# Robust Recursive Least-Squares Finite Impulse Response Predictor in Linear Discrete-Time Stochastic Systems with Uncertain Parameters

SEIICHI NAKAMORI
Kagoshima University
Department of Technical Education
1-20-6, Korimoto, Kagoshima
JAPAN
k6165161@kadai.jp

*Abstract:* - This paper newly proposes the robust RLS Wiener FIR prediction algorithm based on the innovation theory for the linear stochastic systems including with parameters. In the robust RLS Wiener predictor, the following information is used. (1) The system matrices for the signal and the degraded signal. (2) The observation matrices for the signal and the degraded signal. (3) The variance of the state for the degraded signal. (4) The cross-variance of the state for the signal with the state. (5) The variance of the observation noise. As a step to obtain the robust RLS Wiener FIR prediction algorithm, this paper presents the robust prediction algorithm of the signal using the covariance information etc. In the predictor, the following information is used. (1) The observation matrices for the signal and the degraded signal. (2) The variance of the state for the degraded signal. (3) The auto-covariance information of the state for the degraded signal. (4) The cross-covariance information of the state for the signal with that for the degraded signal. (5) The variance of the observation noise. The estimation accuracy of the proposed robust RLS Wiener FIR predictor is superior to the existing RLS Wiener FIR predictor.

*Key–Words:* Robust RLS Wiener FIR predictor, discrete-time stochastic systems, auto-covariance function, innovation approach, uncertain parameters

Received: January 24, 2020. Revised: April 2 , 2020. Re-revised: April 8 , 2020. Accepted: April 19, 2020. Published: April 24, 2020.

## 1 Introduction

Finite impulse response (FIR) filter is known in the areas of the digital filter and the filter for signal or state estimations. Concerning the digital filter, Wong, et al. [1], based on stochastic computation, proposes the finite impulse response digital filter with an improved scaling scheme. Nazaripouya, et al. [2] designs the digital FIR filter by using the convex and quasi-convex optimization methods. The digital FIR filter has the properties of minimum-phase, minimum-length, lower group delay with fewer design parameters and faster convergence in comparison with the existing design techniques.

From the aspects of the theory and applications, the robust prediction and filtering techniques have been investigated, e.g. [3]-[5]. In [3], by introducing an iteratively re-weighted least-squares optimization criterion, the robust Kalman filter is designed. The robust filter is applied to a problem in vision. In [4], three different methods are proposed by designing the robust Kalman filter for outliers in the one-step-ahead prediction of the wind speed. In [5], for multi-sensor systems with the uncertainty parameters, a new robust Kalman prediction technique is proposed for compensating parametric uncertainty by fictitious noise. The approach is reduced to the robust Kalman prediction problem for the system with the uncertain noise variances, and the local and centralized robust Kalman predictors are proposed.

The recursive least-squares (RLS) Wiener estimators use the complete information of the state-space model but the information of the input matrix and the input noise variance [6]. For the discrete-time stochastic systems with the uncertain parameters, in the estimation of the signal, the robust RLS Wiener estimators [7] and the robust RLS Wiener finite impulse response filter [6] are proposed. The estimation accuracy of the robust

RLS Wiener estimators [7] are superior to the robust Kalman filter [9] and the RLS Wiener filter [6]. The FIR Kalman filter [10]-[14] is less sensitive to the uncertainties in the state-space model. In Nakamori [8] it is shown that, as the finite interval increases, the mean square-value (MSV) of the estimation errors by the robust RLS Wiener FIR filter [8] becomes gradually small and approaches that by the robust RLS Wiener filter [7].

Nakamori [15] proposes the RLS Wiener FIR prediction and filtering algorithms based on the innovation approach in linear discrete-time stochastic systems. Apart from the filter and smoother, the predictor is useful in the prediction of the air pollution levels etc. [16]. Since the robust RLS Wiener FIR prediction problem is not referred in Nakamori [8], this paper newly proposes the robust RLS Wiener FIR prediction algorithm in Theorem 3, based on the innovation theory, for the linear discrete-time stochastic systems with the uncertain parameters. It is assumed that the signal process is fitted to the auto-regressive (AR) model of the finite order. Also, the degraded signal, caused by the uncertain parameters in the observation and system matrices, is fitted to the AR model of the finite order. Theorem 1 proposes the equation, which the optimal impulses response function satisfies, in the robust RLS FIR prediction problem. Theorem 3 proposes the robust RLS Wiener FIR prediction algorithm which uses the following information. (1) The system matrices for the signal $z(k)$ and the degraded signal $\breve{z}(k)$. (2) The observation matrices for the signal and the degraded signal. (3) The variance $\breve{K}(k,k)$ of the state $\breve{x}(k)$ for the degraded signal. (4) The cross-variance $K_{x\breve{x}}(k,k)$ of the state $x(k)$ for the signal with the state $\breve{x}(k)$. (5) The variance of the observation noise. As a step to the predictor in Theorem 3, Theorem 2 presents the robust RLS FIR prediction algorithm of the signal. The predictor In Theorem 2 uses the following information. (1) The observation matrices for the signal and the degraded signal. (2) The variance of the state for the degraded signal. (3) The auto-covariance information of the state for the degraded signal. (4) The cross-covariance information of the state for the signal with that for the degraded signal. (5) The variance of the observation noise.

The prediction characteristics of the robust RLS Wiener FIR predictor are shown in comparison with those by the RLS Wiener FIR predictor [15] and the robust RLS Wiener FIR filter [8]. The estimation accuracy of the proposed robust RLS Wiener FIR predictor is superior by far to that of the RLS Wiener FIR predictor [15].

In this paper, the typos in the robust RLS Wiener FIR filter [8] are also corrected.

# 2 Robust least-squares FIR prediction problem

Let an m-dimensional observation equation and an n-dimensional state equation be given by

$$\begin{aligned} &\breve{y}(k) = \breve{z}(k) + v(k), \\ &\breve{z}(k) = \bar{H}(k)\bar{x}(k), \\ &\bar{H}(k) = H + \Delta H(k), \\ &\bar{x}(k+1) = \bar{\Phi}(k)\bar{x}(k) + \Gamma w(k), \\ &\bar{\Phi}(k) = \Phi + \Delta\Phi(k), \\ &E[v(k)v^T(s)] = R\delta_K(k-s), \\ &E[w(k)w^T(s)] = Q\delta_K(k-s), \end{aligned} \tag{1}$$

in linear discrete-time stochastic systems with the uncertain quantities $\Delta H(k)$ and $\Delta\Phi(k)$. $v(k)$ and $w(k)$ are the white observation and input noises with the variances $R$ and $Q$ respectively. Their auto-covariance functions are given in (1) by use of the Kronecker delta function $\delta_K(k-s)$. The state equation for $\bar{x}(k+1)$ includes the uncertain quantity $\Delta\Phi(k)$ additionally to the system matrix $\Phi(k)$. Also, the observation matrix $\bar{H}(k)$ contains the uncertain quantity $\Delta H(k)$. Hence, $\breve{z}(k)$ shows the deviated trajectory from the nominal trajectory of the signal $z(k)$ generated by the precise state-space model (2). In (1), as the sum of the degraded signal $\breve{z}(k)$ and the observation noise $v(k)$, the degraded observed value $\breve{y}(k)$ is obtained. Compared with (1), the precise state-space model is given by

$$\begin{aligned} &y(k) = z(k) + v(k), \\ &z(k) = Hx(k), \\ &x(k+1) = \Phi x(k) + \Gamma w(k). \end{aligned} \tag{2}$$

In (2), $z(k)$ is the signal to be estimated. $H$ is an $m$ by $n$ observation matrix, $x(k)$ is the state. The observation noise $v(k)$ and the input noise $w(k)$ have the same auto-covariance functions as those in (1). It is assumed that the sequences of the signal and the observation noise are statistically independent and have zero means. This paper, based on the innovation approach, newly designs the robust RLS FIR predictor using the covariance information in Theorem 2 and the robust RLS Wiener FIR predictor in Theorem 3 for estimating the signal $z(k)$ with the degraded observed value $\breve{y}(k)$. Here, both the robust predictors do not use any information on the uncertain quantities $\Delta\Phi(k)$

and $\Delta H(k)$.

Suppose that the sequence of the degraded signal $\breve{z}(k)$ is fitted to the AR model of the finite order $N$ as

$$\begin{aligned}&\breve{z}(k) = -a_1\breve{z}(k-1) - a_2\breve{z}(k-2)\cdots\\&-a_N\breve{z}(k-N) + \breve{e}(k),\\&E[\breve{e}(k)\breve{e}^T(s)] = \breve{Q}\delta_K(k-s).\end{aligned} \tag{3}$$

Let the degraded signal $\breve{z}(k)$ be represented with the state $\breve{x}(k)$ by

$$\begin{aligned}&\breve{z}(k) = \breve{H}\breve{x}(k),\\&\breve{x}(k) = \begin{bmatrix}\breve{x}_1(k)\\ \breve{x}_2(k)\\ \vdots\\ \breve{x}_{N-1}(k)\\ \breve{x}_N(k)\end{bmatrix}\\&= \begin{bmatrix}\breve{z}(k)\\ \breve{z}(k+1)\\ \vdots\\ \breve{z}(k+N-2)\\ \breve{z}(k+N-1)\end{bmatrix},\\&\breve{H} = [I_{m\times m} \quad 0 \quad 0 \quad \cdots \quad 0 \quad 0].\end{aligned} \tag{4}$$

