On the asymptotic behavior of linearly constrained filters for robust multi-channel signal processing

Paul Chauchat, J. Vilà-Valls, E. Chaumette

On the Asymptotic Behavior of Linearly Constrained Filters for Robust Multi-Channel Signal Processing

Paul Chauchat[b], Jordi Vilà-Valls[a], Eric Chaumette[a]

[a] *ISAE-SUPAERO/University of Toulouse, DEOS Dept., Toulouse, France.*
[b] *IETR, CentraleSupelec, Cesson-Sévigné, Rennes, France*

Abstract

The Kalman filter (KF) is known to loose its optimality properties when the model used does not perfectly match the true system. Extending the use of linear constraints to this filter recently proved to be efficient to mitigate a large class of parametric model mismatch, through the linearly constrained Kalman filter and minimum variance filter (LCKF and LCMVF). However, the asymptotic performances of these new filters are still an open question. In this work, we bring a first answer to the latter problem in the case of measurement model mismatch. We show that both LCKF and LCMVF are equivalent to unconstrained filters in which the directions of the constraints are cancelled by projection, allowing a better understanding of their asymptotic properties. In particular, the steady-state mean square error, when it exists, is derived. The consistency of the filters with respect to nonlinear mismatches is also improved via new constraints. An array processing example is provided to assess the derived formulas, the consistency and performance of the filters.

Keywords: Linearly constrained filtering, asymptotic performance, model mismatch, robust adaptive beamforming.

Email addresses: paul.chauchat@centralesupelec.fr (Paul Chauchat), jordi.vila-valls@isae.fr (Jordi Vilà-Valls), eric.chaumette@isae.fr (Eric Chaumette)

This paper is the preprint of the accepted paper (DOI:10.1016/j.sigpro.2022.108500)

1. Introduction

State estimation of linear systems is a fundamental task, especially in signal processing [1, 2], for which the minimum variance distorsionless response filter (MVDRF) and Kalman filter (KF) play a major role. However, for these filters to minimize the mean square error (MSE), four conditions are needed: i) known system matrices, ii) known noise means and covariance matrices, iii) known inputs, and iv) perfect filter initialisation (for the KF). This is hardly the case in practice, and there thus exists an inherent model mismatch, which may induce a performance breakdown of the filter [3]. In order to directly mitigate parametric system model mismatch, the use of linear constraints [4, 5] was recently leveraged, yielding a new class of linearly constrained MVDRF (LCMVF) and linearly constrained KF (LCKF) [6]. While their potential use and extensions cover various aspects [7, 8], their asymptotic performance is still an open question, since standard tools [9, 10] rely on assumptions i)-iv). We give a first answer in this work, focusing on measurement mismatch. The main theoretical contributions are as follow

• We present a way to increase the consistency of the LCKF and LCMVF for nonlinear mismatch, using additional constraints.

• We show that the LCKF and LCMVF are equivalent to standard filters applied to a system using a projected measurement. This allows applying the classical tools to study the asymptotic behavior. For time-invariant systems, the steady-state MSE (if it exists) is also exhibited.

• Numerical experiments on a uniform linear array example, similar to [6], explain the results observed in this reference, illustrate the impact of higher-order constraints on mitigation capabilities and achievable MSE. A trade-off appears between the acceptable range of mismatch and the achievable MSE of the filters.

2. Improving the LCMVF with Nonlinear Measurement Mismatches

2.1. The LCMVF for Sequential Estimation

Consider the following discrete linear system, with process and observation matrices $\mathbf{F}_{k-1} \in \mathbb{C}^{m_{k-1} \times m_k}$, $\mathbf{H}_k \in \mathbb{C}^{p_k \times m_k}$, inputs $\mathbf{u}_{k-1} \in \mathbb{C}^{m_{k-1}}$, $\mathbf{c}_k \in \mathbb{C}^{p_k}$, and where $\mathbf{w}_{k-1}, \mathbf{v}_k$ are zero-mean noises of known covariance $\mathbf{C}_{\mathbf{w_{k-1}}}, \mathbf{C}_{\mathbf{v_k}}$.

$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{u}_{k-1} + \mathbf{w}_{k-1}, \qquad \mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{c}_k + \mathbf{v}_k. \tag{1}$$

In this work we consider the LCKF and the class of LCMVFs (which differ only at their initialisation) derived in [6] to mitigate potential mismatches of (1). We focus on measurement mismatch only. For the sake of simplicity, we consider an observation matrix depending on a scalar parameter $\mathbf{H}_k = \mathbf{H}_k(\theta_k)$, for which we only have access to $\widehat{\mathbf{H}}_k = \mathbf{H}_k(\widehat{\theta}_k)$. The extension to all the measurement model mismatches presented in [6] is straightforward.

