Exponential Asymptotic Stability of Time-Varying Inverse Prediction Error Filters

Roberto López-Valcarce, *Student Member, IEEE*, Soura Dasgupta, *Fellow, IEEE*, Roberto Tempo, *Fellow, IEEE*, and Minyue Fu, *Senior Member, IEEE*

***Abstract*—It is a classical result of linear prediction theory that as long as the minimum prediction error variance is nonzero, the transfer function of the optimum linear prediction error filter for a stationary process is minimum phase, and therefore, its inverse is exponentially stable. Here, extensions of this result to the case of nonstationary processes are investigated. In that context, the filter becomes time varying, and the concept of "transfer function" ceases to make sense. Nevertheless, we prove that under mild condition on the input process, the inverse system remains exponentially stable. We also consider filters obtained in a deterministic framework and show that if the time-varying coefficients of the predictor are computed by means of the recursive weighted least squares algorithm, then its inverse remains exponentially stable under a similar set of conditions.**

I. Introduction

THIS PAPER considers the exponential asymptotic stability (e.a.s.) of inverse linear predictors for zero-mean nonstationary processes with both known and unknown second order statistics. We explain the underlying ideas by first referring to the better understood case of stationary processes when the input statistics are known (the "given covariance" case). For such a process $u(\cdot)$, many signal processing techniques are based on predicting the current value by a linear combination of the m previous measurements. The corresponding forward prediction error is given by

$$f_m(k) = u(k) + \sum_{i=1}^{m} a_i^m u(k-i). \tag{1}$$

The coefficients a_i^m are generally obtained to minimize the mean squared error $E[f_m^2(k)]$ and are uniquely determined by the second-order statistics of $u(\cdot)$.

The linear predictor has found many applications [12], [13]. For example, in differential pulse code modulation (DPCM), it is used to reduce the quantization noise in an encoded signal [5]. In the linear predictive coding (LPC) scheme of speech analysis/synthesis, it is used to reconstruct the speech waveform [9].

Manuscript received May 21, 1999; revised November 17, 1999. The work of R. López-Valcarce was supported by a fellowship from Fundación Pedro Barrié de la Maza (340056). The work of S. Dasgupta was supported in part by the National Science Foundation under grants ECS-9350346, 9970105 and CCR-9973133. The associate editor coordinating the review of this paper and approving it for publication was Prof. Hideaki Sakai.

R. López-Valcarce and S. Dasgupta are with the Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA 52242 USA.

R. Tempo is with CENS-CNR, Politecnico di Torino, Torino, Italy.

M. Fu is with the Department of Electrical and Computer Engineering, University of Newcastle, Callaghan, Australia.

Publisher Item Identifier S 1053-587X(00)04945-X.

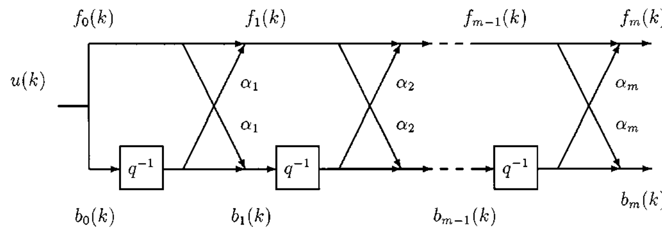

Fig. 1. Lattice implementation of the prediction error filter for stationary input (All-zero lattice).

In seismic signal processing, it is used for deconvolution [14]. Other applications include spectral estimation [7], adaptive filtering [4], and high precision analog to digital converters.

The classical solution to the problem (1) is through the celebrated Levinson–Durbin recursions [8], [12], [13], which simultaneously find both the optimal forward predictor and the optimal backward linear predictor

$$b_m(k) = u(k-m) + \sum_{i=1}^{m} c_i^m u(k-m+i) \tag{2}$$

that minimizes $E[b_m^2(k)]$, $b_m(k)$ being the backward prediction error. Due to the stationarity of $u(\cdot)$, it turns out that $c_i^m = a_i^m$, $1 \leq i \leq m$. Moreover, the polynomial transfer function $1 + a_1^m z^{-1} + \cdots + a_m^m z^{-m}$ from $u(k)$ to $f_m(k)$ is minimum phase (i.e., all its roots lie strictly inside the unit circle) if the minimized value $E[f_m^2(k)]$ is strictly positive [8].

A side benefit of the Levinson–Durbin recursions is that a set of parameters α_i, $1 \leq i \leq m$, which are known as reflection coefficients, are obtained. These can be used to implement the predictor in lattice form as shown in Fig. 1, where each $f_i(k)$ and $b_i(k)$ represents the corresponding optimal signal. This structure presents several advantages over the direct-form implementation of the difference equation (1), such as higher performance under similar hardware constraints, orthogonality of its internal signals, and high modularity (the lattice predictor of order m effectively includes in its structure all predictors of order less than m as well). The minimum-phase property of the forward predictor is equivalent to having $|\alpha_i| < 1$, $1 \leq i \leq m$, [8].

In many applications, such as LPC speech processing, one transmits/stores only $f_m(k)$ and either the reflection coefficents or the direct-form coefficients [5]. The original signal is then asymptotically recovered by implementing the inverse filter either in direct form or in the lattice form of Fig. 2. In either case, the stability of the inverse predictor becomes a crucial issue. Due to the minimum-phase property, this inverse system is exponentially stable.

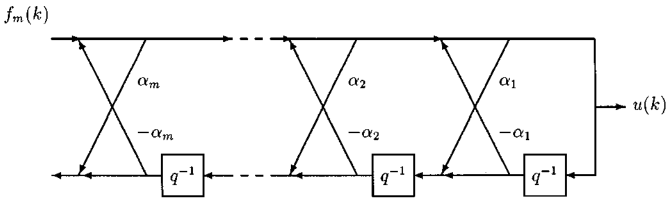

Fig. 2. Lattice implementation of the inverse of the prediction error filter for stationary input (All-pole lattice).

In most real-life problems, however, the process $u(\cdot)$ is nonstationary. As the statistics of $u(\cdot)$ change, so do the the predictor coefficients, leading to a linear time-varying filter. The question that immediately arises is whether the corresponding time-varying inverse system remains exponentially stable. Some approaches try to force stability by checking whether the "frozen" prediction filters obtained along the way are minimum phase and projecting the unstable roots into the stability region if necessary [11]. However, stability of a time-varying system does not, in general, admit a tractable characterization in terms of the poles of the frozen transfer functions. For example, if the reflection coefficients of the recursive filter in Fig. 2 change over time, this time-varying system cannot be guaranteed to be stable, even if the coefficients remain magnitude bounded by one at all times [2].

We explore the stability of time-varying inverse predictors under two settings. In the first, we assume the availability of the time-varying second-order statistics of $u(\cdot)$ and provide sufficient conditions under which the time-varying inverse predictor, which is obtained by the sample by sample minimization of the prediction error variance, is e.a.s. Both direct-form and lattice implementations are studied.

The second setting addresses the practical situation where the statistics of $u(\cdot)$ are unknown, and the predictor coefficients are obtained by solving a weighted least squares minimization problem with each new data sample $u(k)$ (the "given data" case). This can be done efficiently by using the recursive weighted least squares algorithm (RWLS). Again, sufficient conditions under which the time-varying inverse predictor remains e.a.s. are presented.

In both cases, the conditions for e.a.s. are particularly mild and parallel the conditions underlying the nonzero minimum prediction error variance condition characterizing the stationary case. The emphasis in this paper on proving exponential asymptotic, as opposed to mere asymptotic stability, is motivated by robustness issues. Exponentially asymptotically stable systems tolerate modest implementational inaccuracies; asymptotically stable systems, in general, do not [10].

Section II examines the optimum predictor for nonstationary processes in the "given covariance" case. Sections III and IV in turn address the two settings described above. Conclusions are in Section V.