Henceforth, the state equation for the state $\breve{x}(k)$ is expressed by

$$\begin{aligned}&\begin{bmatrix}\breve{x}_1(k+1)\\ \breve{x}_2(k+1)\\ \vdots\\ \breve{x}_{N-1}(k+1)\\ \breve{x}_N(k+1)\end{bmatrix}\\&= \begin{bmatrix}0 & I_{m\times m} & 0 & \cdots & 0\\ 0 & 0 & I_{m\times m} & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & I_{m\times m}\\ -\breve{a}_N & \breve{a}_{N-1} & \breve{a}_{N-2} & \cdots & -\breve{a}_1\end{bmatrix}\begin{bmatrix}\breve{x}_1(k)\\ \breve{x}_2(k)\\ \vdots\\ \breve{x}_{N-1}(k)\\ \breve{x}_N(k)\end{bmatrix}\\&+ \begin{bmatrix}0\\ 0\\ \vdots\\ 0\\ I_{m\times m}\end{bmatrix}\zeta(k),\\&\zeta(k) = \breve{e}(k+N),\\&E[\zeta(k)\zeta^T(s)] = \breve{Q}\delta_K(k-s).\end{aligned} \tag{5}$$

Let $\breve{K}(k,s) = \breve{K}(k-s)$ represent the auto-covariance function of the state $\breve{x}(k)$ in wide-sense stationary stochastic systems [17]. $\breve{K}(k,s)$ is expressed in the semi-degenerate kernel form of

$$\begin{aligned}&\breve{K}(k,s) = \begin{cases}A(k)B^T(s), 0 \le s \le k,\\ B(k)A^T(s), 0 \le k \le s,\end{cases}\\&A(k) = \breve{\Phi}^k, B^T(s) = \breve{\Phi}^{-s}\breve{K}(s,s).\end{aligned} \tag{6}$$

Here, $\breve{\Phi}$ is the system matrix for the state $\breve{x}(k)$. From (5), the system matrix $\breve{\Phi}$ is expressed by

$$\breve{\Phi} = \begin{bmatrix}0 & I_{m\times m} & 0 & \cdots & 0\\ 0 & 0 & I_{m\times m} & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & I_{m\times m}\\ -\breve{a}_N & -\breve{a}_{N-1} & -\breve{a}_{N-2} & \cdots & -\breve{a}_1\end{bmatrix}. \tag{7}$$

By putting $K_{\breve{z}}(k,s) = K_{\breve{z}}(k-s) = E[\breve{z}(k)\breve{z}^T(s)]$, the auto-variance function $\breve{K}(k,k)$ of the state $\breve{x}(k)$ is expressed by

$$\begin{aligned}&\breve{K}(k,k) = E\left[\begin{bmatrix}\breve{z}(k)\\ \breve{z}(k+1)\\ \vdots\\ \breve{z}(k+N-2)\\ \breve{z}(k+N-1)\end{bmatrix}\right.\\&\times [\breve{z}^T(k) \quad \breve{z}^T(k+1) \quad \cdots\\&\left.\breve{z}^T(k+N-2) \quad \breve{z}^T(k+N-1)]\right]\\&= \begin{bmatrix}K_{\breve{z}}(0) & K_{\breve{z}}(-1) & \cdots & K_{\breve{z}}(-N+2) & K_{\breve{z}}(-N+1)\\ K_{\breve{z}}(1) & K_{\breve{z}}(0) & \cdots & K_{\breve{z}}(-N+3) & K_{\breve{z}}(-N+2)\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ K_{\breve{z}}(N-2) & K_{\breve{z}}(N-3) & \cdots & K_{\breve{z}}(0) & K_{\breve{z}}(-1)\\ K_{\breve{z}}(N-1) & K_{\breve{z}}(N-2) & \cdots & K_{\breve{z}}(1) & K_{\breve{z}}(0)\end{bmatrix}.\end{aligned} \tag{8}$$

With $K_{\breve{z}}(k-s)$, the Yule-Walker equation for the AR parameters is formulated as

$$\hat{K}(k,k)\begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_{N-1}^T \\ a_N^T \end{bmatrix} = -\begin{bmatrix} K_{\breve{z}}^T(1) \\ K_{\breve{z}}^T(2) \\ \vdots \\ K_{\breve{z}}^T(N-1) \\ K_{\breve{z}}^T(N) \end{bmatrix},$$
$$\hat{K}(k,k) = \begin{bmatrix} K_{\breve{z}}(0) & K_{\breve{z}}(1) & \cdots & K_{\breve{z}}(N-2) & K_{\breve{z}}(N-1) \\ K_{\breve{z}}^T(1) & K_{\breve{z}}(0) & \cdots & K_{\breve{z}}(N-3) & K_{\breve{z}}(N-2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ K_{\breve{z}}^T(N-2) & K_{\breve{z}}^T(N-3) & \cdots & K_{\breve{z}}(0) & K_{\breve{z}}(1) \\ K_{\breve{z}}^T(N-1) & K_{\breve{z}}^T(N-2) & \cdots & K_{\breve{z}}^T(1) & K_{\breve{z}}(0) \end{bmatrix}. \tag{9}$$

Let $K_{x\breve{x}}(k,s) = K_{x\breve{x}}(k-s) = E[x(k)\breve{x}^T(s)]$ be the cross-covariance function of the state $x(k)$ with the state $\breve{x}(s)$ in wide-sense stationary stochastic systems. $K_{x\breve{x}}(k,s)$ is expressed in the functional form of

$$K_{x\breve{x}}(k,s) = \alpha(k)\beta^T(s), 0 \le s \le k,$$
$$\alpha(k) = \Phi^k, \beta^T(s) = \Phi^{-s}K_{x\breve{x}}(s,s). \tag{10}$$

Here, from (2), $\Phi$ is the system matrix for the state $x(k)$.

Under the above prerequisites on the signal and the degraded signal, etc., Theorem 1, based on the innovation theory, presents the equation, which the optimal impulse response function satisfies, in the robust RLS FIR prediction problem.

**Theorem 1** Let the $l$-step ahead FIR prediction estimate $\hat{x}(k+l|k-L+1)$ of the state $x(k+l)$ be expressed by

$$\hat{x}(k+l|k-L+1) = \sum_{i=k-L+1}^{k} g(k,i)\nu(i),$$
$$\nu(i) = \breve{y}(i) - \breve{H}\breve{\Phi}\hat{\breve{x}}(i-1|i-1-L+1), \tag{11}$$

in terms of the innovation process $\{\nu(i), k-L+1 \le i \le k\}$. In (11), $g(k,i)$ represents the time-varying impulse response function and $\hat{\breve{x}}(i-1|i-1-L+1)$ is the FIR filtering estimate of the state $\breve{x}(i-1)$. Let the FIR filtering estimate $\hat{\breve{x}}(k|k-L+1)$ of $\breve{x}(k)$ be given by

$$\hat{\breve{x}}(k|k-L+1) = \sum_{i=k-L+1}^{k} g_0(k,i)\nu(i) \tag{12}$$

as a linear combination of the impulse response function $g_0(k,i)$ and the innovation sequence $\{\nu(i), k-L+1 \le i \le k\}$. Then the optimal impulse response function $g(k,s)$ satisfies

$$g(k,s)\Lambda(s) = K_{x\breve{x}}(k+l,s)\breve{H}^T - \sum_{i=s-1-L+1}^{s-1} g(k,i)\Lambda(i)g_0^T(s-1,i) \times \breve{\Phi}^T\breve{H}^T,$$
$$K_{x\breve{x}}(k+l,s)\breve{H}^T = K_{x\breve{z}}(k+l,s). \tag{13}$$

In (13), $K_{x\breve{x}}(k+l,s)\breve{H}^T$ is equivalent to the cross-covariance function of the state $x(k+l)$ with the degraded signal $\breve{z}(s)$, $K_{x\breve{z}}(k+l,s)$.

**Proof**

Consider the estimation problem, which minimizes the MSV

$$J = E[||x(k+l) - \hat{x}(k+l|k-L+1)||^2] \tag{14}$$

of the FIR prediction errors. From an orthogonal projection lemma [17]

$$x(k+l) - \sum_{i=k-L+1}^{k} g(k,i)\nu(i) \perp \nu(s),$$
$$k-L+1 \le s \le k, \tag{15}$$

the impulse response function $g(k,i)$ satisfies the Wiener-Hopf equation

$$E[x(k+l)\nu^T(s)] = \sum_{i=k-L+1}^{k} g(k,i)E[\nu(i)\nu^T(s)],$$
$$k-L+1 \le s \le k. \tag{16}$$

In (15), '$\perp$' denotes the notation of the orthogonality. Let the covariance function of the innovation process be given by

$$E[\nu(i)\nu^T(s)] = \Lambda(i)\delta_K(i-s). \tag{17}$$

From (16), (17) and the expression for the innovation process $\nu(s)$, $g(k,s)$ satisfies

$$\begin{aligned} g(k,s)\Lambda(s) &= E[x(k+l)v^T(s)] \\ &= E[x(k+l)(\breve{y}(s) \\ &-\breve{H}\breve{\Phi}\hat{\breve{x}}(s-1|s-1-L+1))^T] \\ &= E[x(k+l)\breve{y}^T(s)] \\ &-E[x(k+l) \\ &\times \hat{\breve{x}}^T(s-1|s-1-L+1)]\breve{\Phi}^T\breve{H}^T. \end{aligned} \tag{18}$$

The term $E[x(k+l)\breve{y}^T(s)]$ is developed as

$$\begin{aligned} &E[x(k+l)\breve{y}^T(s)] \\ &= E[x(k+l)(\breve{z}(s)+v(s))^T] \\ &= E[x(k+l)\breve{z}^T(s)] \\ &= E[x(k+l)\breve{x}^T(s)]\breve{H}^T \\ &= K_{x\breve{x}}(k+l,s)\breve{H}^T. \end{aligned} \tag{19}$$

From (18) and (19), the optimal impulse response function $g(k,s)$ satisfies (13).

