Multirate adaptive inferential estimation

W. Lu
D.G. Fisher

Indexing terms: Adaptive control, Mathematical techniques, Estimation

Abstract: An adaptive inferential algorithm is developed for estimation and control of multirate systems. The output y is measured J times slower than the secondary process output v and the input u, but an output estimate y_e is produced at each sampling interval of v and u. Compared with previous work on multirate inferential systems, the proposed algorithm has a more formal theoretical basis. For example, the output y is related to the secondary output v not only through external stochastic disturbances but also through the internal system structure. Convergence properties are formally proven for the case of zero external stochastic disturbances, and a simplified algorithm is proposed for practical applications. Simulated results illustrate the convergence properties of the algorithm and the improvement obtained in simple feedback control systems.

1 Introduction

This paper combines some of the concepts used in multirate output estimation [5, 6, 9] where the output is sampled J times slower than the input with those of inferential control, e.g. Reference 8, where a secondary output is used to improve the estimate of the primary-output values. The resulting multirate adaptive inferential estimation and control system is shown in Fig. 1. The

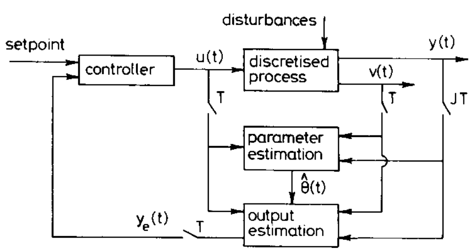

Fig. 1 *Multirate inferential estimation and control system*

primary-process output $y(t)$ is sampled with a period JT while the input $u(t)$ and secondary output $v(t)$ are sampled at the desired control interval T. These measurements are sent in parallel to a parameter-identification algorithm and an output-estimation algorithm. The output estimates $y_e(t)$ are produced every sampling interval T, and can be used as shown in Fig. 1 as the basis for control. If y_e is a good estimate of y, performance should improve because the control action can be implemented with period T rather than JT.

The estimation and control system shown in Fig. 1 was first proposed by Guilandoust *et al.* [3] who proposed two approaches: state space and input–output. However, their state-space approach requires that the process be completely observable from the secondary output $v(t)$. This requirement is severely restrictive since in most cases the dynamic modes of the primary output $y(t)$ are not all included in the dynamic modes of $v(t)$. Their input–output approach does not require this observability assumption. It directly assumes that the process has two input–output models, one for $y(t)$ against $u(t)$, another for $v(t)$ against $u(t)$. However, the formulation requires that the same white-noise term be present in each of the two models to relate $y(t)$ with $v(t)$ and to obtain the working equation of the algorithm. This approach does not adequately reflect the link between y and v, e.g. the link is only via the external white-noise disturbance and if this disturbance vanished then there would be no theoretical basis for the working equation. Furthermore, it is difficult to interpret the physical meaning of the working equation, e.g. the relationship between the order of the polynomials in the working equation and the characteristics of the actual process is not clearly defined.

This paper formulates the working equation based on the framework of linear models [10]. The working equation defines a more fundamental, inferential relationship from v to y via the internal system structure. The proposed approach avoids the need for the limiting assumptions made by Guilandoust *et al.* [3]; quantitatively defines the relationship between the working equation and the original process model plus the external disturbances; permits formal proof of the output-convergence properties; and provides a solid theoretical background for extending the result to multi-input/multi-output cases.

When J, i.e. the output sampling interval, is increased, the number of parameters to be identified increases proportionally. For cases when the number of estimated parameters must be reduced, a simplified algorithm is proposed which works well in simulations but lacks a formal proof of convergence.

2 Models for multirate inferential estimation

In the following discussion, models are derived first for multirate systems without external disturbances and then for systems with deterministic and/or stochastic disturbances. It is assumed that the process shown in Fig. 1 is completely observable from $v(t)$ plus $y(t)$, which is much

Paper 8383D (C8), first received 27th February 1989 and in revised form 27th June 1991

Dr. Fisher is with the Department of Chemical Engineering, University of Alberta, Edmonton, Canada T6G 2G6
Dr. Lu was with the University of Alberta and is now with Control Systems Group, Cominco Ltd., Trail, BC, Canada V1R 4L8

less restrictive than assuming that it is completely observable from $v(t)$ alone [3]. The process is of order n with an observability index nv from $v(t)$. To simplify the notation, it is assumed that the input sampling interval $T = 1$ and t is also used to indicate discrete time.

2.1 Case 1: no disturbances

By the observability assumption, the system can be represented as [10]

$$x(t+1) = \begin{bmatrix} 0 & \cdots & 0 & -a_1 & 0 & \cdots & 0 & -\bar{a}_1 \\ & I_{nv-1} & & \vdots & \vdots & & \vdots & \vdots \\ & & & -a_{nv} & 0 & \cdots & 0 & -\bar{a}_{nv} \\ 0 & \cdots & 0 & -a_{nv+1} & 0 & \cdots & 0 & -\bar{a}_{nv+1} \\ \vdots & & \vdots & \vdots & & I_{ny-1} & & \vdots \\ 0 & \cdots & 0 & -a_n & & & & -\bar{a}_n \end{bmatrix} x(t) + \begin{bmatrix} b_1 \\ \vdots \\ b_{nv} \\ b_{nv+1} \\ \vdots \\ b_n \end{bmatrix} u(t) \quad (1)$$

$$v(t) = x_{nv}(t) \quad (2)$$

$$y(t) = hx_{nv}(t) + x_n(t) \quad (3)$$

where $ny = n - nv$. The corresponding input–output relation from $u(t)$ and $v(t)$ to $y(t)$ can be expressed as

$$A(q^{-1})[y(t) - hv(t)] = B(q^{-1})u(t) + \bar{C}(q^{-1})v(t) \quad (4)$$

where

$$A(q^{-1}) = 1 + \bar{a}_n q^{-1} + \bar{a}_{n-1} q^{-2} + \cdots + \bar{a}_{nv+1} q^{-ny} = \prod_{i=1}^{ny} [1 - (\lambda_i q)^{-1}] \quad (5)$$

The λ_i are the roots of $A(q^{-1})$,

$$B(q^{-1}) = b_n q^{-1} + b_{n-1} q^{-2} + \cdots + b_{nv+1} q^{-ny} \quad (6)$$

and

$$\bar{C}(q^{-1}) = -a_n q^{-1} - a_{n-1} q^{-2} - \cdots - a_{nv+1} q^{-ny} \quad (7)$$

Let

$$C(q^{-1}) = A(q^{-1})h + \bar{C}(q^{-1}) = h + (\bar{a}_n h - a_n) q^{-1} + \cdots + (\bar{a}_{nv+1} h - a_{nv+1}) q^{-ny} \quad (8)$$

