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 uncertain 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 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


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 uncertain noise variance $s$, 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 impulse 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 \( \tilde{z}(k) \). (2) The observation matrices for the signal and the degraded signal. (3) The variance \( K(k,k) \) of the state \( \tilde{x}(k) \) for the degraded signal. (4) The cross-variance \( K_{xx}(k,k) \) of the state \( x(k) \) for the signal with the state \( \tilde{x}(k) \). (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

\[
\begin{align*}
\hat{y}(k) &= \tilde{z}(k) + v(k), \\
\tilde{z}(k) &= \tilde{H}(k)\tilde{x}(k), \\
\tilde{H}(k) &= H + \Delta H(k), \\
\tilde{x}(k + 1) &= \tilde{F}(k)\tilde{x}(k) + \Gamma w(k), \\
\tilde{F}(k) &= F + \Delta F(k), \\
E[v(k)v^T(s)] &= R\delta_K(k-s), \\
E[w(k)w^T(s)] &= Q\delta_K(k-s),
\end{align*}
\]

linear discrete-time stochastic systems with the uncertain quantities \( \Delta H(k) \) and \( \Delta F(k) \). \( v(k) \) 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 \( \tilde{x}(k+1) \) includes the uncertain quantity \( \Delta F(k) \) additionally to the system matrix \( F(k) \). Also, the observation matrix \( \tilde{H}(k) \) contains the uncertain quantity \( \Delta H(k) \).

Hence, \( \tilde{z}(k) \) 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 \( \tilde{z}(k) \) and the observation noise \( v(k) \), the degraded observed value \( \hat{y}(k) \) is obtained. Compared with (1), the precise state-space model is given by

\[
\begin{align*}
y(k) &= z(k) + v(k), \\
z(k) &= Hx(k), \\
x(k + 1) = Fx(k) + \Gamma w(k).
\end{align*}
\]

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 \( \hat{y}(k) \). Here, both the robust predictors do not use any information on the uncertain quantities \( \Delta F(k) \) and \( \Delta H(k) \).
Suppose that the sequence of the degraded signal \( \hat{z}(k) \) is fitted to the AR model of the finite order \( N \) as
\[
\hat{z}(k) = -a_1 \hat{z}(k-1) - a_2 \hat{z}(k-2) \ldots -a_N \hat{z}(k-N) + \hat{e}(k),
\]
\[
E[\hat{e}(k)\hat{e}(s)^T] = \hat{Q}\delta_K(k-s).
\] (3)

Let the degraded signal \( \hat{z}(k) \) be represented with the state \( \hat{x}(k) \) by
\[
\hat{z}(k) = \hat{H}\hat{x}(k),
\]
\[
\hat{x}(k) = \begin{bmatrix}
\hat{x}_1(k) \\
\hat{x}_2(k) \\
\vdots \\
\hat{x}_{N-1}(k) \\
\hat{x}_N(k)
\end{bmatrix}, 
\]
\[
\hat{H} = [I_{m \times m} \ 0 \ \ldots \ 0].
\]

Henceforth, the state equation for the state \( \hat{x}(k) \) is expressed by
\[
\begin{bmatrix}
\hat{x}_1(k+1) \\
\hat{x}_2(k+1) \\
\vdots \\
\hat{x}_{N-1}(k+1) \\
\hat{x}_N(k+1)
\end{bmatrix} =
\begin{bmatrix}
0 & I_{m \times m} & 0 \\
0 & 0 & I_{m \times m} \\
\vdots & \vdots & \vdots \\
-\hat{a}_N & -\hat{a}_{N-1} & -\hat{a}_{N-2} \\
\vdots & \vdots & \vdots \\
0 & I_{m \times m} & \hat{a}_1 \\
\vdots & \vdots & \vdots \\
0 & 0 & \hat{a}_1
\end{bmatrix}
\begin{bmatrix}
\hat{x}_1(k) \\
\hat{x}_2(k) \\
\vdots \\
\hat{x}_{N-1}(k) \\
\hat{x}_N(k)
\end{bmatrix} + \begin{bmatrix}
\zeta(k), \\
0 \\
0 \\
\vdots \\
0
\end{bmatrix}
\] (5)

Let \( \hat{K}(k,s) = \hat{R}(k-s) \) represent the auto-covariance function of the state \( \hat{x}(k) \) in wide-sense stationary stochastic systems [17]. \( \hat{R}(k,s) \) is expressed in the semi-degenerate kernel form of
\[
\hat{R}(k,s) = \begin{cases}
\hat{A}(k)\hat{B}^T(s), & 0 \leq s \leq k, \\
\hat{B}(k)\hat{A}^T(s), & 0 \leq k \leq s,
\end{cases}
\]
\[
\hat{A}(k) = \hat{\Phi}^k, \hat{B}^T(s) = \hat{\Phi}^{-s}\hat{K}(s,s).
\] (6)

