Numerical Robustness of Predictor-Based Least Squares Algorithm

Kazushi IKEDA Youhua WANG Kenji NAKAYAMA
Dept. Elec. & Comp. Eng., Kanazawa Univ.
2-40-20 Kodatsuno, Kanazawa 920 Japan
kazushi@t.kanazawa-u.ac.jp

Abstract: The numerical property of the predictor-based least squares (PLS) algorithm which provides the same least squares solutions as the RLS algorithm is studied. This paper theoretically proves that the backward PLS (BPLS) algorithm is statistically stable. First, the eigenvalues of the transition matrix is verified to be within or on the unit circle of the complex plain regardless of the input signals. Then, the expectation of the transition matrix is shown to have the eigenvalues within the unit circle. This means that the BPLS algorithm is statistically stable.

1. Introduction

In solving the least squares problem for transversal adaptive filters, the recursive least squares (RLS) algorithm is well known. However, it is reported that divergence phenomena may occur when the arithmetic precision is not enough or the input signal is ill-conditioned[5, 6, 1].

Another approach for solving the least squares problem is to use the fast least squares (FLS) algorithm. The principle of the algorithm is different from that of the RLS algorithm in that the relation of the forward and backward predictors and the gain vector are exploited, which result in a fast convergence rate with much less computation. However, the numerical instability of the FLS algorithm is so serious that they cannot be continuously used in real applications, especially under finite-precision implementation[2, 4].

Although the algorithm we called the predictor-based least squares (PLS) algorithm is easily derived from the FLS algorithm, it is reported to be much more stable than the RLS, FLS, or other stabilized versions[7]. Ref.[8] show that three main instability sources encountered in both the RLS and the FLS algorithms, including the unstable behavior of the conversion factor, the loss of symmetry, and the loss of positive definiteness of the inverse correlation matrix, do not exist in the PLS algorithm. Nevertheless, the stability has not been guaranteed theoretically yet.

This paper gives the theoretical proof of the stability of the backward PLS (BPLS) algorithm. First, it is shown that the eigenvalues of the transition matrix are necessarily within or on the unit circle of the complex plain regardless of the input signal or the tap-weight vectors of the predictors. The property, however, cannot guarantee the stability of the BPLS algorithm because this system is time-varying. Then, we evaluate the expectation of the transition matrix and show that its eigenvalues are within the unit circle. This means that the BPLS algorithm is statistically stable.

2. Backward PLS Algorithm

In the FLS algorithm[3], the gain vector $\boldsymbol{k}_M(n)$ is obtained using two order-update equations,

$$\boldsymbol{k}_{M+1}(n) = \begin{bmatrix} 0 \\ \boldsymbol{k}_M(n-1) \end{bmatrix} + \frac{f_{M+1}(n)}{F_{M+1}(n)} \boldsymbol{a}_{M+1}(n), \quad (1)$$

$$\boldsymbol{k}_{M+1}(n) = \begin{bmatrix} \boldsymbol{k}_M(n) \\ 0 \end{bmatrix} + \frac{b_{M+1}(n)}{B_{M+1}(n)} \boldsymbol{c}_{M+1}(n). \quad (2)$$

In deriving $\boldsymbol{k}_M(n)$ from $\boldsymbol{k}_M(n-1)$, $\boldsymbol{k}_{M+1}(n)$ is used therefore reversely order-updated. This is said that one of the causes of instability of the FLS algorithm[8]. In order to avoid the above, the forward (backward, resp.) PLS algorithm uses Eq. (1) (Eq. (2)) to get $\boldsymbol{k}_M(n)$, which leads the necessity of the gain vectors $\boldsymbol{k}_m(n)$ and the predictors $\boldsymbol{a}_m(n)$ ($\boldsymbol{c}_m(n)$) for all $m = 1, \ldots, M$ and it makes the computational load order of M^2 although that of the FLS algorithm is order of M[7].