If the filter is updated at step k with an arbitrary gain $\mathbf{L}$, $\widehat{\mathbf{x}}_{k|k}(\mathbf{L}) = \widehat{\mathbf{x}}_{k|k-1} + \mathbf{L}\left(\mathbf{y}_k - \widehat{\mathbf{H}}_k\widehat{\mathbf{x}}_{k|k-1}\right)$, then its error becomes

$$\widehat{\mathbf{x}}_{k|k}(\mathbf{L}) - \mathbf{x}_k = \underbrace{(\mathbf{I} - \mathbf{L}\widehat{\mathbf{H}}_k)(\mathbf{F}_{k-1}(\widehat{\mathbf{x}}_{k-1|k-1} - \mathbf{x}_{k-1}) - \mathbf{w}_{k-1}) + \mathbf{L}\mathbf{v}_k}_{\boldsymbol{\alpha}(\mathbf{L})} + \boldsymbol{\epsilon}_k(\mathbf{L}).$$

$\boldsymbol{\alpha}(\mathbf{L})$ is the standard term of Kalman filtering theory, and $\boldsymbol{\epsilon}_k(\mathbf{L}) = \mathbf{L}(\mathbf{H}_k - \widehat{\mathbf{H}}_k)\mathbf{x}_k$. Thus, for the filter to be unbiased and for its variance to be consistent, it is imposed that $\boldsymbol{\epsilon}_k(\mathbf{L}) = \mathbf{0}$. Of course, this cannot be done in general without exploiting a mismatch model. Here, the considered mismatch model is a parametric one, which leads to linear constraints for the gain which are detailed in Section 2.2. The LCKF is thus closely related to the gain-constrained Kalman filter [11]. In this case, focusing on measurement model mismatch, this yields

$$\mathbf{L}_k = \underset{\mathbf{L}}{\operatorname{argmin}} \left\{ \mathbb{E}\left[(\widehat{\mathbf{x}}_{k|k}(\mathbf{L}) - \mathbf{x}_k)(\cdot)^H \right] \right\} \text{ s.t. } \mathbf{L}\boldsymbol{\Delta}_k = \mathbf{0}. \tag{2}$$

It turns out that (2) has a closed-form solution, leading to the following recursion

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{F}_{k-1}^H + \mathbf{C}_{\mathbf{w}_{k-1}}, \tag{3a}$$

$$\mathbf{S}_{k|k-1} = \widehat{\mathbf{H}}_k\mathbf{P}_{k|k-1}\widehat{\mathbf{H}}_k^H + \mathbf{C}_{\mathbf{v}_k}, \quad \mathbf{K}_k = \mathbf{P}_{k|k-1}\widehat{\mathbf{H}}_k^H(\mathbf{S}_{k|k-1})^{-1}, \tag{3b}$$

$$\mathbf{L}_k = \mathbf{K}_k - \mathbf{K}_k\boldsymbol{\Delta}_k\boldsymbol{\Psi}_k^{-1}\boldsymbol{\Delta}_k^H\left(\mathbf{S}_{k|k-1}\right)^{-1}, \quad \boldsymbol{\Psi}_k = \boldsymbol{\Delta}_k^H(\mathbf{S}_{k|k-1})^{-1}\boldsymbol{\Delta}_k \tag{3c}$$

$$\mathbf{P}_{k|k} = \left(\mathbf{I} - \mathbf{K}_k\widehat{\mathbf{H}}_k\right)\mathbf{P}_{k|k-1} + \mathbf{K}_k\boldsymbol{\Delta}_k\boldsymbol{\Psi}_k^{-1}\boldsymbol{\Delta}_k^H\mathbf{K}_k^H \tag{3d}$$

where, $\mathbf{K}_k$ is the unconstrained Kalman gain. If no prior exists for $\widehat{\mathbf{x}}_{0|0}, \mathbf{P}_{0|0}$, the LCKF is replaced with the LCMVF, initialised with a least-squares approach, the so-called Fisher Initialisation, with $\widetilde{\boldsymbol{\Delta}}_1 = \begin{pmatrix}\widehat{\mathbf{H}}_1 & \boldsymbol{\Delta}_1\end{pmatrix}$, and $\widetilde{\mathbf{T}}_1 = \begin{pmatrix}\mathbf{I} & \mathbf{0}\end{pmatrix}$:

$$\mathbf{P}_{1|1} = \widetilde{\mathbf{T}}_1(\widetilde{\boldsymbol{\Delta}}_1^H\mathbf{C}_{\mathbf{v}_1}^{-1}\widetilde{\boldsymbol{\Delta}}_1)^{-1}\widetilde{\mathbf{T}}_1^H, \quad \widehat{\mathbf{x}}_{1|1} = \widetilde{\mathbf{T}}_1(\widetilde{\boldsymbol{\Delta}}_1^H\mathbf{C}_{\mathbf{v}_1}^{-1}\widetilde{\boldsymbol{\Delta}}_1)^{-1}\widetilde{\boldsymbol{\Delta}}_1^H\mathbf{C}_{\mathbf{v}_1}^{-1}\mathbf{z}_1. \tag{4}$$

The LCMVF can also be seen as an MVDRF with additional constraints [6].

