An H_∞ Optimization and Its Fast Algorithm for Time-Variant System Identification

Kiyoshi Nishiyama

Abstract—In some estimation or identification techniques, a forgetting factor ρ has been used to improve the tracking performance for time-varying systems. However, the value of ρ has been typically determined empirically, without any evidence of optimality. In our previous work, this open problem is solved using the framework of H_∞ optimization. The resultant H_∞ filter enables the forgetting factor ρ to be optimized through a process noise that is determined by the filter Riccati equation. This paper seeks to further explain the previously derived H_∞ filter, giving an H_∞ interpretation of its tracking capability. Additionally, a fast algorithm of the H_∞ filter, called the fast H_∞ filter, is presented when the observation matrix has a shifting property. Finally, the effectiveness of the derived fast algorithm is illustrated for time-variant system identification using several computer simulations. Here, the fast H_∞ filter is shown to outperform the well known least-mean-square algorithm and the fast Kalman filter in convergence rate.

Index Terms—Fast algorithm, FKF, FTF, H_∞ filter, Kalman filter, LMS, RLS, system identification.

I. INTRODUCTION

SYSTEM identification is an experimental approach to the modeling of an unknown system or a plant and is of great importance in the fields of signal processing, communications, and control systems [1]–[3]. It involves the following steps: experimental planning, the selection of a model structure, parameter estimation, and model validation. In many cases, the unknown system is modeled as an adaptive filter whose parameters (tap weights) are adjusted such that its output will perfectly match that of the unknown system, except for observation noise, as illustrated in Fig. 1.

Let the available input signal at time k be denoted by the set of samples $u_k, u_{k-1}, \dots, u_{k-N+1}$, where N is the number of adjustable parameters in the model. This input signal is applied simultaneously to the plant and the model. Let their respective outputs be denoted by y_k and $\hat{z}_k$. The plant output y_k serves the purpose of a desired response for the adaptive filtering algorithm to adjust the model parameters. When a finite impulse response (FIR) filter model is employed for the plant, the model output is given by

$$\hat{z}_k = \sum_{i=0}^{N-1} \hat{h}_k(i) u_{k-i}, \quad k = 0, 1, 2, \dots \quad (1)$$

where $\{\hat{h}_k(i)\}_{i=0}^{N-1}$ are the adjustable model parameters. The difference between the plant output and the model output $e_k =$

Manuscript received October 7, 2002; revised June 7, 2003. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Paul D. Fiore.

The author is with the Department of Computer and Information Science, Faculty of Engineering, Iwate University, Ueda, Morioka, Japan (e-mail: nishiyama@cis.iwate-u.ac.jp).

Digital Object Identifier 10.1109/TSP.2004.826156

Fig. 1. System identification.

$y_k - \hat{z}_k$ is applied to an adaptive filtering algorithm. For the parameter estimation, the least-mean-square (LMS) algorithm is widely used due to its low computational cost. However, it still has the major limitations of a relatively slow rate of convergence and a sensitivity to the whiteness of inputs [3].

Efficient estimation algorithms have been developed in the past based mainly on the minimization of the L_2 -norm of the estimation error [4], [5]. In such an approach, the Kalman filter has played a central role among the minimum variance estimators. Indeed, the adaptive filtering algorithm with high tracking performance is successfully derived by applying the Kalman filter to the following state-space model:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{w}_k, \quad y_k = z_k + v_k = \mathbf{H}_k \mathbf{x}_k + v_k \quad (2)$$

where the state vector $\mathbf{x}_k$ is set to the impulse response of unknown system as $\mathbf{x}_k = [h_k(0), h_k(1), \dots, h_k(N-1)]^T$, the observation matrix $\mathbf{H}_k$ consists of the coming inputs as $\mathbf{H}_k = [u_k, u_{k-1}, \dots, u_{k-N+1}]$, and the system noise $\mathbf{w}_k$ is assumed to be a zero mean, white Gaussian process [3]. However, the standard Kalman filter required computations of at least $O(N^2)$ for each iteration. Fortunately, a fast algorithm with computational complexity of $O(N)$ per iteration, called the fast Kalman filter (FKF), has been developed for the N -dimensional state-space model such that $\mathbf{x}_{k+1} = \mathbf{x}_k$, $y_k = \mathbf{H}_k \mathbf{x}_k + v_k$ in which the observation matrix $\mathbf{H}_k$ has a shifting property, i.e., $\mathbf{H}_{k+1}$ is obtained from $\mathbf{H}_k$ by shifting its elements and adding a new one [6]. Some fast algorithms have also been derived for general time-invariant state-space models with system noise $\mathbf{w}_k$ [7]. Furthermore, over the past few years, the FKF algorithm has been extended to a class of structured time-variant state-space models $\mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{G}_k \mathbf{w}_k$, $y_k = \mathbf{H}_k \mathbf{x}_k + v_k$ such that there exist matrices $\{\Psi_k\}$ satisfying $\mathbf{H}_k = \mathbf{H}_{k+1} \Psi_k$, $\mathbf{F}_{k+1} \Psi_k = \Psi_{k+1} \mathbf{F}_k$, and $\mathbf{G}_{k+1} = \Psi_{k+1} \mathbf{G}_k$ [8]–[11].

Focusing on the performance, the standard FKF provides extremely fast convergences for identification of time-invariant FIR systems at a reasonable computational requirement, whose order is equal to that of the LMS algorithm. Nevertheless, its performance is still unsatisfactory because the standard FKF fails to track time-variant FIR systems. Under such circumstances, a fast algorithm with the computational complexity of $O(N)$, which is more suitable to time-variant system identification, has been developed based on the N -dimensional state equation such that $\mathbf{x}_{k+1} = \lambda^{-1/2}\mathbf{x}_k$. This algorithm is called the fast transversal filter (FTF) algorithm, which corresponds to a computationally reduced version of the recursive least square (RLS) algorithm exponentially weighted with λ for shift-structured models [9]–[11]. Here, λ is called a forgetting factor, emphasizing more recent input data and discounting very old input data. Additionally, the FTF can be quite compactly expressed as the square root Chandrasekhar filter [8]–[11]. However, the FTF algorithm, whose convergence is relatively susceptible to correlation of tap inputs, has required an empirical choice of λ for satisfactory tracking performance. Likewise, fast algorithms for the structured state-space models with system noise satisfying $\mathbf{H}_k = \mathbf{H}_{k+1}\Psi$ and $\mathbf{G}_{k+1} = \Psi\mathbf{G}_k$ will also involve a similar problem associated with the driving matrix $\mathbf{G}_k$ since it is rare in the tracking problems that $\mathbf{G}_k$ is given *a priori*.