Then

$$A(q^{-1})y(t) = B(q^{-1})u(t) + C(q^{-1})v(t) \quad (9)$$

Multiplying both sides of eqn. 9 by

$$\prod_{i=1}^{ny} [1 + (\lambda_i q)^{-1} + \cdots + (\lambda_i q)^{2-J} + (\lambda_i q)^{1-J}]$$

results in the equivalent form

$$A_J(q^{-J})y(t) = B_J(q^{-1})u(t) + C_J(q^{-1})v(t) \quad (10)$$

where

$$A_J(q^{-J}) = 1 + a_{J1} q^{-J} + a_{J2} q^{-2J} + \cdots + a_{Jny} q^{-nyJ} \quad (11)$$

$$B_J(q^{-1}) = b_{J1} q^{-1} + b_{J2} q^{-2} + \cdots + b_{Jm} q^{-m} \quad (12)$$

$$C_J(q^{-1}) = C_{J0} + C_{J1} q^{-1} + C_{J2} q^{-2} + \cdots + C_{Jm} q^{-m} \quad (13)$$

and $m = J \times ny$.

The working equation (eqn. 10) can also be derived directly from the original continuous model of the process with a discretised input [5].

The stability property of $A_J(q^{-J})$ follows that of $A(q^{-1})$. If the parameters of $A(q^{-1})$ are real, so are those of $A_J(q^{-J})$.

2.2 Case 2: deterministic and stochastic disturbances

If the dynamic modes of the deterministic disturbances do not result in pole-zero cancellation with the dynamic modes of the process, the composite system, i.e. the process system plus deterministic disturbances, can be represented by

$$x(t+1) = \begin{bmatrix} 0 & \cdots & 0 & -a_1 & 0 & \cdots & 0 & -\bar{a}_1 \\ & I_{nv-1} & & \vdots & \vdots & & \vdots & \vdots \\ & & & -a_{nv} & 0 & \cdots & 0 & -\bar{a}_{nv} \\ 0 & \cdots & 0 & -a_{nv+1} & 0 & \cdots & 0 & -\bar{a}_{nv+1} \\ \vdots & & \vdots & \vdots & & I_{ny-1} & & \vdots \\ 0 & \cdots & 0 & -a_n & & & & -\bar{a}_n \end{bmatrix} x(t) + \begin{bmatrix} b_1 \\ \vdots \\ b_{nv} \\ b_{nv+1} \\ \vdots \\ b_n \end{bmatrix} u(t) + \begin{bmatrix} r_1 \\ \vdots \\ r_{nv} \\ r_{nv+1} \\ \vdots \\ r_n \end{bmatrix} \bar{w}(t) \quad (14)$$

$$v(t) = x_{nv}(t) + \eta_v(t) \quad (15)$$

$$y(t) = hx_{nv}(t) + x_n(t) + \eta_y(t) \quad (16)$$

where the augmented state variable x includes the dynamics of deterministic disturbances. The order of the composite system (eqn. 14) is greater than or equal to the order of the process but is still represented by n. The stochastic disturbances $\bar{w}(t)$, $\eta_v(t)$ and $\eta_y(t)$ are assumed to be white Gaussian sequences with finite variances.

As in the zero-disturbance case, the input–output relationship between $u(t)$, $\bar{w}(t)$, $v(t)$ and $y(t)$ can be obtained as follows:

$$A(q^{-1})[y(t) - hv(t) + h\eta_v(t) - \eta_y(t)] = B(q^{-1})u(t) + \bar{C}(q^{-1})v(t) + R(q^{-1})\bar{w}(t) \quad (17)$$

where A, B and $\bar{C}$ have the form of eqns. 5–7 and

$$R(q^{-1}) = r_n q^{-1} + r_{n-1} q^{-2} + \cdots + r_{nv+1} q^{-ny} \quad (18)$$

Eqn. 17 can be rewritten (in a form similar to eqn. 9) as

$$A(q^{-1})y(t) = B(q^{-1})u(t) + C(q^{-1})v(t) + D(q^{-1})z(t) \quad (19)$$

where $z(t)$ is white and has a finite variance. In eqn. 19 $R(q^{-1})\bar{w}(t) - A(q^{-1})h\eta_v(t) + A(q^{-1})\eta_y(t)$ has been replaced by $D(q^{-1})z(t)$ using the representation and spectral factorisation principle [1]. The polynomial $D(q^{-1})$ is defined by

$$D(q^{-1}) = d_0 + d_1 q^{-1} + \cdots + d_{ny} q^{-ny} \quad (20)$$

Like eqn. 10, an equivalent form of eqn. 19 is

$$A_J(q^{-J})y(t) = B_J(q^{-1})u(t) + C_J(q^{-1})v(t) + D_J(q^{-1})z(t) \quad (21)$$

where

$$D_J(q^{-1}) = d_{J0} + d_{J1}q^{-1} + d_{J2}q^{-2} + \cdots + d_{Jm}q^{-m} \quad (22)$$

If the deterministic disturbances result in pole-zero cancellation with the process, the observability property of the composite system is lost. This situation, called input-zero blocking, is not discussed in this paper.

The derivation of the working equation (eqn. 10 or 21) is easily extended to multi-input/multi-output cases. For example, if there are several secondary measured variables rather than a single $v(t)$ the same approach can be used to derive an appropriate working equation. On the other hand, if $v(t)$ is not available, i.e. if $nv = 0$, the working equation naturally reduces to the single-input/single output case treated by Lu and Fisher [5, 6]. Note that it is almost impossible to make either the extension to the multi-input/multi-output case or the reduction to the single-input/single-output case when using the approach of Guilandust *et al.* [3].