Since the two are very similar, only the backward PLS (BPLS) algorithm is studied in this paper. For convenience of analysis, we write the BPLS algorithm below:

$$\psi_m(n) = \boldsymbol{c}_m(n-1)^T \boldsymbol{u}_m(n), \quad (3)$$

$$B_m(n) = \lambda B_m(n-1) + \gamma_m(n)\psi_m(n)^2, \quad (4)$$

$$\boldsymbol{c}_m(n) = \boldsymbol{c}_m(n-1) - \psi_m(n) \begin{bmatrix} \boldsymbol{k}_{m-1}(n) \\ 0 \end{bmatrix}, \quad (5)$$

$$\gamma_{m+1}(n) = \frac{\lambda B_m(n-1)}{B_m(n)} \gamma_m(n), \quad (6)$$

$$\boldsymbol{k}_m(n) = \begin{bmatrix} \boldsymbol{k}_{m-1}(n) \\ 0 \end{bmatrix} + \frac{\gamma_m(n)\psi_m(n)}{B_m(n)} \boldsymbol{c}_m(n), \quad (7)$$

$$\alpha(n) = d(n) - \boldsymbol{w}_M(n-1)^T \boldsymbol{u}_M(n), \quad (8)$$

$$\boldsymbol{w}_M(n) = \boldsymbol{w}_M(n-1) + \boldsymbol{k}_M(n)\alpha(n), \quad (9)$$

$$m = 1, \ldots, M,$$

where $\psi_m(n)$ is the backward a priori prediction error, $B_m(n)$ is the minimum power of the backward prediction error, $\gamma_m(n)$ is the conversion factor, $\boldsymbol{k}_m(n)$ is the gain vector, $\boldsymbol{c}_m(n)$ is the tap-weight vector of the backward predictor, $\alpha(n)$ is the a priori estimation error, $\boldsymbol{u}_m(n)$ is the tap-input vector, $d(n)$ is the desired signal, and $\boldsymbol{w}_M(n)$ is the tap-weight vector of the adaptive filter. The backward a posteriori prediction error $b_m(n)$ in Eq. (2) is equivalent to $\gamma_m(n)\psi_m(n)$ in the above definition.

To initialize the BPLS algorithm at time $n = 0$, set $\boldsymbol{c}_m(0) = [\boldsymbol{0}_{m-1}^T, 1]^T$, $B_m(0) = \delta$, $\boldsymbol{k}_m(0) = \boldsymbol{0}_m$, $\gamma_m(0) = 1$ for $m = 1, 2, \ldots, M$, where δ is a small positive constant and $\boldsymbol{0}_m$ is the $1 \times m$ vector all of whose elements are zero. And at each iteration $n \geq 1$, generate the first-order variables as $\gamma_1(n) = 1$ and $\boldsymbol{k}_0(n)$ is null. Then, all of the variables are derived with Eqs. (3)-(9) when the input signal $u(i), i = 1, \ldots$ are given.

3. Eigenvalues of Transition Matrix

In this section, it is shown that the eigenvalues of the transition matrix are between 0 and 1 or equal to 1 regardless of the input signal, and this property remains even if the tap-weight vectors $\boldsymbol{c}_m(n)$ of the backward predictors include numerical errors.

First, we show the following lemma:

Lemma 1

$$\forall m, n, 0 < \boldsymbol{u}_m(n)^T \boldsymbol{k}_m(n) < 1$$

and the property holds even when $\boldsymbol{c}_m(n)$ *includes errors.*

Substituting Eq. (5) to Eq. (7),

$$\boldsymbol{k}_m(n) = \left(1 - \frac{\gamma_m(n)\psi_m(n)^2}{B_m(n)}\right) \begin{bmatrix} \boldsymbol{k}_{m-1}(n) \\ 0 \end{bmatrix} + \frac{\gamma_m(n)\psi_m(n)}{B_m(n)} \boldsymbol{c}_m(n-1) \quad (10)$$

is derived. Therefore,

$$1 - \boldsymbol{u}_m(n)^T \boldsymbol{k}_m(n) = a_m(n)\left(1 - \boldsymbol{u}_{m-1}(n)^T \boldsymbol{k}_{m-1}(n)\right) \quad (11)$$

where

$$a_m(n) = \left(1 - \frac{\gamma_m(n)\psi_m(n)^2}{B_m(n)}\right).$$

Eq. (11) means that $0 < \boldsymbol{u}_m(n)^T \boldsymbol{k}_m(n) < 1$ when $0 < a_m(n) < 1$ and $0 < \boldsymbol{u}_{m-1}(n)^T \boldsymbol{k}_{m-1}(n) < 1$, which are inductively shown to be satisfied from Eq. (4) and

$$0 < \boldsymbol{u}_1(n)^T \boldsymbol{k}_1(n) = \frac{\gamma_1(n)\psi_1(n)^2}{B_1(n)} < 1.$$

Since the property of $\boldsymbol{c}_m(n)$ is not used at all in the derivation above,

$$\forall m, n, 0 < \boldsymbol{u}_m(n)^T \boldsymbol{k}_m(n) < 1$$

holds even when $\boldsymbol{c}_m(n)$ has numerical errors.