II. Preliminaries

In this section, we provide the initial machinery to set up the part of the problem that concerns optimal prediction in the "given covariance" case. This includes a review of the general form of the prediction error filter, the state-space representations of the inverse filter, the assumptions underlying the pertinent stability analysis, and a discussion on the concept of perfect predictability.

A. General Form of the Predictor

We examine now the structure and properties of the prediction filters in the general case. Let $u(\cdot)$ be a zero-mean (nonstationary) process with autocorrelation coefficients $r_n(k) = E[u(k)u(k-n)]$, and let the parameters $a_i^m(k)$, $c_i^m(k)$ minimize $E[f_m^2(k)]$, $E[b_m^2(k)]$, respectively, where

$$f_m(k) = u(k) + \sum_{i=1}^{m} a_i^m(k)u(k-i) \tag{3}$$

$$b_m(k) = u(k-m) + \sum_{i=1}^{m} c_i^m(k)u(k-m+i). \tag{4}$$

In addition, let $\mathcal{F}_m(k)$ and $\mathcal{B}_m(k)$ be the minimized values of $E[f_m^2(k)]$ and $E[b_m^2(k)]$, respectively.

By the orthogonality principle, it is readily seen that the prediction errors must satisfy for all k

$$E[f_m(k)u(k-i)] = 0, \qquad i = 1, \cdots, m \tag{5}$$

$$E[b_m(k)u(k-j)] = 0, \qquad j = 0, \cdots, m-1. \tag{6}$$

From this, upon defining the $m \times m$ autocorrelation matrix

$$\mathbf{R}_m(k) \triangleq \begin{bmatrix} r_0(k) & r_1(k) & \cdots & r_{m-1}(k) \\ r_1(k) & r_0(k-1) & \cdots & r_{m-2}(k-1) \\ \vdots & \vdots & \ddots & \vdots \\ r_{m-1}(k) & r_{m-2}(k-1) & \cdots & r_0(k-m+1) \end{bmatrix} \tag{7}$$

and the coefficient vectors

$$\mathbf{a}_m(k) \triangleq [a_1^m(k) \quad \cdots \quad a_m^m(k)]' \tag{8}$$

$$c_m(k) \triangleq [c_m^m(k) \quad \cdots \quad c_1^m(k)]' \tag{9}$$

it follows that the predictor coefficients are the solutions to the following time-dependent normal equations

$$\mathbf{R}_m(k-1)\mathbf{a}_m(k) = - \begin{bmatrix} r_1(k) \\ r_2(k) \\ \vdots \\ r_m(k) \end{bmatrix} \tag{10}$$

$$\mathbf{R}_m(k)\mathbf{c}_m(k) = - \begin{bmatrix} r_m(k) \\ r_{m-1}(k-1) \\ \vdots \\ r_1(k-m+1) \end{bmatrix}. \tag{11}$$

Equations (3)–(5) yield the following expressions for $\mathcal{F}_m(k)$ and $\mathcal{B}_m(k)$:

$$\mathcal{F}_m(k) = r_0(k) + \sum_{i=1}^{m} a_i^m(k)r_i(k) \tag{12}$$

$$\mathcal{B}_m(k) = r_0(k-m) + \sum_{i=1}^{m} c_i^m(k)r_i(k-m+i). \tag{13}$$

The extended Levinson–Durbin recursions [12], which we review next, still provide a means to solving for the $(i+1)$th-order

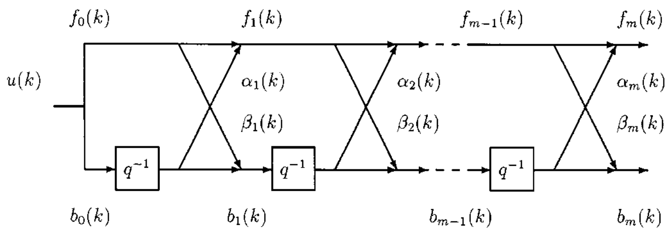

Fig. 3. Time-varying, asymmetric lattice implementation of the prediction error filter for nonstationary input.

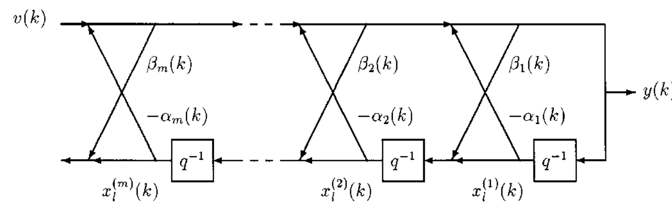

Fig. 4. Time-varying, asymmetric lattice implementation of the inverse of the prediction error filter for nonstationary input.

predictors, given those of order i. To this end, with the convention $f_0(k) = b_0(k) = u(k)$, define the cross-correlation coefficient

$$\Delta_i(k) = E[b_i(k-1)f_i(k)], \qquad i \geq 0 \tag{14}$$

and, for $i \geq 1$, the reflection coefficients

$$\alpha_i(k) = \begin{cases} -\dfrac{\Delta_{i-1}(k)}{\mathcal{B}_{i-1}(k-1)}, & \text{if } \mathcal{B}_{i-1}(k-1) > 0 \\ 0, & \text{if } \mathcal{B}_{i-1}(k-1) = 0 \end{cases} \tag{15}$$

$$\beta_i(k) = \begin{cases} -\dfrac{\Delta_{i-1}(k)}{\mathcal{F}_{i-1}(k)}, & \text{if } \mathcal{F}_{i-1}(k) > 0 \\ 0, & \text{if } \mathcal{F}_{i-1}(k) = 0. \end{cases} \tag{16}$$

Then, one has $a_1^1(k) = \alpha_1(k)$, $c_1^1(k) = \beta_1(k)$ and for $i > 1$

$$\mathbf{a}_i(k) = \begin{bmatrix} \mathbf{a}_{i-1}(k) \\ 0 \end{bmatrix} + \alpha_i(k) \begin{bmatrix} \mathbf{c}_{i-1}(k-1) \\ 1 \end{bmatrix} \tag{17}$$

$$\mathbf{c}_i(k) = \begin{bmatrix} 0 \\ \mathbf{c}_{i-1}(k-1) \end{bmatrix} + \beta_i(k) \begin{bmatrix} 1 \\ \mathbf{a}_{i-1}(k) \end{bmatrix}. \tag{18}$$

Therefore, for $i \geq 1$, the prediction errors obey

$$f_i(k) = f_{i-1}(k) + \alpha_i(k) b_{i-1}(k-1) \tag{19}$$

$$b_i(k) = b_{i-1}(k-1) + \beta_i(k) f_{i-1}(k) \tag{20}$$

which show how the prediction error filters admit a time-varying lattice realization, as shown in Fig. 3. Note that this lattice structure is, in general, asymmetric, i.e., $\alpha_i(k) \neq \beta_i(k)$, by contrast to the stationary case. Observe also that now, $\alpha_i(k)$, $\beta_i(k)$ need not be bounded in magnitude by one individually. Rather, by using (14) and the Cauchy-Schwarz inequality, we have

$$\Delta_i^2(k) \leq \mathcal{F}_i(k)\mathcal{B}_i(k-1). \tag{21}$$

Then, it follows from (15) and (16) that $0 \leq \alpha_i(k)\beta_i(k) \leq 1$ for all k and for all $i \geq 1$. This, of course, reduces in the stationary case to the well-known property $|\alpha_i| \leq 1$.