(Q.E.D.)

Starting with (13), the robust RLS FIR prediction algorithm using the covariance information etc. is presented in Theorem 2. Then Theorem 3 proposes the robust RLS Wiener FIR prediction algorithm.

# 3 Robust RLS FIR predictor using covariance information and robust RLS Wiener FIR predictor

Theorem 2 proposes the robust RLS FIR prediction algorithm using the covariance information $\breve{K}(k,s)$ of the state $\breve{x}(k)$ for the degraded signal $\breve{z}(k)$ and the cross-covariance information $K_{x\breve{x}}(k,s)$ of the state $x(k)$ for the signal $z(k)$ with the state $\breve{x}(s)$ for the degraded signal $\breve{z}(s)$, etc.

**Theorem 2** Let the state equation and the observation equation, which contains the uncertain quantities $\Delta\Phi$ and $\Delta H$ respectively, be given by (1). Let $H$ represent the observation matrix for the signal $z(k)$. Let $\breve{\Phi}$ and $\breve{H}$ represent the system and observation matrices respectively for the degraded signal $\breve{z}(k)$, fitted to the AR model (3). Let the variance $\breve{K}(k,k)$ of the state $\breve{x}(k)$ for the degraded signal $\breve{z}(k)$ be given. Let the auto-covariance function $\breve{K}(k,s)$ of $\breve{x}(k)$ be expressed by (6) in terms of $A(k)$ and $B(s)$. Let the cross-covariance function $K_{x\breve{x}}(k,s)$ of $x(k)$ with $\breve{x}(s)$ be given by (10) in terms of $\alpha(k)$ and $\beta(s)$. Let the variance of the white observation noise $v(k)$ be $R$. Then the robust RLS estimation algorithm for the $l$-step ahead FIR prediction estimate $\hat{z}(k+l|k-L+1)$ of the signal $z(k+l)$ consists of (20)-(40) in linear discrete-time stochastic systems.

$l$-step ahead FIR prediction estimate of the signal $z(k+l)$: $\hat{z}(k+l|k-L+1)$

$$\begin{aligned} &\hat{z}(k+l|k-L+1) \\ &= H\hat{x}(k+l|k-L+1) \end{aligned} \tag{20}$$

$l$-step ahead FIR prediction estimate of the state $x(k+l)$: $\hat{x}(k+l|k-L+1)$

$$\begin{aligned} &\hat{x}(k+l|k-L+1) \\ &\qquad = \alpha(k+l)e(k) \end{aligned} \tag{21}$$

FIR filtering estimate of the signal $z(k)$: $\hat{z}(k|k-L+1)$

$$\hat{z}(k|k-L+1) = H\hat{x}(k|k-L+1) \tag{22}$$

FIR filtering estimate of the state $x(k)$: $\hat{x}(k|k-L+1)$

$$\hat{x}(k|k-L+1) = \alpha(k)e(k) \tag{23}$$

Initial condition of $\hat{x}(k|k-L+1)$ at $k=L$: $\hat{x}(L|1)$

FIR filtering estimate of the state $\breve{x}(k)$: $\hat{\breve{x}}(k|k-L+1)$

$$\hat{\breve{x}}(k|k-L+1) = A(k)e_0(k) \tag{24}$$

Initial condition of $\hat{\breve{x}}(k|k-L+1)$ at $k=L$: $\hat{\breve{x}}(L|1)$

Recursive equation for $e(k)$:

$$\begin{aligned} e(k) &= e(k-1) + J(k)(\breve{y}(k) \\ &-\breve{H}A(k)e_0(k-1)) \\ &-J(k-L)(\breve{y}(k-L) \\ &-\breve{H}A(k-L)e_0(k-L-1)) \end{aligned} \tag{25}$$

Initial condition of $e(k)$ at $k=L$: $\overline{e}(L)$

Recursive equation for $e_0(k)$:

$$\begin{aligned} e_0(k) &= e_0(k-1) + J_0(k)(\breve{y}(k) \\ &-\breve{H}A(k)e_0(k-1)) \\ &-J_0(k-L)(\breve{y}(k-L) \\ &-\breve{H}A(k-L)e_0(k-L-1)) \end{aligned} \tag{26}$$

Initial condition of $e_0(k)$ at $k=L$: $\overline{e}_0(L)$

Equation for $J(k)$:

$$\begin{aligned} J(k) &= [\beta^T(k)\breve{H}^T \\ &-r(k-1)A^T(k-1)\breve{\Phi}^T\breve{H}^T]\Lambda^{-1}(k) \end{aligned} \tag{27}$$

$$\begin{aligned} r(k) &= r(k-1) + J(k)\Lambda(k)J_0^T(k) \\ &-J(k-L)\Lambda(k-L)J_0^T(k-L) \end{aligned} \tag{28}$$

Equation for $J_0(k)$:

$$J_0(k) = [B^T(k)\breve{H}^T - r_0(k)A^T(k-1)\breve{\Phi}^T\breve{H}^T]\Lambda^{-1}(k) \tag{29}$$

Recursive equation for $r_0(k)$:

$$r_0(k) = r_0(k-1) + J_0(k)\Lambda(k)J_0^T(k) - J_0(k-L)\Lambda(k-L)J_0^T(k-L) \tag{30}$$

Initial condition of $r_0(k)$ at $k = L$: $\overline{r}_0(L)$

Equation for $\Lambda(k)$:

$$\Lambda(k) = \breve{H}\breve{K}(k,k)\breve{H}^T + R - \breve{H}A(k)r_0(k-1)A^T(k)\breve{H}^T \tag{31}$$

Initial condition of FIR filtering estimate $\hat{x}(k|k-L+1)$ of $x(k)$ at $k = L$: $\hat{x}(L|1)$

$$\hat{x}(L|1) = \alpha(L)\overline{e}(L) \tag{32}$$

Recursive equation for $\overline{e}(L)$:

$$\overline{e}(L) = \overline{e}(L-1) + \overline{J}(L)(\breve{y}(L) - \breve{H}A(L)\overline{e}_0(L-1)), \quad \overline{e}(0) = 0 \tag{33}$$

Equation for $\overline{J}(L)$:

$$\overline{J}(L) = (\beta^T(L)\breve{H}^T - \overline{r}(L-1)A^T(L-1)A^T(1)\breve{H}^T) \times \overline{\Lambda}^{-1}(L) \tag{34}$$

Recursive equation for $\overline{r}(L)$:

$$\overline{r}(L) = \overline{r}(L-1) + \overline{J}(L)\overline{\Lambda}(L)\overline{J}_0^T(L), \quad \overline{r}(0) = 0 \tag{35}$$

Initial condition of FIR filtering estimate $\hat{\breve{x}}(k|k-L+1)$ of $\breve{x}(k)$: at $k = L$: $\hat{\breve{x}}(L|1)$

$$\hat{\breve{x}}(L|1) = A(L)\overline{e}_0(k) \tag{36}$$

Recursive equation for $\overline{e}_0(L)$:

$$\overline{e}_0(L) = \overline{e}_0(L-1) + \overline{J}_0(L)(\breve{y}(L) - \breve{H}A(L)\overline{e}_0(L-1)), \quad \overline{e}_0(0) = 0 \tag{37}$$

Equation for $\overline{J}_0(L)$:

$$\overline{J}_0(L) = (B^T(L)\breve{H}^T - \overline{r}_0(L-1)A^T(L)\breve{H}^T)\overline{\Lambda}^{-1}(L) \tag{38}$$

Recursive equation for $\overline{r}_0(L)$:

$$\overline{r}_0(L) = \overline{r}_0(L-1) + \overline{J}_0(L)\overline{\Lambda}_0(L)\overline{J}_0^T(L), \quad \overline{r}_0(0) = 0 \tag{39}$$

Equation for $\overline{\Lambda}(L)$:

$$\overline{\Lambda}(L) = \breve{H}\breve{K}(L,L)\breve{H}^T + R - \breve{H}A(L)\overline{r}_0(L-1)A^T(L)\breve{H}^T \tag{40}$$

Proof of Theorem 2 is deferred to Appendix A.

Based on the robust RLS FIR prediction algorithm in Theorem 2, Theorem 3 presents the robust RLS Wiener FIR prediction algorithm.

**Theorem 3** Let the state and the observation equations, including the uncertain quantities $\Delta\Phi$ and $\Delta H$ be given by (1). Let $\Phi$ and $H$ represent the system and observation matrices respectively for the signal $z(k)$. Let $\breve{\Phi}$ and $\breve{H}$ represent the system and observation matrices respectively for the degraded signal $\breve{z}(k)$, which is fitted to the AR model (3). Let the variance $\breve{K}(k,k)$ of the state $\breve{x}(k)$ for the degraded signal $\breve{z}(k)$ and the cross-variance $K_{x\breve{x}}(k,k)$ of the state $x(k)$ for the signal $z(k)$ with the state $\breve{x}(k)$ be given. Let the variance of the white observation noise $v(k)$ be $R$. Then the robust RLS Wiener estimation algorithm for the $l$-step ahead FIR prediction estimate $\hat{z}(k+l|k-L+1)$ of the signal $z(k+l)$ consists of (41)-(57) in linear discrete-time stochastic systems.