3 Estimation algorithm

3.1 Definitions

The notation used to describe the process in eqns. 10 and 21 can be simplified by dropping the J subscript. The inferential model of the process then becomes

$$A(q^{-J})y(t) = B(q^{-1})u(t) + C(q^{-1})v(t) + D(q^{-1})z(t) \quad (23)$$

where

$$A(q^{-J}) = 1 + a_1q^{-J} + a_2q^{-2J} + \cdots + a_nq^{-nJ} \quad (24)$$

$$B(q^{-1}) = b_1q^{-1} + b_2q^{-2} + \cdots + b_mq^{-m} \quad (25)$$

$$C(q^{-1}) = c_0 + c_1q^{-1} + c_2q^{-2} + \cdots + c_mq^{-m} \quad (26)$$

$$D(q^{-1}) = d_0 + d_1q^{-1} + d_2q^{-2} + \cdots + d_mq^{-m}$$

$$m = J \times n \quad (27)$$

Here n plays the role of ny in the previous section. It is further assumed that the original state-space representation (eqns. 14–16) has all its eigenvalues inside the stable region. Let

$$\phi(t-1)^T = [-y(t-J), -y(t-2J), \ldots, -y(t-nJ), u(t-1), u(t-2), \ldots, u(t-m), v(t), v(t-1), \ldots, v(t-m)] \quad (28)$$

and

$$\theta_0^T = [a_1, \ldots, a_n, b_1, \ldots, b_m, c_0, c_1, \ldots, c_m] \quad (29)$$

Then

$$y(t) = \phi(t-1)^T\theta_0 + D(q^{-1})z(t) \quad (30)$$

Next, define the following:

A posteriori model output

$$\bar{y}(t) = \bar{\phi}(t-1)^T\hat{\theta}(t) \quad (31)$$

where

$$\bar{\phi}(t-1)^T = [-\bar{y}(t-J), -\bar{y}(t-2J), \ldots, -\bar{y}(t-nJ), u(t-1), u(t-2), \ldots, u(t-m), v(t), v(t-1), \ldots, v(t-m)] \quad (32)$$

with the initial values

$$\bar{\phi}(-1) = \phi(-1)$$

(available from the measurement data)

$$\bar{\phi}(-1+i) = \text{arbitrary} \quad \text{for } i = 1, 2, \ldots, J-1$$

$$\hat{\theta}(t)^T = [\hat{a}_1(t), \ldots, \hat{a}_n(t), \hat{b}_1(t), \ldots, \hat{b}_m(t), \hat{c}_0(t), \ldots, \hat{c}_m(t)] \quad (33)$$

i.e. $\hat{\theta}(t)$ is the estimate of θ_0 at time t.

A posteriori model output error

$$\eta(t) = y(t) - \bar{y}(t) \quad (34)$$

A priori model output

$$\hat{y}(t) = \bar{\phi}(t-1)^T\hat{\theta}(t-1) \quad (35)$$

A priori model output error

$$e(t) = y(t) - \hat{y}(t) \quad (36)$$

Generalised a posteriori output error

$$\bar{\eta}(t) = L(q^{-J})\eta(t) \quad (37)$$

where

$$L(q^{-J}) = 1 + l_1q^{-J} + \cdots + l_{l'}q^{-l'J} \quad (38)$$

is a fixed moving-average filter.

Generalised a priori output error

$$\bar{v}(t) = e(t) + [L(q^{-J}) - 1]\eta(t) \quad (39)$$

3.2 Estimation algorithm

The parameter estimation algorithm is given by

$$\hat{\theta}(Jt) = \hat{\theta}(Jt-J) + P(Jt-2)\bar{\phi}(Jt-1)\bar{v}(Jt) \div [1 + \bar{\phi}(Jt-1)^TP(Jt-2)\bar{\phi}(Jt-1)] \quad (40)$$

$$\hat{\theta}(Jt+i) = \hat{\theta}(Jt) \quad (i = 1, 2, \ldots, J-1) \quad (41)$$

$$P[J(t+1)-2] = P(Jt-2) - P(Jt-2)\bar{\phi}(Jt-1)\bar{\phi}(Jt-1)^TP(Jt-2) \div [1 + \bar{\phi}(Jt-1)^TP(Jt-2)\bar{\phi}(Jt-1)] \quad (42)$$

$$\hat{\theta}(0) = \text{arbitrary} \quad (43)$$

$$P(-2) > 0 \quad (44)$$

The regressor $\bar{\phi}(t)$ has been defined in eqn. 32. At each unity time step, the estimated outputs $\bar{y}(t)$ and $\hat{y}(t+1)$ can be calculated by using eqns. 31 and 35 although the output is measured only every J sampling intervals.

3.3 Convergence at the output sampling instants (for $z(t) \equiv 0$)

The first step is to define the convergence properties of the output estimates at the output (slow) sampling interval JT.

Theorem 1: Consider the algorithm (eqns. 40–44) applied to the inferential model in eqn. 23 with $z(t) \equiv 0$; then, provided that the system $H(q^{-J}) = [L(q^{-J})/A(q^{-J}) - 1/2]$ is very strictly passive:

$$\lim_{N\to\infty} \sum_{t=1}^{N} \eta(Jt)^2 < \infty \quad (45)$$

which implies that

$$\lim_{N\to\infty} \sum_{t=1}^{N} \bar{\eta}(Jt)^2 < \infty \tag{46}$$

and

$$\lim_{t\to\infty} |y(Jt) - \bar{y}(Jt)| = 0 \tag{47}$$

Also

$$\lim_{N\to\infty} \sum_{t=1}^{N} \bar{\phi}(Jt-1)^T P(Jt-2)\bar{\phi}(Jt-1)\bar{\eta}(Jt)^2 < \infty \tag{48}$$

which implies that

$$\lim_{N\to\infty} \sum_{t=1}^{N} \|\hat{\theta}(Jt) - \hat{\theta}[J(t-s)]\|^2 < \infty \tag{49}$$

where s is any finite integer.

If $\{u(t)\}$ is bounded, then

$$\lim_{t\to\infty} \bar{v}(Jt) = 0 \tag{50}$$

and

$$\lim_{t\to\infty} |y(Jt) - \hat{y}(Jt)| = 0 \tag{51}$$

Proof: Define

$$b(t) = -\bar{\phi}(t-1)^T \tilde{\theta}(t) \tag{52}$$

where

$$\tilde{\theta}(t) = \hat{\theta}(t) - \theta_0 \tag{53}$$

Combining eqns. 30 and 31 [and noting that $z(t) \equiv 0$] gives

$$A(q^{-J})\eta(t) = b(t) \tag{54}$$

or more particularly

$$A(q^{-J})\eta(Jt) = b(Jt) \tag{55}$$

Then eqns. 37 and 55 give

$$A(q^{-J})\bar{\eta}(Jt) = L(q^{-J})b(Jt) \tag{56}$$

Multiplying eqn. 40 by $\bar{\phi}(Jt-1)^T$ and then subtracting from $y(Jt)$ gives

$$\eta(Jt) = e(Jt) - \bar{\phi}(Jt-1)^T P(Jt-2)\bar{\phi}(Jt-1)\bar{v}(Jt) \div [1 + \bar{\phi}(Jt-1)^T P(Jt-2)\bar{\phi}(Jt-1)] \tag{57}$$

Combining eqn. 57 with eqns. 37 and 39 yields

$$\bar{\eta}(Jt) = \bar{v}(Jt) \div [1 + \bar{\phi}(Jt-1)^T P(Jt-2)\bar{\phi}(Jt-1)] \tag{58}$$

