Optimal Operation of Integrated Processing Systems

Part I: Open-Loop On-Line Optimizing Control

CARLOS E. GARCIA
MANFRED MORARI

Chemical Engineering Department
University of Wisconsin
Madison, WI 53706

A new algorithm for the continuous tracking of the optimal economic operating conditions of an integrated chemical plant has been developed. The method does not make use of fundamental models of the plant but is based on an on-line search using experimental moves of the independent variables. The results of these moves provide the data for the identification of a dynamic process model. This model allows one to determine the gradient of the objective function and thus guides the next move of the independent variable eventually leading to the optimum. The advantages of the new technique over other methods are demonstrated in simulation examples:

1) The speed of tracking is equally fast or faster than the relaxation speed of the system after a step change.
2) The noise insensitivity allows safe tracking of the optimum despite significant measurement errors.

As a key feature a decentralized form of the algorithm is developed making it suitable for distributed microprocessor implementation.

SCOPE

Due to modelling inaccuracies and the influence of unmeasurable process disturbances the economically optimum operating conditions of a process can generally not be predicted but are determined during the operation by experimentation. This procedure encounters difficulties of two kinds: 1) Long experimentation times do not allow one to track a drifting optimum; and 2) Measurement noise hides the interdependencies of input and output variables and makes the reliable determination of the optimal search direction difficult. Starting from the early work of Draper and Li (1951) over Evolutionary Operation (Box and Draper, 1969) to the application of modern optimization techniques, the large number of attempts toward a solution which are reviewed below, gives an indication of the importance of the problem. The new approach suggested here is based on the identification of a dynamic system model and uses measurements during the transient and not only at steady state. This makes the algorithm faster and less noise sensitive. Simulation studies clearly indicate the superiority over previous methods. While only the open loop unconstrained search is treated in this paper, the second part of this series will elucidate the modifications necessary when constraints are encountered and how the optimizing controller can be interfaced with a regulatory controller.

CONCLUSIONS AND SIGNIFICANCE

A new decentralized form of optimizing controller for integrated processing systems was developed and test results obtained from the simulation of a two CSTR example demonstrated its superiority with respect to speed and noise insensitivity. Both the speed of convergence and the degree of noise suppression can be influenced by several adjustable parameters of the optimizing controller. In general it was found exactly as for the tuning of filters that superior noise suppression slows down the parameter estimation and thus the speed of movement of the optimizing controller has to be reduced to avoid divergence. Good conservative a priori estimates of the controller parameters are suggested for which stability is virtually guaranteed. On-line tuning can be explored to accelerate the convergence. The robustness, reliability, speed and decentralized structure of the optimizing controller make it a likely candidate for complex industrial applications.

INTRODUCTION

Increasing energy and raw materials costs have forced the modern engineer to make better use of the available capacity of existing plants without much additional investment. Aside from methods which involve modifications of the process itself, as for example, energy integration schemes, other ways are being sought to exploit the available degrees of freedom and increase optimality. For instance, by continuously maintaining the plant at its steady-state optimum despite changing environmental conditions it is possible to achieve an important performance improvement.

0001-1541-81-4914-0960-$2.00.

Converter catalyst deactivation, heat exchanger equipment fouling, ambient temperature drifts and changes in feedstock quality offer a few examples of the disturbances which can have a lasting economic impact on the operation. Continuous tracking and driving the process to its best operating conditions when such changes occur are termed optimizing control.

To be applicable in an industrial environment in a nontrivial situation any optimizing controller (OC) should:

1) Perform on-line experiments in order to account for unmeasured disturbances and model inaccuracies
2) Find the optimum faster than the period of important process disturbances
3) Be able to deal with noisy measurements
4) Be able to handle multivariable, large scale systems

To present a method robust enough to take all these factors into consideration is the subject of this paper. We shall now examine currently used approaches to optimizing control and point out their useful features as well as their disadvantages.

PRESENT APPROACHES TO OPTIMIZING CONTROL

Off-Line Methods

Many techniques reported in the literature rely on the model prediction of optimal inputs. Key measurements are fed to a simulator or model and a search is performed off-line, making them fast but necessarily inaccurate (e.g., Webb et al., 1978; Prett and Gillette, 1979; Duncanson and Youle, 1970; Shah and Stillman, 1970; Davis et al., 1974).

On-Line Methods

1. Methods using a continuously varying perturbation signal: Early approaches to on-line optimizing control employed either ramp signals (Draper and Li, 1951) or sinusoidal inputs (Box and Chanmugam, 1962; Kotnour et al., 1966) to detect changes in performance. Besides requiring undesirable perturbations from an operating point of view these methods are generally noise sensitive and essentially limited to systems with one degree of freedom. Extensions even to two degrees encounter complex stability problems (White et al., 1968) and no applications have been reported.

2. Direct search methods: Multivariable systems have been handled by implementing either gradient or pattern-type search methods from optimization theory directly on the plant. Edler, et al., (1970) compare the performance of different techniques. Being essentially steady-state procedures, measurements should be taken only after the process has settled after each change in the manipulated variables, which makes the OC very slow. On the other hand, if the manipulated variables are perturbed without awaiting the end of the transient, complex stability problems arise (Sawaragi et al., 1971) which lie beyond a possible theoretical analysis.

3. Methods to handle measurement noise: In stochastic approximation the numerical on-line computed gradient is filtered and used to determine input moves. Although the scheme can be shown to converge in probability to the plant optimum (Saridis, 1974) discouraging application results have made it necessary to reduce it to a trial and error technique with the noise level limiting the achievable optimality (Ahlgren and Stevens, 1966). If replications of measurements are averaged better results are obtained (Luecke, 1970) at the expense of slowing the procedure.

4. Optimizing control using dynamic model identification: Desiring speed on one hand and accuracy on the other the most promising approach appears to be the recursive identification of a dynamic model of the system as proposed by Bamberger and Isermann (1978). Instead of using only the steady-state information, parameters in a simple dynamic input-output model are estimated on the basis of the transient response. In a second step the steady-state version of the model is used to determine how the manipulated variables should be varied in order to improve the economic performance. Then the procedure is repeated with a process identification at the new operating point followed by an optimization step and so on until the economically optimal point is reached. The appealing features are evident:

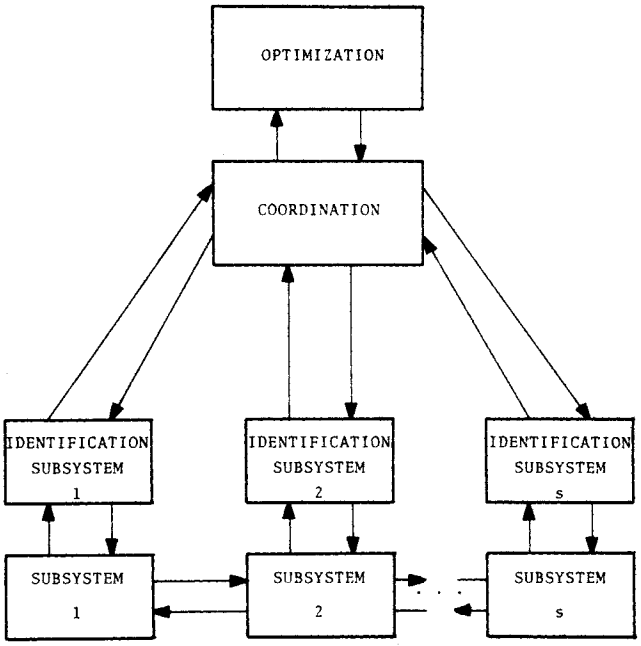

Figure 1. Structure of hierarchical optimizing controller for interconnected systems; centralized optimization with decentralized identification.