B. Inverse Filter

As in the stationary case, the inverse forward predictor can be implemented by direct realization of the input–output time-varying difference equation

$$y(k) = v(k) - \sum_{i=1}^{m} a_i^m(k) y(k-i) \tag{22}$$

where $v(\cdot)$ and $y(\cdot)$ are the input and output signals, respectively. Alternatively, one could use the recursive lattice structure depicted in Fig. 4. The state-space equations of the direct-form inverse filter are

$$\mathbf{x}_d(k+1) = \mathbf{F}_d(k)\mathbf{x}_d(k) + \mathbf{G}_d v(k) \tag{23}$$

$$y(k) = \mathbf{H}_d(k)'\mathbf{x}_d(k) + v(k) \tag{24}$$

where the matrices $\mathbf{F}_d(k)$, $\mathbf{G}_d$ and $\mathbf{H}_d(k)$ are given by

$$\mathbf{F}_d(k) \triangleq \begin{bmatrix} -a_1^m(k) & -a_2^m(k) & \cdots & -a_{m-1}^m(k) & -a_m^m(k) \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \tag{25}$$

$$\mathbf{G}_d \triangleq [1 \quad 0 \quad \cdots \quad 0]' \tag{26}$$

$$\mathbf{H}_d(k) \triangleq [-a_1^m(k) \quad -a_2^m(k) \quad \cdots \quad -a_m^m(k)]'. \tag{27}$$

The state vector $\mathbf{x}_d(k)$ is given by

$$\mathbf{x}_d(k) = [y(k-1) \quad \cdots \quad y(k-m)]'.$$

Let $\mathbf{F}_l(k)$ be the $m \times m$ matrix with elements given by

$$[\mathbf{F}_l(k)]_{i,j} = \begin{cases} -\alpha_j(k), & i = 1 \\ -\beta_{i-1}(k)\alpha_j(k), & 2 \leq i \leq j \leq m \\ 1 - \beta_j(k)\alpha_j(k), & i = j+1 \\ 0, & i > j+1 \end{cases} \tag{28}$$

and also define

$$\mathbf{G}_l(k) \triangleq [1 \quad \beta_1(k) \quad \cdots \quad \beta_{m-1}(k)]' \tag{29}$$

$$\mathbf{H}_l(k) \triangleq [-\alpha_1(k) \quad -\alpha_2(k) \quad \cdots \quad -\alpha_m(k)]' \tag{30}$$

one can show that the state-space equations for the lattice inverse filter are

$$\mathbf{x}_l(k+1) = \mathbf{F}_l(k)\mathbf{x}_l(k) + \mathbf{G}_l(k)v(k) \tag{31}$$

$$y(k) = \mathbf{H}_l(k)'\mathbf{x}_l(k) + v(k). \tag{32}$$

Here, the ith element of $\mathbf{x}_l(k)$ is simply the output of the $(m-i+1)$th delay element in Fig. 4. Then, it is readily seen by choosing $v(k)$ as $f_m(k)$ and by reversing the recursions (19) and (20) that

$$\mathbf{x}_l(k) = [b_0(k-1) \quad \cdots \quad b_{m-1}(k-1)]'.$$

Thus, from the fact that when $v(k) = f_m(k)$ then $y(k) = u(k)$, one has from (3) and (4) that

$$\mathbf{x}_l(k) = \mathbf{T}_{dl}(k)\mathbf{x}_d(k) \tag{33}$$

where $\mathbf{T}_{dl}(k)$ is a lower triangular matrix given by

$$\mathbf{T}_{dl}(k) \triangleq \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ c_1^1(k-1) & 1 & 0 & \cdots & 0 \\ c_2^2(k-1) & c_1^2(k-1) & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{m-1}^{m-1}(k-1) & c_{m-2}^{m-1}(k-1) & c_{m-3}^{m-1}(k-1) & \cdots & 1 \end{bmatrix}. \tag{34}$$

It follows that

$$\mathbf{F}_l(k) = \mathbf{T}_{dl}(k+1)\mathbf{F}_d(k)\mathbf{T}_{dl}(k)^{-1} \tag{35}$$

$$\mathbf{G}_l(k) = \mathbf{T}_{dl}(k+1)\mathbf{G}_d \tag{36}$$

$$\mathbf{H}_l(k)' = \mathbf{H}_d(k)'\mathbf{T}_{dl}(k)^{-1} \tag{37}$$

which give the relations between the direct-form and the lattice-form state-space representations.

C. Assumptions

In this subsection, we motivate and state the standing assumptions for Section III. The minimum requirement for stability is that the coefficients $a_i(\cdot)$ be bounded. Since these are solutions to the normal equations (10), it is clear that the assumption below on $\mathbf{R}_m(\cdot)$ suffices for their boundedness.

Assumption 1: Consider $\mathbf{R}_m(k)$ defined in (7). Then, the matrix sequence $\mathbf{R}_m(\cdot)$ is uniformly positive definite (u.p.d.), i.e., there exist positive constants γ_1, γ_2 such that

$$\gamma_1 \mathbf{I} \le \mathbf{R}_m(k) \le \gamma_2 \mathbf{I} \quad \text{for all } k. \tag{38}$$

This requires some elaboration. In the stationary case, $\mathbf{R}_m(\cdot)$ is constant, and therefore , it is either singular or nonsingular for all time k. Its singularity will simply imply the existence of a perfect mth-order predictor for $u(\cdot)$ (see also the next subsection), which may well be bounded. However, in that case, the forward predictor will have zeros on the unit circle, precluding the exponential stability of the inverse filter.

In the nonstationary case, however, there is the added possibility of either $\mathbf{R}_m(\cdot)$ becoming occasionally singular or its smallest eigenvalue going to zero asymptotically. In such a case, we argue that the predictor coefficients may be unbounded. Indeed, consider the case $m = 1$. Then

$$a_1^1(k) = -\frac{r_1(k)}{r_0(k-1)}.$$

While $r_1(k) \le r_0(k)$, it is not true, in general, that $r_1(k) \le r_0(k-1)$. Now, suppose that $r_1(k) = r_0(k)/2$ for all k but that

$$r_0(k) = \begin{cases} 1, & k \text{ even} \\ 1/2^k, & k \text{ odd.} \end{cases}$$

Then, it is readily seen that $a_1^1(k) = -2^{k-1}$ for even k, and therefore, it becomes unbounded. Thus, even though $\mathbf{R}_1(\cdot)$ is nonsingular at all k and bounded, an unbounded $a_1^1(\cdot)$ results. An assumption such as Assumption 1 is needed to ensure a bounded predictor. Observe that this assumption does not force $\mathbf{R}_{m+1}(\cdot)$, which is needed later, to be u.p.d.

D. Perfect Predictability

Recall that in the stationary case, one needs $\mathcal{F}_m = E[f_m^2(k)] > 0$ in order to have all the roots of $1 + a_1^m z^{-1} + \cdots + a_m^m z^{-m}$ strictly inside the unit circle [8]. If $\mathcal{F}_m = 0$, then at least one root lies on the unit circle, and at least one reflection coefficient has magnitude one. In that case, the inverse filter is not e.a.s. Note that $\mathcal{F}_m = 0$ is possible only if the $(m+1)$-dimensional autocorrelation matrix is singular. In view of this, it seems reasonable to expect that in the nonstationary case, some conditions should be imposed on $\mathcal{F}_m(\cdot)$ (or $\mathbf{R}_{m+1}(\cdot)$) in order to have e.a.s. The difficulty that arises is that these quantities are no longer constant: The forward prediction error variance can occasionally become zero, even though it may be positive at other times. Now, moreover, $\mathcal{F}_m(k) \ne \mathcal{B}_m(k)$ in general, which suggests that one may become zero even while the other does not. Next, we present some properties that illustrate this perfect predictability problem. First, a formal definition is needed.

Definition 1: The process $u(\cdot)$ is forward (resp. backward) perfectly predictable of order m at time k_0 if $\mathcal{F}_m(k_0) = 0$ (resp. $\mathcal{B}_m(k_0) = 0$).