2.2. Increasing the Consistency for Nonlinear Measurement Mismatches

For mismatched observation models, we know that $\boldsymbol{\epsilon}(\mathbf{L}_k) = \mathbf{L}_k(\mathbf{H}_k - \widehat{\mathbf{H}}_k)\mathbf{x}_k$. Recursion (3) is only valueable if the LCKF covariance correctly reflects the system's MSE, i.e., if $\mathbf{L}_k\boldsymbol{\Delta}_k = \mathbf{0}$ ensures $\boldsymbol{\epsilon}(\mathbf{L}_k) = \mathbf{0}$. One of the main limitations of the LCKF as derived until now is that it relied on linearisation, and was thus only consistent for linear and small nonlinear mismatch [6]. In this section, a way to extend the consistency domain of the LCKF is presented. Consider the Taylor expansion of the parametric mismatch

$$\mathbf{H}_k(\theta_k) = \mathbf{H}_k(\widehat{\theta}_k) + \sum_m \frac{1}{m!}\left.\frac{\partial^m\mathbf{H}_k}{\partial\theta^m}\right|_{\widehat{\theta}_k}(\theta_k - \widehat{\theta}_k)^m. \tag{5}$$

Then, adding the constraints $\mathbf{L}_k\left.\frac{\partial^m\mathbf{H}_k}{\partial\theta^m}\right|_{\widehat{\theta}_k} = \mathbf{0}$ for $m \geq 0$ would properly yield an unbiased estimator. Obviously, unless the expansion is finite and short enough, this implies $\mathbf{L}_k = \mathbf{0}$. However, this means that, until $rank(\boldsymbol{\Delta}_k) = p_k - 1$, additional constraints associated to a longer Taylor expansion can be added, extending the range of mitigable mismatch. This will also lower the filter's performance. This trade-off is illustrated in Section 5 on a multi-channel example, which involves a scalar mismatch in an observation model of large dimension.

3. Studying the Asymptotic Behavior with Measurement Mismatch

3.1. Reinterpreting the Constraint $\mathbf{L}_k\mathbf{\Delta}_k = \mathbf{0}$

Theorem 1. *Consider an LCKF with constraint* $\mathbf{L}_k\mathbf{\Delta}_k = \mathbf{0}$ *for all* k*. Let* $\mathbf{\nabla}_k$ *be a basis of Span*$(\mathbf{\Delta}_k)^\perp$*. Then the LCKF is equivalent to the standard KF associated with* (1)*, where the observation model is replaced by*

$$\widetilde{\mathbf{y}}_k = \mathbf{\nabla}_k^H \mathbf{y}_k \tag{6}$$

Proof. Only the update must be checked. (3) can be written

$$= \mathrm{P}_{k|k-1} - \mathbf{P}_{k|k-1}\widehat{\mathbf{H}}_k^H \mathbf{S}_{k|k-1}^{-1} \mathbf{\Pi}^\perp_{\mathbf{\Delta}_k} \widehat{\mathbf{H}}_k \mathbf{P}_{k|k-1}.$$

Notice that $\mathbf{\Pi}^\perp_{\mathbf{\Delta}_k}$ is a projector on $\mathrm{Span}(\mathbf{\Delta}_k)^\perp$. Then we can check that $\mathbf{S}_{k|k-1}^{-1}\mathbf{\Pi}^\perp_{\mathbf{\Delta}_k} = \mathbf{\nabla}_k(\mathbf{\nabla}_k^H \mathbf{S}_{k|k-1}\mathbf{\nabla}_k)^{-1}\mathbf{\nabla}_k^T$. Indeed, both coincide on the columns of $(\mathbf{\Delta}_k, \mathbf{S}_{k|k-1}\mathbf{\nabla}_k)$, which form a basis of $\mathbb{R}^{p_k}$. Thus, we can replace in (7) and get

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1}(\mathbf{\nabla}_k^H \widehat{\mathbf{H}}_k)^H (\mathbf{\nabla}_k^H \mathbf{S}_{k|k-1}\mathbf{\nabla}_k)^{-1} \mathbf{\nabla}_k^H \widehat{\mathbf{H}}_k \mathbf{P}_{k|k-1} \tag{8}$$

which exactly coincides with the standard KF update using (6). □

Theorem 1 means that the LCKF cancels information contained in the mismatched direction. This seems intuitive, but the LCKF has an advantage over directly using the projected observation model (6): it avoids computing $\mathbf{\nabla}_k$.

3.2. Consequence for the Asymptotic Behavior of the LCKF

3.2.1. General Stability Conditions

An important open question is the asymptotic stability of the LCKF, i.e., how the constraints may affect the convergence properties of the KF. Theorem 1 substantially eases this task, as it allows using all the tools which were designed for the KF, such as [9, 10]: most of the practioner's work will be to check the

impact of $\mathbf{\Delta}_k$ on controlability and observability of the system. Since these stability results heavily rely on the particular structure of the Kalman gain, lost for the LCKF, this would have been impossible simply based on (3).

