre-computed and its behaviour can be assessed in advance independently of the prior knowledge on the initial state.

Index Terms—State estimation, unbiased filter, filtering, smoothing, minimum mean-squared error upper bound.

I. INTRODUCTION

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

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

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \quad (1b)$$

where the time index $k \geq 1$, \mathbf{x}_k is the P -dimensional state vector, \mathbf{y}_k is the N -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 containing elements with finite modulus. The process and the measurement noise sequences have zero-mean values and the initial state has a finite 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 $\hat{\mathbf{x}}_{k|l} \triangleq \hat{\mathbf{x}}_{k|l}(\mathbf{y}_1, \dots, \mathbf{y}_l)$ 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 the process noise sequence, the measurement noise sequence

and the initial state \mathbf{x}_0 verify certain uncorrelation conditions [2, §9.1][3, §7.1][4, §8.2] and are Gaussian, the minimum variance or minimum mean-squared error (MSE) filter estimate for LDSS models has a convenient recursive predictor/corrector format, so-called the Kalman filter (KF). Even if the noise is non-Gaussian, the KF is the linear least-mean-squares (LLMS) filter (LLMSF) estimate.

Interestingly enough, the Kalman (stochastic) filtering problem has been related to several variants of deterministic least-squares estimation, i.e. the estimation of \mathbf{x}_0 for a LDSS model (1a-1b) where $\mathbf{x}_k = \mathbf{x}_{k-1} = \dots = \mathbf{x}_0$ and \mathbf{x}_0 is deterministic and unknown [2, §2.6][3, §3][4, §4][5, §4.12][6]. In the deterministic framework, under Gaussian measurement noise, the weighted least-squares estimator (WLSE) is identical to the maximum likelihood estimator (MLE). Even if the measurement noise is non-Gaussian, the WLSE coincides with the best linear unbiased estimator (BLUE) [4, §4] aka the linear minimum variance distortionless response estimator (LMVDRE) [7, §6][8, §5.6], which is a particular case of constrained linear estimation [2, p98]. The LMVDRE is a well known sub-optimal technique in the deterministic framework used to circumvent the problem of the large computational cost of the MLE [7] and its sensitivity to imperfect, incomplete, or erroneous knowledge about the observation model [9]. For instance, in the fields of radar, sonar, and wireless communication, it is common place to design a LMVDRE for the most studied estimation problem: that of separating the components of data formed from a linear superposition of individual signals to noisy data [7][8][9]. This is the reason why, sometimes, the LMVDRE is also called a deconvolution filter [7, §6][8, §5.6].

The main novel contribution of the present paper is the proof that in the stochastic framework, for the fairly general subset of LDSS models for which \mathbf{F}_k , $k \geq 1$, are invertible and \mathbf{H}_1 is full rank, the LMVDR filter (LMVDRF) shares the same recursion as the KF, except at time $k = 1$. This surprising result is worth knowing for users of the KF since, each time one initializes the KF recursion at time $k = 1$ with the WLSE associated to the measurement model (1b), i.e. $\hat{\mathbf{x}}_{1|1} = \mathbf{P}_{1|1} \mathbf{H}_1^H \mathbf{C}_{\mathbf{v}_1}^{-1} \mathbf{y}_1$ and $\mathbf{P}_{1|1} = (\mathbf{H}_1^H \mathbf{C}_{\mathbf{v}_1}^{-1} \mathbf{H}_1)^{-1}$, because the first and second order statistics of the initial state \mathbf{x}_0 are not available, one does not implement any longer the KF but the LMVDRF. Although the LMVDRF is sub-optimal in terms of MSE, it has a number of merits: a) it does not depend on the prior knowledge (first and second order statistics) on the initial state, b) it is an upper bound on the performance of the KF whatever the prior knowledge on the initial state, c) if the LDSS is time-invariant, it shares the same discrete algebraic

Eric Chaumette, Benoit Priot, François Vincent, Gael Pages and Arnaud Dion are with ISAE-SUPAERO, Université de Toulouse, 31055 Toulouse Cedex 4, France. Emails: eric.chaumette@isae.fr, benoit.priot@isae.fr, françois.vincent@isae.fr, gael.pages@isae.fr, arnaud.dion@isae.fr.

Riccati equation (DARE) as the KF. These features are quite interesting for filtering performance analysis and design of a LDSS system since they allow to synthesize an infinite impulse response (IIR) distortionless filter (a deconvolver) which performance is robust to an unknown initial state. However, it seems likely that, since the LMVDRF shares the same recursion as the KF, it also shares the same sensitivity to modelling error [3, §10][10] or uncertainties in the system matrices [11][12][13]. The robustification of the KF to the presence of mismodeling has been reinvestigated lately by using unbiased finite impulse response (UFIR) [14], p-shift FIR [15][16] or minimum variance UFIR [17] 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 LMVDRF 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.

The rest of the paper is organized as follows. Notations and signal model (joint proper complex) are introduced in Section II. In Section III, for sake of clarity, we give the main points of background knowledge on LLMS and LMVDR estimators required to discuss the filtering equations in the next Section. In section IV, we first highlight the general assumptions required on LDSS models to obtain a LLMSF satisfying the same predictor/corrector format as the KF (without extension of the state and measurement equations); then we demonstrate that the LMVDRF shares the same recursion as the KF, except at time $k = 1$. This result allows to show in Section V, that the two filters share as well, for the so-called "standard LDSS model", the same possible extensions (prediction, smoothing, filtering with fading memory, constrained filtering,...) and alternate formulations (information filtering, innovations approach,...). Last, the LMVDRF's properties and performance (in comparison with the KF) are illustrated with an example in Section VI.

II. NOTATIONS AND SIGNAL MODEL

The notational convention adopted is as follows: we shall use italic, small boldface and capital boldface letters to denote respectively scalars, column vectors and matrices. $\mathcal{M}_{\mathbb{C}}(N, P)$ denotes the vector space of complex matrices with N rows and P columns. The scalar/matrix/vector transpose, conjugate and transpose conjugate are respectively indicated by the superscripts T , $*$ and H . \mathbf{I} is the identity matrix. $[\mathbf{A} \ \mathbf{B}]$ denotes the matrix resulting from the horizontal concatenation of \mathbf{A} and \mathbf{B} . $S = \text{span}\{\mathbf{A}\}$ denotes the linear span of the set of the column vectors of \mathbf{A} . $\mathbf{A} \geq \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is positive semi-definite. If $\mathbf{C}^H = \mathbf{C}$ and $\mathbf{C} > \mathbf{0}$, $\Pi_{\mathbf{A}}^{\mathbf{C}} = \mathbf{A}(\mathbf{A}^H \mathbf{C} \mathbf{A})^{-1} \mathbf{A}^H \mathbf{C}$ denotes the orthogonal projection matrix on $\text{span}\{\mathbf{A}\}$ and $\mathbf{A} \perp^{\mathbf{C}} \mathbf{B}$ means that \mathbf{A} and \mathbf{B} are orthogonal, both for the Hermitian inner product defined by \mathbf{C} . S^{\perp} denotes the orthogonal complement of the subspace S for the canonical Hermitian inner product. $E[\cdot]$ denotes the expectation operator. If \mathbf{x} is a random vector, $\mathbf{m}_{\mathbf{x}} \triangleq E[\mathbf{x}]$ is the expectation of \mathbf{x} . If \mathbf{x} and \mathbf{y} are two complex random vectors: a) $\mathbf{R}_{\mathbf{x}} = E[\mathbf{x}\mathbf{x}^H]$, $\mathbf{R}_{\mathbf{y}} = E[\mathbf{y}\mathbf{y}^H]$ and $\mathbf{R}_{\mathbf{x},\mathbf{y}} = E[\mathbf{x}\mathbf{y}^H]$ are respectively the

correlation matrices of \mathbf{x} , of \mathbf{y} and the cross-correlation matrix of \mathbf{x} and \mathbf{y} ; b) $\mathbf{C}_{\mathbf{x}} = \mathbf{R}_{\mathbf{x}} - \mathbf{m}_{\mathbf{x}}\mathbf{m}_{\mathbf{x}}^H$, $\mathbf{C}_{\mathbf{y}} = \mathbf{R}_{\mathbf{y}} - \mathbf{m}_{\mathbf{y}}\mathbf{m}_{\mathbf{y}}^H$ and $\mathbf{C}_{\mathbf{x},\mathbf{y}} = \mathbf{R}_{\mathbf{x},\mathbf{y}} - \mathbf{m}_{\mathbf{x}}\mathbf{m}_{\mathbf{y}}^H$ are respectively the covariance matrices of \mathbf{x} , of \mathbf{y} and the cross-covariance matrix of \mathbf{x} and \mathbf{y} ; c) if $\mathbf{R}_{\mathbf{y}}$ is invertible, then $\mathbf{R}_{\mathbf{x}|\mathbf{y}} \triangleq \mathbf{R}_{\mathbf{x}} - \mathbf{R}_{\mathbf{x},\mathbf{y}}\mathbf{R}_{\mathbf{y}}^{-1}\mathbf{R}_{\mathbf{x},\mathbf{y}}^H$; d) If $\mathbf{C}_{\mathbf{y}}$ is invertible, then $\mathbf{C}_{\mathbf{x}|\mathbf{y}} \triangleq \mathbf{C}_{\mathbf{x}} - \mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1}\mathbf{C}_{\mathbf{x},\mathbf{y}}^H$.

