Hindawi
Mathematical Problems in Engineering
Volume 2018, Article ID 5303694, 15 pages
https://doi.org/10.1155/2018/5303694

Research Article

Receding Horizon Least Squares Estimator with Application to Estimation of Process and Measurement Noise Covariances

Vladimir Shin,[1] Rebbecca T. Y. Thien,[2] and Yoonsoo Kim[2]

[1]*Department of Information and Statistics, Research Institute of Natural Science, Gyeongsang National University, Jinju 52828, Republic of Korea*
[2]*Department of Aerospace and Software Engineering, Research Center for Aircraft Parts Technology, Gyeongsang National University, Jinju 52828, Republic of Korea*

Correspondence should be addressed to Yoonsoo Kim; yoonsoo@gnu.ac.kr

Received 3 July 2018; Accepted 17 October 2018; Published 19 November 2018

Academic Editor: Gabriel Luque

This paper presents a noise covariance estimation method for dynamical models with rectangular noise gain matrices. A novel receding horizon least squares criterion to achieve high estimation accuracy and stability under environmental uncertainties and experimental errors is proposed. The solution to the optimization problem for the proposed criterion gives equations for a novel covariance estimator. The estimator uses a set of recent information with appropriately chosen horizon conditions. Of special interest is a constant rectangular noise gain matrices for which the key theoretical results are obtained. They include derivation of a recursive form for the receding horizon covariance estimator and iteration procedure for selection of the best horizon length. Efficiency of the covariance estimator is demonstrated through its implementation and performance on dynamical systems with an arbitrary number of process and measurement noises.

1. Introduction

Filtering and system identification are powerful techniques for building models of complex control systems in time and frequency domains [1–5]. The Kalman filtering and its variations are signal processing techniques in widely used applications, such as navigation, target tracking, vehicle state estimation, communications engineering, air traffic control, and many others [6–9]. As the well-known standard Kalman filter (KF) equations do not directly depend on the original process and measurement noise covariances Q and R, respectively. They depend on the transformed covariances,

$$\begin{aligned} \widetilde{Q} &= GQG^T, \\ \text{and } \widetilde{R} &= DRD^T, \end{aligned} \tag{1}$$

where $G = G_k$ and $D = D_k$ are the rectangular gain matrices associated with process and measurement noises, respectively [1, 2, 6, 7]. In this paper, the original $\{Q, R\}$ and transformed $\{\widetilde{Q}, \widetilde{R}\}$ noise covariances shall be denoted by "ONC" and "TNC", respectively.

The optimality of KF depends on the accuracy of TNC, which directly depend on the prior assumptions on the ONC. Inadequacy of the prior values for the TNC can lead to unexpected results and the KF divergence [1, 2, 7]. The adaptive filtering is one of the approaches to prevent the divergence problem when precise knowledge on the noise covariances is not available.

Adaptive Kalman filtering (AKF) is not the subject of this paper, so we present only its brief survey. The AKF represents a promising strategy for dynamical adjustment of the TNC $\{\widetilde{Q}, \widetilde{R}\}$ for online estimation. Two popular types of the AKF algorithms are the innovation-based estimation approach [10–13] and the adaptive fading filtering approach [14, 15]. The latter is a type of covariance scaling method, into which suboptimal fading factors are incorporated. The innovation-based estimation approach calculates $\widetilde{Q}$ and/or $\widetilde{R}$, assuming that the innovation sequence of the KF represents a white noise. In this approach, one of the covariance matching, the correlation, and the maximum likelihood methods is used. In the covariance matching method at first the sample error covariance of a state is computed based on a moving window

and then $\widetilde{Q}$ is estimated using the calculated sample covariance. In the same way, $\widetilde{R}$ is computed, but using the sample covariance of the KF innovation sequence. In the correlation method, the noise covariance matrices are estimated based on the sample autocorrelations between the innovations by exploiting the relations between the estimation error covariance and the innovation covariance [11, 16, 17]. The drawback of the covariance matching and the correlation methods is that they do not guarantee the positive definiteness of the calculated covariances and require a large window of data. Lastly, implementation of the maximum likelihood method is very complicated which makes it impractical.

As noted the AKF procedures estimate not the ONC, but the TNC. Therefore the estimation of the ONC remains open. Knowledge of the ONCs is necessary to study the quality of a system model. The resulting ONCs help to understand how much environmental disturbance affects the behavior of dynamics and to evaluate the accuracy of sensor measurements. However, if the process G or measurement D ("noise gain matrix") is rectangular, we then have both theoretical and practical challenges in calculating the ONC. As will be shown shortly, popular approaches including the usage of the standard least squares approach and current data may lead to unsatisfactory results. To the best of the authors' knowledge, there are no existing results for the noise covariance estimation within the Kalman filtering framework utilizing not only current data, but finite-memory information from the most recent time interval (receding horizon).

The primary aim of this paper is to expand the previous results concerning the standard least squares covariance estimation to a new robust receding horizon covariance estimation. We propose a novel general estimation method for calculation of the ONCs with rectangular noise gain matrices G and D. The method combines the least squares (LS) optimization and receding horizon strategy. The proposed estimator can be integrated into the AKF for simultaneous estimation of both transformed $\{\widetilde{Q}, \widetilde{R}\}$ and original $\{Q, R\}$ noise covariances.

The main contributions of the paper are listed in the following:

(1) A novel receding horizon least squares optimization criterion is proposed. Using this criterion, an optimal covariance estimator for an arbitrary structure of noise gain matrices is derived.

(2) Constant noise gain matrices are comprehensively investigated, including the derivation of a recursive procedure for the optimal covariance estimator with a horizon interval and a stopping condition during the iteration process.

(3) Performance of the proposed estimator is illustrated on theoretical and practical examples for proving its correctness and efficiency.

This paper is organized as follows. In Section 2, two motivation examples within the Kalman filtering framework are demonstrated. Section 3 presents the standard LS noise covariance estimator and describes consequences of processing current (single) data only. In Section 4, a novel receding horizon least squares estimator (RHLSE) is proposed. It is based on the generalized LS approach and the novel receding horizon LS criterion. Here, the cardinal equations for determining the RHLSE and its equivalent modifications are derived and discussed. In Section 5, the algorithm of searching for the best horizon length improving the accuracy of the RHLSE is proposed. Section 6 demonstrates the practical usage of the RHLSE in adaptive Kalman filtering. In Section 7, performance and effectiveness of the proposed estimator are studied via a theoretical example with rectangular noise gain matrices and a motion with random velocity. Finally, we conclude the paper in Section 8. The list of main notations is given in Table 1.

2. Motivation Examples with Rectangular Noise Gain Matrices

To begin, two typical motivation examples of a dynamical system with rectangular noise gain matrices are presented.

Let us describe the noise estimation problem of present interest via two motivation examples. Suppose we have a linear discrete-time system,

$$\begin{aligned} x_{k+1} &= F_k x_k + G_k w_k, \quad w_k \sim (0, Q), \\ y_k &= H_k x_k + D_k v_k, \quad v_k \sim (0, R). \end{aligned} \tag{2}$$

Then the TNCs $\widetilde{Q}_k = G_k Q G_k^T$ and $\widetilde{R}_k = D_k R D_k^T$ represent covariance matrices of the linearly transformed noises $\widetilde{w}_k = G_k w_k$ and $\widetilde{v}_k = D_k v_k$ with the rectangular gain matrices G_k and D_k, respectively, i.e., $\widetilde{Q}_k = \mathrm{cov}(\widetilde{w}_k)$ and $\widetilde{R}_k = \mathrm{cov}(\widetilde{v}_k)$.

Motivation 1 ("colored process noise"). Consider the scalar linear system with colored process noise:

$$\begin{aligned} x_{k+1} &= \theta x_k + \eta_k, \\ y_k &= x_k + v_k, \end{aligned} \tag{3}$$

where $v_k \sim \mathbb{N}(0, r)$ is a measurement white noise.

Suppose that η_k is a colored process noise generated by the autoregressive equation of order p, $AR(p)$,

$$\eta_k = a_1 \eta_{k-1} + a_2 \eta_{k-2} \ldots + a_p \eta_{k-p} + w_k, \tag{4}$$

where $w_k \sim \mathbb{N}(0, q)$ is a white noise.

Equation (4) may be considered as some system shaping the colored noise η_k from a white noise w_k. Then the following is a state-space representation of the $AR(p)$ model (4):

$$\varepsilon_{k+1} = \widetilde{F} \varepsilon_k + \widetilde{G} w_k, \tag{5}$$

where

$$\varepsilon_k = \begin{bmatrix} \eta_{k-1} \\ \eta_{k-2} \\ \vdots \\ \eta_{k-p} \end{bmatrix} \in \mathbb{R}^p,$$

Table 1: List of main notations.