3.2.2. Steady-State Covariance

Consider a time-invariant system, i.e. $\mathbf{F}_{k-1} = \mathbf{F}$, $\widehat{\mathbf{H}}_k = \widehat{\mathbf{H}}$, $\mathbf{C}_{\mathbf{w}_{k-1}} = \mathbf{C}_{\mathbf{w}}$, $\mathbf{C}_{\mathbf{v}_k} = \mathbf{C}_{\mathbf{v}}$. If the predictor covariance $\mathbf{P}_{k|k-1}$ converged to a steady-state value $\mathbf{P}_\infty$, it would be a solution of, using $\mathbf{S} = \widehat{\mathbf{H}}\mathbf{P}\widehat{\mathbf{H}}^H + \mathbf{C}_{\mathbf{v}}$,

$$\mathbf{P} = \mathbf{F}\left(\mathbf{P} - \mathbf{P}\widehat{\mathbf{H}}^H\left(\mathbf{S}^{-1} + \mathbf{S}^{-1}\mathbf{\Delta}(\mathbf{\Delta}^H\mathbf{S}^{-1}\mathbf{\Delta})^{-1}\mathbf{\Delta}^H\mathbf{S}^{-1}\right)\widehat{\mathbf{H}}\mathbf{P}\right)\mathbf{F}^H + \mathbf{C}_{\mathbf{w}}. \quad (9)$$

This is not a Riccati equation, because of the term $(\mathbf{\Delta}^H\mathbf{S}^{-1}\mathbf{\Delta})^{-1}$. This is illustrated in Section 4.2, where, for a scalar state, (9) becomes a polynomial of degree three. To the best of our knowledge, solving (9) is an open question in general, but an important one, since explaining and predicting the plateaus reported for instance in [6] would greatly help the practitioner for system design. However, Theorem 1 allows replacing (9) with a proper Riccati equation

$$\mathbf{P} = \mathbf{F}\left(\mathbf{P} - \mathbf{P}\widehat{\mathbf{H}}^T\mathbf{\nabla}(\mathbf{\nabla}^H(\widehat{\mathbf{H}}\mathbf{P}\widehat{\mathbf{H}}^H + \mathbf{C}_{\mathbf{v}})\mathbf{\nabla})^{-1}\widehat{\mathbf{H}}\mathbf{P}\right)\mathbf{F}^H + \mathbf{C}_{\mathbf{w}} \quad (10)$$

which will, in Section 4.2, induce a second-order polynomial.

Remark 1. *Note that the filter's initialisation plays no role in* (9)*, meaning that, asymptotically, the LCMVF and the LCKF show the same performances (which is also true for their unconstrained counterparts).*

4. Explaining and Improving Linearly Constrained Array Processing

Consider the uniform linear multi-channel array signal processing problem with a miscalibration issue, as in [6]. The signal is obtained by N sensors equally spaced at the assumed distance $\hat{d} = \lambda/2$ (half-wavelength), but in reality at $d \neq \hat{d}$. The impinging random signal source x, assumed Gaussian complex circular, is at a broadside angle $\alpha = 10°$, subject to a spatially and temporally

white noise. Due to fluctuation of the propagation medium, x is also assumed partially coherent with constant correlation factor f and received mean power:

$$x_k = fx_{k-1} + w_{k-1}, \quad C_{w_{k-1}} = (1-f^2)C_{x_{k-1}}, \tag{11}$$

$$\mathbf{y}_k = \mathbf{h}(d,\alpha)x + \mathbf{v}_k, \quad \mathbf{h}(d,\alpha) = \left(e^{j2\pi\frac{(n-1)d\sin(\alpha)}{\lambda}}\right)_{1\leq n\leq N}, \quad \mathbf{C}_{\mathbf{v_k}} = r\mathbf{I}. \tag{12}$$

Moreover, a jamming source x_J might be present, at a known broadside angle α_J, thus modifying (12) as

$$\mathbf{y}_k^{\text{JAM}} = \mathbf{y}_k + \mathbf{h}(d,\alpha_J)x_J. \tag{13}$$

The goal here is to mitigate the mismatch on d, the presence of the jammer, or both at once. They had only been treated separately in [6].

4.1. Mitigation of the Mismatches

Jammer only. In this case, we assume that the array is properly calibrated, i.e. $\hat{d} = d$, but a jammer is present. This induces a purely linear mismatch, as the only unknown is x_J, and it is entirely mitigated by cancelling $\mathbf{h}(d,\alpha_J)$ [6], i.e.

$$\mathbf{\Delta} = \mathbf{h}(d,\alpha_J) \tag{14}$$

Miscalibration only. The mismatch on d has a nonlinear effect on (12). Following Section 2.2, we consider M constraints, where $1 \leq M \leq N-1$ is user-chosen:

$$\mathbf{\Delta}^M = \left[\frac{\partial\mathbf{h}}{\partial d}\Big|_{\widehat{d},\alpha} \quad \cdots \quad \frac{\partial^M\mathbf{h}}{\partial d^M}\Big|_{\widehat{d},\alpha}\right], \left(\frac{\partial^m\mathbf{h}}{\partial d^m}\Big|_{\widehat{d},\alpha}\right)_i = (i-1)\left(\frac{\partial^{m-1}\mathbf{h}}{\partial d^{m-1}}\Big|_{\widehat{d},\alpha}\right)_i \tag{15}$$