A. Joint proper complex signals

As in [2, §3] and [8, §5.1], we adopt a joint proper (proper and cross-proper) complex signals assumption for the set of vector $(\mathbf{x}_0, \{\mathbf{w}_k\}, \{\mathbf{v}_k\})$ which allows to resort to standard estimation in the mean square error (MSE) sense defined on the Hilbert space of complex random variables with finite second-order moment. A proper complex random variable is uncorrelated with its complex conjugate [8], and a zero mean proper complex random vector is said to be second-order circular [2, §3.2.5]. Indeed, in diverse fields such as communications, control, and signal processing, many practical problems of interest modelize the observation vector as complex and consisting of a bandpass signal which is the output of an Hilbert filtering leading to an "in-phase" real part associated to a "quadrature" imaginary part. These observation vectors belongs to the general class of analytic signals for which a fundamental result states that wide-sense stationary analytic signals, and also complex baseband representations of wide-sense stationary real bandpass signals, must be proper [8, §2.6.1]. Besides its capability to characterize the statistical behavior of the underlying physics of these problems of interest, the proper complex signals assumption also simplify computations and, in many aspects, make complex random signals look and behave like real random signals. Indeed, 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 transpose²[2, §3.2.5][8, §5.4.1].

B. Equivalent linear observation model

Here, $\mathbf{F}_{k-1} \in \mathcal{M}_{\mathbb{C}}(P, P)$ and $\mathbf{H}_k \in \mathcal{M}_{\mathbb{C}}(N, P)$. First, as (1a) can be rewritten as:

$$\mathbf{x}_k = \mathbf{B}_{k,0}\mathbf{x}_0 + \sum_{l=0}^{k-1} \mathbf{B}_{k,l+1}\mathbf{w}_l, \mathbf{B}_{k,l} = \begin{cases} \mathbf{F}_{k-1}\mathbf{F}_{k-2}\dots\mathbf{F}_l, & k > l \\ \mathbf{I} & k = l \\ \mathbf{0} & k < l \end{cases}$$

an equivalent form of (1b) is:

$$\mathbf{y}_k = \mathbf{A}_k\mathbf{x}_0 + \mathbf{n}_k, \mathbf{A}_k = \mathbf{H}_k\mathbf{B}_{k,0}, \mathbf{n}_k = \sum_{l=0}^{k-1} \mathbf{H}_k\mathbf{B}_{k,l+1}\mathbf{w}_l + \mathbf{v}_k. \quad (2)$$

Second, (2) can be extended on a horizon of k points from the first observation as:

$$\bar{\mathbf{y}}_k = \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_k \end{pmatrix} = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_k \end{bmatrix} \mathbf{x}_0 + \begin{pmatrix} \mathbf{n}_1 \\ \vdots \\ \mathbf{n}_k \end{pmatrix} = \bar{\mathbf{A}}_k\mathbf{x}_0 + \bar{\mathbf{n}}_k, \quad (3)$$

¹Note that if \mathbf{x} and \mathbf{y} are (proper) normal complex random vectors, then $\mathbf{C}_{\mathbf{x}|\mathbf{y}}$ is exactly the covariance matrix of \mathbf{x} given \mathbf{y} .

²The case where $(\mathbf{x}_0, \{\mathbf{w}_k\}, \{\mathbf{v}_k\})$ are improper complex random vector or hybrid vector (real and complex components) is addressed in Section V-D.

where $\bar{\mathbf{y}}_k, \bar{\mathbf{n}}_k \in \mathcal{M}_{\mathbb{C}}(Nk, 1)$ and $\bar{\mathbf{A}}_k \in \mathcal{M}_{\mathbb{C}}(Nk, P)$.

III. BACKGROUND ON LINEAR/AFFINE ESTIMATORS

Let us recall that the hypothesis of zero-mean process and measurement noise sequences is equivalent to the assumption of non-zero but known mean values³: it suffices to center the random variables before processing. Since this transformation is reversible, there is no loss of information in making such a preprocessing [2, p84].

A. Linear least-mean-squares estimator (LLMSE)

Let us consider two complex random vectors \mathbf{x} and \mathbf{y} . The error between the signal \mathbf{x} and the affine estimator $\hat{\mathbf{x}} \triangleq \hat{\mathbf{x}}(\mathbf{y}) = \mathbf{K}\mathbf{y} + \mathbf{a}$, $\mathbf{K} \in \mathcal{M}_{\mathbb{C}}(\dim(\mathbf{x}), \dim(\mathbf{y}))$, $\mathbf{a} \in \mathcal{M}_{\mathbb{C}}(\dim(\mathbf{x}), 1)$, is $\mathbf{e} \triangleq \mathbf{e}(\mathbf{y}, \mathbf{x}) = \hat{\mathbf{x}}(\mathbf{y}) - \mathbf{x}$ and the error covariance matrix is:

$$\mathbf{P}(\mathbf{K}, \mathbf{a}) = E[\mathbf{e}\mathbf{e}^H] = E[(\hat{\mathbf{x}}(\mathbf{y}) - \mathbf{x})(\hat{\mathbf{x}}(\mathbf{y}) - \mathbf{x})^H] \quad (4)$$

As:

$$\mathbf{e} = \mathbf{K}(\mathbf{y} - \mathbf{m}_{\mathbf{y}}) - (\mathbf{x} - \mathbf{m}_{\mathbf{x}}) + (\mathbf{a} - (\mathbf{m}_{\mathbf{x}} - \mathbf{K}\mathbf{m}_{\mathbf{y}})) \quad (5)$$

then, if $\mathbf{C}_{\mathbf{y}}$ is invertible, (4) can be rewritten as [8, p121]:

$$\mathbf{P}(\mathbf{K}, \mathbf{a}) = \mathbf{C}_{\mathbf{x}|\mathbf{y}} + (\mathbf{K} - \mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1})\mathbf{C}_{\mathbf{y}}(\mathbf{K} - \mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1})^H + (\mathbf{a} - (\mathbf{m}_{\mathbf{x}} - \mathbf{K}\mathbf{m}_{\mathbf{y}}))(\mathbf{a} - (\mathbf{m}_{\mathbf{x}} - \mathbf{K}\mathbf{m}_{\mathbf{y}}))^H. \quad (6)$$

This quadratic form in \mathbf{a} is positive semidefinite, so $\mathbf{P}(\mathbf{K}, \mathbf{a}) \geq \mathbf{C}_{\mathbf{x}|\mathbf{y}} + (\mathbf{K} - \mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1})\mathbf{C}_{\mathbf{y}}(\mathbf{K} - \mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1})^H$ with equality for:

$$\mathbf{a}^b = \arg \min_{\mathbf{a}} \{\mathbf{P}(\mathbf{K}, \mathbf{a})\} = \mathbf{m}_{\mathbf{x}} - \mathbf{K}\mathbf{m}_{\mathbf{y}}, \quad (7)$$

yielding:

$$\mathbf{P}(\mathbf{K}, \mathbf{a}^b) = \mathbf{C}_{\mathbf{x}|\mathbf{y}} + (\mathbf{K} - \mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1})\mathbf{C}_{\mathbf{y}}(\mathbf{K} - \mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1})^H,$$

where the superscript b is used in the following to remind the reader that the value under consideration (scalar, vector, matrix) is the "best" one according to an unambiguous criterion previously defined (here the MSE). Then (5) reduces to $\mathbf{e} = \mathbf{K}(\mathbf{y} - \mathbf{m}_{\mathbf{y}}) - (\mathbf{x} - \mathbf{m}_{\mathbf{x}})$ leading to:

$$\mathbf{K}^b = \arg \min_{\mathbf{K}} \{\mathbf{P}(\mathbf{K}, \mathbf{a}^b)\} = \mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1}, \quad \mathbf{a}^b = \mathbf{m}_{\mathbf{x}} - \mathbf{K}^b\mathbf{m}_{\mathbf{y}}, \quad (8)$$

and⁴:

$$\hat{\mathbf{x}}^b - \mathbf{m}_{\mathbf{x}} = \widehat{\mathbf{x} - \mathbf{m}_{\mathbf{x}}}^b = \mathbf{C}_{\mathbf{x},\mathbf{y}}\mathbf{C}_{\mathbf{y}}^{-1}(\mathbf{y} - \mathbf{m}_{\mathbf{y}}) \quad (9)$$

where:

$$\mathbf{P}(\mathbf{K}^b, \mathbf{a}^b) = \mathbf{C}_{\mathbf{x}|\mathbf{y}}, \quad (10)$$

As \mathbf{K}^b (8) is also the solution of the linear LMSE after the centering of \mathbf{x} and \mathbf{y} (9), we follow [2] hereinafter and only use the term linear to designate both linear and affine LMSE.

³The case of unknown or unavailable mean values is addressed Section V-C

⁴Note that if \mathbf{x} and \mathbf{y} are (proper) normal complex random vectors, then $\hat{\mathbf{x}}^b$ is the mean value of \mathbf{x} given \mathbf{y} , i.e. the minimum MSE estimator.