$\mathbb{R}^n$	Set of n-dimensional real column-vectors
$\mathbb{R}^{n\times m}$	Set of $n \times m$ real matrices
$O_{n\times m}$	$n \times m$ zero matrix
A^T	Transpose of matrix
I_n	Identity matrix of size $n \times n$
A^{-1}	Inverse of $n \times n$ matrix A
A^+	Moore–Penrose pseudoinverse of $n \times m$ matrix A
$tr(A)$	Trace of $n \times n$ matrix A
$[A, B]$	Row block matrix consisting of A and B
$A > 0$ $(A \geq 0)$	Positive definite (semidefinite) matrix
$\mathbb{N}(m, C)$	Normal distribution with mean m and covariance matrix C
$w_k \sim \mathbb{N}(0, Q)$	Normal process with zero-mean and covariance matrix Q
$w_k \sim (0, Q)$	Random process with zero-mean and covariance matrix Q
$\boldsymbol{E}(*)$	Expectation operator
$\mathrm{cov}(x_k)$	Covariance (covariance matrix) of random vector x_k
$\mathrm{cov}(x_k, y_k)$	Cross-covariance of x_k and y_k
$\|A\|_F$	Frobenius norm (F-norm) of matrix, $\|A\|_F = \sqrt{tr(AA^T)}$
$\|A\|_1$	1-norm of matrix, $\|A\|_1 = \max_j \sum_{i=1}^n \|a_{ij}\|$, $A \in \mathbb{R}^{n\times m}$
$\|A\|_\infty$	Infinity norm (I-norm) of matrix, $\|A\|_\infty = \max_i \sum_{j=1}^n \|a_{ij}\|$, $A \in \mathbb{R}^{n\times m}$
$\|A\|_{max}$	Maximum norm (M-norm), $\|A\|_{max} = \max_{i,j}\|a_{ij}\|$
$\|A\|_2$	Spectral norm (S-norm), $\|A\|_2 = \sqrt{\max[\lambda(A^TA)]}$

$$\widetilde{F} = \begin{bmatrix} a_1 & a_2 & \cdots & a_{p-1} & a_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \in \mathbb{R}^{p\times p},$$

$$\widetilde{G} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^p. \tag{6}$$

We combine (3) and (6) to obtain a system model with $(p+1)$–dimensional augmented state X_k and a scalar process noise $w_k \sim \mathbb{N}(0, q)$,

$$\begin{aligned} X_{k+1} &= F_k X_k + G_k w_k, \\ y_k &= h_k^T X_k + v_k, \end{aligned} \tag{7}$$

where

$$X_k = \begin{bmatrix} x_k \\ \varepsilon_k \end{bmatrix} \in \mathbb{R}^{p+1},$$

$$F_k = \begin{bmatrix} \theta & \widetilde{G}^T \\ 0_{p\times 1} & \widetilde{F} \end{bmatrix} \in \mathbb{R}^{(p+1)\times(p+1)},$$

$$G_k = \begin{bmatrix} 0 \\ \widetilde{G} \end{bmatrix} \in \mathbb{R}^{p+1},$$

$$h_k = \begin{bmatrix} 1 \\ 0_{p\times 1} \end{bmatrix} \in \mathbb{R}^{p+1},$$

$$w_k, v_k \in \mathbb{R}. \tag{8}$$

Motivation 2 ("composite process white noise and colored measurement noise"). Consider the following discrete state-space model,

$$x_{k+1} = Fx_k + G_1\xi_{1,k} + G_2\xi_{2,k}, \tag{9a}$$

$$y_k = Cx_k + \eta_k, \tag{9b}$$

$$\eta_k = B\eta_{k-1} + e_{k-1}, \tag{9c}$$

where $x_k \in \mathbb{R}^n$, $y_k \in \mathbb{R}^m$, and $\xi_{1,k} \in \mathbb{R}^{r_1}$, $\xi_{2,k} \in \mathbb{R}^{r_2}$ are two uncorrelated process white noises with covariances Q_1 and Q_2, respectively; and $e_k \in \mathbb{R}^m, e_k \sim \mathbb{N}(0, R)$.

The state equation (9a) can be written in standard form with one composite process white noise $w_k \in \mathbb{R}^r$, and rectangular gain matrix $G \in \mathbb{R}^{n\times r}$,

$$x_{k+1} = Fx_k + Gw_k, \tag{10}$$

where $G = [G_1, G_2]$, $w_k^T = [\xi_{1,k}^T, \xi_{2,k}^T]$, $w_k \in \mathbb{R}^r$, $r = r_1 + r_2$.

Following [2] we define an auxiliary measurement $z_k = y_{k+1} - By_k$ described by the new measurement equation with white noise v_k,

$$z_k = Hx_k + Dv_k, \tag{11}$$

where $H = CF - BC, D = [CG, I_m], v_k^T = [w_k^T, e_k^T]$, $v_k \in \mathbb{R}^{r+m}$.

Combining (10) and (11) we obtain the standard structure of the system model (2) with rectangular noise gain matrices $G \in \mathbb{R}^{n\times r}$ and $D \in \mathbb{R}^{m\times(r+m)}$.

3. Least Squares (LS) Estimation of Noise Covariance Using Single Data

As noted above, the purpose of this paper is to calculate the unknown ONCs Q and R based on the available TNCs $\widetilde{Q}_k$ and $\widetilde{R}_k$ obtained by the AKF algorithm. In this section, we present a simple LS solution to this problem.

For a clear presentation, let us first rewrite (1) in the following form:

$$A_{kn\times p} X_{kp\times p} A^T_{kp\times n} = B_{kn\times n}, \tag{12}$$

where

$$X_k = X_k^T \in \mathbb{R}^{p\times p} \text{ is an unknown ONC;} \tag{13a}$$

$$B_k = B_k^T \in \mathbb{R}^{n\times n} \text{ is a known TNC (single data);} \tag{13b}$$

$$A_k \in \mathbb{R}^{n\times p} \text{ is a noise gain matrix, } A_k = G_k \text{ or } D_k. \tag{13c}$$

In a special case where the gain matrix $A_k \in \mathbb{R}^{n\times n}$ is square and nonsingular, the LS solution to (12) takes the form $X_k = A_k^{-1} B_k A_k^{-T}$, where the known matrix B_k represents a current TNC at time instant k. But in practice, the gain matrix $A_k \in \mathbb{R}^{n\times p}$ is rectangular, as the two motivation examples illustrated. In this rectangular case, the LS solution to (12) can be obtained using a pseudoinverse matrix.

The linear matrix equation (12) has been considered by many authors [18–21]. They have studied general solutions to (12) for special solution structures, e.g., symmetric, triangular, or diagonal X using matrix decomposition techniques such as the singular value decomposition (SVD), the generalized SVD, and the canonical correlation decomposition.

Theorem 3 (see [18]). *The least squares solution to matrix equation (12) takes the form*

$$X_k^{LS} \triangleq X_k^{LS}(B_k) = A_k^{+} B_k A_k^{+T}. \tag{14}$$

Note that the Moore–Penrose pseudoinverse A_k^{+} always exists and is unique.

However, the known TNC B_k in practice is usually obtained from an experiment. It is calculated by using the AKF. Since the LS solution $X_k^{LS}(B_k)$ uses only the single (approximately calculated) TNC B_k, it is likely to be unstable or inaccurate. Therefore to overcome this disadvantage we propose a more realistic idea to use the set of data (TNCs), $B_k, B_{k-1}, \ldots, B_{k-\ell+1}$, for the calculation of unknown covariance X. This set of data is obtained over the most recent time interval (receding horizon), $[k-\ell+1, k]$, saving in a finite memory. Using these data

$$B_{k-\ell+1}^{k} = \{B_{k-\ell+1}, B_{k-\ell+2}, \ldots, B_k\}, \tag{15}$$

we can improve the LS solution (14), and, as a result, one can expect that the obtained receding horizon LS solution $X_k^{LS}(B_{k-\ell+1}^{k})$ will be more robust against processing and measurement errors than the LS solution $X_k^{LS}(B_k)$ which utilizes only the current data B_k.

4. Receding Horizon Least Squares Estimator (RHLSE)

In this section, we propose the novel *receding horizon least squares estimator* referred to as *RHLSE*.

In literature, the estimation problem in dynamical systems having various uncertainties has been solved in different ways, including finite-memory (receding horizon or sliding window) estimations [7, 22]. As a result, the finite-memory estimators are known to be more robust against model uncertainties and numerical errors than the standard LS estimator or Kalman filter [23–25]. For this reason, this receding horizon strategy is chosen to improve the LS solution (14).

According to the strategy proposed above, the unknown solution X_k minimizes a new receding horizon LS criterion representing the sum of squared residuals over the receding horizon $t \in [k-\ell+1, k]$, i.e.,

$$\min_{X_k} J_{k,\ell}(X_k) \tag{16}$$

with

$$\begin{aligned} J_{k,\ell}(X_k) &\triangleq \left\| A_k X_k A_k^T - B_k \right\|_F^2 \\ &\quad + \left\| A_{k-1} X_k A_{k-1}^T - B_{k-1} \right\|_F^2 + \cdots \\ &\quad + \left\| A_{k-\ell+1} X_k A_{k-\ell+1}^T - B_{k-\ell+1} \right\|_F^2 \\ &= \sum_{t=k-\ell+1}^{k} \left\| A_t X_k A_t^T - B_t \right\|_F^2 \end{aligned} \tag{17}$$

where $\ell \geq 1$ is the receding horizon length (HL).

Theorem 4 (RHLSE for time-varying noise gain matrices). *The solution $X_k \triangleq X_{k,\ell}^{RHLSE}$ of the receding horizon least squares optimization problem represented by (17) satisfies the equation*

$$\sum_{t=k-\ell+1}^{k} \left(M_t X_{k,\ell}^{RHLSE} M_t \right) = \sum_{t=k-\ell+1}^{k} \left(A_t^T B_t A_t \right), \tag{18}$$

where $M_t = A_t^T A_t$.

The proof of Theorem 4 is given in Appendix.

Next consider the important case in which the noise gain matrix $A_k \equiv A$ is time-invariant. Then the receding horizon LS problem (17) is rewritten as

$$\min_{X_k} J_{k,\ell}(X_k) \tag{19}$$

with

$$J_{k,\ell}(X_k) = \sum_{t=k-\ell+1}^{k} \left\| A X_k A^T - B_t \right\|_F^2, \tag{20}$$

and we get the following result.

Theorem 5 (RHLSE for time-invariant noise gain matrices). *The optimal receding horizon least squares optimization problem for constant noise gain matrices represented by (20) has the explicit solution,*

$$X_{k,\ell}^{RHLSE}\left(B_{k-\ell+1}^{k}\right) = A^{+}\overline{B}_{k,\ell}A^{+T}, \tag{21}$$

where $\overline{B}_{k,\ell}$ is the horizon average covariance,

$$\overline{B}_{k,\ell} = \frac{1}{\ell}\sum_{t=k-\ell+1}^{k} B_t. \tag{22}$$

The proof of Theorem 5 is given in Appendix.

As we see from (21) and (22), the explicit solution $X_{k,\ell}^{RHLSE}$ depends only on the single data representing the horizon average covariance $\overline{B}_{k,\ell}$. Therefore we get the following.

Corollary 6. *The RHLSE (21) represents the LS estimator (14) calculated for the horizon average covariance, $\overline{B}_{k,\ell}$, i.e.,*

$$X_{k,\ell}^{RHLSE}\left(B_{k-\ell+1}^{k}\right) = X_{k}^{LS}\left(\overline{B}_{k,\ell}\right) = A^{+}\overline{B}_{k,\ell}A^{+T}. \tag{23}$$

5. Selection of Best Horizon Length for Time-Invariant Noise Gain Matrices

The optimal noise covariance (21) is sequentially calculated for each HL (iteration), $\ell = 1, 2, \ldots, \ell^{*}$. Here we propose a stopping criterion for selecting the best HL, ℓ^{*}. This stopping criterion is based on the errors between iterations. Let us first represent the solutions (21) and (22) in recursive form.