Both mismatches at once. If both a miscalibration and a jammer are present, a naive approach would be to simply concatenate (14) and (15). However, (14) also involves d. Thus, the Taylor expansion of $\mathbf{h}(\hat{d},\alpha_J)$ must also be considered:

$$\mathbf{\Delta}^{M_1,M_2} = \left[\mathbf{h}(\hat{d},\alpha_J) \quad \frac{\partial\mathbf{h}}{\partial d}\Big|_{\widehat{d},\alpha} \quad \frac{\partial\mathbf{h}}{\partial d}\Big|_{\widehat{d},\alpha_J} \quad \cdots \quad \frac{\partial^{M_1}\mathbf{h}}{\partial d^{M_1}}\Big|_{\widehat{d},\alpha} \quad \frac{\partial^{M_2}\mathbf{h}}{\partial d^{M_2}}\Big|_{\widehat{d},\alpha_J}\right] \tag{16}$$

4.2. Predicted Steady-State Covariance

Although Riccati equations may be hard to solve in general, the fact that the state is scalar here gives (10) a much gentler form. Let $\mathbf{\nabla}_M$ be an orthonormal

basis of Span$(\mathbf{\Delta}^M)^\perp$, and denote $\widetilde{\mathbf{h}} = \mathbf{\nabla}_M^H \mathbf{h}(\widehat{d}, \alpha)$. Then, if it exists, the steady-state covariance of the KF satisfies, since $\mathbf{C_v} = r\mathbf{I}$ and $\mathbf{\nabla}_M$ is orthogonal

$$p = f^2 p - f^2 p^2 \widetilde{\mathbf{h}}^H (\widetilde{\mathbf{h}} p \widetilde{\mathbf{h}}^H + \mathbf{\nabla}_M^H \mathbf{C_v} \mathbf{\nabla}_M)^{-1} \widetilde{\mathbf{h}} + C_w \tag{17a}$$

$$= f^2 p - f^2 p^2 \widetilde{\mathbf{h}}^H (p \widetilde{\mathbf{h}} \widetilde{\mathbf{h}}^H + r\mathbf{I})^{-1} \widetilde{\mathbf{h}} + C_w. \tag{17b}$$

In the particular case of $M = 1$, $\widetilde{\mathbf{h}}\widetilde{\mathbf{h}}^H$ is of rank one, thus we have $(p\widetilde{\mathbf{h}}\widetilde{\mathbf{h}}^H + r\mathbf{I})^{-1} = \frac{1}{r}\mathbf{I} - \frac{1}{r+p\|\widetilde{\mathbf{h}}\|^2}\frac{p}{r}\widetilde{\mathbf{h}}\widetilde{\mathbf{h}}^H$. Thus, injecting in (17), we readily obtain

$$p = f^2 p - f^2 \frac{\|\widetilde{\mathbf{h}}\|^2}{r + p\|\widetilde{\mathbf{h}}\|^2} p^2 + C_w \Leftrightarrow p^2 + \left((1-f^2)\widetilde{r} - C_w\right) p - \widetilde{r} C_w = 0, \tag{18}$$

with $\widetilde{r} = \frac{r}{\|\widetilde{\mathbf{h}}\|^2}$. Therefore, the steady-state variance always exists, and it is obtained as the only positive root of this degree 2 polynomial:

$$p = \frac{1}{2}\left(-(1-f^2)\widetilde{r} + C_w + \sqrt{\left((1-f^2)\widetilde{r} - C_w\right)^2 + 4\widetilde{r}C_w}\right). \tag{19}$$

Note that there is, in practice, no need to compute $\mathbf{\nabla}_M$, since it only appears in $\|\widetilde{\mathbf{h}}\|^2$, and we have $\|\mathbf{\nabla}_M^H \widehat{\mathbf{h}}\|^2 = \|(\mathbf{I} - \mathbf{\Delta}(\mathbf{\Delta}^H\mathbf{\Delta})^{-1}\mathbf{\Delta}^H)\widehat{\mathbf{h}}\|^2$, with $\widehat{\mathbf{h}} = \mathbf{h}(\widehat{d}, \alpha)$.