Observe from (7), (12), and (13) that

$$\mathcal{F}_m(k) = [1 \quad \mathbf{a}_m(k)']\mathbf{R}_{m+1}(k)\begin{bmatrix} 1 \\ \mathbf{a}_m(k) \end{bmatrix} \tag{39}$$

$$\mathcal{B}_m(k) = [\mathbf{c}_m(k)' \quad 1]\mathbf{R}_{m+1}(k)\begin{bmatrix} \mathbf{c}_m(k) \\ 1 \end{bmatrix}. \tag{40}$$

Then, $\mathcal{F}_m(k_0) = 0$ or $\mathcal{B}_m(k_0) = 0$ implies $\mathbf{R}_{m+1}(k_0)$ singular (the converse is not true though). Therefore, if the matrix sequence $\mathbf{R}_{m+1}(\cdot)$ is u.p.d., then $\mathcal{F}_m(k) \ge \gamma_1$, $\mathcal{B}_m(k) \ge \gamma_1$ for all k, and perfect predictability of order m cannot occur.

In fact, for the first stage that achieves perfect prediction, the following property holds (a proof is given in the Appendix).

Property 1: Consider the sequences $\mathcal{F}_m(\cdot)$, $\mathcal{B}_m(\cdot)$, $\alpha_m(\cdot)$, $\beta_m(\cdot)$ defined in (12)–(16), respectively. Under Assumption 1 the following statements are equivalent.

1) $\mathcal{F}_m(k_0) = 0$.
2) $\mathcal{B}_m(k_0) = 0$.
3) $\alpha_m(k_0)\beta_m(k_0) = 1$.

Therefore, forward and backward perfect predictability are equivalent. In the next section, we present sufficient conditions for the e.a.s. of the inverse system. It turns out that these conditions indeed involve the frequency with which the process $u(\cdot)$ becomes perfectly predictable.

III. Stability of the Inverse Filter

Recall the definition of e.a.s. for time-varying systems.

Definition 2: The linear time-varying system $\mathbf{x}(k+1) = \mathbf{A}(k)\mathbf{x}(k)$ is e.a.s. if there exist $0 \le d < 1$ and $c > 0$ such that for every k_0 and every bounded initial condition $\|\mathbf{x}(k_0)\| < \infty$, the resulting state vector sequence $\mathbf{x}(\cdot)$ satisfies the exponential bound

$$\|\mathbf{x}(k)\| < cd^{k-k_0}\|\mathbf{x}(k_0)\|, \qquad \text{for all } k \ge k_0.$$

We will also use the concept of uniform stabilizability of a pair $[\mathbf{A}(\cdot), \mathbf{b}(\cdot)]$. The definition is rather technical and can be found in [1]. Loosely speaking, uniform stabilizability of $[\mathbf{A}(\cdot), \mathbf{b}(\cdot)]$ reduces to saying that in the system

$$\mathbf{x}(k+1) = \mathbf{A}(k)\mathbf{x}(k) + \mathbf{b}(k)v(k)$$

the state trajectories that cannot be controlled from the input must be exponentially decaying, and vice versa. It was shown in [1] that uniform stabilizability is equivalent to the ability to stabilize the system by means of state feedback (hence, its name). Since this is the property of interest here, we shall use the following definition.

Definition 3: The pair $[\mathbf{A}(\cdot), \mathbf{b}(\cdot)]$ is uniformly stabilizable if there exists a bounded gain sequence $\mathbf{g}(\cdot)$ such that the system $\mathbf{x}(k+1) = [\mathbf{A}(k) + \mathbf{b}(k)\mathbf{g}(k)']\mathbf{x}(k)$ is e.a.s.

We are now in position to study the stability of the inverse predictor. Consider first the direct-form implementation of (23) and (24). The key observation is stated next.

Lemma 1: The pair $[\mathbf{F}_d(\cdot), \mathbf{G}_d]$ defined in (7) and (12) satisfies the following time-varying equation for all k:

$$\mathbf{R}_m(k) - \mathbf{F}_d(k)\mathbf{R}_m(k-1)\mathbf{F}_d(k)' = \mathcal{F}_m(k)\mathbf{G}_d\mathbf{G}_d' \quad (41)$$

where $\mathbf{R}_m(k)$ and $\mathcal{F}_m(k)$ are given in (7), (12), respectively.

Proof: Define $\mathbf{e}_i$ as the $m \times 1$ unit vector with 1 in the ith location, and let

$$t_{i,j} \triangleq \mathbf{e}_i'[\mathbf{R}_m(k) - \mathbf{F}_d(k)\mathbf{R}_m(k-1)\mathbf{F}_d(k)']\mathbf{e}_j$$

i.e., $t_{i,j}$ is the i, j element of the left-hand side of (41). Then, from (7), (10), and (12), we immediately have

$$\begin{aligned} t_{1,1} &= r_0(k) - \mathbf{a}_m(k)'\mathbf{R}_m(k-1)\mathbf{a}_m(k) \\ &= r_0(k) + \mathbf{a}_m(k)' \begin{bmatrix} r_1(k) \\ \vdots \\ r_m(k) \end{bmatrix} \\ &= \mathcal{F}_m(k) \\ &= \mathbf{e}_1'\mathcal{F}_m(k)\mathbf{G}_d\mathbf{G}_d'\mathbf{e}_1. \end{aligned}$$

For $i > 1$, from (7), (10), and (25), one has

$$\begin{aligned} t_{i,1} &= r_{i-1}(k) - \mathbf{e}_{i-1}'\mathbf{R}_m(k-1)\mathbf{a}_m(k) \\ &= r_{i-1}(k) - \mathbf{e}_{i-1}' \begin{bmatrix} r_1(k) \\ \vdots \\ r_m(k) \end{bmatrix} \\ &= 0. \end{aligned}$$

Finally, for $i \geq j > 1$, from (7) and (25)

$$\begin{aligned} t_{i,j} &= r_{i-j}(k-j+1) - \mathbf{e}_{i-1}'\mathbf{R}_m(k-1)\mathbf{e}_{j-1} \\ &= r_{i-j}(k-j+1) - r_{i-j}(k-j+1) \\ &= 0. \end{aligned}$$

Note that $\mathbf{e}_i'\mathcal{F}_m(k)\mathbf{G}_d\mathbf{G}_d'\mathbf{e}_j = 0$ unless $i = j = 1$. This, together with the symmetry of the matrices in (41), proves the result. ■

Note that (41) resembles a Lyapunov equation [1]. This fact will be used in the Appendix to prove the following main result of this section.

Theorem 1: Under Assumption 1, the system (23) is e.a.s. if there exists an integer S and a constant $\epsilon > 0$ such that for all k, there exists n_k satisfying the following.

1) $n_k \geq k$.
2) $n_k + m - 1 \leq k + S$.
3) $\mathcal{F}_m(n_k + i) \geq \epsilon$ for $i = 0, 1, \cdots m-1$.

The conditions in Theorem 1 impose in a sense a limitation on "how often" the process $u(\cdot)$ may become perfectly predictable. They boil down to saying that in any time window of a fixed size S, there should be m consecutive time instants in which the forward prediction error variance is bounded away from zero (recall that m is the order of the system). These conditions are satisfied if, for example, $\mathbf{R}_{m+1}(\cdot)$ is u.p.d. and are a natural extension to the standard stability condition for the stationary case.

Next, we present two examples that should clarify the role of the conditions in Theorem 1.