Recently, a measure that differs from the L_2 -norm has been introduced into optimal estimation, providing an unifying approach for robust estimation [12], [13]. This measure is the H_∞ -norm (or the maximum energy gain) of the operation that relates the exogenous disturbances (the initial state and noises) to the estimation error. As a solution to the suboptimal H_∞ estimation problem that keeps the H_∞ -norm less than a prescribed positive value, an H_∞ filter has been derived for time-variant state-space models using a game theory approach [14], [15]. The H_∞ filtering problem has also been solved as a linear estimation in an indefinite-metric space, called the Krein space [16], [17], [19]. Furthermore, it has been proved that the LMS algorithm is H_∞ -optimal [18]. The H_∞ filters have attracted much attention in the field of robust estimation [12]–[19], since they attempt to optimally estimate the required combination of the states against the worst disturbances. However, their computational complexity becomes intractable as the size of the state vector grows large. To solve this problem, some fast array algorithms have been developed using J-unitary transformations [19]. Unfortunately, they have been applicable only for time-invariant state-space models.

In our previous work [20], an H_∞ optimization scheme was developed to derive a new fast filtering algorithm for time-variant system identification. The resultant fast algorithm possesses the capability of successfully tracking the time-varying system with a computational complexity of $O(N)$. The key idea was to incorporate a forgetting factor, which is defined as $\rho = 1 - \gamma_f^{-2}$, into the H_∞ optimization process through a system noise depending on the Riccati variables. Fortunately, the H_∞ optimization allows the minimization of the maximum energy gain and the optimal choice of the forgetting factor ρ to be simultaneously achieved by means of so-called γ -iteration.

In this paper, after the coloring of the process noise $\mathbf{w}_k$ in [20] was moved to a driving matrix $\mathbf{G}_k$, the meaning of the modification will be clarified from viewpoints of the H_∞ criterion. Furthermore, the forgetting factor ρ is redefined by a more general form

$\rho = 1 - \chi(\gamma_f)$, where $\chi(\gamma_f)$ is a monotonically decreasing function of γ_f such that $\chi(1) = 1$ and $\chi(\infty) = 0$. In such a setting, $\chi(\gamma_f)$ is a function to be designed by the user like a weight function in the H_∞ control. The optimal determination of $\chi(\gamma_f)$ will be left as a further problem to be solved in the future. In addition, a scalar existence condition for critically checking the validity of the fast H_∞ filter is given, which makes it easy to use the filter in practical situations. Finally, the superiority of the derived fast H_∞ filter over the fast Kalman filter and the LMS in terms of tracking performance is illustrated for identification of unknown FIR systems using several computer simulations.

This paper is organized as follows. Section II recalls the modified H_∞ filter for time-variant system identification and somewhat extends it. In addition, the relationship of the H_∞ filter with other adaptive filtering algorithms is discussed. Section III attempts to explain the mechanism of the tracking capability of the H_∞ filter from the viewpoint of the criterion. Section IV presents a fast algorithm of the modified H_∞ filter when the observation matrix $\mathbf{H}_k$ has a shifting property. Section V investigates some important points of our study using computer simulations. Finally, Section VI summarizes the results and discusses unanswered questions and possible future work.

II. MODIFIED H_∞ FILTER

A. Solution to the Modified H_∞ Filtering Problem

In some estimation or identification techniques, a forgetting factor ρ is used to improve the tracking performance for time-varying systems. However, the value of ρ has been typically determined empirically, with no evidence of optimality. Recently, in [20], we have found that this problem can be solved in the framework of H_∞ optimization by using a process noise $\mathbf{w}_k$ that is determined by the filter Riccati equation. Modifying the proof in [20] by shifting the coloring of the process noise $\mathbf{w}_k$ to a driving matrix $\mathbf{G}_k$ and replacing the forgetting factor ρ with $\rho = 1 - \chi(\gamma_f)$, we obtain the following theorem. Here, to simplify the presentation, we will confine our attention to the real-valued case.

Theorem 1: (Modified H_∞ Filter) For the following state-space model:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{G}_k \mathbf{w}_k, \quad \mathbf{w}_k, \mathbf{x}_k \in \mathcal{R}^N \quad (3)$$

$$y_k = \mathbf{H}_k \mathbf{x}_k + v_k, \quad y_k, v_k \in \mathcal{R} \quad (4)$$

$$z_k = \mathbf{H}_k \mathbf{x}_k, \quad z_k \in \mathcal{R}, \quad \mathbf{H}_k \in \mathcal{R}^{1 \times N}, \quad k=0, 1, \dots, L \quad (5)$$

one possible level- γ_f modified H_∞ filter to achieve

$$\sup_{\mathbf{x}_0, \{\mathbf{w}_i\}, \{v_i\}} \frac{\sum_{i=0}^k \|e_{f,i}\|^2}{\rho} \cdot \frac{1}{\|\mathbf{x}_0 - \hat{\mathbf{x}}_0\|_{\Sigma_0^{-1}}^2 + \sum_{i=0}^k \|\mathbf{w}_i\|^2 + \frac{\sum_{i=0}^k \|v_i\|^2}{\rho}} < \gamma_f^2 \quad (6)$$

is represented by

$$\hat{z}_k|_k = \mathbf{H}_k \hat{\mathbf{x}}_k|_k \quad (7)$$

$$\hat{\mathbf{x}}_{k+1}|_{k+1} = \hat{\mathbf{x}}_k|_k + \mathbf{K}_{s,k+1} (y_{k+1} - \mathbf{H}_{k+1} \hat{\mathbf{x}}_k|_k) \quad (8)$$

$$\mathbf{K}_{s,k+1} = \hat{\Sigma}_{k+1}|_k \mathbf{H}_{k+1}^T \cdot (\mathbf{H}_{k+1} \hat{\Sigma}_{k+1}|_k \mathbf{H}_{k+1}^T + \rho)^{-1} \quad (9)$$

and

$$\hat{\Sigma}_{k+1|k} = \frac{\left(\hat{\Sigma}_{k|k-1} - \hat{\Sigma}_{k|k-1} \begin{bmatrix} \mathbf{H}_k^T & \mathbf{H}_k^T \end{bmatrix} \cdot \mathbf{R}_{e,k}^{-1} \begin{bmatrix} \mathbf{H}_k \\ \mathbf{H}_k \end{bmatrix} \hat{\Sigma}_{k|k-1} \right)}{\rho} \quad (10)$$

where

$$\begin{aligned} e_{f,i} &= \tilde{z}_{i|i} - \mathbf{H}_i \mathbf{x}_i, \quad \hat{\mathbf{x}}_{k+1|k} = \hat{\mathbf{x}}_{k|k} \\ \mathbf{R}_{e,k} &= \mathbf{R}_k + \begin{bmatrix} \mathbf{H}_k \\ \mathbf{H}_k \end{bmatrix} \hat{\Sigma}_{k|k-1} \begin{bmatrix} \mathbf{H}_k^T & \mathbf{H}_k^T \end{bmatrix}, \quad \mathbf{R}_k = \begin{bmatrix} \rho & 0 \\ 0 & -\rho\gamma_f^2 \end{bmatrix} \\ \hat{\Sigma}_{0|-1} &= \Sigma_0, \quad 0 < \rho = 1 - \chi(\gamma_f) \leq 1, \quad \gamma_f > 1 \end{aligned} \quad (11)$$