Substituting eqn. 58 into eqn. 40 gives

$$\hat{\theta}(Jt) = \hat{\theta}(Jt-J) + P(Jt-2)\bar{\phi}(Jt-1)\bar{\eta}(Jt) \tag{59}$$

Subtracting θ_0 from both sides yields

$$\tilde{\theta}(Jt) - P(Jt-2)\bar{\phi}(Jt-1)\bar{\eta}(Jt) = \tilde{\theta}(Jt-J) \tag{60}$$

Let

$$V(Jt) = \tilde{\theta}(Jt)^T P[J(t+1)-2]^{-1}\tilde{\theta}(Jt) \tag{61}$$

and then from eqns. 60 and 61

$$[\tilde{\theta}(Jt) - P(Jt-2)\bar{\phi}(Jt-1)\bar{\eta}(Jt)]^T P(Jt-2)^{-1} \times [\tilde{\theta}(Jt) - P(Jt-2)\bar{\phi}(Jt-1)\bar{\eta}(Jt)]^T = V(Jt-J) \tag{62}$$

or

$$\tilde{\theta}(Jt)^T P(Jt-2)^{-1}\tilde{\theta}(Jt) - 2\tilde{\theta}(Jt)^T\bar{\phi}(Jt-1)\bar{\eta}(Jt) + \bar{\phi}(Jt-1)^T P(Jt-2)\bar{\phi}(Jt-1)\bar{\eta}^2(Jt) = V(Jt-J) \tag{63}$$

Using the inversion lemma (lemma 3.3.4 of Reference 2),

$$\begin{aligned}\tilde{\theta}(Jt)^T P(Jt-2)^{-1}\tilde{\theta}(Jt) &= \tilde{\theta}(Jt)^T\{P[J(t+1)-2]^{-1} - \bar{\phi}(Jt-1)\bar{\phi}(Jt-1)^T\}\tilde{\theta}(Jt)\\ &= V(Jt) - \tilde{\theta}(Jt)^T\bar{\phi}(Jt-1)\bar{\phi}(Jt-1)^T\tilde{\theta}(Jt)\end{aligned} \tag{64}$$

Combining eqns. 52, 62 and 63,

$$V(Jt) = V(Jt-J) - 2[\bar{\eta}(Jt) - b(Jt)/2]b(Jt) - \bar{\phi}(Jt-1)^T P(Jt-2)\bar{\phi}(Jt-1)\bar{\eta}^2(Jt) \tag{65}$$

The remainder of the proof is the same as that of theorem 3.5.1 in Reference 2, except that the very strictly passive condition is used for the relation between $[\bar{\eta}(Jt) - b(Jt)/2]$ and $b(Jt)$, where the input sequence is $b(0)$, $b(J)$, $b(2J)$, ... and the output sequence is $[\bar{\eta}(0) - b(0)/2]$, $[\bar{\eta}(J) - b(J)/2]$, $[\bar{\eta}(2J) - b(2J)/2]$, ...

3.4 Convergence at the input sampling instants (for $z(t) \equiv 0$)

Using the results from theorem 1, it is now possible to define the convergence properties of the output estimate at each input sampling interval T, i.e. at the output inter-sampling points.

Theorem 2: Under the same conditions as theorem 1:

$$\|\hat{\theta}(Jt) - \theta_0\|^2 \leqslant \kappa_1\|\hat{\theta}(0) - \theta_0\|^2 \quad \forall t > 0 \tag{66}$$

where

$$\kappa_1 = \lambda_{max}[P(-2)^{-1} + \phi(-1)\phi(-1)^T]/\lambda_{min}[P(-2)^{-1}],$$

and $\lambda_{max}(\cdot)$ and $\lambda_{min}(\cdot)$ represent the maximum and minimum eigenvalues of A in eqn. 23.

There exists a positive number ε such that if $\{u(t)\}$ is bounded then $\|\hat{\theta}(0) - \theta_0\|^2 < \varepsilon$ implies

$$(a) \quad |y(t) - \hat{y}(t)| \leqslant \delta\left[\limsup_{t\to\infty} \|\hat{\theta}(Jt) - \theta_0\|\right] \times \left[\limsup_{t\to\infty} |u(t) - u(t-1)|\right] + \Delta(t) \quad \forall t > 0 \tag{67}$$

where $\Delta(t)$ is a sequence satisfying $\lim_{t\to\infty} \Delta(t) = 0$, δ and the two limit superiors are finite numbers.

$$(b) \quad \lim_{t\to\infty} |y(t) - \hat{y}(t)| = 0 \tag{68}$$

provided that $\lim_{t\to\infty} \hat{\theta}(Jt) = \theta_0$ or $\lim_{t\to\infty} [u(t) - u(t-1)] = 0$.

Proof: Using eqn. 65,

$$V(Jt) = V(0) - 2\sum_{i=1}^{t} b(Ji)[\bar{\eta}(Ji) - b(Ji) \div 2] - \sum_{i=1}^{t} \bar{\phi}(Ji-1)^T P(Ji-2)\bar{\phi}(Ji-1)\bar{\eta}^2(Ji) \tag{69}$$

From eqn. 56

$$[\bar{\eta}(Jt) - b(Jt)/2] = [L(q^{-J})/A(q^{-J}) - 1/2]b(Jt) = H(q^{-J})b(Jt) \quad (70)$$

By the very strictly passive assumption of $H(q^{-J})$

$$\sum_{i=1}^{t} b(Ji)[\bar{\eta}(Ji) - b(Ji)/2] \geqslant -b(0)[\bar{\eta}(0) - b(0)/2] \quad (71)$$

Considering the initial conditions of $\bar{\phi}$,

$$\begin{aligned}\bar{\eta}(0) &= L(q^{-J})\eta(t)\Big|_{t=0} = \eta(0)\\ &= y(0) - \bar{y}(0)\\ &= \phi(-1)^T\theta_0 - \bar{\phi}(-1)^T\hat{\theta}(0)\\ &= \phi(-1)^T[\theta_0 - \hat{\theta}(0)]\\ &= -\tilde{\theta}(0)^T\phi(-1)\end{aligned} \quad (72)$$

But

$$b(0) = -\tilde{\theta}^T(0)\bar{\phi}(-1) = -\tilde{\theta}^T(0)\phi(-1) \quad (73)$$

and therefore

$$\sum_{i=1}^{t} b(Ji)[\bar{\eta}(Ji) - b(Ji)/2] \geqslant -b^2(0)/2 = -\tilde{\theta}^T(0)\phi(-1)\phi(-1)^T\tilde{\theta}(0)/2 \quad (74)$$

Substituting eqn. 74 into eqn. 69 and using the definition of V gives

$$\tilde{\theta}(Jt)^T P[J(t+1) - 2]^{-1}\tilde{\theta}(Jt) \leqslant \tilde{\theta}^T(0)[P(-2) + \phi(-1)\phi(-1)^T]^{-1}\tilde{\theta}(0) \quad (75)$$

where the positive definite property of P is used. By the inversion lemma (lemma 3.3.4 of Reference 2) it is easy to verify that

$$\lambda_{\min}\{P[J(t+1) - 2]^{-1}\} \geqslant \lambda_{\min}[P(-2)^{-1}] \quad \forall t \geqslant 0 \quad (76)$$

Then eqns. 75 and 76 immediately yield eqn. 66.