B. LMVDR estimator (LMVDRE)

We adopt the notation used in the deterministic framework for the LMVDRE [7, §6][8, §5.6] to stress the fact that the LMVDRE is different from the LLMSE. Let us consider three complex random vectors \mathbf{x} , \mathbf{y} and \mathbf{v} verifying: $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}$, where $\mathbf{H} \in \mathcal{M}_{\mathbb{C}}(\dim(\mathbf{y}), \dim(\mathbf{x}))$ is full rank. The error between the signal \mathbf{x} and the affine estimator $\hat{\mathbf{x}} \triangleq \hat{\mathbf{x}}(\mathbf{y}) = \mathbf{W}^H\mathbf{y} + \mathbf{a}$, $\mathbf{W} \in \mathcal{M}_{\mathbb{C}}(\dim(\mathbf{y}), \dim(\mathbf{x}))$, $\mathbf{a} \in \mathcal{M}_{\mathbb{C}}(\dim(\mathbf{x}), 1)$, is $\mathbf{e} \triangleq \mathbf{e}(\mathbf{y}, \mathbf{x}) = \hat{\mathbf{x}}(\mathbf{y}) - \mathbf{x}$ and the error covariance matrix is given by (4). In this setting, \mathbf{W} can be seen as a "state-former" in the same way as a beamformer in array processing or a frequency-bin former in spectral analysis [7, §6][8, §5.6]. We look for a distortionless "state-former", that is \mathbf{W}^b and \mathbf{a}^b such that:

$$\{\mathbf{W}^b, \mathbf{a}^b\} = \arg \min \{\mathbf{P}(\mathbf{W}^H, \mathbf{a})\} \text{ s.t. } \mathbf{W}^H\mathbf{H} = \mathbf{I}. \quad (11)$$

Since (7) is valid whatever $\mathbf{K} = \mathbf{W}^H$, then \mathbf{e} (5) reduces to:

$$\begin{aligned} \mathbf{e} &= \mathbf{W}^H(\mathbf{y} - \mathbf{m}_{\mathbf{y}}) - (\mathbf{x} - \mathbf{m}_{\mathbf{x}}) \\ &= (\mathbf{W}^H\mathbf{H} - \mathbf{I})(\mathbf{x} - \mathbf{m}_{\mathbf{x}}) + \mathbf{W}^H(\mathbf{v} - \mathbf{m}_{\mathbf{v}}) \\ &= \mathbf{W}^H(\mathbf{v} - \mathbf{m}_{\mathbf{v}}), \end{aligned}$$

leading to the following well known solution of (11) [3, p83][8, §5.6]:

$$\mathbf{W}^b = \mathbf{C}_{\mathbf{v}}^{-1}\mathbf{H}(\mathbf{H}^H\mathbf{C}_{\mathbf{v}}^{-1}\mathbf{H})^{-1}, \quad \mathbf{a}^b = -(\mathbf{W}^b)^H\mathbf{m}_{\mathbf{v}}, \quad (12)$$

and:

$$\mathbf{P}(\mathbf{W}^b, \mathbf{a}^b) = (\mathbf{H}^H\mathbf{C}_{\mathbf{v}}^{-1}\mathbf{H})^{-1}. \quad (13)$$

Additionally, since we have restricted \mathbf{W} to the subset satisfying $\mathbf{W}^H\mathbf{H} = \mathbf{I}$, we can assert that:

$$\mathbf{P}(\mathbf{W}^b, \mathbf{a}^b) \geq \mathbf{P}(\mathbf{K}^b, \mathbf{a}^b). \quad (14)$$

Then again, as \mathbf{W}^b (12) is also the solution of the linear MVDRE after the centering of \mathbf{x} , \mathbf{y} and \mathbf{v} , we only use the term linear to designate both linear and affine MVDRE.

IV. LINEAR FILTERS FOR LDSS MODELS

A. LLMSE for LDSS models

It has been known for ages [1][19] that when \mathbf{x}_0 is also zero-mean, the LLMSE of \mathbf{x}_k based on measurements up to and including time k is simply (8):

$$\hat{\mathbf{x}}_{k|k} = [\mathbf{G}_{k-1}^b \mathbf{K}_k^b] \bar{\mathbf{y}}_k \mid [\mathbf{G}_{k-1}^b \mathbf{K}_k^b] = \mathbf{C}_{\mathbf{x}_k, \bar{\mathbf{y}}_k} \mathbf{C}_{\bar{\mathbf{y}}_k}^{-1}, \quad (15)$$

where $\mathbf{G}_{k-1}^b \in \mathcal{M}_{\mathbb{C}}(P, N(k-1))$ and $\mathbf{K}_k^b \in \mathcal{M}_{\mathbb{C}}(P, N)$. What is less known, to the best of our knowledge (and not disclosed in recent monographs [2-5] or papers [14-17]), is the following general form of the LLMSE [18, Section II]:

$$\begin{aligned} \hat{\mathbf{x}}_{k|k} &= \mathbf{F}_{k-1}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{K}_k^b(\mathbf{y}_k - \mathbf{H}_k\mathbf{F}_{k-1}\hat{\mathbf{x}}_{k-1|k-1}) \\ &\quad + (\mathbf{I} - \mathbf{K}_k^b\mathbf{H}_k) \left(\mathbf{C}_{\mathbf{w}_{k-1}, \bar{\mathbf{y}}_{k-1}}, \mathbf{C}_{\bar{\mathbf{y}}_{k-1}}^{-1} \right) \bar{\mathbf{y}}_{k-1} \\ &\quad - \mathbf{K}_k^b \left(\mathbf{C}_{\mathbf{v}_k, \bar{\mathbf{y}}_{k-1}}, \mathbf{C}_{\bar{\mathbf{y}}_{k-1}}^{-1} \right) \bar{\mathbf{y}}_{k-1} \end{aligned} \quad (16)$$

for $k \geq 2$, where \mathbf{K}_k^b is given by [18, (13)] ($\mathbf{K}_k^b = \mathbf{W}_k^H$). One can notice that (16) has two additional terms in comparison with the recursive predictor/corrector form:

$$\hat{\mathbf{x}}_{k|k} = \mathbf{F}_{k-1}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{K}_k^b(\mathbf{y}_k - \mathbf{H}_k\mathbf{F}_{k-1}\hat{\mathbf{x}}_{k-1|k-1}) \quad (17)$$

introduced by Kalman [1]. Therefore, the general assumptions required to obtain the Kalman form (17) of the LLMSF for LDSS are:

$$\mathbf{C}_{\mathbf{w}_{k-1}, \bar{\mathbf{y}}_{k-1}} = \mathbf{0}, \quad \mathbf{C}_{\mathbf{v}_k, \bar{\mathbf{y}}_{k-1}} = \mathbf{0}, \quad k \geq 2. \quad (18)$$

For any LDSS model satisfying (18), then the associated LLMSF has a recursive predictor/corrector format (17), aka the KF, where \mathbf{K}_k^b must be computed according to the following recursion for $k \geq 1$ [18, Section II]:

$$\begin{aligned} \mathbf{P}_{k|k-1} &= \mathbf{F}_{k-1} \mathbf{P}_{k-1|k-1} \mathbf{F}_{k-1}^H + \mathbf{C}_{\mathbf{w}_{k-1}} \\ &\quad + \mathbf{F}_{k-1} \mathbf{C}_{\mathbf{w}_{k-1}, \mathbf{x}_{k-1}}^H + \mathbf{C}_{\mathbf{w}_{k-1}, \mathbf{x}_{k-1}} \mathbf{F}_{k-1}^H \end{aligned} \quad (19a)$$

$$\begin{aligned} \mathbf{S}_{k|k} &= \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^H + \mathbf{C}_{\mathbf{v}_k} + \mathbf{H}_k \mathbf{C}_{\mathbf{v}_k, \mathbf{x}_k}^H + \mathbf{C}_{\mathbf{v}_k, \mathbf{x}_k} \mathbf{H}_k^H \\ \mathbf{K}_k^b &= (\mathbf{P}_{k|k-1} \mathbf{H}_k^H + \mathbf{C}_{\mathbf{v}_k, \mathbf{x}_k}^H) \mathbf{S}_{k|k}^{-1} \end{aligned} \quad (19b)$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k^b \mathbf{H}_k) \mathbf{P}_{k|k-1} - \mathbf{K}_k^b \mathbf{C}_{\mathbf{v}_k, \mathbf{x}_k} \quad (19c)$$

provided that $\mathbf{P}_{0|0} = \mathbf{C}_{\mathbf{x}_0}$ and $\hat{\mathbf{x}}_{0|0} = \mathbf{0}^5$. Let us remind that: $\mathbf{P}_{k|k-1} = E[(\hat{\mathbf{x}}_{k|k-1} - \mathbf{x}_k)(\hat{\mathbf{x}}_{k|k-1} - \mathbf{x}_k)^H] = \mathbf{C}_{\mathbf{x}_k | \bar{\mathbf{y}}_{k-1}}$, $\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_{k-1} \hat{\mathbf{x}}_{k-1|k-1}$, is the a priori estimate of \mathbf{x}_k , and $\mathbf{S}_{k|k} = E[\varepsilon_k \varepsilon_k^H] = \mathbf{C}_{\mathbf{y}_k | \bar{\mathbf{y}}_{k-1}}$, $\varepsilon_k = \mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1}$, is the innovations vector. Note that the so-called "standard LDSS model" mentioned in monographs [2, §9.1][3, §7.1][4, §8.2], which satisfies:

$$\begin{aligned} \mathbf{C}_{\mathbf{x}_0, \mathbf{w}_k} &= \mathbf{0}, \quad \mathbf{C}_{\mathbf{x}_0, \mathbf{v}_k} = \mathbf{0}, \quad \mathbf{C}_{\mathbf{w}_l, \mathbf{w}_k} = \mathbf{C}_{\mathbf{w}_k} \delta_k^l, \\ \mathbf{C}_{\mathbf{v}_l, \mathbf{v}_k} &= \mathbf{C}_{\mathbf{v}_k} \delta_k^l, \quad \mathbf{C}_{\mathbf{w}_l, \mathbf{v}_k} = \mathbf{C}_{\mathbf{w}_{k-1}, \mathbf{v}_k} \delta_k^{l+1}, \end{aligned} \quad (20)$$

and which has been regarded so far as leading to the general form of the KF (without extension of the state and measurement equations) including correlated process and measurement noise, is in fact a special case of (18) yielding the following simplified expressions of (19a-19c):

$$\begin{aligned} \mathbf{P}_{k|k-1} &= \mathbf{F}_{k-1} \mathbf{P}_{k-1|k-1} \mathbf{F}_{k-1}^H + \mathbf{C}_{\mathbf{w}_{k-1}} \\ \mathbf{S}_{k|k} &= \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^H + \mathbf{C}_{\mathbf{v}_k} + \mathbf{H}_k \mathbf{C}_{\mathbf{v}_k, \mathbf{w}_{k-1}}^H + \mathbf{C}_{\mathbf{v}_k, \mathbf{w}_{k-1}} \mathbf{H}_k^H \\ \mathbf{K}_k^b &= (\mathbf{P}_{k|k-1} \mathbf{H}_k^H + \mathbf{C}_{\mathbf{v}_k, \mathbf{w}_{k-1}}^H) \mathbf{S}_{k|k}^{-1} \\ \mathbf{P}_{k|k} &= (\mathbf{I} - \mathbf{K}_k^b \mathbf{H}_k) \mathbf{P}_{k|k-1} - \mathbf{K}_k^b \mathbf{C}_{\mathbf{v}_k, \mathbf{w}_{k-1}} \end{aligned}$$

However, a thorough characterization of the subset of LDSS models compliant with (18) is out of the scope of the paper and is left for future research.