Next, we define the transition matrix and show that all of its eigenvalues are between 0 and 1 or equal to 1. Let $\left(\boldsymbol{k}_m(n)^T, \boldsymbol{0}_{j-m}^T\right)^T$ and $\left(\boldsymbol{c}_m(n)^T, \boldsymbol{0}_{j-m}^T\right)^T$ be denoted by $\boldsymbol{k}_m^j(n)$ and $\boldsymbol{c}_m^j(n)$, respectively. When $j < m$, $\boldsymbol{c}_m^j(n)$ is defined as the vector whose jth element is equal to that of $\boldsymbol{c}_m(n)$. Then, Eq. (5) and Eq. (7) are rewritten as

$$\boldsymbol{c}_m^m(n) = \boldsymbol{c}_m^m(n-1) - \psi_m(n)\boldsymbol{k}_{m-1}^m(n), \quad (12)$$

$$\boldsymbol{k}_m^m(n) = \boldsymbol{k}_{m-1}^m(n) + \frac{\gamma_m(n)\psi_m(n)}{B_m(n)} \boldsymbol{c}_m^m(n). \quad (13)$$

Substituting Eq. (3) to Eq. (12),

$$\boldsymbol{c}_{m+1}(n) = (E_{m+1} - \boldsymbol{k}_m^{m+1}(n)\boldsymbol{u}_{m+1}(n)^T)\boldsymbol{c}_{m+1}(n-1) \quad (14)$$

is derived where E_{m+1} is $(m+1) \times (m+1)$ identity matrix. Because the $(m+1)$th element of $\boldsymbol{c}_{m+1}(n)$ is constantly unity, the essential transition formula of $\boldsymbol{c}_{m+1}(n)$ is

$$\begin{aligned} \boldsymbol{c}_{m+1}^m(n) &= (E_m - \boldsymbol{k}_m(n)\boldsymbol{u}_{m+1}(n)^T)\boldsymbol{c}_{m+1}(n-1) \\ &= (E_m - \boldsymbol{k}_m(n)\boldsymbol{u}_m(n)^T)\boldsymbol{c}_{m+1}^m(n-1) \\ &\quad -u(n-m)\boldsymbol{k}_m(n) \end{aligned} \quad (15)$$

Let the optimal predictor be denoted by $\boldsymbol{c}_{m+1}(\infty)$, then the above is rewritten as

$$\begin{aligned} &\boldsymbol{c}_{m+1}^m(n) - \boldsymbol{c}_{m+1}^m(\infty) \\ =\ &(E_m - \boldsymbol{k}_m(n)\boldsymbol{u}_m(n)^T) \\ &(\boldsymbol{c}_{m+1}^m(n-1) - \boldsymbol{c}_{m+1}^m(\infty)) \\ &-\boldsymbol{k}_m(n)\left(\boldsymbol{u}_{m+1}(n)^T \boldsymbol{c}_{m+1}(\infty)\right). \end{aligned} \quad (16)$$

Since $\boldsymbol{c}_{m+1}(\infty)$ is the optimal backward predictor, the last term statistically vanishes. In the same way, the transition formula of $\boldsymbol{w}_M(n)$ is also written as

$$\begin{aligned} &\boldsymbol{w}_M(n) - \boldsymbol{w}_M(\infty) \\ =\ &(E_m - \boldsymbol{k}_M(n)\boldsymbol{u}_M(n)^T) \\ &(\boldsymbol{w}_M(n-1) - \boldsymbol{w}_M(\infty)) \\ &+\boldsymbol{k}_M(n)\left(\boldsymbol{u}_M(n)^T \boldsymbol{w}_M(\infty) - d(n)\right). \end{aligned} \quad (17)$$

Then, let us define the transition matrix as $E_m - \boldsymbol{k}_m(n)\boldsymbol{u}_m(n)^T$ and consider the eigenvalues of the transition matrix in the following.