Example 1: Here, we show that the conditions in Theorem 1 are not necessary, in general. Consider the process $u(\cdot)$, which is defined as follows:

$$u(k) = \begin{cases} u(k-1), & \text{if } k \bmod 3 = 0 \\ w(k), & \text{else} \end{cases}$$

where $w(\cdot)$ is a zero-mean, unit-variance white process. Now, observe that

$$[r_1(k), r_2(k), r_3(k)] = \begin{cases} [1, 0, 0], & \text{if } k \bmod 3 = 0 \\ \mathbf{0}, & \text{else.} \end{cases}$$

Further, for $k \bmod 3 = 0$, one has $\mathbf{R}_3(k-1) = \mathbf{I}$. Thus, from (10), the optimum forward prediction error of order $m = 3$ is

$$f_3(k) = \begin{cases} u(k) - u(k-1), & \text{if } k \bmod 3 = 0 \\ u(k), & \text{else.} \end{cases} \quad (42)$$

The corresponding minimum variance is

$$\mathcal{F}_3(k) = \begin{cases} 0, & \text{if } k \bmod 3 = 0 \\ 1, & \text{else.} \end{cases}$$

The prediction error filter is periodically time varying with period 3. Accordingly, the inverse filter is e.a.s. iff all the eigenvalues of the state transition matrix $\mathbf{F}_d(0)\mathbf{F}_d(1)\mathbf{F}_d(2)$ have magnitude strictly less than one. From (42), we have

$$\mathbf{F}_d(0) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\mathbf{F}_d(1) = \mathbf{F}_d(2) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Thus, $\mathbf{F}_d(0)\mathbf{F}_d(1)\mathbf{F}_d(2)$ is the zero matrix, and the inverse filter is e.a.s. Note, however, that during any three consecutive time instants, $\mathcal{F}_3(k)$ becomes zero once so that the conditions of Theorem 1 do not hold. This shows that these conditions, although sufficient, need not be necessary for e.a.s. ■

Example 2: Let $T > 1$; consider now the process $u(\cdot)$, which is defined as

$$u(k) = \begin{cases} u(k-T), & \text{if } k \bmod T = 0 \\ w(k), & \text{else} \end{cases} \quad (43)$$

where again, $w(\cdot)$ is a zero-mean, unit-variance white process. Now, observe that

$$[r_1(k), \cdots, r_T(k)] = \begin{cases} [0, \cdots, 0, 1], & \text{if } k \bmod T = 0 \\ \mathbf{0}, & \text{else.} \end{cases}$$

Further, for $k \bmod T = 0$, one has $\mathbf{R}_T(k-1) = \mathbf{I}$. Thus, from (10), the optimum forward prediction error of order $m = T$ is

$$f_T(k) = \begin{cases} u(k) - u(k-T), & \text{if } k \bmod T = 0 \\ u(k), & \text{else} \end{cases}$$

and the corresponding variance is

$$\mathcal{F}_T(k) = \begin{cases} 0, & \text{if } k \bmod T = 0 \\ 1, & \text{else.} \end{cases}$$

Again, the prediction error filter is periodically time varying only now with period T. Further, now

$$\mathbf{F}_d(0) = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

and

$$\mathbf{F}_d(1) = \cdots = \mathbf{F}_d(T-1) = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}.$$

Note that $\mathbf{F}_d(1)$ is simply a shift matrix, i.e., for any $[\eta_1, \cdots, \eta_T]$

$$\mathbf{F}_d(1)[\eta_1 \quad \eta_2 \quad \cdots \quad \eta_T]' = [0 \quad \eta_1 \quad \cdots \quad \eta_{T-1}]'.$$

Thus

$$\mathbf{F}_d(1) \cdots \mathbf{F}_d(T-1) = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \end{bmatrix}$$

and the state transition matrix reduces to

$$\mathbf{F}_d(0)\mathbf{F}_d(1) \cdots \mathbf{F}_d(T-1) = \text{diag}(1, 0, \cdots, 0).$$

Hence, the inverse system is not e.a.s. since $\lambda = 1$ is an eigenvalue. Observe that again, the conditions of Theorem 1 are violated since during any T consecutive time instants, $\mathcal{F}_T(k)$ must become zero once. Note that T can be arbitrarily large. This example illustrates how the violation of the conditions of Theorem 1 may cause instability. ■

Consider now the lattice form of the inverse predictor, given by (31) and (32). As shown by (35), the feedback matrices of the lattice and direct forms are related through the matrix sequence $\mathbf{T}_{dl}(\cdot)$ given in (34). Under Assumption 1, the elements of $\mathbf{T}_{dl}(\cdot)$ are bounded from above since they satisfy the normal equations (11). Therefore, $\mathbf{T}_{dl}(\cdot)'\mathbf{T}_{dl}(\cdot)$ is bounded, which implies that e.a.s. of the direct-form inverse is equivalent to e.a.s. of the lattice-form inverse. Nevertheless, it is instructive to present the analog of (41) for the lattice form.

By the orthogonality conditions (6), the random variables $b_i(k)$, $i \geq 0$ are mutually orthogonal. Then

$$\mathbf{P}_m(k) \triangleq E\left\{ \begin{bmatrix} b_0(k) \\ \vdots \\ b_{m-1}(k) \end{bmatrix} [b_0(k) \quad \cdots \quad b_{m-1}(k)] \right\} = \text{diag}\left(\mathcal{B}_0(k) \quad \cdots \quad \mathcal{B}_{m-1}(k)\right). \tag{44}$$

Note from (33) that the matrix sequence $\mathbf{P}_m(\cdot)$ satisfies

$$\mathbf{P}_m(k) = \mathbf{T}_{dl}(k+1)\mathbf{R}_m(k)\mathbf{T}_{dl}(k+1)'. \tag{45}$$

Premultiplying in (41) by $\mathbf{T}_{dl}(k+1)$, postmultiplying by $\mathbf{T}_{dl}(k+1)'$, and using (35) and (36) and (45), one obtains the following Lyapunov-like equation for the lattice structure:

$$\mathbf{P}_m(k) - \mathbf{F}_l(k)\mathbf{P}_m(k-1)\mathbf{F}_l(k)' = \mathcal{F}_m(k)\mathbf{G}_l(k)\mathbf{G}_l(k)' \tag{46}$$

for all k. This could have been used to directly show the e.a.s. of the lattice structure using the same technique as in the proof of Theorem 1. It is appealing, though not entirely surprising, that the Lyapunov matrices featuring in the stability analysis of the direct and lattice realizations are the autocorrelation matrix and a diagonal matrix, respectively. Such is also the case for stationary $u(\cdot)$.

IV. Stability of the RWLS Inverse Filter

In this section, we analyze the "given data" case, in which the statistics of the input process are unknown, and one attempts to obtain the filter that is optimum in a weighted least squares (WLS) sense based on the observed data. In the nonrecursive WLS scheme known as the autocorrelation method, a time-invariant prediction error filter is obtained from a *finite* data register and is minimum phase [4] as long as the minimum value of the underlying cost function is nonzero. However, in recursive implementations, the input data sequence need not be finite in principle, and the optimum filter is updated on a sample-by-sample basis. The inverse system again becomes time varying, rendering its stability analysis nontrivial.

If the samples $u(0), \cdots u(k)$ are available, then one seeks the coefficients $a_j^m(k)$ that minimize

$$J_D(k) \triangleq \sum_{i=0}^{k} \lambda^{k-i} \left[u(i) + \sum_{j=1}^{m} a_j^m u(i-j) \right]^2 \tag{47}$$

where $0 < \lambda \leq 1$ is a forgetting factor. The data are prewindowed according to $u(j) = 0,\ j < 0$. As k increases, this sequence of problems can be efficiently solved by the recursive weighted least squares (RWLS) algorithm [4], [17], [18], which computes the optimum coefficients at time k from those at time $k-1$. We will consider the RWLS algorithm with soft constrained initialization (see [4, Tab. 13.1]). This starting procedure is commonly used in practice; it sets the initial value of the inverse of an autocorrelation matrix defined below at $\mathbf{P}(0) = \delta^{-1}\mathbf{I}$ with $0 < \delta \ll 1$. Due to this initialization, which is used to avoid a singular autocorrelation matrix, the coefficients obtained by the RWLS algorithm minimize the cost function

$$\overline{J}_D(k) \triangleq \lambda^k \delta \sum_{j=1}^{m} [a_j^m]^2 + \sum_{i=0}^{k} \lambda^{k-i} \left[u(i) + \sum_{j=1}^{m} a_j^m u(i-j) \right]^2 \tag{48}$$

rather than (47) [17].