The recursive formula for the horizon average covariance (22) takes the form

$$\overline{B}_{k,\ell} = \frac{\ell-1}{\ell}\overline{B}_{k,\ell-1} + \frac{1}{\ell}B_{k-\ell+1}, \text{âNĺ}\ell = 1, 2, \ldots, \ell_{max}. \tag{24}$$

Using (21) and (24) the explicit solution $X_{k,\ell}^{RHLSE}$ is also rewritten in recursive form,

$$\begin{aligned} X_{k,\ell}^{RHLSE} &= A^{+}\overline{B}_{k,\ell}A^{+T} \\ &= A^{+}\left[\frac{\ell-1}{\ell}\overline{B}_{k,\ell-1} + \frac{1}{\ell}B_{k-\ell+1}\right]A^{+T} \\ &= \frac{\ell-1}{\ell}X_{k,\ell-1}^{RHLSE} + \frac{1}{\ell}X_{k-\ell+1}^{LS}, \text{âNĺ}\ell \\ &= 1, 2, \ldots, \ell_{max}, \end{aligned} \tag{25}$$

where

$$X_{k-\ell+1}^{LS} = A^{+}B_{k-\ell+1}A^{+T}. \tag{26}$$

For the stopping criterion based on the norm of iteration error

$$\Delta X_{\ell,\ell-1} \triangleq X_{k,\ell}^{RHLSE} - X_{k,\ell-1}^{RHLSE}, \text{âNĺ}\ell \geq 2, \tag{27}$$

we have the following result.

Theorem 7. *Let $X_{k,\ell}^{RHLSE}$ be the RHLSE (21). Then the Frobenius norm of the iteration error $\Delta X_{\ell,\ell-1}$ is calculated as*

$$\begin{aligned} \left\|\Delta X_{\ell,\ell-1}\right\|_F &= \frac{1}{\ell}\sqrt{tr\left(\widetilde{M}^2\Delta B_{k,\ell-1}^2\right)}, \text{âNĺ}\ell \geq 2, \\ \widetilde{M} &= A^{+T}A^{+}, \\ \Delta B_{k,\ell-1} &= \overline{B}_{k,\ell-1} - B_{k-\ell+1}. \end{aligned} \tag{28}$$

The proof of Theorem 7 is given in Appendix.

Using (28), the best HL ℓ^{*} is defined as

$$\ell^{*} = \min\left\{\ell : \left\|\Delta X_{\ell,\ell-1}\right\|_F < \varepsilon\right\}, \text{âNĺ}\ell \geq 2, \tag{29}$$

where ε is a given tolerance.

Remark 8. As noted in Section 3, the initial covariance $X_{k,1}^{RHLSE} = X_{k}^{LS}(B_k) = A^{+}B_kA^{+T}$, calculated at $\ell = 1$, may be inaccurate. Numerical examples given below confirm this fact. Therefore to improve the covariance $X_{k,1}^{RHLSE}$, we need to continue the iterative process $X_{k,\ell}^{RHLSE}$, $\ell \geq 2$, using the stopping condition (29).

Remark 9. In general, the stopping condition (29) does not guarantee the convergence of the iterative process (25); that is, the inequality $\|X - X_{k,\ell}^{RHLSE}\|_F < \rho\|X - X_{k,\ell-1}^{RHLSE}\|_F$, $\rho \in (0,1)$ may not hold. Besides, the number of iterations is limited, $\ell \leq \ell_{max}$. However, (29) seems to be still useful for stopping the iterative process, as will be demonstrated by examples.

Remark 10. The stopping condition (29) can be carried out in real time. In fact, the term $tr(\widetilde{M}^2\Delta B_{k,\ell-1}^2)$ depends on the TNCs ($\overline{B}_{k,\ell-1}$ and $B_{k-\ell+1}$) for a given horizon. Therefore, the complexity of the real-time implementation of the stopping condition is not an issue.

6. Application of RHLSE

In this section, an application of the RHLSE to the estimation of process and measurement noise covariances is considered.

6.1. Kalman Filtering. The Kalman filtering framework involves estimation of the state of a discrete-time linear dynamical system with additive white Gaussian noise,

$$\begin{aligned} x_{k+1} &= F_kx_k + G_kw_k, \\ y_k &= H_kx_k + D_kv_k, \end{aligned} \tag{30}$$

where state $x_k \in \mathbb{R}^{n_x}$, measurement $y_k \in \mathbb{R}^{n_y}$, process noise $w_k \in \mathbb{R}^{n_w}$, $w_k \sim \mathbb{N}(0, Q)$, and measurement noise $v_k \in \mathbb{R}^{n_v}$, $v_k \sim \mathbb{N}(0, R)$. All matrices $F_k, G_k, H_k, D_k, Q \geq 0$ and $R > 0$ are with appropriate dimensions. We assume that the initial state $x_0 \sim \mathbb{N}(\overline{x}_0, C_0)$, and process and measurement noises w_k, v_k are mutually uncorrelated with constant covariances (ONCs), $Q \in \mathbb{R}^{n_w \times n_w}$ and $R \in \mathbb{R}^{n_v \times n_v}$, respectively.

As mentioned in Introduction the KF equations depend on the known TNCs $\widetilde{Q}_k = G_k Q G_k^T$ and $\widetilde{R}_k = D_k R D_k^T$, rather than the ONCs $\{Q, R\}$. We have

Prediction step:

$$\begin{aligned} \widehat{x}_k^- &= F_{k-1}\widehat{x}_{k-1}, \quad \widehat{x}_0 = \overline{x}_0, \\ P_k^- &= F_{k-1}P_{k-1}F_{k-1}^T + \widetilde{\mathbf{Q}}_{\boldsymbol{k-1}}, \\ P_0 &= C_0, \end{aligned} \tag{31}$$

Correction step:

$$\begin{aligned} K_k &= P_k^- H_k^T \left(H_k P_k^- H_k^T + \widetilde{\boldsymbol{R}}_{\boldsymbol{k}}\right)^{-1}, \\ \widehat{x}_k &= \widehat{x}_k^- + K_k\left(y_k - H_k\widehat{x}_k^-\right), \\ P_k &= \left(I_{n_x} - K_k H_k\right)P_k^-. \end{aligned}$$

6.2. Adaptive Kalman Filtering. In a real problem the ONCs and, as a consequence, the TNCs are unknown. In this case the AKF simultaneously estimates both the state $\widehat{x}_k$ and TNCs $\{\widetilde{Q}, \widetilde{R}\}$. The general scheme for adaptive estimation of the TNCs can be presented in Table 2.

Then the AKF equations for the system model (30) are given by

$$\begin{aligned} \textit{Initial values:}\quad & \widehat{x}_0 = \overline{x}_0, \\ & P_0 = C_0, \\ & Q_0 > 0, \\ & R_0 > 0. \\ & \widetilde{\mathbf{Q}}_{\mathbf{0}} = G_0 Q_0 G_0^T, \\ & \widetilde{\boldsymbol{R}}_{\mathbf{0}} = D_0 R_0 D_0^T. \\ & \widehat{x}_k^- = F_{k-1}\widehat{x}_{k-1}, \\ & P_k^- = F_{k-1}P_{k-1}F_{k-1}^T + \widetilde{\mathbf{Q}}_{\boldsymbol{k-1}}, \\ & \widetilde{y}_k = y_k - H_k\widehat{x}_k^-, \\ & C_{\widetilde{y}_k} = \frac{1}{k_0}\sum_{t=k-k_0+1}^{k}\widetilde{y}_t\widetilde{y}_t^T, \\ & \widetilde{\boldsymbol{R}}_{\boldsymbol{k}} = C_{\widetilde{y}_k} - H_k P_k^- H_k^T, \\ & K_k = P_k^- H_k^T\left(H_k P_k^- H_k^T + \widetilde{\boldsymbol{R}}_{\boldsymbol{k}}\right)^{-1}, \\ & \widehat{x}_k = \widehat{x}_k^- + K_k\left(y_k - H_k\widehat{x}_k^-\right), \\ & P_k = \left(I_{n_x} - K_k H_k\right)P_k^-, \\ & \widetilde{x}_k = \widehat{x}_k - \widehat{x}_k^-, \\ & \widetilde{\mathbf{Q}}_{\boldsymbol{k}} = \frac{1}{n_0}\sum_{t=k-n_0+1}^{k}\widetilde{x}_t\widetilde{x}_t^T. \end{aligned} \tag{32}$$

To estimate the unknown ONCs $\{Q, R\}$ using the available TNCs $\{\widetilde{Q}_k, \widetilde{R}_k\}$ we apply the RHLSE. Equation (21) for the corresponding estimates $\widehat{R}_{k,\ell} \triangleq R_{k,\ell}^{RHLSE}$ and $\widehat{Q}_{k,\ell} \triangleq Q_{k,\ell}^{RHLSE}$ are

$$\begin{aligned} \widehat{R}_{k,\ell} &= D^{+}\overline{B}_{k,\ell}^{r}D^{+T}, \\ \overline{B}_{k,\ell}^{r} &= \frac{1}{\ell}\sum_{t=k-\ell+1}^{k}\widetilde{R}_t, \\ \widehat{Q}_{k,\ell} &= G^{+}\overline{B}_{k,\ell}^{q}G^{+T}, \\ \overline{B}_{k,\ell}^{q} &= \frac{1}{\ell}\sum_{t=k-\ell+1}^{k}\widetilde{Q}_t. \end{aligned} \tag{33}$$

Here $\widetilde{Q}_t$ and $\widetilde{R}_t$ are calculated in (32), and the best HLs $\ell^* = \ell^r$ for $\widehat{R}_{k,\ell}$ and $\ell^* = \ell^q$ for $\widehat{Q}_{k,\ell}$ are determined by (29), i.e.,

$$\begin{aligned} \ell^r &= \min\left\{\ell : \left\|\Delta R_{\ell,\ell-1}\right\|_F < \varepsilon\right\}, \quad \ell \geq 2, \\ \left\|\Delta R_{\ell,\ell-1}\right\|_F &= \frac{1}{\ell}\sqrt{tr\left(\widetilde{M}_D^2\Delta\widetilde{R}_{k,\ell-1}^2\right)}, \\ \widetilde{M}_D &= D^{+T}D^{+}, \\ \Delta\widetilde{R}_{k,\ell-1} &= \overline{B}_{k,\ell-1}^{r} - \widetilde{R}_{k-\ell+1}, \end{aligned} \tag{34}$$

and

$$\begin{aligned} \ell^q &= \min\left\{\ell : \left\|\Delta Q_{\ell,\ell-1}\right\|_F < \varepsilon\right\}, \quad \ell \geq 2, \\ \left\|\Delta Q_{\ell,\ell-1}\right\|_F &= \frac{1}{\ell}\sqrt{tr\left(\widetilde{M}_G^2\Delta\widetilde{Q}_{k,\ell-1}^2\right)}, \\ \widetilde{M}_G &= G^{+T}G^{+}, \\ \Delta\widetilde{Q}_{k,\ell-1} &= \overline{B}_{k,\ell-1}^{q} - \widetilde{Q}_{k-\ell+1}, \end{aligned} \tag{35}$$

respectively.