Theorem 1 greatly simplifies the study of the steady-state. Indeed, the initial equation (9) leads to a polynomial of degree three, which shares the positive root of (19), with $\phi = \|\widehat{\mathbf{h}}\|^2\|\mathbf{\Delta}\|^2, \psi = \|\widehat{\mathbf{h}}\|^2\|\mathbf{\Delta}\|^2 - (\mathbf{\Delta}^H\widehat{\mathbf{h}})^2, \mu = \phi + \psi$

$$\psi p^3 - (\psi C_w + f^2\widetilde{r}\phi - \mu\widetilde{r})p^2 - (\mu\widetilde{r}C_w + \phi(f^2-1)\widetilde{r}^2)p - \phi\widetilde{r}^2 C_w = 0. \tag{20}$$

5. Numerical Experiments

Numerical simluations validate the steady-state covariance formulae ((18) in general, (19) for $M = 1$, and, incidentally, (20)), and show the influence of extended constraints $\mathbf{\Delta}^M$ for $M \geq 1$. We used $C_x = 1$, a correlation factor $f = 1 - 10^{-4}$, and observation variance $r = 1$. Each time, 10^4 Monte Carlo runs of $K = 10^3$ successive measurements were computed. We first reproduce experiment 4.2.1 of [6], and compare the asymptotic variance with the one given by (19), represented by a horizontal line in Figure 1. (17) correctly predicted $\mathbf{P}_\infty$, as in all other reported cases (note that, this may not coincide with the actual MSE of the system as we show thereafter).

5.1. Mitigation Capability of Large Nonlinear Mismatches

We investigate the mitigation capability of taking into account a growing number of Taylor expansion terms (i.e., with increasing M) for a large mismatch of $d = \hat{d} + \lambda/10$, that is, a 20% mismatch. The predicted and emprirical MSE (denoted as (LCKF) and (MC) respectively) of the first- to fifth-order filters are shown on Figure 2. We can clearly see that this mismatch is far too heavy to be handled by an LCKF using only a first-order constraint, which becomes biased. However, the higher the order, the less biased the filter is, and the more accurate the filter estimated MSE is, the fit being very good for the fifth order.

5.2. Impact of the Constraints on the MSE

Naturally, when more constraints are considered, the achievable MSE is also larger. The next question is thus whether this steady-state MSE grows indefinitely or also has a finite limit. Note that, due to the chosen model (12) for $\mathbf{h}$, $\|\widetilde{\mathbf{h}}\|^2$ only involves N. We thus solved (17) for $1 \leq M < N$, with various values of N, as reported in Figure 3. It turns out that the steady-state filters' MSE indeed has a finite upper limit, i.e., as long as the Taylor expansion holds, one can guarantee a given accuracy once the filter has converged. What was however much less expected, is that this value does not seem to depend on N. This is very important, as it means that, if the expected mismatch is large and one needs additional constraints, the gains of having a larger array could be lost. Moreover, note that, when considering a large number of constraints, a practical numerical issue may arise, because $(\mathbf{\Delta}^M)^H \mathbf{\Delta}^M$ becomes severely ill-conditioned as M grows. In particular, we had to resort to the variable arithmetic precision tools of Matlab to compute the last values of Figure 3.

Remark 2. *In this particular case, the limit MSE is that of an ideal KF (i.e., not mismatched) with only one sensor (i.e., $N = 1$). Indeed, we have $\begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}^H \in Span(\mathbf{\Delta}^{N-1})^{\perp}$. Thus, the measurement is projected on its first component. This is not true for other models, although a limit still exists.*

5.3. How Many Constraints are Needed?

The sensitivity of (12) to d increases with N, so the validity domain of a Taylor expansion of given length decreases. Thus, one can expect that more constraints will be needed as N grows in order to mitigate a given mismatch range. This, along with the results of Sections 5.2 and 5.1, brings up an important practical question: for given mismatch range and array size, which order of the Taylor expansion is needed ? An answer is given in Figure 4, which reports the number of terms of the expansion required so that the LCKF predicted MSE deviates from the empirical one from less than 5%. That is, experiments similar to Section 5.1, with three mismatch levels, 5%, 10% and 20%, and array sizes ranging from 10 to 50, were run with growing M until the MSE is correctly estimated. The resulting M values are reported here. The number of terms needed increases with N, but slowly, and the mismatch. A sixth-order expansion is needed to properly mitigate the heaviest mismatch, with $N = 50$. For the lowest mismatch, only two terms are needed for all the considered cases.

5.4. Handling Miscalibration and Jamming

Finally, the ability of linear constraints to mitigate both jammer and miscalibration as exposed in Section 4.1 is evaluated, in particular the failure of the naive method. A low miscalibration $d = 0.98\hat{d}$ is considered, with a jammer located at $\alpha_J = \alpha + \alpha_{3dB}$, where α_{3dB} denotes the bandwidth. The jammer to noise power is 60 dB and it is always activated. We compare the performance of three filters: the naive approach of simply concatenating the constraints (15) and (14) (i.e., taking $M_1 = 1$ and $M_2 = 0$); the LCMVF with $M_1 = 1$ and $M_2 = 2$; and the LCKF with the same parameters, starting with $\widehat{x}_0 = 0,\ p_0 = 1$. The esimated and empirical MSE of each filter are plotted in Figure 5. The naive approach obviously fails, because the miscalibration, although small, is amplified by the jammer to noise ratio. Taking second order derivatives into account for the jammer allows perfectly estimating the MSE. And, as expected, the LCKF and LCMVF are asymptotically equivalent.