and the driving matrix $\mathbf{G}_k$ is generated by

$$\mathbf{G}_k \mathbf{G}_k^T = \frac{\chi(\gamma_f)}{\rho} \left(\hat{\Sigma}_{k|k-1} - \hat{\Sigma}_{k|k-1} \begin{bmatrix} \mathbf{H}_k^T & \mathbf{H}_k^T \end{bmatrix} \cdot \mathbf{R}_{e,k}^{-1} \begin{bmatrix} \mathbf{H}_k \\ \mathbf{H}_k \end{bmatrix} \hat{\Sigma}_{k|k-1} \right). \quad (12)$$

In addition, the parameter γ_f should be adjusted to be sufficiently small for robustness, as long as the following existence condition is satisfied:

$$\hat{\Sigma}_{i|i-1}^{-1} + \frac{1 - \gamma_f^{-2}}{\rho} \mathbf{H}_i^T \mathbf{H}_i > 0, \quad i = 0, \dots, k. \quad (13)$$

Now, we shall discuss Theorem 1 in brief. Once γ_f is given, since the weight parameter $\rho = 1 - \chi(\gamma_f)$ in the criterion is known, the modified H_∞ filter and the existence condition can be derived using the standard H_∞ estimation scheme [19]. Indeed, if the existence condition of (13) is satisfied for a given γ_f , then the maximum energy gain, which is the left-hand side of (6), never exceeds γ_f^2 , and the level- γ_f robust estimate of z_k or $\mathbf{x}_k$ is obtained from (7)–(10). Similarly, the algorithm may be repeated for a smaller γ_f (larger $\mathbf{G}_k$) as long as the existence condition holds true. The resulting smallest γ_f provides the optimal driving matrix $\mathbf{G}_k$ for tracking the time-varying state-space model of interest so that the H_∞ filter achieves the highest robustness to time-variations in the state $\mathbf{x}_k$. This also provides the optimal ρ via (11).

B. Key Ideas

The key ideas to derive the modified H_∞ filter in Theorem 1 are

- 1) to define a forgetting factor ρ as a function of the upper limit γ_f formed by

$$\rho = 1 - \chi(\gamma_f) \quad (14)$$

where $\chi(\gamma_f)$ is a monotonically decreasing function of γ_f such that $\chi(1) = 1$ and $\chi(\infty) = 0$;

- 2) to introduce a driving matrix $\mathbf{G}_k$ generated by (12), which becomes known step by step as the Riccati recursion is carried out.

The process leading to these key ideas is described in the following. If a forgetting factor ρ is directly introduced into the H_∞ filtering algorithm, as was done for the Kalman filter, then the factor ρ remains in the algorithm as a tuning parameter to be determined by trial and error. Here, we present a forgetting factor ρ as a function of γ_f so that ρ can be governed by the H_∞ optimization. A candidate for the function was chosen as

$\rho = 1 - \chi(\gamma_f)$ with a monotonically decreasing function χ of γ_f , by taking into account that $0 < \rho \leq 1$ and $\rho = 1$ when $\gamma_f = \infty$, and then weighting $\|e_{f,i}\|$ and $\|v_i\|$ in the H_∞ criterion. However, weighting only the criterion does not suffice for the H_∞ optimization of ρ since ρ still disappears in the Riccati recursion. To overcome this problem, we introduced a driving matrix $\mathbf{G}_k$ so that the system noise has the same effect as the forgetting factor in the Riccati equation, and then, we added system noise term ($\mathbf{w}_k$) to the denominator of the H_∞ criterion (the maximum energy gain). This enabled the control of ρ from the H_∞ criterion. In such a setting, it follows that the resultant H_∞ filtering algorithm reduces to that of the Kalman filter without system noise as γ_f tends to infinity (see Section II-C).

C. Relationship with Other Algorithms

1) *Kalman Filter:* As seen in Theorem 1, the formula of the H_∞ filter is very similar to that of the Kalman filter. However, for the H_∞ filter to exist, the additional condition (13) must be satisfied. In addition, the indefinite matrix $\mathbf{R}_k = \begin{bmatrix} \rho & 0 \\ 0 & -\rho\gamma_f^2 \end{bmatrix}$ appears in the Riccati equation, whereas it does not appear in the Kalman filter.

To clarify the relationship between the H_∞ filter and Kalman filter, applying the matrix inversion lemma to $\mathbf{R}_{e,k}^{-1}$ and increasing γ_f to ∞ give

$$\begin{aligned} \mathbf{R}_{e,k}^{-1} &= \lim_{\gamma_f \rightarrow \infty} \left(\begin{bmatrix} \rho & 0 \\ 0 & -\rho\gamma_f^2 \end{bmatrix} + \begin{bmatrix} \mathbf{H}_k \\ \mathbf{H}_k \end{bmatrix} \hat{\Sigma}_{k|k-1} \begin{bmatrix} \mathbf{H}_k^T & \mathbf{H}_k^T \end{bmatrix} \right)^{-1} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} \mathbf{H}_k \\ 0 \end{bmatrix} \left(\mathbf{H}_k^T \mathbf{H}_k + \hat{\Sigma}_{k|k-1}^{-1} \right)^{-1} \begin{bmatrix} \mathbf{H}_k^T & 0 \end{bmatrix}. \end{aligned} \quad (15)$$

Substituting the last equation into (10) provides the modified H_∞ filter when $\gamma_f = \infty$:

$$\begin{aligned} \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_{s,k} (y_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1}) \\ \hat{\mathbf{x}}_{k+1|k} &= \hat{\mathbf{x}}_{k|k} \end{aligned} \quad (16)$$

$$\mathbf{K}_{s,k} = \hat{\Sigma}_{k|k-1} \mathbf{H}_k^T \left(\mathbf{H}_k \hat{\Sigma}_{k|k-1} \mathbf{H}_k^T + 1 \right)^{-1} \quad (17)$$

$$\begin{aligned} \hat{\Sigma}_{k|k} &= \hat{\Sigma}_{k|k-1} - \hat{\Sigma}_{k|k-1} \mathbf{H}_k^T \left(\mathbf{H}_k \hat{\Sigma}_{k|k-1} \mathbf{H}_k^T + 1 \right)^{-1} \\ &\quad \times \mathbf{H}_k \hat{\Sigma}_{k|k-1} \\ \hat{\Sigma}_{k+1|k} &= \hat{\Sigma}_{k|k}. \end{aligned} \quad (18)$$

This algorithm completely agrees with that of the Kalman filter when the system noise is absent and the variance of observation noise is 1. Note that although the H_∞ filter when $\gamma_f = \infty$ is identical to that of the Kalman filter, the meanings of optimality are different.