7. Simulation Examples

In this section, we first test the efficiency of the proposed RHLSE and the stopping criterion on a theoretical example for which the true ONC is known (Section 7.1). And then, performance of the RHLSEs (33)-(35) shall be demonstrated with the example of a moving object (Section 7.2).

7.1. Theoretical Example. Let us estimate a random constant $\theta \sim \mathbb{N}(m_\theta, \sigma_\theta^2)$ given two single sensor measurements $y_{1,k}$ and $y_{2,k}$ of θ corrupted by their own uncorrelated Gaussian white noises (measurement errors) $v_{1,k}$ and $v_{2,k}$, respectively. We assume that the two sensors contain a common environmental Gaussian white noise η_k, which is uncorrelated with $v_{1,k}$

Table 2: Adaptive estimation of transformed noise covariances.

	R estimation	Q estimation
Sequence	Innovation sequence $\tilde{y}_k = y_k - H_k\hat{x}_k^-$	Residual sequence $\tilde{x}_k = \hat{x}_k - \hat{x}_k^-$
Sample covariance	$C_{\tilde{y}_k} = \frac{1}{k_0}\sum_{t=k-k_0+1}^{k} \tilde{y}_t\tilde{y}_t^T$	$C_{\tilde{x}_k} = \frac{1}{k_0}\sum_{t=k-k_0+1}^{k} \tilde{x}_t\tilde{x}_t^T$
Estimated noise covariance (TNC)	$\widetilde{R}_k = C_{\tilde{y}_k} - H_kP_k^-H_k^T$	$\widetilde{Q}_k = C_{\tilde{x}_k}$

and $v_{2,k}$. Then the dynamic equations describing this situation are

$$\begin{aligned} &\text{State:} \quad x_{k+1} = x_k, \\ &\qquad k = 0, 1, \ldots; \ x_0 = \theta \sim \mathbb{N}\left(m_\theta, \sigma_\theta^2\right), \\ &\text{Sensor 1:} \quad y_{1,k} = x_k + v_{1,k} + \eta_k, \\ &\qquad v_{1,k} \sim \mathbb{N}\left(0, r_1\right), \ \eta_k \sim \mathbb{N}\left(0, \sigma_\eta^2\right), \\ &\text{Sensor 2:} \quad y_{2,k} = x_k + v_{2,k} + \eta_k, \quad v_{1,k} \sim \mathbb{N}\left(0, r_2\right). \end{aligned} \tag{36}$$

Here we calculate the estimation accuracy of the proposed RHLSE in terms of the receding HL, $\ell \geq 1$, for the two measurement models.

Model 1 ("measurement model without environmental noise η_k").

$$\begin{aligned} y_{1,k} &= x_k + v_{1,k}, \\ y_{2,k} &= x_k + v_{2,k}, \end{aligned} \tag{37}$$

or in matrix form

$$\begin{aligned} Y_k &= Hx_k + D^{(1)}V_k^{(1)}, \\ Y_k &= \begin{bmatrix} y_{1,k} \\ y_{2,k} \end{bmatrix}, \\ H &= \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \\ D^{(1)} &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \\ V_k^{(1)} &= \begin{bmatrix} v_{1,k} \\ v_{2,k} \end{bmatrix} \sim \mathbb{N}\left(0, R^{(1)}\right). \end{aligned} \tag{38}$$

Here the ONC and TNC take the form

$$\begin{aligned} R^{(1)} &= \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix}, \\ \widetilde{R}^{(1)} &= D^{(1)}R^{(1)}D^{(1)^T} = \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix}, \end{aligned} \tag{39}$$

respectively.

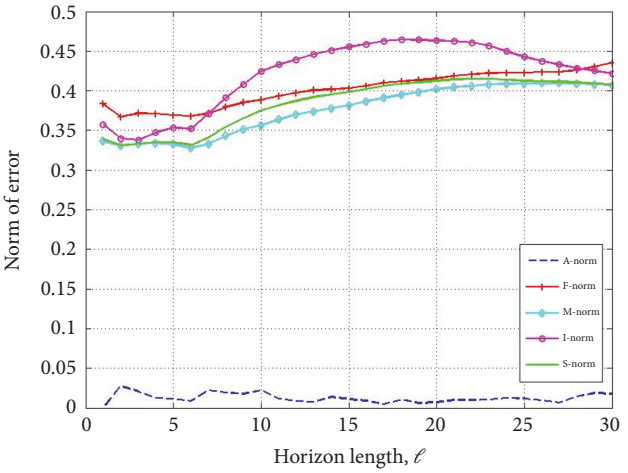

Figure 1: Norms of true and iteration errors for Model 1.

For the model (36)-(39) the measurement noise gain matrix $D \triangleq D^{(1)} = I_2$, and $D^+ \triangleq D^{(1)+} = I_2$ and $\widetilde{M}_D = D^{+T}D^+ = I_2$. Then using (33) and (34) the RHLSE $\widehat{R}_{k,\ell} \triangleq R_{k,\ell}^{RHLSE}$ for the measurement noise covariance $X \triangleq R^{(1)}$ and the best HL $\ell^* \triangleq \ell^r$ take the following form,

$$\begin{aligned} \widehat{R}_{k,\ell} &= \overline{B}_{k,\ell}^r = \frac{1}{\ell}\sum_{t=k-\ell+1}^{k} \widetilde{R}_t^{(1)}, \quad \ell = 2, 3, \ldots, \ell^*, \\ \ell^* &= \min\left\{\ell : \left\|\Delta\widetilde{R}_{\ell,\ell-1}\right\|_F < \varepsilon\right\}, \end{aligned} \tag{40}$$

where

$$\begin{aligned} \left\|\Delta\widetilde{R}_{\ell,\ell-1}\right\|_F &= \frac{1}{\ell}\sqrt{tr\left(\Delta\widetilde{R}_{k,\ell-1}^2\right)}, \quad \ell \geq 2, \\ \Delta\widetilde{R}_{k,\ell-1} &= \overline{B}_{k,\ell-1}^r - \widetilde{R}_{k-\ell+1}^{(1)}, \\ \widetilde{R}_t^{(1)} &= C_{\tilde{y}_t} - HP_t^-H^T, \\ C_{\tilde{y}_t} &= \frac{1}{t-t_0}\sum_{i=t_0+1}^{t} \widetilde{Y}_i\widetilde{Y}_i^T, \quad t = k-\ell+1, , \ldots, k, \\ \widetilde{Y}_i &= Y_i - H\hat{x}_{i-1}, \quad i = t_0+1, \ldots, t, \end{aligned} \tag{41}$$

and t_0 is chosen empirically to give some statistical smoothing.

The simulation results are illustrated in Table 3 and Figure 1 for $k = 50$, $r_1 = 0.1$, $r_2 = 0.8$, and $\ell_{max} = 30$. Our

Table 3: Norms of true and iteration errors for Model 1.

HL ℓ	Absolute and relative norms for iteration error $\Delta R_{\ell,\ell-1}$		Norm of true error $\|*\| = \|E_\ell^{true}\|$			
	A-norm $\|\Delta R_{\ell,\ell-1}\|_F$	R-norm $\frac{\|\Delta R_{\ell,\ell-1}\|_F}{\|\widehat{R}_{k,\ell}\|_F}$	F-norm $\|*\|_F$	M-norm $\|*\|_{max}$	I-norm $\|*\|_\infty = \|*\|_1$	S-norm $\|*\|_2$
I	II	III	IV	V	VI	VII
1	- - -	- - -	0.3832	0.3365	0.3572	0.3392
2	0.0276	0.0417	0.3674	0.3308	0.3398	**0.3312**
3	0.0207	0.0316	0.3717	0.3330	**0.3381**	0.3331
4	0.0122	0.0185	0.3712	0.3341	0.3472	0.3351
5	0.0114	0.0171	0.3693	0.3325	0.3534	0.3350
6	**0.0085**	**0.0129**	**0.3680**	**0.3279**	0.3520	**0.3314**
7	0.0220	0.0330	0.3716	0.3330	0.3713	0.3410
8	0.0194	0.0287	0.3793	0.3434	0.3912	0.3544
9	0.0180	0.0263	0.3852	0.3515	0.4080	0.3654
10	0.0222	0.0318	0.3887	0.3565	0.4245	0.3748
11	0.0114	0.0162	0.3933	0.3637	0.4331	0.3816
12	0.0084	0.0118	0.3974	0.3700	0.4391	0.3871
13	0.0076	0.0106	0.4008	0.3740	0.4459	0.3919
14	0.0137	0.0188	0.4017	0.3778	0.4511	0.3953
15	0.0110	0.0149	0.4033	0.3817	0.4553	0.3986
16	0.0090	0.0121	0.4063	0.3868	0.4590	0.4025
17	0.0046	0.0062	0.4103	0.3912	0.4625	0.4064
18	0.0103	0.0138	0.4120	0.3951	0.4647	0.4091
19	0.0062	0.0083	0.4135	0.3982	0.4642	0.4107
20	0.0073	0.0097	0.4155	0.4017	0.4631	0.4124
21	0.0102	0.0138	0.4188	0.4045	0.4628	0.4144
22	0.0099	0.0135	0.4209	0.4064	0.4608	0.4151
23	0.0103	0.0142	0.4223	0.4080	0.4567	0.4152
24	0.0125	0.0175	0.4226	0.4086	0.4496	0.4138
25	0.0117	0.0165	0.4230	0.4090	0.4428	0.4127
26	0.0096	0.0136	0.4236	0.4095	0.4372	0.4120
27	0.0068	0.0097	0.4238	0.4097	0.4328	0.4114
28	0.0143	0.0207	0.4264	0.4092	0.4284	0.4105
29	0.0188	0.0280	0.4307	0.4082	0.4251	0.4093
30	0.0180	0.0275	0.4353	0.4071	0.4217	0.4079

point of interest are the values of different norms of the true error $E_\ell^{true} \triangleq X - X_{k,\ell}^{RHLSE} = R^{(1)} - \widehat{R}_{k,\ell}$, and the best HL ℓ^*. In Table 3, a sign $\|*\|$ represents one of the five considered matrix norms, $\|*\|_F$, $\|*\|_M$, $\|*\|_\infty$, $\|*\|_1$, and $\|*\|_2$ (Note: $\|*\|_\infty = \|*\|_1$ for symmetric matrices). The columns II and III show the absolute norm (A-norm) and the relative norm (R-norm) of the iteration error $\Delta R_{\ell,\ell-1} = \widehat{R}_{k,\ell} - \widehat{R}_{k,\ell-1}$, respectively.