In this deterministic framework, the autocorrelation coefficients are defined as

$$r_{jD}(k) = \sum_{i=0}^{k} \lambda^{k-i} u(i) u(i-j), \qquad j \geq 0 \tag{49}$$

where we assume that $u(k) = 0$ for all $k < 0$. The deterministic autocorrelation matrix $\mathbf{R}_{mD}(k)$ is defined to bear the same relationship to $r_{jD}(k)$ as $\mathbf{R}_m(k)$ does to $r_j(k)$ in (7). If, in addition, we introduce the vector $\mathbf{r}_D(k) = [r_{1D}(k) \quad \cdots \quad r_{mD}(k)]'$, then the coefficients $\mathbf{a}(k) = [a_1^m(k) \quad \cdots \quad a_m^m(k)]'$ recursively computed by the RWLS algorithm satisfy the following modified normal equations:

$$\mathbf{M}(k)\mathbf{a}(k) = -\mathbf{r}_D(k) \tag{50}$$

with

$$\mathbf{M}(k) = \delta\lambda^k \mathbf{I} + \mathbf{R}_{mD}(k-1). \tag{51}$$

A typical assumption made on the deterministic autocorrelation matrix $\mathbf{R}_{mD}(\cdot)$ is the condition of persistent excitation (p.e.); see [6].

Assumption 2: The matrix sequence $\mathbf{R}_{mD}(\cdot)$ is u.p.d. for all $k \geq k_0$.

Since one assumes $u(k) = 0$ for all $k < 0$ [17], one cannot have $\mathbf{R}_{mD}(0)$ positive definite. In fact

$$\mathbf{R}_{mD}(k) = \sum_{i=0}^{k} \lambda^{k-i}\mathbf{u}(i)\mathbf{u}(i)' \tag{52}$$

with $\mathbf{u}(k) = [u(k)\ u(k-1) \cdots u(k-m+1)]'$. Thus, k_0 in Assumption 2 must exceed $m-1$. In addition, observe from (52) that this ensures that $\mathbf{M}(k)$ is u.p.d. for all $k \geq 0$.

Therefore, the $\mathbf{a}(k)$ provided by (50) obeys for all $k \geq 0$ and for some N

$$\|\mathbf{a}(k)\| \leq N. \tag{53}$$

Note also that Assumption 2 holds iff there exist positive ϵ_1, ϵ_2 and integer L such that for all k

$$\epsilon_1 \mathbf{I} \leq \sum_{i=k}^{k+L} \mathbf{u}(i)\mathbf{u}(i)' \leq \epsilon_2 \mathbf{I} \tag{54}$$

(see [3]). We wish to show that the system

$$\mathbf{x}_d(k+1) = \mathbf{F}_d(k)\mathbf{x}_d(k) \tag{55}$$

is e.a.s., with $\mathbf{F}_d(k)$ defined as in (25). Consider now instead the vector $\overline{\mathbf{a}}(k)$ defined by

$$\mathbf{R}_{mD}(k-1)\overline{\mathbf{a}}(k) = -\mathbf{r}_D(k) \qquad \text{for all } k \geq k_0 \tag{56}$$

and the corresponding system

$$\overline{\mathbf{x}}_d(k+1) = \overline{\mathbf{F}}_d(k)\overline{\mathbf{x}}_d(k) \qquad \text{for all } k \geq k_0 \tag{57}$$

where with $\overline{\mathbf{a}}(k) = [\overline{a}_1^m(k) \quad \cdots \quad \overline{a}_m^m(k)]'$, the matrix $\overline{\mathbf{F}}_d(k)$ is given by

$$\overline{\mathbf{F}}_d(k) = \begin{bmatrix} -\overline{a}_1^m(k) & -\overline{a}_2^m(k) & \cdots & -\overline{a}_{m-1}^m(k) & -\overline{a}_m^m(k) \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}. \tag{58}$$

The following lemma establishes the relationship between the e.a.s. of (57) and that of (55).

Lemma 2: Consider (57) under (56) and (58) and (55) under (25), (50), and (51). Suppose Assumption 2 holds and that (57) is e.a.s. for all $k \geq k_0$, i.e., there exist c and $0 \leq d < 1$ such that for all $k_1 \geq k_0$, $k \geq k_1$ and $\overline{\mathbf{x}}_d(k_1)$

$$\|\overline{\mathbf{x}}_d(k)\| \leq c d^{k-k_1} \|\overline{\mathbf{x}}_d(k_1)\|. \tag{59}$$

Then, (55) is e.a.s. for all $k \geq 0$.

Proof: Because of (50), (51), and (56)

$$\begin{aligned}\overline{\mathbf{a}}(k) - \mathbf{a}(k) &= \left\{[\delta\lambda^k \mathbf{I} + \mathbf{R}_{mD}(k-1)]^{-1} \right. \\ &\quad \left. -\mathbf{R}_{mD}^{-1}(k-1)\right\} \mathbf{r}_D(k) \\ &= -\delta\lambda^k [\delta\lambda^k \mathbf{I} + \mathbf{R}_{mD}(k-1)]^{-1} \\ &\quad \cdot \mathbf{R}_{mD}^{-1}(k-1)\mathbf{r}_D(k).\end{aligned}$$

Because of (54) (which also ensures the boundedness of $\mathbf{r}_D(\cdot)$) and Assumption 2, there exists γ such that for all $k \geq 0$, $\|\overline{\mathbf{a}}(k) - \mathbf{a}(k)\| \leq \gamma\lambda^k$. Thus

$$\|\overline{\mathbf{F}}_d(k) - \mathbf{F}_d(k)\| \leq \overline{\gamma}\lambda^k \qquad \text{for all } k \geq k_0$$

for some constant $\overline{\gamma}$. Thus, if (57) is e.a.s. for all $k \geq k_0$, then so is (55) under (25) and (50), [19]. Further, since $\mathbf{F}_d(\cdot)$ is bounded, this proves that (55) under (25) and (50) is e.a.s. for all $k \geq 0$. ■

Thus, instead of directly proving that (55) is e.a.s. under (50), it suffices to consider (56)–(58) for all $k \geq k_0$. Observe that the value of (47) under (56) is

$$J_*(k) = r_{0D}(k) + \overline{a}(k)'\mathbf{r}_D(k).$$

Further, as in Section III, one has

$$\mathbf{R}_{mD}(k) - \overline{\mathbf{F}}_d(k)\mathbf{R}_{mD}(k-1)\overline{\mathbf{F}}_d(k)' = J_*(k)\mathbf{G}_d\mathbf{G}_d'.$$

Noting the correspondence between this equation and (41), one arrives at the result below, where now, $J_*(k)$ plays the role of $\mathcal{F}_m(k)$ in Section III.

Theorem 2: Suppose that Assumption 2 holds, $0 < \lambda < 1$, and there exist an integer S and a constant $\epsilon > 0$ such that for all k, there exists n_k satisfying the following.

1) $n_k \geq k$.
2) $n_k + m - 1 \leq k + S$.
3) $J_*(n_k + i) \geq \epsilon$ for $i = 0, 1, \cdots m-1$.

Then, the inverse prediction error filter (23) with coefficients obtained by the RWLS algorithm is e.a.s.