B. LMVDRF for LDSS Models

1) *Initial state estimation:* From (3), the LMVDRE of the initial state \mathbf{x}_0 is given by (12)(13) for $k \geq 2$:

$$\begin{aligned} \hat{\mathbf{x}}_{0|k}^d &= \begin{pmatrix} \mathbf{D}_{k-1}^b \\ \mathbf{W}_{0|k}^b \end{pmatrix}^H \bar{\mathbf{y}}_k \mid \begin{pmatrix} \mathbf{D}_{k-1}^b \\ \mathbf{W}_{0|k}^b \end{pmatrix} = \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{A}}_k \mathbf{P}_{0|k}^d, \\ \mathbf{P}_{0|k}^d &= \left(\bar{\mathbf{A}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{A}}_k \right)^{-1}, \end{aligned}$$

where $\mathbf{D}_{k-1}^b \in \mathcal{M}_{\mathbb{C}}(N(k-1), P)$, $\mathbf{W}_k^b \in \mathcal{M}_{\mathbb{C}}(N, P)$ and the superscript d is used to remind the reader that the estimator

⁵From (8), it is straightforward to prove [2][3] that the case of a non-zero mean initial state \mathbf{x}_0 is addressed by simply setting $\hat{\mathbf{x}}_{0|0} = \mathbf{m}_{\mathbf{x}_0}$.

is a "distortionless" estimator. Actually, $\hat{\mathbf{x}}_{0|k}^d$ has a recursive predictor/corrector format [18, Section III.A] if and only if $\mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \mathbf{n}_k} \in \text{Span}(\bar{\mathbf{A}}_{k-1})$, i.e. for $k \geq 2$:

$$\hat{\mathbf{x}}_{0|k}^d = \hat{\mathbf{x}}_{0|k-1}^d + \begin{pmatrix} \mathbf{W}_{0|k}^b \end{pmatrix}^H \left(\mathbf{y}_k - \mathbf{A}_k \hat{\mathbf{x}}_{0|k-1}^d \right) \Leftrightarrow \quad (21)$$

$$\mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \mathbf{n}_k} = -\bar{\mathbf{A}}_{k-1} \boldsymbol{\lambda}_{k-1} \quad (22)$$

Then $\mathbf{W}_{0|k}^b$ and $\mathbf{P}_{0|k}^d$ satisfy as well for $k \geq 2$:

$$\begin{aligned} \mathbf{S}_{0|k}^d &= \mathbf{A}_k \mathbf{P}_{0|k-1}^d \mathbf{A}_k^H + \mathbf{C}_{\mathbf{n}_k} + \mathbf{A}_k \boldsymbol{\lambda}_{k-1} + \boldsymbol{\lambda}_{k-1}^H \mathbf{A}_k^H \\ \mathbf{W}_{0|k}^b &= \left(\mathbf{S}_{0|k}^d \right)^{-1} \left(\mathbf{A}_k \mathbf{P}_{0|k-1}^d + \boldsymbol{\lambda}_{k-1}^H \right) \\ \mathbf{P}_{0|k}^d &= \left(\mathbf{I} - \left(\mathbf{W}_{0|k}^d \right)^H \mathbf{A}_k \right) \mathbf{P}_{0|k-1} - \left(\mathbf{W}_{0|k}^d \right)^H \boldsymbol{\lambda}_{k-1}^H \end{aligned}$$

For $k = 1$, provided that \mathbf{H}_1 is full rank, (12)(13) simply yield:

$$\hat{\mathbf{x}}_{0|1}^d = \mathbf{P}_{0|1}^d \mathbf{H}_1^H \mathbf{C}_{\mathbf{v}_1}^{-1} \mathbf{y}_1, \quad \mathbf{P}_{0|1}^d = (\mathbf{H}_1^H \mathbf{C}_{\mathbf{v}_1}^{-1} \mathbf{H}_1)^{-1}. \quad (23)$$

If at first sight assumption (22) seems rather obscure, it is however a key property to derive the main contribution of the present paper introduced in the next section.

2) *Current state estimation:* It appears that condition (22) is fulfilled under the general conditions of existence (18) of the KF for a fairly general class of LDSS models. Indeed, let us consider the following equivalent form of (3), provided that \mathbf{F}_k , $k \geq 1$, are invertible:

$$\begin{aligned} \bar{\mathbf{y}}_k &= \bar{\mathbf{A}}_k' \mathbf{x}_k + \bar{\mathbf{n}}_k', \quad \bar{\mathbf{A}}_k' = \begin{pmatrix} \bar{\mathbf{A}}_{k-1}' \mathbf{F}_{k-1}^{-1} \\ \mathbf{H}_k \end{pmatrix}, \\ \bar{\mathbf{n}}_k' &= \begin{pmatrix} \bar{\mathbf{n}}_{k-1}' - \bar{\mathbf{A}}_{k-1}' \mathbf{F}_{k-1}^{-1} \mathbf{w}_{k-1} \\ \mathbf{v}_k \end{pmatrix}, \quad \forall k \geq 2, \end{aligned} \quad (24)$$

and $\bar{\mathbf{y}}_1 = \mathbf{y}_1 = \bar{\mathbf{A}}_1' \mathbf{x}_1 + \bar{\mathbf{n}}_1'$, $\bar{\mathbf{A}}_1' = \mathbf{A}_1' = \mathbf{H}_1$, $\bar{\mathbf{n}}_1' = \mathbf{n}_1' = \mathbf{v}_1$. Then, according to (12)(13), the LMVDRE of the current state \mathbf{x}_k is given by:

$$\hat{\mathbf{x}}_{k|k}^d = \begin{pmatrix} \mathbf{D}_{k-1}^b \\ \mathbf{W}_{k|k}^b \end{pmatrix}^H \bar{\mathbf{y}}_k \mid \begin{pmatrix} \mathbf{D}_{k-1}^b \\ \mathbf{W}_{k|k}^b \end{pmatrix} = \mathbf{C}_{\bar{\mathbf{n}}_k'}^{-1} \bar{\mathbf{A}}_k' \mathbf{P}_{k|k}^d, \quad (25a)$$

$$\mathbf{P}_{k|k}^d = \left(\bar{\mathbf{A}}_k'^H \mathbf{C}_{\bar{\mathbf{n}}_k'}^{-1} \bar{\mathbf{A}}_k' \right)^{-1}. \quad (25b)$$

However we can reparameterize the equivalent observation model (24), formally, as:

$$\bar{\mathbf{y}}_k = \begin{pmatrix} \bar{\mathbf{y}}_{k-1} \\ \mathbf{y}_k \end{pmatrix} = \bar{\mathbf{A}}_k \mathbf{x}_0 + \bar{\mathbf{n}}_k, \quad (26a)$$

where:

$$\begin{aligned} \bar{\mathbf{A}}_k &\triangleq \begin{pmatrix} \bar{\mathbf{A}}_{k-1} \\ \mathbf{A}_k \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{A}}_{k-1}' \mathbf{F}_{k-1}^{-1} \\ \mathbf{H}_k \end{pmatrix} \\ \bar{\mathbf{n}}_k &\triangleq \begin{pmatrix} \bar{\mathbf{n}}_{k-1} \\ \mathbf{n}_k \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{n}}_{k-1}' - \bar{\mathbf{A}}_{k-1}' \mathbf{w}_{k-1} \\ \mathbf{v}_k \end{pmatrix}, \\ \mathbf{x}_0 &\triangleq \mathbf{x}_k \end{aligned} \quad (26b)$$

and compute $\hat{\mathbf{x}}_{k|k}^d$ and $\mathbf{P}_{k|k}^d$ according to (see previous section):

$$\hat{\mathbf{x}}_{k|k}^d \triangleq \hat{\mathbf{x}}_{0|k}^d = \mathbf{P}_{k|k}^d (\mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{A}}_k)^H \bar{\mathbf{y}}_k, \quad (27a)$$

$$\mathbf{P}_{k|k}^d \triangleq \mathbf{P}_{0|k}^d = \left(\bar{\mathbf{A}}_k^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} \bar{\mathbf{A}}_k \right)^{-1}. \quad (27b)$$

Then, according to (26a-26b):

$$\mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \mathbf{n}_k} = \mathbf{C}_{\bar{\mathbf{n}}'_{k-1}, \mathbf{v}_k} - \bar{\mathbf{A}}_{k-1} \mathbf{C}_{\mathbf{w}_{k-1}, \mathbf{v}_k}.$$

However, (18) applied to (24) leads to:

$$\mathbf{C}_{\bar{\mathbf{y}}_{k-1}, \mathbf{v}_k} = \bar{\mathbf{A}}'_{k-1} \mathbf{C}_{\mathbf{x}_{k-1}, \mathbf{v}_k} + \mathbf{C}_{\bar{\mathbf{n}}'_{k-1}, \mathbf{v}_k} = \mathbf{0}.$$

Thus, under (18):

$$\bar{\mathbf{A}}_{k-1} \mathbf{F}_{k-1} \mathbf{C}_{\mathbf{x}_{k-1}, \mathbf{v}_k} + \mathbf{C}_{\bar{\mathbf{n}}'_{k-1}, \mathbf{v}_k} = \mathbf{0}.$$