Here, \( \hat{\Phi} \) is the system matrix for the state \( \hat{x}(k) \). From (5), the system matrix \( \hat{\Phi} \) is expressed by
\[
\hat{\Phi} =
\begin{bmatrix}
0 & I_{m \times m} & 0 \\
0 & 0 & I_{m \times m} \\
\vdots & \vdots & \vdots \\
-\hat{a}_N & -\hat{a}_{N-1} & -\hat{a}_{N-2} \\
\vdots & \vdots & \vdots \\
0 & 0 & 0 \\
\cdots & \cdots & \cdots \\
0 & I_{m \times m} & 0 \\
\cdots & \cdots & \cdots \\
0 & 0 & I_{m \times m}
\end{bmatrix}
\] (7)

By putting \( \hat{K}_g(s) = \hat{K}_g(k-s) = E[\hat{z}(k)\hat{z}^T(s)] \), the auto-variance function \( \hat{K}(k,k) \) of the state \( \hat{x}(k) \) is expressed by
\[
\hat{K}(k,k) = E\begin{bmatrix}
\hat{z}(k) \\
\hat{z}(k+1) \\
\vdots \\
\hat{z}(k+N-2) \\
\hat{z}(k+N-1)
\end{bmatrix} \times
\begin{bmatrix}
\hat{z}^T(k) \\
\hat{z}^T(k+1) \\
\vdots \\
\hat{z}^T(k+N-2) \\
\hat{z}^T(k+N-1)
\end{bmatrix}
\] (8)

With \( \hat{K}_g(k-s) \), the Yule-Walker equation for the AR parameters is formulated as
Let $\hat{K}(k,s) = K_{xx}(k-s) = E[x(k)\hat{x}(s)]$ be the cross-covariance function of the state $x(k)$ with the state $\hat{x}(s)$ in wide-sense stationary stochastic systems. $K_{xx}(k,s)$ is expressed in the functional form of

$$K_{xx}(k,s) = \alpha(k)\beta^T(s), 0 \leq s \leq k, \quad \alpha(k) = \Phi^k, \beta^T(s) = \Phi^{-s}K_{xx}(s,s).$$

(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

$$\hat{x}(k+1) = \sum_{i=k-L+1}^{k} g(k,i)v(i),$$

(11)

where $g(k,i) = \frac{y^{(i)}}{H^T\hat{x}^{(i-1)}} - H^T\hat{x}^{(i-1)}L+1$, and in terms of the innovation process $\{v(i), k-L+1 \leq i \leq k\}$. In (11), $g(k,i)$ represents the time-varying impulse response function and $\hat{x}(i-1) | i-1 - L + 1)$ is the FIR filtering estimate of the state $\hat{x}(i-1)$. Let the FIR filtering estimate $\hat{x}(k|k-L+1) of \hat{x}(k)$ be given by

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

(12)

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

$$g(k,s)\Lambda(s) = K_{xx}(k+l,s)H^T$$

(13)

$$- \sum_{i=s-1}^{s-1} g(k,i)\Lambda(i)g^T_0 (s-1, i) \times \Phi^T H^T, \quad K_{xx}(k+l, s)H^T = K_{xx}(k+l, s).$$

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

**Proof**

Consider the estimation problem, which minimizes the MSV

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

(14)

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

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

(15)

$k-L+1 \leq s \leq k,$

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

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

(16)

$k-L+1 \leq s \leq k.$

From (16), (17) and the expression for the innovation process $v(s)$, $g(k,s)$ satisfies
Let the variance degraded signal and observation matrices respectively for the signal (1). Let quantities observation equation, which contains the uncertain

\[
\begin{align*}
&g(k,s)A(s) = E[x(k + l)u^T(s)] \\
&= E[x(k + l)(\bar{y}(s) - \bar{H}\bar{R}(s - 1)s - 1 + L + 1)^T] \\
&= E[x(k + l)\bar{y}^T(s)] \\
&= E[x(k + l)\bar{x}T(s)] \\
&\times \bar{x}T(s - 1)\bar{R}(s - 1 + L + 1)]\bar{D}.T. \\
\end{align*}
\]

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

\[
\begin{align*}
&= E[x(k + l)(\bar{x}(s) + \nu(s))] \\
&= E[x(k + l)\bar{x}T(s)] \\
&= E[x(k + l)\bar{x}T(s)]B^T \\
&= K_{xx}(k + l,s)B^T. \\
\end{align*}
\]