Comments on Table 3 and Figure 1 are in order.

(1a) The values of norms in Table 3 confirm the known inequalities for matrix norms $\|A\|$ of the symmetric matrix $A = E_\ell^{true} \in \mathbb{R}^{n\times n}$, i.e.,

$$\|A\|_{max} \le \|A\|_2 \le \|A\|_F \le \sqrt{n}\,\|A\|_2\,, \qquad \|A\|_2 \le \|A\|_\infty\,. \tag{42}$$

(1b) The best HL for each of the presented norms is the following, $\ell_A^* = \ell_R^* = \ell_F^* = \ell_M^* = 6$; $\ell_I^* = 3$; and

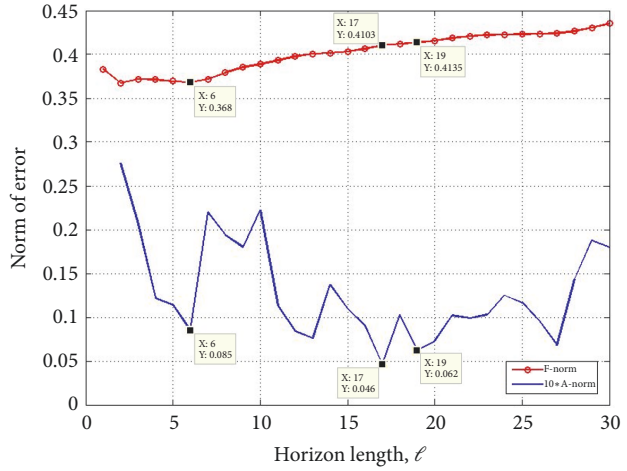

FIGURE 2: Selection of best HL for Model 1: comparison of true error (F-norm) and absolute iteration error (A-norm).

$\ell_S^* = 2$ or 6. So a small HL, $2 \le \ell \le 6$, is suitable for Model 1 in which the sizes of the ONC and TNC are equal, $R^{(1)}, \widetilde{R}^{(1)} \in \mathbb{R}^{2\times 2}$.

(1c) The values of all norms of the true error E_ℓ^{true} confirm Remark 8 (see columns IV~VII). They show that the initial HL $\ell = 1$ is not optimal, i.e., the single ONC $\widehat{R}_{k,1}$ is not good.

(1d) The curves of all norms of the true error have a similar shape. The curves slightly fluctuate and then increase around a small HL, $\ell < 8$.

(1e) The curves of the A- and R-norms are similar. Therefore, only one of them is shown in Figure 1.

(1f) As noted in Remark 9, fulfillment of the condition $\|\Delta R_{\ell,\ell-1}\|_F < \varepsilon$ does not guarantee the convergence. In Table 3 we can observe that the small iteration errors $\|\Delta R_{19,18}\|_F < \|\Delta R_{20,19}\|_F < \|\Delta R_{6,5}\|_F = 0.0085 < \varepsilon = 0.01$, correspond to the large true errors, $\|E_{20}^{true}\|_F > \|E_{19}^{true}\|_F > \|E_6^{true}\|_F$. This fact is also illustrated in Figure 2, where F-norm ($\|E_\ell^{true}\|_F$) and scaled A-norm ($10\|\Delta X_{\ell,\ell-1}\|_A$) are plotted for a clear comparison.

Next, we consider more complicated Model 2, in which the noise gain matrix $D^{(2)}$ is rectangular, and the sizes of the original $R^{(2)}$ and transformed $\widetilde{R}^{(2)}$ measurement covariances are different. In this case (12) represents an undefined system of equations.

Model 2 ("measurement model with environmental noise η_k").

$$\begin{aligned} y_{1,k} &= x_k + v_{1,k} + \eta_k, \\ y_{2,k} &= x_k + v_{2,k} + \eta_k, \end{aligned} \tag{43}$$

or in matrix form

$$\begin{aligned} Y_k &= Hx_k + D^{(2)}V_k^{(2)}, \\ Y_k &= \begin{bmatrix} y_{1,k} \\ y_{2,k} \end{bmatrix}, \\ H &= \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \\ D^{(2)} &= \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \\ V_k^{(2)} &= \begin{bmatrix} v_{1,k} \\ v_{2,k} \\ \eta_k \end{bmatrix} \sim \mathbb{N}\left(0, R^{(2)}\right). \end{aligned} \tag{44}$$

Here the ONC and TNC take the form

$$\begin{aligned} R^{(2)} &= \begin{bmatrix} r_1 & 0 & 0 \\ 0 & r_2 & 0 \\ 0 & 0 & \sigma_\eta^2 \end{bmatrix}, \\ \widetilde{R}^{(2)} &= D^{(2)}R^{(2)}D^{(2)^T} = \begin{bmatrix} r_1 + \sigma_\eta^2 & \sigma_\eta^2 \\ \sigma_\eta^2 & r_2 + \sigma_\eta^2 \end{bmatrix}, \end{aligned} \tag{45}$$

respectively.

For the model (43)-(45),

$$\begin{aligned} D &\triangleq D^{(2)} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \\ D^+ &\triangleq D^{(2)+} = \frac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \\ 1 & 1 \end{bmatrix}, \end{aligned} \tag{46}$$

and the RHLSE (33) takes the form $\widehat{R}_{k,\ell} = D^+\overline{B}_{k,\ell}^r D^{+T}$.

Setting up $k = 50$, $r_1 = 0.1$, $r_2 = 0.8$, and $\sigma_\eta^2 = 0.3$, we calculate the norms of iteration error $\Delta R_{\ell,\ell-1}$ and true error E_ℓ^{true}. The simulation results are illustrated in Table 4 and Figure 3.

In general, comments (1a)~(1f) made for Model 1 are the same for Table 4 and Figure 3, but there are some minor changes concerning (1b) and (1d).

(2b) The best HLs for the norms are little different, i.e., $\ell_A^* = \ell_R^* = 9$, $\ell_F^* = 5$, $\ell_M^* = \ell_I^* = 6$, $\ell_S^* = 7$. So a small HL, $\ell \le 9$, is also suitable for Model 2, when the noise gain matrix D is rectangular.

(2d) The curves of all norms of the true error have a similar shape. The curves increase and then slightly decrease around a medium HL, $9 < \ell < 20$.

Here is a summary of the simulation results presented for Models 1 and 2.

Table 4: Norms of true and iteration errors for Model 2.

HL ℓ	Absolute and relative norms for iteration error $\Delta R_{\ell,\ell-1}$		Norm of true error $\|*\| = \|E_\ell^{true}\|$			
	A-norm $\|\Delta R_{\ell,\ell-1}\|_F$	R-norm $\frac{\|\Delta R_{\ell,\ell-1}\|_F}{\|\widehat{R}_{k,\ell}\|_F}$	F-norm $\|*\|_F$	M-norm $\|*\|_{max}$	I-norm $\|*\|_\infty = \|*\|_1$	S-norm $\|*\|_2$
I	II	III	V	VI	VII	VIII
1	- - -	- - -	0.1813	0.1151	0.2303	0.1724
2	0.0858	0.1636	0.1230	0.0651	0.1303	0.1197
3	0.0726	0.1248	0.0752	0.0478	0.0957	0.0727
4	0.0582	0.0926	0.0549	0.0317	0.0635	0.0541
5	0.0105	0.0165	**0.0532**	0.0277	0.0555	0.0501
6	0.0095	0.0147	0.0543	**0.0272**	**0.0544**	0.0476
7	0.0146	0.0222	0.0608	0.0280	0.0561	**0.0457**
8	0.0237	0.0359	0.0713	0.0317	0.0634	0.0520
9	**0.0089**	**0.0136**	0.0786	0.0364	0.0728	0.0582
10	0.0207	0.0319	0.0894	0.0475	0.0950	0.0726
11	0.0110	0.0171	0.0990	0.0544	0.1088	0.0825
12	0.0115	0.0181	0.1024	0.0601	0.1202	0.0908
13	0.0147	0.0236	0.1100	0.0678	0.1357	0.1020
14	0.0162	0.0267	0.1179	0.0753	0.1507	0.1130
15	0.0122	0.0203	0.1241	0.0806	0.1613	0.1210
16	0.0079	0.0133	0.1308	0.0852	0.1705	0.1280
17	0.0084	0.0143	0.1319	0.0868	0.1736	0.1305
18	0.0112	0.0195	0.1335	0.0884	0.1768	0.1332
19	0.0143	0.0255	0.1361	0.0896	0.1793	0.1361
20	0.0125	0.0227	0.1378	0.0897	0.1795	0.1376
21	0.0092	0.0167	0.1395	0.0900	0.1801	0.1395
22	0.0099	0.0176	0.1360	0.0878	0.1757	0.1358
23	0.0103	0.0181	0.1320	0.0852	0.1705	0.1312
24	0.0109	0.0188	0.1269	0.0817	0.1635	0.1250
25	0.0115	0.0194	0.1241	0.0794	0.1589	0.1206
26	0.0166	0.0274	0.1207	0.0757	0.1515	0.1141
27	0.0122	0.0197	0.1187	0.0726	0.1453	0.1091
28	0.0150	0.0237	0.1165	0.0684	0.1369	0.1027
29	0.0123	0.0191	0.1181	0.0662	0.1325	0.0996
30	0.0206	0.0313	0.1216	0.0625	0.1250	0.0952

(i) We propose to use the RHLSE $X_{k,\ell}^{RHLSE} = A^+\overline{B}_{k,\ell}A^{+T}$ with a small HL, $2 \le \ell \le 9$, for both square and rectangular noise matrices $A = G$ or D. Usage of short ($\ell = 1$) and long ($\ell \gg 1$) horizon lengths may lead to unsatisfactory results.