Although we have considered here the transversal filter computed by the standard RWLS algorithm, these stability results also apply well to least-squares filters obtained by order recursive algorithms, in view of the equivalence of the direct-form and the lattice implementations of the inverse system. For example, several lattice LS algorithms in the literature can be shown to be equivalent to the QR decomposition-based least squares lattice (QRD-LSL) algorithm [17], [18]. The QRD-LSL algorithm with zero initialization of the state variables produces a filter that minimizes at every k the cost function (47) [18]. Therefore, our approach readily applies to those algorithms as well.

V. Conclusions

We have examined the stability properties of time-varying inverse prediction error filters. Our analysis provides sufficient conditions on the nonstationary input sequence for exponential asymptotic stability of these systems (in both, the "given covariance" and the "given data" cases and, for both, the direct-form and lattice structures. The key condition in both cases is that the input sequence should not become perfectly predictable (i.e., yield an arbitrarily small prediction error variance) arbitrarily often. These properties are natural extensions of the minimum-phase character of the time-invariant prediction error filters obtained in the given covariance case for stationary processes and in the given data case when the coefficients are obtained from a pre- and post-windowed finite data register (autocorrelation method).

As a future line of research, it would seem worthwhile to explore the connections of the results here presented to the work of Sayed *et al.* [15], [16], where the properties of displacement equations similar to (41) and (46), and their connections to interpolation problems, are investigated.

APPENDIX

A. Proof of Property 1

First, observe that from the direct evaluation of $\mathcal{F}_m(k)$, $\mathcal{B}_m(k)$ using (14)–(16) and (19) and (20), one finds that for $m \geq 1$

$$\mathcal{F}_m(k) = \mathcal{F}_{m-1}(k) - \alpha_m^2(k)\mathcal{B}_{m-1}(k-1) \quad (60)$$
$$\mathcal{B}_m(k) = \mathcal{B}_{m-1}(k-1) - \beta_m^2(k)\mathcal{F}_{m-1}(k). \quad (61)$$

Now, suppose that $\mathcal{F}_m(k_0) = 0$. Then

$$\mathcal{F}_{m-1}(k_0) - \alpha_m^2(k_0)\mathcal{B}_{m-1}(k_0-1) = 0.$$

Hence

$$\alpha_m^2(k_0) = \frac{\mathcal{F}_{m-1}(k_0)}{\mathcal{B}_{m-1}(k_0-1)}.$$

The last step is valid since $\mathbf{R}_m(\cdot)$ u.p.d. implies $\mathcal{F}_{m-1}(k) > 0$, $\mathcal{B}_{m-1}(k) > 0$ for all k. Given the definition (15) of $\alpha_m(\cdot)$, it follows that $\Delta_{m-1}^2(k_0) = \mathcal{F}_{m-1}(k_0)\mathcal{B}_{m-1}(k_0-1)$, and therefore, from (15) and (16)

$$\alpha_m(k)\beta_m(k) = \frac{\Delta_{m-1}^2(k_0)}{\mathcal{F}_{m-1}(k_0)\mathcal{B}_{m-1}(k_0-1)} = 1.$$

Now, using the definition (18) of $\beta_m(\cdot)$ and the fact that $\mathcal{F}_m(k_0) > 0$, from (61)

$$\begin{aligned}\mathcal{B}_m(k_0) &= \mathcal{B}_{m-1}(k-1) - \frac{\Delta_{m-1}^2(k_0)}{\mathcal{F}_{m-1}^2(k_0)}\mathcal{F}_{m-1}(k_0)\\ &= \mathcal{B}_{m-1}(k-1) - \mathcal{B}_{m-1}(k-1)\\ &= 0.\end{aligned}$$

The implication $\mathcal{B}_m(k_0) = 0 \Rightarrow \alpha_m(k_0)\beta_m(k_0) = 1 \Rightarrow \mathcal{F}_m(k_0) = 0$ is similarly proved. ■

B. Proof of Theorem 1

Let $\hat{\mathbf{F}}_d(k) = \mathbf{F}_d(-k)'$. By the duality theorem [1], the system $\mathbf{x}(k+1) = \mathbf{F}_d(k)\mathbf{x}(k)$ is e.a.s. if and only if the system $\hat{\mathbf{x}}(k+1) = \hat{\mathbf{F}}_d(k)\hat{\mathbf{x}}(k)$ is e.a.s.; define now $\hat{\mathbf{R}}_m(k) = \mathbf{R}_m(-k)' = \mathbf{R}_m(-k)$ and $\hat{\mathcal{F}}_m(k) = \mathcal{F}_m(-k)$. Then, (41) becomes

$$\hat{\mathbf{R}}_m(k) - \hat{\mathbf{F}}_d(k)'\hat{\mathbf{R}}_m(k+1)\hat{\mathbf{F}}_d(k) = \hat{\mathcal{F}}_m(k)\mathbf{G}_d\mathbf{G}_d' \quad (62)$$

which is a time-varying Lyapunov equation. Accordingly, the extended lemma of Lyapunov [1] guarantees exponential stability of $\hat{\mathbf{x}}(k+1) = \hat{\mathbf{F}}_d(k)\hat{\mathbf{x}}(k)$ provided that we have the following.

1) The sequence matrix $\hat{\mathbf{R}}_m(\cdot)$ is bounded and positive semidefinite.
2) $\hat{\mathbf{F}}_d(\cdot)$, $\hat{\mathcal{F}}_m(\cdot)$ are bounded.
3) The pair $[\mathbf{F}_d(\cdot), \mathcal{F}_m^{1/2}(\cdot)\mathbf{G}_d]$ is uniformly stabilizable.

Condition 1 is a consequence of Assumption 2. Condition 2 follows from (39) and the fact that $\mathbf{R}_m(k)$ u.p.d. implies the boundedness of the $a_i^m(\cdot)$'s. Therefore, it remains to be shown that the pair $[\mathbf{F}_d(\cdot), \mathcal{F}_m^{1/2}(\cdot)\mathbf{G}_d]$ is uniformly stabilizable. To do so, introduce the following gain sequence:

$$\mathbf{g}(k) = \begin{cases}\mathcal{F}_m^{-1/2}(k)\mathbf{a}_m(k), & \text{if } \mathcal{F}_m(k) \geq \epsilon\\ \mathbf{0}, & \text{if } \mathcal{F}_m(k) < \epsilon.\end{cases}$$

Observe that $\mathbf{g}(\cdot)$ is bounded. We claim now that the system $\tilde{\mathbf{x}}(k+1) = \tilde{\mathbf{F}}_d(k)\tilde{\mathbf{x}}(k)$ is e.a.s., where

$$\tilde{\mathbf{F}}_d(k) = \mathbf{F}_d(k) + \mathcal{F}_m^{1/2}(k)\mathbf{G}_d\hat{\mathbf{g}}(k)'.$$

To see this, let $\mathbf{Z}$ be the $m \times m$ shift matrix with ones in the first subdiagonal and zeros elsewhere. Since all the eigenvalues of this matrix are zero, one has $\mathbf{Z}^m = 0$ from the Cayley–Hamilton theorem. It is clear that

$$\tilde{\mathbf{F}}_d(k) = \begin{cases}\mathbf{Z}, & \text{if } \mathcal{F}_m(k) \geq \epsilon\\ \mathbf{F}_d(k), & \text{if } \mathcal{F}_m(-k) < \epsilon.\end{cases}$$

Now, let $\Phi(k,l)$, $k \geq l$ be the state transition matrix of the system $\tilde{\mathbf{x}}(k+1) = \tilde{\mathbf{F}}_d(k)\tilde{\mathbf{x}}(k)$, i.e.

$$\Phi(k,l) = \begin{cases}\mathbf{I}, & \text{if } k = l\\ \tilde{\mathbf{F}}_d(k-1)\tilde{\mathbf{F}}_d(k-2)\cdots\tilde{\mathbf{F}}_d(l), & \text{if } k > l.\end{cases}$$

Then, $\mathbf{x}(k) = \Phi(k,l)\mathbf{x}(l)$ for $k \geq l$.