The asymptotical agreement suggests that the Kalman filter has quite a large H_∞ norm. In other words, the maximum energy gain of the Kalman filtering algorithm has no upper bound. This leads to the property that the Kalman filter must be more sensitively affected by variations in the initial state vector and observation noise.

2) *LMS Algorithm:* The optimal gain $\gamma_{f,op}^2$ for the modified H_∞ filtering problem has yet to be obtained since it is not guaranteed that $\gamma_{f,op} \geq 1$ when $\chi(\gamma_f)$ is a monotonically decreasing function of γ_f .

Thus, consider the time-invariant state-space model without system noise in Theorem 1. If the input signal $\{\mathbf{u}_i\}_{i=0}^{\infty}$ is exciting, and

$$0 < \mu < \inf_k \frac{1}{\mathbf{u}_k^T \mathbf{u}_k} \quad (19)$$

then the minimum H_{∞} norm of the transfer operator from the disturbances $\mu^{-1/2}(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|0})$ and $\{v_i\}_{i=0}^{\infty}$ to the predicted error $\{\hat{d}_{i|k-1} - \mathbf{u}_i^T \hat{\mathbf{x}}_{i|k-1}\}_{i=0}^{\infty}$ is

$$\gamma_{p,op} = 1. \quad (20)$$

This fact yields $\hat{\Sigma}_{k|k-1} = \mu \mathbf{I}$ for all k [18], [19].

On the other hand, the H_{∞} *a priori* filter for the stationary model is represented by

$$\hat{\mathbf{z}}_{k|k-1} = \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1} \quad (21)$$

$$\hat{\mathbf{x}}_{k+1|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_{a,k} (y_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1}) \quad (22)$$

$$\mathbf{K}_{a,k} = \tilde{\Sigma}_{k|k-1} \mathbf{H}_k^T \left(1 + \mathbf{H}_k \tilde{\Sigma}_{k|k-1} \mathbf{H}_k^T \right)^{-1} \quad (23)$$

whose existence condition is

$$\tilde{\Sigma}_{k|k-1}^{-1} = \hat{\Sigma}_{k|k-1}^{-1} - \gamma_p^{-2} \mathbf{H}_k^T \mathbf{H}_k > 0, \quad k = 0, \dots, N. \quad (24)$$

Thus, rewriting the H_{∞} *a priori* filter equation for $\mathbf{H}_k = \mathbf{u}_k^T$ and $\hat{\Sigma}_{k|k-1} = \mu \mathbf{I}$ yields the LMS algorithm

$$\begin{aligned} \hat{\mathbf{x}}_{k+1|k} &= \hat{\mathbf{x}}_{k|k-1} + \tilde{\Sigma}_{k|k-1} \mathbf{u}_k \left(1 + \mathbf{u}_k^T \tilde{\Sigma}_{k|k-1} \mathbf{u}_k \right)^{-1} \\ &\quad \times (y_k - \mathbf{u}_k^T \hat{\mathbf{x}}_{k|k-1}) \\ &= \hat{\mathbf{x}}_{k|k-1} + \left(\tilde{\Sigma}_{k|k-1}^{-1} + \mathbf{u}_k \mathbf{u}_k^T \right)^{-1} \mathbf{u}_k (y_k - \mathbf{u}_k^T \hat{\mathbf{x}}_{k|k-1}) \\ &= \hat{\mathbf{x}}_{k|k-1} + \mu \mathbf{u}_k (y_k - \mathbf{u}_k^T \hat{\mathbf{x}}_{k|k-1}). \end{aligned} \quad (25)$$

The above discussion indicates that the LMS algorithm is H_{∞} -optimal but is not a solution of the modified H_{∞} filtering problem, which is basically formulated for the time-varying state-space model driven by system noise. It should be noted here that when the input signal is not exciting, the LMS algorithm corresponding to $\gamma = 1$ will be suboptimal.

3) *RLS Algorithm*: Replacing $\mathbf{R}_{e,k}^{-1}$ with

$$\begin{bmatrix} \left(\rho + \mathbf{H}_k \hat{\Sigma}_{k|k-1} \mathbf{H}_k^T \right)^{-1} & 0 \\ 0 & 0 \end{bmatrix} \quad (26)$$

in (10), we readily find that the Riccati equation of the modified H_{∞} filter presented in Theorem 1 reduces to

$$\hat{\Sigma}_{k+1|k} = \frac{(\hat{\Sigma}_{k|k-1} - \mathbf{K}_{s,k} \mathbf{H}_k \hat{\Sigma}_{k|k-1})}{\rho} \quad (27)$$

which is identical to that of the RLS algorithm [3].

Comparing the inverse of $\mathbf{R}_{e,k}$ in (11) with (26) will give new insight to the relationship between the modified H_{∞} filter and the RLS.

III. MEANING OF THE MODIFIED H_{∞} CRITERION

A. H_{∞} Interpretation of the Tracking Capability

Suppose that the system state $\mathbf{x}_i$, which is fixed to $\mathbf{x}_0$, suddenly starts to change at a certain time i' ($< k$), achieves the

different state by next time ($i' + 1$), and remains constant after that time (i.e., $\mathbf{x}_i = \mathbf{x}_0$, $i \leq i'$, $\mathbf{x}_i = \mathbf{x}_{i'}$, $i > i'$) and that v_i is negligible. In this situation, the inequality of (6) reduces to

$$\sup_{\mathbf{x}_0, \mathbf{w}_{i'}} \frac{\sum_{i=0}^k \|e_{f,i}\|^2}{\|\mathbf{x}_0 - \hat{\mathbf{x}}_{0|0}\|_{\Sigma_0^{-1}}^2 + \|\mathbf{w}_{i'}\|^2} < \gamma_f^2 \quad (28)$$