(ii) In parallel with testing the efficiency of the stopping criterion we also study the behavior of the mean square errors (MSEs) for the covariance estimates $X_{k,\ell}^{RHLSE} \in \mathbb{R}^{2\times 2}$ (Model 1) and $X_{k,\ell}^{RHLSE} \in \mathbb{R}^{3\times 3}$ (Model 2),

$$MSE_{k,ij} \triangleq \boldsymbol{E}\left[\left(X_{ij} - X_{k,\ell;ij}^{RHLSE}\right)^2\right], \quad i,j = 1,\ldots,p, \tag{47}$$

where X_{ij} is the true simulated value, $X_{k,\ell;ij}^{RHLSE}$ is the RHLSE estimate of the element X_{ij}, $k(= 0,1,\ldots)$ is the time instant, and $\ell = \ell^*$ is the best HL,

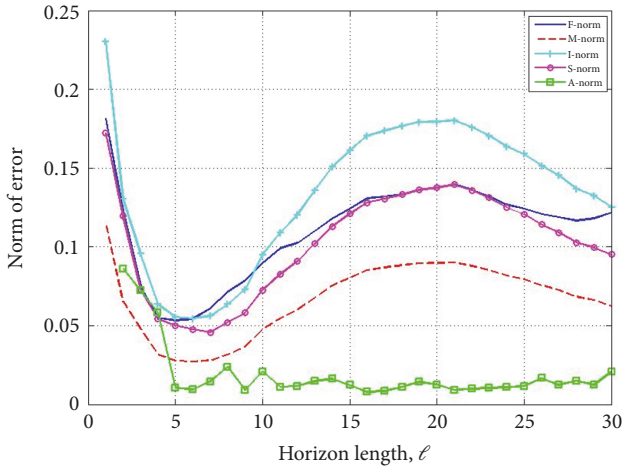

Figure 3: Norms of true and iteration errors for Model 2.

$X = [X_{ij}]$, $X_{k,\ell}^{RHLSE} = [X_{k,\ell;ij}^{RHLSE}]$. To calculate the MSEs the Monte-Carlo simulation with 1000 runs was used, i.e., the expectation operator in (47) is taken for the 1000 error data at each time instant k. We observe that the relative errors, $\Delta MSE_{k,ii}^{(1)}$, for Model 1,

$$\begin{aligned} \Delta MSE_{k,ii}^{(1)} &= \frac{MSE_{k,ii}}{X_{ii}} \times 100\%, \quad i = 1, 2, \\ X_{11} &= r_1, \\ X_{22} &= r_2, \\ \ell^* &= 6, \end{aligned} \tag{48}$$

range from 10.8% to 13.7% within the time zone $k \in [0, 50]$, and then they decrease. At $k > 500$ the maximum values of the $\Delta MSE_{k,11}^{(1)}$ and $\Delta MSE_{k,22}^{(1)}$ reach the values of 7.4% and 8.8%, respectively. Similar results, but slightly worse due to the rectangularity of the noise gain matrix $D^{(2)}$, are obtained for Model 2. For this model the relative errors,

$$\begin{aligned} \Delta MSE_{k,ii}^{(2)} &= \frac{MSE_{k,ii}}{X_{ii}} \times 100\%, \quad i = 1, 2, 3, \\ X_{11} &= r_1, \\ X_{22} &= r_2, \\ X_{33} &= \sigma_\eta^2, \\ \ell^* &= 6, \end{aligned} \tag{49}$$

range from 12% to 16.5% within the time zone $k \in [0, 50]$. At $k > 500$ the maximum values of the $\Delta MSE_{k,11}^{(2)}$, $\Delta MSE_{k,22}^{(2)}$, and $\Delta MSE_{k,33}^{(2)}$ amount to 7.1%, 8.4%, and 9.2%, respectively. Thus, this analysis of the MSEs shows that the RHLSE is suitable for practical applications.

7.2. Numerical Example: 1D Tracking with Random Velocity. In this section, the proposed RHLSE is used in adaptive filtering with unknown ONC (Q and R).

Let the true target state at discrete time t_k be defined as $x_k = [p_k \; v_k]^T$, where p_k and v_k denote the position and velocity in Cartesian coordinates, respectively. Assume that the constant velocity is subjected to a random disturbance w_k. The measured position can be defined as $y_k = p_k + \xi_k + \eta_k$, where ξ_k is a sensory error and η_k is an environmental noise. The target state dynamics is then modeled as follows:

$$\begin{aligned} x_{k+1} &= Fx_k + Gw_k, \quad x_k \in \mathbb{R}^2, \\ y_k &= Hx_k + De_k, \quad y_k \in \mathbb{R}, \end{aligned}$$

where

$$\begin{aligned} &w_k \in \mathbb{R}, \\ &w_k \sim \mathbb{N}(0, Q), \\ &e_k \in \mathbb{R}^2, \\ &e_k \sim \mathbb{N}(0, R), \\ &x_k = \begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix} = \begin{bmatrix} p_k \\ v_k \end{bmatrix}, \\ &F = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}, \\ &G = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \\ &e_k = \begin{bmatrix} \xi_k \\ \eta_k \end{bmatrix}, \\ &H^T = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \\ &D^T = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \\ &R = \operatorname{diag}\{r_\xi \;\; r_\eta\}, \\ &x_0 \sim \mathbb{N}(\overline{x}_0, C_0), \\ &\Delta t \ll 1. \end{aligned} \tag{50}$$

Simulation results with usage of the AKF (32) and RHLSE (12) are illustrated in Figures 4–8 for the following model parameters:

$$\begin{aligned} \Delta t &= 0.005, \\ \overline{x}_0 &= [4 \;\; 10]^T, \\ C_0 &= \operatorname{diag}\{10 \;\; 2.5\}, \\ Q &= 0.25, \\ \ell_{max} &= 100, \end{aligned}$$

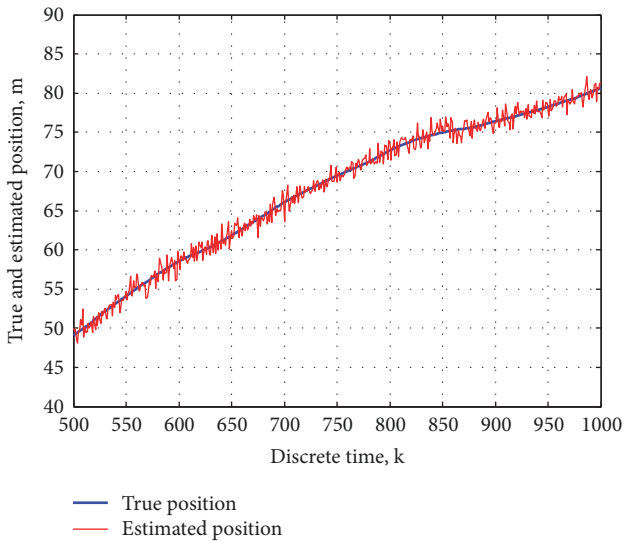

Figure 4: True and estimated position, $k \in [500, 1000]$.

$$k \in [0; 1000],$$
$$r_\xi = 0.64,$$
$$r_\eta = 0.09. \tag{51}$$

Figures 4 and 5 show the position and velocity estimates compared to the true values. The AKF appears to converge, in the sense that the estimation error tends to decrease, but slowly for velocity because the sensor measures only the position.

The key value in adaptive filtering is the TNC $\widetilde{R}_k = C_{\widetilde{y}_k} - H_k P_k^- H_k^T$. This covariance is plotted in Figure 6.

As we observe, the covariance $\widetilde{R}_k$ is not seriously changing for $k \geq 500$. For this reason we fixed the receding horizon interval $[k-\ell+1, k] = [500-\ell+1, 500]$ for further estimation of the ONCs R and Q by formulas (33). We have

$$\widehat{R}_{k,\ell} = D^+ \overline{B}^r_{k,\ell} D^{+T},$$
$$\widehat{Q}_{k,\ell} = G^+ \overline{B}^q_{k,\ell} G^{+T},$$
$$D^+ = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, \; G^+ = [0 \;\; 1], \; k = 500, \; \ell \geq 1. \tag{52}$$

The best HLs ℓ^r and ℓ^q are calculated based on the iteration errors $\Delta\widetilde{R}_{\ell,\ell-1}$ and $\Delta\widetilde{Q}_{\ell,\ell-1}$, respectively,

$$\ell^r = \min\left\{\ell : \left\|\Delta\widetilde{R}_{\ell,\ell-1}\right\|_F < \varepsilon = 10^{-5}\right\},$$
$$\left\|\Delta\widetilde{R}_{\ell,\ell-1}\right\|_F = \frac{1}{\ell}\sqrt{tr\left(\widetilde{M}_D^2 \Delta\widetilde{R}^2_{k,\ell-1}\right)}, \quad \widetilde{M}_D = \frac{1}{2}, \tag{53}$$

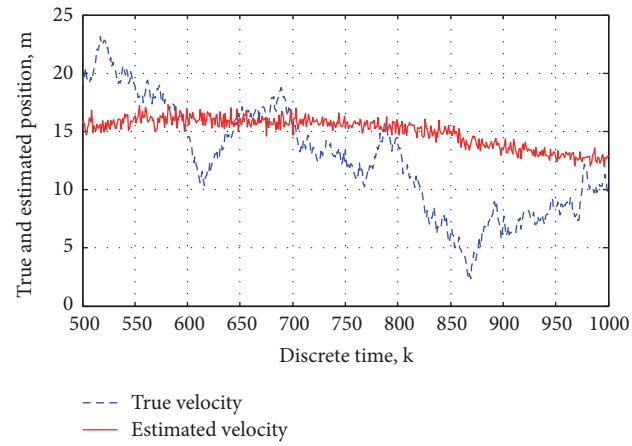

Figure 5: True and estimated velocity, $k \in [500, 1000]$.