$l$-step ahead FIR prediction estimate of the signal $z(k+l)$: $\hat{z}(k+l|k-L+1)$

$$\hat{z}(k+l|k-L+1) = H\hat{x}(k+l|k-L+1) \tag{41}$$

$l$-step ahead FIR prediction estimate of the state $x(k+l)$: $\hat{x}(k+l|k-L+1)$

$$\hat{x}(k+l|k-L+1) = \Phi^l\hat{x}(k|k-L+1) \tag{42}$$

FIR filtering estimate of the signal $z(k)$: $\hat{z}(k|k-L+1)$

$$\hat{z}(k|k-L+1) = H\hat{x}(k|k-L+1) \tag{43}$$

FIR filtering estimate of the state $x(k)$: $\hat{x}(k|k-L+1)$

$$\begin{aligned}\hat{x}(k|k-L+1) &= \Phi\hat{x}(k-1|k-1-L+1) + G(k)(\breve{y}(k) - \breve{H}\breve{\Phi}\hat{\breve{x}}(k-1|k-1-L+1)) \\ &\quad - \Phi^L G(k-L)(\breve{y}(k-L) - \breve{H}\breve{\Phi} \times \hat{\breve{x}}(k-L-1|k-L-1-L+1))\end{aligned} \tag{44}$$

Initial condition of $\hat{x}(k|k-L+1)$ at $k = L$: $\hat{x}(L|1)$

FIR filtering estimate of the state $\breve{x}(k)$: $\hat{\breve{x}}(k|k-L+1)$

$$\hat{\hat{x}}(k|k-L+1) \\ = \breve{\Phi}\hat{\hat{x}}(k-1|k-1-L+1) \\ +G_0(k)(\breve{y}(k) \\ -\breve{H}\breve{\Phi}\hat{\hat{x}}(k-1|k-1-L+1)) \\ -\Phi^L G_0(k-L)(\breve{y}(k-L)-\breve{H}\breve{\Phi} \\ \times \hat{\hat{x}}(k-L-1|k-L-1-L+1)) \quad (45)$$

Initial condition of $\hat{\hat{x}}(k|k-L+1)$ at $k=L$: $\hat{\hat{x}}(L|1)$

FIR filter gain for $\hat{x}(k|k-L+1)$: $G(k)$

$$G(k) = [K_{x\breve{x}}(k,k)\breve{H}^T \\ -\Phi S(k-1)\breve{\Phi}^T\breve{H}^T]\Lambda^{-1}(k) \quad (46)$$

FIR filter gain for $\hat{\hat{x}}(k|k-L+1)$: $G_0(k)$

$$G_0(k) = [\breve{K}(k,k)\breve{H}^T \\ -\breve{\Phi}S_0(k-1)\breve{\Phi}^T\breve{H}^T]\Lambda^{-1}(k) \quad (47)$$

Equation for $\Lambda(k)$:

$$\Lambda(k) = R + \breve{H}\breve{K}(k,k)\breve{H}^T \\ -\breve{H}\breve{\Phi}S_0(k-1)\breve{\Phi}^T\breve{H}^T \quad (48)$$

Recursive equation for $S(k)$:

$$S(k) = \Phi S(k-1)\breve{\Phi}^T \\ +G(k)\Lambda(k)G_0^T(k) \\ -\Phi^L G(k-L)\Lambda(k-L)G_0^T(k-L) \\ \times (\Phi^T)^L \quad (49)$$

Initial condition of $S(k)$ at $k=L$: $\overline{S}(L)$

Recursive equation for $S_0(k)$:

$$S_0(k) = \breve{\Phi}S_0(k-1)\breve{\Phi}^T \\ +G_0(k)\Lambda(k)G_0^T(k) \\ -\breve{\Phi}^L G_0(k-L)\Lambda(k-L)G_0^T(k-L) \\ \times (\breve{\Phi}^T)^L \quad (50)$$

Initial condition of $S_0(k)$ at $k=L$: $\overline{S}_0(L)$

Recursive equation for $\hat{x}(L|1)$:

$$\hat{x}(L|1) = \Phi\hat{x}(L-1|1) \\ +\overline{G}(L)(\breve{y}(L)-\breve{H}\breve{\Phi}\hat{\hat{x}}(L-1|1)), \\ \hat{x}(0|1) = 0 \quad (51)$$

Filter gain for $\hat{x}(L|1)$ in (51): $\overline{G}(L)$

$$\overline{G}(L) = [K_{x\breve{z}}(L,L) \\ -\Phi\overline{S}(L-1)\breve{\Phi}^T\breve{H}^T]\overline{\Lambda}^{-1}(L), \\ K_{x\breve{z}}(L,L) = K_{x\breve{x}}(L,L)\breve{H}^T \quad (52)$$

Recursive equation for $\hat{\hat{x}}(L|1)$:

$$\hat{\hat{x}}(L|1) = \breve{\Phi}\hat{\hat{x}}(L-1|1) \\ +\overline{G}_0(L)(\breve{y}(L)-\breve{H}\breve{\Phi}\hat{\hat{x}}(L-1|1)), \\ \hat{\hat{x}}(0|1) = 0 \quad (53)$$

Filter gain for $\hat{x}(L|1)$ in (53): $\overline{G}_0(L)$

$$\overline{G}_0(L) = [\breve{K}(L,L)\breve{H}^T \\ -\breve{\Phi}\overline{S}_0(L-1)\breve{\Phi}^T\breve{H}^T]\overline{\Lambda}^{-1}(L) \quad (54)$$

Equation for $\overline{\Lambda}(L)$:

$$\overline{\Lambda}(L) = R + \breve{H}\breve{K}(L,L)\breve{H}^T \\ -\breve{H}\breve{\Phi}\overline{S}_0(L-1)\breve{\Phi}^T\breve{H}^T \quad (55)$$

Recursive equation for $\overline{S}(L)$:

$$\overline{S}(L) = \Phi\overline{S}(L-1)\breve{\Phi}^T \\ +\overline{G}(L)\overline{\Lambda}(L)\overline{G}_0^T(L), \\ \overline{S}(0) = 0 \quad (56)$$

Recursive equation for $\overline{S}_0(L)$:

$$\overline{S}_0(L) = \breve{\Phi}\overline{S}_0(L-1)\breve{\Phi}^T \\ +\overline{G}_0(L)\overline{\Lambda}(L)\overline{G}_0^T(L), \\ \overline{S}_0(0) = 0 \quad (57)$$

Proof of Theorem 3 is deferred to Appendix B.

Necessary conditions on the stabilit y of the robust RLS Wiener FIR prediction and filtering algorithms are as follows.

(1) All the real parts in the eigenvalues of the matrix $\Phi$ are negative.

(2) All the real parts in the eigenvalues of the matrix $\breve{\Phi} - G_0(k)\breve{H}\breve{\Phi}$ are negative.

(3)$R + \breve{H}[\breve{K}(k,k) - \breve{\Phi}S_0(k-1)\breve{\Phi}^T]\breve{H}^T > 0$

(4) All the real parts in the eigenvalues of the matrix $\breve{\Phi} - \overline{G}_0(k)\breve{H}\breve{\Phi}$ are negative.

(5)$R + \breve{H}[\breve{K}(k,k) - \breve{\Phi}\overline{S}_0(k-1)\breve{\Phi}^T]\breve{H}^T > 0$

Section 4 proposes the re cursive algorithm for the prediction error variance function of the robust RLS Wiener FIR predictor presented in Theorem 3. Also, the exi stence of the robust RLS Wiener FIR prediction estimate $\hat{z}(k+l|k-L+1)$ of the signal $z(k+l)$ is shown.

# 4 Prediction error variance function of signal

Let the variance function of t he FIR prediction error $z(k+l) - \hat{z}(k+l|k-L+1)$ be denoted by $\tilde{P}_z(k+l)$ . Let the auto-covariance functi on $K(k,s)$ of the state $x(k)$ be expressed by

$$K(k,s) \\ = \begin{cases} A_x(k)B_x^T(s), 0 \le s \le k, \\ B_x(k)A_x^T(s), 0 \le k \le s, \end{cases} \\ A_x(k) = \alpha(k) = \Phi^k, \\ B_x^T(s) = \Phi^{-s}K(s,s). \quad (58)$$

From (16) and (A.10), the FIR prediction error variance function $\tilde{P}_z(k)$ is formulated as

$$\begin{aligned}
&\tilde{P}_z(k+l) = H(K(k+l,k+l)\\
&-E[\hat{x}(k+l|k-L+1)\\
&\times \hat{x}^T(k+l|k-L+1)])H^T\\
&= H(K(k,k)\\
&-E[x(k+l)\hat{x}^T(k+l|k-L+1)])\\
&\times H^T\\
&= H(K(k,k)\\
&-\Phi^{k+l}\underline{r}(k)(\Phi^T)^{k+l})H^T,\\
&\underline{r}(k) = \sum_{i=k-L+1}^{k} J(i)\Lambda(i)J^T(i).
\end{aligned} \tag{59}$$

$\underline{r}(k)$ is calculated recursively by

$$\begin{aligned}
&\underline{r}(k) = \underline{r}(k) + J(k)\Lambda(k)J^T(k)\\
&-J(k-L)\Lambda(k-L)J^T(k-L),\\
&\underline{r}(0) = 0.
\end{aligned} \tag{60}$$

Hence, the RLS Wiener FIR prediction error variance function $\tilde{P}_z(k+l)$ is calculated by (27) ~ (31), (59) and (60) recursively. Since $\tilde{P}_z(k+l)$ is positive-semidefinite, the RLS Wiener FIR prediction variance of the signal $z(k+l)$, $HE[\hat{x}(k+l|k-L+1)\hat{x}^T(k+l|k-L+1)]H^T$, is upper bounded by $HK(k,k)H^T$ and lower bounded by the zero matrix as

$$\begin{aligned}
&0 \le\\
&HE[\hat{x}(k+l|k-L+1)\\
&\times \hat{x}^T(k+l|k-L+1)]H^T\\
&\le HK(k,k)H^T.
\end{aligned} \tag{61}$$

This validates the existence of the robust RLS Wiener FIR prediction estimate $\hat{z}(k+l|k-L+1)$ of the signal $z(k+l)$.