See $\boldsymbol{u}_m(n)$ as a vector in $\boldsymbol{R}^m$ and consider the orthogonal complement of $\boldsymbol{u}_m(n)$, whose $m-1$ basis vectors are denoted by $\boldsymbol{v}_1(n), \ldots, \boldsymbol{v}_{m-1}(n)$. Since $\boldsymbol{u}_m(n)^T \boldsymbol{v}_i(n) = 0$,

$$\left(E_m - \boldsymbol{k}_m(n)\boldsymbol{u}_m(n)^T\right)\boldsymbol{v}_i(n) = \boldsymbol{v}_i(n) \quad (18)$$

is satisfied for any $\boldsymbol{v}_i(n)$, which means that $\boldsymbol{v}_i(n)$ is an eigenvector of the matrix $E_m - \boldsymbol{k}_m(n)\boldsymbol{u}_m(n)^T$ and its eigenvalue is unity. And since

$$\begin{aligned} &\left(E_m - \boldsymbol{k}_m(n)\boldsymbol{u}_m(n)^T\right)\boldsymbol{k}_m(n) \\ =\ &\left(1 - \boldsymbol{u}_m(n)^T \boldsymbol{k}_m(n)\right)\boldsymbol{k}_m(n), \end{aligned} \quad (19)$$

$\boldsymbol{k}_m(n)$ is an eigenvector of the matrix $E_m - \boldsymbol{k}_m(n)\boldsymbol{u}_m(n)^T$ and its eigenvalue is $1 - \boldsymbol{u}_m(n)^T\boldsymbol{k}_m(n)$. $\boldsymbol{k}_m(n)$ and $\boldsymbol{v}_i(n), i = 1, \ldots, m-1$ are linearly independent because $\boldsymbol{u}_m(n)^T\boldsymbol{k}_m(n) \neq 0$, which concludes that the eigenvalues of $E_m - \boldsymbol{k}_m(n)\boldsymbol{u}_m(n)^T$ are only 1 and $1 - \boldsymbol{u}_m(n)^T\boldsymbol{k}_m(n)$. From Lemma 1, the following theorem is derived:

Theorem 1 *For any tap-input vector $\boldsymbol{u}_m(n)$, the eigenvalues of $E_m - \boldsymbol{k}_m(n)\boldsymbol{u}_m(n)^T$ are laid within or on the unit circle.*

Since only the property that $0 < \boldsymbol{u}_m(n)\boldsymbol{k}_m(n) < 1$ is used in the derivation above, the theorem holds even when $\boldsymbol{c}_m(n)$ has numerical errors. This means that the BPLS algorithm is robust against numerical errors.

4. Expectation of Transition Matrix

Because the system described by Eq. (16) is time-varying, its stability cannot be guaranteed even though the eigenvalues of the transition matrix are within the unit circle. In other words, the stability of the BPLS algorithm cannot be derived from Theorem 1. In this section, we evaluate the expectation of the transition matrix and show that its eigenvalues are laid within the unit circle. This means that the BPLS algorithm is statistically stable, where we assume that the tap-input vector $\boldsymbol{u}_m(n)$ obeys a colored Gaussian distribution.

In case of $m = 1$, $\boldsymbol{c}_1(n)$ is constantly equal to 1 and then stable. So, in order to prove the stability of $\boldsymbol{c}_m(n)$ for any m, it is enough to show that $\boldsymbol{c}_{m+1}(n)$ is stable when so are $\boldsymbol{c}_i(n), i = 1, \ldots, m$.

Let $C_m(n)$, $D_m(n)$, and $\Psi_m(n)$ be defined as

$$C_m(n) = (\boldsymbol{c}_1^m(n), \ldots, \boldsymbol{c}_m^m(n)), \quad (20)$$

$$D_m(n) = \operatorname{diag}\{d_1(n), \ldots, d_m(n)\}, \quad (21)$$

$$d_i(n) = \frac{\gamma_{m+1}(n)}{\gamma_{i+1}(n)} \frac{\gamma_i(n)}{B_i(n)}, \quad (22)$$

$$\begin{aligned} \Psi_m(n) &= (\psi_1(n), \ldots, \psi_m(n))^T \\ &= C_m(n-1)^T\boldsymbol{u}_m(n), \quad (23) \end{aligned}$$