- Periodic updating of the parameters through experimentation provides the adaptive element with respect to disturbance change that we seek.
- The need to wait for steady-state is ruled out since a dynamic model is identified each time.
- Most identification schemes have been shown to work well with low signal to noise ratios in industrial applications (Gustavsson, 1972) allowing one to reduce the undesired plant upsets to a minimum. Moreover, extension to multivariate systems can be made without much difficulty.

NEW OPTIMIZING CONTROL TECHNIQUE

Our goal is the development of an on-line optimizing control method for structured processing systems consisting of interactive subsystems. An example would be a refinery consisting of several subsections run by different groups of operating personnel. The optimizing control scheme should find the optimal operating conditions for the integrated system while taking maximum advantage of the subsection structure. Because of the advantages discussed above the optimizing control will proceed via the identification of a dynamic model.

It is natural to carry out the identification in a decentralized fashion and the optimization in a centralized fashion as shown in Figure 1. With the necessary coordination each subsystem is identified separately. Then, the optimizer is supplied with the subsystem models and chooses changes in the inputs to the system such that the overall economic objective is improved.

Model Description

Since the optimization routine to be applied requires only local information, namely gradients, a linear approximation of the general nonlinear plant at an operating point is sufficient for our purposes. Each subsystem will be described by the discrete time input-output model

$$A^i(q)y^i(k) = q^{\tau_i}(B^i(q)m^i(k) + C^i(q)u^i(k)\} + v^i(k) \quad i = 1, \ldots, s \tag{1}$$

where $y^i \in R^{r_i}$ is the measured subsystem output, $m^i \in R^{l_i}$ the noise free optimization inputs and u^i are the noise corrupted interconnection inputs coming from neighboring subsystems

$$u = Ly \tag{2}$$

with L an interconnection matrix. $A^i(q)$, $B^i(q)$ and $C(q)$ are polynomial matrices in the backward shift operator q;

$$qy(k+1) = y(k)$$

whose order and structure depend on the individual subsystem state dimension and observability properties (Valis, 1970; Rowe, 1970). The vector of stationary correlated noise $v^i(k)$ combines the effects of higher order modelling error terms and measurement noise. Due to the interconnections it will in general be correlated with other subsystem noise vectors $v_j (j \neq i)$.

It will be assumed that the orders of the polynomial matrices as well as the multiple delays represented by τ_i are obtained previous to the identification either from past knowledge of the system or from preliminary experiments. It is generally accepted that second order—deadtime models describe chemical processes with sufficient accuracy for control purposes. Thus experimentation to determine the order is usually not justified. The identification method to be derived will estimate the parameters in the polynomials $A^i(q)$, $B^i(q)$ and $C^i(q)$ presuming that y^i and u^i are measured with m^i known noise free. Then, the steady-state parts of these locally valid dynamic models are easily obtained by setting $q = 1$,

$$A^i y^i = B^i m^i + C^i u^i \quad i = 1, \ldots, s \tag{3}$$

with

$$u = Ly$$

or in integrated form to be used by the central optimizer:

$$Ay = Bm \tag{4}$$

Optimization Schemes

The general static optimization problem to be considered is

$$\min_m \quad P(y(m), m) \tag{5}$$

where we assume the system equations to be expressed implicitly through $y(m)$. The case when m is subjected to additional inequality constraints which might become active during the search, falls into the category of constraint control and will be considered in Part II.

A gradient search algorithm is employed to solve Eq. 5. Using subscripts to denote iteration numbers the operating points m_l are updated as follows,

$$m_{l+1} = m_l - \mu S_l \nabla_m P|_l \tag{6}$$

where μ is a fixed stepsize, S_l a positive definite matrix and $\nabla_m P|_l$ is the gradient of the objective function at m_l,

$$\nabla_m^T P(y,m)|_l = \frac{\partial}{\partial m^T} P(y,m)\Big|_l + \frac{\partial}{\partial y^T} P(y,m)\Big|_l \left(\frac{dy}{dm^T}\right)_l. \tag{7}$$

Note that the Jacobian of the plant equations is obtained from the static form of the locally valid identified model for the overall process and computed from Eq. 4 as

$$\left(\frac{dy}{dm^T}\right)_l = (A^{-1}B)_l. \tag{8}$$

Depending on the choice of S_l algorithms with diverse convergence characteristics can be obtained. If the inverse Hessian at m_l is used the Newton search results. Alternatively, an approximation to the inverse Hessian by first order terms yields a Quasi-Newton algorithm while the familiar steepest descent follows by setting $S_l = I$, the identity matrix. The Quasi-Newton techniques were found to be very noise sensitive and to perform worse than steepest descent in our simulation runs and therefore only steepest descent was employed.

Since the gradients for our algorithm are computed as easily as function values, the usual one dimensional minimization along the search direction to determine the stepsize is not needed. It has been shown that for a fixed μ within a certain interval determined by the properties of the Hessian of $P(y(m), m)$ convergence is obtained (Polak, 1971).

We finish this discussion with three important remarks:

1) Since we never await the steady-state after a change in a manipulated variable, the process always lags behind the search and due to nonlinearities $\nabla_m P|_l$ does not correspond to m_l.

2) Search directions are determined by using only gradient values, thus from the optimization point of view the identification of a more complex model (e.g., involving quadratic terms) is not justified.

3) The estimates used in the evaluation of the Jacobian are corrupted with an estimation error. Thus the search direction will vary within a certain confidence interval and will slow down the convergence of the algorithm.

We now present the parameter identification scheme for composite systems which is the backbone of the optimizing controller described.

Identification of Interconnected Systems

For a method to be useful within the optimization framework described in the previous section it should:

- Converge relatively fast
- Provide unbiased estimates
- Deal with time varying parameters
- Handle small signal to noise ratios
- Be transformable into a simple recursive form.