From eqn. 54,

$$A(q^{-J})\eta(t) = b(t) \quad (77)$$

or equivalently, using eqn. 41,

$$A(q^{-J})\eta(Jt + i) = -\bar{\phi}(Jt + i - 1)\tilde{\theta}(Jt) \qquad i = 0, 1, 2, \ldots, j - 1 \quad (78)$$

From eqn. 78

$$\begin{aligned}A(q^{-J})[\eta(Jt + i) - \eta(Jt)] &= -[\bar{\phi}(Jt + i - 1) - \bar{\phi}(Jt - 1)]\tilde{\theta}(Jt)\\ &\qquad i = 1, 2, \ldots, J - 1\end{aligned} \quad (79)$$

But

$$\begin{aligned}&\bar{\phi}(Jt + i - 1)\\ &= \{[F_1(Jt - J, q^{-1})q^{-J}, F_1(Jt - 2J, q^{-1})q^{-2J}, \ldots,\\ &\qquad F_1(Jt - nJ, q^{-1})q^{-nJ}]u(Jt + i)\\ &\quad + [F_2(Jt - J, q^{-1})q^{-J}, F_2(Jt - 2J, q^{-1})q^{-2J}, \ldots,\\ &\qquad F_2(Jt - nJ, q^{-1})q^{-nJ}]v(Jt + i),\\ &\qquad [q^{-1}, q^{-2}, \ldots, q^{-m}]u(Jt + i),\\ &\qquad [1, q^{-1}, q^{-2}, \ldots, q^{-m}]v(Jt + i)\}\\ &\qquad\qquad i = 1, 2, \ldots, J - 1\end{aligned} \quad (80)$$

where

$$F_1(t, q^{-1}) = \hat{B}(t, q^{-1})/\hat{A}(t, q^{-J})$$

$$F_2(t, q^{-1}) = \hat{C}(t, q^{-1})/\hat{A}(t, q^{-J})$$

$$\hat{A}(t, q^{-J}) = 1 + \hat{a}_1(t)q^{-J} + \hat{a}_2(t)q^{-2J} + \cdots + \hat{a}_n(t)q^{-nJ}$$

$$\hat{B}(t, q^{-1}) = \hat{b}_1(t)q^{-1} + \hat{b}_2(t)q^{-2} + \cdots + \hat{b}_m(t)q^{-m}$$

and

$$\hat{C}(t, q^{-J}) = \hat{c}_0(t) + \hat{c}_1(t)q^{-J} + \hat{c}_2(t)q^{-2J} + \cdots + \hat{c}_n(t)q^{-nJ}$$

By the very strictly passive assumption, $A(q^{-J})$ is asymptotically stable. Thus there exists $\varepsilon_1 > 0$ such that if $\|\hat{\theta}(0) - \theta_0\|^2 < \varepsilon_1$, then $\hat{A}(0, q^{-J})$ is also asymptotically stable. From the first part of theorem 2 it is concluded that $\hat{A}(Jt, q^{-J})$ $\forall t \geqslant 0$ has uniformly all its eigenvalues strictly inside the unit circle for $\varepsilon = \varepsilon_1/\kappa_1$. Also from eqn. 49, eqn. 80 is slowly time varying. Therefore bounded $\{u(t)\}$ and $\{v(t)\}$ imply bounded $\{\bar{\phi}(t)\}$ and the existence of $0 < M_1, M_2, M < \infty$ such that

$$\begin{aligned}&\|\bar{\phi}(Jt + i - 1) - \bar{\phi}(Jt - 1)\|\\ &\quad \leqslant M_1|u(Jt + i) - u(Jt)| + M_2|v(Jt + i) - v(Jt)|\\ &\quad \leqslant M|u(Jt + i) - u(Jt)|\\ &\qquad\qquad i = 1, 2, \ldots, J - 1\end{aligned} \quad (81)$$

since $\bar{\phi}(kJ + i - 1) - \bar{\phi}(kJ - 1)$ against $u(kJ + i) - u(kJ)$ and $v(kJ + i) - v(kJ)$, $i = 1, \ldots, J - 1$ also satisfies eqn. 80 and the relationship between $v(t)$ and $u(t)$ is linear time invariant and asymptotically stable by the assumption on the eigenvalues of the original process. Considering that $\lim_{t\to\infty} \eta(Jt) = 0$ (from eqn. 47) and $A(q^{-J})$ is asymptotically stable, eqns. 79 and 81 yield

$$\begin{aligned}|\eta(Jt + i)| &\leqslant \delta_1\left[\limsup_{t\to\infty} \|\tilde{\theta}(Jt)\|\right]\\ &\quad \times \left[\limsup_{t\to\infty} |u(Jt + i) - u(Jt)|\right] + \Delta(Jt + i)\\ &\qquad\qquad i = 1, 2, \ldots, J - 1\end{aligned} \quad (82)$$

where $\Delta(Jt + i)$, $i = 1, 2, \ldots, J - 1$ are some sequences satisfying $\lim_{t\to\infty} \Delta(Jt + i) = 0$, δ_1 and the two limit superiors are finite positive numbers. Note that, for $i = 1, 2, \ldots, J - 1$, $\lim_{t\to\infty} \sup|u(Jt + i) - u(Jt)| \leqslant (J - 1) \lim_{t\to\infty} \sup|u(t) - u(t - 1)|$. Part (*a*) is obtained by letting $|y(Jt) - \hat{y}(Jt)| = \Delta(Jt)$, $\delta = (J - 1)\delta_1$ noting that

$$\begin{aligned}\eta(Jt + i) &= y(Jt + i) - \bar{y}(Jt + i)\\ &= y(Jt + i) - \bar{\phi}(Jt + i - 1)^T\hat{\theta}(Jt + i)\\ &= y(Jt + i) - \bar{\phi}(Jt + i - 1)^T\hat{\theta}(Jt)\\ &= y(Jt + i) - \hat{y}(Jt + i)\\ &\qquad\qquad i = 1, 2, \ldots, J - 1\end{aligned} \quad (83)$$

and using eqns. 51 and 82. Part (*b*) is obvious.