because the system noise $\mathbf{w}_i$ arises only when the state changes ($\mathbf{x}_{i'+1} - \mathbf{x}_{i'} = \mathbf{G}_{i'} \mathbf{w}_{i'} \neq 0$). Such a change may be regarded as the worst case for robust estimation in comparison with gradual state transitions. The effect of the noise $\mathbf{w}_{i'}$ propagates through $\mathbf{G}_{i'}$ to the level- γ_f H_{∞} filtering algorithm, improving its tracking performance. If the state estimate cannot follow the changed state accurately enough, the maximum energy gain will immediately exceed γ_f^2 due to a rapid increase of the numerator of (28). In this case, γ_f must be reset to a larger value. Setting $\gamma_f = \infty$ yields the zero driving matrix $\mathbf{G}_k = \mathbf{0}$ so that the term of the system noise disappears in the Riccati recursion and does not affect the estimation of $\mathbf{x}_k$. The H_{∞} filtering algorithm in the limit coincides with that of the Kalman filter with $\mathbf{G}_k = \mathbf{0}$, which will completely fail to follow such a change of the state. This means that the maximum energy gain of the Kalman filter is not bounded, allowing an endless increase of the estimation error energy. Contrarily, if γ_f is chosen as small as the existence condition allows, the H_{∞} filter with the optimal $\mathbf{G}_k$ (or ρ) will be able to very quickly track the changed state. This is because the sum of the squared error $\|e_{f,i}\|^2$ must be subject to the inequality of (28), which prevents the maximum energy gain from growing over the small γ_f^2 . Hence, the problem of determining the smallest γ_f for robustness is equivalent to finding the optimal driving matrix $\mathbf{G}_k$ in a class of (12), which enables the H_{∞} filter to most quickly track the time-varying state. This fact will provide us an interpretation of its tracking capability from the viewpoint of the H_{∞} criterion.

B. Analysis of Driving Matrix

According to [19], the H_{∞} filter is equivalent to the Krein-space Kalman filter. Thus, assuming that $\mathbf{w}_k$ is a zero-mean, white noise with the covariance of the identity matrix $\mathbf{I}$ in the Krein space, we can obtain

$$\hat{\Sigma}_{k+1|k} - \hat{\Sigma}_{k|k} = \mathbf{G}_k \mathbf{G}_k^T \quad (29)$$

from the state equation of (3). On the other hand, from Theorem 1, the following simple relationship between $\mathbf{G}_k$ and $\hat{\Sigma}_{k+1|k}$ holds.

Corollary 1: The driving matrix $\mathbf{G}_k$ satisfies

$$\mathbf{G}_k \mathbf{G}_k^T = \chi(\gamma_f) \hat{\Sigma}_{k+1|k}. \quad (30)$$

Proof: Substituting (10) into (12), we readily see it. ■

Substituting (30) into (29) provides

$$\hat{\Sigma}_{k+1|k} - \hat{\Sigma}_{k|k} = \chi(\gamma_f) \hat{\Sigma}_{k+1|k}. \quad (31)$$

Similarly, we can calculate the error covariance matrix for the unforced dynamical model:

$$\mathbf{x}_{k+1} = \rho^{-\frac{1}{2}} \mathbf{x}_k \quad (32)$$

which holds the key to the formulation of a general framework for deriving the RLS family of adaptive filtering algorithm [3], [9]. The result yields

$$\hat{\Sigma}_{k+1|k} - \hat{\Sigma}_{k|k} = (1 - \rho)\hat{\Sigma}_{k+1|k} = \chi(\gamma_f)\hat{\Sigma}_{k+1|k}. \quad (33)$$

Comparing (33) with (31), we find that the driving matrix $\mathbf{G}_k$ in the H_∞ filter is chosen so that the Riccati update for the noise-driven model of (3) agrees with that for the unforced model of (32).

Such a constructed $\mathbf{G}_k$ has the capability to express any difference $\mathbf{x}_{k+1} - \mathbf{x}_k$ of the state vector with a linear combination $\mathbf{G}_k \mathbf{w}_k$ because the matrix $\hat{\Sigma}_{k+1|k}$ is nonsingular. This $\mathbf{G}_k$ generated by (12) is reasonable candidate driving matrix.

IV. FAST ALGORITHM FOR THE MODIFIED H_∞ FILTER

A. Shifting Property of Observation Matrices

Throughout this section, we assume that the $1 \times N$ observation matrix $\mathbf{H}_k = [H_k(1), H_k(2), \dots, H_k(N)]$ in the state-space model (3)–(5) has a shifting property such that $\mathbf{H}_{k+1}$ is obtained from $\mathbf{H}_k$ by shifting its elements and adding a new one, i.e., $\mathbf{H}_{k+1} = [u_{k+1}, H_k(1), \dots, H_k(N-1)]$, where u_{k+1} is a new outcome of the observation at time $k+1$. This is expressed in the matrix form by

$$\underline{\mathbf{H}}_k = \underline{\mathbf{H}}_{k+1} \underline{\Psi} \quad (34)$$

using a full-length observation matrix $\underline{\mathbf{H}}_k$ and a lower triangular shift matrix $\underline{\Psi}$, where

$$\underline{\mathbf{H}}_k = [H_k(1), H_{k-1}(1), \dots, H_0(1), \mathbf{0}_{1 \times (L-k)}]^T \in \mathcal{R}^{1 \times (L+1)}$$

$$\underline{\Psi} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 1 & 0 \end{bmatrix} \in \mathcal{R}^{(L+1) \times (L+1)} \quad (35)$$

in which $\mathbf{0}_{1 \times (L-k)}$ denotes the $1 \times (L-k)$ zero matrix, and $L(> N)$ is the length of observation data. This shift-structured model without system noise $\mathbf{w}_k$ is often used for echo cancellers and system identification [21], although it appears to be a very specialized model. It should be noted here that the specifically constructed driving matrix $\mathbf{G}_k$ does not satisfy $\underline{\mathbf{G}}_{k+1} = \underline{\Psi} \underline{\mathbf{G}}_k$, where $\underline{\mathbf{G}}_k = \text{diag}\{\mathbf{G}_k, \mathbf{0}\} \in \mathcal{R}^{(L+1) \times (L+1)}$. Surprisingly, even in such a case, the computational complexity required to compute the H_∞ filter gain $\mathbf{K}_{s,k}$ for each iteration can be reduced from $\mathcal{O}(N^2)$ to $\mathcal{O}(N)$.

B. Fast Algorithm

This filter gain $\mathbf{K}_{s,k}$ can be easily obtained in a less computationally demanding manner than in (9). This is shown in the summary of a fast algorithm [20] to implement the filter presented in Theorem 1 as follows.

H_∞ Fast Algorithm (Fast H_∞ Filter): When the observation matrix $\mathbf{H}_k$ has a shifting property, the modified H_∞ filter for the N -dimensional state-space model (3)–(5) with $\hat{\Sigma}_{0|-1} = \varepsilon_0 \mathbf{I}$, $\varepsilon_0 > 0$ can be recursively performed at the computational complexity of $\mathcal{O}(N)$ per iteration as follows.

Step 0) Take the initial conditions for the recursions as

$$\mathbf{K}_0 = \mathbf{0}_{N \times 2}, \quad \mathbf{A}_{-1} = \mathbf{0}_{N \times 1}, \quad S_{-1} = \frac{1}{\varepsilon_0}$$

$$\mathbf{D}_{-1} = \mathbf{0}_{N \times 1}, \quad \hat{\mathbf{x}}_{0|0} = \mathbf{0}_{N \times 1}$$

where ε_0 is set to a relatively large positive number, and $\mathbf{0}_{N \times M}$ denotes the $N \times M$ zero matrix.