Since for any k there exists an n_k obeying the conditions of the theorem, one has

$$\tilde{\mathbf{F}}(n_k + i) = \mathbf{Z}, \qquad i = 0, 1, \cdots m-1.$$

Then, for all $n \geq k + S$, one has

$$\begin{aligned}\Phi(n,k) &= \Phi(n, n_k+m)\Phi(n_k+m, n_k)\Phi(n_k,k)\\ &= \Phi(n, n_k+m)\tilde{\mathbf{F}}_d(n_k+m-1)\cdots\tilde{\mathbf{F}}_d(n_k)\Phi(n_k,k)\\ &= \Phi(n, n_k+m)\mathbf{Z}^m\Phi(n_k,k)\\ &= \mathbf{0}\end{aligned}$$

because $\mathbf{Z}^m = \mathbf{0}$. Therefore, any state trajectory of $\tilde{\mathbf{x}}(k+1) = \tilde{\mathbf{F}}_d(k)\tilde{\mathbf{x}}(k)$ becomes zero after at most S time instants, showing that the system is e.a.s. and that the pair $[\mathbf{F}_d(\cdot), \mathcal{F}_m^{1/2}(\cdot)\mathbf{G}_d]$ is uniformly stabilizable. This concludes the proof. ■

REFERENCES

[1] B. D. O. Anderson and J. B. Moore, "Detectability and stabilizability of time-varying discrete-time linear systems," *SIAM J. Contr. Optim.*, vol. 19, no. 1, pp. 20–32, Jan. 1981.

[2] S. Dasgupta, M. Fu, and C. Schwarz, "Robust relative stability of time invariant and time varying lattice filters," *IEEE Trans. Signal Processing*, vol. 46, pp. 2088–2100, Aug. 1998.

[3] S. Dasgupta and Y. Huang, "Asymptotically convergent modified recursive least-squares with data dependent updating and forgetting factor for systems with bounded noise," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 383–391, June 1987.

[4] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[5] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[6] C. R. Johnson, Jr., *Lectures on Adaptive Parameter Estimation*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[7] S. M. Kay, *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[8] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 63, pp. 561–580, Apr. 1975.

[9] J. D. Markel and A. H. Gray, Jr., *Linear Prediction of Speech*. New York: Springer-Verlag, 1976.

[10] B. D. O. Anderson, R. R. Bitmead, C. R. Johnson, Jr., P. V. Kokotovic, R. L. Kosut, I. M. Y. Mareels, L. Praly, and B. D. Riedle, *Stability of Adaptive Systems: Passivity and Averaging Analysis*. Cambridge, MA: MIT Press, 1986.

[11] R. L. Moses and D. Liu, "Determining the closest stable polynomial to an unstable one," *IEEE Trans. Signal Processing*, vol. 39, pp. 901–906, Apr. 1991.

[12] S. J. Orfanidis, *Optimum Signal Processing: An Introduction*, 2nd ed. New York: MacMillan, 1988.

[13] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*, 2nd ed. New York: MacMillan, 1992.

[14] E. A. Robinson and S. Treitel, *Geophysical Signal Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1980.

[15] A. H. Sayed, T. Constantinescu, and T. Kailath, "Time-variant displacement structure and interpolation problems," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 960–976, May 1994.

[16] A. H. Sayed, H. Lev-Ari, and T. Kailath, "Time-variant displacement structure and triangular arrays," *IEEE Trans. Signal Processing*, vol. 42, pp. 1052–1062, May 1994.

[17] A. H. Sayed and T. Kailath, "A state-space approach to adaptive RLS filtering," *IEEE Signal Processing Mag.*, vol. 11, pp. 18–60, July 1994.

[18] A. H. Sayed and T. Kailath, "Recursive least-squares adaptive filters," in *The Digital Signal Processing Handbook*. Piscataway, NJ: CRC-IEEE, 1998.

[19] J. L. Willems, *Stability Theory of Dynamical Systems*. London, U.K.: Nelson, 1970.

Roberto López-Valcarce (S'96) was born in Spain in 1971. He received the engineering degree from Universidad de Vigo, Vigo, Spain, in 1995. He received the M.S. degree in electrical engineering from the University of Iowa (UI), Iowa City, in 1998. He is currently pursuing the Ph.D. degree at UI.

From 1995 to 1996, he was a Systems Engineer with Intelsis Sistemas Inteligentes. He was the recipient of a Fundación Pedro Barrié de la Maza fellowship for continuation of studies at UI. His research interests are in adaptive signal processing and communications.

Soura Dasgupta (M'87–SM'93–F'98) was born in 1959 in Calcutta, India. He received the B.E. degree in electrical engineering from the University of Queensland, Australia, Brisbane, in 1980 and the Ph.D. in systems engineering from the Australian National University, Canberra, in 1985.

He is currently Dean's Diamond Professor of Electrical and Computer Engineering at the University of Iowa, Iowa City. In 1981, he was a Junior Research Fellow in the Electronics and Communications Sciences Unit at the Indian Statistical Institute, Calcutta. He has held visiting appointments at the University of Notre Dame, University of Iowa, Univesité Catholique de Louvain-La-Neuve, and the Australian National University. His research interests are in controls, signal processing and neural networks.

Dr. Dasgupta was a corecipient of the Gullimen Cauer Award for the best paper published in the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS in 1990 and 1991 and was a Presidential Faculty Fellow. Between 1988 and 1991, he served as an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He is an Associate Editor for the IEEE Control Systems Society Conference Editorial Board.

Roberto Tempo (F'00) was born in Cuorgné, Italy, in 1956. He received the B.S. degree in electrical engineering from Politechnico di Torino, Torino, Italy, in 1980.

From 1981 to 1983, he was with the Dipartimento di Automatica e Informatica, Politechnico di Torino. In 1984, he joined the National Research Council of Italy (CNR) at the research institute CENS (prsently IRITI), Torino, where he is currently a Director of Research of Systems and Computer Engineering. He has held visiting and research positions at Columbia University, New York, NY, the German Aerospace Research Organization, Oberpfaffenhofen, Germany, and the University of Wisconsin, Madison. His main research areas include robustness analysis and control of uncertain systems and identification of complex systems subject to bounded errors.

Dr. Tempo received the "Outstanding Paper Prize Award" from the International Federation of Automatic Control for a paper published in *Automatica.* in 1993. He is currently an Editor of *Automatica* and an Associate Editor of IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He served as Program Vice-Chair for short papers of the IEEE Conference on Decision and Control, Phoenix, AZ, in 1999. He is an appointed member of the Board of Governors of the Control Systems Society. He is also a member of the EUCA (European Union of Control Associations) Council.

Minyue Fu (SM'94) received the B.S. in electrical engineering from the China University of Science and Technology, Hefei, in 1982 and the M.S. and Ph.D. degrees in electrical engineering from the University of Wisconsin, Madison, in 1983 and 1987, respectively.

From 1983 to 1987, he held a teaching assistantship and a research assistantship at the University of Wisconsin, Madison. He worked as a Computer Engineering Consultant at Nicolet Instruments, Inc., Madison, during 1987. From 1987 to 1989, he served as an Assistant Professor in the Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI. During the summer of 1989, he was employed by the Universite Catholque de Louvain, Louvain, Belgium, as a Maitre de Conferences Invited. He joined the Department of Electrical and Computer Engineering, the University of Newcastle, Callaghan, Australia, in 1989. Currently, he is an Associate Professor and Head of the Department. His main research interests include control systems and signal processing.

Dr. Fu has been an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL and an Associate Editor for the Conference Editor Board of the IEEE Control Systems Society. He is currently an Associate Editor for the journal *Optimization and Engineering*.