Theorems 1 and 2 are extensions of the results in Lu and Fisher [6]. Here the secondary output $v(t)$ is included in the algorithm and the proofs are more fully developed.

The convergence results of theorems 1 and 2 are applicable to cases with $z(t) \equiv 0$ and can be applied to the convergence analysis of any adaptive servocontrol using the multirate inferential estimation algorithm.

The estimation algorithm (eqns. 40–44) can be applied to processes operating in noisy, stochastic disturbance environments. However, in general, if external stochastic disturbances are present then there is model mismatch since the parameters of the external stochastic disturbance model $D(q^{-1})$ are not identified. This makes the convergence analysis quite difficult.

Note that only output convergence is proven and no conclusion is made about parameter convergence to the true parameter vector. It is not difficult to observe from the derivation in the previous section that the parameterisation of the inferential equation [e.g. eqn. 23 with $z(t) \equiv 0$] is, in general, not unique, i.e. θ_0 can be anything belonging to an equivalence class set in the parameter vector space. It would be desirable, where parameter convergence is important, to use some improved algorithm with structurally constrained inferential working equations so that only a unique convergence point in the parameter vector space exists for identification.

3.5 Simplified algorithm

The algorithm (eqns. 40–44) has one practical disadvantage: the number of parameters to be estimated increases linearly with J. Since an exact model match cannot be achieved by the algorithm when $D(q^{-1})z(t) \neq 0$, there is no reason not to simplify the algorithm in such cases. Decreasing the number of parameters in the algorithm would, in general, increase the model mismatch and result in poorer performance. However, in some applications reducing the number of parameters to be estimated improves the numerical conditioning of the estimation algorithm and reduces the variance of the output estimate, thus resulting in better overall performance. Based on this observation (cf. parsimony principle [4]), it is proposed that the algorithm defined by eqns. 40–44 be simplified by reducing the number of $\hat{b}$ and $\hat{c}$ parameters. One extreme case, considered here as a demonstration example, is to set $\hat{b}_{i+1}(t) = \hat{c}_i(t) \equiv 0$ for $i \neq 0, J, 2J, \ldots, (n-1)J$. The number of parameters to be estimated in the proposed simplified algorithm is $3n$ and is therefore independent of J. The algorithm could be simplified or modified in other ways, for particular applications. However the simplified algorithm considered here is similar in form to the original algorithm and can achieve an exact model match (if $\hat{c}_{nJ}$ is not set to zero and included in the parameter estimate vector) at the output sampling intervals if $D(q^{-1})z(t) = 0$ and the input u is kept constant within each primary output sampling interval, i.e. over J intervals.

3.6 Predictive estimation

A one-step-ahead prediction of $y(t)$ can be calculated from the *a priori* mode (eqn. 13):

$$\hat{y}(t+1) = \bar{\phi}(t)^T\hat{\theta}(t) \tag{84}$$

In general, to predict $y(t)$ k steps ahead:

$$y_e(t+k) = \phi_e^T(t-1+k)\hat{\theta}(t) \tag{85}$$

where

$$\phi_e^T(t-1+k) = [-\bar{y}_e(t-J+k), -\bar{y}_e(t-2J+k), \ldots, -\bar{y}_e(t-nJ+k), \ldots] \tag{86}$$

and

$$\begin{aligned} \bar{y}_e(\tau) &= \bar{y}(\tau) \qquad \text{if } \tau \leqslant t \\ \bar{y}_e(\tau) &= y_e(\tau) \qquad \text{if } t < \tau \end{aligned} \tag{87}$$

4 Simulation results

The simulated process is given by

$$y = \left(\frac{-5.9}{7.8s+1} + \frac{\alpha}{3.6s+1}\right)e^{-ks}u + \frac{-6.34}{19.2s+1}e^{-ks}w \tag{88}$$

$$v = \frac{-16.877}{3.6s+1}u + \frac{-18.0}{9s+1}w \tag{89}$$

This model is a modified version of the linearised distillation column model described by Patke *et al.* [7], where y is the composition of one component of the overhead, u is the reflux rate, v is a temperature measured at an appropriate stage in the column and w is a disturbance in the feed. The process output $y(t)$ was scaled to a comparable numerical magnitude with respect to that of u and v, and for convenience the model was considered to be dimensionless. The term $\{\alpha/(3.6s+1)\}e^{-ks}u$ in eqn. 88 was added to the original model of Patke *et al.* [7] to illustrate different modes of coupling between $v(t)$ and $y(t)$. (In the following simulation examples $\alpha = 0.0$ or $\alpha = 0.5$.)

4.1 Open-loop output estimation without disturbances ($J = 10$)

For the process defined by eqns. 88 and 89 the time delay $k = 0$, the disturbance $w = 0$ and the input u is a PRBS sequence passed through a zero-order hold. The sampling interval for u and v is one time unit but the sampling interval for the output y is 10 units, i.e. $J = 10$. The appropriate working equation, as defined in Section 2, is

$$(1 + a_1q^{-10})y(t) = \sum_{i=1}^{10} b_i q^{-i}u(t) + \sum_{i=0}^{10} c_i q^{-i}v(t) \tag{90}$$

For the simplified algorithm, the following working equation is used:

$$(1 + a_1q^{-10})y(t) = b_1q^{-1}u(t) + c_0 v(t) \tag{91}$$

Since a zero-order hold is used, the algorithm (eqns. 40–44) with the working equation (eqn. 90) (full algorithm) allows an exact model match to the simulated plant model.

If the simplified algorithm is used, there is model mismatch because of the insufficient number of estimated parameters.

When using the full algorithm, it is not necessary to use $v(t)$ even if $\alpha \neq 0$. However, if $v(t)$ is not used the assumptions about the state representation of the process change accordingly, e.g. the observability index with respect to $v(t)$ becomes zero. Therefore the structure of the working equation has to be reformulated (see Lu and Fisher [5, 6]), i.e. the working equation cannot be obtained by simply dropping the $v(t)$ terms in eqn. 90. The output estimate will still converge to the real output and the number of parameters to be estimated does not change [6]. The advantage of using $v(t)$ is that each common mode shared by y and v will, in general, reduce the length of the data window by J and improve the numerical conditioning of the algorithm.

When the simplified algorithm is used, the secondary measurement $v(t)$ is necessary even if $a = 0$ because it partially compensates for the information lost by omitting $J - 1$ values of $u(t)$ in the regressor.