Table 5: Best HLs for measurement noise covariance.

ℓ^r	$\left\|\Delta\widetilde{R}_{\ell,\ell-1}\right\|_F$	$\left\|E^R_\ell\right\|_F$
6	3.9237E-06	0.0070463
46	9.39816E-07	0.0070258
52	2.8749E-06	0.0069196
54	2.1627E-07	0.0070807
58	1.0062E-06	0.0070107

and

$$\ell^q = \min\left\{\ell : \left\|\Delta\widetilde{Q}_{\ell,\ell-1}\right\|_F < \varepsilon = 2 \bullet 10^{-5}\right\},$$
$$\left\|\Delta\widetilde{Q}_{\ell,\ell-1}\right\|_F = \frac{1}{\ell}\sqrt{tr\left(\widetilde{M}_G^2 \Delta\widetilde{Q}^2_{k,\ell-1}\right)}, \quad \widetilde{M}_G = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{54}$$

As indicated earlier in Remark 9, the best HLs ℓ^r and ℓ^q do not guarantee a minimum of the true errors E^R_ℓ and E^Q_ℓ, respectively. Therefore, in addition we calculate the norms of the true errors, $\|E^R_\ell\|_F$ and $\|E^Q_\ell\|_F$, which confirm that the proposed idea of choosing the best HL using the iterative error (53) or (54) leads to good results. Figure 7 presents the norms of the true error $\|E^R_\ell\|_F$ and the iteration error $\|\Delta\widetilde{R}_{\ell,\ell-1}\|_F$ for measurement covariance, respectively. Analogously, Figure 8 presents the norms of the true error $\|E^Q_\ell\|_F$ and iteration error $\|\Delta\widetilde{Q}_{\ell,\ell-1}\|_F$ for process covariance, respectively. We observe that as well as in the theoretical example (Section 7) the obtained best HLs $\ell^r = 6$ and $\ell^q = 5$ are located within a small horizon interval. In Figure 8, both curves show one value for the best HL, $\ell^q = 5$. But in Figure 7 we observe five horizon lengths $\ell^r = 6, 46, 52, 54\ 58$, which satisfy condition $\|\Delta\widetilde{R}_{\ell,\ell-1}\|_F < 10^{-5}$ within the horizon interval $\ell \in [1; 100]$, and the last four lie in the middle of the horizon interval (see Table 5). Table 5 shows that the iteration errors, $\|\Delta\widetilde{R}_{\ell,\ell-1}\|_F$, are different; however, the true errors $\|E^R_\ell\|_F$ remain almost the same. Hence, it can be concluded that the small length $\ell^r = 6$ is acceptable.

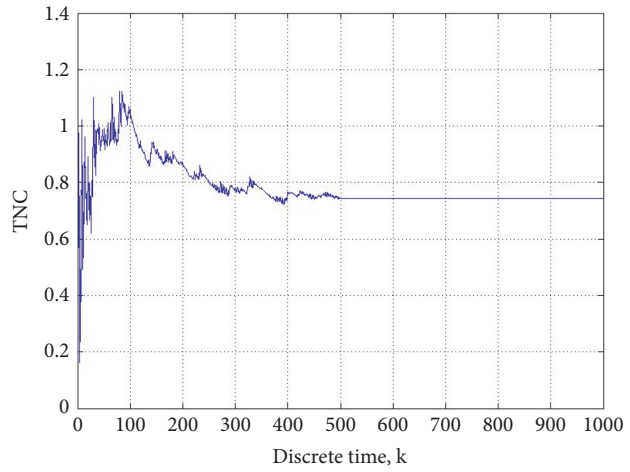

Figure 6: Transformed measurement noise covariance (variance) $\widetilde{R}_k$.

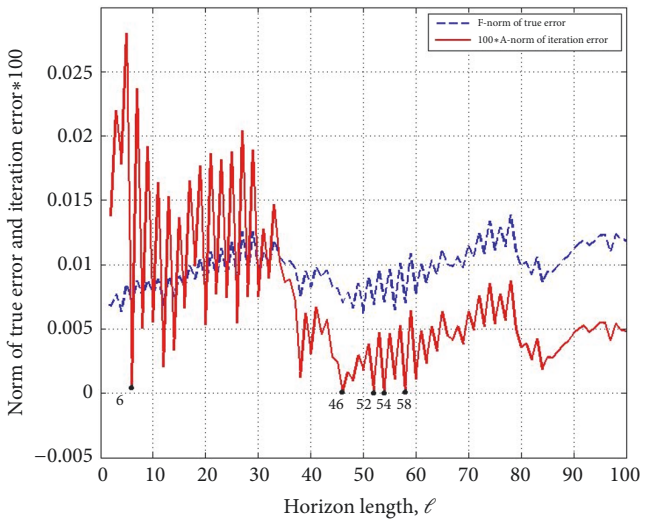

Figure 7: Best HL $\ell^r = 6$ based on minimum of iteration error $\Delta\widetilde{R}_{\ell,\ell-1}$: comparison of the norms of true error and iteration error (scaled by 100 times).

To compare the relative MSEs,

$$\begin{aligned}
\Delta MSE_k^{(Q)} &= \frac{MSE_k^{(Q)}}{Q} \times 100\%, \\
\Delta MSE_k^{(\xi)} &= \frac{MSE_k^{(\xi)}}{r_\xi} \times 100\%, \\
\Delta MSE_k^{(\eta)} &= \frac{MSE_k^{(\eta)}}{r_\eta} \times 100\%, \\
\ell^q &= 5, \\
\ell^r &= 6,
\end{aligned} \tag{55}$$

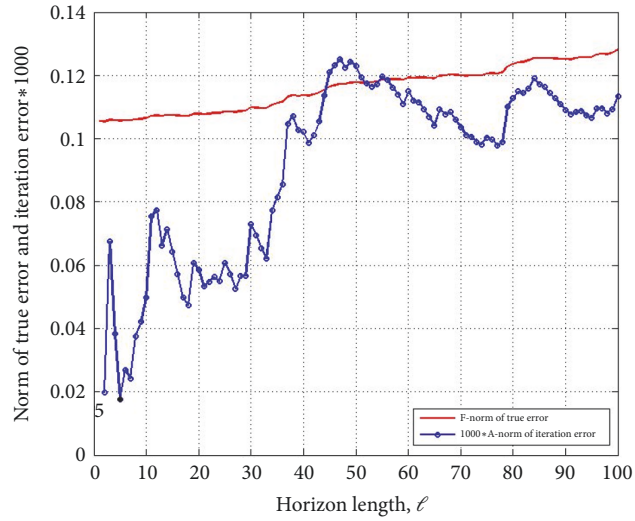

Figure 8: Best HL $\ell^q = 5$ based on minimum of iteration error $\Delta\widetilde{Q}_{\ell,\ell-1}$: comparison of the norms of true error and iteration error (scaled by 1000 times).

the Monte-Carlo simulation with 1000 runs was performed. The numerical values of all relative errors (55) within the time interval $k \in [0, 300]$ range from 13.8% to 16.7%. And for large time instants, $k > 500$, they do not change much. In this case, the maximum values of the errors $\Delta MSE_k^{(Q)}$, $\Delta MSE_k^{(\xi)}$, and $\Delta MSE_k^{(\eta)}$ amount to 14.5%, 12.7%, and 13%, respectively.

8. Conclusion

In some application problems, knowledge of noise covariances is necessary to study quality of the system and measurement model. In order to estimate arbitrary process and measurement noise covariances, a receding horizon estimation criterion with a generalized least squares approach is proposed (Theorem 4).

Special attention is given for time-invariant noise gain matrices (Theorem 5). In this case the proposed RHLSE is comprehensively investigated, including the derivation of compact matrix form for recursive RHLSE (25) and an effective stopping condition for iterative estimation process (Theorem 7). The computational procedure based on the Moore–Penrose pseudoinverse for evaluation of the best noise covariances is recommended.

In view of the importance of noise covariances in practice, the proposed RHLSE is illustrated on theoretical and practical dynamical models with different types of noises. The examples show that the RHLSE with selection of the best HL within a small horizon interval yields reasonably good estimation accuracy.

Simulation analysis and comparison of the proposed estimator by several norms of estimation error show that usage of the receding horizon strategy is a very effective approach for the achievement of robustness against any uncertainties and errors.

Appendix

Proof of Theorem 4. Using the notations $M_t = A_t^T A_t$ and $M_t^T = M_t$, $X^T = X$, the criterion (17) can be rewritten as

$$\begin{aligned} J_{k,\ell}(X) &= \sum_{t=k-\ell+1}^{k} \left\| A_t X A_t^T - B_t \right\|_F^2 \\ &= \sum_{t=k-\ell+1}^{k} tr\left[\left(A_t X A_t^T - B_t \right) \left(A_t X A_t^T - B_t^T \right) \right] \\ &= \sum_{t=k-\ell+1}^{k} \left[tr\left(A_t X M_t X A_t^T \right) - tr\left(B_t A_t X A_t^T \right) \right. \\ &\quad \left. - tr\left(A_t X A_t^T B_t^T \right) + tr\left(B_t B_t^T \right) \right]. \end{aligned} \tag{A.1}$$

Differentiating each summand of the functional $J_{k,\ell}(X)$ with respect to X using formulas of the trace and matrix derivatives, and $B_t^T = B_t$,

$$\begin{aligned} \frac{\partial}{\partial X} tr(CXD) &= C^T D^T, \\ \frac{\partial}{\partial X} tr\left(CXDXC^T \right) &= 2C^T CXD, \end{aligned} \tag{A.2}$$

we get

$$\frac{\partial J_{k,\ell}}{\partial X} = 2 \sum_{t=k-\ell+1}^{k} \left(A_t^T A_t X M_t - A_t^T B_t A_t \right). \tag{A.3}$$

Then setting the result to zero, $\partial J_{k,\ell}/\partial X = 0$, we obtain

$$\sum_{t=k-\ell+1}^{k} \left(M_t X_{k,\ell}^{RHLSE} M_t \right) = \sum_{t=k-\ell+1}^{k} \left(A_t^T B_t A_t \right). \tag{A.4}$$

This completes the proof. □

Proof of Theorem 5. Since the matrices $A_k \equiv A$ and $M_k \equiv M$ are constant, and after simple manipulations the general solution (18) takes the form

$$\begin{aligned} &\sum_{t=k-\ell+1}^{k} M X_{k,\ell}^{RHLSE} M = \sum_{t=k-\ell+1}^{k} A^T B_t A \Longleftrightarrow \\ &M \left(\sum_{t=k-\ell+1}^{k} X_{k,\ell}^{RHLSE} \right) M = A^T \left(\sum_{t=k-\ell+1}^{k} B_t \right) A \Longleftrightarrow \\ &\ell \left(M X_{k,\ell}^{RHLSE} M \right) = A^T \left(\sum_{t=k-\ell+1}^{k} B_t \right) A \Longleftrightarrow \\ &M X_{k,\ell}^{RHLSE} M = A^T \overline{B}_{k,\ell} A \Longleftrightarrow \\ &X_k^{RHLSE} = \left(M^+ A^T \right) \overline{B}_{k,\ell} \left(A M^+ \right). \end{aligned} \tag{A.5}$$

Next, using the equality of pseudoinverse $M^+ A^T = (A^T A)^+ A^T = A^+$ [18], we obtain (21).

