Excitation Conditions for Signed Regressor Least Mean Squares Adaptation

W. A. SETHARES, MEMBER, IEEE, IVEN M. Y. MAREELS, MEMBER, IEEE, BRIAN D. O. ANDERSON, FELLOW, IEEE, C. RICHARD JOHNSON, JR., SENIOR MEMBER, IEEE, AND ROBERT R. BITMEAD, SENIOR MEMBER, IEEE

Abstract —**The stability of the signed regressor variant of least mean square (LMS) adaptation is found to be heavily dependent on the characteristics of the input sequence. Averaging theory is used to derive a persistence of excitation condition which guarantees exponential stability of the signed regressor algorithm. Failure to meet this condition (which is not equivalent to persistent excitation for LMS) can result in exponential instability, even with the use of leakage. This new persistence of excitation condition is then interpreted in both deterministic and stochastic settings.**

I. Introduction

IN HIGH DATA rate applications of adaptive filters such as speech processing, echo cancellation, and adaptive equalization, it is often important to maximize the speed at which the filters operate and/or to minimize the hardware requirements of the adaptive mechanism. One approach is to reduce the numerical complexity in the adaptive algorithm by coarsely quantizing certain signals. For example, in the well-known LMS adaptive filter [21] the coefficient update term is a scaled version of the product of two signals. Coarse (one bit) quantization of either of these signals converts the multiplications of LMS into single bit operations which are faster and simpler to implement. One such computationally simplified version of LMS is the signed regressor (SR) algorithm.

To be specific, consider the following adaptive finite impulse response (FIR) filtering task. With output $y(k)$ and input $u(k)$, an FIR filter can be defined by

$$y(k+1) = \sum_{i=1}^{n} b_i u(k-i+1) = X_k^T \theta^* \qquad (1.1)$$

where $\theta^* = (b_1, b_2, \cdots, b_n)^T$ represents the unknown parameters of the underlying FIR process and $X_k = (u(k), u(k-1), \cdots, u(k-n+1))^T$ is the regressor vector. The output $y(k+1)$ is to be estimated by

$$\hat{y}(k+1) = \sum_{i=1}^{n} \hat{b}_i(k) u(k-i+1) = \hat{\theta}_k^T X_k \qquad (1.2)$$

where $\hat{\theta}_k = [\hat{b}_1(k), \hat{b}_2(k), \cdots, \hat{b}_n(k)]^T$ is the parameter estimate vector. The error between the output $y(k)$ and the estimated output $\hat{y}(k)$ is called the prediction error. The traditional LMS algorithm is

$$\hat{\theta}_{k+1} = \hat{\theta}_k + \mu X_k [y(k) - \hat{y}(k)]. \qquad (1.3)$$

With $|u_k| < M$ and $0 < \mu < (2/nM^2)$, it is easily proven that $\hat{y}(k) \to y(k)$.

The use of one bit quantization on the input regressor replaces (1.3) by the signed regressor (SR) algorithm

$$\hat{\theta}_{k+1} = \hat{\theta}_k + \mu \operatorname{sgn}(X_k)[y(k) - \hat{y}(k)] \qquad (1.4)$$

where μ is positive and small. This update was first suggested in [19]. Note that the "sgn" function applied to a vector is an element by element operation. Thus

$$\operatorname{sgn}(X_k) = [\operatorname{sgn}(u(k)), \operatorname{sgn}(u(k-1)), \cdots, \operatorname{sgn}(u(k-n+1))]^T$$

where

$$\operatorname{sgn}(a) = \begin{cases} 1, & a > 1 \\ 0, & a = 0 \\ -1, & a < 0. \end{cases}$$

1.1. Previous Investigations

In the original examination of the SR algorithm in [19], the parameter errors of (1.4) were shown to be mean convergent under the assumptions that the entries of X_k are jointly Gaussian, that X_k is independent of X_j for $j < k$, and that μ is chosen small with respect to the maximum eigenvalue of $E\{XX^T\}$. The present paper extends this result by relaxing the independence assumption and considering other distributions.

More recent investigations in [13] and [4] focus on the degradation of convergence rate of the SR versus standard LMS. Unfortunately, the behavioral degradation can be more dramatic than just a decrease in the convergence rate. While the LMS algorithm moves the parameter estimates parallel to the regressor X (which leads on the

Manuscript received June 23, 1986; revised April 20, 1987 and January 19, 1988. The work of W. A. Sethares, I. M. Y. Mareels, and C. R. Johnson, Jr. was conducted while they were visiting the Australian National University, Canberra, ACT, Australia. The work of W. A. Sethares and C. R. Johnson, Jr. was supported by the National Science Foundation under Grant ECS-85-06149, under Grant INT-8513400, and under Grant MIP-8608787. I. M. Y. Mareels's work was supported by the National Fund for Scientific Research, Belgium. This paper was recommended by W. B. Mikhael and S. Stearns, Guest Editors for IEEE Transactions on Circuits and Systems, July 1987.

W. A. Sethares was with the School of Electrical Engineering, Cornell University, Ithaca, NY. He is now with the Department of Electrical and Computer Engineering, University of Wisconsin at Madison, Madison, WI 53706.

I. M. Y. Mareels was with the Research School of Physical Science, Australian National University, Canberra, ACT, Australia. He is now with the Department of System Engineering, University of Ghent, Belgium.

B. D. O. Anderson and R. R. Bitmead are with the Department of Systems Engineering Research School of Physical Science, Australian National University, Canberra, ACT, Australia.

C. R. Johnson, Jr. is with the School of Electrical Engineering, Cornell University, Ithaca, NY 14853.

IEEE Log Number 8820491.

average, in the direction of "steepest descent" of the squared error surface), the SR algorithm moves the parameter estimates parallel to the $\operatorname{sgn}(X)$ vector. As suggested in [8], the danger of instantaneous misalignment between $\operatorname{sgn}(X)$ and X is that the signed regressor parameter updates might actually climb (rather than descend) the error surface. A similar concern underlies the construction in [11] of a three periodic sequence that causes the parameter estimates of a related algorithm (the sign–sign variant) to diverge. The present paper documents and explains similar examples of divergence by the SR algorithm when excited by inputs that would cause LMS to converge.

Bershad [3] argues that the alignment of $\operatorname{sgn}(X)$ and X on the average is more important than occasional instantaneous misalignment, citing the convergence results in [19] for Gaussian inputs. Claasen and Mecklenbrauker [9] reply that input distributions which are not Gaussian may cause nonconvergent behavior given the "average" geometrical construct. Both are correct. The present paper confirms that Gaussian inputs are stabilizing and finds (non-Gaussian) distributions that cause the parameter estimate equation to be unstable.

Despite these potential problems, the SR algorithm has been found to be useful in certain applications. Cowan and Grant [10], for instance, constructed a hardware processor using the signed regressor scheme to implement a 64-point adaptive transversal filter that was used to successfully filter speech from periodic background noise.

1.2. A New Persistence of Excitation Condition

Given the uncertainty expressed in the literature regarding which inputs cause convergence and which cause misbehavior of the SR algorithm, this paper seeks conditions on the input $u(k)$ which guarantee that the parameter estimates in (1.2), when estimated by the SR update (1.4) will converge to the unknown parameters of the underlying FIR process in (1.1). Moreover, this convergence must be *robust* to small nonidealities such as measurement noise, mismodeling errors and slow variations of the desired parameterization. Typically, such conditions are called "persistence of excitation" conditions. The best known is the persistency of excitation condition for LMS (PE for LMS), which involves the positivity of a summed outer product of the regressor sequence, and which guarantees convergence of the parameter estimates $\hat{\theta}_k$ of (1.3) to the desired value θ^* [5]. The exponential character of this convergence imparts a degree of robustness to sufficiently small nonidealities [2].

This paper shows that the PE for LMS condition does not necessarily guarantee boundedness of the parameter estimates for the SR algorithm, but that a similar condition, involving the summed outer product of $\operatorname{sgn}(X)$ and X, does guarantee that $\hat{\theta}_k$ of (1.4) converges exponentially to θ^*. This new condition is called persistence of excitation for the signed regressor algorithm (PE for SR). The bulk of the paper examines the PE for SR condition and compares it to the standard PE for LMS condition. For instance, any sequence that is PE for SR is also PE for LMS, while the reverse implication is false. In certain cases, dissatisfaction of the PE for SR condition will actually cause exponential divergence of the parameter estimates $\hat{\theta}$ of (1.4). This instability should not be confused with divergence that is due to an inappropriately large stepsize, since it will occur for any positive stepsize μ less than some fixed μ^*. Such divergence is impossible for LMS.

To give an idea of the difficulty of the task of distinguishing inputs which are PE for SR from inputs that destabilize the SR algorithm, consider the following periodic input sequences to the three dimensional signed regressor algorithm.

Example 1: $u(k)=$ the 4 periodic sequence $\{1,1,1,-3\}$.

Example 2: $u(k)=$ the 5 periodic sequence $\{1,8,1,-5,-5\}$.

One of these fulfills the PE for SR condition and the parameter estimates $\hat{\theta}$ of (1.4) converge to θ^*. The other fails the PE for SR condition, and $\hat{\theta}$ diverges to $\pm\infty$ for any parameter set b_i, $i=1,2,3$, and any initial estimate (excluding an estimate which is exactly b_i), no matter how small the stepsize μ is chosen. Can you guess which example is divergent and which is convergent? (The answer will appear at the end of Section II.) Since both are PE for LMS, both cause exponential convergence of $\hat{\theta}$ of LMS (1.3) to θ^*. This divergent example also falsifies the "proof" of convergence of the SR algorithm in, e.g., [16], which does not involve any explicit constraints on the input other than boundedness.