Note that with the approach of Guilandanst *et al.* [3] their working equation cannot be formulated for this case with $w(t) = 0$, since the existence of a sustained external stochastic disturbance is assumed by their approach.

The performance of the full algorithm is shown in Fig. 2 ($\alpha = 0$) and Fig. 3 ($\alpha = 0.5$) and shows excellent con-

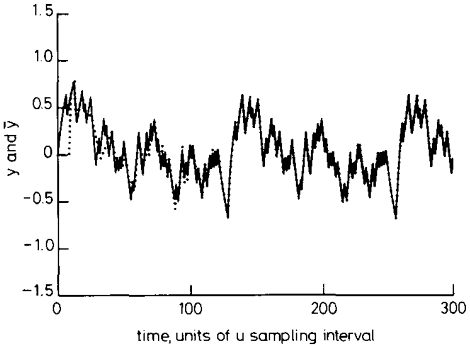

Fig. 2 *Output estimation with full algorithm*
u = PRBS; $\alpha = 0$; ——— y; ········ $\bar{y}$

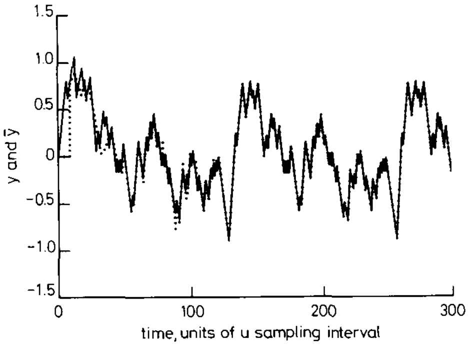

Fig. 3 *Output estimation with full algorithm*
u = PRBS; $\alpha = 0.5$; ——— y; ········ $\bar{y}$

vergence of the output estimates to the real output. The output estimates $\bar{y}(t)$ in Figs. 4 and 5 produced by the simplified algorithm are not as good as the corresponding estimates in Figs. 2 and 3, but are still a good approximation of $y(t)$. The sacrifice in output estimation accuracy may be worthwhile since the number of estimated parameters is significantly reduced (from 22 to 3).

4.2 PI feedback control with disturbances ($J = 10$)

The disturbance $w(t)$ is defined by the following series of step changes:

$$w(t) = +0.7 \quad 100(i-1) \leqslant t < 100i \quad i = 1, 3, 5, \ldots$$
$$w(t) = -0.7 \quad 100i \leqslant t \leqslant 100(i+1) \quad i = 1, 3, 5, \ldots \tag{92}$$

The control objective is to maintain the output at the desired value $y = 0$ using u as the control variable. A conventional proportional-integral feedback controller is chosen for simplicity so that

$$u(It) = k_c\, y_b(It) + (k_c\, I/T_i) \sum_{i=0}^{t} y_b(Ii) \quad t = 1, 2, \ldots \tag{93}$$

When eqn. 93 is used with the adaptive inferential estimation algorithms, a small perturbation signal is added to u to improve the excitation. y_b and I ($I = 1$ or J) are selected for each specific case as described below.

Ideal case: The best control should be obtained when all measurements are available at the desired control interval, i.e. $I = 1$ and $y_b(It) = y(t)$ in eqn. 93. Fig. 6 shows the closed loop response with and without the process time delay ($k = 5$). The controller constants $K_c = 0.2$ and $T_i = 10$ were obtained by trial and error tuning.

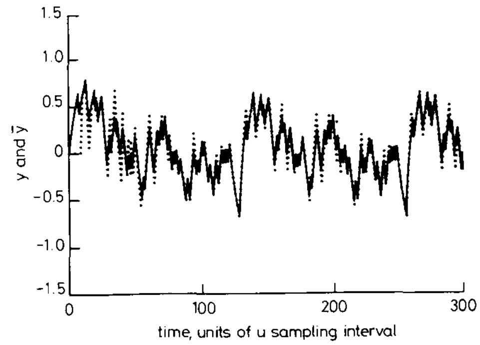

Fig. 4 *Output estimation with simplified algorithm*
u = PRBS; $\alpha = 0$; ——— y; ········ $\bar{y}$

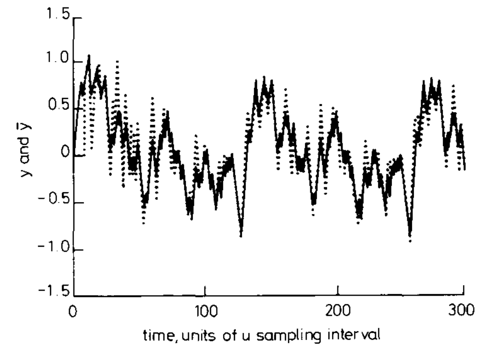

Fig. 5 *Output estimation with simplified algorithm*
u = PRBS; $\alpha = 0.5$; ——— y; ········ $\bar{y}$

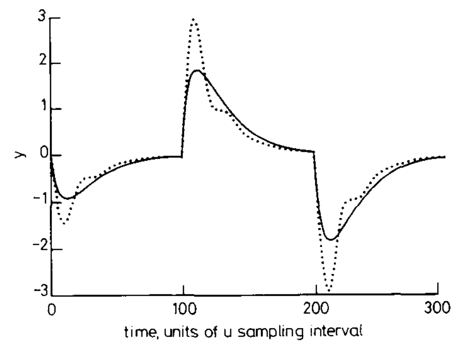

Fig. 6 *Closed loop PI control, ideal case (using the process output $y(t)$ at the fast control sampling intervals)*
——— no delay
········ with delay

Practical case: With the conventional proportional-integral control scheme if $J = 10$ the control interval must be increased from one time unit to 10, i.e. $I = J = 10$ and $y_b(It) = y(It)$ in eqn. 93. As expected, the control performance (solid line in Fig. 7) becomes oscillatory and control detuning is required (dotted line, T_i increased to 60). If a time delay $k = 5$ is included, the performance degradation is even more severe (results not shown).