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

\begin{align*}
&z(k + l|k - L + 1) = \hat{z}(k + l|k - L + 1) \\
&= H\bar{x}(k + l|k - L + 1) + \nu(k) \\
&= \alpha(k + l)e(k) \\
&= a(k + l)e(k) \quad (21)
\end{align*}

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

\[
\hat{z}(k|k - L + 1) = H\bar{x}(k|k - L + 1)
\]

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

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

\[
\hat{x}(k|k - L + 1) = A(k)e_0(k) \quad (24)
\]

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

Recursive equation for \(e(k)\):

\[
\begin{align*}
e(k) &= e(k - 1) + J(k)\gamma(k) \\
&- H\alpha(k)e_0(k - 1) \\
&- J(k - L)\gamma(k) - \bar{H}A(k)e_0(k - L - 1) \\
\end{align*}
\]

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

Recursive equation for \(\tilde{e}_0(k)\):

\[
\begin{align*}
&\tilde{e}_0(k) = e_0(k - 1) + J_0(k)\gamma(k) \\
&- H\alpha(k)e_0(k - 1) \\
&- J_0(k - L)\gamma(k - L) \\
&- \bar{H}A(k - L)e_0(k - L - 1) \\
\end{align*}
\]

Initial condition of \(\tilde{e}_0(k)\) at \(k = L: \tilde{e}_0(L)\)

**Equation for \(J(k)\):**

\[
\begin{align*}
&J(k) = \beta^T(k)\bar{H} \\
&- r(k - 1)A^T(k - 1)\beta^T(k)\tilde{H}\Lambda^{-1}(k) \\
&= r(k) = r(k - 1) + J(k)\Lambda(k)J_0^T(k) \\
&- J(k - L)\Lambda(k - L)J_0^T(k - L) \\
&- J(k - L)\Lambda(k - L)J_0^T(k - L)
\end{align*}
\]

**Equation for \(J_0(k)\):**

\[
\begin{align*}
&J_0(k) = \beta(k)A^T(k)\bar{H} \\
&- r(k - 1)A^T(k - 1)\beta(k)\tilde{H}\Lambda^{-1}(k) \\
&= r(k) = r(k - 1) + J_0(k)\Lambda(k)J_0^T(k) \\
&- J_0(k - L)\Lambda(k - L)J_0^T(k - L) \\
&- J_0(k - L)\Lambda(k - L)J_0^T(k - L)
\end{align*}
\]
Equation for \( r_0(k) \):
\[
J_0(k) = [B^T(k) \bar{H}^T - r_0(k)A^T(k - 1) \Phi^T \bar{H}^T] \Lambda^{-1}(k)
\] (29)

Recursive equation for \( r_0(k) \):
\[
r_0(k) = r_0(k - 1) + J_0(k) \Lambda(k) J_0^T(k)
\] (30)

Initial condition of \( r_0(k) \) at \( k = L \):
\[
\bar{T}_0(L)
\] (31)

Equation for \( \Lambda(k) \):
\[
\Lambda(k) = \bar{H} \bar{K}(k, k) \bar{H}^T + R
\] (32)

Recursive equation for \( \overline{\sigma}(L) \):
\[
\overline{\sigma}(L) = \overline{\sigma}(L - 1) + J(L) \overline{\sigma}(L) J^T(L),
\] (33)

Equation for \( \overline{J}(L) \):
\[
\overline{J}(L) = (B^T(L) \bar{H}^T - \overline{\sigma}(L - 1) A^T(L - 1) A^T(1) \bar{H}^T) \Lambda^{-1}(L)
\] (34)

Recursive equation for \( \overline{\tau}(L) \):
\[
\overline{\tau}(L) = \overline{\tau}(L - 1) + J_0(L) \overline{\tau}(L) J_0^T(L),
\] (35)

Initial condition of \( \overline{\tau}(L) \) at \( k = L \):
\[
\overline{T}_0(L)
\] (36)

Equation for \( \overline{J}_0(L) \):
\[
\overline{J}_0(L) = (B^T(L) \bar{H}^T - \overline{\tau}(L - 1) A^T(L) \bar{H}^T) \Lambda^{-1}(L)
\] (37)

Recursive equation for \( \overline{\tau}_0(L) \):
\[
\overline{\tau}_0(L) = \overline{\tau}_0(L - 1) + \overline{J}_0(L) \overline{\tau}_0(L) \overline{J}_0^T(L),
\] (38)

Equation for \( \overline{\Lambda}(L) \):
\[
\overline{\Lambda}(L) = \bar{H} \bar{K}(L, L) \bar{H}^T + R
\] (39)