Extensive studies can be found in the literature on the estimation of parameters for discrete linear models. The survey paper by Åstrom and Eykhoff (1971) contains a comprehensive list of references on both theory and implementation. Also, Eykhoff's book (1974) is a good source for methods and applications. Comparative studies have been reported (Isermann et al., 1974; Söderström et al., 1978) which evaluate the characteristics of the different techniques.

Parameter estimation of a dynamic model via least squares (*LS*) is the simplest scheme available. Its handy recursive form makes it very appealing for on-line implementation, but it is well known to yield biased estimates when the noise is correlated, and hence becomes inconvenient for our purposes. From among the several techniques directed to correct this flaw, which do not require additional noise identification and retain the *LS* structure, we have chosen to use the instrumental variables (*IV*) approach.

Although it can be shown to be a non-efficient estimator, observed parameter variances in simulation studies of *IV* are not significantly different from their minimum achievable values, the Cramer-Rao lower bounds (Rowe, 1970) and we only foresee mild effects on the sensitivity of the search.

We will limit the following discussion to SISO systems for simplicity in notation. The scheme for interconnected systems will be presented afterwards.

IV Method for SISO. For $N + p$ measurements modelled by:

$$A(q)y(k) = B(q)m(k) + v(k) \tag{9}$$

Define

$$Y_N(k) = (y(k) \quad y(k+1) \ldots y(N+k-1))^T$$

$$V_N(k) = (v(k) \quad v(k+1) \ldots v(N+k-1))^T$$

$$M_N(k) = (m(k)m(k+1) \ldots m(N+k-1))^T$$

$$\Phi_N(y,m) = (Y_N(0) \quad Y_N(-1) \ldots Y_N(1-p)\ M_N(0) \ldots M_N(1-p))$$

$$\theta = (a_1 \ldots a_p \quad b_1 \ldots b_p)^T \tag{10}$$

Then Eq. 9 is equivalent to

$$Y_N(1) = \Phi_N(y,m)\theta + V_N(1) \tag{11}$$

We multiply on the left by a $N \times 2p$ matrix Z_N to be defined later and solve for θ

$$\theta = \hat{\theta}_N - (Z_N^T\Phi_N(y,m))^{-1}Z_N^T V_N(1) \tag{12}$$

where $\hat{\theta}_N$ is the estimate of θ:

$$\hat{\theta}_N = (Z_N^T\Phi_N(y,m))^{-1}Z_N^T Y_N(1) \tag{13}$$

Theorem. (Wong and Polak, 1967): $\hat{\theta}_N$ is a consistent estimator of θ if the matrix of "instrumental variables" Z_N is selected such that the following limits in probability hold:

$$p \lim_{N\to\infty} \frac{1}{N} Z_N^T V_N(1) = 0 \tag{14}$$

$$p \lim_{N\to\infty} \frac{1}{N} Z_N^T \Phi_N(y,m) = R \text{ (nonsingular)} \tag{15}$$

(Limit in probability is weaker than the usual limit definition.)

Note that for $Z_N = \Phi_N(y,m)$ the instrumental variable method is equal to the least squares method. We emphasize that

$$p \lim_{N\to\infty} \frac{1}{N} \Phi_N^T(y,m)\ V_N(1) \neq 0$$

in general because y and v are correlated. Therefore, *LS* usually yields biased estimates.

Among the different choices for the *IV* matrix the "bootstrap method" has been found to yield good results (Wong and Polak, 1967; Smets, 1970):

$$Z_N = \Phi_N(\hat{x},m)$$

where

$$\hat{x}(k) = \hat{a}_1\hat{x}(k-1) + \ldots + \hat{a}_p\hat{x}(k-p) + \hat{b}_1 m(k-1) + \ldots + \hat{b}_p m(k-p) \tag{16}$$

Thus instead of the measured outputs, the model predicted outputs are used in Z_N where $\hat{A}$, $\hat{B}$ now employ the most recent estimates $\hat{\theta}_N$ which are continuously updated as a result of the identification. Note that the convergence of the bootstrap estimator has not been demonstrated here or elsewhere although it has been shown to work well in simulation results.

Extension to Identification of Interconnected Systems. The main difficulty in applying the *IV* technique directly in the identification of individual subsystems arises from the noisy signal u^i as an input to the model. For simplicity, let us demonstrate this by considering a very simple example. The same argument will apply to general systems as well. If no plant noise is assumed, a 2×2 linear system is described by:

$$\begin{aligned} x_1(k+1) &= a_{11}x_1(k) + a_{12}x_2(k) + b_1 m_1(k) \\ x_2(k+1) &= a_{21}x_1(k) + a_{22}x_2(k) + b_2 m_2(k) \end{aligned} \tag{17}$$

with measurements

$$\begin{aligned} y_1(k) &= x_1(k) + e_1(k) \\ y_2(k) &= x_2(k) + e_2(k) \end{aligned}$$

where $e_1(k)$, $e_2(k)$ are mutually independent white noise sequences of measurement errors

$$\text{i.e., } E[e_1(k)\, e_2(k-l)] = 0 \quad \forall l$$

$$E[e_i(k)\, e_i(k-l)] = \Sigma_i \delta(l), \quad i = 1,2$$

After decomposition in two one-dimensional subsystems, the individual input-output relationships are:

Subsystem 1:

$$y_1(k+1) = a_{11}y_1(k) + b_1 m_1(k) + a_{12}u_1(k) + v_1(k+1) \tag{18}$$

Subsystem 2:

$$y_2(k+1) = a_{22}y_2(k) + b_2 m_2(k) + a_{21}u_2(k) + v_2(k+1)$$

where $u_1(k) = y_2(k)$, $u_2(k) = y_1(k)$. The noise has the following dynamics

$$\begin{aligned} v_1(k+1) &= e_1(k+1) - a_{11}e_1(k) - a_{12}e_2(k) \\ v_2(k+1) &= e_2(k+1) - a_{21}e_1(k) - a_{22}e_2(k) \end{aligned} \tag{19}$$

where we can see the noise correlation among subsystems.