This completes the proof. □

Proof of Theorem 7. Since $\overline{B}_{k,\ell} = \overline{B}_{k,\ell}^T = B + \Delta\overline{B}_{k,\ell}$, and using (25) we have

$$\begin{aligned} \left\| X_{k,\ell}^{RHLSE} - X_{k,\ell-1}^{RHLSE} \right\|_F^2 &= \left\| \frac{1}{\ell} \left(X_{k-\ell+1}^{LS} - X_{k,\ell-1}^{RHLSE} \right) \right\|_F^2 \\ &= \frac{1}{\ell^2} \left\| A^+ \left(\overline{B}_{k,\ell-1} - B_{k-\ell+1} \right) A^{+T} \right\|_F^2, \end{aligned} \tag{A.6}$$

or

$$\begin{aligned} \left\| \Delta X_{\ell,\ell-1} \right\|_F^2 &= \frac{1}{\ell^2} \left\| A^+ \Delta B_{k,\ell-1} A^{+T} \right\|_F^2 \\ &= \frac{1}{\ell^2} tr\left[A^+ \Delta B_{k,\ell-1} A^{+T} \left(A^+ \Delta B_{k,\ell-1} A^{+T} \right)^T \right] \\ &= \frac{1}{\ell^2} tr\left(A^+ \Delta \overline{B}_{k,\ell-1} \widetilde{M} \Delta \overline{B}_{k,\ell-1} A^{+T} \right), \\ &\qquad \widetilde{M} = A^{+T} A^+. \end{aligned} \tag{A.7}$$

Using the cyclic property of trace for arbitrary matrices we get

$$\begin{aligned} &tr\left(A^+ \Delta \overline{B}_{k,\ell-1} \widetilde{M} \Delta \overline{B}_{k,\ell-1} A^{+T} \right) \\ &\quad = tr\left(A^{+T} A^+ \Delta \overline{B}_{k,\ell-1} \widetilde{M} \Delta \overline{B}_{k,\ell-1} \right) \\ &\quad = tr\left(\widetilde{M} \Delta \overline{B}_{k,\ell-1} \widetilde{M} \Delta \overline{B}_{k,\ell-1} \right). \end{aligned} \tag{A.8}$$

Next using the cyclic property of trace of the product of three symmetric matrices $tr(L_1 L_2 L_3) = tr(L_1 L_3 L_2)$, we obtain

$$\begin{aligned} &tr\left(\underbrace{\widetilde{M} \Delta \overline{B}_{k,\ell-1}}_{L_1} \underbrace{\widetilde{M}}_{L_2} \underbrace{\Delta \overline{B}_{k,\ell-1}}_{L_3} \right) \\ &\quad = tr\left(\widetilde{M} \Delta \overline{B}_{k,\ell-1} \Delta \overline{B}_{k,\ell-1} \widetilde{M} \right) = tr\left(\widetilde{M} \Delta \overline{B}_{k,\ell-1}^2 \widetilde{M} \right) \\ &\quad = tr\left(\widetilde{M}^2 \Delta \overline{B}_{k,\ell-1}^2 \right). \end{aligned} \tag{A.9}$$

Substituting (A.8) and (A.9) into (A.7) we get

$$\left\| \Delta X_{\ell,\ell-1} \right\|_F^2 = \frac{1}{\ell^2} tr\left(\widetilde{M}^2 \Delta \overline{B}_{k,\ell-1}^2 \right). \tag{A.10}$$

This completes the proof. □

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the fund of research promotion program, Gyeongsang National University, 2017, the National

Research Foundation of Korea (NRF) funded by the Ministry of Education (Grant no. NRF-2018R1D1A3A03000717), and the Ministry of Science and ICT (Grant no. NRF-2017R1A5A1015311).

References

[1] R. G. Brown and P. Y. Hwang, *Introduction to Random Signals and Applied Kalman Filtering with Matlab Exercises*, John Wiley & Sons, New York, NY, USA, 4th edition, 2012.

[2] D. Simon, *Optimal State Estimation. Kalman, H-Infinity, and Nonlinear Approaches*, John Wiley & Sons, Hoboken, NJ, USA, 2006.

[3] N. H. Nguyen and K. l. Dogancay, "Improved pseudolinear Kalman filter algorithms for bearings-only target tracking," *IEEE Transactions on Signal Processing*, vol. 65, no. 23, pp. 6119–6134, 2017.

[4] M. V. Kulikova and J. V. Tsyganova, "Improved discrete-time Kalman filtering within singular value decomposition," *IET Control Theory & Applications*, vol. 11, no. 15, pp. 2412–2418, 2017.

[5] Y. S. Shmaliy, S. Zhao, and C. K. Ahn, "Unbiased finite impulse response filtering: an iterative alternative to Kalman filtering ignoring noise and initial conditions," *IEEE Control Systems Magazine*, vol. 37, no. 5, pp. 70–89, 2017.

[6] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*, John Wiley & Sons, New York, NY, USA, 2001.

[7] M. S. Grewal, A. P. Andrews, and C. G. Bartone, *Global Navigation Satellite Systems, Inertial Navigation, and Integration*, John Wiley & Sons, New Jersey, NJ, USA, 3rd edition, 2013.

[8] Y. Singh and R. Mehra, "Relative study of measurement noise covariance R and process noise covariance Q of the Kalman filter in estimation," *IOSR Journal of Electrical and Electronics Engineering*, vol. 10, no. 6, pp. 112–116, 2015.

[9] A. Almagbile, J. Wang, and W. Ding, "Evaluating the Performances of Adaptive Kalman Filter Methods in GPS/INS Integration," *Journal of Global Positioning Systems*, vol. 9, no. 1, pp. 33–40, 2010.

[10] A. H. Mohamed and K. P. Schwarz, "Adaptive Kalman filtering for INS/GPS," *Journal of Geodesy*, vol. 73, no. 4, pp. 193–203, 1999.

[11] R. K. Mehra, "On the Identification of Variances and Adaptive Kalman Filtering," *IEEE Transactions on Automatic Control*, vol. AC-15, no. 2, pp. 175–184, 1970.

[12] C. Hide, T. Moore, and M. Smith, "Adaptive Kalman filtering for low-cost INS/GPS," *Journal of Navigation*, vol. 56, no. 1, pp. 143–152, 2003.

[13] G. Chen, Ed., *Approximate Kalman filtering*, vol. 2 of *Series in Approximations and Decompositions*, World Scientific Publishing Co., Inc., River Edge, NJ, USA, 1993.

[14] Q. Xia, M. Rao, Y. Ying, and X. Shen, "Adaptive fading Kalman filter with an application," *Automatica*, vol. 30, no. 8, pp. 1333–1338, 1994.

[15] C. Jiang, S. Zhang, and Q. Zhang, "Adaptive Estimation of Multiple Fading Factors for GPS/INS Integrated Navigation Systems," *Sensors*, vol. 17, no. 6, pp. 1–18, 2017.

[16] W. Ding, J. Wang, C. Rizos, and D. Kinlyside, "Improving adaptive kalman estimation in GPS/INS integration," *The Journal of Navigation*, vol. 60, no. 3, pp. 517–529, 2007.

[17] M. Oussalah and J. D. Schutter, "Adaptive Kalman filter for noise identification," in *Proceedings of the International Seminar on Modal Analysis*, pp. 1225–1232, Kissimmee, FL, USA. KU Leuven, Belgium, 2001.

[18] A. J. Laub, *Matrix Analysis for Scientists Engineers*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa, USA, 2004.

[19] K.-W. E. Chu, "Symmetric solutions of linear matrix equations by matrix decompositions," *Linear Algebra and its Applications*, vol. 119, pp. 35–50, 1989.

[20] H. Dai, "On the symmetric solutions of linear matrix equations," *Linear Algebra and its Applications*, vol. 131, pp. 1–7, 1990.

[21] F. J. H. Don, "On the symmetric solutions of a linear matrix equation," *Linear Algebra and its Applications*, vol. 93, pp. 1–7, 1987.

[22] A. H. Jazwinski, "Limited Memory Optimal Filtering," *IEEE Transactions on Automatic Control*, vol. 13, no. 5, pp. 558–563, 1968.

[23] A. Alessandri, M. Baglietto, and G. Battistelli, "Receding-horizon estimation for switching discrete-time linear systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 11, pp. 1736–1748, 2005.

[24] W. H. Kwon, P. S. Kim, and P. Park, "A receding horizon Kalman FIR filter for discrete time-invariant systems," *IEEE Transactions on Automatic Control*, vol. 44, no. 9, pp. 1787–1791, 1999.

[25] B. Kwon and S. Han, "Least-Mean-Square Receding Horizon Estimation," *Mathematical Problems in Engineering*, vol. 2012, Article ID 631759, 19 pages, 2012.