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 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 \( \Phi \) and \( H \) represent the system and observation matrices respectively for the degraded signal \( \tilde{z}(k) \), which is fitted to the AR model (3). Let the variance \( \bar{K}(k, k) \) of the state \( \tilde{x}(k) \) for the degraded signal \( \tilde{z}(k) \) and the cross-variance \( K_{x\tilde{x}}(k, k) \) of the state \( x(k) \) for the signal \( z(k) \) with the state \( \tilde{x}(k) \) 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 \( \hat{x}(k + l)|k + l - 1 \) 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 \( z(k + l) \):
\[
\hat{z}(k + l)|k + l - 1 = \Phi \hat{x}(k + l)|k + l - 1
\] (40)

FIR filtering estimate of the signal \( z(k) \):
\[
\hat{x}(k + 1) = A \hat{x}(k + 1)
\] (41)

FIR filtering estimate of the state \( x(k) \):
\[
\hat{x}(k + 1) = \Phi \hat{x}(k + 1)
\] (42)

FIR filtering estimate of the state \( x(k) \):
\[
\hat{x}(k + 1) = \Phi \hat{x}(k + 1)
\] (43)

FIR filtering estimate of the state \( x(k) \):
\[
\hat{x}(k + 1) = \Phi \hat{x}(k + 1)
\] (44)
\[ \hat{x}(k|k - L + 1) = \Phi \hat{x}(k - 1|k - 1 - L + 1) + G_o(k) \hat{y}(k) - \bar{H} \Phi \hat{x}(k - 1|k - 1 - L + 1) - \Phi^2 G_o(k - L) \hat{y}(k - L) + \Phi \hat{S}_o(k - 1|k - 1 - L + 1) \]  
(45)

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

FIR filter gain for \( \hat{x}(k|k - L + 1) \): 
\[ G(k) = [K_{xx}(k,k)H^T - \Phi \hat{S}_o(k - 1|k - 1 - L + 1)]^{-1} \]  
(46)

FIR filter gain for \( \hat{x}(k|k - L + 1) \): 
\[ G_o(k) = [\bar{K}(k,k)H^T - \Phi \hat{S}_o(k - 1|k - 1 - L + 1)]^{-1} \]  
(47)

Equation for \( \Lambda(k) \): 
\[ \Lambda(k) = R + \bar{H} \bar{K}(k,k)H^T - \Phi \hat{S}_o(k - 1|k - 1 - L + 1) \]  
(48)

Recursive equation for \( S(k) \): 
\[ S(k) = \Phi S(k - 1|k - 1) - \Phi^2 G_o(k - L) \hat{y}(k - L) \]  
(49)

Recursive equation for \( S_o(k) \): 
\[ S_o(k) = \Phi^2 G_o(k - L) \hat{y}(k - L) \]  
(50)

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

Recursive equation for \( S_o(k) \) at \( k = L \): 
\[ \bar{S}_o(L) = \Phi^2 G_o(k - L) \hat{y}(k - L) \]  
(51)

Filter gain for \( \hat{x}(L|1) \) in (51): 
\[ \bar{G}(L) = [K_{xx}(L,L) - \Phi \hat{S}(L - 1|L - 1)]^{-1} \]  
(52)

Recursive equation for \( \hat{x}(L|1) \): 
\[ \hat{x}(L|1) = \Phi \hat{x}(L - 1|1) + \bar{G}(L)(\hat{y}(L) - \bar{H} \Phi \hat{x}(L - 1|1) \]  
(54)

Equation for \( \bar{S}(L) \): 
\[ \bar{S}(L) = \Phi \bar{S}(L - 1) - \Phi^2 G_o(L - 1) \bar{H} \]  
(55)

Recursive equation for \( \bar{S}_o(L) \): 
\[ \bar{S}_o(L) = \Phi \bar{S}_o(L - 1) - \Phi^2 G_o(L - 1) \bar{H} \]  
(56)

Recursive equation for \( \bar{S}_o(L) \): 
\[ \bar{S}_o(L) = \Phi \bar{S}_o(L - 1) - \Phi^2 G_o(L - 1) \bar{H} \]  
(57)

Proof of Theorem 3 is deferred to Appendix B.

Necessary conditions on the stability of the robust RLS Wiener FIR predictor presented in Theorem 3. Also, the existence of the robust RLS Wiener FIR predictor estimate \( \hat{x}(k + l|k - L + 1) \) of the signal \( z(k + l|k - L + 1) \) is shown.