Assuming that the bootstrap estimator is used for the identification of the first subsystem, we examine the following limit:

$$\lim_{N\to\infty} \frac{1}{N} Z_N^{1T} V_N^1(1) = \lim_{N\to\infty} \frac{1}{N} \begin{bmatrix} \sum_{k=1}^{N} \hat{x}_1(k-1)\, v_1(k) \\ \sum_{k=1}^{N} m_1(k-1)\, v_1(k) \\ \sum_{k=1}^{N} u_1(k-1)\, v_1(k) \end{bmatrix} \tag{20}$$

where

$$\hat{x}_1(k+1) = \hat{a}_{11}\hat{x}_1(k) + \hat{b}_1 m_1(k) + \hat{a}_{12}u_1(k)$$

and

$$u_1(k) = a_{22}u_1(k-1) + b_2 m_2(k-1) + a_{21}y_1(k-1) + v_2(k)$$

First we may note that the *IV* $\hat{x}_1(k-1)$ is not noise free and although for this special case it is uncorrelated to $v_1(k)$, for a more general model it would be. Moreover, since m_1 is known noise free the second element in vector (Eq. 20) vanishes asymptotically while the dependence of $u_1(k-1)$ on $v_2(k-1)$ yields a non-zero last element. Specifically, Eq. 20 reduces to:

$$\lim_{N\to\infty} \frac{1}{N} Z_N^{1T} V_N^1(1) = \begin{bmatrix} 0 \\ 0 \\ -a_{12}\Sigma_2 \end{bmatrix} \tag{21}$$

and we observe that the bias is proportional to the interconnection parameter. Hence, condition (Eq. 14) is not satisfied and unbiasedness cannot be ensured.

In the *IV* scheme for the integrated process consistency is achieved by using $\hat{x}_2$ instead of u_1 in the matrix Z_N. But $\hat{x}_2$ is the *IV* generated by the identification algorithm of the second subsystem. Therefore it is natural to expect that the bias can be eliminated by supplying the local estimator with the instrumental variable of the neighboring subsystem. This suggests a coordination structure as indicated in Figure 2. Each individual estimator uses all its available measurements plus information

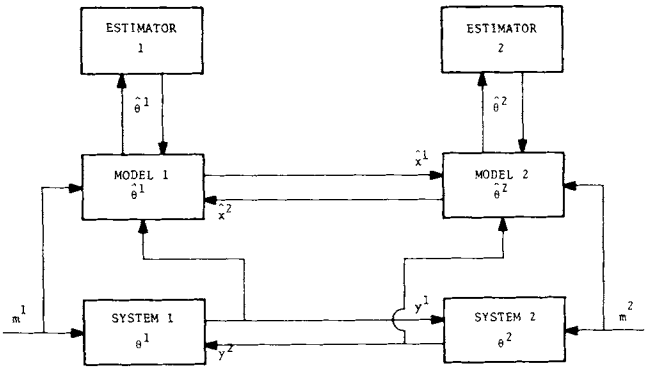

Figure 2. A decomposition approach to the *IV* identification of large interconnected systems.

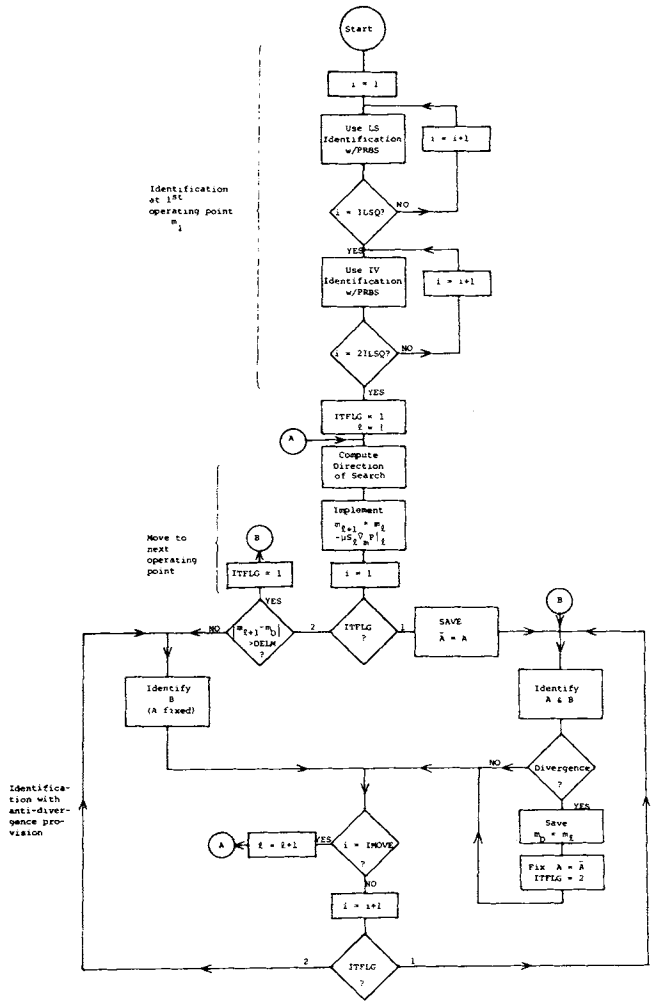

Figure 3. Flowchart of the new optimizing control technique.

from the coordinator, whose function is basically the transfer of instrumental variables.

The analysis can be extended to any number of subsystems of arbitrary dimension and order since the problem structure is basically the same. For the development of the *IV* procedure as applied to multivariable systems the reader is referred to the paper by Rowe (1970).

The decomposition is not without tradeoffs. It requires the "main" interconnection variables to be measured thus possibly increasing the expenditures for instruments. The establishment of quantitative decomposition criteria along the ideas put forward by Morari et al. (1980) will be a topic for future investigation.

SUMMARY OF NEW TECHNIQUE

After specifying both the optimization and identification schemes to be used by the controller, we proceed to discuss the proper selection of tuning parameters, and measures to prevent divergence of the *IV* algorithm. Figure 3 illustrates the complete algorithm.

Tuning Parameters

As described below, there are three tuning parameters. In most circumstances a functional relationship can be established *a priori* between two of them, further simplifying the implementation.

1. Search stepsize (μ): As previously mentioned the upper bound on the allowed fixed optimization algorithm stepsize depends on the particular response surface and has to be found on-line in principle. However, from our limited experience it appears that the scheme is not very sensitive to its choice unless an unreasonably large value is employed.

2. Samples between imput moves (IMOVE): It is desirable to keep IMOVE as small as possible in order to move rapidly toward the optimum. However, we are limited by the fact that after each input change enough time has to elapse for the dynamic system to move and to allow an adaptation of the parameters. Thus high noise levels or long system response times will require a large value of IMOVE.

3. Forgetting factor (λ): In order for the *IV* estimation method to be implementable on-line, the "one-shot" estimation Eq. 13 must be expressed in a recursive form so as to update the parameters as new samples are available. For a SISO model the estimate at sampling time $N+1$ is obtained from the estimate at time N as follows (Goodwin and Payne, 1977):

$$\hat{\theta}_{N+1} = \hat{\theta}_N + K_{N+1}(y(N+1) - \phi^T_{N+1}\hat{\theta}_N)$$
$$K_{N+1} = P_{N+1}\, z_{N+1}$$
$$P_{N+1} = [P_N - P_N z_{N+1}[\lambda + \phi^T_{N+1}P_N z_{N+1}]^{-1}\, \phi^T_{N+1}P_N]/\lambda \tag{22}$$

and

$$P_N = (Z^T_N \Phi_N)^{-1}$$

where z_{N+1}, ϕ_{N+1} are the vectors containing the latest samples and $0 < \lambda \leqq 1$ is the forgetting factor which is included to make the estimator track time varying parameters. Intuitively, λ can be interpreted as the fraction of past data used in the present estimation, and any past data point $j < N$ contributes as a fraction λ^{N-j+1} to the estimate at $N+1$. Equal weighting is given to all samples when $\lambda = 1$.