1.3. An Overview of this Paper

Section II defines PE for SR. Averaging theory is used to prove that periodic regressor sequences which are PE for SR cause exponential convergence of the parameter estimates of the signed regressor algorithm. Theorem 1 also gives conditions under which the parameter estimates diverge exponentially. The puzzle of Examples 1 and 2 is then unraveled.

Having presented an analytical test that distinguishes destabilizing from stabilizing inputs to the SR algorithm, the focus shifts in Section III to a search for signal classes that satisfy the PE for SR condition and those that do not. Section III weaves a path through eight lemmas to answer the following questions: Is it always possible to find a zero mean t-periodic input sequence which is PE for SR? Is it always possible to find a zero mean t-periodic input sequence which causes the parameter estimates of the signed regressor algorithm to diverge? The answer to the first question is "yes," for t greater than the order n of the algorithm (for $n\geqslant 3$). Unfortunately, the answer to the second question is also "yes," for $t\geqslant 3$ and $n\geqslant 3$. Theorem 2 contains the precise statements. This result may be discomforting. In essence, it shows that the stability and instability properties of the signed regressor algorithm are intimately tied to surprisingly subtle characteristics of the input sequence. The focus is primarily on zero mean signals since sign changes are necessary for the PE for SR condition, since zero average signals are present in a wide variety of adaptive filtering applications, and to place the deterministic analysis on an equal footing with the stochastic analysis of later sections.

For LMS, the addition of a leakage factor provides an exponential "safety net" from which the parameter estimates cannot escape, even with (bounded) noises or nonidealities such as unmodeled dynamics or small nonlinearities. Perhaps the most surprising result for the signed regressor algorithm is presented in Section IV, where it is shown that exponential divergence of the parameter estimates is possible for certain inputs, even with the use of leakage.

Section V considers the signed regressor algorithm when excited by stochastic inputs. A stochastic PE for SR condition is derived in terms of the expectation of the outer product of $\operatorname{sgn}(X)$ and X, and Section VI shows that several common stochastic processes fulfill this condition. In particular, the SR algorithm is stable when excited by colored Gaussian inputs, and when excited by any process with independent and identically distributed moments. An example of a stochastic process which fails the stochastic PE for SR condition is given, and this process appears to cause divergent parameter estimates in simulation studies. Without doubt, the stability properties of the SR algorithm are crucially linked to the characteristics of the input/regressor sequence.

II. Persistency of Excitation for Deterministic Signals

In the ideal case (with no measurement noise or unmodeled disturbances), the LMS algorithm gives rise to the error system

$$\theta_{k+1} = \theta_k - \mu X_k X_k^T \theta_k \tag{2.1}$$

where $\theta_k = \theta^* - \hat{\theta}_k$ is the parameter error. Using the standard Lyapunov function $\theta^T\theta$ with suitably small μ, (2.1) can easily be shown to be Lyapunov stable. If there is a time interval t and an α such that

$$\lambda_{\min}\left\{\frac{1}{t}\sum_{i=j}^{j+t-1} X_i X_i^T\right\} > \alpha > 0, \qquad \text{for every } j$$

then the LMS algorithm is said to be persistently excited (PE for LMS), and the parameter estimates converge exponentially to their actual values, guaranteeing a certain robustness even in the nonideal case [5]. Equivalently, PE for LMS implies that (2.1) is exponentially asymptotically stable to the equilibrium $\theta = 0$. In this context, exponential stability means that there exists an $\alpha \in (0,1)$ and an $N < \infty$ such that $\|\theta_k\| \leqslant N\|\theta_0\|\alpha^k \ \forall k$.

The one step transition matrix for the error equation associated with standard LMS (2.1) is $I - \mu XX^T$. For the signed regressor algorithm (1.4), the error system is

$$\theta_{k+1} = \theta_k - \mu \operatorname{sgn}(X_k) X_k^T \theta_k \tag{2.2}$$

where $\theta_k = \theta^* - \hat{\theta}_k$ is the parameter error vector, and the one step transition matrix is $I - \mu \operatorname{sgn}(X)X^T$. Since the stability properties of LMS depend on the average of the outer product of XX^T, it is reasonable to conjecture that the behavior of the SR algorithm should depend on the average value of $\operatorname{sgn}(X)X^T$, at least when the adaptive gain μ is small.

The most straightforward results are obtained when the X_k sequence (or, equivalently, the $u(k)$ sequence) is periodic. Note that it is possible to cope with the almost periodic case by following the analysis in [1]. Accordingly, let $u(k)$ be a t-periodic sequence and define

$$M_t = \frac{1}{t}\sum_{i=j}^{j+t-1} \operatorname{sgn}(X_i) X_i^T. \tag{2.3}$$

Definitions: If $\operatorname{Re}\lambda_i(M_t) > 0$ for $i = 1, 2, \cdots, n$, then the regressor vector X_k will be said to be *persistently exciting for the signed regressor algorithm* (PE for SR). The notation $\operatorname{Re}\lambda_i(M_t)$ indicates the real part of the ith eigenvalue of the matrix M_t. Since X_k is composed of shifted versions of the scalar input u_k, the input sequence itself may also be called PE for SR. The matrix M_t will be called the *excitation matrix* for the signed regressor algorithm. △△△

The idea behind persistence of excitation (for both LMS and SR) is that it defines the class of signals *for a particular algorithm* which guarantee parameter convergence in the ideal case, and that this convergence is robust to small nonidealities. Exponential convergence is one way to guarantee robustness in the nonideal case [2].

The following theorem shows that the error system for the signed regressor algorithm is exponentially asymptotically stable when the input is PE for SR. Since the excitation matrix M_t can have eigenvalues with negative real parts, it is reasonable to conjecture that the SR algorithm will be exponentially unstable for any stepsize when M_t has an eigenvalue with negative real parts. This conjecture is true. This is in sharp contrast to LMS where Lyapunov stability always holds.

Theorem 1: Consider the signed regressor algorithm (1.4) and the associated error equation (2.2). If X_k is a t-periodic sequence that is PE for SR, that is, if $\operatorname{Re}\lambda_i(M_t) > 0$ for every i, then there exists a μ^* such that (2.2) is exponentially asymptotically stable for every $0 < \mu < \mu^*$. If X_k is not persistently exciting, and $\operatorname{Re}\lambda_i(M_t) < 0$ for some i, then $\exists \mu^*$ such that for every $0 < \mu < \mu^*$, (2.2) is exponentially unstable.

Proof: Because the input is t-periodic, the stability properties are determined by the eigenvalues of the t-step transition matrix for (2.2) which is given by

$$A(l+t-1, l) = \prod_{j=l}^{l+t-1}\left(I - \mu \operatorname{sgn}(X_j) X_j^T\right).$$

By adding and subtracting $\mu t M_t$, where M_t is the excitation matrix, this can be approximated as

$$A(l+t-1, l) = I - \mu t M_t + B(l+t-1, l)$$

where $B(\cdot,\cdot)$ is $o(\mu t)$

and

$$B(l+t-1,l) = \mu^2 \sum_{k_1=l}^{l+t-1} \sum_{k_2=k_1}^{l+t-1} \operatorname{sgn}(X_{k_1}) X_{k_1}^T \operatorname{sgn}(X_{k_2}) X_{k_2}^T + \cdots + \mu^t \operatorname{sgn}(X_{l+t-1}) X_{l+t-1}^T \cdots \operatorname{sgn}(X_l) X_l^T.$$

The norm of the error $B(\cdot,\cdot)$ (following the logic of [7]) may be bounded as

$$\|B(l+t-1,l)\| \leqslant \sum_{k=0}^{t-2} t/(2+k)! \mu^{2+k} t^{2+k} \cdot \|\operatorname{sgn}(X)\|^{2+k} \|X\|^{2+k} \leqslant \frac{1}{2}(\mu t \|\operatorname{sgn}(X)\| \|X\|)^2 \cdot \exp(\mu t \|\operatorname{sgn}(X)\| \|X\|)$$

if μt is sufficiently small. It follows that as $\mu \to 0$,

$$\lambda_i[A(l+t-1,l)] \to \lambda_i[I - \mu t M_t] = 1 - \mu t \lambda_i(M_t).$$

Accordingly, if $\operatorname{Re}\lambda_i(M_t) > 0$ for all i, $|\lambda_i[A(l+t-1,l)]| < 1$ for all i when μ is suitably small. If $\operatorname{Re}\lambda_i(M_t) < 0$ for some i, then $|\lambda_i[A(l+t-1,l)]| > 1$. The stability and instability results then follow. △△△

This theorem explains (and proves) the behavior of the SR algorithm in the examples of the introduction. For Example 1, the input is PE for SR since

$$M_4 \equiv \frac{1}{4}\sum_{i=1}^{4} \operatorname{sgn}(X_i) X_i^T = \frac{1}{4}\begin{pmatrix} 6 & -2 & -2 \\ -2 & 6 & -2 \\ -2 & -2 & 6 \end{pmatrix}$$

which has eigenvalues 1/2, 2, and 2. For Example 2, the matrix M_t has an eigenvalue at -0.86.

The next section presents results which help to classify input sequences that are PE for SR, and input sequences for which the excitation matrix has eigenvalues with negative real parts, causing instability of the SR algorithm.

III. Interpretation of Deterministic PE for SR

The persistency of excitation condition for LMS has an intuitively appealing meaning in terms of the number of sinusoids present in the input. Unfortunately, no such simple interpretation (in terms of spectral complexity) is possible for the PE for SR condition, since an input which consists of any number of sinusoids in conjunction with a large dc bias (i.e., any input with no sign changes) has a rank one excitation matrix. Another interpretation of PE for LMS is as a spanning condition on a t-periodic regressor vector sequence X_k, that is, $\Sigma_{k=j}^{j+t} X_k X_k^T > 0$ if and only if the vectors $X_j, X_{j+1}, \cdots, X_{j+t}$ span $\mathbb{R}^n$ for every j. Unfortunately, no such simple interpretation on the spanning properties of $\operatorname{sgn}(X_j)$ or X_j is equivalent to PE for SR. What, then, can be said about the PE for SR condition?

Lemmas 1 and 2 find that spanning conditions on $\operatorname{sgn}(X)$ and X are necessary (but not sufficient) for exponential stability of the signed regressor algorithm. These show that any sequence that is PE for SR is also PE for LMS, while examples 3 and 4 show that the reverse implications are, in general, false. Lemma 3 (and surrounding discussion) demonstrate that for the simple one- and two-dimensional SR algorithms, the spanning property on $\operatorname{sgn}(X)$ is sufficient for stability.

Lemma 4 provides a technical result (relating the eigenvalues of $M_t + M_t^T$ to the eigenvalues of M_t) which is used in Lemma 5 to show that there are t-periodic zero mean sequences which are PE for SR whenever $t > n \geqslant 3$, where n is the dimension of the algorithm. Lemmas 6, 7, and 8 then establish the instability results. First, a single class of 3-periodic inputs is shown to destabilize any SR algorithm of dimension 3 or higher. This is extended to show that for almost any $t > n \geqslant 3$, such destabilizing inputs exist. The results of all the lemmas are finally gathered together in Theorem 2, and extended to include certain nonzero mean sequences.

3.1. Two Necessary Spanning Conditions

For t-periodic inputs, a simple condition which is necessary (but not sufficient) for exponential stability of the signed regressor algorithm is that

$$\frac{1}{t}\sum_{i=1}^{t} \operatorname{sgn}(X_i)\operatorname{sgn}(X_i)^T > 0. \tag{3.1}$$

This requires that the sign of the regressor vector span $\mathbb{R}^n$ every period, and is a relatively easy condition to check. Formally, we have the following.

Lemma 1: Any sequence that is PE for SR must fulfill (3.1).

Proof: Suppose that condition (3.1) does not hold. Then there is a nonzero constant vector c such that $c^T \operatorname{sgn}(X_i) = 0$ for every $i \in [1,t]$. This implies that

$$c^T M_t = c^T \frac{1}{t}\sum_{i=1}^{t} \operatorname{sgn}(X_i) X_i^T = \frac{1}{t}\sum_{i=1}^{t} c^T \operatorname{sgn}(X_i) X_i^T = 0$$

where M_t is the excitation matrix. Hence c^T is a null vector of M_t, M_t is singular, and X_k is not PE for SR. △△△

Another condition which is necessary for exponential stability of the SR algorithm is that the regressor vector span $\mathbb{R}^n$ every period.

Lemma 2: Any sequence that is PE for SR is also PE for LMS. In symbols,

$$\operatorname{Re}\lambda_j\left(\sum_{i=1}^{t} \operatorname{sgn}(X_i) X_i^T\right) > 0, \qquad \forall j$$

$$\text{implies that } \sum_{i=1}^{t} X_i X_i^T > 0.$$

Proof: By contradiction. Suppose $\lambda_i(\Sigma X X^T) = 0$ for some i. Then there exists a nonzero vector c such that $c^T X_k = 0 \ \forall k \in [1,t]$ which implies $(\Sigma \operatorname{sgn}(X) X^T)c = 0$ which implies that $\operatorname{Re}\lambda_i(\Sigma \operatorname{sgn}(X) X^T) = 0$. △△△

To see that the reverse implications do not hold (and hence that the spanning conditions of Lemmas 1 and 2 are not, in general, sufficient for exponential stability), consider the following simple periodic examples:

Example 3: $n=2$, $t=2$, $u(k)=$ the two periodic sequence $\{1/2,1\}$ for which $\Sigma XX^T>0$ but $\Sigma \operatorname{sgn}(X)X^T$ and $\Sigma \operatorname{sgn}(X)\operatorname{sgn}(X)^T$ each have a zero eigenvalue.

Example 4: $n=7$, $t=7$, $u(k)=$ the seven periodic sequence $\{2,4,4,2,-1,-10,-1\}$ for which $\Sigma \operatorname{sgn}(X)\operatorname{sgn}(X)^T>0$ but ΣXX^T has a zero eigenvalue and $\Sigma \operatorname{sgn}(X)X^T$ has two negative eigenvalues.

It is thus strictly more difficult to persistently excite the signed regressor algorithm than to persistently excite LMS.

We now consider several simple cases.

3.2. One- and Two-Dimensional SR

In one dimension, the PE for SR requirement is trivially satisfied for all nonvanishing inputs since $\Sigma \operatorname{sgn}(X_i)X_i^T=\Sigma|u_i|$. In two dimensions, the necessary spanning condition of Lemma 1 on $\operatorname{sgn}(X)$ is also sufficient for exponential stability of the SR algorithm.

Lemma 3: The error equation (2.2) for the two-dimensional signed regressor algorithm (1.4) is exponentially stable for any t-periodic input sequence provided that $\Sigma_{i=1}^{t}\operatorname{sgn}(X_i)\operatorname{sgn}(X_i)^T>0$, and provided that μ is sufficiently small.

Proof: Direct computation shows that

$$tM_t=\sum_{i=1}^{t}\operatorname{sgn}(X_i)X_i^T=\begin{vmatrix}\sum_{i=1}^{t}|u_i| & \sum_{i=1}^{t-1}\operatorname{sgn}(u_i)u_{i+1}+\operatorname{sgn}(u_t)u_1\\ \sum_{i=1}^{t-1}\operatorname{sgn}(u_{i+1})u_i+\operatorname{sgn}(u_1)u_t & \sum_{i=1}^{t}|u_i|\end{vmatrix}=\begin{vmatrix}\alpha & \beta\\ \delta & \alpha\end{vmatrix}$$

If $\Sigma \operatorname{sgn}(X)\operatorname{sgn}(X)^T$ has a zero eigenvalue, then either

(i) u_i has the same sign for every i,
(ii) u_i alternates in sign, or
(iii) u_i is identically zero,

since any other sign pattern will cause $\operatorname{sgn}(X_1)$, $\operatorname{sgn}(X_2),\cdots,\operatorname{sgn}(X_t)$ to span $\mathbb{R}^2$. Case (iii) is trivial, since M_t is the zero matrix. If either (i) or (ii) holds, then $\beta=\delta$ and $|\beta|=\alpha$ implying that M_t has a zero eigenvalue. If neither (i) nor (ii) hold, then for some i and j, the sign of $\operatorname{sgn}(u_i)u_{i+1}$ is different from the sign of $\operatorname{sgn}(u_j)u_{j+1}$ and so $|\beta|<\alpha$. Similarly, $|\delta|<\alpha$. M_t is therefore, diagonally dominant which implies that all eigenvalues have positive real parts. △△△

Actually, somewhat more is true. Dasgupta and Johnson [12] found a Lyapunov function which proves that the two-dimensional SR algorithm is never unstable. Unfortunately, the stability properties of the one- and two-dimensional algorithms do not extend to higher dimensions.

3.3. Stability Results for n-Dimensional SR

The input sequences of Examples 2 and 4 cause instability of the $n=3$ and $n=7$ dimensional SR algorithm. Is such instability generic? Do there exist input sequences which are PE for SR for arbitrary t and n? Lemmas 4 and 5 guarantee (by construction) that there are t-periodic sequences which are PE for SR whenever $t>n$.

If $t<n$, then $\Sigma \operatorname{sgn}(X)\operatorname{sgn}(X)^T$ has a zero eigenvalue, since at least two rows are identical. If $t=n$, it is also impossible for a zero mean sequence to be PE for SR. To see this, define $1_t=[1,1,\cdots,1]^T\in\mathbb{R}^t$ and suppose that $\Sigma u_i=0$ (where, as usual, the sum is taken over one period). Then, $1_t^TX_j=0$ for every j. Hence

$$tM_t1_t=\left\{\sum_{i=1}^{t}\operatorname{sgn}(X_i)X_i^T\right\}1_t=\sum_{i=1}^{t}\operatorname{sgn}(X_i)\{X_i^T1_t\}=0 \tag{3.2}$$

and M_t has a zero eigenvalue corresponding to the eigenvector 1_t. This shows that zero mean sequences require $t>n$ in order to be PE for SR.

To proceed, the following technical lemma is useful. It translates the PE for SR condition on M_t to a sufficient (but not necessary) condition on the symmetric matrix $M_t+M_t^T$.

Lemma 4: If $M_t+M_t^T>0$, then $\operatorname{Re}\lambda_{\min}(M_t)>0$.

Proof: Consider the system $\dot{x}=M_tx$ which has a candidate Lyapunov function $V=x^Tx$. Then $\dot{V}=x^T\dot{x}+\dot{x}^Tx=x^T(M_t+M_t^T)x>0$. Hence every mode of M_t is unstable and so $\operatorname{Re}\lambda(M_t)>0$. The reverse implication is false, even with zero mean periodic X_i. An example is $n=4$, $t=7$, with periodic input $\{-20,0.1,0.1,0.1,-0.1,19.9,-0.1\}$. △△△

Now suppose that the period t is greater than the dimension n of the regressor. Define the extended information vector $Z_k=[u_k,u_{k-1},\cdots,u_{k-t+1}]^T$. Then

$$R_t=\frac{1}{t}\sum_{i=1}^{t}\operatorname{sgn}(Z_i)Z_i^T \text{ contains } M_t=\frac{1}{t}\sum_{i=1}^{t}\operatorname{sgn}(X_i)X_i^T$$

as its leading $n\times n$ square submatrix. Consider a zero mean sequence with the following sign pattern: $u_1>0$, and $u_i<0$ for $i=2,3,\cdots t$. Then $(1/2)(R_t+R_t^T)\in\mathbb{R}^{t\times t}$ can be written directly as

$$\frac{1}{t}\begin{pmatrix}2\sum_{i=2}^{t}|u_i| & u_2+u_t & u_3+u_{t-1} & \cdots & u_t+u_2\\ u_2+u_t & 2\sum_{i=2}^{t}|u_i| & u_2+u_t & \cdots & u_{t-1}+u_3\\ \cdot & \cdot & \cdot & \cdots & \cdot\\ \cdot & \cdot & \cdot & \cdots & \cdot\\ \cdot & \cdot & \cdot & \cdots & \cdot\\ u_t+u_2 & u_{t-1}+u_3 & u_{t-2}+u_4 & \cdots & 2\sum_{i=2}^{t}|u_i|\end{pmatrix} \tag{3.3}$$

where $(1/2)(M_t + M_t^T) \in \mathbb{R}^{n\times n}$, the leading $n \times n$ square submatrix, is diagonally dominant and hence positive definite. Lemma 4 then implies that $\mathrm{Re}\,\lambda(M_t) > 0$, and so any zero mean periodic sequence with this sign pattern will be PE for SR of any dimension up to $t-1$. This proves the following.

Lemma 5: Consider the n-dimensional signed regressor algorithm (1.4).

(i) There are periodic zero mean sequences of every period $t > n$ which are PE for SR and
(ii) There are no periodic zero mean sequences of period $t \leqslant n$ which are PE for SR. △△△

3.4. Instability Results for n-Dimensional SR

Statement (ii) of the previous lemma shows that short zero mean periodic sequences may cause nonconvergent behavior of the SR algorithm. The situation is even worse if the zero mean assumption is relaxed. The next result demonstrates a class of "short" destabilizing sequences.

Lemma 6: Consider the 3-periodic input sequence u_1, u_2, u_3 with $u_1 > 0$, $u_2 < 0$, and $u_3 < 0$. If $|u_2| + |u_3| < |u_1|$, then the error system (2.2) of the n-dimensional signed regressor algorithm (1.4) will be unstable for any $n > 2$, provided μ is suitably small.

Proof: Since $\mathrm{Re}\,\lambda_i(\Sigma\,\mathrm{sgn}(X)X^T) < 0$ if and only if $\mathrm{Re}\lambda_i(\Sigma\,\mathrm{sgn}(cX)(cX)^T) < 0$ for any $c \neq 0$, it suffices to consider the normalized input sequence $u_1 = 1$, $u_2 = -a$, $u_3 = -b$ with a and b positive. In the three-dimensional case, the eigenvalues of $M_3 = \Sigma\,\mathrm{sgn}(X)X^T$ are the roots of the polynomial $s^3 - 3\alpha s^2 + (3\alpha - 3\delta\beta) + 3\delta\beta\alpha - \beta^3 - \delta^3 - \alpha^3$ where $\alpha = 1 + a + b$, $\beta = -1 - a - b$ and $\delta = -1 + a - b$. By the Routh test, M_3 will have negative eigenvalues whenever $\beta^3 + \delta^3 + \alpha^3 - 3\delta\beta\alpha < 0$. In terms of a and b, this is equivalent to $a^3 + b^3 + 3ab - 1 < 0$. This cubic can be factored as $(a + b - 1)(a^2 + b^2 - ab + a + b + 1)$. The second of these factors is positive for every positive a and b, since it is equal to $(a-b)^2 + ab + a + b + 1$. The inequality is then true whenever $a + b - 1 < 0$. This proves the result in three dimensions.

The n-dimensional excitation matrix M_n contains M_3 as its leading three-dimensional submatrix and has rank at most 3 since the i, jth element of M_n is equal to the $i(\mathrm{mod}\,3)$, $j(\mathrm{mod}\,3)$th element of M_n where $i(\mathrm{mod}\,3) \in \{1,2,3\}$.

Let

$$T_{ij} = \begin{Bmatrix} 1, & \text{if } i = j \\ -1, & \text{if } i - j = 3k \text{ for some integer } k \\ 0, & \text{otherwise} \end{Bmatrix}$$

and partition M_3 in columns as and $[C_1|C_2|C_3]$. Then $(TM_nT^{-1})_{ij} = 0$ for $i > 3$ and the leading three-dimensional minor of TM_nT^{-1} is $[\alpha C_1|\beta C_2|\delta C_3]$ where

$$\alpha = \left[\frac{n+2}{3}\right], \quad B = \left[\frac{m+1}{3}\right], \quad \delta = \left[\frac{n}{3}\right]$$

and $[*]$ denotes the greatest integer $\leqslant *$. But $\det[\alpha C_1|\beta C_2|\delta C_3] = \alpha\beta\delta\det(M_3)$ and thus M_n will have negative determinant whenever M_3 has negative determinant. Thus M_n has an eigenvalue with negative real part whenever $|u_1| > |u_2| + |u_3|$. △△△

Lemma 1 showed that any sequence that is PE for SR must fulfill (3.1), that is, must have $\Sigma\,\mathrm{sgn}(X)\,\mathrm{sgn}(X)^T > 0$. This is essentially a requirement on the sign pattern of the input. For t-periodic sequences, there are t^2 possible sign patterns, but only a few of these need to be examined. Observe:

(i) $\Sigma\,\mathrm{sgn}(X)\,\mathrm{sgn}(X^T)$ and $\Sigma\,\mathrm{sgn}(X)X^T$ are even functions of X. Thus an input with the sign pattern $+ + + -$ will have the same stability properties as one with the pattern $- - - +$.
(ii) $\Sigma\,\mathrm{sgn}(X)\,\mathrm{sgn}(X)^T = \Sigma P\,\mathrm{sgn}(X)\,\mathrm{sgn}(X)^T P^T$ and $\Sigma\,\mathrm{sgn}(X)X^T = \Sigma P\,\mathrm{sgn}(X)X^T P^T$ where P is the unitary matrix:

$$P = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

This shows that the stability properties of a sign pattern are invariant under circular shifts, for instance, the sign pattern $+ - - -$ will have the same properties as $- - + -$. We will say that two sign patterns are distinct if they cannot be equated by any combination of shifts and multiplications by -1.

3.5. The Three-Dimensional SR

This subsection considers in detail the stability/instability properties of the three-dimensional SR algorithm when excited by t-periodic zero mean sequences. Lemma 5 shows that no sequence of period less than four can be PE for SR. There are only three distinct sign patterns of four periodic sequences; $+ + + -$, $+ - + -$, and $+ + - -$. It is easy to check that the second and third patterns have $\mathrm{Re}\,\lambda(\Sigma\,\mathrm{sgn}(X)\,\mathrm{sgn}(X)^T) = 0$, and so cannot be PE for SR by Lemma 1. The first pattern is PE for SR, as implied by the proof of Lemma 5. This is how Example 1 was generated.

The situation for five periodic sequences is more complicated, and at the same time more typical of the general case. There are only three distinct sign patterns for which $\Sigma\,\mathrm{sgn}(X)\,\mathrm{sgn}(X)^T > 0$. These are $+ + + + -$, $+ + + - -$, and $+ + - + -$. The first is PE for SR (again by the proof of Lemma 5). Both other patterns may have eigenvalues with negative real part, leading to instability. For $u_1, u_2, u_3 > 0$ and $u_4, u_5 < 0$ with $\Sigma_{i=1}^5 u_i = 0$, it is easy to calculate that

$$(1/2)\sum_{i=1}^{5}\left[\mathrm{sgn}(X_i)X_i^T + X\,\mathrm{sgn}(X_i^T)\right]$$
$$= \begin{pmatrix} 2(u_1+u_2+u_3) & u_2 & u_4+u_5-u_2 \\ u_2 & 2(u_1+u_2+u_3) & u_2 \\ u_4+u_5-u_2 & u_2 & 2(u_1+u_2+u_3) \end{pmatrix}.$$

Let $\alpha = 2(u_1 + u_2 + u_3)$ and $\beta = u_2$. Then $u_4 + u_5 - u_2 = -\beta - 1/2\alpha$. The above matrix can be rewritten as

$$\alpha \begin{pmatrix} 1 & \delta & -1/2-\delta \\ \delta & 1 & \delta \\ -1/2-\delta & \delta & 1 \end{pmatrix} \quad \text{where } \delta = \beta/\alpha.$$

Note that $0 < \delta < 1/2$. For $\delta < 1/4$, this matrix is diagonally dominant, and hence the input is PE for SR. Larger δ, however, can cause this matrix to have negative eigenvalues. For instance, $\delta = 0.4$ generates the divergent example of the introduction. The third sign pattern can also be reduced to a single parameter family of matrices which can be readily examined.

To generalize this unstable example to t-periodic input sequences (with $t > 5$), consider the $j + k$ periodic zero mean input sequence $a_1, a_2, \cdots, a_j, b_1, b_2, \cdots, b_k$ where $a_i > 0$ and $b_j < 0$ for every i. Let $\alpha = \Sigma_i |a_i| + \Sigma_i |b_i|$. The excitation matrix can be written

$$\frac{1}{t}\begin{pmatrix} \alpha & \alpha - 2a_1 + 2b_1 & \alpha - 2a_1 - 2a_2 + 2b_1 + 2b_2 \\ \alpha - 2a_j + 2b_k & \alpha & \alpha - 2a_1 + 2b_1 \\ \alpha - 2a_j - 2a_{j-1} + 2b_k + 2b_{k-1} & \alpha - 2a_j + 2b_k & \alpha \end{pmatrix}.$$

This matrix is independent of the particular values of a_3 through a_{j-2} and b_3 through b_{k-2} because of the zero mean property. For instance, the input sequence $\{1, 2, 2, 2, 1, -1/2, -2.5, -2, -2.5, -1/2\}$ generates the symmetric matrix

$$M(\hat{\alpha}) = \frac{1}{10}\sum_{i=1}^{10} \operatorname{sgn}(X_i) X_i^T = \begin{pmatrix} \hat{\alpha} & \hat{\alpha} - p & \hat{\alpha} - q \\ \hat{\alpha} - p & \hat{\alpha} & \hat{\alpha} - p \\ \hat{\alpha} - q & \hat{\alpha} - p & \hat{\alpha} \end{pmatrix}$$

where $\hat{\alpha} = 1.6$, $p = 0.3$, and $q = 1.2$. $M(\hat{\alpha})$ has a negative real eigenvalue. Now the input sequence can be lengthened by adding any number of intermediate zero mean terms (a_3 through a_{j-2} and b_3 through b_{k-2}). The only effect on the matrix M will be to increase the value of $\hat{\alpha}$.

Lemma 7: There are zero mean t-periodic sequences for every $t > 8$ which destabilize the three dimensional SR algorithm.

Proof: Direct computation shows that $\det M(\alpha) = (4pq - q^2)\alpha - 2qp^2$. Since p, q, and α are positive, $\det M(\alpha) < 0$ whenever $4pq - q^2 \leqslant 0$. This implies that $M(\alpha)$ has at least one eigenvalue with negative real part. Note that it is easy to choose a_1, a_2, a_{j-1}, a_j, b_1, b_2, b_{k-1}, and b_k so that $M(\alpha)$ is symmetric. △△△

These divergent results may be generalized to the nth order signed regressor algorithm.

Lemma 8: Consider again the $j + k$ periodic zero mean input sequence as above. If the n-dimensional excitation matrix is symmetric and has a negative eigenvalue, then either,

(i) the $(n+1)$-dimensional excitation matrix has a negative eigenvalue, or
(ii) there is a $j + k + 4$ periodic zero mean input sequence for which the $(n+1)$-dimensional excitation matrix is symmetric and has a negative eigenvalue.

Proof: The hypothesis forces the n-dimensional excitation matrix to have the following Toeplitz form:

$$M^n(\alpha) = \frac{1}{j+k}\sum_{i=1}^{j+k} \operatorname{sgn}(X_i) X_i^T = \frac{1}{j+k}\begin{pmatrix} \alpha & \alpha - \beta_1 & \alpha - \beta_2 & \cdots & \alpha - \beta_{n-1} \\ \alpha - \beta_1 & \alpha & \alpha - \beta_1 & \cdots & \alpha - \beta_{n-2} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \alpha - \beta_{n-1} & & & & \alpha \end{pmatrix}$$

where $\alpha = \Sigma_i |a_i| + \Sigma_i |b_i|$ and the β_i depend on $a_1, \cdots, a_{n-1}$, $a_{j-n+2}, \cdots, a_j$, $b_1, \cdots, b_{n-1}$, and $b_{k-n+2}, \cdots, b_k$, but do not depend on $a_n, \cdots, a_{j-n+1}$ or $b_n, \cdots, b_{k-n+1}$. Let $M^{n+1}(\alpha)$ be the $(n+1)$-dimensional excitation matrix. Then $M^n(\alpha)$ is the principle leading $n \times n$ submatrix of $M^{n+1}(\alpha)$. Clearly, $M^{n+1}(\alpha)$ cannot be positive definite, since all minors of a positive definite matrix must have positive determinant. If $M^{n+1}(\alpha)$ is symmetric, this establishes (i). If not, then add 4 new terms to the input sequence, in the a_{n+1}, a_{j-n}, b_{n+1}, and b_{k-n} positions. Pick these terms so that their sum is 0. This implies that $M^{n+1}(\alpha)$ is symmetric. Then $M^n(\alpha)$ again has a negative eigenvalue, and is the n dimensional leading submatrix of $M^{n+1}(\alpha)$. This implies (ii). △△△

A simple inductive argument, beginning with $n = 3$, and using Lemmas 7 and 8 then shows that for every n, there are t-periodic sequences of sufficient length which cause instability of the SR algorithm.

3.6. Consolidation

The stability and instability results of the previous sections may now be gathered together.

Theorem 2: Consider the n dimensional signed regressor algorithm (1.4) excited by t-periodic zero mean input sequences with suitably small stepsize.

(i) The case $n = 1$ is exponentially stable whenever the input is nonvanishing.
(ii) The case $n = 2$ is exponentially stable whenever $\Sigma \operatorname{sgn}(X) \operatorname{sgn}(X)^T > 0$.
(iii) In general, for $n > 2$, there exist t-periodic zero mean sequences of any length $t > n$ which are PE for SR, and, there exist t-periodic zero mean sequences of sufficient length which destabilize the algorithm and lead to exponential divergence of the parameter estimates.
(iv) Moreover, the above examples of convergence and divergence are robust with respect to the zero mean assumption, that is, the same results are true if the

hypothesis that the input sequence has zero mean is replaced by the hypothesis that the input sequence has ϵ mean, for sufficiently small ϵ.

Proof: The only part of the theorem which remains unproven is the extension to ϵ mean. In Lemma 5, an extension to ϵ mean will not disturb the diagonal dominance, provided $n\epsilon$ is smaller than the smallest u_i. (There may, however, be periodic ϵ mean sequences with $t \leqslant n$ which are persistently exciting.) In Lemma 7, introduction of ϵ mean will perturb the values of p and q by ϵ. For small enough ϵ, this will not change the sign of $(4pq - q^2)\alpha - 2p^2q$. Similarly, in Lemma 8, perturbing the values of β_i by ϵ will not effect the sign of the eigenvalues of $M_n(\alpha)$. △△△

It is difficult to succinctly interpret the engineering significance of the persistence of excitation condition for the signed regressor algorithm. If a particular application (for example, adaptive FIR identification) allows a free choice of input sequences, then the deterministic conditions derived here could be used to choose an input sequence for a particular order regressor so as to guarantee exponential convergence. In many applications, the input sequence is not directly under user control. It is then important to have an idea of the probability that an input will destabilize the algorithm. Sections V and VI show that if certain statistical properties characterize the input, then stability can be assured.

IV. Leakage

A leakage factor Λ is often incorporated into adaptive algorithms in order to add robustness during quiescent periods when the degree of persistence of excitation becomes small [14]. For the standard LMS, this is

$$\hat{\theta}_{k+1} = (1-\Lambda)\hat{\theta}_k + \mu X_k e_k$$

and the parameter error equation is

$$\theta_{k+1} = \left[(1-\Lambda)I - \mu X_k X_k^T\right]\theta_k + \Lambda\theta^*.$$

Since ΣXX^T is always at least positive semi-definite, the one step transition matrix is uniformly a contraction, with magnitude depending on Λ. This means that initial condition effects die away exponentially. The parameter error θ_k does not, however, tend to zero, due to the $\Lambda\theta^*$ term. The limiting bias in θ_k is proportional to Λ; this is the penalty for guaranteeing stability in the presence of bounded measurement noise and unmodeled dynamics, see [20].

For the signed regressor algorithm, the addition of a leakage factor does not always guarantee stability. The error equation for the SR algorithm with leakage is

$$\theta_{k+1} = \left((1-\Lambda)I - \mu\,\mathrm{sgn}(X_k)X_k^T\right)\theta_k + \Lambda\theta^*. \quad (4.1)$$

The stability of (4.1) (for t-periodic inputs) is determined by the eigenvalues of the matrix $R_t = \Lambda I + \mu M_t$ where M_t is the excitation matrix $\Sigma\,\mathrm{sgn}(X)X^T$. It is easy to imagine that if M_t has an eigenvalue with negative real part, and if Λ is small compared to μ, then R_t may also have an eigenvalue with negative real part. The signed regressor algorithm will then be exponentially unstable, despite the addition of a leakage term.

Theorem 3: Consider the signed regressor algorithm with leakage and the associated error system (4.1). Suppose that a t-periodic input sequence causes the excitation matrix M_t to have an eigenvalue with negative real part (as in part (iii) of Theorem 2). Then $\exists\mu^*$ and a Λ^* such that $\forall 0 < \mu < \mu^*$ and $\forall 0 < \Lambda < \Lambda^*$, (4.1) is exponentially unstable.

Proof: The eigenvalues of R_t can be shown to determine the stability/instability of (4.1) by mimicking the proof of Theorem 1. Let P be the matrix of eigenvectors that transforms M_t into Jordan form. Then R_t is similar to

$$P^{-1}(\Lambda I + \mu M_t)P = \Lambda I + \mu\,\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) = \mathrm{diag}(\Lambda + \mu\lambda_1, \cdots, \Lambda + \mu\lambda_n)$$

where the λ_i are the eigenvalues of M_t. Thus whenever $\Lambda + \mu\lambda_i < 0$ for some i, R_t will have an eigenvalue with negative real part. △△△

V. Persistency of Excitation for Stochastic Inputs

In the deterministic case, the appropriate persistence of excitation condition for the signed regressor algorithm was that the eigenvalues of $\Sigma\,\mathrm{sgn}(X)X^T$ have, on the average, positive real parts. The analogous condition for stochastic inputs involves the eigenvalues of $E\{\mathrm{sgn}(X)X^T\}$. Suppose that $u(k)$ and hence X_k are stationary random processes with finite moments up to order m.

Definitions: If $\mathrm{Re}\,\lambda_i(E\{\mathrm{sgn}(X)X^T\}) > 0$ for $i = 1, 2, \cdots, n$ then the regressor vector X_k will be said to be *persistently exciting for the signed regressor algorithm* (PE for SR). Since X_k is composed of shifted versions of the scalar input $u(k)$, the input sequence itself may also be called PE for SR. The matrix $E\{\mathrm{sgn}(X)X^T\}$ will be called the *excitation matrix* for the signed regressor algorithm.

As in the deterministic case, persistence of excitation implies exponential stability.

Theorem 4: The error system (2.2) for the signed regressor algorithm (1.4) is exponentially stable when:

(i) the real parts of the eigenvalues of $E\{\mathrm{sgn}(X)X^T\}$ are positive,
(ii) the stepsize is small enough,
(iii) the input process fulfills a mixing condition (stated precisely in (5.4) below).

Proof: Recall that the system (2.2) has the m-term transition matrix

$$A(k+m, k) = \prod_{l=k}^{k+m-1}\left(I - \mu\,\mathrm{sgn}(X_l)X_l^T\right). \quad (5.1)$$

The assumption on the moments of X_k is that there are constants M and $\delta < \infty$ such that $E\|(Z_k)^i\|_P < M\delta^i$, $i = 1, 2, \cdots, m$ where $Z_k = \mathrm{sgn}(X_k)X_k^T$ and the P norm is the induced norm

$$\|A\|_P = \max_{x^TPx=1} x^TA^TPAx$$

where P is defined to be the unique positive definite solution to the Lyapunov equation $E\{\text{sgn}(X)X^T\}^T P + PE\{\text{sgn}(X)X^T\} = I$. Henceforth, all norms in this proof are P norms. (5.1) can be rewritten as

$$A(k+m,m) = I - \mu m E\{\text{sgn}(X_k)X_k^T\} + \mu m P_k + Q_k \tag{5.2}$$

where

$$P_k = \frac{1}{m}\sum_{i=1}^{m} E\{\text{sgn}(X)X^T\} - Z_{k+i-1}$$

and

$$Q_k = \mu^2 \sum_{i=1}^{m}\sum_{j=1}^{m} Z_{k+j-1}Z_{k-i+1} - \mu^3 \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{l=1}^{m} Z_{k+l-1}Z_{k+j-1}Z_{k+i-1} + \cdots + (-\mu)^m Z_{k+m-1}\cdots Z_k.$$

The proof uses three results from [6]. First, a sufficient condition for exponential convergence of a system with transition matrix $A(k+m,k)$ is that there exist an $m>0$ and an induced norm $\|\cdot\|$ such that

$$E\{\|A(k+m,k)\|\} < 1. \tag{5.3}$$

In this context, exponential convergence means the existence of an ω-dependent almost surely finite N and an ω-independent $\alpha \in (0,1)$ for which $\|\theta_k\| < N\alpha^k$.

Second, by virtue of the assumptions on the moments of Z_k, and provided $\mu m\delta < 1$, the norm of the expected value of Q_k can be bounded by $E\|Q_k\| < 3M(\mu m \delta)^2$.

The third is the following: Let B_l be a stationary process with mean $E\{B\}$, and let $f^{ij}(\lambda)$ be the spectral density of the $(i-j)$th component of $B_l - E\{B\}$. If $f^{ij}(\lambda)$ exists and is twice differentiable at $\lambda = 0$, then

$$E\left\|\frac{1}{m}\sum_{l=k}^{k+m-1}(B_l - E\{B\})\right\| = O(m^{-1/2}) \tag{5.4}$$

uniformly in k.

To continue to proof, let $B_l = \text{sgn}(X_l)X_l^T$ and $E\{B\} = E\{\text{sgn}(X)X^T\}$. Taking norms on (5.2) shows that

$$E\|A(k+m,k)\| \le \|I - \mu m E\{\text{sgn}(X)X^T\}\| + \mu m O(m^{-1/2}) + 3M(\mu m\delta)^2. \tag{5.5}$$

Pick an arbitrary $\epsilon > 0$. Choose m so that $O(m^{-1/2}) < \epsilon$, and then choose μ so that $\mu m = \epsilon$. Then

$$\begin{aligned} E\|A(k+m,k)\| &\le \|I - \epsilon E\{\text{sgn}\, XX^T\}\| + (3M\delta^2+1)\epsilon^2 \\ &\le \max_{x^TPx=1} x^T(I-\epsilon E\{B\})^T P (I - \epsilon E\{B\})x + (3M\delta^2+1)\epsilon^2 \\ &= \max_{x^TPx=1} x^TPx - \epsilon x^T(E\{B\}^TP + PE\{B\})x + \epsilon^2 x^T E\{B\}^T P E\{B\}x + (3M\delta^2+1)\epsilon^2. \end{aligned} \tag{5.6}$$

Since $E\{B\}^TP + PE\{B\} = I$, and since $\min_{x^TPx=1} x^Tx = \frac{1}{\lambda_{\max}(P)}$, this becomes

$$= 1 - \frac{\epsilon}{\lambda_{\max}(P)} + (3M\delta^2 + 1 + \|E\{B\}\|)\epsilon^2.$$

Hence

$$\|A(k+m,k)\| \le 1 - \frac{\epsilon}{2\lambda_{\max}(P)}, \qquad \forall 0 < \epsilon < \frac{1}{2(3M\delta^2 + 1 + \|E\{B\}\|)\lambda_{\max}(P)}.$$

The desired stability follows. ΔΔΔ

There is a strong analogy between the stability results of Theorem 1 for deterministic systems and Theorem 4 for stochastic signals. It was fairly straightforward to extend Theorem 1 to give conditions for exponential instability in the deterministic case, but the proof of exponential instability with stochastic inputs is more subtle, and has so far eluded us. We therefore propose the following.

Instability conjecture: Under conditions (ii) and (iii) of Theorem 3, if some eigenvalue of the excitation matrix $E\,\text{sgn}\,XX^T$ has a negative real part, then the signed regressor algorithm will be exponentially unstable. ΔΔΔ

Condition (ii) of Theorem 4 is easy to fulfill since the stepsize is a design parameter. Many stochastic processes, especially those which are asymptotically decorrelated, fulfill condition (iii). The next section gives several examples. The next section also examines the excitation matrix $E\{\text{sgn}(X)X^T\}$ and shows that some random processes fulfill condition (i) while others fulfill the hypothesis of the instability conjecture.

VI. Interpretation of Stochastic PE for SR

Just as it is not immediately obvious what classes of deterministic inputs are PE for SR, it is not immediately clear what stochastic processes fulfill the excitation conditions of Theorem 4. This section demonstrates that certain stochastic processes are PE for SR while other processes fulfill the hypothesis of the instability conjecture. These latter can be shown (via simulation) to be unstable.

Suppose that u_k is stationary, independent, and zero mean, with any distribution (subject only to the existence of the mth moments, as above). Then X_k is n-dependent, and so the correlation function satisfies

$$R_l^{ij} = E\{[B_k - E\{B\}]_{ij}^T [B_{k+l} - E\{B\}]_{ij}^T\} = 0 \quad \text{for } l > n. \tag{6.1}$$

This shows that

$$\frac{d^2}{d\lambda^2} f^{ij}(\lambda)\Big|_{\lambda=0} = -\sum_{l=-\infty}^{\infty} l^2 R_l^{ij} e^{-jl\lambda}\Big|_{\lambda=0} = -\sum_{l=-\infty}^{\infty} l^2 R_l^{ij} = -\sum_{l=n}^{n} l^2 R_l^{ij} \tag{6.2}$$

which is clearly finite. Thus condition (iii) of Theorem 4 holds. Direct calculation shows that the excitation matrix

$E\{\operatorname{sgn}(X)X^T\}$ is equal to $E(|u_k|)I$. Hence Theorem 4 shows that for small enough μ, θ_k converges to zero exponentially fast.

A somewhat more realistic situation than the i.i.d. case is to suppose that the input u_k is generated by a white Gaussian noise source w_k passed through a stable linear filter F. The spectral density of u_k is then the spectral density of w_k multiplied by the square of the frequency response of F, and so R_l^{ij} is a weighted sum of decaying exponentials. This implies that (6.2) is summable and so condition (i) of Theorem 4 is again fulfilled. Gathering these results together, we have

Theorem 5: Suppose that the input to the signed regressor algorithm (1.4) is either

(i) generated by a white, zero mean Gaussian noise source passed through a stable linear filter, or

(ii) independent and identically distributed with finite moments.

Then, with small enough stepsize, the signed regressor algorithm is exponentially stable.

Proof: The only part that remains to show is that for case (i), $E\{\operatorname{sgn}(X)X^T\}$ has eigenvalues with positive real parts. For u_k and u_j jointly Gaussian, and zero mean

$$E\{\operatorname{sgn}(u_k)u_j\} = E\{\operatorname{sgn}(u_k)E(u_j|u_k)\}$$
$$= E\left\{\operatorname{sgn}(u_k)\frac{Eu_ku_j}{Eu_k^2}u_k\right\} = \frac{Eu_ku_j}{Eu_k^2}E|u_k|.$$

Since u_k is stationary, let $\alpha = E|u|/Eu^2$. Then the excitation matrix can be written $E(\operatorname{sgn}(X)X^T) = \alpha EXX^T$. It is well known that the covariance matrix EXX^T is positive definite when u is the output of a stable linear filter driven by white Gaussian noise. Thus the excitation matrix has eigenvalues with positive real parts, and Theorem 4 may be applied. △△△

In the deterministic case, some periodic sequences stabilize the SR algorithm and other sequences were destabilizing. This same dichotomy appears to be present in the stochastic case. Theorem 5 demonstrated stability when the input could be characterized by certain distributions and correlations. Other distributions exist, however, which cause $E\{\operatorname{sgn}(X)X^T\}$ to have eigenvalues with negative real parts. Such distributions fulfill the conditions of the instability conjecture.

Consider a Markov process u_k with l states, taking values $s = (s_1, s_2, \cdots, s_l)$. Let the one step transition probabilities be defined by the matrix P as $p_{ij} = p(u_{k+1} = s_i|u_k = s_j)$ and let the limiting probabilities $q = (q_1, q_2, \cdots, q_l)$ be the eigenvector of P associated with the eigenvalue $\lambda = 1$. Let $Q = \operatorname{diag}(q_i)$, for $i = 1, 2, \cdots, l$. The m-state transition matrix is then P^m and the excitation matrix can be calculated from

$$E\operatorname{sgn}(u_k)u_k = \sum_i \operatorname{sgn}(s_i)s_ip(u_k = s_i) = \operatorname{sgn}(s)^TQs$$

and

$$\begin{aligned} E\operatorname{sgn}(u_{k+m})u_k &= \sum_{i,j}\operatorname{sgn}(s_i)s_jp(u_{k+m} = s_i|u_k = s_j)p(u_j = s_j) \\ &= \operatorname{sgn}(s)^TP^mQs. \end{aligned} \tag{6.3}$$

Similarly, $E[\operatorname{sgn}\{u_{k-m}\}u_k] = s^TP^mQ\operatorname{sgn}(s)$. The excitation matrix $E\{\operatorname{sgn}(X)X^T\}$ thus has the Toeplitz form:

$$\begin{pmatrix} \operatorname{sgn}(s)^TQs & \operatorname{sgn}(s)^TPQs & \operatorname{sgn}(s)^TP^2Qs & \cdots & \operatorname{sgn}(s)^TP^{n-1}Qs \\ s^TPQ\operatorname{sgn}(s) & \operatorname{sgn}(s)^TQs & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ s^TP^{n-1}Q\operatorname{sgn}(s) & s^TP^{n-2}Q\operatorname{sgn}(s) & \cdots & \cdots & \operatorname{sgn}(s)^TQs \end{pmatrix}.$$

To be specific, consider a Markov process with two states taking values s_1 and s_2, with transition probabilities α and β. The transition matrix P is given by

$$P = \begin{pmatrix} \alpha & \beta \\ 1-\alpha & 1-\beta \end{pmatrix}.$$

The limiting probability (the eigenvector corresponding to $\lambda_1 = 1$) is $\xi_1 = (\beta/(1-\alpha+\beta), (1-\alpha)/(1-\alpha+\beta))^T$. The second eigenvalue is $\lambda_2 = \alpha - \beta$, with eigenvector $\xi_2 = (1, -1)^T$. Let $\Lambda = \{\xi_1|\xi_2\}$. Then the m-term probability matrix is

$$\begin{aligned} P^m &= \left(\Lambda\operatorname{diag}(\lambda_1, \lambda_2)\Lambda^{-1}\right)^m = \Lambda\operatorname{diag}\left(1, (\alpha-\beta)^m\right)\Lambda^{-1} \\ &= \frac{1}{1-\alpha+\beta} \\ &\quad\cdot\begin{pmatrix} \beta + (1-\alpha)(\alpha-\beta)^m & \beta(1-(\alpha-\beta)^m) \\ (1-\alpha)(1-(\alpha-\beta)^m) & (1-\alpha)+\beta(\alpha-\beta)^m \end{pmatrix}. \end{aligned}$$

The excitation matrix may be calculated directly from (6.3). It is easy to choose particular values of $(s_1, s_2, \alpha, \beta)$ for which $E\{\operatorname{sgn}(X)X^T\}$ has a negative real eigenvalue. For instance, the excitation matrix generated by $(1, -0.1, 0, 0.8)$ and $(1, -0.15, 0.2, 0.3)$ each have negative eigenvalues, and in simulations these cause "divergence" of the signed regressor algorithm (parameter estimates that overflow the numerical capabilities of the computer) in accordance with the instability conjecture. Other value sets, such as $(1, -0.1, 0.4, 0.8)$ have all positive eigenvalues and hence are convergent by Theorem 5. These examples were computed and simulated for the $n=6$ dimensional algorithm.

One situation which is easy to analyze is the case when $\alpha=\beta$, since then $P^m=P$ for every m, and $PQ=QP^T$. Hence all the off-diagonal terms of (6.3) are equal. Then the n-dimensional excitation matrix is $(E\,\mathrm{sgn}(X)X^T)_n=(c-d)I_n+dB_n$ where $c=E|u|$, $d=E\,\mathrm{sgn}(u_k)u_j$ for $k\neq j$, and $B_n=1_n1_n^T$. This excitation matrix has one eigenvalue equal to $dn+(c-d)$ and $n-1$ eigenvalues of $(c-d)$. Thus if $c>0>d$, the dimension n can always be chosen large enough so that one eigenvalue is negative. If $c>d>0$, then all eigenvalues are positive for any n. If s_1 and s_2 have opposite signs, then (6.3) shows that $d=(2\alpha-1)\{|s_1|\alpha-|s_2|(1-\alpha)\}$. Thus, the excitation matrix has negative eigenvalues whenever

(i) $\alpha>1/2$ and $|s_1|>\dfrac{(1-\alpha)}{\alpha}|s_2|$, or

(ii) $\alpha<1/2$ and $|s_1|<\dfrac{(1-\alpha)}{\alpha}|s_2|$.

The SR algorithm will be stable whenever $d\geqslant 0$.

One common situation is when the Markov process consists of two states with values s^* and $-s^*$. Then $d=|s^*|(1-2\alpha)^2$ and the algorithm is stable.

When $\alpha=\beta$, the condition for zero mean is that $s_1\alpha+s_2(1-\alpha)=0$, and so $d=0$. When $\alpha\neq\beta$, and the process is zero mean, the excitation matrix can be computed as

$$c\begin{pmatrix}1 & \delta & \delta^2 & \cdots & \delta^{n-1}\\ \delta & 1 & \delta & \cdots & \delta^{n-2}\\ \cdot & \cdot & \cdot & \cdots & \cdot\\ \cdot & \cdot & \cdot & \cdots & \cdot\\ \cdot & \cdot & \cdot & \cdots & \cdot\\ \delta^{n-1} & \delta^{n-2} & \delta^{n-3} & \cdots & 1\end{pmatrix}=cD_n$$

where $\delta=\alpha-\beta$ and $c=\dfrac{|s_1|\beta+|s_2|(1-\alpha)}{1-\alpha+\beta}$. D_n has determinant $(1-\delta^2)^{n-1}$. Since $-1<\delta<1$, D_n is positive definite for every n. (This may be verified by row reduction.) Thus no two state zero mean Markov process can cause divergence of the SR algorithm.

More complex zero mean Markov processes, however, can cause the excitation matrix to have negative eigenvalues. For instance, the zero mean process with states $(1,-0.1,-11)$ and transition probabilities

$$\begin{pmatrix}0.02 & 0.1 & 0.85\\ 0.2 & 0.1 & 0.1\\ 0.78 & 0.8 & 0.05\end{pmatrix}$$

has a six-dimensional excitation matrix $E\,\mathrm{sgn}(X)X^T$ with negative eigenvalues.

Theorem 6: Suppose that the input to the n-dimensional signed regressor algorithm (1.2) is a Markov process on a l-dimensional state space. For $l\geqslant 2$, and n large enough, there are processes with states $s_1,s_2,\cdots,s_l$ and transition matrices P that stabilize the algorithm. Other states and transition probabilities cause $E\{\mathrm{sgn}(X)X^T\}$ to have eigenvalues with negative real parts.