### 4 Prediction error variance function of signal

Let the variance function of the FIR prediction error \( z(k + l|k - L + 1) \) be denoted by \( \bar{P}_e(k + l) \). Let the auto-covariance function \( K(k,s) \) of the state \( x(k) \) be expressed by 
\[ K(k,s) = \begin{cases} A_x(k)B_x^T(s), & 0 \leq s \leq k, \\ B_x(k)A_x^T(s), & 0 \leq k \leq s, \end{cases} \]  
(58)

\[ A_x(k) = \alpha(k) = \Phi^k, \]  
\[ B_x^T(s) = \Phi^{-s} K(s,s). \]  
(59)
From (16) and (A.10), the FIR prediction error variance function \( \hat{P}_2(k) \) is formulated as

\[
\hat{P}_2(k + l) = H(K(k + l, k + l) - E[\hat{x}(k)\hat{x}^T(k + l|k - L + 1)])H^T
\]

\[
= H(K(k, k) - E[\hat{x}(k + l)\hat{x}^T(k + l|k - L + 1)])H^T
\]

\[
= H(K(k, k) - \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\]

\[
r(k) = \sum_{i=k-L+1}^{k} j(i)\Lambda(i)^T = \Phi^{k+1}T(k)(\Phi^T)^{k+1}H^T,
\( K(k, s) = K(k - s) \) represents the auto-covariance function of the state \( \hat{x}(k) \) in wide-sense stationary stochastic systems. \( K(k, s) \) is expressed in the semi-degenerate kernel form (6). \( \Phi \) in (6) represents the system matrix for the state \( \hat{x}(k) \). Also, from the auto-covariance function \( K_x(s, s) = E[\hat{x}(k)\hat{x}(s)] \) of the degraded signal \( \hat{z}(k) \), the auto-variance function \( K(k, k) \) represents the system matrix for the state \( \hat{x}(k) \) is expressed as

\[
\begin{bmatrix}
\hat{z}(k) \\
\hat{z}(k + 1) \\
\vdots \\
\hat{z}(k + N - 2) \\
\hat{z}(k + N - 1)
\end{bmatrix}
\]

\times \begin{bmatrix}
\hat{z}(k) \\
\hat{z}(k + 1) \\
\vdots \\
\hat{z}(k + N - 2) \\
\hat{z}(k + N - 1)
\end{bmatrix}
\]

\begin{bmatrix}
K_x(0) & K_x(1) & \cdots & K_x(N - 2) & K_x(N - 1) \\
K_x(1) & K_x(0) & \cdots & K_x(N - 3) & K_x(N - 2) \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
K_x(N - 2) & K_x(N - 1) & \cdots & K_x(0) & K_x(1) \\
K_x(N - 1) & K_x(N - 2) & \cdots & K_x(1) & K_x(0)
\end{bmatrix} \tag{66}
\]

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

\[
K_{xz}(k, s) = \Phi^{k-s}K_{x}(s, s),
\]

\( 0 \leq s \leq k \),

\[
K_{x}(k, k) = \begin{bmatrix}
K_{x}(k, k) \\
K_{x}(k + 1, k) \\
\vdots \\
K_{x}(k + N - 2, k) \\
K_{x}(k + N - 1, k)
\end{bmatrix}
\] \tag{67}

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

**Fig.1** Signal \( z(k + l) \) and robust RLS Wiener FIR prediction estimate \( \hat{z}(k + l | k - L + 1) \),

\( L = 200 \), \( l = 3 \) 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 $L$, $50 \leq L \leq 500$, and MSVs of filtering errors $z(k) - \hat{z}(k|1)$, $1 \leq k \leq L$, by robust RLS Wiener FIR filter [8] vs. $L$, $50 \leq L \leq 500$, 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 $N(0,0.7^2)$. 

Seiichi Nakamori
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. $k$, $1 \leq k \leq 500$, 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 $L$, $50 \leq L \leq 500$, and the MSVs of the filtering errors $z(k) - \hat{z}(k|1)$, $1 \leq k \leq L$, by the robust RLS Wiener FIR filter [8] vs. $L$, $50 \leq L \leq 500$, 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 $50 \leq L \leq 200$, 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 $L$, $50 \leq L \leq 500$, 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 $50 \leq L \leq 200$, 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 $L$, $50 \leq L \leq 500$, in the case $s$ of $l = 1$ and
By introducing an equation

\[ g_0(k,s)A_0(s) = E[\bar{x}(k)u^T(s)] = E[\bar{x}(k)v(s)] - H\delta^2(s - 1) \]

where

\[ -H\delta^2(s - 1) = [H \delta^2(s - 1)]^T \]

and

\[ E[\bar{x}(k)v(s)] = E[\bar{x}(k)v^T(s)] \]

have been shown in section 5. The estimation errors of 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

\[ \sum_{k=L}^{N} \sum_{l=1}^{L} (z(k + l) - \hat{z}(k + l|k - L + 1))^2 / 1001. \]