The following remarks result from our own experience with simulation cases and other published results (e.g., Smets, 1970; Eykhoff, 1974).

(i) $\lambda < 1$ always increases the variance of the estimates

(ii) For a fixed noise level, as λ is decreased, faster parameter convergence is achieved.

(iii) For fixed λ, higher noise levels slow down the convergence.

λ could be chosen independent of the other tuning parameters according to these considerations but it is reasonable to presume some dependence on IMOVE. After IMOVE samples the parameter estimates should be independent of the measurements taken before the last input variation. Thus we suggest to choose λ such that IMOVE is approximately twice the asymptotic memory length ($1/1-\lambda$) of the estimator (Clarke and Gawthrop, 1975)

$$\lambda \approx 1 - \frac{2}{\text{IMOVE}} \tag{23}$$

This rule worked very well for all our example runs.

Divergence of the Recursive Algorithm

It is certain that if the information matrix $Z^T_N\Phi_N$ is singular the estimate $\hat{\theta}_N$ will diverge. Let us examine how it may occur. For simplicity, consider the MISO system:

$$y(k) = \sum_{i=1}^{p} a_i y(k-i) + \sum_{i=1}^{p} b^1_i m_1(k-i) \ldots \sum_{i=1}^{p} b^m_i m_m(k-i) \tag{24}$$

According to the analysis in Söderström (1974) the following equivalence holds for the noise free bootstrap estimator:

$$Z^T_N\Phi_N = \hat{D}MD^T \tag{25}$$

where M is a square matrix of delayed inputs which is positive

definite if the $m_j(k)$ sequences are persistently exciting, D is a matrix of true system parameters for the model above and $\hat{D}$ is a matrix of same structure but for the model parameters. One can show that positive definiteness of M and D, $\hat{D}$ of full rank are sufficient for nonsingularity of the information matrix.

It is known (Woodside, 1971) that for LS estimation the information matrix is always of full rank when the observations are noisy. Though it is not possible to conclude rigorously that the same will be true for IV (the information matrix is nonsymmetric and depends on the parameter estimates) no divergence difficulties were encountered in the simulation examples when sufficient noise was added to the measurements. Therefore divergence is unlikely to occur in real life situations. Nevertheless we will analyze the reasons and provide counter measures.

1. Input sequence not persistently exciting: Since it is assumed that sufficient excitation is provided by the input moves, persistent perturbations of limited amplitude such as PRBS are recommended only at the start-up until fair estimates are obtained. Also, diminishing excitation upon convergence can be avoided by choosing a larger stepsize μ which keeps the OC varying within a small neighborhood of the optimum due to inaccuracies in the estimation.

2. Rank deficiency of $\hat{D}$ or D: Matrices $\hat{D}$ or D may become rank deficient if some parameters in the model and/or in the real system vanish. Consider the case when the output y in model (Eq. 24) is optimized. At the steady-state optimum the following must occur:

$$\sum_{i=1}^{p} b_i^j = 0 \quad j = 1, \ldots, m$$

Matrices $\hat{D}$ and D exhibit the following structure:

$$D = \left[\begin{array}{ccc|c|ccc} 0 & b_1^1 & b_2^1 \ldots b_p^1 & & 0 & b_1^m \ldots & b_p^m \\ 0 & 0 & b_1^1 b_2^1 \ldots b_p^1 & 0 \ldots & 0 & & \\ & 0 & b_1^1 b_2^1 b_p^1 & & & b_1^m \ldots & b_p^m \\ \hline & & Q_1 & & & Q_m & \end{array}\right]$$

where $Q_1 \ldots Q_m$ are full rank matrices of the coefficients a_i^j. For the simple case $p = m = 1$,

$$D = \begin{bmatrix} 0 & b_1 \\ 1 & a_1 \end{bmatrix}$$

and trivially, $b_1 = 0 \Rightarrow$ rank $D < 2$. Higher order models require all coefficients b_i^j to vanish for rank deficiency. The same situation is expected to occur for MIMO systems.

For the recursive algorithm (Eq. 22) it can be demonstrated by induction that

$$P_N = (Z_N^T \Lambda_N \Phi_N + \lambda^N \epsilon\, I)^{-1} \tag{26}$$

where

$$P_0 = \frac{1}{\epsilon} I$$

and

$$\Lambda_N = \text{diag}(\lambda^{N-1} \lambda^{N-2} \ldots \lambda\ 1).$$

Consequently, for $\lambda = 1$, by choosing ϵ sufficiently large the recursion can be made stable even when the "one-shot" computation, i.e., $Z_N^T \Phi_N$ becomes singular, is expected to diverge. However, we require $\lambda < 1$ for the algorithm to track time varying parameters implying that for some large N, P_N becomes singular for any ϵ.

In order to resolve the conflict between speed of adaptation and preventing divergence at the optimum we found the following modification helpful. Since the autoregressive part of the model reflects mainly the dynamics of the system, as the search converges, the parameters in $A(q)$ remain constant, while the magnitude and direction of the gradient is mainly determined by $B(q)$. It can easily be shown that by fixing $A(q)$ and estimating only $B(q)$ close to the optimum, the resulting information matrix is symmetric and of full rank, yielding a stable recursive estimator. The model to be identified is then

$$\bar{y}(k+1) = A(q)y(k+1) = B(q)m(k+1)$$

where $\bar{A}(q)$ contains the best estimated values before divergence.

By continuously identifying B, a change in plant conditions can be detected and the plant driven to a new optimum. Once it moves away, more than a certain pre-specified distance (DELM) the original scheme is resumed to regain speed.

EXAMPLE

A simple reaction system was simulated on the digital computer in order to demonstrate the basic features of the proposed OC. Two CSTR's in series in which an exothermic autocatalytic reaction is taking place are shown in Figure 4. The units interact in both directions due to the recycle of a fraction α of the product stream into the first reactor. It is assumed that a servo controller has been set up such that the individual reactor temperatures can be specified the OC. T_1 and T_2 are assumed to be known noise free. The reaction is

$$A + B \underset{k_-}{\overset{k_+}{\rightleftharpoons}} 2B$$

with rate $= k_+ C_A C_B - k_- C_B^2$ ($k_\pm = A_\pm \exp(-E_\pm/RT)$). The physical and design constants of the system are

$$\alpha = 0.5$$

Tank 1 mean residence time = 30 min.