Therefore, under (18):

$$\mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \mathbf{n}_k} = -\bar{\mathbf{A}}_{k-1} \mathbf{F}_{k-1} \mathbf{C}_{\mathbf{x}_{k-1}, \mathbf{v}_k} - \bar{\mathbf{A}}_{k-1} \mathbf{C}_{\mathbf{w}_{k-1}, \mathbf{v}_k}$$

that is:

$$\mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \mathbf{n}_k} = -\bar{\mathbf{A}}_{k-1} \boldsymbol{\lambda}_{k-1}, \quad \boldsymbol{\lambda}_{k-1} = \mathbf{C}_{\mathbf{x}_k, \mathbf{v}_k},$$

which is a sufficient condition (22) to obtain a predictor/corrector format for the LMVDRE of the current state (21):

$$\hat{\mathbf{x}}_{k|k}^d \triangleq \hat{\mathbf{x}}_{0|k}^d = \hat{\mathbf{x}}_{0|k-1}^d + \left(\mathbf{W}_{0|k}^b \right)^H \left(\mathbf{y}_k - \mathbf{A}_k \hat{\mathbf{x}}_{0|k-1}^d \right) \quad (28)$$

where $\mathbf{A}_k = \mathbf{H}_k$ (26b). Moreover some additional calculus detailed in [18, Section III.B] allow to show that under (18) $\hat{\mathbf{x}}_{0|k-1}^d = \mathbf{F}_{k-1} \hat{\mathbf{x}}_{k-1|k-1}^d$, (28) finally becoming for $k \geq 2$:

$$\hat{\mathbf{x}}_{k|k}^d = \mathbf{F}_{k-1} \hat{\mathbf{x}}_{k-1|k-1}^d + \left(\mathbf{W}_{k|k}^b \right)^H \left(\mathbf{y}_k - \mathbf{H}_k \mathbf{F}_{k-1} \hat{\mathbf{x}}_{k-1|k-1}^d \right) \quad (29)$$

where:

$$\begin{aligned} \mathbf{P}_{k|k-1}^d &= \mathbf{F}_{k-1} \mathbf{P}_{k-1|k-1}^d \mathbf{F}_{k-1}^H + \mathbf{C}_{\mathbf{w}_{k-1}} \\ &+ \mathbf{F}_{k-1} \mathbf{C}_{\mathbf{x}_{k-1}, \mathbf{w}_{k-1}} + \mathbf{C}_{\mathbf{x}_{k-1}, \mathbf{w}_{k-1}}^H \mathbf{F}_{k-1}^H \end{aligned} \quad (30)$$

$$\begin{aligned} \mathbf{S}_{k|k}^d &= \mathbf{H}_k \mathbf{P}_{k|k-1}^d \mathbf{H}_k^H + \mathbf{C}_{\mathbf{v}_k} + \mathbf{H}_k \mathbf{C}_{\mathbf{x}_k, \mathbf{v}_k} + \mathbf{C}_{\mathbf{x}_k, \mathbf{v}_k}^H \mathbf{H}_k^H \\ \mathbf{W}_{k|k}^b &= \left(\mathbf{S}_{k|k}^d \right)^{-1} \left(\mathbf{H}_k \mathbf{P}_{k|k-1}^d + \mathbf{C}_{\mathbf{x}_k, \mathbf{v}_k}^H \right) \end{aligned} \quad (31a)$$

$$\mathbf{P}_{k|k}^d = \left(\mathbf{I} - \left(\mathbf{W}_{k|k}^b \right)^H \mathbf{H}_k \right) \mathbf{P}_{k|k-1}^d - \left(\mathbf{W}_{k|k}^b \right)^H \mathbf{C}_{\mathbf{x}_k, \mathbf{v}_k}^H \quad (32)$$

For $k = 1$, provided that \mathbf{H}_1 is full rank, then (23):

$$\hat{\mathbf{x}}_{1|1}^d = \mathbf{P}_{1|1}^d \mathbf{H}_1^H \mathbf{C}_{\mathbf{v}_1}^{-1} \mathbf{y}_1 \quad (33a)$$

$$\mathbf{P}_{1|1}^d = \left(\mathbf{H}_1^H \mathbf{C}_{\mathbf{v}_1}^{-1} \mathbf{H}_1 \right)^{-1} \quad (33b)$$

It is a bit of a surprise to obtain exactly the same recursion for the KF and the LMVDRF under the same conditions of existence (18), since the two solutions (8) and (12) do not look that similar in the general case. Anyway, this result highlights the key role of the initial condition at $k = 1$ for a recursive predictor/corrector filter. Indeed, despite they share the same recursion, the two filters have different properties:

- the LMVDRF is suboptimal in the MSE sense (14): $\forall \{ \mathbf{m}_{\mathbf{x}_0}, \mathbf{C}_{\mathbf{x}_0} \} : \mathbf{P}_{k|k} \leq \mathbf{P}_{k|k}^d$,
- since $\mathbf{P}_{k|k}^d$ (32)(33b) does not depend on $\{ \mathbf{m}_{\mathbf{x}_0}, \mathbf{C}_{\mathbf{x}_0} \}$, therefore $\mathbf{P}_{k|k}^d$ is an upper bound on the KF performance,
- as the LMVDRF does not depend on the prior knowledge on the initial state \mathbf{x}_0 (first and second order statistics), its behaviour can be assessed in advance and it can be pre-computed.

Thus, it is worth having in mind that anytime a KF is initialized with a WLSE as in (33a)(33b) because $\mathbf{C}_{\mathbf{x}_0}$ and $\mathbf{m}_{\mathbf{x}_0}$ are not known, it becomes a LMVDRF. However, the awareness of this implicit filter change is less important if the LDSS model is time-invariant, since $\mathbf{P}_{k|k-1}^d$ and $\mathbf{P}_{k|k-1}$ shares the same discrete algebraic Riccati equation (DARE). Therefore for the standard LDSS model (20), if (\mathbf{F}, \mathbf{H}) is detectable and $(\mathbf{F} - \mathbf{C}_{\mathbf{w}, \mathbf{v}} \mathbf{C}_{\mathbf{v}}^{-1} \mathbf{H})$ is stabilizable [3, §7.3], the LMVDRF and the KF share the same steady-state gain and performance. Last, for completeness, let us recall that for the standard LDSS model (20), (33a)(33b) are also the asymptotic expressions of $\hat{\mathbf{x}}_{1|1}$ and $\mathbf{P}_{1|1}$ in the KF when $\mathbf{C}_{\mathbf{x}_0} = \alpha \mathbf{C}$ and α tends to infinity [2, §3.4.2], i.e. a KF with "no a priori knowledge" is actually a LMVDRF.

V. LMVDR PREDICTOR, SMOOTHER AND FURTHER CONSIDERATIONS

A. LMVDR predictor and smoother

If the KF and the LMVDRF can share the same convenient recursive predictor/corrector format (17), it is first because this format propagates both (global) unbiasedness and distortionless properties. Indeed, it is known that rewriting (17) as [3]:

$$\begin{aligned} \hat{\mathbf{x}}_{k|k} - \mathbf{x}_k &= \left(\mathbf{I} - \mathbf{K}_{k|k}^b \mathbf{H}_k \right) \mathbf{F}_{k-1} \left(\hat{\mathbf{x}}_{k-1|k-1} - \mathbf{x}_{k-1} \right) \\ &- \left(\mathbf{I} - \mathbf{K}_{k|k}^b \mathbf{H}_k \right) \mathbf{w}_{k-1} + \mathbf{K}_{k|k}^b \mathbf{v}_k \end{aligned}$$

allows to check the unbiasedness property propagation:

$$E \left[\hat{\mathbf{x}}_{k-1|k-1} - \mathbf{x}_{k-1} \right] = \mathbf{0} \Rightarrow E \left[\hat{\mathbf{x}}_{k|k} - \mathbf{x}_k \right] = \mathbf{0}. \quad (34)$$

In the LMVDRF case, let us remind that (25a):

$$\hat{\mathbf{x}}_{k-1|k-1} = \bar{\mathbf{W}}_{k-1}^H \bar{\mathbf{y}}_{k-1} = \mathbf{x}_{k-1} + \mathbf{v}'_{k-1}$$

where (24) $\bar{\mathbf{y}}_{k-1} = \bar{\mathbf{A}}'_{k-1} \mathbf{x}_{k-1} + \bar{\mathbf{n}}'_{k-1}$, $\mathbf{v}'_{k-1} = \bar{\mathbf{W}}_{k-1}^H \bar{\mathbf{n}}'_{k-1}$ and $\bar{\mathbf{W}}_{k-1}^H \bar{\mathbf{A}}'_{k-1} = \mathbf{I}$. Therefore (17) can be rewritten as:

$$\begin{aligned} \hat{\mathbf{x}}_{k|k} &= \left(\mathbf{W}_{k|k}^b \right)^H \mathbf{H}_k \mathbf{x}_k + \left(\mathbf{W}_{k|k}^b \right)^H \mathbf{v}_k + \\ &\left(\mathbf{I} - \left(\mathbf{W}_{k|k}^b \right)^H \mathbf{H}_k \right) \mathbf{F}_{k-1} \left(\mathbf{x}_{k-1} + \mathbf{v}'_{k-1} \right) \\ &= \mathbf{x}_k + \mathbf{v}'_k \end{aligned}$$

where:

$$\mathbf{v}'_k = \left(\mathbf{W}_{k|k}^b \right)^H \mathbf{v}_k + \left(\mathbf{I} - \left(\mathbf{W}_{k|k}^b \right)^H \mathbf{H}_k \right) \left(\mathbf{F}_{k-1} \mathbf{v}'_{k-1} - \mathbf{w}_{k-1} \right),$$

demonstrating the distortionless property propagation:

$$\hat{\mathbf{x}}_{k-1|k-1} = \mathbf{x}_{k-1} + \mathbf{v}'_{k-1} \Rightarrow \hat{\mathbf{x}}_{k|k} = \mathbf{x}_k + \mathbf{v}'_k. \quad (35)$$