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

\[ J(s)A(s) = \beta^T(s)H^T \]

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

\[ g(k,s) = \alpha(k + l)f(s). \] (A.2)

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

\[ g_0(k,s)A_0(s) = E[\bar{x}(k)u^T(s)] = E[\bar{x}(k)v(s)] - H\delta^2(s - 1) \]

and

\[ E[\bar{x}(k)v(s)] = E[\bar{x}(k)v^T(s)] \]

By subtracting

\[ \sum_{i=0}^{k-1} J(i)A(i)g_0^T(s - 1,i) \]

we obtain

\[ \bar{x}(k + l|k - L + 1) \]

and

\[ \alpha(k + l) \sum_{i=k-L+1}^{k} J(i)v(i) \]

By subtracting \( e(k - 1) \) from \( e(k) \), it
follows that
e(k) - e(k - 1) = f(k)(y(k) - e(k - 1))
-\hat{H}\bar{\Phi}\tilde{x}(k - 1|k - 1 - L + 1)
-J(k - L)(y(k) - L)
-\hat{H}\bar{\Phi}
\times \tilde{x}(k - L - 1|k - L - 1 + 1). \tag{A.11}

Let the FIR filtering estimate \(\hat{x}(k|k - L + 1)\) of \(x(k)\) be given by
\[
\hat{x}(k|k - L + 1) = \sum_{i=k-L+1}^{k} g_0(k,i)v(i), \tag{A.12}
\]
\[
v(k) = \hat{y}(i) - \hat{H}\bar{\Phi}\tilde{x}(i - 1|i - 1 - L + 1).
\]
Also, by introducing
\[
r_0(k) = \sum_{i=k-L+1}^{k} j_0(i)\Lambda_0(i)j_0^T(i), \tag{A.13}
\]
\[
(A.4) \text{ is rewritten as}
\]
\[
J_0(k)\Lambda_0(k) = B^T(k)H^T
-\hat{r}(k)A^T(k - 1)\bar{\Phi}^TH^T. \tag{A.14}
\]
Subtracting \(r_0(k - 1)\) from \(r_0(k)\), we have
\[
r_0(k) - r_0(k - 1) = J_0(k)\Lambda_0(k)j_0^T(k)
-J_0(k - L)\Lambda_0(k - L)j_0^T(k - L). \tag{A.15}
\]
By introducing
\[
e_0(k) = \sum_{i=k-L+1}^{k} j_0(i)v(i) \tag{A.16}
\]
from (A.5), the FIR filtering estimate \(\hat{x}(k|k - L + 1)\) of \(x(k)\) is given by
\[
\hat{x}(k|k - L + 1) = \hat{A}(k)e_0(k). \tag{A.17}
\]
By subtracting \(e_0(k - 1)\) from \(e_0(k)\), it follows that
\[
e_0(k) - e_0(k - 1)
= J_0(k)(\hat{y}(k) - \hat{H}\hat{A}(k)e_0(k - 1))
-J_0(k - L)(\hat{y}(k) - L)
-\hat{H}\hat{A}(k - L)e_0(k - L - 1),
\hat{A}(k)e_0(k)
\times \hat{x}(k - L - 1|k - L - 1 + 1). \tag{A.18}
\]
From (A.13), (A.16) and (A.17), the variance \(\Lambda(k)\) of the innovation process \(v(k)\) is expressed by
\[
\Lambda(k) = E[v(k)v^T(k)]
= E[\hat{y}(k)\hat{y}(k)]
-\hat{H}\bar{\Phi}\tilde{x}(k - 1|k - 1 - L + 1)\times(\hat{y}(k)
-\hat{H}\bar{\Phi}\tilde{x}(k - 1|k - 1 - L + 1))^T
\]
\[
\hat{H}\hat{R}(k,k)\hat{H}^T + R
-\hat{H}\hat{\Phi}E[\tilde{x}(k - 1|k - 1 - L + 1)\times\tilde{x}^T(k - 1|k - 1 - L + 1)]\hat{\Phi}^TH^T
\]
\[
= \hat{H}\hat{R}(k,k)\hat{H}^T + R
-\hat{H}\hat{A}(k)r_0(k - 1)A^T(k)\hat{H}^T. \tag{A.19}
\]
Let the initial condition of the FIR filtering estimate of \(x(k)\) at \(k = L\) be \(\hat{x}(L|1)\).
\[
\hat{x}(L|1) = \sum_{i=1}^{L} \bar{v}(L, i)v(i), \tag{A.20}
\]
\[
\bar{v}(i) = \hat{y}(i) - \hat{H}\bar{\Phi}\tilde{x}(i - 1|1)
\]
Let the variance of the innovation process \(\bar{v}(L)\) be \(\bar{\Lambda}(L)\).
\[
\bar{\Lambda}(L) = E[\bar{v}(L)\bar{v}^T(L)]
= E[(\hat{y}(L) - \hat{H}\hat{\Phi}\tilde{x}(L - 1|1))
\times(\hat{y}(k) - \hat{H}\hat{\Phi}\tilde{x}(k - 1|1))^T]
\]
\[
= \hat{H}\hat{R}(L,L)\hat{H}^T + R
-\hat{H}\hat{\Phi}E[\tilde{x}(L - 1|1)\tilde{x}^T(L - 1|1)]\hat{\Phi}^TH^T
\]
\[
= \hat{H}\hat{R}(L,L)\hat{H}^T + R
-\hat{H}\hat{A}(L)r_0(L - 1)A^T(L)\hat{H}^T. \tag{A.21}
\]
Here,
\[
r_0(L) = \sum_{i=1}^{L} j_0(i)\bar{\Lambda}(L)j_0^T(i). \tag{A.22}
\]
Let the initial condition of the FIR filtering estimate of \(\hat{x}(k)\) at \(k = L\) be \(\hat{x}(L|1)\).
By introducing (A.26) is rewritten as
\[
\hat{\zeta}(L|1) = \sum_{i=1}^{l} g_{0}(L, i) \hat{v}(i),
\]  
(A.23)
\[
\bar{v}(k) = \hat{y}(k) - \tilde{H} \hat{\Phi}(k-1|1)
\]