Tank 2 mean residence time = 25 min.

$E_+/R = 17786.$ $E_-/R = 23523°K$

$A_+ = 9.73 \times 10^{22}$ $A_- = 3.1 \times 10^{30}$ l/gmol-s

resulting in an overall time constant of approximately 40 min.

We desire to find the temperatures T_1 and T_2 which maximize C_{B2} the concentration of B at the second reactor. The important disturbance for optimization is C_{A0}, the feed concentration of A which has a significant effect on the optimum $\hat{C}^*_{B2}$ as shown in the following table:

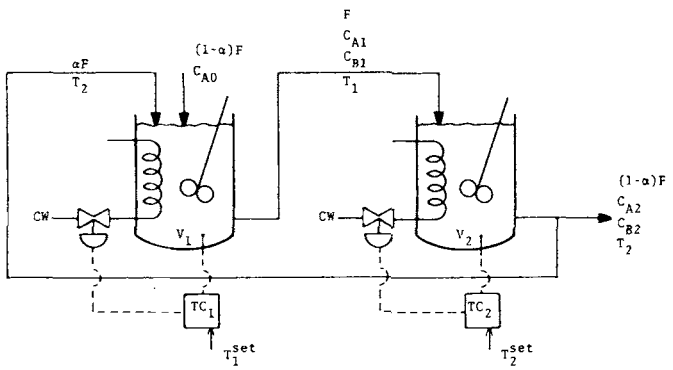

Figure 4. Two CSTR system used in the example.

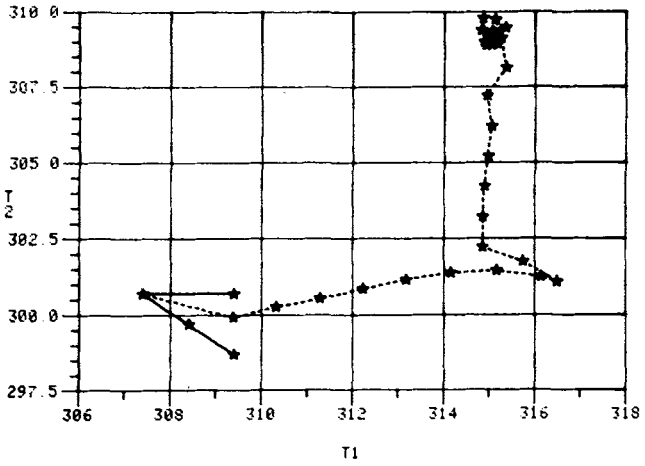

Figure 5. Sequence of set point changes for the noise free search (μ = 5000, IMOVE = 15, λ = 0.9).

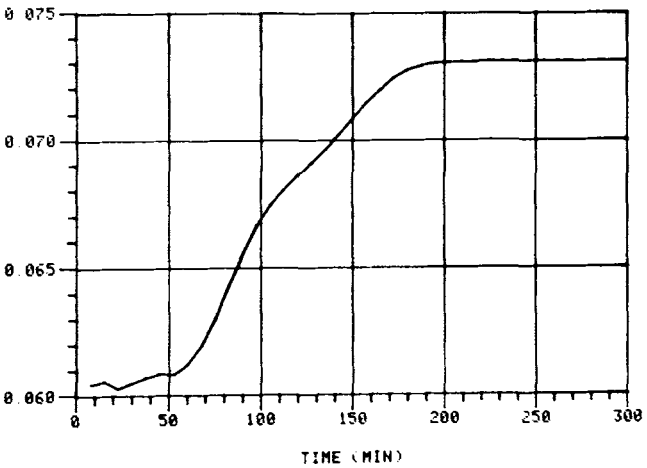

Figure 6. Record of the objective C_{B2} during the search for the optimum shown in Figure 5.

$C_{A0}\left(\frac{\text{gmol}}{\text{l}}\right)$	$C^*_{B2}\left(\frac{\text{gmol}}{\text{l}}\right)$	T^*_1(°K)	T^*_2(°K)
0.1	0.0731	315.0	309.0
0.06	0.0420	317.0	312.0
0.20	0.1536	312.0	306.0

The proposed large scale systems approach was used where each reactor is considered as a subsystem, modelled as a linear first order system:

Reactor 1:

$$x_1(k+1) = a_{11}x_1(k) + a_{12}u_1(k) + b_1m_1(k) + d_1$$

Reactor 2:

$$x_2(k+1) = a_{22}x_2(k) + a_{21}u_2(k) + b_2m_2(k) + d_2$$

where

$$x_1 = u_2 = C_{B1} \qquad x_2 = u_1 = C_{B2}$$
$$m_1 = T_1 \qquad m_2 = T_2$$

and the parameters d_1, d_2 are included to account for the nonzero steady-state.

For all the runs a sampling time of 0.5 minutes is used. Short sampling periods are suggested since the convergence rate of the identification actually depends on the amount of data processed per unit time. Due to the generally imprecise values of the parameters at the beginning of the search, a maximum bound on the input steps (STMX = 1 degree) is specified. As suggested by Smets (1970) in case no a priori information about parameter values is at hand the *LS* algorithm is run during the first 40 iterations (ILSQ = 40) and the search is started after 80. During this initial period, PRBS of 1° amplitude in each temperature are employed. All the following runs start at (T_1, T_2) = (308, 299)°K for C_{A0} = 0.1.

The Deterministic Case

In Figures 5 and 6 the noise free case for λ = 0.90, IMOVE = 15 and μ = 5000 is shown. Convergence is achieved in about four system time constants which is not much longer than the time required for a system to reach a new steady-state after a step change. Also, divergence was avoided by stopping the identification of the autoregressive part of the model close to the optimum as described above.

Effect of Noise

If the output signals C_{B1} and C_{B2} are corrupted with severe measurement noise of variance $\sigma^2 = 9 \times 10^{-6}$, excellent performance of the *OC* is observed (Figure 7). The need to set the tuning parameters to λ = 0.97, IMOVE = 40 brought an increase in the search time. Nevertheless, no parameter divergence was found to occur for this noise level, showing the robustness of the proposed *OC* to measurement errors. At this noise level the *OC* is unable to find the optimum when *LS* identification is used showing the detrimental effect of estimation bias on the search (Figure 8).