By noticing that the KF and the LMVDRF both satisfy:

$$\begin{aligned} \hat{\mathbf{x}}_{k|k} - \mathbf{x}_k &= \mathbf{K}_{k|k}^b \mathbf{v}_k + \\ &\left(\mathbf{I} - \mathbf{K}_{k|k}^b \mathbf{H}_k \right) \left(\mathbf{F}_{k-1} \left(\hat{\mathbf{x}}_{k-1|k-1} - \mathbf{x}_{k-1} \right) - \mathbf{w}_{k-1} \right) \end{aligned}$$

they both can be derived as:

$$\left(\mathbf{W}_{k|k}^b \right)^H = \mathbf{K}_{k|k}^b = \arg \min \{ \mathbf{P}_{k|k} (\mathbf{K}_k) \},$$

where:

$$\begin{aligned} \mathbf{P}_{k|k}(\mathbf{K}_k) &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^H \\ &\quad - (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{C}_{\mathbf{x}_k, \mathbf{v}_k} \mathbf{K}_k^H - \mathbf{K}_k \mathbf{C}_{\mathbf{x}_k, \mathbf{v}_k}^H (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^H \\ &\quad + \mathbf{K}_k \mathbf{C}_{\mathbf{v}_k} \mathbf{K}_k^H, \quad (36) \end{aligned}$$

and $\mathbf{P}_{k|k-1}$ is given by (19a), leading to general form of the Joseph stabilized version of the covariance measurement update equation.

1) *Standard LDSS model*: As the LMVDRF only differs from the KF in the initial values of $\hat{\mathbf{x}}_{1|1}$ and $\mathbf{P}_{1|1}$, the LMVDRF can always be regarded as a particular instantiation of the KF for the following initial conditions: $\mathbf{m}_{\mathbf{x}_0} = \mathbf{0}$, $\mathbf{C}_{\mathbf{v}_1, \mathbf{w}_0} = -\mathbf{H}_1^H (\mathbf{H}_1^H \mathbf{C}_{\mathbf{v}_1}^{-1} \mathbf{H}_1)^{-1}$, $\mathbf{F}_0 \mathbf{C}_{\mathbf{x}_0} \mathbf{F}_0^H + \mathbf{C}_{\mathbf{w}_0} = 2 (\mathbf{H}_1^H \mathbf{C}_{\mathbf{v}_1}^{-1} \mathbf{H}_1)^{-1}$, since, then:

$$\mathbf{S}_{1|1} = \mathbf{C}_{\mathbf{v}_1}, \mathbf{K}_1^b = (\mathbf{H}_1^H \mathbf{C}_{\mathbf{v}_1}^{-1} \mathbf{H}_1)^{-1} \mathbf{H}_1^H \mathbf{C}_{\mathbf{v}_1}^{-1}, \hat{\mathbf{x}}_{1|1} = \mathbf{K}_{1|1}^b \mathbf{y}_1.$$

Therefore, for the standard LDSS model, the two filters share: a) the same possible extensions: colored process noise, colored measurement noise, prediction, smoothing, filtering with fading memory, constrained filtering,..., b) the same alternate formulations: information filtering, innovations approach,... [2][3]. It seems likely that the two filters also share the same drawbacks as sensitivity to modelling error [3, §10][10] or uncertainties in the system matrices [12][13]. However, a thorough comparison of the robustness [3, §10.4][11] of the two filters is far beyond the scope of this paper and is left for future research.

2) *Non-standard LDSS model*: As mentioned in Section IV-A, (18) defines the general assumptions required to obtain the filter recursive predictor/corrector format (without extension of the state and measurement equations) for LDSS models and encompass the standard LDSS model (20). To the best of our knowledge, the innovations process properties has been derived only for the standard LDSS model [2, §9] and for the standard linear continuous space state model [2, §16][20]:

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{z}(t) + \mathbf{v}(t), \quad \mathbf{z}(t) = \mathbf{H}(t) \mathbf{x}(t) \\ \dot{\mathbf{x}}(t) &= \mathbf{F}(t) \mathbf{x}(t) + \mathbf{w}(t) \end{aligned}$$

where $t \geq 0$, $\mathbf{x}(0)$, $\mathbf{z}(t)$, $\mathbf{v}(t)$, $\mathbf{w}(t)$ are zero mean process verifying:

$$\begin{aligned} \mathbf{C}_{\mathbf{x}, \mathbf{w}}(0, t) &= \mathbf{0}, \mathbf{C}_{\mathbf{w}}(t, s) = \mathbf{C}_{\mathbf{w}}(t) \delta(t - s), \\ \mathbf{C}_{\mathbf{v}}(t, s) &= \mathbf{C}_{\mathbf{v}}(t) \delta(t - s), \mathbf{C}_{\mathbf{w}, \mathbf{v}}(t, s) = \mathbf{C}_{\mathbf{w}, \mathbf{v}}(t) \delta(t - s), \\ \forall s < t : \mathbf{C}_{\mathbf{v}, \mathbf{z}}(t, s) &= \mathbf{0}, \forall s < t : \mathbf{C}_{\mathbf{w}, \mathbf{y}}(t, s) = \mathbf{0}, \end{aligned}$$

and $\mathbf{C}_{\mathbf{x}, \mathbf{y}}(t, s) = E[\mathbf{x}(t) \mathbf{y}(s)^H]$. Therefore generalization of the innovations process properties to (18) can not easily be drawn from existing results and is not addressed in the present paper. However some straightforward extensions are available, at least the ones that are based on an explicit use of the predictor/corrector format (17)(19a)(19b)(19c)(36) such as augmented system techniques, forward-backward recursion,..., [3] allowing to take into account colored process noise, colored measurement noise, prediction, smoothing, filtering with fading memory, constrained filtering, reduced order-filtering, ...

B. The deterministic least-squares problem

If $\mathbf{x}_k = \mathbf{x}_{k-1} = \dots = \mathbf{x}_0$ where \mathbf{x}_0 is deterministic and unknown, then the LMVDRE and the WLSE are identical (duality) [2, §3.4][4, §4], that is:

$$\hat{\mathbf{x}}_{0|k}^d = \arg \min \left\{ (\bar{\mathbf{y}}_k - \bar{\mathbf{A}}_k \mathbf{x}_0)^H \mathbf{C}_{\bar{\mathbf{n}}_k}^{-1} (\bar{\mathbf{y}}_k - \bar{\mathbf{A}}_k \mathbf{x}_0) \right\} \quad (37)$$

Then one can easily notice that usual assumptions on the recursive WLSE [2, §2.2.2][5, §4.12][6]:

$$\mathbf{w}_k = \mathbf{0}, \quad \mathbf{C}_{\mathbf{v}_k, \mathbf{v}_k} = \mathbf{C}_{\mathbf{v}_k} \delta_k^l, \quad (38)$$

leads to (2):

$$\bar{\mathbf{n}}_k = \bar{\mathbf{v}}_k, \quad \mathbf{C}_{\bar{\mathbf{n}}_{k-1}, \bar{\mathbf{n}}_k} = \mathbf{C}_{\bar{\mathbf{v}}_{k-1}, \bar{\mathbf{v}}_k} = \mathbf{0},$$

which is a particular case of (22) and can be regarded as an alternative proof of the recursive predictor/corrector format of the WLSE in that particular case, i.e.:

$$\begin{aligned} \hat{\mathbf{x}}_{0|k}^d &= \arg \min \left\{ \sum_{l=1}^k (\mathbf{y}_l - \mathbf{A}_l \mathbf{x}_0)^H \mathbf{C}_{\mathbf{v}_l}^{-1} (\mathbf{y}_l - \mathbf{A}_l \mathbf{x}_0) \right\} \\ &= \hat{\mathbf{x}}_{0|k-1}^d + \left(\mathbf{W}_{0|k}^b \right)^H \left(\mathbf{y}_k - \mathbf{A}_k \hat{\mathbf{x}}_{0|k-1}^d \right) \end{aligned}$$

where:

$$\begin{aligned} \mathbf{W}_{0|k}^b &= \left(\mathbf{A}_k \mathbf{P}_{0|k-1}^d \mathbf{A}_k^H + \mathbf{C}_{\mathbf{v}_k} \right)^{-1} \mathbf{A}_k \mathbf{P}_{0|k-1}^d, \\ \mathbf{P}_{0|k}^d &= \left(\mathbf{I} - \left(\mathbf{W}_{0|k}^d \right)^H \mathbf{A}_k \right) \mathbf{P}_{0|k-1}. \end{aligned}$$

The extension of this result to the regularized WLSE (RWLSE) [2, §2.4][6]:

$$\begin{aligned} \hat{\mathbf{x}}_{0|k}^d &= \arg \min \left\{ (\mathbf{x}_0 - \mathbf{c})^H \mathbf{\Lambda}_0^{-1} (\mathbf{x}_0 - \mathbf{c}) \right. \\ &\quad \left. + \sum_{l=1}^k (\mathbf{y}_l - \mathbf{A}_l \mathbf{x}_0)^H \mathbf{C}_{\mathbf{v}_l}^{-1} (\mathbf{y}_l - \mathbf{A}_l \mathbf{x}_0) \right\} \end{aligned}$$

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

$$\mathbf{y}_0 = \mathbf{H}_0 \mathbf{x}_0 + \mathbf{v}_0, \quad \mathbf{C}_{\mathbf{v}_0} = \mathbf{\Lambda}_0, \quad \mathbf{y}_0 = \mathbf{c}, \quad \mathbf{H}_0 = \mathbf{I},$$

and by starting the recursion at time $k = 0$:

$$\hat{\mathbf{x}}_{0|0}^d = \mathbf{P}_{0|0}^d \mathbf{H}_0^H \mathbf{C}_{\mathbf{v}_0}^{-1} \mathbf{y}_0 = \mathbf{c}, \quad \mathbf{P}_{0|0}^d = (\mathbf{H}_0^H \mathbf{C}_{\mathbf{v}_0}^{-1} \mathbf{H}_0)^{-1} = \mathbf{\Lambda}_0,$$

instead of time $k = 1$ (23). In the light of the above, the RWLSE is primarily a special case of the LMVDRE, and its relation to the KF highlighted in [6] is actually purely formal.