Step 1) Determine $\mathbf{A}_k$ and S_k recursively as

$$\begin{aligned} \tilde{\mathbf{e}}_k &= \mathbf{c}_k + \mathbf{C}_k \mathbf{A}_{k-1} \in \mathcal{R}^{2 \times 1} \\ \mathbf{A}_k &= \mathbf{A}_{k-1} - \mathbf{K}_k \mathbf{W}_k \tilde{\mathbf{e}}_k \in \mathcal{R}^{N \times 1} \\ \mathbf{e}_k &= \mathbf{c}_k + \mathbf{C}_k \mathbf{A}_k \in \mathcal{R}^{2 \times 1} \\ S_k &= \rho S_{k-1} + \mathbf{e}_k^T \mathbf{W}_k \tilde{\mathbf{e}}_k \in \mathcal{R} \end{aligned}$$

where

$$\mathbf{C}_k = \begin{bmatrix} \mathbf{H}_k \\ \mathbf{H}_k \end{bmatrix} \in \mathcal{R}^{2 \times N}, \quad \mathbf{W}_k = \rho \mathbf{R}_k^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & -\gamma_f^{-2} \end{bmatrix} \in \mathcal{R}^{2 \times 2}$$

and $\mathbf{c}_k \in \mathcal{R}^{2 \times 1}$ is the first row of $\mathbf{C}_{k+1} = [\mathbf{c}_k, \dots, \mathbf{c}_{k-N+1}]$, assuming that $\mathbf{c}_{k-i} = \mathbf{0}_{2 \times 1}$ for $k-i < 0$. In addition, $0 < \rho = 1 - \chi(\gamma_f) \leq 1$, and $\gamma_f > 1$.

Step 2) Calculate $\check{\mathbf{K}}_k$ as

$$\check{\mathbf{K}}_k = \begin{bmatrix} S_k^{-1} \mathbf{e}_k^T \\ \mathbf{K}_k + \mathbf{A}_k S_k^{-1} \mathbf{e}_k^T \end{bmatrix} \in \mathcal{R}^{(N+1) \times 2}.$$

Step 3) Partition $\check{\mathbf{K}}_k$ as

$$\check{\mathbf{K}}_k = \begin{bmatrix} \mathbf{m}_k \\ \boldsymbol{\mu}_k \end{bmatrix}, \quad \mathbf{m}_k \in \mathcal{R}^{N \times 2}, \quad \boldsymbol{\mu}_k \in \mathcal{R}^{1 \times 2}.$$

Step 4) Determine the gain matrix $\mathbf{K}_{k+1}$ at next time step ($k+1$) using $\mathbf{m}_k$ and $\boldsymbol{\mu}_k$, and then, obtain the filter gain $\mathbf{K}_{s,k+1}$ from $\mathbf{K}_{k+1}$ as

$$\begin{aligned} \boldsymbol{\eta}_k &= \mathbf{c}_{k-N} + \mathbf{C}_{k+1} \mathbf{D}_{k-1} \in \mathcal{R}^{2 \times 1} \\ \mathbf{D}_k &= [\mathbf{D}_{k-1} - \mathbf{m}_k \mathbf{W}_k \boldsymbol{\eta}_k] [1 - \boldsymbol{\mu}_k \mathbf{W}_k \boldsymbol{\eta}_k]^{-1} \in \mathcal{R}^{N \times 1} \\ \mathbf{K}_{k+1} &= \mathbf{m}_k - \mathbf{D}_k \boldsymbol{\mu}_k \in \mathcal{R}^{N \times 2} \end{aligned}$$

$$\tilde{\mathbf{K}}_{k+1}(i) = \rho \mathbf{K}_{k+1}(i, 1)$$

$$g_{k+1} = \rho + \gamma_f^{-2} \mathbf{H}_{k+1} \tilde{\mathbf{K}}_{k+1}$$

$$\mathbf{K}_{s,k+1} = g_{k+1}^{-1} \tilde{\mathbf{K}}_{k+1} \in \mathcal{R}^{N \times 1}.$$

Step 5) Update the filter equation of the H_∞ filter as

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k|k} + \mathbf{K}_{s,k+1} (y_{k+1} - \mathbf{H}_{k+1} \hat{\mathbf{x}}_{k|k}) \in \mathcal{R}^{N \times 1}$$

using the filter gain $\mathbf{K}_{s,k+1}$.

Step 6) Increase time index k ($k \rightarrow k+1$), and return to Step 1.

The key point here is that each algorithm step has complexity of $\mathcal{O}(N)$. In addition, the fast algorithm when $\gamma_f = \infty$ agrees with that of the fast Kalman filter [20]. This means that the fast algorithm given in this section could be regarded as an extended version of the fast Kalman filter in the sense of H_∞ optimization.

C. Scalar Existence Condition of the H_∞ Filter

To exhibit the robust performance for the modified H_∞ filter when $\rho = 1 - \chi(\gamma_f)$, the parameter γ_f must be tuned as small

as possible, satisfying $\hat{\Sigma}_{i|i-1}^{-1} + ((1 - \gamma_f^{-2})/\rho)\mathbf{H}_i^T \mathbf{H}_i > 0$, $i = 0, \dots, k$. In general, however, it requires a considerable computational cost to check the positive-definite constraint of (13), especially for large N . Therefore, it may be more favorable in this case to use the alternative existence condition that the matrices $\mathbf{R}_k$ and $\mathbf{R}_{e,k}$ have the same inertia [16], [17], [19], which is the number of its positive, negative, and zero eigenvalues.

Fortunately, the inertia condition for the H_∞ filter can be reduced to a simpler form. The following Lemma (proved in [20]) gives an existence condition for a given γ_f at the computational requirement of $O(N)$ per iteration.

Lemma 1: (Scalar Existence Condition) The H_∞ filter exists for a given γ_f if and only if the following inequality holds:

$$-\varrho \hat{\Xi}_i + \rho \gamma_f^2 > 0, \quad i = 0, \dots, k \quad (36)$$

where

$$\varrho = 1 - \gamma_f^2, \quad \hat{\Xi}_i = \frac{\mathbf{H}_i \tilde{\mathbf{K}}_i}{1 - \frac{1 - \gamma_f^2}{\rho} \mathbf{H}_i \tilde{\mathbf{K}}_i} = \frac{\rho \mathbf{H}_i \mathbf{K}_{s,i}}{1 - \mathbf{H}_i \mathbf{K}_{s,i}}. \quad (37)$$

The existence condition can provide a lower bound of the forgetting factor in the sense of H_∞ optimization. In many cases, it is observed that the optimal γ_f lies in the range from one to forty. A binary search on γ_f , testing if Lemma 1 is satisfied, will determine a suboptimal γ_f to high precision in very small number of iterations. Hence, the computational complexity including the determination of γ_f will also become $O(N)$.