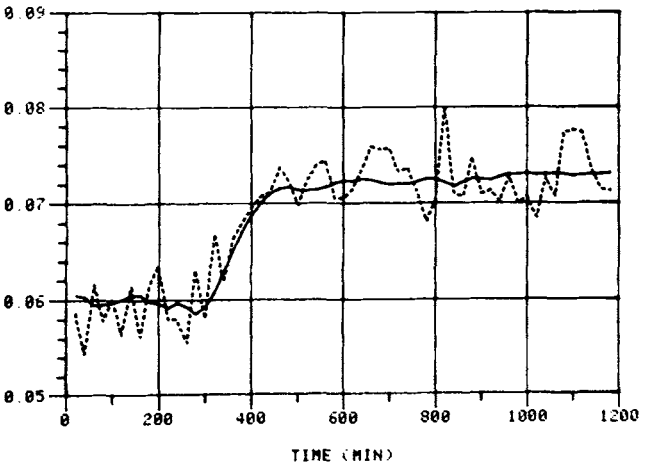

Figure 7. Search for the optimum under severe measurement noise ($\sigma^2 = 9 \times 10^{-6}$, μ = 7500, IMOVE = 40, λ = 0.97); ——— actual objective, ---- measured objective.

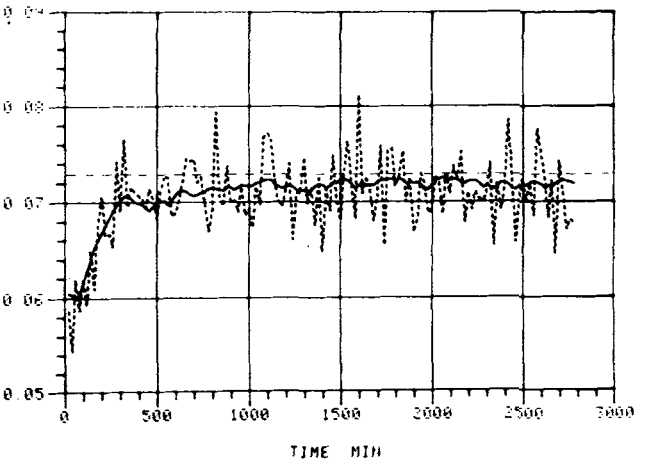

Figure 8. The *OC* using *LS* identification is unable to find the optimum for same parameters and noise level as in Figure 7.

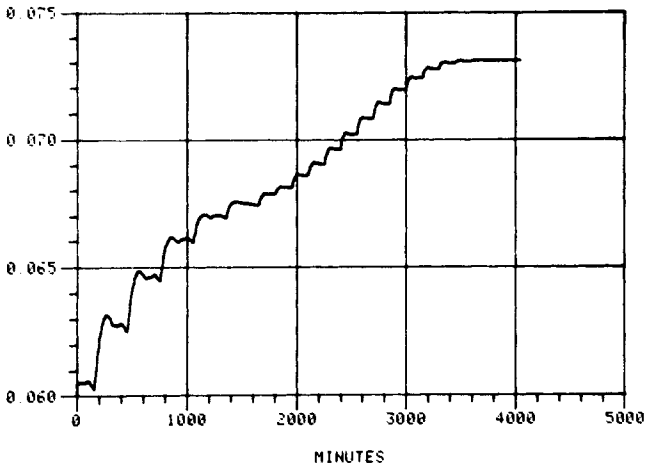

Figure 9. Noise free direct gradient search (IMOVE = 100).

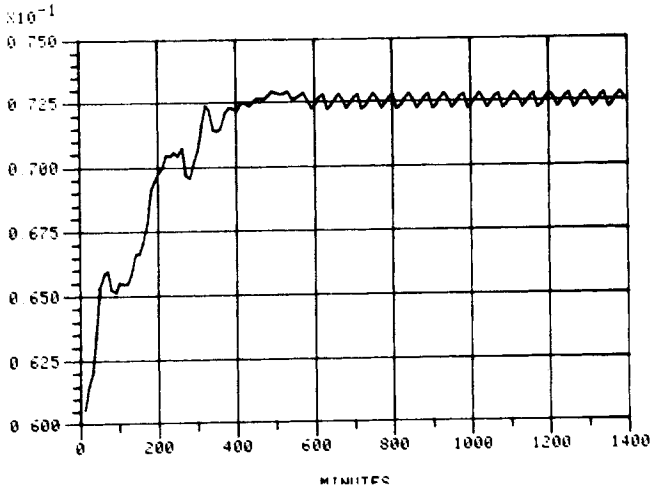

Figure 10. Results for a SIMPLEX optimizing controller (IMOVE = 15).

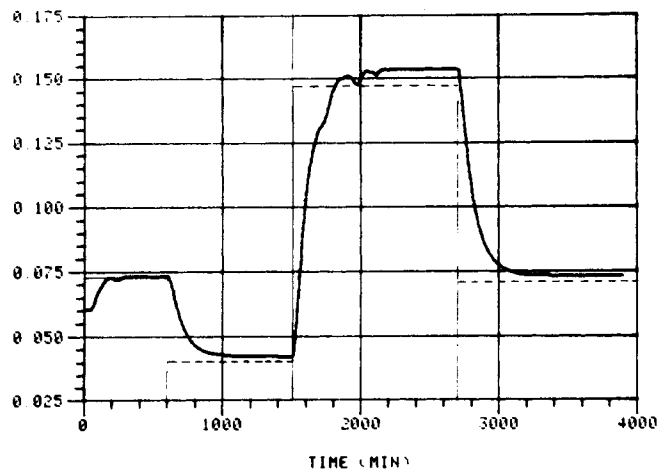

Figure 11. Tracking of a drifting optimum (---- steady-state optimum without optimizing controller).

Direct Search vs. New OC

To make a fair comparison the same time between setpoint changes as in the above identification runs should be allowed. It was found, however, that direct search on the noise-free system with three points used to compute the gradient did not converge for values of IMOVE less than 100 (Figure 9), thus it took over 15 times as long to reach the optimum as compared to the new *OC*. The SIMPLEX pattern search appears less sensitive and finds the optimum when IMOVE = 15 although the oscillations about the solution are quite undesirable (Figure 10). Even for the noise-free case these classical methods are quite slow and further simulation results showed the speed of convergence and stability properties to deteriorate significantly when noise is added.

Tracking of Optimum Changes

The real test for our *OC* is to observe its ability to follow plant upsets. To demonstrate this, a sequence of step changes in C_{A0} was imposed on the system after achieving convergence at the optimum for $C_{A0} = 0.10$, as follows:

C_{A0}(gmol/l)	TIME (min)
0.06	600
0.20	1500
0.10	2700

and the corresponding optima have been given previously. Fig. 11 demonstrates the ability of the *OC* to rapidly adapt the system operating conditions to important disturbance changes. Without *OC* the loss in performance would have been approximately 4%.

ACKNOWLEDGMENT

Support of this research through NSF grant ENG-7906353 is gratefully acknowledged.