respectively. Using Eq. (10) recursively, it is derived that

$$\begin{aligned} & \boldsymbol{k}_m(n) \\ =\ & \frac{\lambda B_m(n-1)}{B_m(n)}\boldsymbol{k}_{m-1}^m(n) \\ & + \frac{\gamma_m(n)}{B_m(n)}\boldsymbol{c}_m(n-1)\boldsymbol{c}_m(n-1)^T\boldsymbol{u}_m(n) \\ =\ & \sum_{i=1}^{m}\left(\prod_{j=i+1}^{m}\frac{\lambda B_j(n-1)}{B_j(n)}\right)\frac{\gamma_i(n)}{B_i(n)} \\ & \cdot\boldsymbol{c}_i(n-1)\boldsymbol{c}_i(n-1)^T\boldsymbol{u}_m(n) \\ =\ & \sum_{i=1}^{m}\frac{\gamma_{m+1}(n)}{\gamma_{i+1}(n)}\frac{\gamma_i(n)}{B_i(n)}\boldsymbol{c}_i(n-1)\boldsymbol{c}_i(n-1)^T\boldsymbol{u}_m(n) \\ =\ & C_m(n-1)D_m(n)C_m(n-1)^T\boldsymbol{u}_m(n). \quad (24) \end{aligned}$$

Therefore, the transition matrix $E_m - \boldsymbol{k}_m(n)\boldsymbol{u}_m(n)^T$ of $\boldsymbol{c}_{m+1}^m(n)$ is written as

$$\begin{aligned} & E_m - \boldsymbol{k}_m(n)\boldsymbol{u}_m(n)^T \\ =\ & E_m - C_m(n-1)D_m(n)C_m(n-1)^T \\ & \cdot\boldsymbol{u}_m(n)\boldsymbol{u}_m(n)^T \\ =\ & C_m(n-1)(E_m - \\ & D_m(n)\Psi_m(n)\Psi_m(n)^T)C_m(n-1)^{-1}. \end{aligned}$$

Then, the eigenvalues of the transition matrix coincide with those of $E_m - D_m(n)\Psi_m(n)\Psi_m(n)^T$. From the assumption that $\boldsymbol{c}_i(n), i = 1, \ldots, m$, and $C_m(n)$ are stable, we only need to consider the eigenvalues of $E_m - D_m(n)\Psi_m(n)\Psi_m(n)^T$. Since $\boldsymbol{c}_i(n)$ is the tap-weight vector of the backward predictor and $\psi_i(n)$ means the backward a priori prediction error by their definitions, when $\boldsymbol{c}_i(n)$ converges, $\mathrm{Ex}\left[\psi_i(n)^2\right]$ approaches to its minimal value, and its derivatives by the jth element $(\boldsymbol{c}_i(n-1))_j$ of $\boldsymbol{c}_i(n-1)$ for $j = 1, \ldots, i-1$ become null, that is,

$$\frac{\partial \mathrm{Ex}\left[\psi_i(n)^2\right]}{\partial(\boldsymbol{c}_i(n-1))_j} = \mathrm{Ex}\left[\psi_i(n)u(n-j)\right] = 0.$$

This means that $\psi_i(n)$ is statistically orthogonal to $u(n-j), j = 1, \ldots, i-1$, and then $\psi_j(n), j = 1, \ldots, i-1$ because $\psi_j(n)$ is a linear combination of $u(n), \ldots, u(n-j)$. Therefore,

$$\mathrm{Ex}\left[\psi_i(n)\psi_j(n)\right] = 0 \quad (25)$$

is satisfied when $i \neq j$. Eq. (25) means that $\psi_i(n), i = 1, \ldots, m$ are statistically independent since $\Psi_m(n)$ obeys a Gaussian distribution by the assumption that $\boldsymbol{u}_m(n)$ is Gaussian.

Because the i,j element of $D_m(n)\Psi_m(n)\Psi_m(n)^T$ is

$$\frac{\gamma_{m+1}(n)}{\gamma_{i+1}(n)}\frac{\gamma_i(n)\psi_i(n)\psi_j(n)}{\lambda B_i(n-1) + \gamma_i(n)\psi_i(n)^2},$$

its expectation is 0 when $i \neq j$ from the independence and otherwise

$$\mathrm{Ex}\left[\frac{\gamma_{m+1}(n)}{\gamma_{i+1}(n)}\frac{\gamma_i(n)\psi_i(n)^2}{\lambda B_i(n-1) + \gamma_i(n)\psi_i(n)^2}\right]$$

which is between 0 and 1 because so are both fractions So, it has been shown that all of the eigenvalues of the expectation of $E_m - D_m(n)\Psi_m(n)\Psi_m(n)^T$ exist between 0 and 1, and so do the eigenvalues of the expectation of

$$C_m(n-1)(E_m - D_m(n)\Psi_m(n)\Psi_m(n)^T)C_m(n-1)^{-1} \quad (26)$$

about $\Psi_m(n)$. That means that $\boldsymbol{c}_{m+1}^m(n)$ is statistically stable, and so is $\boldsymbol{c}_{m+1}(n) = \left(\boldsymbol{c}_{m+1}^m(n)^T, 1\right)^T$.