Here, \(g_{0}(L, s)\) satisfies
\[
\begin{align*}
&g_{0}(L, s) \overline{A}(s) = K(L, s) \tilde{H}^{T} \\
&\quad - \sum_{i=1}^{s-1} g_{0}(L, i) \overline{A}(i) g_{0}^{T}(s-1, i) \Phi^{T} \tilde{H}^{T},
\end{align*}
\]  
(A.24)
\[
\overline{g}(L, s) \text{ in (A.20) satisfies}
\begin{align*}
&\overline{g}(L, s) \overline{A}(s) = K_{xx}(L, s) \tilde{H}^{T} \\
&\quad - \sum_{i=1}^{s-1} \overline{g}(L, i) \overline{A}(i) \overline{g}^{T}(s-1, i) \Phi^{T} \tilde{H}^{T},
\end{align*}
(A.25)
\[
K_{xx}(L, s) \tilde{H}^{T} = K_{xx}(L, s).
\]

By introducing
\[
\begin{align*}
&\bar{f}(s) \overline{A}(s) = \beta^{T}(s) \tilde{H}^{T} \\
&\quad - \sum_{i=1}^{s-1} \bar{f}(i) \overline{A}(i) \bar{g}^{T}(s-1, i) \Phi^{T} \tilde{H}^{T},
\end{align*}
\]  
(A.26)
\[
\overline{g}(L, s) \text{ is given by}
\begin{align*}
&\overline{g}(L, s) = \alpha(L) \bar{f}(s).
\end{align*}
\]  
(A.27)

By introducing
\[
\begin{align*}
&\bar{\tau}(L) = \sum_{i=1}^{l} \bar{f}(i) \overline{A}(i) \bar{g}(i)^{T},
\end{align*}
\]  
(A.28)
\[
(A.26) \text{ is rewritten as}
\begin{align*}
&\bar{f}(L) \overline{A}(L) = \beta^{T}(L) \tilde{H}^{T} \\
&\quad - \bar{\tau}(L-1) \alpha^{T}(L-1) \Phi^{T} \tilde{H}^{T},
\end{align*}
(A.29)
\[
\tilde{\Phi}^{T} = \alpha^{T}(1).
\]

By substituting (A.27) into (A.20) and introducing
\[
\begin{align*}
&\bar{\zeta}(L) = \sum_{i=1}^{l} \bar{f}(i) \overline{A}(i) \bar{v}(i),
\end{align*}
\]  
(A.30)
\[
\hat{\zeta}(L|1) = \alpha(L) \bar{\zeta}(L)
\]  
(A.31)
is obtained. Subtracting \(\bar{\tau}(k-1)\) from \(\bar{\tau}(k)\), we obtain
\[
\begin{align*}
&\bar{\tau}(k) = \bar{\tau}(k-1) + \hat{f}(k) \overline{A}(k) \bar{g}(k)^{T},
\end{align*}
\]  
(A.32)
\[
\bar{\tau}(0) = 0.
\]