# 5 A numerical simulation example

Let a scalar observation equation for the signal $z(k)$ and the state equation for $x(k)$ be described by

$$\begin{aligned}
&y(k) = z(k) + v(k),\\
&z(k) = Hx(k),\\
&H = [1 \quad 0],\\
&x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix},\\
&x(k+1) = \Phi x(k) + \Gamma w(k),\\
&\Phi = \begin{bmatrix} 0 & 1 \\ -a_2 & -a_1 \end{bmatrix},\\
&a_1 = -0.1, a_2 = -0.8,\\
&\Gamma = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\\
&E[v(k)v(s)] = R\delta_K(k-s),\\
&E[w(k)w(s)] = Q\delta_K(k-s),\\
&Q = 0.5^2.
\end{aligned} \tag{62}$$

From (2) it is noted that the signal $z(k)$ is generated by the second-order AR model. Let us consider to calculate the $l$-step ahead prediction estimate of the signal $z(k+l)$ with the degraded observed value $\breve{y}(k)$, which is generated by the state-space model (63) including the uncertain quantities $\Delta H(k)$ and $\Delta\Phi(k)$.

$$\begin{aligned}
&\breve{y}(k) = \breve{z}(k) + v(k),\\
&\breve{z}(k) = \bar{H}(k)\bar{x}(k),\\
&\bar{x}(k) = \begin{bmatrix} \bar{x}_1(k) \\ \bar{x}_2(k) \end{bmatrix},\\
&\bar{H}(k) = H + \Delta H(k)\\
&= [1+\Delta_3(k) \quad 0],\\
&\Delta H(k) = [\Delta_3(k) \quad 0],\\
&\Delta_3(k) = 0.1,\\
&\bar{x}(k+1) = \bar{\Phi}(k)\bar{x}(k) + \Gamma w(k),\\
&\bar{\Phi}(k) = \Phi + \Delta\Phi(k),\\
&\Delta\Phi(k) = \begin{bmatrix} 0 & 0 \\ \Delta_2(k) & \Delta_1(k) \end{bmatrix},\\
&\Delta_1(k) = 0.01, \Delta_2(k) = -0.1
\end{aligned} \tag{63}$$

Without any usages of a priori information of $\Delta H(k)$ and $\Delta\Phi(k)$, the robust RLS Wiener FIR predictor calculates the prediction estimate of the signal recursively. The degraded signal $\breve{z}(k)$ is fitted to the AR model of the $N$-th order.

$$\begin{aligned}
&\breve{z}(k) = -\breve{a}_1\breve{z}(k-1) - \breve{a}_2\breve{z}(k-2)\\
&- \cdots - \breve{a}_N\breve{z}(k-N) + \breve{e}(k),\\
&N = 10\\
&E[\breve{e}(k)\breve{e}(s)] = \breve{Q}\delta_K(k-s).
\end{aligned} \tag{64}$$

From (4) and (64), $\breve{z}(k)$ is expressed with the 1 by $N$ observation vector $\breve{H}$ as

$$\begin{aligned}
&\breve{z}(k) = \breve{H}\breve{x}(k),\\
&\breve{H} = [1 \quad 0 \quad 0 \quad \cdots \quad 0 \quad 0].
\end{aligned} \tag{65}$$

In the simulation example, the state equation for $\breve{x}(k)$ in (5) corresponds to the case of $m = 1$.

$\breve{K}(k,s)=\breve{K}(k-s)$ represents the auto-covariance function of the state $\breve{x}(k)$ in wide-sense stationary stochastic systems. $\breve{K}(k,s)$ is expressed in the semi-degenerate kernel form (6). $\breve{\Phi}$ in (6) represents the sy stem matrix for the state $\breve{x}(k)$. Also, from the auto-covariance functi on $K_{\breve{z}}(k-s)=K_{\breve{z}}(s-k)=E[\breve{z}(k)\breve{z}(s)]$ of the degraded signal $\breve{z}(k)$, the auto-variance function $\breve{K}(k,k)$ of the state $\breve{x}(k)$ is expressed as

$$\begin{aligned}\breve{K}(k,k)&=E\left[\begin{bmatrix}\breve{z}(k)\\ \breve{z}(k+1)\\ \vdots\\ \breve{z}(k+N-2)\\ \breve{z}(k+N-1)\end{bmatrix}\times[\breve{z}(k)\quad \breve{z}(k+1)\quad\cdots\quad \breve{z}(k+N-2)\quad \breve{z}(k+N-1)]\right]\\ &=\begin{bmatrix}K_{\breve{z}}(0) & K_{\breve{z}}(1) & \cdots & K_{\breve{z}}(N-2) & K_{\breve{z}}(N-1)\\ K_{\breve{z}}(1) & K_{\breve{z}}(0) & \cdots & K_{\breve{z}}(N-3) & K_{\breve{z}}(N-2)\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ K_{\breve{z}}(N-2) & K_{\breve{z}}(N-3) & \cdots & K_{\breve{z}}(0) & K_{\breve{z}}(1)\\ K_{\breve{z}}(N-1) & K_{\breve{z}}(N-2) & \cdots & K_{\breve{z}}(1) & K_{\breve{z}}(0)\end{bmatrix}.\end{aligned}\tag{66}$$

Let $K_{z\breve{z}}(k,s)=E[z(k)\breve{z}(s)]$ represent the cross-covariance function of the signal $z(k)$ with the degraded signal $\breve{z}(s)$. From (4) and (65), the cross-covariance function $K_{x\breve{x}}(k,s)$ is expressed as

$$\begin{aligned}&K_{x\breve{x}}(k,s)=\Phi^{k-s}K_{x\breve{x}}(s,s),\\ &0\le s\le k,\\ &K_{x\breve{x}}(k,k)\\ &=\begin{bmatrix}K_{z\breve{z}}(k,k) & K_{z\breve{z}}(k,k+1) & \cdots & K_{z\breve{z}}(k,k+N-2) & K_{z\breve{z}}(k,k+N-1)\\ K_{z\breve{z}}(k+1,k) & K_{z\breve{z}}(k+1,k) & \cdots & K_{z\breve{z}}(k+1,k+N-2) & K_{z\breve{z}}(k+1,k+N-1)\end{bmatrix}.\end{aligned}\tag{67}$$

The AR param eters $\breve{a}_1,\breve{a}_2,\cdots,\breve{a}_{N-1},\breve{a}_N$ in (64) are calculated by the Yule-Walker equation

$$\breve{K}(k,k)\begin{bmatrix}\breve{a}_1\\ \breve{a}_2\\ \vdots\\ \breve{a}_{N-1}\\ \breve{a}_N\end{bmatrix}=-\begin{bmatrix}K_{\breve{z}}(1)\\ K_{\breve{z}}(2)\\ \vdots\\ K_{\breve{z}}(N-1)\\ K_{\breve{z}}(N)\end{bmatrix}.\tag{68}$$

By substituting $H$, $\breve{H}$, $\Phi$, $\breve{\Phi}$, $K_{x\breve{x}}(k,k)$, $\breve{K}(k,k)=\breve{K}(L,L)$ and $R$ into the robust RLS Wiener FIR prediction algorithm of Theorem 3, the prediction estimates are calculated recursively. In evaluating $\breve{\Phi}$ in (7), $\breve{K}(k,k)$ in (66) and $K_{x\breve{x}}(k,k)$ in (67), the 2,000 number of signal an d degraded signal data are used.

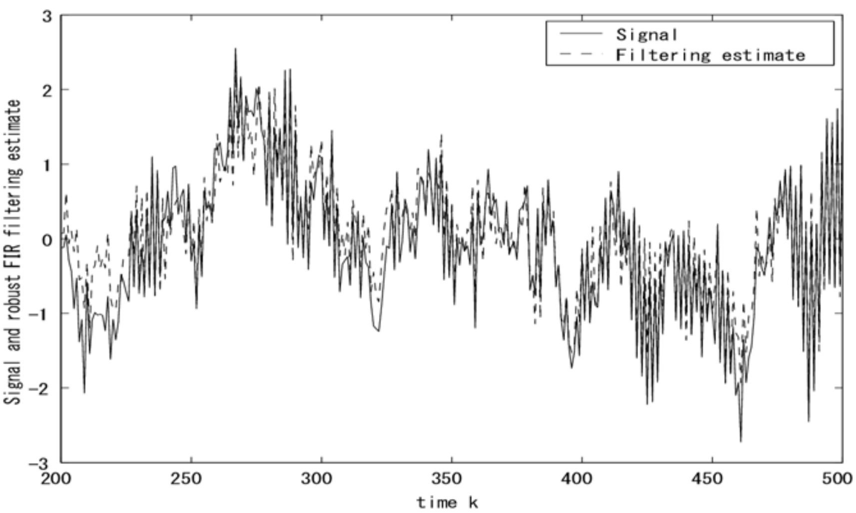


Fig.1 Signal $z(k+l)$ and robust RLS Wiener FIR prediction estimate $\hat{z}(k+l\mid k-L+1)$, $L=200$, $l=3$ vs. $k$ for white Gaussian observation noise $N(0,0.3^2)$.

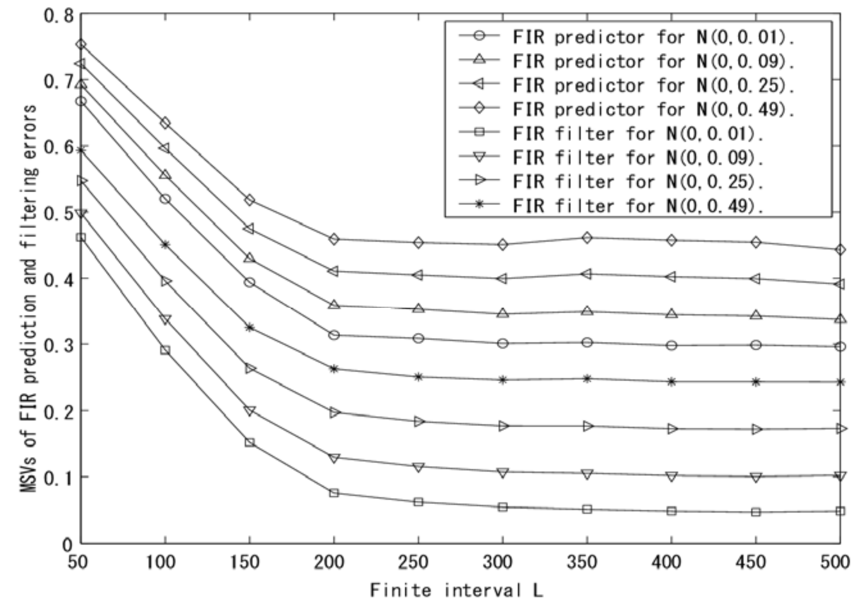


Fig.2 MSVs of one-step ahead prediction errors $z(k+1)-\hat{z}(k+1|k-L+1)$ by robust RLS Wiener FIR predictor in Theorem 3 vs. finite interval $L$, $50 \le L \le 500$, and MSVs of filtering errors $z(k)-\hat{z}(k|1)$, $1 \le k \le L$, by robust RLS Wiener FIR filter [8] vs. $L$, $50 \le L \le 500$, for white Gaussian observation noises $N(0,0.1^2)$, $N(0,0.3^2)$, $N(0,0.5^2)$ and $N(0,0.7^2)$.

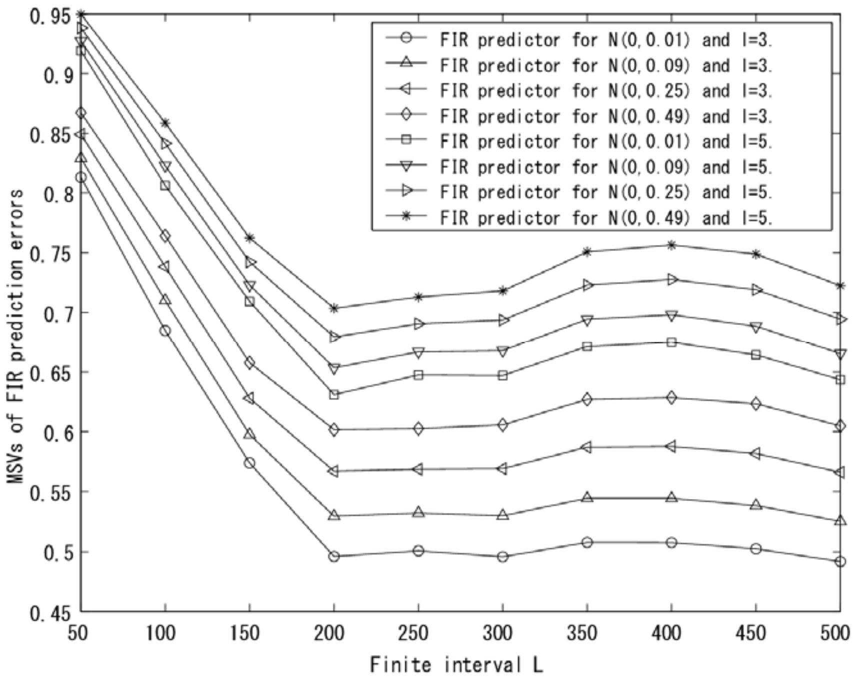


Fig.3 MSVs of robust RLS Wiener prediction errors $z(k+l)-\hat{z}(k+l|k-L+1)$ vs. finite interval $L$, in the cases of $l=3$ and $l=5$, for white Gaussian observation noises $N(0,0.1^2)$, $N(0,0.3^2)$, $N(0,0.5^2)$ and $N(0,0.7^2)$.

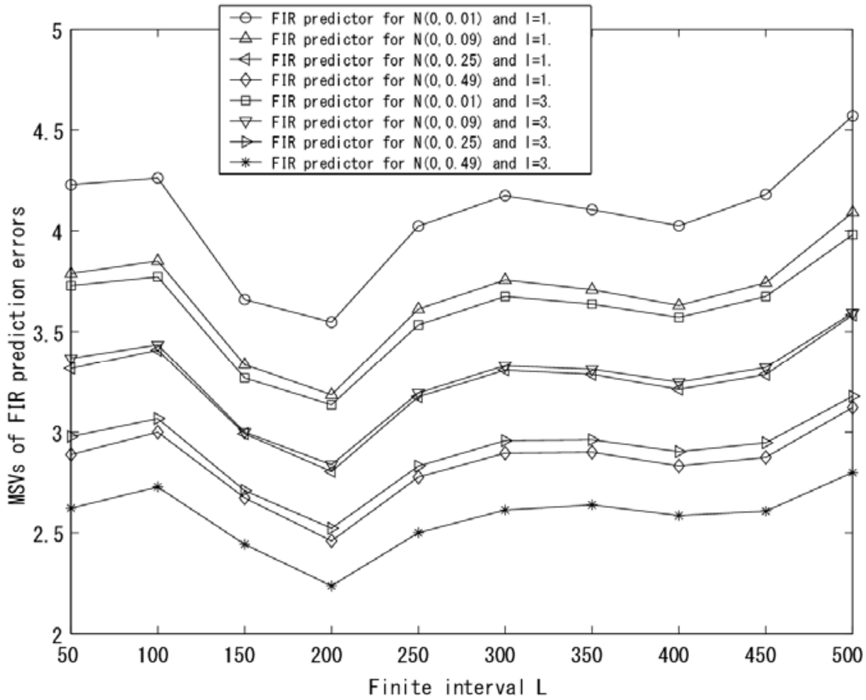


Fig.4 MSVs of prediction errors $z(k+l)-\hat{z}(k+l\,|\,k-L+1)$ by RLS Wiener FIR predictor [15] vs. finite interval $L$, in the cases of $l=1$ and $l=3$, for white Gaussian observation noises $N(0,0.1^2)$, $N(0,0.3^2)$, $N(0,0.5^2)$ and $N(0,0.7^2)$.

Fig.1 illustrates the signal $z(k+l)$ and the robust RLS Wiener FIR prediction estimate $\hat{z}(k+l|k-L+1)$, $L=200$, $l=3$ vs. $k$, $1 \le k \le 500$, for the white Gaussian observation noise $N(0,0.3^2)$. Fig.2 shows the MSVs of the one-step ahead prediction errors $z(k+1)-\hat{z}(k+1|k-L+1)$ of the signal by the robust RLS Wiener FIR predictor in Theorem 3 vs. the finite interval $L$, $50 \le L \le 500$, and the MSVs of the filtering errors $z(k)-\hat{z}(k|1)$, $1 \le k \le L$, by the robust RLS Wiener FIR filter [8] vs. $L$, $50 \le L \le 500$, for the white Gaussian observation noises $N(0,0.1^2)$, $N(0,0.3^2)$, $N(0,0.5^2)$ and $N(0,0.7^2)$. As the variance of the observation noise becomes large, the estimation accuracies of the robust RLS Wiener FIR predictor and the robust RLS Wiener FIR filter become degraded respectively. For each observation noise variance, the estimation accuracy of the robust RLS Wiener FIR filter is superior to that of the robust RLS Wiener FIR predictor. For $50 \le L \le 200$, as $L$ becomes large, the MSVs of the robust RLS Wiener FIR prediction and filtering errors become small steeply. At $L=500$, the MSVs of the robust RLS Wiener FIR prediction and filtering errors attain the smallest values for each observation noise. Fig.3 shows the MSVs of the prediction errors $z(k+l)-\hat{z}(k+l|k-L+1)$ of the signal by the robust RLS Wiener FIR predictor vs. the finite interval $L$, $50 \le L \le 500$, in the cases of $l=3$ and $l=5$, for the white Gaussian observation noises $N(0,0.1^2)$, $N(0,0.3^2)$, $N(0,0.5^2)$ and $N(0,0.7^2)$. The MSV of the prediction errors for $l=3$ is smaller than that for $l=5$ for each observation noise. For $50 \le L \le 200$, as $L$ becomes large, the MSVs of the prediction errors become small steeply. Fig.4 shows the MSVs of the prediction errors $z(k+l)-\hat{z}(k+l|k-L+1)$ of the signal by the RLS Wiener FIR predictor [15] vs. the finite interval $L$, $50 \le L \le 500$, in the cases of $l=1$ and

$l=3$, for the white Gaussian observation noises $N(0,0.1^2)$, $N(0,0.3^2)$, $N(0,0.5^2)$ and $N(0,0.7^2)$. From Fig.2-Fig. 4, the MSVs of the estimation errors by the robust RLS Wiener FIR predictor, for $l=1$ and $l=3$, are smaller than those by the RLS Wiener FIR predictor [15]. Here, the MSV of the FIR prediction errors is evaluated by $\sum_{k=L}^{1000+L}(z(k+l)-\hat{z}(k+l|k-L+1))^2/1001$.

## 6 Conclusions

This paper has newly proposed the robust RLS Wiener FIR prediction algorithm in Theorem 3, based on the innovation theory, for the linear discrete-time stochastic systems with the uncertain parameters. As a step to Theorem 3, Theorem 2 has presented the robust prediction algorithm of the signal using the covariance information etc. Also, in section 4, the recursive algorithm for the prediction error variance function has been proposed.

The prediction characteristics of the robust RLS Wiener FIR predictor have been shown in section 5. The estimation accuracy of the proposed robust RLS Wiener FIR predictor is by far superior to that of the RLS Wiener FIR predictor, but is inferior to that of the robust RLS Wiener FIR filter.

## Appendix A: Proof of Theorem 2

By introducing an equation

$$J(s)\Lambda(s)=\beta^T(s)\breve{H}^T-\sum_{i=s-1-L+1}^{s-1}J(i)\Lambda(i)g_0^T(s-1,i)\times\breve{\Phi}^T\breve{H}^T, \quad (A.1)$$

from (13) and (A.1), the optimal impulse response function $g(k,s)$ satisfies

$$g(k,s)=\alpha(k+l)J(s). \quad (A.2)$$

Likewise $g(k,s)$ in (18), it is seen that $g_0(k,s)$ satisfies

$$\begin{aligned}g_0(k,s)\Lambda_0(s)&=E[\breve{x}(k)v^T(s)]\\&=E[\breve{x}(k)(\breve{y}(s)-\breve{H}\breve{\Phi}\hat{\breve{x}}(s-1|s-1-L+1))^T]\\&=E[\breve{x}(k)\breve{y}^T(s)]-\sum_{i=s-L+1}^{s-1}E[\breve{x}(k)v^T(i)]g_0^T(s,i)\breve{\Phi}^T\breve{H}^T\\&=A(k)B^T(s)\breve{H}^T-\sum_{i=s-L+1}^{s-1}g_0(k,i)\Lambda_0(i)g_0^T(s-1,i)\times\breve{\Phi}^T\breve{H}^T.\end{aligned} \quad (A.3)$$

By introducing

$$J_0(s)\Lambda_0(s)=B^T(s)\breve{H}^T-\sum_{i=s-L+1}^{s-1}J_0(i)\Lambda_0(i)g_0^T(s-1,i)\times\breve{\Phi}^T\breve{H}^T, \quad (A.4)$$

$g_0(k,s)$ satisfies

$$g_0(k,s)=A(k)J_0(s). \quad (A.5)$$

By substituting (A.5) into (A.1), and introducing

$$r(k)=\sum_{i=k-L+1}^{k}J(i)\Lambda(i)J_0^T(i), \quad (A.6)$$

$$J(k)\Lambda(k)=\beta^T(k)\breve{H}^T-r(k-1)A^T(k-1)\breve{\Phi}^T\breve{H}^T \quad (A.7)$$

is obtained. Subtracting $r(k-1)$ from $r(k)$, we get

$$\begin{aligned}&r(k)-r(k-1)=J(k)\Lambda(k)J_0^T(k)\\&-J(k-L)\Lambda(k-L)J_0^T(k-L),\\&r(0)=0.\end{aligned} \quad (A.8)$$

Substituting (A.2) into (11) and introducing $e(k)$, given by

$$e(k)=\sum_{i=k-L+1}^{k}J(i)v(i), \quad (A.9)$$

we obtain

$$\begin{aligned}&\hat{x}(k+l|k-L+1)\\&=\alpha(k+l)\sum_{i=k-L+1}^{k}J(i)v(i),\\&=\alpha(k+l)e(k).\end{aligned} \quad (A.10)$$

By subtracting $e(k-1)$ from $e(k)$, it

follows that

$$e(k) - e(k-1) = J(k)(\breve{y}(k) \\ -\breve{H}\breve{\Phi}\hat{\breve{x}}(k-1|k-1-L+1)) \\ -J(k-L)(\breve{y}(k-L) \\ -\breve{H}\breve{\Phi} \\ \times \hat{\breve{x}}(k-L-1|k-L-1-L+1)). \quad (A.11)$$

Let the FIR filtering esti mate $\hat{\breve{x}}(k|k-L+1)$ of $\breve{x}(k)$ be given by

$$\hat{\breve{x}}(k|k-L+1) \\ = \sum_{i=k-L+1}^{k} g_0(k,i)\upsilon(i), \\ \upsilon(i) = \breve{y}(i) \\ -\breve{H}\breve{\Phi}\hat{\breve{x}}(i-1|i-1-L+1). \quad (A.12)$$

Also, by introducing

$$r_0(k) \\ = \sum_{i=k-L+1}^{k} J_0(i)\Lambda_0(i)J_0^T(i), \quad (A.13)$$

(A.4) is rewritten as

$$J_0(k)\Lambda_0(k) = B^T(k)\breve{H}^T \\ -r_0(k)A^T(k-1)\breve{\Phi}^T\breve{H}^T. \quad (A.14)$$

Subtracting $r_0(k-1)$ from $r_0(k)$, we have

$$r_0(k) - r_0(k-1) \\ = J_0(k)\Lambda_0(k)J_0^T(k) \\ -J_0(k-L)\Lambda_0(k-L)J_0^T(k-L). \quad (A.15)$$

By introducing

$$e_0(k) = \sum_{i=k-L+1}^{k} J_0(i)\upsilon(i), \quad (A.16)$$

from (A.5), the FIR f iltering estimate $\hat{\breve{x}}(k|k-L+1)$ of $\breve{x}(k)$ is given by

$$\hat{\breve{x}}(k|k-L+1) \\ = A(k)e_0(k). \quad (A.17)$$

By subtracting $e_0(k-1)$ from $e_0(k)$, it follows that

$$e_0(k) - e_0(k-1) \\ = J_0(k)(\breve{y}(k) - \breve{H}A(k)e_0(k-1)) \\ -J_0(k-L)(\breve{y}(k-L) \\ -\breve{H}A(k-L)e_0(k-L-1)), \\ A(k)e_0(k-1) \\ = \breve{\Phi}\hat{\breve{x}}(k-1|k-1-L+1), \\ A(k-L)e_0(k-L-1) \\ = \breve{\Phi} \\ \times \hat{\breve{x}}(k-L-1|k-L-1-L+1). \quad (A.18)$$

From (A.13), (A.16) and (A.17), the variance $\Lambda(k)$ of the innovation process $\upsilon(k)$ is expressed by

$$\Lambda(k) = E[\upsilon(k)\upsilon^T(k)] \\ = E[(\breve{y}(k) \\ -\breve{H}\breve{\Phi}\hat{\breve{x}}(k-1|k-1-L+1)) \\ \times (\breve{y}(k) \\ -\breve{H}\breve{\Phi}\hat{\breve{x}}(k-1|k-1-L+1))^T] \\ = \breve{H}\breve{K}(k,k)\breve{H}^T + R \\ -\breve{H}\breve{\Phi}E[\hat{\breve{x}}(k-1|k-1-L+1) \\ \times \hat{\breve{x}}^T(k-1|k-1-L+1)]\breve{\Phi}^T\breve{H}^T \\ = \breve{H}\breve{K}(k,k)\breve{H}^T + R \\ -\breve{H}A(k)r_0(k-1)A^T(k)\breve{H}^T. \quad (A.19)$$

Let the initial condition of the FIR filtering estimate of $x(k)$ at $k = L$ be $\hat{x}(L|1)$.

$$\hat{x}(L|1) = \sum_{i=1}^{L} \overline{g}(L,i)\overline{\upsilon}(i), \\ \overline{\upsilon}(i) = \breve{y}(i) - \breve{H}\breve{\Phi}\hat{\breve{x}}(i-1|1) \quad (A.20)$$

Let the variance of the in novation process $\overline{\upsilon}(L)$ be $\overline{\Lambda}(L)$.

$$\overline{\Lambda}(L) = E[\overline{\upsilon}(L)\overline{\upsilon}^T(L)] \\ = E[(\breve{y}(L) - \breve{H}\breve{\Phi}\hat{\breve{x}}(L-1|1)) \\ \times (\breve{y}(k) - \breve{H}\breve{\Phi}\hat{\breve{x}}(L-1|1))^T] \\ = \breve{H}\breve{K}(L,L)\breve{H}^T + R \\ -\breve{H}\breve{\Phi}E[\hat{\breve{x}}(L-1|1) \\ \times \hat{\breve{x}}^T(L-1|1)]\breve{\Phi}^T\breve{H}^T \\ = \breve{H}\breve{K}(L,L)\breve{H}^T + R \\ -\breve{H}A(L)\overline{r}_0(L-1)A^T(L)\breve{H}^T \quad (A.21)$$

Here,

$$\overline{r}_0(L) = \sum_{i=1}^{L} \overline{J}_0(i)\overline{\Lambda}(i)\overline{J}_0^T(i). \quad (A.22)$$

Let the initial condition of the FIR filtering estimate of $\hat{\breve{x}}(k)$ at $k = L$ be $\hat{\breve{x}}(L|1)$.

$$\hat{\breve{x}}(L|1) = \sum_{i=1}^{L} \overline{g}_0(L,i)\overline{v}(i),$$
$$\overline{v}(k) = \breve{y}(k) - \breve{H}\breve{\Phi}\hat{\breve{x}}(k-1|1) \quad \text{(A.23)}$$

Here, $\overline{g}_0(L,s)$ satisfies

$$\overline{g}_0(L,s)\overline{\Lambda}(s) = \breve{K}(L,s)\breve{H}^T - \sum_{i=1}^{s-1} \overline{g}_0(L,i)\overline{\Lambda}(i)\overline{g}_0^T(s-1,i) \times \breve{\Phi}^T\breve{H}^T. \quad \text{(A.24)}$$

$\overline{g}(L,s)$ in (A.20) satisfies

$$\overline{g}(L,s)\overline{\Lambda}(s) = K_{x\breve{x}}(L,s)\breve{H}^T - \sum_{i=1}^{s-1} \overline{g}(L,i)\overline{\Lambda}(i)\overline{g}_0^T(s-1,i)\breve{\Phi}^T\breve{H}^T, \quad \text{(A-25)}$$
$$K_{x\breve{x}}(L,s)\breve{H}^T = K_{x\breve{z}}(L,s).$$

By introducing

$$\overline{J}(s)\overline{\Lambda}(s) = \beta^T(s)\breve{H}^T - \sum_{i=1}^{s-1} \overline{J}(i)\overline{\Lambda}(i)\overline{g}_0^T(s-1,i)\breve{\Phi}^T\breve{H}^T, \quad \text{(A.26)}$$

$\overline{g}(L,s)$ is given by

$$\overline{g}(L,s) = \alpha(L)\overline{J}(s). \quad \text{(A.27)}$$

By introducing

$$\overline{r}(L) = \sum_{i=1}^{L} \overline{J}(i)\overline{\Lambda}(i)\overline{J}_0^T(i), \quad \text{(A.28)}$$

(A.26) is rewritten as

$$\overline{J}(L)\overline{\Lambda}(L) = \beta^T(L)\breve{H}^T - \overline{r}(L-1)A^T(L-1)\breve{\Phi}^T\breve{H}^T, \quad \text{(A.29)}$$
$$\breve{\Phi}^T = A^T(1).$$

By substituting (A.27) i nto (A.20) and introducing

$$\overline{e}(L) = \sum_{i=1}^{L} \overline{J}(i)\overline{v}(i), \quad \text{(A.30)}$$

$$\hat{x}(L|1) = \alpha(L)\overline{e}(L) \quad \text{(A.31)}$$

is obtained, Subtracting $\overline{r}(k-1)$ from $\overline{r}(k)$, we obtain

$$\overline{r}(k) = \overline{r}(k-1) + \overline{J}(k)\overline{\Lambda}(k)\overline{J}_0^T(k),$$
$$\overline{r}(0) = 0. \quad \text{(A.32)}$$

By subtracting $\overline{e}(L-1)$ from $\overline{e}(L)$, it follows that

$$\overline{e}(L) = \overline{e}(L-1) + \overline{J}(L)(\breve{y}(L) - \breve{H}\breve{\Phi}\hat{\breve{x}}(L-1|1)), \quad \text{(A.33)}$$
$$\overline{r}(0) = 0.$$

By introducing

$$\overline{J}_0(s)\overline{\Lambda}(s) = B^T(s)\breve{H}^T - \sum_{i=1}^{s-1} \overline{J}_0(i)\overline{\Lambda}(i)\overline{g}_0^T(s-1,i)\breve{\Phi}^T\breve{H}^T, \quad \text{(A.34)}$$

$\overline{g}_0(k,s)$ is given by

$$\overline{g}_0(L,s) = A(L)\overline{J}_0(s). \quad \text{(A.35)}$$

From (A.22), (A.25)

$$\overline{J}_0(L)\overline{\Lambda}(L) = B^T(L)\breve{H}^T - \overline{r}_0(L-1)A^T(L)\breve{H}^T \quad \text{(A.36)}$$

is obtained. Subtracting $\overline{r}_0(L-1)$ from $\overline{r}_0(L)$, we obtain

$$\overline{r}_0(L) = \overline{r}_0(L-1) + \overline{J}_0(L)\overline{\Lambda}(L)\overline{J}_0^T(L), \quad \text{(A.37)}$$
$$\overline{r}_0(0) = 0.$$

By substituting (A.35) i nto (A.23) and introducing

$$\overline{e}_0(L) = \sum_{i=1}^{L} \overline{J}_0(i)\overline{v}(i), \quad \text{(A.38)}$$

$$\hat{\breve{x}}(L|1) = A(L)\overline{e}_0(L) \quad \text{(A.39)}$$

is obtained. Subtracting $\overline{e}_0(L-1)$ from $\overline{e}_0(L)$, we obtain

$$\overline{e}_0(L) = \overline{e}_0(L-1) + \overline{J}_0(L)(\breve{y}(L) - \breve{H}A(L)\overline{e}_0(L-1)), \quad \text{(A.40)}$$
$$A(L)\overline{e}_0(L-1) = \breve{\Phi}\hat{\breve{x}}(L-1|1),$$
$$\overline{e}_0(0) = 0.$$

(Q.E.D.)

## Appendix B: Proof of Theorem 3

By substituting (A.11) i nto (A.10) and introducing

$$G(k) = \alpha(k)J(k), \quad \text{(B.1)}$$

(42) and (44) are clear. By substituting (A.18) into (A.17) and introducing

$$G_0(k) = A(k)J_0(k), \quad (B.2)$$

(45) is obtained. By substituting (A.7) into (B.1), using (10) and introducing

$$S(k) = \alpha(k)r(k)\alpha^T(k), \quad (B.3)$$

(46) is obtained. By substituting (A.14) into (B.2), using (6) and introducing

$$S_0(k) = A(k)r_0(k)A^T(k), \quad (B.4)$$

(47) is obtained. From (A-19) and (B.4), by using $A(k) = \breve{\Phi}^k$, (48) is obtained. By substituting (A.8) into (B.3) and using (B.1) with $\alpha(k) = \Phi^k$ and $A(k) = \breve{\Phi}^k$, (49) is obtained. By substituting (A.15) into (B.4) and using (B.2) with $A(k) = \breve{\Phi}^k$, (50) is obtained.

By substituting (A.33) into (A.31) and introducing

$$\overline{G}(L) = \alpha(L)\overline{J}(L), \quad (B.5)$$

(51) is obtained. By substituting (A.29) into (B.5) and introducing

$$\overline{S}(L) = \alpha(L)\overline{r}(L)A^T(L), \quad (B.6)$$

(52) is obtained. By substituting (A.40) into (A.39) and introducing

$$\overline{G}_0(L) = A(L)\overline{J}_0(L), \quad (B.7)$$

(53) is obtained. By substituting (A.36) into (B.7) and introducing

$$\overline{S}_0(L) = A(L)\overline{r}_0(L)A^T(L), \quad (B.8)$$

(54) is obtained. From (A.21) and (B.8), (55) is obtained. By substituting (A.32) into (B.6), (56) is obtained. By substituting (A.37) into (B.8), (57) is obtained.

(Q.E.D.)

***References:***

[1] M. M. Wong, D. Wong, C. Zhang, I. Hijazin, Stochastic inner product core for digital FIR filters, *WSEAS Trans. Systems and control*, Vol. 12, 2017, pp.246-252

[2] H. Nazaripouya, P. Chu, H. Pota, R. Gadh, Design of minimum-length, minimum-phase, low-group-delay FIR Filter using convex optimization method, *WSEAS Trans. Signal Processing*, Vol. 14, 2018, pp. 170-178.

[3] R. P. N. Rao, Robust Kalman filters for prediction, recognition, and learning, *Technical Report 645*, The University of Rochester, Computer Science Department, 1996, pp.1-14.

[4] C. D. Zuluaga, M. A. Alvarez and E. Giraldo, Short-term wind speed prediction based on robust Kalman filtering: An experimental comparison, *Applied Energy*, Vol.156, 2015, pp.321-330.

[5] X. Wang, W. Liu and Z. Deng, Robust centralized fusion steady-state Kalman predictor with uncertain parameters, *Proceedings of the 2015 Chinese Intelligent Automation Conference*, pp. 23-31.

[6] S. Nakamori, Recursive estimation technique of signal from output measurement data in linear discrete-time systems, *IEICE Trans. Fundamentals of Electronics, Communication and Computer Sciences*, Vol.E78-A, No.5, 1995, pp. 600–607.

[7] S. Nakamori, Robust RLS Wiener signal estimators for discrete-time stochastic systems with uncertain parameters, *Frontiers in Signal Processing*, Vol.3, No.1, 2019, pp.1–18.

[8] S. Nakamori, Robust RLS Wiener FIR filter for signal estimation in linear discrete-time stochastic systems with uncertain parameters, *Frontiers in Signal Processing*, Vol.3, No.2, 2019, pp.19-36.

[9] X. Zhu, Y. C. Soh and L. Xie, Design and analysis of discrete-time robust Kalman filters, *Automatica*, Vol.38, No.6, 2002, pp.1069–1077.

[10] A. H. Jazwinski, Limited memory optimal filtering. *IEEE Trans. Automatic Control*, Vol.13, No.10, 1968, pp.558–563.

[11] W. H. Kwon and S. Han, *Receding Horizon Control: Model Predictive*

*Control for State Models*, Springer-Verlag, 2005.

[12] S. Zhao, Y. S. Shmaliy, P. Shi and C. K. Ahn, Fusion Kalman/UFIR filter for state estimation with uncertain parameters and noise statistics. *IEEE Trans. Industrial Electronics*, Vol.64, No.4, 2017, pp.3075–3083.

[13] S. Zhao, Y. S. S hmaliy and F. Liu, Fast Kalman-like optimal unbiased FIR filtering with applications, *IEEE Trans. Signal Processing*, Vol.64, No.9, 2016, pp.2284–2297.

[14] S. Zhao, Y. S. S hmaliy and F. Liu, Fast Kalman-like optimal FIR filter for time-variant systems with improved robustness, *ISA Transactions*, Vol.80, 2018, pp.160–168.

[15] S. Nakamori, RLS Wiener FIR predictor and filter based on innovation approach in line ar discrete-time stochastic systems, *Frontiers in Signal Processing*, Vol.1, No.2, 2017, pp.49–61.

[16] S. Nakamori and A. Hataji, Design o f new predictor usin g covariance information and its application to prediction of air pollut ion levels, *Electronics and Communications in Japan (Part I: Communications)*, Vol.63, No.9, 1980, pp.29-38.

[17] A. P. Sage and J. L. Melsa, *Estimation Theory with Applications to Communications and Control*, McGraw-Hill, 1971.