6. Conclusion

This paper studied the asymptotic behavior and consistency of the LCMVF and LCKF in the case of observation-only mismatches. It turns out that these filters are equivalent to unconstrained ones with projected measurements cancelling the directions of the constraints. This allows using all the existing asymptotic stability results for the KF and MVDRF, applied to the system with projected observations. In particular, the steady-state covariance is exhibited. The asymptotic consistency of the filters is also improved for nonlinear mismatches, thanks to additional constraints based on higher-order terms. Numerical experiments validate the derived formulas, and show the trade-off between the mismatch range in which the filter is consistent, and the accuracy of its estimate. Future work will investigate the general case with process model mismatch.

Acknowledgments

Supported by the DGA/AID projects 2019.65.0068.00.470.75.01 and 2021.65.0070.00.470.75.01.

References

[1] H. L. V. Trees, Optimum Array Processing, Wiley-Interscience, USA, 2002.

[2] P. S. R. Diniz, Adaptive Filtering: Algorithms and Practical Implementation, Springer Science+Business Media, 2006.

[3] J. Vilà-Valls, et al., Modelling mismatch and noise statistics uncertainty in linear MMSE estimation, in: Proc. of EUSIPCO, 2019.

[4] O. Frost, An algorithm for linearly constrained adaptive array processing, Proceedings of the IEEE 60 (8) (1972) 926–935.

[5] L. Griffiths, C. Jim, An alternative approach to linearly constrained adaptive beamforming, IEEE Trans. Antennas Propag. 30 (1) (1982) 27–34.

[6] J. Vilà-Valls, et al., Recursive linearly constrained Wiener filter for robust multi-channel signal processing, Signal Processing 167 (2020)

[7] P. Chauchat, et al., Robust information filtering under model mismatch for large-scale dynamic systems, IEEE Control Syst. Lett. (2021)

[8] E. Hrustic, et al., Robust linearly constrained extended Kalman filter for mismatched nonlinear systems, Int. J. Robust Nonlin 31 (3) (2021)

[9] J. Deyst, C. Price, Conditions for asymptotic stability of the discrete minimum-variance linear estimator, IEEE Trans. Automat. Contr. 13 (6) (1968) 702–705.

[10] P. Bougerol, Kalman filtering with random coefficients and contractions, SIAM J Control Optim 31 (1993) 942–959.

[11] B. O. S. Teixeira, et al., Gain-constrained Kalman filtering for linear and nonlinear systems, IEEE Trans. Signal Process. 56 (9) (2008) 4113–4123.

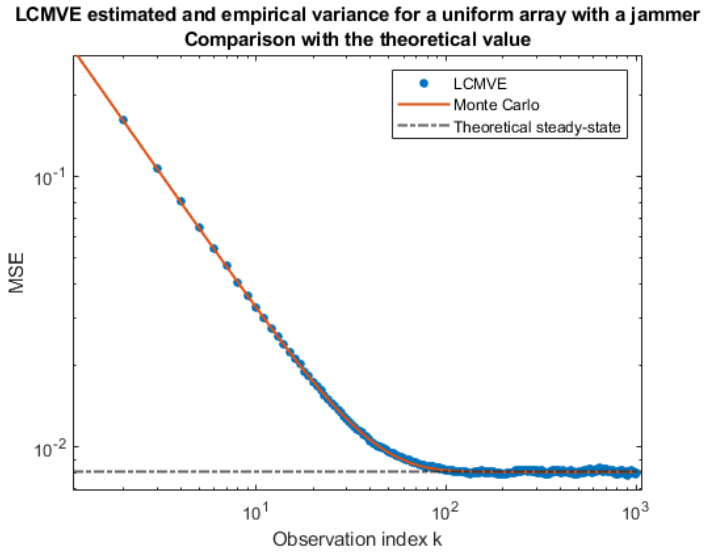

Figure 1: LCMVF estimated and empirical (Monte Carlo) variance for a uniform linear array in the presence of a jammer, as studied in [6]. The steady-state MSE is correctly predicted by (19).

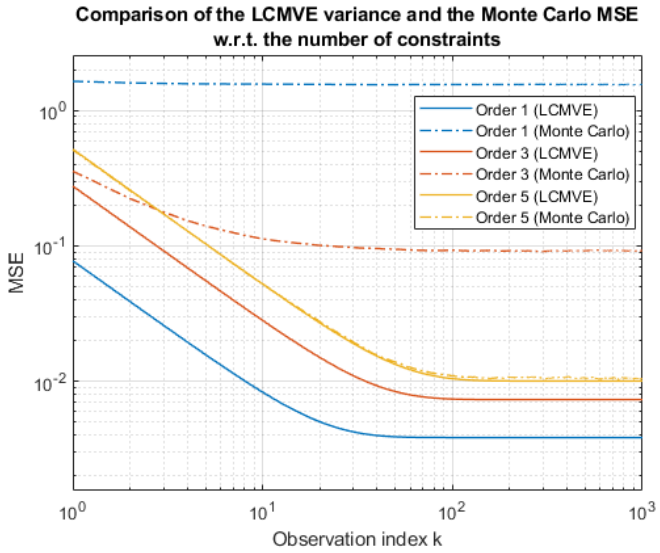

Figure 2: Evolution of the LCMVF variance and the empirical MSE, for different orders of the Taylor expansion used in the constraints, for a 20% mismatch on d. The first-order fails completely, but adding constraints reduces the gap, and from to the fifth-order, the LCKF correctly captures the state's MSE.