By subtracting \(\bar{v}(L-1)\) from \(\bar{v}(L)\), it follows that
\[
\begin{align*}
&\bar{v}(L) = \bar{v}(L-1) \\
&\quad + \bar{f}(L) \overline{A}(L) \bar{g}(L) \hat{\Phi}(k-1|1),
\end{align*}
\]  
(A.33)
\[
\bar{v}(0) = 0.
\]

By introducing
\[
\begin{align*}
&\bar{f}_{0}(s) \overline{A}(s) = B^{T}(s) \tilde{H}^{T} \\
&\quad - \sum_{i=1}^{s-1} \bar{f}_{0}(i) \overline{A}(i) \bar{g}_{0}^{T}(s-1, i) \Phi^{T} \tilde{H}^{T},
\end{align*}
\]  
(A.34)
\[
\bar{g}_{0}(k, s) \text{ is given by}
\begin{align*}
&\bar{g}_{0}(L, s) = A(L) \bar{f}_{0}(s).
\end{align*}
\]  
(A.35)
From (A.22), (A.25)
\[
\begin{align*}
&\bar{f}_{0}(L) \overline{A}(L) = B^{T}(L) \tilde{H}^{T} \\
&\quad - \bar{\tau}_{0}(L-1) A^{T}(L-1) \Phi^{T} \tilde{H}^{T},
\end{align*}
\]  
(A.36)
is obtained. Subtracting \(\bar{\tau}_{0}(L-1)\) from \(\bar{\tau}_{0}(L)\), we obtain
\[
\begin{align*}
&\bar{\tau}_{0}(L) = \bar{\tau}_{0}(L-1) \\
&\quad + \bar{f}_{0}(L) \overline{A}(L) \bar{g}_{0}(L),
\end{align*}
\]  
(A.37)
\[
\bar{\tau}_{0}(0) = 0.
\]

By substituting (A.35) into (A.23) and introducing
\[
\begin{align*}
&\bar{\zeta}_{0}(L) = \sum_{i=1}^{l} \bar{f}_{0}(i) \bar{v}(i),
\end{align*}
\]  
(A.38)
\[
\hat{\zeta}(L|1) = A(L) \bar{\zeta}_{0}(L)
\]  
(A.39)
is obtained. Subtracting \(\bar{\zeta}_{0}(L-1)\) from \(\bar{\zeta}_{0}(L)\), we obtain
\[
\begin{align*}
&\bar{\zeta}_{0}(L) = \bar{\zeta}_{0}(L-1) \\
&\quad + \bar{f}_{0}(L) \overline{A}(L) \bar{g}_{0}(L) \\
&\quad - \tilde{H} A(L) \bar{\tau}_{0}(L-1),
\end{align*}
\]  
(A.40)
\[
\begin{align*}
&\bar{\zeta}_{0}(0) = 0.
\end{align*}
\]
\[\text{(Q.E.D.)}\]

**Appendix B: Proof of Theorem 3**

By substituting (A.11) into (A.10) and introducing
\[
G(k) = \alpha(k)J(k),
\]  
(B.1)
(42) and (44) are clear. By substituting (A.18) into (A.17) and introducing
\[ G_0(k) = A(k)^T_0(k), \quad (B.2) \]
(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
\[ S_0(k) = A(k)^T_0(k) A_0(k), \quad (B.4) \]
(47) is obtained. From (A-19) and (B.4), by using \( A(k) = \Phi^k \), (48) is obtained. By substituting (A.8) into (B.3) and using (B.1) with \( \alpha(k) = \Phi^k \) and \( A(k) = \Phi^k \), (49) is obtained. By substituting (A.15) into (B.4) and using (B.2) with \( A(k) = \Phi^k \), (50) is obtained.

By substituting (A.33) into (A.31) and introducing
\[ \overline{G}(L) = \alpha(L) \overline{f}(L), \quad (B.5) \]
(51) is obtained. By substituting (A.29) into (B.5) and introducing
\[ \overline{S}(L) = \alpha(L) \overline{v}(L) \overline{A}^T(L), \quad (B.6) \]
(52) is obtained. By substituting (A.40) into (A.39) and introducing
\[ \overline{G}_0(L) = A(L) \overline{f}_0(L), \quad (B.7) \]
(53) is obtained. By substituting (A.36) into (B.7) and introducing
\[ \overline{S}_0(L) = A(L) \overline{v}_0(L) \overline{A}^T(L), \quad (B.8) \]
(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:


