[9] J. L. Talmon and A. J. W. van den Boom, "On the estimation of transfer function parameters of process- and noise dynamics using a single-stage estimator," in *Proc. 3rd Int. Fed. Automat. Contr. Symp. Identification and System Parameter Estimation*, 1973.
[10] R. L. Kashyap, "Estimation of parameters in a partially whitened representation of a stochastic process," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 13–21, Feb. 1974.
[11] L. Ljung, T. Söderström, and I. Gustavsson, "Counterexamples to general convergence of a commonly used recursive identification method," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 643–652, Oct. 1975.
[12] J. Holst, "Local convergence properties for some recursive parameter estimation algorithms," Dep. Automat. Contr., Lund Inst. Technol., Lund, Sweden, 1977.
[13] K. J. Åström and B. Wittenmark, "On self-tuning regulators," *Automatica*, vol. 9, pp. 185–199, 1973.
[14] L. Ljung and B. Wittenmark, "Analysis of a class of adaptive regulators," in *Proc. Int. Fed. Automat. Contr. Symp. Stochastic Control*, Sept. 1974.
[15] K. J. Åström, U. Borisson, L. Ljung, and B. Wittenmark, "Theory and applications of adaptive regulators based on recursive identification," in *Proc. 6th Int. Fed. Automat. Contr. World Congr.*, 1975; also *Automatica*, vol. 13, 1977, to be published.
[16] K. J. Åström, *Introduction to Stochastic Control Theory*. New York: Academic, 1970.
[17] L. Ljung and B. Wittenmark, "Asymptotic properties of self-tuning regulators," Dep. Automat. Contr., Lund Inst. Technol., Lund, Sweden, Rep. 7407, Feb. 1974.
[18] G. Ledwich, "Adaptive identification, filtering and control," M. S. thesis, Dep. Elec. Eng., Univ. Newcastle, N.S.W., Australia, June 1976. (See also G. Ledwich and J. B. Moore, "Multivariable self-tuning filters," presented at the 4th Conf. Differential Games and Control Theory, July 1976.)
[19] L. Ljung, "Convergence of an adaptive filter algorithm," *Int. J. Contr.*, to be published.
[20] L. Ljung and B. Wittenmark, "On a stabilizing property of adaptive regulators," in *Int. Fed. Automat. Control Symp. Identification Preprints*, Sept. 1976.
[21] I. D. Landau, private communication.
[22] C. Doncarli, private communication.
[23] K. Y. Wong and E. Polak, "Identification of linear discrete time systems using the instrumental variable method," *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 707–718, 1967.
[24] P. C. Young, "An instrumental variable method for real-time identification of a noisy process," *Automatica*, vol. 6, pp. 271–287, 1970.
[25] L. Ljung, "The extended Kalman filter as a parameter estimator for linear systems," to be published.
[26] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Statist.*, vol. 22, pp. 400–407, 1951.
[27] G. Lüders and K. S. Narendra, "A new canonical form for an adaptive observer," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 117–119, Apr. 1974.
[28] E. J. Hannan, "The convergence of some recursions," *Ann. Statist.*, vol. 4, no. 6, pp. 1258–1270, 1976.
[29] ——, private communication.

Lennart Ljung (S'74–M'75) was born on September 13, 1946 in Malmö, Sweden. He received the B.A., M.S., and Ph.D degrees in 1967, 1970, and 1974, respectively, all from Lund University, Sweden.

From 1970 to 1976 he held various teaching and research positions at the Department of Automatic Control, Lund Institute of Technology, Sweden. In 1972 he spent six months with the Laboratory for Adaptive Systems at the Institute for Control Problems (IPU) in Moscow, USSR. In 1974–1975 he was Research Associate at the Information Systems Laboratory at Stanford University.

Since 1976 he is Professor of Automatic Control at the Department of Electrical Engineering, Linköping University, Linköping, Sweden. His scientific interests include aspects on identification, estimation, and adaptive control.

Analysis of Recursive Stochastic Algorithms

LENNART LJUNG, MEMBER, IEEE

Abstract—**Recursive algorithms where random observations enter are studied in a fairly general framework. An important feature is that the observations may depend on previous "outputs" of the algorithm. The considered class of algorithms contains, e.g., stochastic approximation algorithms, recursive identification algorithms, and algorithms for adaptive control of linear systems.**

It is shown how a deterministic differential equation can be associated with the algorithm. Problems like convergence with probability one, possible convergence points and asymptotic behavior of the algorithm can all be studied in terms of this differential equation. Theorems stating the precise relationships between the differential equation and the algorithm are given as well as examples of applications of the results to problems in identification and adaptive control.

Manuscript received February 23, 1976; revised March 2, 1977. Paper recommended by R. Monopoli, Chairman of the Adaptive, Learning Systems, Pattern Recognition Committee. A major part of this work has been supported by the Swedish Board for Technical Development under Contract 733546.

The author is with the Department of Electrical Engineering, Linköping University, Linköping, Sweden.

I. Introduction

RECURSIVE algorithms, where stochastic observations enter are common in many fields. In the control and estimation literature such algorithms are widely discussed, e.g., in connection with adaptive control, (adaptive) filtering and on-line identification. The convergence analysis of the algorithms is not seldom difficult. As a rule, special techniques for analysis are used for each type of application and often the convergence properties have to be studied only by simulation.

In this paper a general approach to the analysis of the asymptotic behavior of recursive algorithms is described. In effect, the convergence analysis is reduced to stability analysis of a deterministic, ordinary differential equation. This technique is believed to be a fairly general tool and to have a wide applicability. Applications to various prob-

lems have been published in [1]–[4] and some theory was presented in [5].

The objective of the present paper is to give a comprehensive presentation of formal results and useful techniques for the convergence analysis, as well as to illustrate with several examples how the techniques can be applied.

In Section II a general recursive algorithm is described and discussed. A heuristic treatment of the convergence problem is given in Section III and this leads to the basic ideas of the present approach. Section IV contains a discussion of the conditions which are imposed on the algorithm in order to prove the formal results. These theorems are given in Sections V and VI. The theorems suggest certain techniques for the convergence analysis, and these aspects are treated in Section VII. Several examples of how the theorems may be used, some of them reviewing previous applications are given in Section VIII.

II. The Algorithm

A general recursive algorithm can be written

$$x(t)=x(t-1)+\gamma(t)Q(t;x(t-1),\varphi(t)), \tag{1}$$

where $x(\cdot)$ is a sequence of n-dimensional column vectors, which are the objects of our interest. We shall refer to $x(\cdot)$ as "the *estimates*," and they could, e.g., be the current estimates of some unknown parameter vector. They could, however, also be parameters that determine a feedback law of an adaptive controller, etc., and we shall be precise about the character of $x(\cdot)$ only in the examples below. The sequence $\gamma(\cdot)$ is assumed throughout the paper to be a sequence of positive scalars. The m-dimensional vector $\varphi(t)$ is an *observation* obtained at time t, and these are the objects that cause $x(t-1)$ to be updated to take new information into account. (The notion "observation" does not have to be taken literally. The variable φ may very well be the result of certain treatment of actual measurements.) The observations are in general functions of the previous estimates $x(\cdot)$ and of a sequence of random vectors $e(\cdot)$. This means that the observation is a random variable, which may be affected by previous estimates. This is the case, e.g., for adaptive systems, when the input signal is determined on the basis of previous estimates. If the experiment designer has some test signal at his disposal, this may be included in $e(\cdot)$.

The function $Q(\cdot;\cdot,\cdot)$ from $R\times R^n\times R^m$ into R^n is a deterministic function with some regularity conditions to be specified below. This function, together with the choice of the scalar "gain" sequence $\gamma(\cdot)$ determine entirely the algorithm.

We shall not work with completely general dependence of $\varphi(t)$ on $x(\cdot)$, but the following structure for the generation of $\varphi(\cdot)$ will be used:

$$\varphi(t)=A(x(t-1))\varphi(t-1)+B(x(t-1))e(t). \tag{2}$$

Here $A(\cdot)$ and $B(\cdot)$ are $m|m$ and $m|r$ matrix functions.

Remark: It is perhaps more natural to think of an observation $\tilde{\varphi}(t)$ as the (lower dimensional) output of a dynamical system like (2), $\tilde{\varphi}(t)=C(x(t-1))\varphi(t)$. However, this case is naturally subsumed in the present one, since $\varphi(t)$ may enter in Q only as the combination $\tilde{\varphi}(t)$.

The assumption (2) seems to be appropriate for many applications. The same results as those below can be obtained also for nonlinear dynamics.

$$\varphi(t)=g(t;\varphi(t-1),x(t-1),e(t)) \tag{3}$$

and the proofs for this case are given in [7].

Throughout this paper it is assumed that the estimates are desired to converge to some "true" or "optimal" value(s). Since $Q(t,x(t-1),\varphi(t))$ is a random variable, with, in general, nonzero variance, convergence can take place only if the noise is rejected by paying less and less attention to the noisy observations, i.e., by letting

$$\gamma(t)\to 0 \qquad \text{as } t\to\infty. \tag{4}$$

In tracking problems, when a time-varying parameter is to be tracked using algorithm (1), this condition is not feasible. Then $\gamma(t)$ usually tends to some small, nonzero value, the size of which depends on what is known about the variability of the tracked parameters and about the noise characteristics. This case is not treated here, but it is reasonable to assume that analysis under the condition (4) also will have some relevance for the case of tracking slowly varying parameters.

Suppose we have a linear, stochastic, discrete-time system, governed by a linear output feedback law, which at time t is determined by $x(t-1)$. Then the behavior of this overall system can be described by (2), with $\varphi(t)$ consisting of lagged inputs and outputs. Therefore the algorithms (1) and (2) can be understood as archetypical for adaptive control of a linear system. Indeed this setup is useful for analysis of certain adaptive controllers, as will be exemplified below, but the basic algorithm (1), (2) also covers several other cases of interest.

III. Heuristic Analysis

The algorithm (1), (2) is fairly complex to analyze, being a time-variant, stochastic, nonlinear difference equation. Notice also that the correction $x(t)-x(t-1)$ depends in general via $\varphi(t)$ implicitly on all old $x(s)$. Therefore, while (1) certainly is recursive from the user's point of view, it is not so for analysis purposes.

In this section we shall illustrate heuristically how a differential equation can be associated with (1), and how it seems reasonable that asymptotic properties of (1) may be studied in terms of this differential equation. The formal analysis and results follow in the next two sections.

Consider

$$x(t)=x(t-1)+\gamma(t)Q(x(t-1),\varphi(t)), \tag{5}$$

where for simplicity we let Q be time independent. As remarked before, $\varphi(t)$ depends on all previous estimates:

$$\varphi(t)=\sum_{j=1}^{t}\left(\prod_{i=j+1}^{t}A(x(i-1))\right)B(x(j-1))e(j). \quad (6)$$

Now, if (2) is exponentially stable, then the first terms in (6) will be very small and, for some M,

$$\varphi(t)\approx\sum_{j=t-M}^{t}\left(\prod_{i=j+1}^{t}A(x(i-1))\right)B(x(j-1))e(j).$$

Moreover, it follows from (5) and (4) that the difference $x(t)-x(t-1)$ becomes smaller as t increases. Therefore, for sufficiently large t, we have $x(k)\approx x(t)$; $t\geqslant k\geqslant t-2M$. Hence,

$$\begin{aligned}\varphi(k)&\approx\sum_{j=k-M}^{k}\left[A(x(t))\right]^{k-j}B(x(t))e(j)\\&\approx\sum_{j=1}^{k}\left[A(x(t))\right]^{k-j}B(x(t))e(j)\triangleq\bar{\varphi}(k;x(t)) \quad (7)\end{aligned}$$

for $t\geqslant k\geqslant t-M$. Furthermore,

$$Q(x(k-1),\varphi(k))\approx Q(x(t),\bar{\varphi}(k;x(t)))=f(x(t))+w(k) \quad (8)$$

where

$$f(x)=EQ(x,\bar{\varphi}(k;x))$$

and hence $w(k)$ is a random variable with zero mean.

Using (8), we can approximately evaluate

$$\begin{aligned}x(t+s)&=x(t)+\sum_{k=t+1}^{t+s}\gamma(k)Q(x(k-1),\varphi(k))\\&\approx x(t)+f(x(t))\sum_{k=t+1}^{t+s}\gamma(k)+\sum_{k=t+1}^{t+s}\gamma(k)w(k)\\&\approx x(t)+f(x(t))\sum_{t+1}^{t+s}\gamma(k), \quad (9)\end{aligned}$$

where the last step should follow since the last term is a zero mean random variable which is dominated by the second term. Expression (9) suggests that the sequence of estimates more or less follows the difference equation

$$x^D(\tau+\Delta\tau)=x^D(\tau)+\Delta\tau f(x^D(\tau)) \quad (10)$$

where $\Delta\tau$ corresponds to

$$\sum_{t+1}^{t+s}\gamma(k).$$

It is useful to interpret (10) as a way of solving the differential equation ($\Delta\tau$ small),

$$\frac{d}{d\tau}x^D(\tau)=f(x^D(\tau)), \quad (11)$$

where the (fictitious) time τ relates to the original time t in (5) by

$$\tau_t=\sum_{k=1}^{t}\gamma(k). \quad (12)$$

We therefore have some reason to believe that the sequence of estimates $x(\cdot)$ asymptotically should follow the trajectories $x^D(\cdot)$ of (11).

(We could also have related (9) to the difference equation

$$x^D(t)=x^D(t-1)+\gamma(t)f(x^D(t-1)), \quad (13)$$

but the differential equation is easier to handle since it is time-invariant.)

It now seems reasonable that asymptotic properties of algorithm (1), (2) may be studied in terms of the differential equation (11). This heuristic discussion is perhaps not very convincing, but along a similar path, though with far more technical labor, formal results to this effect can be proven. These are given below.

IV. Assumptions on the Algorithm

In order to prove the formal results, certain regularity conditions on the functions Q, A, and B and on the driving "noise" term e, have to be introduced. Some of these are fairly technical, but it is believed that none is very restrictive. Several sets of assumptions are possible, and we shall give a few. In particular, there is a possibility to treat the sequence $e(\cdot)$ either in a stochastic or in a deterministic framework.

We shall start by giving a formal definition of $\bar{\varphi}$ used in the previous section. Let

$$D_S=\{x|A(x) \text{ has all eigenvalues strictly inside the unit circle}\}.$$

Then for each $x\in D_S$, there exists a $\lambda=\lambda(x)$, such that

$$|A(x)^k|<C\lambda(x)^k; \qquad \lambda(x)<1. \quad (14)$$

Take $\bar{x}\in D_S$ and define the random variables $\bar{\varphi}(t,\bar{x})$ and $v(t,\lambda,c)$, $\lambda<1$, by

$$\bar{\varphi}(t,\bar{x})=A(\bar{x})\bar{\varphi}(t-1,\bar{x})+B(\bar{x})e(t); \qquad \bar{\varphi}(0,\bar{x})=0 \quad (15)$$

$$v(t,\lambda,c)=\lambda v(t-1,\lambda,c)+c|e(t)|; \qquad v(0,\lambda,c)=0. \quad (16)$$

Let D_R be an open, connected subset of D_S. The regularity conditions will be assumed to be valid in D_R. Now, the first set of assumptions is the following.

A.1: $e(\cdot)$ is a sequence of independent random variables (not necessarily stationary or with zero means).

A.2: $|e(t)|<C$ with probability one (w.p.1) all t.

A.3: The function $Q(t,x,\varphi)$ is continuously differentiable w.r.t x and φ for $x\in D_R$. The derivatives are, for fixed x and φ, bounded in t.

A.4: The matrix functions $A(\cdot)$ and $B(\cdot)$ are Lipschitz continuous in D_R.

A.5: $\lim_{t\to\infty} EQ(t,\bar{x},\bar{\varphi}(t,\bar{x}))$ exists for $\bar{x}\in D_R$ and is denoted by $f(\bar{x})$. The expectation is over $e(\cdot)$.

A.6: $\sum_1^\infty \gamma(t)=\infty$

A.7: $\sum_1^\infty \gamma(t)^p<\infty$ for some p.

A.8: $\gamma(\cdot)$ is a decreasing sequence.

A.9: $\lim_{t\to\infty}\sup[1/\gamma(t)-1/\gamma(t-1)]<\infty$.

These conditions will be referred to as "assumptions A." A.1 introduces the stochastic structure into the setup. While A.2 certainly is most reasonable for all practical purposes, it is somewhat unattractive from a theoretical point of view, since it excludes, e.g., the common Gaussian models for noise. Below (assumptions B) are given conditions which allow more general noise. Conditions A.3 and A.4 are reasonable regularity properties, and A.5 is the basic assumption that makes it possible to associate (1), (2) with a differential equation. Condition A.6 makes it possible for the algorithm (1) to move the estimate to the desired limit, regardless of the initial value, and it is thus obviously necessary. A.7 gives a condition on how fast $\gamma(t)$ must tend to zero. This is considerably less restrictive than the usually given condition

$$\sum_1^\infty \gamma^2(t)<\infty. \tag{17}$$

Conditions A.8 and A.9 are motivated by technical arguments in the proofs, but they have so far not appeared to be restrictive. For example, it is easy to see that the sequence $\gamma(t)=Ct^{-\alpha}$ satisfies A.6–A.9 for $0<\alpha\leqslant 1$.

If we would like to alleviate A.2, further regularity conditions on Q are required. This gives us our second set of assumptions. ($\mathcal{B}(\bar{x},\rho)$ denotes a ρ-neighborhood of $\bar{x}$, i.e., $\mathcal{B}(\bar{x},\rho)=\{x\,|\,|\bar{x}-x|<\rho\}$.)

B.1: $e(\cdot)$ is a sequence of independent random variables (not necessarily stationary or with zero means).

B.2: $E|e(t)|^p$ exists and is bounded in t for each $p>1$.

B.3: The function $Q(t,x,\varphi)$ is Lipschitz continuous in x and φ: $|Q(t,x_1,\varphi_1)-Q(t,x_2,\varphi_2)|<\mathcal{K}_1(x,\varphi,\rho,v)\{|x_1-x_2|+|\varphi_1-\varphi_2|\}$ for $x_i\in\mathcal{B}(x,\rho)$ for some $\rho=\rho(x)>0$ where $x\in D_R$; $\varphi_i\in\mathcal{B}(\varphi,v)$, $v\geqslant 0$.

B.4: $|\mathcal{K}_1(x,\varphi_1,\rho,v_1)-\mathcal{K}_1(x,\varphi_2,\rho,v_2)|\leqslant \mathcal{K}_2(x,\varphi,\rho,v,w)\cdot\{|\varphi_1-\varphi_2|+|v_1-v_2|\}$ for $\varphi_i\in\mathcal{B}(\varphi,w)$ and $v_i\in\mathcal{B}(v,w)$.

B.5: $A(\cdot)$ and $B(\cdot)$ are Lipschitz continuous in D_R.

B.6: $\lim_{t\to\infty} EQ(t,\bar{x},\bar{\varphi}(t,\bar{x}))$ exists for $\bar{x}\in D_R$ and is denoted by $f(\bar{x})$. The expectation is over $e(\cdot)$.

B.7: For $x\in D_R$, the random variables $Q(t,x,\bar{\varphi}(t,x))$, $\mathcal{K}_1(x,\bar{\varphi}(t,x),\ \rho(x),\ v(t,\lambda,c))$ and $\mathcal{K}_2(x,\bar{\varphi}(t,x),\ \rho(x),\ v(t,\lambda,c),\ v(t,\lambda,c))$ have bounded p-moments for all $p>1$, and all $\lambda<1$, $c<\infty$. Here $\bar{\varphi}(\cdot,x)$ and $v(\cdot,\lambda,c)$ are the random variables defined by (15) and (16).

B.8: $\sum_1^\infty \gamma(t)=\infty$.

B.9: $\sum_1^\infty \gamma(t)^p<\infty$ for some p.

B.10: $\gamma(\cdot)$ is a decreasing sequence.

B.11: $\lim_{t\to\infty}\sup[1/\gamma(t)-1/\gamma(t-1)]<\infty$.

Conditions B.4, B.3, and B.7 admittedly look somewhat complex, but they are as a rule easy to check in a given situation, especially since $Q(t,x,\varphi)$ is a simple function of x and φ in most applications. The conditions B.3 and B.4 effectively require that $Q(t,x,\varphi)$ is twice continuously differentiable and B.7 implies that Q and its derivatives must not increase too rapidly with φ and v.

In these two cases the algorithm (1), (2) is treated directly in a stochastic framework, due to assumption A.1 = B.1. In certain cases it may not be suitable to treat $e(\cdot)$ in (2) as a sequence of random variables. Naturally the algorithm (1), (2) still makes sense, even if $e(\cdot)$ is a given, deterministic sequence. Convergence of (1) will than depend, among other things, on the properties of this sequence $e(\cdot)$. In such a case we may work with the following assumptions. Let $\mathcal{K}_1$ be defined as in B.3 and let $\bar{\varphi}(\cdot,\bar{x})$ and $v(\cdot,\lambda,c)$ be given by (15), (16). Introduce the quantities $z(\cdot,\bar{x})$ and $k_v(\cdot,\bar{x},\lambda,c)$ by

$$z(t,\bar{x})=z(t-1,\bar{x})+\gamma(t)\big[Q\big(t,\bar{x},\bar{\varphi}(t,\bar{x})\big)-z(t-1,\bar{x})\big] \qquad z(0,\bar{x})=0 \tag{18a}$$

$$k_v(t,\bar{x},\lambda,c)=k_v(t-1,\bar{x},\lambda,c)+\gamma(t)\cdot\big[\mathcal{K}_1\big(\bar{x},\bar{\varphi}(t,\bar{x}),\rho(\bar{x}),v(t,\lambda,c)\big)\cdot\big(1+v(t,\lambda,c)\big)-k_v(t-1,\bar{x},\lambda,c)\big]; \qquad k_v(0,\bar{x},\lambda,c)=0. \tag{18b}$$

Notice that for the common choice $\gamma(t)=1/t$, (18a) implies that

$$z(t,\bar{x})=\frac{1}{t}\sum_{k=1}^{t} Q\big(k;\bar{x},\bar{\varphi}(k,\bar{x})\big)$$

and analogously for $k_v(t,\bar{x},\lambda,c)$. The assumptions then are as follows.

C.1: The function $Q(t,x,\varphi)$ is Lipschitz continuous in x and φ: $|Q(t,x_1,\varphi_1)-Q(t,x_2,\varphi_2)|<\mathcal{K}_1(x,\varphi,\rho,v)\{|x_1-x_2|+|\varphi_1-\varphi_2|\}$ for $x_i\in\mathcal{B}(x,\rho)$ for some $\rho=\rho(x)>0$ where $x\in D_R$; $\varphi_i\in\mathcal{B}(\varphi,v)$, $v\geqslant 0$.

C.2: The matrix functions $A(\cdot)$ and $B(\cdot)$ are Lipschitz continuous in D_R.

C.3: $z(t,\bar{x})$ as defined by (18a) converges for all $\bar{x}\in D_R$ as $t\to\infty$. Denote the limit by $f(\bar{x})$.

C.4: $k_v(t,\bar{x},\lambda,c)$ defined by (18b) converges to a finite limit as $t\to\infty$ for all $\bar{x}\in D_R$, $\lambda<1$ and $c<\infty$.

C.5: $\sum_1^\infty \gamma(t)=\infty$.

C.6: $\gamma(t)\to 0$ as $t\to\infty$.

C.7: If the matrices A and B in (2) do not depend on x, then Lipschitz continuity w.r.t to φ is not required in C.1 and we may take $c=0$ in C.4. That is, we may take $v\equiv 0$ in C.1 and C.4.

When these assumptions are used, no stochastic framework has to be introduced. The statements about the behavior of $x(\cdot)$ to be given below are true as long as $e(\cdot)$ is such that C.3 and C.4 hold. If a stochastic framework *is* imposed and C.3, C.4 hold with probability one, then the statements about $x(\cdot)$ will be true w.p.1. This is, essentially, the approach taken in [5] and [8], which also contain a detailed study of algorithms like (18) (esp. [8, ch. 4]). There several different sets of conditions implying

convergence of (18) are given. In fact, conditions B imply that C.3 and C.4 hold w.p.1. It may in this context be remarked that there is actually a tradeoff between condition A.7 = B.9 and conditions B.2 and B.7. The largest p for which B.2 and B.7 need to hold is twice the p for which A.7 holds. Therefore, if $\gamma(\cdot)$ *is* subject to (17), then only the fourth moments of e, Q, and $\mathcal{H}_i$ have to be bounded. This is discussed further in [5] and [8], and we shall not pursue it here.

V. Main Theorems

The function $f(\cdot)$ defined in A.5, B.6, or C.3 is the basic object of interest. As the heuristic discussion in Section III indicated the differential equation

$$\frac{d}{d\tau}x^D(\tau)=f\big(x^D(\tau)\big) \tag{19}$$

will be relevant for the asymptotic behavior of the algorithm (1), (2). The exact relationships between (19) and (1), (2) are given in three theorems. The first one concerns convergence of (1).

Theorem 1: Consider the algorithm (1), (2) subject to assumptions A, B or C. Let $\bar{D}$ be a compact subset of D_R such that the trajectories of (19) that start in $\bar{D}$ remain in a closed subset $\bar{D}_R$ of D_R for $\tau>0$. Assume that

1) there is a random variable C such that

$$x(t)\in\bar{D} \text{ and } |\varphi(t)|<C \quad \text{infinitely often (i.o.) w.p.1} \tag{20}$$

2) the differential equation (19) has an invariant set D_c with domain of attraction $D_A\supset\bar{D}$. (21)

Then $x(t)\to D_c$ with probability one as $t\to\infty$. □

Remarks: By (20) is meant that there exists with probability one a subsequence t_k, possibly depending on the realization ω, such that $x(t_k)\in\bar{D}$ and $|\varphi(t_k)|<C(\omega)$, $k=1, 2,\cdots$. This condition, which we may call the "boundedness condition," is further discussed in Section VI.

An invariant set of a differential equation is a set such that the trajectories remain in there for $-\infty<\tau<\infty$. The domain of attraction of an invariant set D_c consists of those points from which the trajectories converge into D_c as τ tends to infinity. It is obviously an open set. See, e.g., [9]. An interesting special case is when the invariant set D_c is just a stationary point of (19) say x^*, with $f(x^*)=0$. Then the theorem proves convergence of $x(t)$ to x^*. By $x(t)\to D_c$ is meant that

$$\inf_{x\in D_c}|x(t)-x|\to 0.$$

The phrase "w.p.1" naturally does not apply in the case with assumptions C. In order to verify the stability condition (21) analytically, usually the Lyapunov theory has to be applied. Theorem 1 can be given a formulation, which does not refer to any differential equation, but directly relates to a Lyapunov function associated with $f(x)$. In that way also the assumption about the trajectories of (19) remaining in $\bar{D}_R$ and boundedness of $\bar{D}$ can be avoided.

Corollary 1: Consider the algorithm (1), (2) subject to assumptions A, B or C. Let $\bar{D}$ be a closed subset of D_R (possibly $\bar{D}=D_R=R^n$) such that (20) holds. Assume that there exists a twice differentiable function $V(x)$ in D_R, such that

$$V'(x)f(x)\leqslant 0 \qquad x\in D_R. \tag{22}$$

Then either

$$x(t)\to D_c=\{x\,|\,x\in D_R \text{ and } V'(x)f(x)=0\} \quad \text{w.p.1 as } t\to\infty$$

or

$\{x(t)\}$ has a cluster point on the boundary of D_R. □

Our second theorem concerns the set of possible convergence points. It can be used to prove failure of convergence by showing that the "desired" or "true" parameter value does not belong to this set.

Theorem 2: Consider algorithm (1), (2) subject to assumptions A or B. Suppose that $x^*\in D_R$ has the property that

$$P\big(x(t)\nrightarrow\mathcal{B}(x^*,\rho)\big)>0 \qquad \text{for all } \rho>0. \tag{23}$$[1]

Furthermore, suppose that

$$Q(t,x^*,\bar{\varphi}(t,x^*)) \text{ has a covariance matrix bounded from below by a strictly positive definite matrix,} \tag{24}$$

and that

$$EQ(t,x,\bar{\varphi}(t,x)) \text{ is continuously differentiable with respect to } x \text{ in a neighborhood of } x^* \text{ and the derivatives converge uniformly in this neighborhood as } t \text{ tends to infinity.}$$

Then

$$f(x^*)=0 \tag{25a}$$

and

$$H(x^*)=\frac{d}{dx}f(x)\Big|_{x=x^*} \text{ has all eigenvalues in the LHP } (\mathrm{Re}\,z\leqslant 0). \quad \square \tag{25b}$$

The matrix $H(x^*)$ defines, of course, the linear differential equation obtained from (19) by linearization around x^*. Therefore this theorem essentially states that the algorithm can converge only to stable stationary points of the differential equation (19).

If $f(x)=-(d/dx)V(x)$, which might be the case if the algorithm is based on criterion-minimization, then $V(x)$ can be chosen as a Lyapunov function for the differential equation (19). Since $(d/d\tau)V(x(\tau))=-|f(x(\tau))|^2$, we see

[1] $P(A)$ = The probability of the event A.

that the stationary points of (19), together with the point $\{\infty\}$, form an invariant set with global domain of attraction. Moreover, if the stationary points are isolated, it follows from Theorem 2 that only stable ones, i.e., local minima, are possible convergence points. It also follows from Theorem 1 that the estimates cannot oscillate between different minima. Collecting all this we obtain a corollary to Theorems 1 and 2.

Corollary 2: Suppose that $D_R = R^n$, that $f(x) = -(d/dx)V(x)$ and that $V(x)$ has isolated stationary points. Assume that $|\varphi(t)| < C$ i.o. w.p.1. Then, w.p.1, $x(t)$ tends either to a local minimum of $V(x)$ (i.e., $V''(x)$ positive semidefinite) or to infinity as t tends to infinity.

Finally, our third theorem relates the trajectories of the differential equation (19) to the paths of the algorithm (1), (2). The result is formulated as follows. Let $x(t)$, $t = t_0, \cdots$, be generated by (1), (2). The values can be plotted with the sample numbers t as the abscissa. It is also possible to introduce as before a fictitious time τ by

$$\tau_t = \sum_{k=1}^{t} \gamma(k). \tag{26}$$

Suppose that the estimates $x(t)$ are plotted against this time τ. See Fig. 1(a). Let $x^D(\tau, \tau_{t_0}, x(t_0))$ be the solution of (19) with initial value $x(t_0)$ at time τ_{t_0}. Also plot this solution in the same diagram, as in Fig. 1(b). Let I be a set of integers. The probability that *all* points $x(t)$, $t \in I$, simultaneously are within a certain distance ϵ from the trajectory is estimated in the following theorem.

Theorem 3: Consider algorithm (1), (2) under assumptions A or B. Assume that $f(x)$ is continuously differentiable, and that (20) holds. Assume that the solutions to (19) with initial conditions in $\bar{D}$ are exponentially stable, and let I be a set of integers, such that $\inf |\tau_i - \tau_j| = \delta_0 > 0$ where $i \neq j$ and $i, j \in I$. Then for any $p \geq 1$ there exist constants K, ϵ_0 and T_0 that depend on p, $\bar{D}$, and δ_0, such that for $\epsilon < \epsilon_0$ and $t_0 > T_0$.

$$P\left\{ \sup_{\substack{t \in I \\ t \geq t_0}} |x(t) - x^D(\tau_t; \tau_{t_0}, x(t_0))| > \epsilon \right\} \leq \frac{K}{\epsilon^{4p}} \sum_{j=t_0}^{N} \gamma(j)^p \tag{27}$$

where $N = \sup i$; $i \in I$, which may be ∞.

Remark: In the proof of Theorem 3 it is assumed that the exponential stability of the solution $x^D(\tau, \tau_{t_0}; x(t_0))$ is ensured by a quadratic Lyapunov function for the (linear and time-varying) variational equation around this solution, cf., e.g., [10].

Although the proof of Theorem 3 provides an estimate of K from given constants, we do not intend to use (27) to obtain numerical bounds for the probability. The point of the theorem is that a connection between the differential equation (19) and the algorithm (1), (2) is established. In particular, we notice that, due to A.7 there is a p such that the RHS of (27) becomes arbitrarily small when t_0 increases, and ϵ is fixed. This means that the estimates stay close to the corresponding trajectory with higher and higher probability as t_0 increases. Another way of interpreting (27) is that the gain sequence $\gamma(\cdot)$ can be scaled so that $x(\cdot)$ stays arbitrarily close to $x^D(\cdot)$, with an arbitrary high degree of probability.

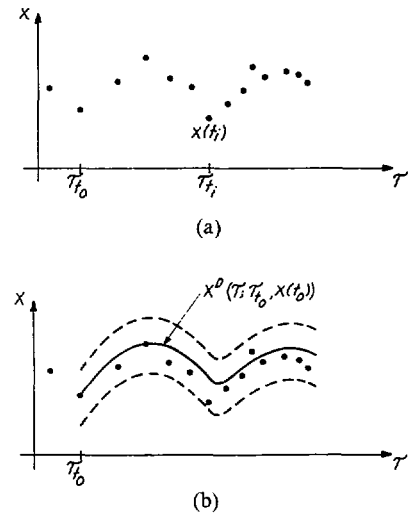

Fig. 1. (a) and (b) To illustrate Theorem 3.

A result that is related to Theorem 3 in the case $A(x) \equiv 0$ is given in [35]. Interesting connections between stochastic-approximation type algorithms and a corresponding differential equation have also recently been made by Kushner [36], [37], using weak convergence theory.

The proof of Theorems 1, 2, and 3 are long and technical. They are given in Appendices I, II, and III, respectively. The idea of the proofs of Theorems 1 and 3 follow the discussion in Section III. However, a considerable amount of technicalities are required to rigorously justify the "approximatively equal" signs.

In Appendix V extensions of the theorems, e.g., to continuous time algorithms and to cases when the limit in A.5 or B.6 does not exist are also commented upon.

VI. The Boundedness Condition

In this section we shall discuss condition (20). The reason why it is required is twofold. First, obviously $x(t)$ must be inside D_R (with $\varphi(t)$ not too large to prevent an immediate jump) for the differential equation to be valid at all, and also inside D_A to get "caught" by a trajectory converging to D_c. Second, and perhaps less obviously, even if $D_R = D_A = R^n$ it may happen that $x(t)$ tends to infinity. The reason for divergence is that if $Q(t, x, \varphi)$ increases rapidly with $|x|$ it may happen that the correction $\gamma(t)Q(t, x(t-1), \varphi(t))$ always is too large even though $\gamma(t)$ tends to zero. Another reason is that the variance of the "noise" $Q(t, x, \varphi) - f(x)$ may increase so fast with $|x|$ that a "random walk" effect becomes predominately close to infinity.

From a practical point of view, the question of boundedness of the estimates may seem uninteresting, since no implementation of (1) will allow that $x(t)$ tends to infinity. It will be kept bounded either by deliberate measures or due to, e.g., overflow in the computer. Now

the measures to keep $x(t)$ in a bounded area may not be completely arbitrary to obtain convergence.

A feature that can be used when $D_R = D_A = R^n$, is to introduce a saturation in $Q(\cdot;\cdot,\cdot)$ so that $|Q(t;x,\varphi)| < \mathcal{K}$. This is further discussed in [7].

Another possibility of preventing $x(t)$ from tending to infinity is to project $x(t)$ into a bounded area if $|x(t)|$ is too large or if $x(t)$ does not belong to a desired area, say D_S. In fact, if $A(x)$ is a known function of x, it is common, and often even necessary to test if $x(t) \in D_S$ and project it into D_S otherwise. We then have an algorithm of the following type:

$$x(t) = \left[x(t-1) + \gamma(t) Q(t, x(t-1), \varphi(t)) \right]_{D_1, D_2} \quad (28)$$

$$\varphi(t) = \begin{cases} A(x(t-1))\varphi(t-1) + B(x(t-1))e(t) \\ \qquad \text{if } x(t-1) \in D_1 \\ \text{a value in a given compact subset of } R^m \\ \qquad \text{if } x(t-1) \notin D_1 \end{cases} \quad (29)$$

where $D_1 \supset D_2$, and

$$[z]_{D_1, D_2} = \begin{cases} z & \text{if } z \in D_1 \\ \text{some value in } D_2 & \text{if } z \notin D_1. \end{cases}$$

It should be clear that D_1, D_2 cannot be chosen arbitrarily. Loosely speaking, the trajectories of (19) that start in D_2 must not leave the area D_1. Otherwise there may be an undesired cluster point on the boundary of D_1. This may be formalized as follows.

Theorem 4: Consider the algorithm (28), (29) subject to assumptions A, B or C. Let $D_1 \subset D_R$ be an open bounded set containing the compact set D_2. Let $\tilde{D} = D_1 \setminus D_2$ (D_1 "minus" D_2). Assume that $D_2 \subset D_A$, with D_A defined as in Theorem 1. Suppose that there exists a twice differentiable function $U(x) \geqslant 0$, defined in a neighborhood of $\tilde{D}$ with properties

$$\sup_{x \in \tilde{D}} U'(x) f(x) < 0, \quad (30)$$

$$U(x) \geqslant C_1 \qquad \text{for } x \notin D_1$$

$$U(x) \leqslant C_2 < C_1 \qquad \text{for } x \in D_2. \quad (31)$$

Then Theorem 1 holds without assumption (20).

The proof of Theorem 4 is given in Appendix IV.

Assumption (30) clearly makes $U(\cdot)$ a Lyapunov function in $\tilde{D}$, while (31) formalizes the intuitive notion of trajectories from D_2 never leaving D_1. We may remark that (30), (31) hold, e.g., if the trajectories of (19) do not intersect the boundary of D_1 "outwards" and D_2 is sufficiently close to D_1.

VII. How to Use the Theorems

The intuitive content of the theorems of Section V is that the algorithm (1), (2),

$$x(t) = x(t-1) + \gamma(t) Q(t, x(t-1), \varphi(t)) \quad (32a)$$

$$\varphi(t) = A(x(t-1))\varphi(t-1) + B(x(t-1))e(t) \quad (32b)$$

can be studied and analyzed in terms of the differential equation

$$\frac{d}{d\tau} x^D(\tau) = f(x^D(\tau)), \quad (33)$$

where

$$f(x) = \lim_{t \to \infty} EQ(t, x, \bar{\varphi}(t,x)). \quad (34)$$

The precise statements about the relations between (32) and (33) of Theorems 1–3 may be summarized in a somewhat looser language as follows.

a) $x(t)$ can converge only to stable stationary points of (33).

b) If $x(\cdot)$ belongs to the domain of attraction of a stable stationary point x^* of (33) i.o. w.p.1, then $x(t)$ converges w.p.1 to x^* as t tends to infinity.

c) The trajectories of (33) are "the asymptotic paths" of the estimates $x(\cdot)$, generated by (32).

These statements are fairly attractive intuitively, and they suggest certain unified techniques to analyze recursive algorithms. We shall illustrate this below, but let us here point out some aspects.

By the result a) the possible convergence points of (32) may be determined and studied. That a possible convergence point must be a zero of (34) is fairly obvious and it may be derived without reference to any differential equation. However, the observation that among these stationary points only stable ones are candidates for being limit points of (32) is a most important complement and it is probably less obvious without the present interpretation in terms of the differential equation. Perhaps the main use of result a) is to prove failure of convergence. It may be remarked that usually an algorithm is constructed so that the desired limit indeed is a stationary point. Consequently the possible lack of convergence is then due to the unstable character of the stationary point, so it is the complement (25b) that is the key result for proving divergence.

Result b) is the result by which convergence can be proved. In many cases it is not easy to find a proper Lyapunov function to prove global stability of (33), and sometimes the RHS of (33) is quite complex. For certain algorithms, though, in particular those arising from criteria-minimization, it is possible to do this analytically, and some examples will be given below.

While analytic treatment of (33) may be difficult, it is always possible to solve it numerically when the dimension of x is not too large. In that way insight can be gained into the global stability properties of the differential equation, the stationary points and their character. In view of result c) the trajectories thus obtained are also relevant for the asymptotic behavior of the algorithm. Therefore, numerical solution of (33) is a valuable complement to simulation of (32). Due to the time scaling (26) in the differential equation, this reveals more rapidly the asymptotic properties and the stationary points of the algorithm. Since the estimates change more and more

slowly, due to (4), it is seldom not difficult to decide from simulations only whether the estimates have settled around a limit value or are just converging slowly. In addition, it might be difficult to tell from a simulation if a certain effect is an inherent feature of the algorithm or just depends on random influence. Numerical solution of (33) may resolve such questions.

The regularity conditions A, B, or C are usually not very difficult to verify as demonstrated below. Notice in particular that the observation φ may have a "dummy" character, since it may have to be extended to fit into the assumptions (2) and A.1. For example, if the sequence $\{e(\cdot)\}$ in (2) is a stationary process with rational spectral power density, then it can be modeled as the output of a stable, linear filter with independent random variables as input. Hence, by extending the vector φ and adjoining this filter to (2), assumption A.1 will hold. Even if this may lead to a vector φ with a large dimension, the complexity of the algorithm and of the analysis is not affected. What matters is the calculation of

$$EQ(t,\bar{x},\bar{\varphi}(t,\bar{x}))$$

and this can often be done without explicit expressions for $A(\cdot), B(\cdot), \varphi(\cdot)$, and even for $Q(\cdot,\cdot,\cdot)$.

In the next section we shall apply the method to a few examples and illustrate how the techniques of the items above may be used.

VIII. Examples

Example 1—Stochastic Approximation Algorithms

Consider the problem of solving

$$E_{\varphi}Q(x,\varphi)=0 \tag{35}$$

for x. Here "E_{φ}" denotes the expectation with respect to φ, while the vector x is considered as a fixed parameter. Quantities $Q(x,\varphi(t))$, $t=1,2,\cdots$, are available for any x, where the distribution of the random vector $\varphi(\cdot)$ does not depend on x. Robbins–Monro [11] proved that under certain assumptions the scheme ("the Robbins–Monro scheme")

$$x(t)=x(t-1)+\gamma(t)Q(x(t-1),\varphi(t)) \tag{36}$$

gives a sequence of estimates that converges to the (a) solution of (35) in the mean square sense. Convergence w.p.1 of (36) has then been studied in several papers, e.g., [12]–[14], and the theorems of these studies do not differ very much conceptually from Theorem 1. The functions used in the convergence theorems of, e.g., [12] or [13] can be interpreted as Lyapunov functions for the differential equation (33) and condition (20) of Theorem 1 is ensured by further conditions on this function, rather than by the more practically oriented Theorem 4, see, e.g., [12, condition A], or [13, condition B, p. 184]. It can also be remarked that the condition

$$E(x-x^*)^TQ(x,\varphi)<0 \tag{37}$$

frequently used in Tsypkin's work, see, e.g., [15], [16], clearly can be understood as a stability condition for (33) with $V(x)=\|x-x^*\|^2$ as the Lyapunov function. Tsypkin's condition

$$EQ^T(x,\varphi)Q(x,\varphi)\leqslant C(1+\|x\|^2) \tag{38}$$

is then a variant of the "boundedness condition."

The convergence results thus obtained are, however, essentially restricted to the case $\varphi(\cdot)$ being independent random variables ($A(\cdot)=0$) and $\gamma(\cdot)$ satisfying (17) which is quite restrictive for control and estimation application. These conditions are inherently tied to the use of martingale theory in the proofs and cannot easily be dispensed with. Our Theorem 1 when applied to (36) is thus more general in that $\varphi(\cdot)$ may be dependent (generated as white noise through a linear filter) and $\gamma(\cdot)$ has only to satisfy A.7. This is satisfied, e.g., for $\gamma(t)=Ct^{-\alpha}$ $0<\alpha\leqslant 1$, while (17) admits only $1/2<\alpha\leqslant 1$. Notice that slowly decreasing gain sequences may be of interest in practice to achieve fast convergence of the sequence of estimates. We must, however, admit that we in return require more regularity of Q and of $e(\cdot)$. On the other hand, nonsmoothness of the involved functions is seldom a problem in applications, and we believe that our version of the convergence theorem is more widely applicable.

In addition, Theorems 2 and 3 are important results for convergence analysis, and we are not aware of similar previous results for the Robbins–Monro scheme.

In many applications it is of interest to minimize a function $E_{\varphi}J(x,\varphi)=P(x)$ with respect to x. If the derivative of J with respect to x can be calculated, the stationary points of $P(x)$ can be found as solutions of

$$E_{\varphi}\left[\frac{\partial}{\partial x}J(x,\varphi)\right]=0.$$

This is a problem that can be solved using the Robbins–Monro scheme and then Corollary 2 of Theorems 1 and 2 is quite useful.

If the derivative of J cannot be calculated it seems natural to replace it with some difference approximation. This was suggested by Kiefer and Wolfowitz [17] and their procedure has also been used for various control and estimation problems. Kushner has in several recent papers discussed interesting variants of this procedure, see e.g., [18], [19]. Our theorems are not directly applicable to the Kiefer–Wolfowitz scheme as they stand, since condition A.3 (or B.3) is not valid. The reason is that the function Q in this case increases to infinity with t. For the case of additive noise to the function to be minimized, however, it can readily be shown that Theorems 1–4 hold anyway. Details are given in [5] and [7].

Stochastic approximation algorithms have been applied to a broad variety of problems in control theory, see e.g., Tsypkin [15], [16], Fu [20], and Saridis *et al.* [21]. The approach is known as "learning systems," and in this framework estimation and identification problems, adaptive control, supervised and unsupervised pattern recognition, etc. can be treated.

An approach that is related to stochastic approximation

is suggested by Aizerman *et al.* [13]. Their "Potential Function Method" can be applied to various problems in machine learning.

Therefore the Robbins–Monro scheme appears in various disguises in many control and estimation algorithms, and consequently the described techniques can be applied to these. A particular example is given below. □

Example 2–An Automatic Classifier

A classifer receives scalar valued signals $\varphi(t)$ which may belong to either of two *a priori* unknown classes A and B. The classifier must find a classification rule, i.e., a number $c(t)$ such that $\varphi(t)$ is classified as A if $\varphi(t) \leqslant c(t)$ and B otherwise. The number $c(t)$ can, e.g., be determined as follows:

$$c(t)=\big(x^A(t)+x^B(t)\big)/2$$

where

$$x^A(t)=\begin{cases} x^A(t-1)+\gamma(t)\big[\varphi(t)-x^A(t-1)\big], \\ \quad \text{if } \varphi(t) \text{ is classified as } A \\ x^A(t-1), \qquad \text{otherwise.} \end{cases} \tag{39}$$

$x^B(t)$ is defined analogously. Clearly, $x^A(t)$ is the mean value of the outcomes classified as A. This scheme is discussed by Tsypkin [22] and Braverman [23].

Let $\varphi(t)$ have the distribution shown in Fig. 2 consisting of two triangular distributions. The probability of outcomes in the left triangle is λ. We assume that $\varphi(\cdot)$ is a sequence of independent random variables. Clearly, it is desirable that the classification rule, the number $c(t)$, should converge to some value between -1 and $+1$. Introduce

$$x(t)=\begin{bmatrix} x^A(t) \\ x^B(t) \end{bmatrix}.$$

Then (39) can be written

$$x(t)=x(t-1)+\gamma(t)Q\big(x(t-1),\varphi(t)\big) \tag{40}$$

where

$$Q(x,\varphi)=\begin{bmatrix} Q^A(x^A,x^B,\varphi) \\ Q^B(x^A,x^B,\varphi) \end{bmatrix}$$

and

$$Q^A(x^A,x^B,\varphi)=\begin{cases} \varphi-x^A, & \text{if } \varphi<\tfrac{1}{2}(x^A+x^B)-\delta \\ 0, & \text{if } \varphi>\tfrac{1}{2}(x^A+x^B)+\delta \end{cases}$$

and Q^B analogously

where the values for $\frac{1}{2}(x^A+x^B)-\delta \leqslant \varphi \leqslant \frac{1}{2}(x^A+x^B)+\delta$ are such that Q^A is a continuously differentiable function of φ and x. Here δ is some small positive number.

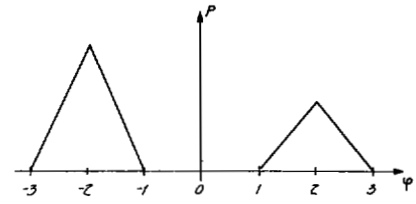

Fig. 2. Probability density function of the random variable to be classified by the automatic classifier.

Clearly, the algorithm (40) together with the observation equation

$$\varphi(t)=e(t) \tag{41}$$

is a simple case of (32). Since $\varphi(\cdot)=e(\cdot)$ is bounded we may use assumptions A. Obviously A.1 to A.5 are satisfied, and let us assume that $\gamma(\cdot)$ is such that A.6–A.9 hold. (Here A.3 holds in virtue of our somewhat artificial modification of Q^A; but this example will illustrate that a heuristic use of the present convergence results will reveal important features of the algorithm.)

$E_\varphi Q(x,\varphi)=f(x)$ is readily computed as follows. For a given x the corresponding classification point is $c(x)=(x^A+x^B)/2$. $f^A(x)$ is then the mean value of the distribution left of the point $c(x)$, minus x^A. $f^B(x)$ is found correspondingly. The algebraic expression for $f(x)$ as a function of x and λ is lengthy and is omitted.

We first note that by construction, the estimates are confined to the area $\bar{D}$: $3>x^B \geqslant x^A>-3$. Therefore condition (20) of Theorem 1 is trivially satisfied. Analytical treatment of the differential equation $\dot{x}=f(x)$ is not easy, but its trajectories can easily be determined by numerical solution and they are shown in Fig. 3 for two choices of λ. For the case $\lambda=0.5$, [Fig. 3(a)] there is convincing evidence that the point $x^*=(-2,2)$ is a stable stationary point with global domain of attraction. Therefore, for $\lambda=0.5$ it follows from Theorem 1 that $x(t)\to x^*$ w.p.1 as $t\to\infty$, which gives a correct classification rule $c^*=0$. The case $\lambda=0.99$ [Fig. 3(b)] corresponds to a common situation where errors that occur rather seldom (1 percent), "outliers," shall be detected. In this case there are two stable stationary points of the differential equation, $x^*=(-2,2)$ and $x^{**}=(-2.3,-1.4)$. There is obviously a nonzero probability that $x(t)$ belongs to the domain of attraction of x^{**} i.o. Therefore Theorem 1 shows that for $\lambda=0.99$, and for any starting value $x(0)$ there is a nonzero probability (that depends on $x(0)$) that $x(t)\to x^{**}$ as $t\to\infty$. This gives an asymptotic classification rule $c^{**}=-1.8$, that classifies 39 percent of the "correct values" as outliers. For this case simulations of the classifier are shown in Fig. 4. In fact, the simulation leading to the undesired value c^{**} appeared only after several (257) attempts and from simulations only it might have been tempting to conclude general convergence to c^*.

In this example it is cumbersome to find a suitable Lyapunov function for the stability problem. However, as seen in Fig. 3 numerical solution of the differential equation yields sufficient insight into the stability properties. Such detailed information can naturally be obtained only if the dimensionality of the problem is small.

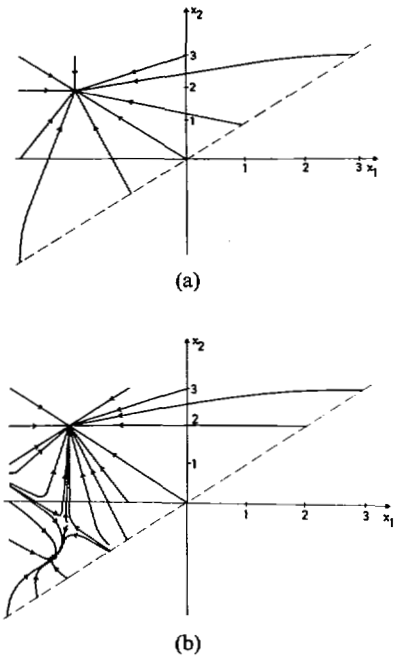

Fig. 3. Trajectories for the ODE that is associated with self-learning classifier (40). (a) $\lambda=0.5$. (b) $\lambda=0.99$.

Example 3—Equation Error Identification Methods

A common way of modeling dynamic systems is as a vector difference equation (VDE),

$$y(t)+A_1y(t-1)+\cdots+A_ny(t-n) = B_1u(t-1)+\cdots+B_nu(t-n), \quad (42)$$

where $y(t)$ and $u(t)$ are column vectors and A_i and B_i are matrices of appropriate dimensions. Introduce

$$\theta=(A_1\cdots A_nB_1\cdots B_n)^T$$
$$\psi(t)=\left(-y(t-1)^T\cdots-y(t-n)^Tu(t-1)\cdots u(t-n)^T\right)^T. \quad (43)$$

Then (42) can be written

$$y(t)=\theta^T\psi(t). \quad (44)$$

We may remark that (44) also covers several other interesting estimation problems, not necessarily related to system identification.

Usually, the true system cannot be described exactly in the form (44). Suppose that it can be described as

$$y(t)=\theta_0^T\psi(t)+v(t) \quad (45)$$

where $v(\cdot)$ is a disturbance that can be modeled as

$$v(t)=D(q^{-1})e_1(t). \quad (46)$$

Here $D(q^{-1})$ is a matrix with rational functions of the backward shift operator q^{-1} as entries and $e_1(\cdot)$ is a stationary sequence of independent random vectors with finite moments. It is assumed that the denominator polynomials in $D(z)$ (z replacing q^{-1}) have all roots outside the unit circle, i.e., $D(q^{-1})$ is an exponentially stable filter.

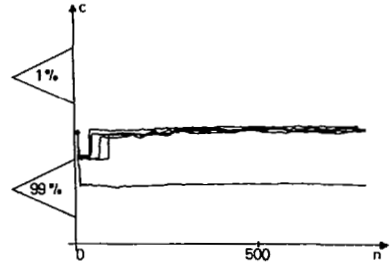

Fig. 4. Simulations of the classifier (40) for the case $\lambda=0.99$.

Even if an exact description of the system is impossible, a θ can be determined that gives a model (44) which describes the recorded data as well as possible. Often θ is determined by minimizing a criterion based on the equation error

$$\|y(t)-\theta^T\psi(t)\|^2. \quad (47)$$

Several algorithms based on the idea of somehow minimizing (47) have been suggested in the literature, see e.g., [24] and also [25] for a comprehensive treatment. The probably best known method of this type is the least squares algorithms, see, e.g., [24]. Then the sum

$$\sum_{t=1}^{N}\|y(t)-\theta^T\psi(t)\|^2 \quad (48)$$

is minimized w.r.t θ to obtain the estimate $\theta(N)$ based on measurement up to time N. An important and well known feature of this method is that the sequence of estimates can be obtained recursively as

$$\theta(t)=\theta(t-1)+\gamma(t)\mathcal{K}(t)\left[y(t)-\theta(t-1)^T\psi(t)\right]^T \quad (49a)$$
$$\mathcal{K}(t)=R^{-1}(t-1)\psi(t) /\left[1+\gamma(t)\left(\psi(t)^TR^{-1}(t-1)\psi(t)-1\right)\right] = R^{-1}(t)\psi(t) \quad (49b)$$
$$R(t)=R(t-1)+\gamma(t)\left[\psi(t)\psi(t)^T-R(t-1)\right] \quad (49c)$$

(usually (49c) is written in terms of $R^{-1}(t)$, which makes it of "Riccati type"). For the minimization of (48) $\gamma(t)$ has to be taken as $1/t$. Other sequences $\gamma(\cdot)$ correspond to criteria where old measurements are discounted, which often is relevant in practice.

Let us assume that the input to the process is determined as

$$u(t)=F(q^{-1})e_2(t)+H\left(q^{-1};\theta(t-1)\right)y(t) \quad (50)$$

where $F(q^{-1})$ and $H(q^{-1},\theta)$ are matrices with rational functions of the backward shift operator q^{-1} as entries. Let $e_2(\cdot)$ be a stationary sequence of random vectors with finite moments, that are mutually independent and also independent of $e_1(\cdot)$. Moreover, $H(q^{-1},\theta)$ is a causal operator that allows output feedback terms in the input.

This feedback law may depend on the current parameter estimate as is further discussed in Example 5.

It is clear that the rational filters in (46) and (50) can be represented in a state space form,

$$z_v(t+1)=A_v z_v(t)+B_v e_1(t+1); \qquad v(t)=(I\ 0\cdots 0)z_v(t) \tag{51}$$

$$z_u(t+1)=A_u(\theta(t))z_u(t)+z_u+B_{ue}e_2(t+1)+B_{uy}y(t+1);$$
$$u(t)=(I\ 0\cdots 0)z_u(t), \tag{52}$$

where $z_v(\cdot)$ and $z_u(\cdot)$ are the corresponding state vectors of appropriate dimensions. We may now form the "observation vector,"

$$\varphi(t)=\left[y(t)^T,\psi(t)^T,z_v(t)^T,z_u(t)^T\right]^T \tag{53}$$

which obeys

$$\varphi(t)=A_\varphi(\theta(t-1))\varphi(t-1)+B_\varphi\begin{pmatrix}e_1(t)\\ e_2(t)\end{pmatrix} \tag{54}$$

where the matrix $A_\varphi(\cdot)$ is formed from (45), (43), (51), (52) in an obvious manner. Its eigenvalues are the poles of the filters $D(q^{-1})$, $F(q^{-1})$ and of the closed loop system which is obtained for (42) with a constant feedback (50) using $\theta(t-1)$. There are also a number of eigenvalues in the origin, arising from the shifting of the vector $\psi(t)$. Notice that $A_\varphi(\theta)$ depends on θ only since the feedback filter $H(q^{-1};\theta)$ does. Let us take

$$x(t)=\left(\theta(t)^T \operatorname{col}^T R(t)\right)^T. \tag{55}$$

Then (49) takes the form

$$x(t)=x(t-1)+\gamma(t)Q(t;x(t-1);\varphi(t)) \tag{56}$$

with an obvious definition of $Q(t;x,\varphi)$ from (49). Therefore the algorithm (49) together with (54) is of the general form (32). Let us check if assumptions B of Section IV are satisfied. Conditions B.1 and B.2 are satisfied due to our assumptions. By straightforward calculations it is readily shown that B.3 is satisfied in the open area $D_R=\{x|R>0\}$, [cf., (55)] e.g., with

$$\mathcal{K}_1(x,\varphi,\rho,v)=(|\theta|+\rho)(1+|\varphi|+v)^2/(1-\rho|R^{-1}|)^2 \tag{57}$$

for $\rho=\rho(x)<1/|R^{-1}|$. Then B.4 will be satisfied with

$$\mathcal{K}_2(x,\varphi,\rho,v,w)=(|\theta|+\rho)(|\varphi|+2w+v)/(1-\rho|R^{-1}|)^2. \tag{58}$$

Condition B.5 is satisfied if the matrix $H(q^{-1};\theta)$ is Lipschitz continuous in θ. For condition B.6 we define

$$\bar{\varphi}(t,\bar{x})=A_\varphi(\bar{\theta})\bar{\varphi}(t-1,\bar{x})+B_\varphi\begin{pmatrix}e_1(t)\\ e_2(t)\end{pmatrix}$$
$$\left(\bar{x}=(\bar{\theta}^T\operatorname{col}^T\bar{R})^T\right).$$

Since $e_i(\cdot)$ are stationary, $\bar{\varphi}(t,\bar{x})$ will approach stationarity exponentially, for all x, such that $\bar{\theta}$ makes the closed-loop system stable. Therefore the limits

$$f(\bar{\theta})=\lim_{t\to\infty}E\bar{\psi}(t,\bar{x})\left[\bar{y}(t,\bar{x})-\bar{\theta}^T\bar{\psi}(t,\bar{x})\right]^T \tag{59a}$$

$$G(\bar{\theta})=\lim_{t\to\infty}E\bar{\psi}(t,\bar{x})\bar{\psi}(t,\bar{x})^T \tag{59b}$$

are well defined where $\bar{y}(t,\bar{x})$ and $\bar{\psi}(t,\bar{x})$ are the corresponding parts of $\bar{\varphi}(t,\bar{x})$, and

$$\lim_{t\to\infty}EQ(t;\bar{x},\bar{\varphi}(t,\bar{x}))=\begin{pmatrix}\bar{R}^{-1}f(\bar{\theta})\\ \operatorname{col}\left(G(\bar{\theta})-\bar{R}\right)\end{pmatrix} \tag{60}$$

so B.6 is satisfied. Moreover, from (57) and (58) it follows that B.7 holds, since all moments of $\bar{\varphi}(t,x)$ and $v(t,\lambda,c)$ exist. Conditions B.8–B.11 about the sequence $\gamma(\cdot)$ are assumed to be satisfied.

The conclusion therefore is that the differential equation

$$\frac{d}{d\tau}\theta(\tau)=R^{-1}(\tau)f(\theta(\tau)) \tag{61a}$$

$$\frac{d}{d\tau}R(\tau)=G(\theta(\tau))-R(\tau) \tag{61b}$$

can be associated with the algorithm (49). In the remaining part of this example, we shall assume that the feedback matrix H *does not depend on* θ (i.e., there is no adaptive feedback), that the matrix $F(z)$ has full rank *a.e.* z and that $e_i(\cdot)$ are full rank processes. (Adaptive feedback is further discussed in Example 5.) This means that the matrix $A_\varphi(\cdot)$ does not depend on θ, $\bar{\varphi}(t,\bar{x})=\varphi(t)$, and $\bar{y}(t,\bar{x})=y(t)$, so the values in (59) are directly defined in terms of input–output covariances. In particular, the matrix G is independent of θ; $G(\theta)=G$.

Introduce

$$r=E\varphi(t)v(t)^T \tag{62}$$

and we have, using (45),

$$f(\theta)=G\cdot(\theta_0-\theta)+r. \tag{63}$$

Hence, (61) can be rewritten as

$$\frac{d}{d\tau}\theta(\tau)=R^{-1}(\tau)G\left[\left(\theta_0+G^{-1}r\right)-\theta(\tau)\right] \tag{64a}$$

$$\frac{d}{d\tau}R(\tau)=G-R(\tau). \tag{64b}$$

With

$$\tilde{\theta}(\tau)=\theta(\tau)-\left(\theta_0+G^{-1}r\right)$$

and

$$V(\tilde{\theta},R)=\tilde{\theta}^TR\tilde{\theta}$$

we have

$$\frac{d}{d\tau}V(\tilde{\theta}(\tau),R(\tau))=-2\tilde{\theta}^TG\tilde{\theta}+\tilde{\theta}^T(G-R(\tau))\tilde{\theta}$$
$$=-\tilde{\theta}^T(\tau)(G+R(\tau))\tilde{\theta}(\tau)\leqslant 0 \quad (65)$$

so that V is a Lyapunov function for (64) or (61) that assures that the stationary point

$$\theta^*=\theta_0+G^{-1}r \quad (66)$$

has a domain of attraction equal to D_R. Therefore condition (21) of Theorem 1 is satisfied with $D_A=D_R$. To check condition (20) we note first that the assumption on full rank and finite moments of $e_i(\cdot)$ implies that $CI>G>\delta I$ for some $\delta>0$, $C<\infty$. Therefore also $(\delta/2)I<R(t)<2CI$ and $|\varphi(t)|<2C$ i.o. w.p.1. We also note that (49a) can for large $\theta(t-1)$ be written

$$\theta(t)\approx\left(I-R^{-1}(t)\psi(t)\psi(t)^T\right)\theta(t-1)$$

which shows that $\theta(t)$ can, w.p.1, not tend to infinity. Hence $x(t)$ belongs to a compact subset of D_R i.o. w.p.1 and condition (20) is satisfied. Theorem 1 now implies that

$$\theta(t)\to\theta^* \qquad \text{w.p.1 as } t\to\infty.$$

In particular, we see that the least squares estimate is consistent only if $r=0$, which essentially is the same as requiring that $v(\cdot)$ is a sequence of uncorrelated random variables, and that the current $u(t)$ is uncorrelated with future $v(s)$, $s\geqslant t$.

Other variants of equation error methods are treated analogously.

These facts are, of course, well known, [24], [25], at least for the case $\gamma(t)=1/t$, but one reason for this example is that the analysis extends into less trivial problems. □

Example 4–Equation Error Methods–Assumptions C

The model (44) and the criterion (48) appear in several contexts, like curve fitting, etc., where a stochastic framework often is not imposed. Let $\gamma(t)=1/t$ and denote

$$\psi(t)y(t)^T=\eta(t) \quad (67a)$$
$$\psi(t)\psi(t)^T=\Lambda(t). \quad (67b)$$

Then the algorithm (49) can be written

$$\Theta(t)=\Theta(t-1)+\frac{1}{t}R^{-1}(t)\left[\eta(t)-\Lambda(t)\Theta(t-1)\right] \quad (68a)$$

$$R(t)=R(t-1)+\frac{1}{t}\left[\Lambda(t)-R(t-1)\right]. \quad (68b)$$

We shall in this example illustrate how assumptions C can be applied to the algorithm (68) to infer convergence directly from the properties of the sequences $\{\eta(t)\}$ and $\{\Lambda(t)\}$. Take in this example the observation to be

$$\varphi(t)=\begin{pmatrix}\eta(t)\\ \operatorname{col}\Lambda(t)\end{pmatrix} \quad (69)$$

and in (32b) $A(x)\equiv 0$, $B(x)\equiv I$. The estimate is as before given by (55). The condition C.1 has already been verified and $\mathcal{K}_1$ can this time be taken as

$$\mathcal{K}_1(x,\varphi,\rho,v)=1+|x|+\rho+|\Lambda|+v. \quad (70)$$

Condition C.2 is trivially satisfied. The variable $z(t,\bar{x})$ defined by (18a) is given by

$$z(t,\bar{x})=\begin{pmatrix}\frac{1}{t}\bar{R}^{-1}\sum_{s=1}^{t}\left[\eta(s)-\Lambda(s)\bar{\Theta}\right]\\ \operatorname{col}\frac{1}{t}\sum_{s=1}^{t}\left[\Lambda(s)-\bar{R}\right]\end{pmatrix}$$
$$=\begin{pmatrix}\bar{R}^{-1}\left\{\left[\frac{1}{t}\sum_{s=1}^{t}\eta(s)\right]-\left[\frac{1}{t}\sum_{s=1}^{t}\Lambda(s)\right]\bar{\Theta}\right\}\\ \operatorname{col}\left[\frac{1}{t}\sum_{s=1}^{t}\Lambda(s)\right]-\operatorname{col}\bar{R}\end{pmatrix} \quad (71)$$

where

$$\bar{x}=\begin{pmatrix}\bar{\Theta}\\ \operatorname{col}\bar{R}\end{pmatrix}.$$

Therefore, if the sequences $\{\eta(s)\}$ and $\{\Lambda(s)\}$ are such that

$$\frac{1}{t}\sum_{s=1}^{t}\eta(s)\to\bar{\eta} \quad (72a)$$

and

$$\frac{1}{t}\sum_{s=1}^{t}\Lambda(s)\to\bar{\Lambda}, \quad (72b)$$

then C.3 holds with

$$f(\bar{x})=\begin{pmatrix}\bar{R}^{-1}\left(\bar{\eta}-\bar{\Lambda}\bar{\Theta}\right)\\ \operatorname{col}\bar{\Lambda}-\bar{R}\end{pmatrix}. \quad (73)$$

Since A and B do not depend on x (see C.7) condition C.4 is satisfied if

$$\frac{1}{t}\sum_{s=1}^{t}|\Lambda(s)| \quad \text{converges} \quad (74)$$

which is implied by (72b) if $\Lambda(s)=\psi(s)\psi(s)^T$. The boundedness condition of Theorem 1 has been verified in

Example 3. The differential equation

$$\frac{d}{d\tau}\Theta(\tau)=R^{-1}(\tau)\left[\bar{\eta}-\bar{\Lambda}\Theta(\tau)\right] \tag{75a}$$

$$\frac{d}{d\tau}R(\tau)=\bar{\Lambda}-R(\tau) \tag{75b}$$

is as shown in Example 3 globally asymptotically stable with stationary point

$$\Theta^*=\bar{\Lambda}^{-1}\bar{\eta} \tag{76}$$

if $\bar{\Lambda}$ is positive definite. Hence Theorem 1 implies that $\Theta(t)$ given by (68) converges to Θ^* defined in (76) for any sequences $\{\Lambda(s)\}$ and $\{\eta(s)\}$ such that (72) holds with a positive definite $\bar{\Lambda}$. This is by no means surprising since (68) in fact only is a way of recursively writing

$$\Theta(t)=\left[\sum_{s=1}^{t}\Lambda(s)\right]^{-1}\sum_{s=1}^{t}\eta(s)$$

using the matrix inversion lemma. However, it illustrates that our method of analysis does not bring in "unnecessary" assumptions, even though it is concerned with a more general algorithm.

Moreover, for a stochastic approximation version of (68),

$$\Theta(t)=\Theta(t-1)+\frac{1}{t}\left[\eta(t)-\Lambda(t)\Theta(t-1)\right] \tag{77}$$

we have exactly the same convergence result under the condition (72), since the associated differential equation then is

$$\frac{d}{d\tau}\Theta(\tau)=\bar{\eta}-\bar{\Lambda}\Theta(\tau). \tag{78}$$

In this case no explicit expression for $\Theta(t)$ is available to infer this result directly. In fact, the algorithm (77) has been studied in [26], [27] using fairly elaborate methods, since it was found that the usual stochastic approximation convergence results could not be applied due to the correlation in the sequences $\{\Lambda(s)\}$ and $\{\eta(s)\}$. ■

These four examples have all been for the case where $A(\cdot)$ in (32b) actually does not depend on x. The convergence part in this case can, at least under further assumptions, not seldom be treated by more conventional statistical methods. When $A(\cdot)$ does depend on x, conventional approaches become much more difficult, and in fact, also in the proof of Theorem 1, a major burden is to keep control over the coupled stability questions in (32a) and (32b). The inclusion of x-dependent A-matrices becomes necessary for more elaborate recursive identification methods in which the observed data are processed through filters that are formed from current parameter estimates. This is the case, e.g., for the extended Kalman filter, the extended least squares method, output error methods, like Landau's scheme [28] etc. Since a companion paper [29] is devoted to the last two schemes we shall not go into any detail here. Further analysis of some recursive identification schemes using the present approach can also be found in [1], [30], and [31].

It was remarked above that the structure (32) can be understood to be patterned after adaptive control of linear systems. We shall therefore conclude with such an example.

Example 5—Self-Tuning Regulators

We shall in this example discuss an application to the self-tuning regulator, described in [32]; see also [2] and [4]. This regulator is based on least squares identification, (49), and the output feedback law is determined from the current parameter estimates. Usually the feedback law is chosen to be a minimum variance regulator, [32], but here it could be a general linear regulator as in (50), where perhaps in most cases F is zero.

In this case the matrix $A_\varphi(\theta)$ in (54) does depend on θ since the feedback term does. However, the point now is that, in contrast to conventional analysis of the least squares algorithm, most of what was said in Example 3 still holds. Up to (61) the development was quite general. This differential equation is valid also in the case of adaptive feedback, although G and r now are functions of θ. If $v(\cdot)$ is a sequence of independent random variables, then $r=0$ and (64) and (65) hold. (Equation (65) does not hold if $r\neq 0$ depends on θ.) We therefore still find that the points defined by

$$D_c=\{\theta \mid f(\theta)=0\}$$

form an invariant set with global domain of attraction. Clearly $\theta_0\in D_c$, and whether D_c contains more points depends on the choice of feedback law and model order. There is a further complication before Theorem 1 can be applied. In this case the area D_S is unknown, i.e., the area of such θ that inserted in a constant feedback law (50) makes the closed-loop system stable ($A_\varphi(\theta)$ in (54) has all eigenvalues inside the unit circle). Therefore we cannot guarantee stability by projecting θ into D_S. Hence condition (20) of Theorem 1 has to be verified by other considerations, e.g., by showing that the overall system has a certain stability property as in [33]. But when this is shown, Theorem 1 proves convergence of $\theta(t)$ into D_c w.p.1. Let us repeat that this holds for the case of arbitrary feedback law, but under the assumption that $v(\cdot)$ is white noise. For general noise $v(\cdot)$ the convergence analysis is more cumbersome, but it can be performed in certain special cases, [2], [4], [29]. We refer also to these papers for more details on how Theorems 1, 2, and 3 can be used in the analysis of self-tuning regulators. Numerical solution of the associated differential equation has turned out to be a valuable tool here, and it has been used in [34] as well as in the references above.

IX. Conclusions

Recursive algorithms like (1) have been analyzed in various contexts. However, we would like to stress again that when φ in (1) is generated as in (2), the analysis becomes more difficult. The reason is that (1) no longer is recursive in x for analysis purposes: the whole history of $x(\cdot)$ enters in each step of (1). Moreover, the coupled stability problems between (1) and (2) are intricate. But the structure (1), (2) is nonetheless common in estimation and control problems; a typical example is adaptive control of linear stochastic systems. The analysis of this case is also known to be usually very difficult.

With the present approach we are able to give a general treatment of (1), (2) under assumptions that do not appear to be restrictive. The examples indicate that the theorems may be applied to rather diverse problems, and perhaps the technique also may serve as a basis for a unified approach to the analysis of adaptive controllers. In addition, an extension is obtained for the conventional convergence results in the simple case where $A(\cdot)$ in (2) is independent of x. We may remark that the analysis is restricted to the asymptotic behavior, convergence, possible convergence points etc. of the algorithm. Two related algorithms which are associated with the same differential equation may differ noticably in transient behavior and convergence rate.

In the described theory, we would like to stress the intuitive content of the theorems and the methodology of analysis as outlined in Section VII. It is no doubt important to appreciate the exact formulations of the theorems and to know the exact conditions under which they are valid. But it is perhaps equally rewarding to use the properly defined differential equation as a general instrument for analysis in a more heuristic fashion. This may be exemplified in Theorem 3, which has a fairly technical formulation and is probably more valuable as a "moral support" for studying the trajectories of the differential equation, than in its literal sense.

Appendices

The proofs of Theorems 1–4 are given in Appendices I–IV. They are slightly compressed versions of the proofs in [7], in that some calculations of technical nature are omitted. Proofs for a simple special case of the algorithm (1), (2) are also given in [6]. These are naturally less technical than the present ones, and reveal perhaps more clearly the underlying ideas. The basic path of the proofs of Theorems 1 and 3 follows the heuristic outline of Section III.

Some notational conventions in the proofs should be noted. "C" will denote any constant, that need not be the same in different parts of the proof. Important dependencies of the constants will be given as arguments, while indexed constants are "global" throughout the proof. "$\mathfrak{B}(\bar{x},\rho)$" denotes as before an open ρ-neighborhood of $\bar{x}$. Realizations in the sample space Ω will be denoted by "ω." The abbreviation "i.o." denotes as before "infinitely often."

Appendix I
Proof of Theorem 1

Outline

In order to prove Theorem 1, we shall first show that the estimates provided by the algorithm locally and asymptotically follow the trajectories of the associated differential equation. This will be done in Lemma 1 under assumptions C and the proof of this follows the intuitive outline of Section III. After that the local behavior of the algorithm is thus established, this is used to prove that all cluster points of the sequence $\{x(t)\}$ must belong to D_c. This is done by means of a Lyapunov function, the existence of which is inferred from the stability condition (21). A possible clusterpoint outside the set D_c would yield a decreasing value of the Lyapunov function along the corresponding trajectory of the differential equation. Since Lemma 1 proves that we follow this trajectory asymptotically, a contradiction is obtained. The value of the Lyapunov function is decreasing outside D_c and it is not possible to return a given point outside D_c infinitely often. This is the intuitive path of proof. The many technicalities tend to obscure the simple idea, and the proof will be structured as much as possible to enhance the basic ideas.

Finally it is proved in Lemma 2 that assumptions A or B imply assumptions C for the case $\gamma(n)=1/n$. An outline of the general proof will also be given, but the details of this are omitted, and the reader is referred to [7] for them.

Lemma 1: Consider the algorithm (1), (2) under assumptions C. Let $\bar{x}\in D_R$ and define the number $m(n,\Delta\tau)$ such that

$$\sum_{n}^{m(n,\Delta\tau)} \gamma(t)\to\Delta\tau \quad \text{as } n\to\infty. \tag{I.1}$$

Assume that

$$x(n)\in\mathfrak{B}(\bar{x},\rho) \tag{I.2a}$$

where

$$\rho=\rho(\bar{x}) \quad \text{is sufficiently small} \tag{I.2b}$$

and that

$$|\varphi(n)|\leqslant C_\varphi. \tag{I.3}$$

Then, there exists a value $\Delta\tau_0=\Delta\tau_0(\bar{x},\rho)$ and a number $N_0=N_0(\bar{x},\rho)$ such that for $\Delta\tau<\Delta\tau_0$ and $n>N_0$

$$x(m(n,\Delta\tau))=x(n)+\Delta\tau f(\bar{x})+q_1(n,m,\bar{x})+q_2(n,m,\bar{x}) \tag{I.4}$$

where $m=m(n,\Delta\tau)$ and

$$q_1(n,m,\bar{x})\to 0 \qquad \text{as } n\to\infty \tag{I.5}$$

and

$$|q_2(n,m,\bar{x})| \leq \Delta\tau \cdot C_1 |x(n)-\bar{x}| + C_2 \Delta\tau^2 \qquad \text{(I.6)}$$

where C_1 and C_2 depend on x and ρ, but not on $n > N_0$ or $\Delta\tau$.

Proof of Lemma 1: The proof is structured into 4 steps. In *Step 1* an explicit expression for $x(j)$ where

$$n \leq j \leq m(n,\Delta\tau) \qquad \text{(I.7)}$$

is derived. This expression shows that if certain terms are small, then $x(j)$ is close to what would be obtained if $Q(t,x(t-1),\varphi(t))$ is replaced by $Q(t,\bar{x},\bar{\varphi}(t,\bar{x}))$ [with $\bar{\varphi}(t,\bar{x})$ defined by (15)]. As a first step to show that these terms indeed are small an expression for $\varphi(t)-\bar{\varphi}(t,\bar{x})$ is derived in *Step 2*. This expression is used in *Step 3* to prove that assumptions C.3 and C.4 imply that the terms are small. A complication so far has been that in order to prove these things it must be assumed that $x(s)$ remains in a small neighborhood of $\bar{x}$ for $n \leq s \leq j$. In *Step 4* it is proved that $x(s)$ actually will remain in this neighborhood up to $s = m(n,\Delta\tau)$ if $\Delta\tau$ is chosen sufficiently small (and dependent only on $\bar{x}$) and n sufficiently large. This will conclude the proof.

Introduce $j = j(n)$ such that

$$n \leq j \leq m(n,\Delta\tau) \qquad \text{(I.8)}$$

and

$$x(k) \in \mathcal{B}(\bar{x},2\rho), \qquad k = n, n+1, \cdots, j-1. \qquad \text{(I.9)}$$

Step 1—An Expression for x(j)

Directly from the algorithm (1) the following expression is obtained:

$$\begin{aligned} x(j) &= x(n) + \sum_{s=n+1}^{j} \gamma(s) Q\big(s; x(s-1), \varphi(s)\big) \\ &= x(n) + \sum_{n+1}^{j} \gamma(s) Q\big(s; \bar{x}, \bar{\varphi}(s;\bar{x})\big) \\ &\quad + \sum_{n+1}^{j} \gamma(s)\big[Q\big(s, x(s-1), \varphi(s)\big) - Q\big(s;\bar{x},\bar{\varphi}(s;\bar{x})\big)\big] \end{aligned} \qquad \text{(I.10)}$$

where $\bar{\varphi}(s;\bar{x})$ is defined by (15). End Step 1.

The first goal is to show that the last term in (I.10) is "small," and in order to do this we first consider $\varphi(s) - \bar{\varphi}(s;\bar{x})$.

Step 2—An Estimate for $|\varphi(t)-\bar{\varphi}(t,\bar{x})|$

We claim that

$$\prod_{k=n}^{t} A(x(k)) \leq C\left[\frac{1+\lambda(\bar{x})}{2}\right]^{t-n} \triangleq C\tilde{\lambda}(\bar{x})^{t-n} \qquad \text{(I.11)}$$

if

$$x(k) \in \mathcal{B}(\bar{x},2\rho); \qquad n \leq k \leq t \quad \text{for sufficiently small}$$

$$\rho = \rho(\bar{x}) \qquad \text{(I.12)}$$

where $\lambda(\bar{x})$ is defined by (14).

The idea obviously is that since the matrix $A(\bar{x})$ is exponentially stable, exponential stability of the time-varying difference equation (2) given by $A(x(k))$ will be guaranteed if $x(k)$ varies in a sufficiently small neighborhood of $\bar{x}$. The formal proof consists of straightforward calculations given in [7], but omitted here. The "sufficiently small" ρ mentioned in (I.2b) refers to the fact that ρ should be so small that (I.12) implies (I.11).

From (2) it follows that

$$\varphi(t) = \left[\prod_{k=n}^{t} A(x(k))\right]\varphi(n) + \sum_{j=n}^{t}\left[\prod_{k=j}^{t} A(x(k))\right] B(j)e(j) \qquad \text{(I.13)}$$

which together with (I.11) and (I.3) implies that

$$|\varphi(t)| \leq C \cdot C_\varphi \tilde{\lambda}^{t-n} + \sum_{j=n}^{t} C\tilde{\lambda}^{t-j}|B(j)||e(j)| \qquad \text{(I.14)}$$

if (I.12) holds. Since the variable

$$\tilde{\varphi}(t) \triangleq \varphi(t) - \bar{\varphi}(t;\bar{x})$$

obeys

$$\begin{aligned} \tilde{\varphi}(t+1) &= A(\bar{x})\tilde{\varphi}(t) + \big[A(x(t)) - A(\bar{x})\big]\varphi(t) \\ &\quad + \big[B(x(t)) - B(x)\big]e(t+1) \end{aligned} \qquad \text{(I.15)}$$

we have that

$$\begin{aligned} \tilde{\varphi}(t) &= A(\bar{x})^{t-n}\tilde{\varphi}(n) \\ &\quad + \sum_{j=n}^{t-1} A(\bar{x})^{t-s}\big[\{A(x(j)) - A(\bar{x})\}\varphi(j) \\ &\quad + \{B(x(j)) - B(\bar{x})\}e(j+1)\big]. \end{aligned} \qquad \text{(I.16)}$$

If K_A and K_B denote the Lipschitz constants of $A(\cdot)$ and $B(\cdot)$, respectively (see assumption C.2), then from (I.16) and (14) it follows that

$$\begin{aligned} |\tilde{\varphi}(t)| &\leq \bar{C} \cdot \lambda(\bar{x})^{t-n}\{|\varphi(n)| + |\bar{\varphi}(n,\bar{x})|\} \\ &\quad + \bar{c}\sum_{j=n}^{t} \lambda(\bar{x})^{t-j}\big[K_A|x(j)-\bar{x}||\varphi(j)| \\ &\quad + K_B|x(j)-\bar{x}||e(j+1)|\big]. \end{aligned} \qquad \text{(I.17)}$$

Clearly, with $v(n,\lambda,c)$ defined by (16),

$$\bar{c}\{|\varphi(n)| + |\bar{\varphi}(n,\bar{x})|\} \leq v(n,\lambda(\bar{x}),c) \qquad \text{(I.18)}$$

for some constant c.

Now, inserting (I.14) and (I.18) in (I.17) gives

$$|\tilde{\varphi}(t)| \leqslant \lambda(\bar{x})^{t-n} v(n,\lambda(\bar{x}),c) + c\cdot\max(K_A,K_B) \cdot \max_{n\leqslant j\leqslant t-1} |x(j)-\bar{x}| \cdot \left\{ \sum_{j=n}^{t-1} \lambda(\bar{x})^{t-j}\tilde{\lambda}^{j-n}\cdot C_\varphi + \sum_{j=n}^{t-1}\sum_{k=n}^{j} \lambda(\bar{x})^{t-j}\tilde{\lambda}^{j-k}|e(k)| + \sum_{j=n}^{t-1} \lambda(\bar{x})^{t-j}|e(j+1)| \right\}.$$

Sum first over j in the double sum and introduce c' and λ such that

$$|k\tilde{\lambda}^k + \lambda^k| < c'\lambda^k \qquad \lambda < 1. \tag{I.19}$$

Then

$$|\varphi(t)-\bar{\varphi}(t,\bar{x})| \leqslant \lambda(\bar{x})^{t-n} v(n,\lambda(\bar{x}),c) + \max_{n\leqslant j\leqslant t-1} |x(j)-x|\cdot v(t,\lambda,c). \tag{I.20}$$

In particular,

$$\lambda(\bar{x})^{t-n} v(n,\lambda(\bar{x}),c) \leqslant v(t,\lambda(\bar{x}),c)$$

and therefore also

$$|\varphi(t)-\bar{\varphi}(t,\bar{x})| \leqslant v(t,\lambda,c) \tag{I.21}$$

where λ can be taken as $(3+\lambda(x))/4$ and c is a constant that is obtained from c_φ, $\bar{c}$ in (14), c in (I.11), c' in (I.19), K_A, K_B, and $\max|B(x)|(x\in\mathcal{B}(\bar{x},\rho))$. However, since assumption C.4 shall hold for any finite c, there is no point in tracing the expression for c through the calculations.

We summarize this step as follows: Choose 2ρ in (I.9) so small that (I.12) implies (I.11) for this ρ. Then (I.20) and (I.21) hold for $n\leqslant t\leqslant j$, where j is defined by (I.8), (I.9).

Note that in case the matrices A and B in (2) do not depend on x, then $\varphi(t)=\bar{\varphi}(t,\bar{x})$ and we may take $v(t,\lambda,c)\equiv 0$ in all calculations to come. This justifies the remark C.7. End Step 2.

We can now return to the expression (I.10) for $x(j)$. From assumption C.1 and (I.21) it follows that

$$\begin{aligned}|Q(s,x(s-1),\varphi(s)) - Q(s,\bar{x},\bar{\varphi}(s;\bar{x}))| &\leqslant \mathcal{K}_1(\bar{x},\bar{\varphi}(s,\bar{x}),2\rho,v(s,\lambda,c)) \cdot\left[|x(s-1)-\bar{x}| + |\varphi(s)-\bar{\varphi}(s;\bar{x})|\right] \\ &\leqslant \max_{n\leqslant k\leqslant s-1}|x(k)-\bar{x}|\cdot\mathcal{K}_1(\bar{x},\bar{\varphi}(s,\bar{x}),2\rho,v(s,\lambda,c)) \cdot\{1+v(x,\lambda,c)\} + \mathcal{K}_1(\bar{x},\bar{\varphi}(s,\bar{x}),2\rho,v(s,\lambda,c)) \cdot\lambda(\bar{x})^{s-n} v(n,\lambda,c)\end{aligned} \tag{I.22}$$

where the last inequality follows from (I.20).

We thus find that

$$x(j) = x(n) + \sum_{s=n}^{j} \gamma(s) f(\bar{x}) + R_1(n,j) + R_2(n,j) \tag{I.23}$$

where

$$R_1(n,j) = \sum_{s=n}^{j} \gamma(s)\left[Q(s,\bar{x},\bar{\varphi}(s;\bar{x})) - f(\bar{x})\right] \tag{I.24}$$

and

$$|R_2(n,j)| \leqslant \max_{n\leqslant k\leqslant j-1}|x(k)-\bar{x}|\cdot q_3(n,j) + q_4(n,j) \tag{I.25}$$

where

$$q_3(n,j) = \sum_{s=n}^{j}\gamma(s)\mathcal{K}_1(\bar{x},\bar{\varphi}(s;\bar{x}),2\rho,v(s,\lambda,c)) \cdot\{1+v(s,\lambda,c)\} \tag{I.26}$$

and

$$q_4(n,j) = \max_{n\leqslant s\leqslant j-1}\left\{\mathcal{K}_1(\bar{x},\bar{\varphi}(s;\bar{x}),2\rho,v(s,\lambda,c)) \cdot v(s,\lambda,c)\cdot\gamma(s)\right\}\cdot\sum_{s=n}^{j}\lambda(\bar{x})^{s-n}. \tag{I.27}$$

Step 3a—$\max_{n\leqslant j\leqslant m(n,\Delta\tau)}|R_1(n,j)|\to 0$ *as* $n\to\infty$

Let the maximum be attained for $j=j^*(n)$.

Solving (18a) we obtain

$$z(j^*(n),\bar{x}) = z(n,\bar{x}) + \sum_{k=n+1}^{j^*(n)} \gamma(k) \cdot\left[Q(k,\bar{x},\bar{\varphi}(k,\bar{x})) - z(k-1,\bar{x})\right].$$

Now, let n tend to infinity. By assumption C.3

$$\lim_{n\to\infty} z(n,\bar{x}) = \lim_{n\to\infty} z(j^*(n),\bar{x}) = f(\bar{x})$$

and hence

$$\begin{aligned}\lim_{n\to\infty}\sum_{k=n+1}^{j^*(n)}\gamma(k)\left[Q(k,\bar{x},\bar{\varphi}(k,x)) - z(k-1,\bar{x})\right] &= \lim_{n\to\infty}\sum_{k=n+1}^{j^*(n)}\gamma(k)\left[Q(k,\bar{x},\bar{\varphi}(k,\bar{x})) - f(\bar{x})\right] \\ &= \lim_{n\to\infty} R_1(n,j^*(n)) = 0\end{aligned}$$

End Step 3a.

Step 3b:

There exists an N_1 such that

$$|q_3(n,j(n))| \leqslant 2\Delta\tau\bar{k}_v \qquad \text{for } n > N_1,$$

where $\bar{k}_v$ is the limit of $k_v(t,\bar{x},\lambda,c)$ defined in C.4. This step is completely analogous to Step 3a, using the definition (18b) and assumption C.4.

End Step 3b.

Step 3c—$q_4(n,j(n))\to 0$ *as* $n\to\infty$

From (18b)

$$\gamma(s)\cdot v(s,\lambda,c)\cdot \mathcal{H}_1(\bar{x},\bar{\varphi}(s,\bar{x}),2\rho,v(s,\lambda,c)) = k_v(s)-k_v(s-1)+\gamma(s)k_v(s-1).$$

The RHS tends to zero according to C.4 and C.6 and the claim is proved. End Step 3c.

Expressions (I.23)–(I.27) together with Step 3 would complete the proof of Lemma 1, if $j(n)$ were equal to $m(n,\Delta\tau)$, that is if $x(k)$ remains in $\mathcal{B}(\bar{x},2\rho)$ for all k up to $m(n,\Delta\tau)$. This is proven in the final step.

Step 4

There exists a $\Delta\tau_0=\Delta\tau_0(\bar{x},\rho)$ and a $N_0=N_0(\bar{x},\rho)$ such that for $\Delta\tau<\Delta\tau_0$ and $n>N_0$, the expression (I.9) will hold for $j=m(n,\Delta\tau)$.

By the definition of $j(n)$

$$x(j-1)\in\mathcal{B}(\bar{x},2\rho) \tag{I.28}$$

and

$$\max_{n\leq k\leq j-1}|x(k)-\bar{x}|<2\rho. \tag{I.29}$$

We shall show that (I.28) and (I.29) imply that also

$$x(j)\in\mathcal{B}(\bar{x},2\rho) \tag{I.39}$$

for sufficiently large n and small $\Delta\tau$. By induction it then follows that

$$x(k)\in\mathcal{B}(\bar{x},2\rho)\qquad k=n,\cdots,m(n,\Delta\tau). \tag{I.31}$$

We have

$$|x(j)-\bar{x}|\leq|x(j)-x(n)|+|x(n)-\bar{x}| \leq|f(\bar{x})|\cdot\sum_{s=n}^{j}\gamma(s)+|R_1(n,j)|+|R_2(n,j)|+\rho \tag{I.32}$$

using (I.23) and (I.2). From (I.24)–(I.27) using Step 3b and (I.29) it follows that

$$\begin{aligned}|x(j)-\bar{x}|&\leq|f(\bar{x})|\cdot\Delta\tau+|R_1(n,j)| +4\rho\cdot\Delta\tau\cdot\bar{k}_v+|q_4(n,j)|+\rho\\ &\leq|f(\bar{x})|\cdot\Delta\tau+\max_{n\leq j\leq m(n,\Delta\tau)}|R_1(n,j)| +4\rho\cdot\Delta\tau\cdot k_v+|q_4(n,m(n,\Delta\tau))|+\rho.\end{aligned} \tag{I.33}$$

Now, choose

$$\Delta\tau<\Delta\tau_0=\min\left\{\frac{\rho}{4|f(\bar{x})|},\frac{1}{16\bar{k}_v}\right\} \tag{I.34}$$

and then n so large that

$$q_5(n,m(n,\Delta\tau))\triangleq\max_{n\leq j\leq m(n,\Delta\tau)}|R_1(n,j)| +|q_4(n,m(n,\Delta\tau))|<\rho/2 \tag{I.35}$$

which is possible according to Steps 3a–3c.

It now follows from (I.33)–(I.35) that

$$|x(j)-\bar{x}|<2\rho$$

and hence that (I.30) and (I.31) hold. End Step 4. From (I.33)–(I.35) it also follows that

$$|x(j)-\bar{x}|\leq|f(\bar{x})|\Delta\tau+q_5(n,m(n,\Delta\tau)) +4\bar{\rho}\Delta\tau\cdot\bar{k}_v+|x(n)-\bar{x}|. \tag{I.36}$$

Since the RHS does not depend on j, this inequality holds for all $n\leq j\leq m(n,\Delta\tau)$ for sufficiently small $\Delta\tau$ and large n.

Therefore, from (I.23)–(I.27) and (I.36)

$$x(m(n,\Delta\tau))=x(n)+\Delta\tau f(x)+q_1(n,m,\bar{x})+q_2(m,n,\bar{x}) \tag{I.37}$$

where

$$|q_1(n,m,\bar{x})|\leq|R_1(n,m)|+q_4(n,m)+q_5(n,m) +2\bar{k}_v\Delta\tau+|f(\bar{x})|\left(\Delta\tau-\sum_{k=n}^{m}\gamma(k)\right) \tag{I.38}$$

and

$$|q_2(n,m,\bar{x})|\leq 2\bar{k}_v\Delta\tau\left[|f(\bar{x})|\cdot\Delta\tau+4p\Delta\tau\cdot\bar{k}_v+|x(n)-\bar{x}|\right] \tag{I.39}$$

for $\Delta\tau<\Delta\tau_0$ and $n>N_0$.

The expression (I.37) with (I.38) and (I.39) is the assertion of the lemma since $q_1(n,m(n,\Delta\tau)\bar{x})$ lends to zero according to Steps 3a–3c and (I.1) [remember that q_5 is defined in (I.35)].

End proof of Lemma 1.

We now proceed with the proof of the main theorem.

Proof of Theorem 1: We consider in the following a sequence ("realization") $\{e(t)\}$ such that assumptions C.3 and C.4 hold and such that condition (20) of Theorem 1 holds.

It follows from the converse stability theorems (see e.g., Krasovskij [39] or Hahn [40]) that the stability assumption (21) of Theorem 1 for $D_c=\{x^*\}$ implies the existence of a function V with properties

1) $V(x)$ is infinitely differentiable;

2) $0\leq V(x)<1$ for $x\in D_A$ and $V(x)=0\Leftrightarrow x=x^*$;

3) $(d/d\tau)V(x(\tau))=V'(x)f(x)\leq 0$ for $x\in D_A$ and equality holds only for $x=x^*$.

[For the case with general invariant set D_c, see Zubov [41] for the proper theorems.]

For convenience in formulations below, we shall let $D_c=\{x^*\}$. An outline of the rest of the proof is as follows:

Step 1: A convergent subsequence $x(n_k)$ tending to $\bar{x}$ is considered. Then $x(n_k)$ is close to $\bar{x}$ infinitely often, and according to Lemma 1, $x(m(n_k,\Delta\tau))$ will approximately be $x(n_k)+\Delta\tau f(\bar{x})$. This means that $V(x(m(n_k,\Delta\tau)))$ is

strictly less than $V(x(n_k))$ if $\bar{x} \neq x^*$. The formal proof is somewhat lengthy and involves several elaborate choices of constants. The result is, however, intuitively clear. The proof of Step 1 extends to (I.45).

Step 2: From the above result it is quite clear that x^* must be a cluster point to $x(n)$, since $V(x(n))$ has a tendency to decrease everywhere in D_A except for $x = x^*$.

Step 3: If there is another cluster point to $\{x(n)\}$ than x^*, say $\hat{x}$, the sequence must move from x^* to $\hat{x}$ infinitely many times. But then $V(x(n))$ is increasing, which contradicts the result of Step 1. Hence, only one cluster point exists and convergence follows.

Step 1: Any cluster point of $\{x(t)\}$ outside D_c yields strictly decreasing values of V.

From assumption (20) of Theorem 1 there exists a subsequence, such that

$$x(\bar{n}_k) \in \bar{D} \text{ and } |\varphi(\bar{n}_k)| < C. \tag{I.40}$$

Since $\bar{D}$ is compact, there exists at least one cluster point to $x(\cdot)$ in $\bar{D}$. Let the cluster point be denoted by $\bar{x}$ and let n_k be a subsequence of $\bar{n}_k$ such that

$$x(n_k) \to \bar{x} \qquad \text{as } k \to \infty. \tag{I.41}$$

Consequently, for arbitrarily small $\epsilon > 0$,

$$|x(n_k) - \bar{x}| < \epsilon \qquad k > K_0(\epsilon). \tag{I.42}$$

Consider now

$$V[x(m(n_k, \Delta\tau))] - V[x(n_k)],$$

where m is defined as in Lemma 1. Denote $n_k = k'$ and $m(n_k, \Delta\tau) = k''$, and use the mean value theorem. This gives

$$\begin{aligned} V[x(k'')] - V[x(k')] &= V'(\xi)[x(k'') - x(k')] \\ &= V'(\bar{x})[x(k'') - x(k')] + [\xi - \bar{x}]^T \\ &\quad \cdot V''(\xi')[x(k'') - x(k')] \end{aligned} \tag{I.43}$$

where ξ and ξ' belong to $\mathcal{B}(\bar{x}, \epsilon + \Delta\tau)$.

Now take $\epsilon < \rho(\bar{x})$, and we can in view of (I.42) apply Lemma 1 to $x(n_k)$, which gives

$$x(k'') - x(k') = \Delta\tau f(\bar{x}) + q_1(k', k'', \bar{x}) + q_2(k', k'', \bar{x})$$

where q_i are subject to (I.5), (I.6).

Insert this into (I.43)

$$V[x(k'')] - V[x(k')] = \Delta\tau V'(\bar{x}) f(\bar{x}) + R_3(\Delta\tau, n_k \bar{x}) \tag{I.44}$$

where

$$\begin{aligned} R_3(\Delta\tau, n_k, \bar{x}) &= (\xi - x)^T V''(\xi')(x(k'') - x(k')) \\ &\quad + V'(\bar{x})\{q_1 + q_2\}. \end{aligned}$$

Now suppose that the cluster point $\bar{x}$ is different from the desired convergence point x^*. Then $V'(\bar{x}) f(\bar{x}) = -\delta$, $\delta > 0$.

Then (I.44) implies that by a proper choice of ϵ, $\Delta\tau$, and K (see [7] for a detailed account), we have

$$V(x(m(n_k, \Delta\tau))) < V(\bar{x}) - \Delta\tau\delta/64 \qquad k > K. \tag{I.45}$$

End Step 1.

This means that if $\bar{x}$ is a cluster point different from x^* the sequence $x(n)$ will i.o. be interior to $\tilde{D} = \{x | x \in \bar{D}_R$ and $V(x) \leq V(\bar{x}) - \Delta\tau\delta/64\}$. The conclusion that $x \in \bar{D}_R$ follows from our assumption that trajectories that start in $\bar{D}_R$ remain in there. This is required in order to apply Lemma 1 to new points in $\tilde{D}$. The set $\tilde{D}$ is compact. Consequently, another cluster point must exist that yields a smaller value of V. Moreover, since $x(t) \in \mathcal{B}(\bar{x}, 2\rho)$ $n_k \leq t \leq m(n_k, \Delta\tau)$ we have from (I.14)

$$|\varphi(k'')| \leq C\tilde{\lambda}^{k''-k'}|\varphi(k')| + C|B| \sum_{k=k'}^{k''} \tilde{\lambda}^{k''-k'}|e(k+1)| \tag{I.46}$$

and hence $|\varphi(m(n_k, \Delta\tau))| < C$ i.o. so the argument can be repeated, again applying Lemma 1 to this new cluster point.

Step 2: Suppose (I.45) holds for any subsequence $\{x(n_k)\}$ that converges to a point different from x^*. Then

$$\liminf_{t \to \infty} V(x(t)) = 0. \tag{I.47}$$

Consider $\inf V(x)$ taken over all cluster points in D_A. Let this value be U. Since the set of cluster points in D_A is compact, there exists a cluster point $\hat{x}$, such that $V(\hat{x}) = U$. If now $U > 0$, $V'(\hat{x}) f(\hat{x})$ will be strictly negative ($= -\hat{\delta}$) and from (I.45) $V(x(k))$ takes a value less than $U - \hat{\delta}\Delta\tau/64$ i.o., which contradicts U being the infinum. Hence, $U = 0$, which means that x^* is a cluster point. End Step 2. □

Step 3: From (I.45) and (I.47) it follows that

$$\limsup_{t \to \infty} V(x(t)) = 0. \tag{I.48}$$

Let $\rho^* = \rho(x^*)$ be the region for which (2) is exponentially stable for $x(k) \in \mathcal{B}(x^*, \rho^*)$, as in (I.11), (I.12).

Suppose that

$$\limsup_{n \to \infty} V(x(n)) = \overline{W} > 0. \tag{I.49}$$

Take $W < \overline{W}$ such that

$$\{x | V(x) \leq W\} \subset B(x^*, \rho^*)$$

and consider the interval $I = [W/3, 2W/3]$. (We may of course choose any subinterval; there is no particular reason to divide it up in thirds.)

Since x^* is a cluster point and since $V(x(n))$ is supposed to have a subsequence tending to W, this interval I is crossed "upwards" and "downwards" infinitely many times by $V(x(n))$.

We shall now proceed to show that (I.49) would imply that there must be a subsequence of $V(x(n))$ that belongs to I, by proving that in $\mathcal{B}(x^*, \rho^*)$ the "step size" $x(n+1) - x(n)$ tends to zero.

First, let $x(\tilde{n}_k)$ be a subsequence tending to x^*, such that $|\varphi(\tilde{n}_k)| < C$. (The existence of such a sequence follows from Step 2 and the stability argument (I.46), using the fact that $\rho = \rho(\bar{x})$ is bounded from below by a positive constant in $\bar{D}$.)

For $t > \tilde{n}_k$, but such that $x(t)$ "remains" in $B(x^*, \rho^*)$ we have, (I.46)

$$|\varphi(t)| \leq C\tilde{\lambda}^{t-\tilde{n}_k}|\varphi(\tilde{n}_k)| + C|B| \sum_{k=n_k}^{t} \tilde{\lambda}^{t-k}|e(k+1)| \leq v(t,\tilde{\lambda},c)$$

with $v(t,\tilde{\lambda},c)$ defined by (16).

Hence,

$$\begin{aligned}|\gamma(t)Q(t,x(t),\varphi(t))| &\leq \gamma(t)|Q(t,x^*,0)| \\ &\quad + \gamma(t)\mathcal{H}_1(x^*,0,\rho^*,v(t,\tilde{\lambda},c))(|x(t)-x^*|+|\varphi(t)|) \\ &\leq \gamma(t)|Q(t,x^*,0)| + \gamma(t)\mathcal{H}_1(x^*,0,\rho^*,v(t,\tilde{\lambda},c)) \\ &\quad \cdot(\rho^* + v(t,\tilde{\lambda},c)) \qquad \text{(I.50)}\end{aligned}$$

where the first inequality follows from assumption C.1.

It follows from Step 3a of the proof of Lemma 1 that the first term of the RHS of (I.50) tends to zero and from Step 3c that the second one tends to zero. Consequently, inside $B(x^*,\rho^*)$ the step size tends to zero, and hence there will be a subsequence of $V(x(n))$ entirely in the interval I. Consider now a special, convergent sequence of "upcrossings," that is a subsequence of this.

Let the sequence n_k' be defined such that

$$V(x(n_k'-1)) < W/3 \qquad \text{(I.51)}$$

$$V(x(n_k')) \geq W/3 \qquad \text{(I.52)}$$

$$V(x(n_k'+s_k)) > 2W/3 \qquad \text{(I.53a)}$$

where

$$s_k \text{ is the first } s \text{ for which } V(x(n_k+s)) \notin I \qquad \text{(I.53b)}$$

$$x(n_k') \to \tilde{x} \qquad \text{as } k \to \infty. \qquad \text{(I.54)}$$

From (I.51) and (I.52) it is clear that $V(\tilde{x}) = W/3$ and let $V'(\tilde{x})f(\tilde{x}) = -\tilde{\delta}$. From (I.45) we now have that

$$V\big(x(m(n_k',\Delta\tau))\big) < W/3 - \tilde{\delta}\Delta\tau/64. \qquad \text{(I.55)}$$

This means that $V(x(n_k'+s_k')) \notin I$ where $s_k' = m(n_k',\Delta\tau) - n_k'$.

Suppose now that there is a $s_k < s_k'$ such that $V(x(n_k'+s_k)) > 2W/3$. However, V is continuous and $x(n_k'+s)$, $(s < m(n_k',\Delta\tau) - n_k')$ will belong to an arbitrarily small neighborhood of $x(n_k')$ for sufficiently small $\Delta\tau$ according to Lemma 1. Therefore (I.55) implies a contradiction to the existence of a subsequence n_k' with properties (I.51)–(I.54). Hence, no interval I may exist, W must be zero and (I.48) follows. End Step 3.

Equation (I.48) implies according to the properties of $V(\cdot)$ that $x_n \to x^*$. This concludes the proof of Theorem 1.

End proof of Theorem 1.

We shall now proceed to prove that assumptions A or B imply assumptions C.3 and C.4. This will be done only for the case $\gamma(t) = 1/t$, which anyway should be the most common one. We may also remark that since the correction term $Q(\cdot,\cdot,\cdot)$ may be time-varying, this case also includes cases where

$$\gamma(t) = \frac{c(t)}{t}; \; c(t) \to c \qquad \text{as } t \to \infty \qquad \text{(I.56)}$$

be redefining the factors of the product

$$\gamma(t)Q(t,x,\varphi).$$

Lemma 2. Suppose the algorithm (1), (2) is subject to assumptions A or B. Assume that $\gamma(t) = 1/t$. Then C.3 holds w.p.1 as well as Steps 3b and 3c of the proof of Lemma 1. If, in addition

$$E\,\mathcal{H}_1(\bar{x},\bar{\varphi}(t,\bar{x}),\rho(\bar{x}),v(t,\lambda,c))(1+v(t,\lambda,c))$$

converges at $t \to \infty$, then also C.4 holds w.p.1.

Proof: For $\gamma(t) = 1/t$ we have from (18a)

$$z(t,\bar{x}) = \frac{1}{t}\sum_{k=1}^{t} Q(k;\bar{x},\bar{\varphi}(k,\bar{x})). \qquad \text{(I.57)}$$

We shall apply the following ergodicity result of Cramér and Leadbetter [38, pp. 94–96] to (I.57).

Let $f(k)$ be a sequence of random variables with zero means and covariances

$$|Ef(k)f(s)| \leq \frac{k^p + s^p}{1+|k-s|^q}; \qquad 0 \leq 2p < q < 1. \qquad \text{(I.58)}$$

Then

$$\frac{1}{t}\sum_{k=1}^{t} f(k) \to 0 \qquad \text{w.p.1 as } t \to \infty \qquad \blacksquare$$

(The proof in [38] is given for the continuous time case, but it goes through without changes for the discrete time case.)

In order to calculate the covariances, introduce for $k > s$ the variable

$$\bar{\varphi}_s^0(k,\bar{x})$$

by

$$\bar{\varphi}_s^0(k+1,\bar{x}) = A(\bar{x})\bar{\varphi}_s^0(k,\bar{x}) + B(\bar{x})e(k+1); \quad \bar{\varphi}_s^0(s,\bar{x}) = 0.$$

This variable is independent of $\bar{\varphi}(s,\bar{x})$ according to assumption A.1 = B.1, and

$$|\bar{\varphi}(k,\bar{x}) - \varphi_s^0(k,\bar{x})| \leq \bar{c}\cdot\lambda(\bar{x})^{k-s}|\bar{\varphi}(s,\bar{x})| \cdot\big(\leq v(k,\lambda(\bar{x}),\bar{c})\big).$$

Hence, using assumption A.3 or B.3

$$\begin{aligned}|Q(k,\bar{x},\bar{\varphi}(k,\bar{x})) - Q(k,\bar{x},\bar{\varphi}_s^0(k,\bar{x}))| \\ \leq \mathcal{H}_1(\bar{x},\bar{\varphi}(k,\bar{x}),0,v(k,\lambda(\bar{x}),\bar{c}))\bar{c}\lambda(\bar{x})^{k-s}|\bar{\varphi}(s,\bar{x})|.\end{aligned}$$

Moreover, from A.2, A.3 (according to which $\mathcal{K}_1$ and $\bar{\varphi}(s,\bar{x})$ are bounded) or B.7,

$$E|Q(k,\bar{x},\bar{\varphi}(k,\bar{x}))-Q(k,\bar{x},\bar{\varphi}_s^0(k,\bar{x}))|^2 \leqslant c\cdot\lambda(\bar{x})^{k-s}. \tag{I.59a}$$

Now

$$\begin{aligned}&|\operatorname{cov}\{Q(k,\bar{x},\bar{\varphi}(k,\bar{x})),Q(s,\bar{x},\bar{\varphi}(s,\bar{x}))\}|\\&\quad=|\operatorname{cov}\{Q(k,\bar{x},\bar{\varphi}_s^0(k,\bar{x}))+[Q(k,\bar{x},\bar{\varphi}(k,\bar{x}))\\&\qquad-Q(k,\bar{x},\bar{\varphi}_s^0(k,\bar{x}))],Q(s,\bar{x},\bar{\varphi}(s,x))\}|\\&\quad=|0+\operatorname{cov}\{[Q(k,\bar{x},\bar{\varphi}(k,\bar{x}))-Q(k,\bar{x},\bar{\varphi}_s^0(k,\bar{x}))],\\&\qquad Q(s,\bar{x},\bar{\varphi}(s,x))\}|\leqslant c\cdot\lambda(\bar{x})^{k-s},\end{aligned} \tag{I.59b}$$

where the second equality follows since $\bar{\varphi}_s^0(k,\bar{x})$ and $\bar{\varphi}(s,\bar{x})$ are independent, and the last inequality follows from Schwarz' inequality, (I.59a), and assumption B.7.

Hence, (I.58) holds and

$$\frac{1}{t}\sum_{k=1}^{t}[Q(k,\bar{x},\bar{\varphi}(k,\bar{x}))-EQ(k,\bar{x},\bar{\varphi}(k,\bar{x}))]\to 0 \quad \text{w.p.1 as } t\to\infty$$

which according to assumption A.5 = B.6 implies that

$$\frac{1}{t}\sum_{k=1}^{t}Q(k,\bar{x},\bar{\varphi}(k,\bar{x}))\to f(\bar{x}) \quad \text{w.p.1 as } t\to\infty. \tag{I.60}$$

Notice that (I.60) holds for any given $\bar{x}$ w.p.1. For a given realization ω outside a null set it does not immediately follow that it holds for all $x\in D_R$. To conclude that, we first note that (I.60) will hold w.p.1 simultaneously over a dense denumerable, subset of D_R. (The union of a denumerable number of null sets is a nullset.) Hence, for a given realization outside a null set (I.60) holds for all $\bar{x}$ in a dense subset of D_R. But since $Q(\cdot,\cdot,\cdot)$ and $f(\cdot)$ are continuous in x, we may extend this set to D_R itself.

The proof of the second part of Lemma 2 is analogous. (Under assumptions A it is trivial.)

End proof of Lemma 2.

The proof for general sequences $\{\gamma(n)\}$, subject to A.8, A.9 is achieved by first proving that

$$E\left|\sum_{k=n}^{m(n,\Delta\tau)}\gamma(k)[Q(k,\bar{x},\bar{\varphi}(k,\bar{x}))-EQ(k,\bar{x},\bar{\varphi}(k,\bar{x}))]\right|^{2p}<C\cdot\gamma(n)^p \tag{I.61}$$

and then proving convergence by the Borel–Cantelli Lemma and assumption A.7.

We shall need the estimate (I.61) in Appendix III, and let us therefore prove (I.61) for $p=1$. The case with general p is proven analogously, though with more technical labor. The details are given in [7]. Denote for short

$$Q_k=Q(k,\bar{x},\bar{\varphi}(k,\bar{x}))$$

and assume that Q_k is a scalar and that $EQ_k=0$. Then

$$\begin{aligned}E\left(\sum_{k=n}^{m}\gamma(k)Q_k\right)^2&=\sum_{k=n}^{m}\sum_{s=n}^{m}\gamma(k)\gamma(s)EQ_kQ_s\\&\leqslant(n)^2C\sum_{k=n}^{m}\sum_{s=n}^{m}\lambda^{|k-s|}\leqslant C\gamma(n)^2(m-n)\end{aligned} \tag{I.62}$$

where the first inequality follows from assumption A.8 and (I.59b). It remains now only to prove that

$$(m(n,\Delta\tau)-n)\leqslant c/\gamma(n). \tag{I.63}$$

Since $\gamma(\cdot)$ is decreasing

$$2\Delta\tau\geqslant\sum_{n}^{m}\gamma(k)\geqslant(m-n)\gamma(m)$$

or

$$(m-n)\leqslant c/\gamma(m). \tag{I.64}$$

Moreover, from assumption A.9

$$\frac{1}{\gamma(n+1)}-\frac{1}{\gamma(n)}<c$$

we have

$$\gamma(n+1)>\gamma(n)(1-c\gamma(n+1))$$

or, upon repetition

$$\gamma(m)>\gamma(n)\prod_{n+1}^{m}(1-c\gamma(j))\sim\gamma(n)\exp\left(-c\cdot\sum_{n+1}^{m}\gamma(j)\right)\sim c'\gamma(n).$$

This, together with (I.64) implies (I.63) and the proof of (I.61) is complete for $p=1$.

Proof of Corollary to Theorem 1: The function $V(x)$ with the property (22), whose existence is assumed in the corollary plays the role of the Lyapunov function throughout the proof of Theorem 1. Since, in the formulation of the corollary there is no guarantee that $x(m(n,\Delta\tau))$ will belong to D_R even if $x(n)$ does, the possibility of a cluster point on the boundary of D_R cannot be excluded. Moreover, if $\bar{D}$ is unbounded the existence of a cluster point in $\bar{D}$ is not guaranteed. But if no cluster point exists, then the sequence $\{x(n)\}$ will tend to infinity, which in that case will be a boundary point of D_R.

End proof of corollary.

APPENDIX II
PROOF OF THEOREM 2

To prove (25a) we first assume that $f(x^*) \neq 0$.

We first note that according to Lemma 2, assumptions A or B imply that (I.4)–(I.6) hold w.p.1.

Let $\Delta\tau(x^*) = \Delta\tau^*$ be "the sufficiently small" $\Delta\tau$ as defined in Lemma 1.

Take

$$\rho^* < \Delta\tau^* |f(x^*)|/4 \tag{II.1}$$

and let $\Omega^* = \{\omega | x(t) \to \mathcal{B}(x^*, \rho^*)\}$ with $P(\Omega^*) = P^* > 0$.

If $x(t)$ converges to $\mathcal{B}(x^*, \rho^*)$, it is in particular inside $\mathcal{B}(x^*, 2\rho^*)$ infinitely often. Therefore, there is a cluster point $\bar{x}$ inside $\mathcal{B}(x^*, 2\rho^*)$. Also, inside $\mathcal{B}(x^*, 2\rho^*)$ the observation equation (2) will be stable for sufficiently small ρ^*, according to Step 2 of the proof of Lemma 1.

Hence, we can select a subsequence $\{x(n_k)\}$ tending to $\bar{x}$, such that $|\varphi(n_k)| < C$, i.e., (I.3) holds. Then Lemma 1 implies that

$$x(m(n_k, \Delta\tau^*)) \approx x(n_k) + f(x^*) \cdot \Delta\tau^*$$

asymptotically as $k \to \infty$ for $\omega \in \Omega^*$ except on a null set. This means according to (II.1) that $x(m(n_k, \Delta\tau^*))$ is outside $\mathcal{B}(x^*, 2\rho^*)$, and the assumed convergence is contradicted.

To illustrate the basic idea of the proof of (25b), consider the special case

$$Q(t, x, \varphi(t+1)) = Ax + e(t+1)$$

where A is an $n|n$ matrix and $e(\cdot)$ is a sequence of independent random variables with zero mean values. Suppose that A has an eigenvalue λ with $\operatorname{Re}\lambda > 0$, and let L be a corresponding left eigenvector. Let $\tau(n) = Lx(n)$ and $\epsilon(n) = Le(n)$. The condition on $\operatorname{cov} Q$ implies that $\{\epsilon(n)\}$ is not identically zero. Then the algorithm (1) can be written

$$\tau(n+1) = \tau(n) + \gamma(n+1)\left[\lambda\tau(n) + \epsilon(n+1)\right]$$

and

$$\tau(m) = \Gamma(n, m) \cdot \left\{ \tau(n) + \sum_{n+1}^{m} \tilde{\beta}_k^m \epsilon(k) \right\}$$

where

$$\Gamma(n, m) = \prod_{n+1}^{m} (1 + \lambda\gamma(k)) \sim \exp\left\{ \lambda \sum_{n+1}^{m} \gamma(k) \right\} \tag{II.2}$$

and

$$\tilde{\beta}_k^m = \gamma(k) \prod_{n+1}^{k} (1 + \lambda\gamma(j))^{-1}.$$

Since $\tau(n)$ and the sum of random variables are independent and $\Gamma(n, m)$ tends to infinity as m increases, it follows that $\tau(m)$ will, with probability one, not tend to zero as m tends to infinity. Hence, $x(m)$ will not converge to 0 $(= x^*)$ with nonzero probability.

The general case is proven by linearization around x^* and the additional terms are taken care of by appropriate approximations. Like in the proof of Lemma 1, this leads to several technicalities, but the basic idea remains the same as above.

To keep the Appendices within reasonable size, these technicalities will not be given here. The full proof can be found in [7]. Proofs for related algorithms are given in [6] and [1].

APPENDIX III
PROOF OF THEOREM 3

The idea of this proof is to apply Lemma 1 to obtain local estimates of how much $x(t)$ differs from the corresponding trajectory, and then linking such estimates together making use of the stability property.

Heuristic Outline of How the Estimates are linked Together

The idea of how the local estimates are extended to global ones can be geometrically expressed as follows, cf. Fig. 5.

Assume that the estimate at time τ_k is in the interval A. The trajectories that start in A belong at time $\tau_k + \Delta\tau_k$ to the interval B which is smaller than A since the trajectory is stable. Now, the estimates obtained by the algorithm differ from the trajectories with a small quantity according to Lemma 1. Denote this distance by C. During the time interval $\Delta\tau_k$, the estimates have not diverged from the nominal trajectory if $A \leqslant B + 2C$.

To achieve this, A and $\Delta\tau_k$ must be chosen with care. The interval $\Delta\tau_k$ must be large enough to let the trajectories converge sufficiently, and small enough to limit second-order effects and the noise influence.

The formal proof will be developed in 5 steps. In Step 1 the details of the application of Lemma 1 are given. Step 2 deals with the implications of the stability assumption. In Step 3 an interval for $\Delta\tau_k$ is selected, corresponding to $\Delta\tau_k$ not being too small nor too large as described in the heuristic outline. In Step 4 it is shown that the estimates stay with the given ϵ-region as depicted in Fig. 5 if certain stochastic variables are less than a given value. Step 5 calculates the probability that they are less than this value.

Step 1—Application of Lemma 1:

Order the set of indices $I = \{n_i\}$ such that

$$n_1 < n_2 < \cdots < n_k < n_{k+1} < \cdots.$$

Denote $\Delta\tau_k = \tau_{n_{k+1}} - \tau_{n_k}$. Then by taking $\bar{x} = x(n_k)$ in Lemma 1 we obtain

$$x(n_{k+1}) = x(n_k) + \Delta\tau_k f\left[x(n_k)\right] + q_1(n_k, n_{k+1}, \bar{x}) + q_2(n_k, n_{k+1}, \bar{x}) \tag{III.1}$$

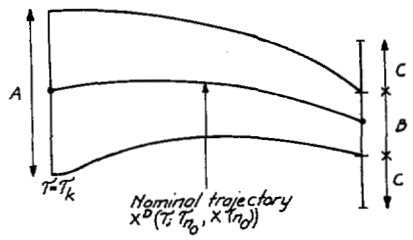

Fig. 5. To illustrate the idea of the proof of Theorem 3.

where $q_2(n_k,n_{k+1},\bar{x})$ is given by (I.6) and (I.39). Since $x_{n_k}=x$,

$$|q_2(n,\Delta\tau_k;\bar{x})|\leq C_2\Delta\tau_k^2. \tag{III.2}$$

Moreover $q_1(n_k,n_{k+1},\bar{x})$ is given by (I.38).

From Lemma 1 we also know that (III.1), (III.2) are valid only if $\Delta\tau_k$ is sufficiently small and n_k so large that

$$|q_1(n_k,n_{k+1},x(n_k))|<\rho/2. \tag{III.3}$$

We first note that the constant C_2 in (III.2) can be taken to be globally valid in $\bar{D}$, according to the expression (I.39) and since $|f(x)|$ and k_v are bounded in $\bar{D}$. Moreover, there is a common lower bound to $\Delta\tau_0(x)$ for $x\in\bar{D}$, which can be realized as follows.

The radius 2ρ, which is so small that (I.12) implies (I.11) depends on $\bar{x}$ and is a measure of how fast $A(x)$ changes in a neighborhood of $\bar{x}$. Since $A(x)$ is Lipschitz continuous in $\bar{D}$ and $\bar{D}$ is compact, the radius $\rho(\bar{x})$ will have a positive lower bound as $\bar{x}$ varies over $\bar{D}$. Denote this by $\bar{\rho}$. Let [cf., (I.34)]

$$\bar{\Delta\tau}_0=\inf_{x\in D}\left(\frac{\bar{\rho}}{|f(\bar{x})|},\frac{1}{16\bar{k}_v}\right) \tag{III.4}$$

which is strictly positive, since $|f(\bar{x})|$ and $\bar{k}_v$ are bounded in $\bar{D}$.

Then (III.1), (III.2) is valid for $\Delta\tau_k<\bar{\Delta\tau}_0$ if (III.3) holds.

By solving the differential equation (19) from τ to $\tau+\Delta\tau$ with $\bar{x}$ as the initial condition we have

$$|x^D(\tau+\Delta\tau;\tau,\bar{x})-(\bar{x}+\Delta\tau\cdot f(\bar{x}))|\leq L\Delta\tau^2 \tag{III.5}$$

where L can be taken globally in $\bar{D}$. Introduce the following abbreviated notation

$$x_i^D(j)=x^D(\tau_{n_j};\tau_{n_i},x(n_i)). \tag{III.6}$$

Equations (III.1), (III.2), and (III.5) now imply in the notation of (III.6)

$$|x(n_{k+1})-x_k^D(k+1)|\leq(c_2+L)\Delta\tau^2+q_1(n_{k+1},n_k,x(n_k)) \tag{III.7}$$

if (III.3) holds. Introduce

$$M=C_2+L. \tag{III.8}$$

The estimate (III.7) with the expression (I.38) for q_1 is the basis for the rest of the proof.

End Step 1.

Step 2—Stability of the Trajectories

According to the assumptions of the theorem there exists a function $V(\Delta x,\tau)$ that is quadratic in Δx and such that

$$\frac{d}{d\tau}V(\Delta x,\tau)<-C|\Delta x|^2 \qquad C>0 \tag{III.9}$$

along solutions of the variational equation of (19). Since we assume $V(\Delta x,\tau)$ to be quadratic in Δx it is no additional loss of generality to assume that

$$V(\Delta x,\tau)=|\Delta x|^2$$

since it is always possible to make a (time-dependent) change of metric.

Then (III.9) implies that

$$|x^D(\tau+\Delta\tau;\tau,x+\Delta x)-x^D(\tau+\Delta\tau;\tau,x)|\leq(1-\lambda\Delta\tau)|\Delta x| \tag{III.10}$$

for some $1>\lambda>0$.

End Step 2.

Step 3—Selection of an Interval for $\Delta\tau_k$

To obtain the upper and lower bounds for $\Delta\tau_k$ discussed under the heuristic outline, let

$$\epsilon_0=\min\left(\frac{4\delta_0 M}{\lambda},\frac{\bar{\Delta\tau}_0 M4}{3\lambda},\frac{4}{\lambda}\sqrt{\rho M}\right). \tag{III.11}$$

Choose $\epsilon<\epsilon_0$. Then automatically

$$\lambda\epsilon/8M<\Delta\tau_k \qquad \text{all } k \tag{III.12}$$

since $\Delta\tau_k>\delta_0$ according to the assumptions of the theorem. An upper bound for $\Delta\tau_k$ can be obtained by possibly extending the set I, cf., beginning of Step 1. Do this so that

$$\frac{\lambda\epsilon}{4M}<\Delta\tau_k<\frac{3\lambda\epsilon}{4M}. \tag{III.13}$$

The resulting set I may depend on ϵ. According to (III.11), automatically $\Delta\tau_k<\bar{\Delta\tau}_0$.

End Step 3.

Step 4—The Estimates Remain in an ϵ-Neighborhood of the Corresponding Trajectory if $q_1(n_{k+1},n_k,x(n_k))$ is Sufficiently Small, all $n_k\in I$

Assume that

$$|q_1(n_{k+1},n_k,x(n_k))|<r(\epsilon) \tag{III.14a}$$

where

$$r(\epsilon)=\frac{\lambda^2\epsilon^2\cdot 3}{M\cdot 16}. \tag{III.14b}$$

According to (III.11),

$$r(\epsilon)<\delta/2. \tag{III.14c}$$

Assume also that

$$|x_0^D(k)-x(n_k)|\leq\epsilon. \tag{III.15}$$

Since $\Delta\tau_k$ now is assumed to be less than $\overline{\Delta\tau}_0$ and since (III.14) is assumed to hold, Lemma 1 can be applied, giving according to (III.7) and (III.14) [notation as in (III.6)]

$$|x(n_{k+1}) - x_k^D(k+1)| \leq M\cdot\Delta\tau_k^2 + r(\epsilon). \quad \text{(III.16)}$$

Now

$$\begin{aligned}|x_0^D(k+1) - x(n_{k+1})| &\leq |x_0^D(k+1) - x_k^D(k+1)| \\ &\quad + |x_k^D(k+1) - x(n_{k+1})| \\ &\leq (1-\lambda\Delta\tau_k)|x_0^D(k) - x(k)| \\ &\quad + M\cdot\Delta\tau_k^2 + r(\epsilon) \\ &\leq (1-\lambda\Delta\tau_k)\epsilon + M\cdot\Delta\tau_k^2 + \frac{\lambda^2\epsilon^2\cdot 3}{M\cdot 16} \\ &= \epsilon + M\left[\left(\Delta\tau_k - \frac{\lambda\epsilon}{4M}\right)\left(\Delta\tau_k - \frac{3\lambda\epsilon}{4M}\right)\right] \leq \epsilon \quad \text{(III.17)}\end{aligned}$$

where the second inequality follows from (III.10) and (III.16), the third from (III.15) and (III.14b), and the last one from (III.13).

The conclusion is that if $x(n_k)$ lies in an ϵ-neighborhood of the trajectory, so will $x(n_{k+1})$ if (III.14) holds. If (III.14) holds for all $n_k \in I$, then

$$\sup_{n_k \in I} |x(n_k) - x_0^D(k)| \leq \epsilon. \quad \text{(III.18)}$$

End Step 4.

It remains now only to estimate the probability that (III.14) holds for all $n_k \in I$.

Step 5—The Probability That (III.14) Holds

The expression for $q_1(n_{k+1}, n_k, x(n_k))$ may according to (I.38) and (I.22) be decomposed as follows

$$\begin{aligned}|q_1(m,n,\bar{x})| &\leq \left| \sum_{k=n}^{m} \gamma(k) \right. \\ &\quad \left. \cdot \left[Q\left(k, \bar{x}, \bar{\varphi}(k,\bar{x})\right) - EQ\left(k,\bar{x},\bar{\varphi}(k,\bar{x})\right)\right] \right| \\ &\quad + \sum_{k=n}^{m} \gamma(k)\mathcal{K}_1\left(\bar{x}, \bar{\varphi}(k,\bar{x}), 2\bar{\rho}, v(k,\lambda,c)\right) \\ &\quad \cdot \lambda^{k-n} v(n,\lambda,c) \\ &\quad + \left| \Delta\tau f(\bar{x}) - \sum_{k=n}^{m} \gamma(k) EQ\left(k,\bar{x},\bar{\varphi}(k,\bar{x})\right)\right|. \quad \text{(III.19)}\end{aligned}$$

The last term in this expression is deterministic and tends to zero according to assumption A.5 = B.6 and according to (I.1) ($m = m(n,\Delta\tau)$). Let N_1 be such that this last term is less than $r(\epsilon)/2$ for $n_k > N_1$.

For the first term of (III.19) the estimate of its $2p$-absolute moment, (I.61) holds. The second term of (III.19) has a mean that is bounded by

$$C\cdot\sum_{n}^{m} \gamma(k)\lambda^{k-n} \leq \frac{c}{1-\lambda}\gamma(n)$$

according to assumptions B.7 (or A.3) and B.10. Analogously to the first term of (III.19) its $2p$-absolute moment around the mean is bounded by $C\cdot\gamma(n)^p$. (Much better estimates are possible, but uninteresting.) Denote the two first terms of (III.19) by $q_1^A(m,n,\bar{x})$ and the last one by $q_1^B(m,n,\bar{x})$. Then

$$q_1(m,n,\bar{x}) = q_1^A(m,n,\bar{x}) + q_1^B(m,n,\bar{x}) \quad \text{(III.20)}$$

where

$$E|q_1^A(n_{k+1}, n_k, x(n_k))|^{2p} \leq C\cdot\gamma(n_k)^p \quad \text{(III.21)}$$

and

$$|q_1^B(n_{k+1}, n_k, x(n_k))| < r(\epsilon)/2, \qquad n_k > N_1. \quad \text{(III.22)}$$

Chebyshev's inequality gives, using (III.21)

$$\begin{aligned}P\left(|q_1^A(n_{k+1}, n_k, x(n_k))| > r(\epsilon)/2\right) \\ \leq \left(\frac{2}{r(\epsilon)}\right)^{2p}\cdot C\cdot\gamma(n_k)^p. \quad \text{(III.23)}\end{aligned}$$

The probability that (III.14) does not hold for some $n_k > N_1$, $n_k \in I$ is thus bounded by

$$\begin{aligned}\left(\frac{2}{r(\epsilon)}\right)^{2p}\cdot C\cdot\sum_{n_k=N_1}^{N}\gamma(n_k)^p \\ \leq \left(\frac{2}{r(\epsilon)}\right)^{2p}\cdot C\cdot\sum_{j=N_1}^{N}\gamma(j)^p. \quad \text{(III.24)}\end{aligned}$$

Combining the conclusion of Step 4 with this result gives the assertion of the theorem. The number T_0 mentioned in the theorem equals N_1 defined by (III.22).

Remark: The inequality in (III.24) is somewhat wasteful. For example, if $\gamma(t) = 1/t$, then with

$$\Delta\tau_k = \sum_{t=n_k}^{n_{k+1}} \frac{1}{t}$$

implies that $n_{k+1} \approx n_k(1+\Delta\tau_k)$ for small $\Delta\tau_k$ and taking $p=1$ in (III.24) an upper bound for the LHS of (III.24) is given by

$$\begin{aligned}\left(\frac{2}{r(\epsilon)}\right)^2\cdot C\cdot\sum_{n_k=N_1}^{N}\frac{1}{N_1\left(1+\overline{\Delta\tau}_0\right)^k} \\ \leq \frac{1}{N_1}\left(\frac{2}{r(\epsilon)}\right)^2\cdot C\cdot\frac{4M}{\lambda\epsilon} \leq \bar{c}/\epsilon^5. \quad \text{(III.25)}\end{aligned}$$

End proof of Theorem 3.

Appendix IV
Proof of Theorem 4

In virtue of the projection we know that $x(t)$ belongs to a compact area i.o. that is part of D_R. We could therefore apply Theorem 1 directly, apart from the fact that the

projection algorithm (28), (29) differs from the algorithm (1), (2) treated in Theorem 1.

It therefore suffices to show that the "projection" takes place at most a finite number of times w.p.1. After the last time $x(t)$ is forced into D_2 the projection algorithm coincides with the basic algorithm (1), (2) and the proof of Theorem 1 is valid.

If indeed, the estimate $x(t)$ were outside D_1 infinitely often, then it would have to pass from D_2 to outside D_1 i.o., i.e., to a higher value of $U(x(t))$ (see (31)) in spite of the force trying to decrease U according to (30). In Step 3 of the proof of Theorem 1 it was proved that this is impossible, and hence the projection facility in (28) is used only a finite number of times. Also, the estimates cannot remain in $\tilde{D}$ from a certain time on, since condition (30) shows that (using Step 2 of the proof of Theorem 1) they will be forced into D_2.

Appendix V
On Some Extensions

Continuous Time Algorithms: It should be clear that the theorems and proofs given here are valid also for continuous time algorithms, with proper and straightforward modifications.

The Case when Assumption A.5 = B.6 does not hold: The basic assumption, defining the RHS of the differential equation is A.5 = B.6.

Let $EQ(t,\bar{x},\bar{\varphi}(t,\bar{x}))$ be denoted by $f(\bar{x},t)$. In the proof it is only used that

$$\sum_{k=n}^{m(n,\Delta\tau)} \gamma(k) f(\bar{x},k) \to \Delta\tau f(\bar{x}) \qquad \text{as } n\to\infty \qquad \text{(V.1)}$$

which, for the case $\gamma(k)=1/k$ is equivalent to

$$\frac{1}{t}\sum_{k=1}^{t} f(\bar{x},k) \to f(\bar{x}) \qquad \text{as } t\to\infty. \qquad \text{(V.2)}$$

Notice that (V.1) or (V.2) very well may hold even if A.5 = B.6 does not hold. It is consequently unnecessarily restrictive to assume convergence of $f(\bar{x},t)$ as $t\to\infty$. Moreover, if (V.1) does not hold, it is sufficient to require the existence of a twice differentiable function $V(s)$ [which plays the role of Lyapunov function for the in this case undefined differential equation] such that

$$V'(\bar{x})\left[\sum_{k=n}^{m(n,\Delta\tau)} \gamma(k) f(\bar{x},k)\right] < -\delta(\bar{x})\cdot\Delta\tau \qquad \text{(V.3)}$$

for all sufficiently large n, where $\delta(\bar{x})>0$ for $\bar{x}\notin D_c$.

More Complex Generation of the Observations: In some cases, like for the extended Kalman filter, which is presently being analyzed using the methods of this paper, a more complex mechanism replaces (2):

$$\varphi(t) = A(z(t-1))\varphi(t-1) + B(z(t-1))e(t) \qquad \text{(V.4a)}$$

$$z(t) = h(z(t-1),x(t)). \qquad \text{(V.4b)}$$

Equation (V.4) if of course just a special case of the general, nonlinear dynamics (3). To treat this case let $\bar{z}(\bar{x})$ denote the limit of the recursion

$$z(t,\bar{x}) = h(z(t-1,\bar{x}),\bar{x}), \qquad t\to\infty \qquad \text{(V.5)}$$

and define $\bar{\varphi}(t,\bar{x})$ through

$$\bar{\varphi}(t,\bar{x}) = A(\bar{z}(\bar{x}))\bar{\varphi}(t-1,\bar{x}) + B(\bar{z}(\bar{x}))e(t). \qquad \text{(V.6)}$$

The theorems will then hold in the same formulations as before, if the following three conditions on (V.4b) hold:

$$|z(t)-z(t,\bar{x})| \le C\cdot \max_{r\le k\le t} |\bar{x}-x(k)| \quad \text{if } z(r)=z(r,\bar{x}) \qquad \text{(V.7)}$$

$$|\bar{z}(\bar{x})-z(t,\bar{x})| \le c\mu^t\cdot|\bar{z}(\bar{x})-z(0,\bar{x})|; \quad \mu<1 \qquad \text{(V.8)}$$

and

$$|z(t)|+|\varphi(t)|<c \text{ i.o.} \qquad \text{(V.9)}$$

To prove this we note that assumption A.4, (V.7) and (V.8) imply that

$$|A(z(j))-A(\bar{z})| \le c \max_{n\le k\le j} |x(k)-\bar{x}| + c\mu^{j-n}|\bar{z}(\bar{x})-z(n)|.$$

If we use this estimate in (I.16) together with (V.9) we arrive at the same expression as in (I.17) and the proof continues without further changes.

Acknowledgment

I would like to express my sincere gratitude to Profs. K. J. Åström, H. J. Kushner and Ya. Z. Tsypkin for important discussions on the subject of this paper. I have also benefitted from the referees' competent suggestions on how to improve the readability of the Appendices.

References

[1] L. Ljung, T. Söderström, and I. Gustavsson, "Counterexamples to general convergence of a commonly used recursive identification method," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 643–652, Oct. 1975.

[2] K. J. Åström, U. Borisson, L. Ljung, and B. Wittenmark, "Theory and applications of adaptive regulators based on recursive identification," in *Proc. 6th IFAC Congress*, Boston, MA, 1975. An extended version of this paper will appear in *Automatica*, Sept. 1977.

[3] L. Ljung and S. Lindahl, "Convergence properties of a method for state estimation in power systems," *Int. J. Contr.*, vol. 22, pp. 113–118, July 1975.

[4] L. Ljung and B. Wittenmark, "Analysis of a class of adaptive regulators," in *Proc. IFAC Symp. on Stochastic Control*, Budapest, Sept. 1974.

[5] L. Ljung, "Convergence of recursive, stochastic algorithms," in *Proc. IFAC Symp. on Stochastic Control*, Budapest, Sept. 1974.

[6] —— "Strong convergence of a stochastic approximation algorithm," *Annals of Statistics*, to be published; also available as Rep. LiTH-ISY-I-0126, Linköping University, Linköping, Sweden.

[7] —— "Theorems for the asymptotic analysis of recursive, stochastic algorithms," Dep. Automat. Contr., Lund Institute of Technology, Lund, Sweden, Rep. 7522.

[8] —— "Convergence of recursive stochastic algorithms," Division of Automatic Control, Lund Institute of Technology, Lund, Sweden, Rep. 7403, Feb. 1974.
[9] W. Hahn, *Stability of Motion*. Berlin, Germany: Springer-Verlag, 1967.
[10] R. W. Brockett, *Finite Dimensional Linear Systems*. New York: Wiley, 1970.
[11] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Stat.*, vol. 22, pp. 400–407, 1951.
[12] J. Blum, "Multidimensional stochastic approximation methods," *Ann. Math. Stat.*, vol. 25, pp. 737–744, 1954.
[13] M. A. Aizermann, E. M. Braverman, and L. I. Rozonoer, "Metod Potentsialnykh funktsij v teorii obuchenija mashin" (The method of potential functions in the theory of machine learning), *Izd. Nauka,* Moscow, 1970 (in Russian).
[14] M. T. Wasan, *Stochastic Approximation*. Cambridge University Press, 1969.
[15] Ya. Z. Tsypkin, *Adaption and Learning in Automatic Systems*. New York: Academic, 1971.
[16] ——, *Foundations of the Theory of Learning Systems*. New York: Academic, 1973.
[17] J. Kiefer and J. Wolfowitz, "Stochastic estimation of the maximum of a regression function," *Ann. Math. Stat.*, vol. 23, pp. 462–466, 1952.
[18] H. J. Kushner, "Stochastic approximation algorithms for the local optimization of functions with non-unique stationary points," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 646–655, 1972.
[19] H. J. Kushner and T. Gavin, "Stochastic approximation type methods for constrained systems: Algorithms and numerical results," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 349–358, 1974.
[20] K. S. Fu, "Learning system theory," in *System Theory*, L. A. Zadeh and E. Polak, Eds. New York: McGraw-Hill, 1969, pp. 425–466.
[21] G. N. Saridis, Z. J. Nikolic, and K. S. Fu, "Stochastic approximation algorithms for systems identification, estimation and decomposition of mixtures," *IEEE Trans. Syst. Sci. Cybern.*, vol. SSC-5, pp. 8–15, 1969.
[22] Ya. Z. Tsypkin, "Self-learning—What is it?," *IEEE Trans. Automat. Contr.*, vol. AC-13, pp. 608–612, 1968.
[23] E. M. Braverman, "The method of potential functions in the problem of training machines to recognize patterns without a teacher," *Automat. Remote Contr.*, vol. 27, pp. 1748–1770, 1966.
[24] K. J. Aström and P. Eykhoff, "System identification—A survey," *Automatica*, vol. 7, pp. 123–164, 1971.
[25] J. M. Mendel, *Discrete Techniques of Parameter Estimation*. New York: Marcel Dekker, 1973.
[26] L. L. Scharf and D. C. Farden, "Optimum and adaptive array processing in frequence-wavenumber space," in *Proc. IEEE Conf. on Decision and Control*, Phoenix, AZ, 1974, pp. 604–609.
[27] D. C. Farden, "Stochastic approximation with correlated data," Dep. Elect. Eng., Colorado State University, Fort Collins, CO, ONR Tech. Rep. 11.
[28] I. D. Landau, "Unbiased recursive identification using model reference adaptive techniques," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 194–202, Apr. 1976.
[29] L. Ljung, "On positive real transfer functions and the convergence of some recursive schemes," this issue, pp. 539–551.
[30] T. Söderström, L. Ljung, and I. Gustavsson, "A comparative study of recursive identification methods," Automatic Control, Lund Institute of Technology, Lund, Sweden, Rep. 7427, 1974.
[31] L. Ljung, "Convergence of an adaptive filter algorithm," *Int. J. Contr.*, 1977, to be published.
[32] K. J. Aström and B. Wittenmark, "On self-tuning regulators," *Automatica*, vol. 9, pp. 185–199, 1973.
[33] L. Ljung and B. Wittenmark, "On a stabilizing property of adaptive regulators," in *Preprints 4th IFAC Symp. Identification*, Part 3, Tbilisi, USSR, Paper 19.1, pp. 380–389, 1976.
[34] K. J. Aström and B. Wittenmark, "Analysis of a self-tuning regulator for nonminimum phase systems," in *Proc. IFAC Symp. on Stochastic Control*, Budapest, 1974.
[35] D. P. Derevitskii and A. L. Fradkov, "Two models for analyzing the dynamics of adaptation algorithms," *Automat. and Remote Contr.*, vol. 35, pp. 59–67, 1974.
[36] H. J. Kushner, "General convergence results for stochastic approximations via weak convergence theory," Brown University, Providence, RI, LCDS Tech. Rep. 76-1, 1976.
[37] ——, "Convergence of recursive adaptive and identification procedures via weak convergence theory," *IEEE Trans. Automat. Contr.*, to be published.
[38] H. Cramér and M. R. Leadbetter, *Stationary and Related Stochastic Processes*. New York: Wiley, 1967.
[39] N. N. Krasovskij, *Stability of Motion*. Stanford, CA: Stanford University Press, 1963.
[40] W. Hahn, *Stability of Motion*. Berlin, Germany: Springer-Verlag, 1967.
[41] V. I. Zubov, *Methods of A. M. Lyapunov and their Applications*. Groningen, Holland: Noordhoff, 1964.

Lennart Ljung (S'74–M'75), for a photograph and biography see page 551 of this issue.