Proof: To show the existence of Markov inputs that generate an excitation matrix with both positive and negative eigenvalues, specialize to the case where all transition probabilities depend on a single parameter α, i.e.,

$$P=\begin{pmatrix}\alpha & \alpha & \alpha & \cdots & \alpha\\ \alpha & \alpha & \alpha & \cdots & \alpha\\ \cdot & \cdot & \cdot & \cdots & \cdot\\ \cdot & \cdot & \cdot & \cdots & \cdot\\ \cdot & \cdot & \cdot & \cdots & \cdot\\ \beta & \beta & \beta & \cdots & \beta\end{pmatrix}\qquad \text{where } \beta=1-(1-l)\alpha.$$

Then $P^m=P$ for every m, and the limiting probabilities are $q=(\alpha,\alpha,\cdots,\beta)^T$. Examining (6.2) shows that the excitation matrix has only two distinct entries and may be written in the form $(c-d)I_n+dB_n$ where $B_n=1_n1_n^T$. As in the two state case, $c>0>d$ implies that $E\{\mathrm{sgn}(X)X^T\}$ has eigenvalues with negative real part, while $c>d>0$ gives eigenvalues with positive real part. From (6.3),

$$c=\alpha\sum_{i=1}^{l}|s_i|+\beta|s_l|$$

$$d=\alpha^2\sum_{i=1}^{l-1}\sum_{j=1}^{l-1}\mathrm{sgn}(s_i)s_j+\alpha\beta\sum_{j=1}^{l-1}\big(\mathrm{sgn}(s_j)s_l+\mathrm{sgn}(s_l)s_j\big)+\beta^2|s_l|$$

and so it is easy to choose the α, β, and s_i to make d either positive or negative. It is then not too difficult to demonstrate that the hypothesis of condition (iii) of Theorem 4 are fulfilled by finite state Markov chains. △△△

VII. Conclusion

This paper has used averaging theory to derive persistence of excitation conditions for the signed regressor algorithm which *are not equivalent* to the standard least mean squares excitation conditions. These new conditions were then interpreted in both the deterministic and stochastic settings. Classes of deterministic inputs were delineated for which the signed regressor algorithm is stable, and contrasted with classes of inputs which destabilize the algorithm. Similarly, in the stochastic case, certain processes cause convergence of the parameter estimates while others cause divergence. Even the use of leakage does not guarantee bounded input/bounded output stability of the signed regressor algorithm.

The behavior of the signed regressor algorithm is heavily dependent on the characteristics of the input. Certain applications, such as processing of speech (which can be modeled as a jointly Gaussian stochastic process) may be ideal candidates for the signed regressor algorithm. In other applications, such as the processing of certain digital data (which can be modeled as finite state-space Markov chains or short almost periodic sequences) the signed regressor algorithm may not be advisable.

References

[1] 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. Reidle, *Stability of Adaptive Systems: Passivity and Averaging Analysis*. Cambridge, MA: MIT Press, 1986.

[2] B. D. O. Anderson and R. M. Johnstone, "Adaptive systems and time-varying plants," *Int. J. Contr.*, vol. 37, pp. 367–377, 1983.

[3] N. J. Bershad, "Comments on 'Comparison of the convergence of two algorithms for adaptive FIR digital filters,'" *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 1604–1606, Dec. 1985.

[4] ——, "On the optimum data nonlinearity in LMS adaptation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 69–76, Feb. 1986.

[5] R. R. Bitmead, "Persistence of excitation and the convergence of adaptive schemes," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 183–191, Mar. 1984.

[6] R. R. Bitmead, B. D. O. Anderson, and T. S. Ng, "Convergence rate determination for gradient-based adaptive estimators," *Automatica*, vol. 22, no. 2, Mar. 1986.

[7] R. R. Bitmead and C. R. Johnson, Jr, "Discrete averaging principles and robust adaptive identification," in *Control and Dynamic Systems: Advances in Theory and Applications*, vol. 25 (C. T. Leondes, Ed.,) New York: Academic, pp. 237–271, 1987.

[8] T. A. C. M. Claasen and W. F. G. Mecklenbrauker, "Comparison of the convergence of two algorithms for adaptive FIR digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 670–678, June 1981.

[9] ——, "Author's reply to 'Comments on "Comparison of the convergence of two algorithms for adaptive FIR digital filters'", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 202–203, Feb. 1986.

[10] C. F. N. Cowan and P. M. Grant, *Adaptive Filters*, (section 7.2), Englewood Cliffs, NJ: Prentice-Hall, 1985.

[11] S. Dasgupta and C. R. Johnson, Jr., "Some comments on the behavior of sign-sign adaptive identifiers," *Syst. Contr. Lett.*, vol. 7, Apr. 1986.

[12] ——, "Sign-sign adaptive identifiers: convergence and robustness properties," presented at the *2nd IFAC Workshop on Adaptive Systems*, Lund, Sweden, 1986.

[13] D. L. Duttweiler, "Adaptive filter performance with nonlinearities in the correlation multiplier," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 578–586, Aug. 1982.

[14] P. A. Ioannou and P. V. Kokotovic, *Adaptive Systems with Reduced Models*. New York: Springer-Verlag, 1983.

[15] N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, (section 6.5.3), Englewood Cliffs, NJ: Prentice-Hall, 1984.

[16] A. A. A. Karim, R. Nagarajan, and K. B. Mirza, "Discrete time MRAS with reduced implementation complexity," *Electron. Lett.*, vol. 17, pp. 809–810, Oct. 1981.

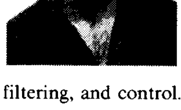
[17] R. Kumar and J. B. Moore, "Adaptive equalization via fast quantized-state methods," *IEEE Trans. Commun.*, vol. COM-29, pp. 1492–1501, Oct. 1981.

[18] R. W. Lucky, "Techniques for adaptive equalization of digital communication systems," *Bell System Tech. J.*, vol. 45, pp. 225–286, Feb. 1966.

[19] J. L. Moschner, "Adaptive filter with clipped input data," Tech. Rep. 6796-1, Stanford Univ. Information Systems Lab. Stanford, CA, June 1970.

[20] W. A. Sethares, D. A. Lawrence, C. R. Johnson, Jr., and R. R. Bitmead, "Parameter drift in LMS adaptive filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, Aug. 1986.

[21] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, Jr., "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, pp. 1151–1162, Aug. 1976.



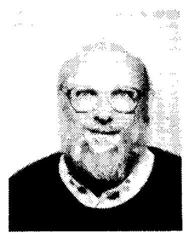
W. A. Sethares (S'84–M'87) received the B.A. degree in mathematics from Brandeis University, Waltham, MA, and the M.S. and Ph.D. degrees in electrical engineering from Cornell University, Ithaca, NY.

He has worked at Raytheon Company as a Systems Engineer and recently has accepted a teaching position in the Department of Electrical and Computer Engineering at the University of Wisconsin in Madison. His current research interests include adaptive systems: identification, filtering, and control.

Iven M. Y. Mareels (SM'86–M'86) received the B.E. degree in electrical and mechanical engineering from the State University Gent, Belgium, in 1982 and the Ph. D. degree in systems engineering from the Australian National University, Canberra, Australia, in 1987.

From 1982 until mid 1984 he was a research assistant of the National Fund for Scientific Research Belgium at the Power Systems Department of the State University Gent. He was at the Systems Engineering Department of the Australian National University from 1984 until 1986. In 1986 he joined as a research associate of the National Fund for Scientific Research Belgium at the Systems engineering Department of the State University Gent. During 1987 he fulfilled his military service, after which he resumed the latter position. His main research interests are in nonlinear systems theory and in adaptive control and identification.

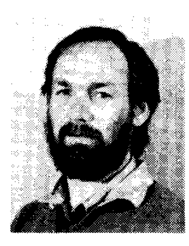
Brian D. O. Anderson (S'62–M'66–SM'74–F'75), for a photograph and biography please see page 384 of the April 1988 issue of this Transactions.



C. Richard Johnson, Jr. (S'74–M'77–SM'82) was born in Macon, GA, in 1950. He received the Ph.D. degree in electrical engineering with minors in engineering-economic systems and art history from Stanford University, CA, in 1977.

He is currently a Professor of Electrical Engineering at Cornell University, Ithaca, NY. His research interests are in adaptive parameter estimation as applied to digital filtering, identification, and control. He has served as an Editor of the *International Journal of Adaptive Control and Signal Processing* and as an Associate Editor for the IEEE Transactions on Acoustics, Speech, and Signal Processing, the IEEE Transactions on Automatic Control, and *Automatica*. He has authored *Lectures on Adaptive Parameter Estimation* (Englewood Cliffs, NJ: Prentice-Hall).

Dr. Johnson has received the C. Holmes MacDonald Outstanding Teacher Award in 1983, and the 1982 Senior (Best Paper) Award of the IEEE Acoustics, Speech, Signal Processing Society. He is a member of Eta Kappa Nu.



Robert R. Bitmead (S'76–M'79–SM'86) was born in Sydney, Australia, in 1954. He received the B.Sc. degree in applied mathematics from the University of Sydney and the M.E. and Ph.D. degrees in electrical engineering from the University of Newcastle, Australia, in 1976, 1977 and 1979, respectively.

During 1979 he held a Post-Doctoral position at the University of Newcastle. From 1980 to 1982 he was a faculty member in Electrical and Electronic Engineering at James Cook University of North Queensland, Townsville, and since 1983 he has been Fellow of Systems Engineering at the Australian National University, Canberra. In 1984 he held a visiting appointment in the School of Electrical Engineering at Cornell University.

He served as Associate Editor of the IEEE Transactions on Automatic Control from 1984 to 1986. He is currently Editor for Adaptive Control Theory of the *International Journal of Adaptive Control and Signal Processing* and is vice chairman of the IFAC committee on the Mathematics of Control. His research interests include adaptive systems, stochastic processes, multivariable control, and signal processing.