C. Unknown mean values

So far, we have followed the didactic approach usually adopted in monographs dealing with linear filtering [2, §3.2.4][3][4][5, §4] which assumes that the mean values of $(\mathbf{x}_0, \{\mathbf{w}_k\}, \{\mathbf{v}_k\})$ are available (theoretically known or estimated). However an analysis of the benefits of the knowledge of the mean values in term of estimation performance is of some interest for LDSS systems design. Indeed, on the one hand, the knowledge of the mean values allows to move from linear LMS/MVDR estimators to affine LMS/MVDR estimators in order to reduce the minimum attainable MSE (10)(13). On the other hand, in some systems, the centering of signals

has a cost in terms of mean values estimation (DC offsets for instances) and of hardware or software implementation.

First, when the mean values of $(\mathbf{x}_0, \{\mathbf{w}_k\}, \{\mathbf{v}_k\})$ are not known, the following modification of the usual definition of proper complex random variables is required:

- a proper complex random variable \mathbf{x} is orthogonal with its complex conjugate \mathbf{x}^* , i.e. $E[\mathbf{x}(\mathbf{x}^*)^H] = E[\mathbf{x}\mathbf{x}^T] = \mathbf{0}$.

- \mathbf{x} and \mathbf{y} are joint proper if: $E[\mathbf{x}\mathbf{x}^T] = \mathbf{0}, E[\mathbf{y}\mathbf{y}^T] = \mathbf{0}$ and $E[\mathbf{x}\mathbf{y}^T] = \mathbf{0}$.

Second, all the results introduced in Section IV hold provided one replaces covariance matrices for correlation matrices and set $\hat{\mathbf{x}}(\mathbf{y}) = \mathbf{K}\mathbf{y}$ or $\hat{\mathbf{x}}(\mathbf{y}) = \mathbf{W}^H\mathbf{y}$. Then, a measure of the loss in MSE sense that occurs, if one does not take the mean values into account for the synthesis of the LLMSE, is [18, Section IV]:

$$\Delta\mathbf{P} = \mathbf{R}_{\mathbf{x}|\mathbf{y}} - \mathbf{C}_{\mathbf{x}|\mathbf{y}} = (1 + \mathbf{m}_{\mathbf{y}}^H \mathbf{C}_{\mathbf{y}}^{-1} \mathbf{m}_{\mathbf{y}})^{-1} \mathbf{a}^b (\mathbf{a}^b)^H \quad (39)$$

As intuitively expected $\Delta\mathbf{P} \geq \mathbf{0}$ with equality if and only if $\mathbf{m}_{\mathbf{x}} = \mathbf{K}^b \mathbf{m}_{\mathbf{y}}$, what is unlikely to be true in the general case. Therefore in the general case, the diagonal element of $\Delta\mathbf{P}$ are strictly positive which is an evidence of the soundness of the centering if we are interested in the estimation of \mathbf{x} only. Indeed, it is worth noting that $\Delta\mathbf{P}$ is a rank one positive matrix and that the improvement only occurs along $\text{span}\{\mathbf{a}^b\}$, which means that for any linear combination of $\mathbf{x} : \lambda^T \mathbf{x} \mid \lambda \in \text{span}\{\mathbf{a}^b\}^\perp$, there is no performance gain obtained by centering the random vectors. As $\dim(\text{span}\{\mathbf{a}^b\}^\perp) = P - 1$, if one is primarily interested only in some linear combinations of $\{\lambda_i^T \mathbf{x}\}$, it is worth checking if $\{\lambda_i\} \in \text{span}\{\mathbf{a}^b\}^\perp$, especially if the random vectors centering is costly.

Last, the common recursive predictor/corrector format for the KF and the LMVDRF ($\hat{\mathbf{x}}_{k|k}^d \triangleq \hat{\mathbf{x}}_{k|k}$, $\mathbf{W}_{k|k}^b \triangleq (\mathbf{K}_{k|k}^b)^H$) is still:

$$\hat{\mathbf{x}}_{k|k} = \mathbf{F}_{k-1} \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{K}_{k|k}^b (\mathbf{y}_k - \mathbf{H}_k \mathbf{F}_{k-1} \hat{\mathbf{x}}_{k-1|k-1})$$

and is obtained for LDSS satisfying:

$$\mathbf{R}_{\mathbf{w}_{k-1}, \bar{\mathbf{y}}_{k-1}} = \mathbf{0}, \mathbf{R}_{\mathbf{v}_k, \bar{\mathbf{y}}_{k-1}} = \mathbf{0}, k \geq 2, \quad (40)$$

where:

$$\begin{aligned} \mathbf{P}_{k|k-1} &= \mathbf{F}_{k-1} \mathbf{P}_{k-1|k-1} \mathbf{F}_{k-1}^H + \mathbf{R}_{\mathbf{w}_{k-1}} \\ &+ \mathbf{F}_{k-1} \mathbf{R}_{\mathbf{w}_{k-1}, \mathbf{x}_{k-1}}^H + \mathbf{R}_{\mathbf{w}_{k-1}, \mathbf{x}_{k-1}} \mathbf{F}_{k-1}^H \end{aligned} \quad (41a)$$

$$\begin{aligned} \mathbf{S}_{k|k} &= \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^H + \mathbf{R}_{\mathbf{v}_k} + \mathbf{H}_k \mathbf{R}_{\mathbf{v}_k, \mathbf{x}_k}^H + \mathbf{R}_{\mathbf{v}_k, \mathbf{x}_k} \mathbf{H}_k^H, \\ \mathbf{K}_{k|k}^b &= (\mathbf{P}_{k|k-1} \mathbf{H}_k^H + \mathbf{R}_{\mathbf{v}_k, \mathbf{x}_k}^H) \mathbf{S}_{k|k}^{-1}, \end{aligned} \quad (41b)$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_{k|k}^b \mathbf{H}_k) \mathbf{P}_{k|k-1} - \mathbf{K}_{k|k}^b \mathbf{R}_{\mathbf{v}_k, \mathbf{x}_k} \quad (41c)$$

with initial conditions:

$$\text{KF: } \hat{\mathbf{x}}_{0|0} = \mathbf{0}, \mathbf{P}_{0|0} = \mathbf{R}_{\mathbf{x}_0} \quad (42a)$$

$$\text{LMVDRF: } \hat{\mathbf{x}}_{1|1} = \mathbf{P}_{1|1} \mathbf{H}_1^H \mathbf{R}_{\mathbf{v}_1}^{-1} \mathbf{y}_1, \mathbf{P}_{1|1} = (\mathbf{H}_1^H \mathbf{R}_{\mathbf{v}_1}^{-1} \mathbf{H}_1)^{-1} \quad (42b)$$

and a KF MSE loss (39) given by ($k \geq 1$):

$$\begin{aligned} \Delta\mathbf{P}_{k|k} &= (1 + e_k)^{-1} \mathbf{a}_k^b (\mathbf{a}_k^b)^H, \\ \mathbf{a}_k^b &= \mathbf{u}_k - \mathbf{K}_k^b \delta_k, e_k = e_{k-1} + \delta_k^H \mathbf{S}_{k|k}^{-1} \delta_k, \\ \mathbf{u}_k &= \mathbf{F}_{k-1} \mathbf{a}_{k-1}^b + \mathbf{m}_{\mathbf{w}_{k-1}}, \delta_k = \mathbf{H}_k \mathbf{u}_k + \mathbf{m}_{\mathbf{v}_k}, \end{aligned} \quad (43)$$

with initial conditions $\mathbf{a}_0^b = \mathbf{m}_{\mathbf{x}_0}$ and $e_0 = 0$.

D. On the use of complex signals

It is worth reminding that complex LDSS models (1a)(1b) is simply the compact form of an equivalent real LDSS models:

$$\begin{aligned} \begin{pmatrix} \mathbf{x}_k^r \\ \mathbf{x}_k^i \end{pmatrix} &= \begin{bmatrix} \mathbf{F}_{k-1}^r & -\mathbf{F}_{k-1}^i \\ \mathbf{F}_{k-1}^i & \mathbf{F}_{k-1}^r \end{bmatrix} \begin{pmatrix} \mathbf{x}_{k-1}^r \\ \mathbf{x}_{k-1}^i \end{pmatrix} + \begin{pmatrix} \mathbf{w}_{k-1}^r \\ \mathbf{w}_{k-1}^i \end{pmatrix} \\ \begin{pmatrix} \mathbf{y}_k^r \\ \mathbf{y}_k^i \end{pmatrix} &= \begin{bmatrix} \mathbf{H}_k^r & -\mathbf{H}_k^i \\ \mathbf{H}_k^i & \mathbf{H}_k^r \end{bmatrix} \begin{pmatrix} \mathbf{x}_k^r \\ \mathbf{x}_k^i \end{pmatrix} + \begin{pmatrix} \mathbf{v}_k^r \\ \mathbf{v}_k^i \end{pmatrix} \end{aligned}$$

where $\mathbf{a}^r, \mathbf{a}^i, \mathbf{A}^r$ and \mathbf{A}^i denotes respectively the real and imaginary part of \mathbf{a} and \mathbf{A} . Therefore, handling proper complex random vectors simply offers the capability of a computation burden reduction without loss of estimation performance in the MSE sense [2, §3.2.5][8, §5.4.1], since the complex LDSS models manipulates vectors with twice less elements and matrices with four time less elements (but twice more additions and multiplications). If the checking of the "proper" property is problematic, one can always use a real LDSS models and all the results released in the present paper provided that one substitutes the matrix/vector transpose conjugate for the matrix/vector transpose.