V. SIMULATION STUDY

The performance of the derived fast H_∞ filter is evaluated for FIR system identification, such as that used in echo cancellers. In this section, the forgetting factor is ordinarily chosen as $\rho = 1 - \gamma_f^{-2}$ for simplicity.

A. Verification of the Equivalence and the Complexity

As an unknown system, we employ an echo path whose time-invariant impulse response $\{h(i)\}$ consists of $\{0.0, 0.008, -0.012, 0.064, 0.013, -0.052, -0.007, 0.039, 0.011, 0.0, -0.002, -0.009, -0.016, -0.013, -0.001, 0.004, 0.015, 0.013, 0.007, 0.0, -0.001, -0.002, -0.001, 0.0\}$ for $i < 24$ and zero otherwise ($24 \leq i < N$). The observed echo is given by

$$y_k = \sum_{i=0}^{N-1} h(i)u_{k-i} + v_k, \quad k = 1, 2, \dots, L \quad (38)$$

where v_k is a stationary, white Gaussian noise with zero mean and standard deviation $\sigma_v = 1.0 \times 10^{-4}$, N denotes the length (tap number) of the adjustable impulse response, and L stands for the length of observation data. Note that the state-space model of the observed echo y_k becomes time-varying due to $\mathbf{H}_k = [u_k, u_{k-1}, \dots, u_{k-(N-1)}]$ with a shifting property such that $\mathbf{H}_{k+1} = [u_{k+1}, H_k(1), H_k(2), \dots, H_k(N-1)]$, provided that $\mathbf{x}_k = [h(0), h(1), \dots, h(N-1)]^T$. The received signal (tap inputs) $\{u_k\}$ is generated by the following autoregressive (AR) model:

$$u_k = \alpha_1 u_{k-1} + \alpha_2 u_{k-2} + w'_k \quad (39)$$

Fig. 2. Comparison between the H_∞ filter and the fast H_∞ filter for $\hat{x}_{48|48}$ when $\rho = 1 - \gamma_f^{-2}$ and $\gamma_f = 1.0 \times 10^5$. (a) Estimate of impulse response by the H_∞ filter. (b) Difference of the estimate between the H_∞ filter and the fast H_∞ filter.

TABLE I
NUMBER OF TAPS IN THE IMPULSE RESPONSE VERSUS THE COMPUTATION TIME

tap number	HF [s]	FHF [s]	FKF [s]	LMS [s]
24	1.37	1.95	1.18	0.46
48	2.77	2.92	1.84	0.72
96	8.56	4.76	3.13	1.23
192	32.5	8.61	5.78	2.29
384	126.6	16.3	11.2	4.41

where $\alpha_1 = 0.7$, $\alpha_2 = 0.1$, and w'_k is a stationary, white Gaussian noise with zero mean and standard deviation $\sigma_{w'} = 0.04$.

Figs. 2(a) and (b) show the estimate of the impulse response at $k = 48$ by the H_∞ filter and the difference of the estimate between the H_∞ filter and the fast H_∞ filter when $\gamma_f = 1.0 \times 10^5$, respectively, where $\hat{\mathbf{x}}_{1|0} = \mathbf{0}$, $\hat{\Sigma}_{1|0} = 20\mathbf{I}$, and the tap number N is set to 48. In addition, since the time-invariant case corresponds to the case when $\gamma_f = \infty$ in the H_∞ filter, γ_f is set to an extremely large number. From the simulation results, it is seen that the H_∞ and the fast H_∞ filters are almost equivalent in performance. Here, note that the results of the fast H_∞ and the fast Kalman filters become close enough to be indistinguishable when $\gamma_f = 1 \times 10^5$.

Table I lists the number N of tap weights versus the computation time on average over four trials, where the computation time stands for the CPU time consumed to compute $\hat{\mathbf{x}}_{k|k}$, $k = 1, \dots, 100$ which is observed using the **etime** command in MATLAB (Ver. 4.2) on EWS (sparc, 60 MHz, 32 MB). From these results, it is clear that the fast H_∞ filter (FHF) is superior in computational complexity to the conventional H_∞ filter (HF). Indeed, the complexity of the FHF is of $O(N)$ per iteration, which is equal to that of the FKF and the LMS algorithm.

Fig. 3. Comparison between the FHF and the LMS on convergence rate for different observation noise levels (a) when $\sigma_v = 10^{-4}$ and (b) when $\sigma_v = 10^{-3}$.

Fig. 4. Comparison of the FHF with the FTF and the RLS on sensitivity to correlation of tap inputs where $\gamma_f = 3.0$, $\rho = 1 - \gamma_f^{-2}$, $\sigma_v = 10^{-4}$, $\sigma_w = 0.04$, $\alpha_1 = 0.7$, and $\alpha_2 = 0.4$.

Fig. 3 gives a comparison between the FHF and the LMS algorithms on convergence rate for different observation noise levels, where the vertical axis denotes the squared norm $\|\mathbf{x}_k - \hat{\mathbf{x}}_k\|_k^2$. Additionally, the selection of $\chi(\gamma_f)$ affected the beginning FHF convergence rate. For example, when $\rho = 1 - \gamma_f^{-n}$, $n > 0$, a decrease of n tends to accelerate the convergence. Furthermore, to examine sensitivity to correlation of tap inputs, the FHF is compared with the FTF and the RLS algorithms for highly colored tap inputs when α_2 is set to 0.4, as shown in Fig. 4. Although the FTF is tuned so that it converges at about the same rate as the FHF until 100 iterations, it diverges very quickly after 140 iterations. When α_2 was reset to 0.1, both the FHF and the FTF soundly converged. From these results, it is seen that the FHF is more robust to variations in the correlation of tap inputs than the FTF and the RLS. This allows the FHF to use a smaller forgetting factor, providing faster convergence.

B. Tracking Performance

In the previous subsection, we implicitly assumed that the underlying system is time-invariant. However, the natural motivation for online parameter estimation (system identification) is that system parameters are time-varying. The derived fast H_∞ filter is more appropriate and efficient in such a time-varying

Fig. 5. Comparison of the FHF with $\rho = 1 - \gamma_f^{-2}$ and $\gamma_f = 2.0$, the FKF, and the LMS with $\mu = 5.0$. (a) Time-varying impulse responses to be tracked. (b) Estimates by the FHF. (c) Estimates by the FKF. (d) Estimates by the LMS.

Fig. 6. Dependency of tracking performance on parameters. (a) Tracking error of the FHF versus γ_f when $\rho = 1 - \gamma_f^{-2}$. (b) Tracking error of the LMS versus μ .

case. To emphasize the effectiveness of the fast H_∞ filter, we suppose that the impulse response of the echo path is suddenly changed at time $k = 81$ and at time $k = 201$, as shown in Fig. 5(a), which represents the target to be tracked. Here, the observation length $L = 300$ is large enough for the transient period to have ended.