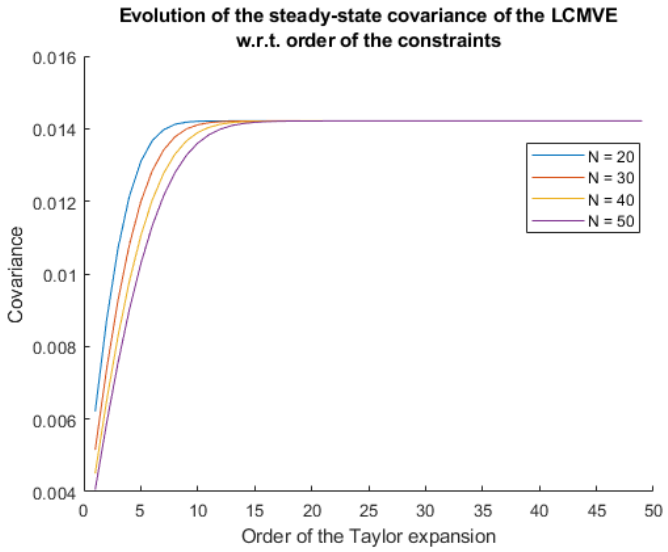

Figure 3: Steady-state variance of the LCMVF computed with (17) for $1 \leq M \leq N-1$, and different values for N. For each N, the steady-state variance tends to a limit as M grows. Unexpectedly, this limit is the same for all N.

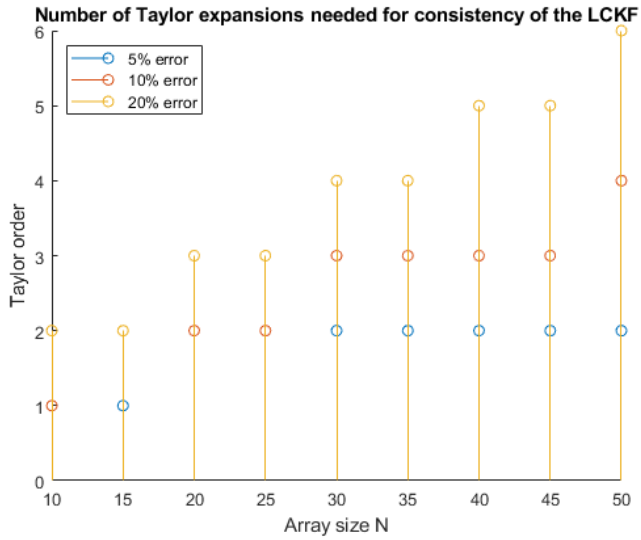

Figure 4: Number M of terms of the Taylor expansion needed to achieve 5% error between the predicted and empirical steady-state MSE, for various mismatch levels. As expected, larger arrays require more terms, since they induce increased nonlinearities. However, M increases much more slowly than N.

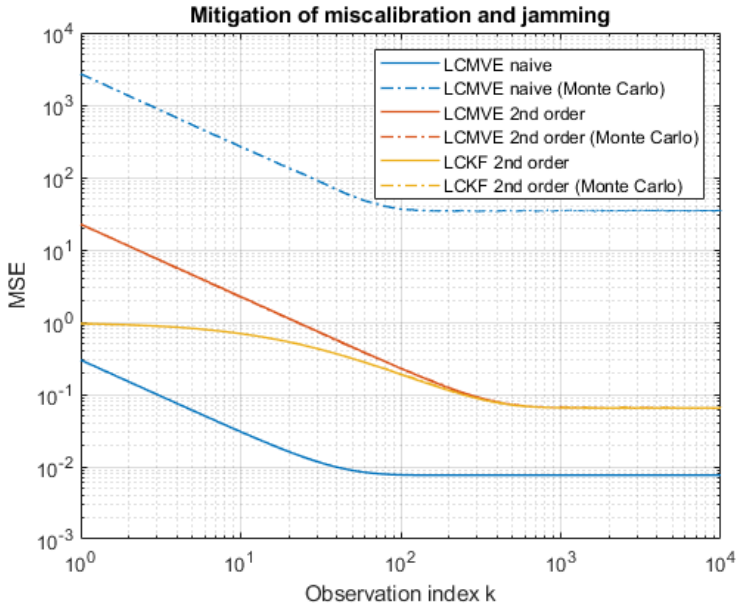

Figure 5: Predicted and empirical MSE of filters in the presence of jamming and miscalibration issues. Merely treating them separately fails, while using constraints up to the 2nd order for the miscalibration in the direction of the jammer cancels the mismatch.