VI. AN ILLUSTRATIVE EXAMPLE

Consider the problem of the estimation of a bias from noisy measurements based on the following time invariant LDSS model:

$$x_k = x_{k-1} + w_{k-1}, y_k = x_k + v_k, \quad (44)$$

where the noise process, the measurement noise and the initial state x_0 are Gaussian and uncorrelated (20): $C_{x_0, w_k} = 0$, $C_{x_0, v_k} = 0$, $C_{w_l, w_k} = \sigma_w^2 \delta_{lk}^l$, $C_{v_l, v_k} = \sigma_v^2 \delta_{lk}^l$, $C_{w_l, v_k} = 0$. Moreover, assume that $x_0 = 0$ and $\sigma_v^2 = 100$.

Figure (1) displays the root-MSE (RMSE) of the LMVDRF, that is $\sqrt{P_{k|k}^d}$ (32), and the RMSEs of the KF, that is $\sqrt{P_{k|k}}$ (19c), for $C_{x_0} = \sigma_{x_0}^2 \in \{10^{-2}, 1, 10^2, 10^4\}$, when $\sigma_w^2 = \frac{1}{4}$. Figure (1) illustrates the properties of the LMVDRF mentioned in Section IV-B:

- the LMVDRF is suboptimal in the MSE sense: $P_{k|k} \leq P_{k|k}^d$,
- $P_{k|k}^d$ is an upper bound on the KF performance,
- if the LDSS model is time-invariant, the LMVDRF and the KF share the same steady-state performance,
- when $\sigma_{x_0}^2$ tends to infinity the KF tends to the LMVDRF.

As already mentioned, although the LMVDRF is sub-optimal in terms of MSE, its performance does not depend on the prior knowledge (first and second order statistics) on the initial state, which can be a valuable feature as illustrated on Figure (2). Indeed in Figure (2) we assume that the true value of $\sigma_{x_0}^2$ ($\sigma_{x_0}^2 = 1$) is not available and that one has to initialize the KF recursion with an assumed value

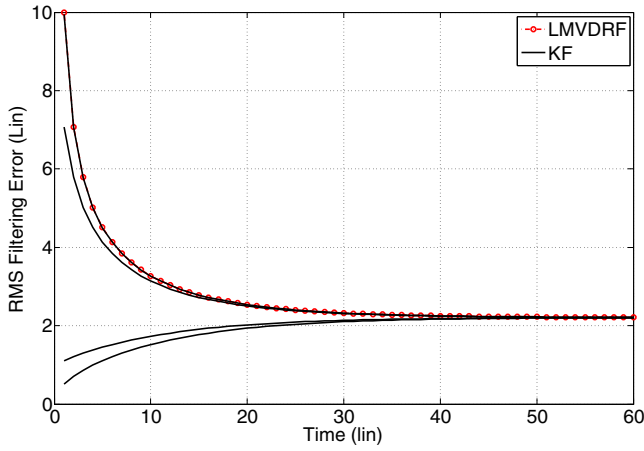


Fig. 1. Comparison of $\sqrt{P_{k|k}^d}$ (32) and $\sqrt{P_{k|k}}$ (19c) for $\sigma_{x_0}^2 \in \{10^{-2}, 1, 10^2, 10^4\}$ and $\sigma_w^2 = \frac{1}{4}$.

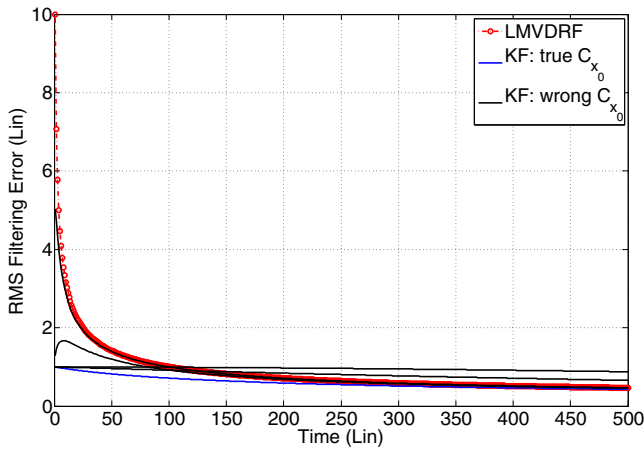


Fig. 2. Comparison of $\sqrt{P_{k|k}^d}$ (32) and $\sqrt{P_{k|k}}$ when $\sigma_{x_0}^2$ is unknown and $\sigma_w^2 = 10^{-4}$.

$\hat{\sigma}_{x_0}^2 \in \{10^{-2}, 10^{-1}, 1, 10, 10^2\}$. In the case of Figure (2), $\sigma_w^2 = 10^{-4}$, $\sqrt{P_{k|k}^d}$ is given by (32) and the empirical $\sqrt{P_{k|k}}$ is assessed with 10^5 Monte-Carlo trials. Thus, in the presence of uncertainties regarding the prior knowledge on the initial state, the LMVDRF may offer better performance than a KF wrongly initialized.

VII. CONCLUSION

For LDSS models, we have identified the general assumptions required to obtain, without extension of the state and measurement equations, a recursive predictor/corrector format of the LLMSE, aka the KF. We have shown that these assumptions, for a fairly general subset of LDSS models, allow the LMVDRF to share exactly the same recursion as the KF, except at time $k = 1$. Interestingly enough, numerous users of the KF, without knowing it, have used a LMVDRF each time they have initialized a KF at time $k = 1$ with a WLSE because the first and second order statistics of the initial state x_0 were not available. Indeed, it is worth being aware of this

result since the LMVDRF is sub-optimal in MSE sense and is an upper bound on the performance of the KF whatever the first and second order statistics of the initial state. In the light of the above, the KF can be looked upon as a "initial state first and second order statistics" matched filter whereas the LMVDRF is a mismatched filter. Once known, these features are quite interesting for estimation performance analysis and design of a LDSS system since they allow to synthesize an IIR distortionless filter whose performance does not depend on the unknown initial state. The LMVDRF may not be the best filter, but its behaviour can be assessed in advance and it can be pre-computed. Additionally we have shown that for the standard LDSS model, the LMVDRF properties (upper bound and pre-computation) can be extended to LMVDR predictors and LMVDR smoothers.

REFERENCES

- [1] R. Kalman, "A new approach to linear filtering and prediction problems", ASME Journal of Basic Engineering, 82: 35-45, 1960
- [2] T. Kailath, A. Sayed, and B. Hassibi, *Linear Estimation*, Prentice-Hall, Upper Saddle River, New Jersey, 2000
- [3] D. Simon, *Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches*, Wiley InterScience, 2006
- [4] B. P. Gibbs, *Advanced Kalman Filtering Least-Squares and Modeling*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2011
- [5] G. A. Einicke, *Smoothing, Filtering and Prediction - Estimating The Past, Present and Future*, InTech, www.intechopen.com, 2012
- [6] A.H. Sayed and T. Kailath, "A state-space approach to adaptive RLS filtering", IEEE Signal Processing Magazine, 11(3): 18-60, 1994
- [7] H.L. Van Trees, *Optimum Array Processing*, Wiley-Interscience, 2002
- [8] P.J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data*, Cambridge University Press 2010
- [9] S. A. Vorobyov, Principles of minimum variance robust adaptive beamforming design, Elsevier Signal Processing, 93: 3264-3277, 2013
- [10] H. Hefkes, "The effect of erroneous models on the Kalman Filter response", IEEE Trans. on AC, 11(13): 541-543, 1966
- [11] D. Peña and L. Guttman, "Bayesian approach to robustifying the Kalman filter", in J. Spall (Ed.) Bayesian analysis of time series and dynamic models (pp. 227-253), New York: Marcel Dekker, 1988
- [12] Y. Theodor and U. Shaked, "Robust discrete-time minimum-variance filtering", IEEE Trans. on SP, 44(2): 181-189, 1996
- [13] S. Kossan and D. Simon, "Kalman Filtering with Uncertain Noise Covariances", Intelligent systems and Control Conference, 2004
- [14] Y. S. Shmaliy, "Linear optimal FIR estimation of discrete time-invariant state-space models", IEEE Trans. SP, 58(6): 3086-2010, 2010
- [15] Y. S. Shmaliy, "An Iterative Kalman-Like Algorithm Ignoring Noise and Initial Conditions", IEEE Trans. SP, 59(6): 2465-2473, 2011
- [16] D. Simon and Y. S. Shmaliy, "Unified forms for Kalman and finite impulse response filtering and smoothing", Automatica, 49: 1892-1899, 2013.
- [17] S. Zhao, Y.S. Shmaliy, B. Huang, F. Liua, "Minimum variance unbiased FIR filter for discrete time-variant systems", Automatica, 53:355-361, 2015
- [18] E. Chaumette, B. Priot, F. Vincent, G. Pages and A. Dion, "Minimum Variance Distortionless Response Estimators For Linear Discrete State-Space Models: Mathematical Appendix", research memorandum, ISAE Supaéro, 2016
- [19] N. Wiener, *The Extrapolation, Interpolation and Smoothing of Stationary Time Series*, John Wiley & Sons, Inc., New York, N.Y., 1949
- [20] T. Kailath, "An innovations approach to least-squares estimation - Part I: Linear filtering in additive white noise", IEEE Trans. on AC, 13(6): 646-655, 1968