Fig. 5(b), (c), and (d) show the tracking results of the changed impulse response by the fast H_∞ filter with $\gamma_f = 2.0$, the fast Kalman filter, and the LMS algorithm with a large step size $\mu = 5.0$, respectively. Here, the step size μ is experimentally chosen to make the convergence as fast as possible. Additionally, in the case when the tap input was replaced with speech signal or white Gaussian noise, similar performance was observed. Disappointingly, the tracking performance of the LMS algorithm was not improved enough, even in the white Gaussian case. From these results, it is obviously seen that the fast H_∞ filter successfully tracks the time-varying impulse response despite its sudden switch, whereas the fast Kalman filter fails to track it due to the stationary assumption, and the LMS algorithm also fails due to the slow convergence behavior.

Fig. 6(a) and (b) show the relationship between the tracking performance and the tuning parameters, where the tracking error e_{track} defined by

$$e_{\text{track}} = \sum_{k=1}^L \sqrt{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^T (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})} \\ = \sum_{k=1}^L \sqrt{\sum_{i=1}^N ((x_k(i) - \hat{x}_{k|k}(i))^2)} \quad (40)$$

is used as a measure of the performance. These results show that the values of γ_f and μ employed in Fig. 5 are reasonable. However, the optimal value of γ_f does not necessarily coincide with the value of γ_f at the minimum error. Indeed, in the case of Fig. 5(a), the optimal value of γ_f was 2.0, and the minimum-error value of γ_f was about 2.4. It should be noted here that the errors in transit periods considerably raise the bottom of the total tracking error of the fast H_∞ filter. In addition, it is observed that the value of γ_f at the minimum error tends to increase as N increases, whereas the value of μ tends to decrease as N increases.

VI. CONCLUSIONS

The mechanism that causes the modified H_∞ filter to quickly track time-varying systems has been clarified from the viewpoint of the H_∞ criterion. In this process, a class of forgetting factors, i.e., a set of $\rho = 1 - \chi(\gamma_f)$, has been found to construct the H_∞ filter. The framework of H_∞ optimization developed here has succeeded in theoretically determining the optimal value of the forgetting factor, as well as minimizing the maximum energy gain. This success led to the fast algorithm with faster tracking capability, called the fast H_∞ filter, whose computational complexity is of $O(N)$ per iteration. Also, a simple method to determine the existence of the fast H_∞ filter for a given γ_f was derived, which also requires $O(N)$ arithmetic operations per iteration. Furthermore, a simple bisection search algorithm for finding the suboptimal γ_f has been developed. Several computer simulations confirmed the superiority of the derived fast H_∞ filter over the fast Kalman filter and the LMS algorithm in terms of tracking performance for unknown time-varying systems, even when the step size of the LMS is experimentally optimized. In a comparison with the FTF and the RLS, the fast H_∞ filter exhibited more stable performance for highly colored tap inputs.

More detailed evaluation, the optimal determination of $\chi(\gamma_f)$, practical applications, and some extension of the fast H_∞ filter will be performed in the future. In particular, it is believed by the author that $\chi(\gamma_f)$ can be optimally determined by balancing the weighting of the H_∞ criterion with the driving of the system noise for the filter tracking performance.

ACKNOWLEDGMENT

The author would like to acknowledge the valuable suggestions and comments from Dr. P. D. Fiore and anonymous reviewers.

REFERENCES

- [1] L. Ljung, *System Identification: Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
- [2] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, First ed. Englewood Cliffs, NJ: Prentice-Hall, 1985.
- [3] S. Haykin, *Adaptive Filter Theory*, Third ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [4] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, First ed. New York: Academic, 1970.
- [5] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, First ed. Englewood Cliffs, NJ: Prentice-Hall, 1979.
- [6] L. Ljung, M. Morf, and D. Falconer, "Fast calculation of gain matrices for recursive estimation schemes," *Int. J. Contr.*, vol. 27, no. 1, pp. 1–19, 1978.
- [7] M. Morf, G. S. Sidhu, and T. Kailath, "Some new algorithms for recursive estimation in constant, linear, discrete-time systems," *IEEE Trans. Automat. Contr.*, vol. 19, pp. 315–323, 1974.
- [8] A. H. Sayed and T. Kailath, "Extended Chandrasekhar recursions," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 619–623, Mar. 1994.
- [9] —, "A state-space approach to adaptive RLS filtering," *IEEE Signal Processing Mag.*, pp. 18–60, 1994.
- [10] T. Kailath and A. H. Sayed, *Fast Reliable Algorithms for Matrices With Structure*, First ed. Philadelphia, PA: SIAM, 1999.
- [11] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, First ed. Englewood Cliffs, NJ: Prentice-Hall, 2000.
- [12] K. M. Nagpal and P. P. Khargonekar, "Filtering and smoothing in an H_∞ setting," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 152–166, Feb. 1991.
- [13] U. Shaked and Y. Theodor, " H_∞ -optimal estimation: a tutorial," in *Proc. IEEE 31st CDC*, 1992, pp. 2278–2286.
- [14] I. Yaesh and U. Shaked, "Nondefinite least squares and its relation to H_∞ -minimum error state estimation," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 1469–1472, Dec. 1991.
- [15] —, "Game theory approach to optimal linear state estimation and its relation to the minimum H_∞ -norm estimation," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 828–831, June 1992.
- [16] B. Hassibi, A. H. Sayed, and T. Kailath, "Linear estimation in Krein spaces—Part I: Theory," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 18–33, Jan. 1996.
- [17] —, "Linear estimation in Krein spaces—Part II: Applications," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 34–49, Jan. 1996.
- [18] —, " H_∞ optimality of the LMS algorithm," *IEEE Trans. Signal Processing*, vol. 44, pp. 267–280, Feb. 1996.
- [19] —, *Indefinite-Quadratic Estimation and Control*, First ed. Philadelphia, PA: SIAM, 1999.
- [20] K. Nishiyama, "Derivation of a fast algorithm of modified H_∞ filters," in *Proc. IEEE Int. Conf. Industrial Electron., Contr., Instrum.*, 2000, pp. 462–467.
- [21] D. W. Lin, "On digital implementation of the fast Kalman algorithms," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 998–1005, May 1984.

Kiyoshi Nishiyama was born in Tokyo, Japan, in 1957. He received the M.E. degree in electrical engineering from Chiba University, Chiba, Japan, in 1985 and the Dr. Eng. degree from Tokyo Institute of Technology in 1991.

He joined the Department of Computer and Information Science, Iwate University, Ueda, Japan, in 1998, where he is presently a Professor. His research interests are in optimization, estimation, and learning theory.