When $\psi_i(n), i = 1, \ldots, m$ are not independent because of the error of $C_m(n-1)$, the expectation of $E_m - D_m(n)\Psi_m(n)\Psi_m(n)^T$ does not become a diagonal matrix. However, if the error is small (for example, order of ε), the eigenvalues of the expectation also move a little (order of ε) and still exist within the unit circle, $c_{m+1}(n)$ is still statistically stable. Then, the stability of $c_m(n)$ for any m has been inductively proven.

5. Simulation Results

Computer simulations are done to confirm that the eigenvalues of the transition matrix are within the unit circle.

A backward predictor with 10 taps is employed for the simulation. Each of the tap-weight vectors of the backward predictors $c_m(n)$ and the gain vectors $k_m(n)$ consists of an 8-bit exponent and a 3-bit mantissa. The input $u(n)$ is made by an AR model $(1, a_1, a_2))$ of white Gaussian noise $N(0,1)$ where $a_1 = -2r_x \cos\theta$, $a_2 = r_x^2$, $r_x = 0.82$, and $\theta = \pi/4$. The initial parameter $\delta = 10$ which is a little large to make the condition worse and the forgetting factor $\lambda = 0.95$ are used, and the eigenvalues of the ensemble average (50 samples) of the transition matrices at time $n = 15$ (early stage) and $n = 50$ (convergence stage) are calculated.

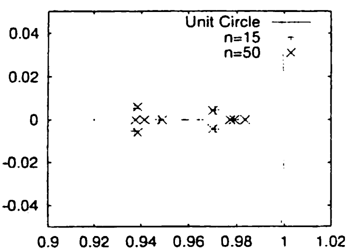

Figure 1: Eigenvalues at $n = 15$ and $n = 50$

Fig. (1) clearly demonstrates that the eigenvalues are within the unit circle in both cases even when the predictors and the gain vectors do not have enough precision and supports the theoretical result given above.

6. Conclusion

The stability of the BPLS algorithm has been proven in this paper. First, the eigenvalues of the transition matrix which represents the essence of the BPLS algorithm are shown to be between 0 and 1 or equal to 1 for any tap-input vector $u_m(n)$ and any backward predictors $c_i(n), i = 1, \ldots, m$. This is one of the reasons that the BPLS algorithm is robust against numerical errors.

To show the stability of the BPLS algorithm more strictly, the expectation of the transition matrix is evaluated and shown to have eigenvalues between 0 and 1. This means that the BPLS algorithm is statistically stable.

The computer simulation results also show that the transition matrix has eigenvalues within the unit circle in average which guarantees the stability of the BPLS algorithm even in short-precision cases. Because the superiority of the BPLS algorithm over the RLS algorithm is theoretically supported, too, it is very promising that the use of the RLS algorithm and its square root versions may be replaced by the BPLS algorithm.

References

[1] G. E. Bottomley and S. T. Alexander, "A theoretical basis for the divergence of conventional recursive least squares filters", *Proc. ICASSP*, pp. 908–911, 1989.

[2] J. M. Cioffi, "Limited-precision effects in adaptive filtering", *IEEE Trans. CAS*, vol. 34, pp. 821–833, 1987.

[3] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, 2nd edition, 1991.

[4] D. W. Lin, "On digital implementation of the fast Kalman algorithm", *IEEE Trans. ASSP*, vol. 32, pp. 998–1005, 1984.

[5] S. Ljung and L. Ljung, "Error propagation properties of recursive least squares adaptation algorithms", *Automatica*, vol. 21, pp. 157–167, 1985.

[6] M. H. Verhaegen, "Round-off error propagation in four generally-applicable, recursive, least-squares estimation schemes", *Automatica*, vol. 25, pp. 437–444, 1989.

[7] Y. Wang and K. Nakayama, "A stable least square algorithm based on predictors and its application to fast Newton transversal filters", *IEICE Trans. Fundamentals*, vol. E78-A, pp. 999–1003, 1995.

[8] Y. Wang and K. Nakayama, "A comparative study of instability in recursive least squares and predictor based least squares algorithms", *IEICE Trans. Fundamentals*, vol. E80-A, pp. in press, 1997.