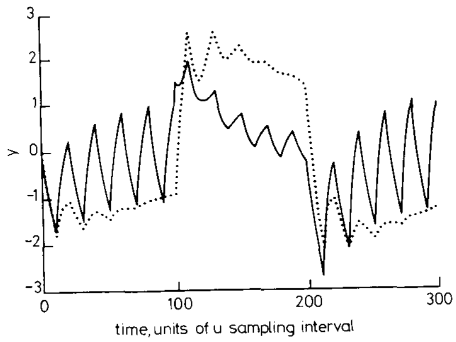

Fig. 7 *Closed loop PI control, practical case (using the process output $y(t)$ at the slow sampling intervals, i.e. every JT intervals)*
——— with delay; $K_c = 0.2$; $T_i = 10$
········ $T_i = 60$

Multirate inferential control with delay compensation: The output is sampled every $J = 10$ intervals but estimates of the output y_e are produced at every control interval, i.e. $I = 1$ and $y(It) = y_e(t + k)$ in eqn. 93. As shown in Figs. 8

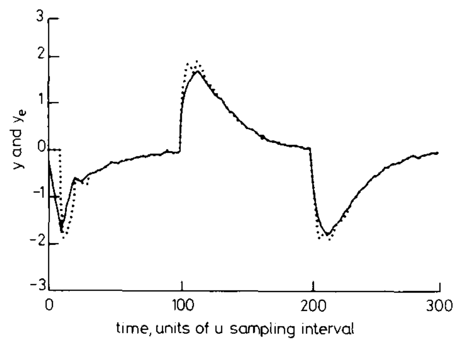

Fig. 8 *Multirate inferential control with no delay*
——— y
········ y_e

($k = 0$) and 9 ($k = 5$), the estimated output values are very close to the true values and control performance is very close to the ideal case plotted in Fig. 6. However the case with time delay (Fig. 9) is not as good as the corresponding case with no delay.

Simplified multirate inferential control: The open-loop output estimation results in Figs. 2–5 showed that the output estimation error increased when the simplified algorithm was used. However, Fig. 10 shows that under closed-loop conditions the simplified algorithm produced output estimates and control performance equal to, or better than the full algorithm (Fig. 8). When a process delay is included (Fig. 11), the output estimates and the control performance are again slightly better than the full algorithm (Fig. 9).

The improved results obtained with the simplified algorithm are example dependent. However, for a given application it is obviously worthwhile to evaluate both the full and various simplified algorithms.

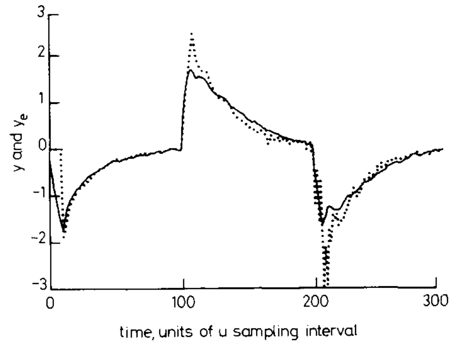

Fig. 9 *Multirate inferential control with delay $k = 5$*
——— y
········ y_e

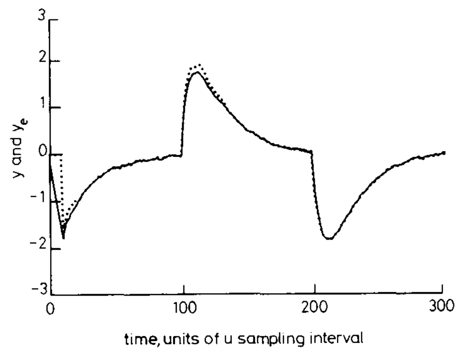

Fig. 10 *Multirate simplified inferential control with no delay*
——— y
········ y_e

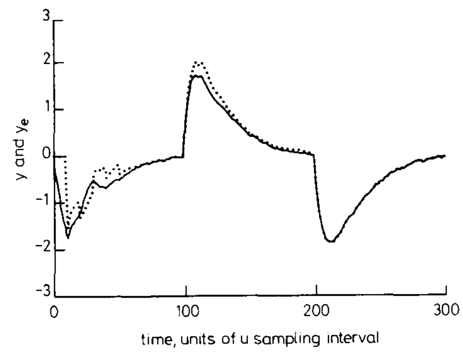

Fig. 11 *Multirate simplified inferential control with delay $k = 5$*
——— y
········ y_e

5 Conclusions

A multirate inferential estimation algorithm based on $\{u(t), v(t), y(Jt), t = 0, 1, 2, \ldots\}$ is derived. The output convergence properties are formally proven for the case without unmeasured external stochastic disturbances.

Multirate inferential control using the output estimates $y_e(t)$ rather than the measured values $y(Jt)$ is significantly better than the comparable conventional single-rate control scheme using $y(Jt)$ and approaches that of the ideal case where the output is measured every sampling interval, i.e. the output values $y(t)$ are used.

A simplified inferential algorithm is presented which actually outperforms the full algorithm in the closed-loop-simulation example. However, the convergence properties are not formally proven.

The algorithm has direct application in the process industries (e.g. distillation columns) where the output measurements $y(Jt)$ (e.g. composition) are available only at intervals J times slower than the desired control interval, and a secondary measurement $v(t)$ (e.g. temperature) is available at every control interval. It can also be used as an alternative to many conventional cascade control loops in which the outer loop operates with a sampling interval JT and the inner loop with T.

6 Acknowledgment

The authors acknowledge the contributions of Dr. A.J. Morris, one of the authors of Reference 3, who was a member of the academic staff at the University of Alberta in 1985–86 and provided some of the initial motivation and direction for this project.

7 References

1 ASTROM, K.J.: 'Introduction to stochastic control theory' (Academic Press, 1970)
2 GOODWIN, G.C., and SIN, K.S.: 'Adaptive filtering prediction and control' (Prentice-Hall, Englewood Cliffs, NJ, 1984)
3 GUILANDOUST, M.T., MORRIS, A.J., and THAM, M.T.: 'Adaptive estimation in inferential process control'. Proceedings of American Control Conference, 1987, pp. 1743–1748
4 LJUNG, L.: 'System identification: theory for the user' (Prentice-Hall, Englewood Cliffs, NJ, 1987)
5 LU, W., and FISHER, D.G.: 'Output estimation with multirate sampling', *Int. J. Control*, 1988, **48**, (1), pp. 149–160
6 LU, W., and FISHER, D.G.: 'Least squares output estimation with multirate sampling', *IEEE Trans.*, 1989, **AC-34**, (6), pp. 669–672
7 PATKE, N.G., DESHPANDE, P.B., and CHOU, A.C.: 'Evaluation of inferential and parallel cascade schemes for distillation control', *Ind. Eng. Chem. Process Des. Dev.*, 1982, **21**, pp. 266–272
8 PARRISH, J.R., and BROSILOW, C.B.: 'Inferential control applications', *Automatica*, 1985, **21**, (5), pp. 527–538
9 SODERSTROM, T.: 'Methods and applications in adaptive control' (Springer-Verlag, 1980)
10 WOLOVICH, W.A.: 'Linear multivariable system' (Springer-Verlag, New York, 1974)