NOTATION

A,B,C = polynomial matrices in linear dynamic models
a,b = parameters in polynomial matrices
$D,\hat{D}$ = matrices of parameters defined in Eq. 25
K_N = recursive *IV* algorithm gain vector
L = interconnection matrix
M_N = vector of input variables
m = manipulated inputs
P = economic objective function
P_N = recursive *IV* algorithm "covariance" matrix
S_l = positive definite matrix in optimization algorithm
s = total number of subsystems
u = interconnection inputs
V_N = vector of errors
v = correlated error
x = state variable
Y_N = vector of outputs
y = measured outputs
Z_N = *IV* matrix
∇_m = gradient operator w.r.t. m

Greek Letters

θ = vector of system parameters
Λ_N = diagonal matrix defined in Eq. 26
λ = forgetting factor
μ = optimization algorithm stepsize
τ = time delay
Φ_N = regressor matrix in *IV* algorithm

Subscripts

l = optimization iteration number
N = identification iteration number

Superscripts

$\hat{}$ = estimated value

LITERATURE CITED

Ahlgren, T. D., and W. F. Stevens, "Process Optimization in the Presence of Error," *IEC Proc. Des. & Dev.*, **5**, 290 (1966).

Åström, K. J., and P. Eykhoff, "System Identification—A Survey," *Automatica*, **7**, 123 (1971).

Bamberger, W., and R. Isermann, "Adaptive On-Line Steady State Optimization of Slow Dynamic Processes," *Automatica*, **14**, 223 (1978).

Box, G. E. P., and J. Chanmugam, "Adaptive Optimization of Continuous Processes," *IEC Fund.*, **1**, 2 (1962).

Box, G. E. P., and N. R. Draper, *Evolutionary Operation*, Wiley, New York (1969).

Clarke, D. W., and P. J. Gawthrop, "Self Tuning Controller," *Proc. IEE*, **122** (9), 929 (1975).

Davis, T. A., D. E. Griffin, and P. U. Webb, "Cat Cracker Optimization and Control," *Chem. Eng. Prog.*, **70**, (11), 53 (1974).

Draper, C. S., and Y. T. Li, "Principles of Optimizing Control Systems and an Application to the Internal Combustion Engine," American Society of Mechanical Engineers (Sept., 1951).

Duncanson, L. A., and P. V. Youle, "On-Line Control of Olefin Plant," *Chem. and Proc. Eng.*, **51**, 49 (May, 1970).

Edler, J., P. N. Nikiforuk, and E. B. Tinker, "A Comparison of the Performance of Techniques for Direct, On-Line Optimization," *The Canadian J. of Chem. Eng.*, **48**, 432 (1970).

Eykhoff, P., *System Identification*, Wiley, London (1974).

Goodwin, G. C., and R. L. Payne, *Dynamic System Identification: Experiment Design and Data Analysis*, Academic Press, New York (1977).

Gustavsson, I., "Comparison of Different Methods for Identification of Industrial Processes," *Automatica*, **8**, 127 (1972).

Isermann, R., U. Baur, W. Bamberger, P. Kneppo, and H. Siebert, "Comparison of Six On-Line Identification and Parameter Estimation Methods," *Automatica*, **10**, 81 (1974).

Kotnour, K. D., G. E. P. Box, and R. J. Altpeter, "A Discrete Predictor-Controller Applied to Sinusoidal Perturbation Adaptive Optimization," *ISA Trans.* **5**, 255 (1966).

Luecke, R. H., "Search for the Optima of Stochastic Functions," *Int. J. Control*, **12**, 609 (1970).

Morari, M., Y. Arkun, G. Stephanopoulos, "Studies in the Synthesis of Control Structures for Chemical Processes, Part I," *AIChEJ.*, **26**, 220 (1980).

Polak, E., Computational Methods in Optimization: A Unified Approach, Academic Press, New York (1971).

Prett, D. M., and R. D. Gillette, "Optimization and Constrained Multivariable Control of a Catalytic Cracking Unit," Paper No. **51C**, AIChE 86th National Meeting (April, 1979).

Rowe, I. H., "A Bootstrap Method for the Statistical Estimation of Model Parameters," *Int. J. Control*, **12**, 721 (1970).

Saridis, G. N., "Stochastic Approximation Methods for Identification and Control—A Survey," *IEEE AC-19*, 798 (1974).

Sawaragi, Y., T. Takamatsu, K. Fukunaga, E. Nakanishi, and H. Tamura, "Dynamic Version of Steady State Optimizing Control of a Distillation Column by Trial Method," *Automatica*, **7**, 509 (1971).

Shah, M. J., and R. F. Stillman, "Computer Control and Optimization of a Large Methanol Plant," **I/EC 62** (12), 59 (1970).

Smets, A. J., "The Instrumental Variable Method and Related Identification Schemes," Eindhoven University of Technology, Dept. of Elect. Eng., TH-Report 70-E-15 (1970).

Söderström, T., "Convergence of Identification Methods Based on the Instrumental Variable Approach," *Automatica*, **10**, 685 (1974).

Söderström, T., L. Ljung, and I. Gustavsson, "A Theoretical Analysis of Recursive Identification Methods," *Automatica*, **14**, 231 (1978).

Valis, J., "On-Line Identification of Multivariable Linear Systems of Unknown Structure from Input Output Data," IFAC Symp. on Identification and Process Parameter Estimation, Prague, paper 1.5 (1970).

Webb, P. V., B. E. Lutter, and R. L. Hair, "Dynamic Optimization of Fluid Cat. Crackers," *Chem. Eng. Prog.*, **74** (6), 72 (1978).

White, J. W., G. J. Foley, and R. J. Altpeter, "Stability of Input Perturbation Extremum-Seeking Systems," *IEC Fund.*, **7**,296 (1968).

Wong, K. Y., and E. Polak, "Identification of Linear Discrete Time Systems Using the Instrumental Variable Method," *IEEE Trans. Autom. Contr.*, **AC-12**, 707 (1967).

Woodside, C. M., "Estimation of the Order of Linear Systems," *Automatica*, **7**, 727 (1971).

Manuscript received June 17, 1980; revision received December 17, and accepted December 23, 1980.

Numerical Solution of Nonequilibrium Multicomponent Mass Transfer Operations

D. S. BILLINGSLEY

and

ARUN CHIRACHAVALA

IBM Corp.
Palo Alto Scientific Center
1530 Page Mill Road
Palo Alto, CA 94304

An improved model and solution procedure for nonequilibrium, multicomponent, two-phase, steady-state mass transfer problems are developed. Processes as diverse as packed column operation and the drying of porous solids are included. Ideas such as "HETP," "HTU" or "mass transfer section" are avoided. "Stiff" systems appear to be handled well.

SCOPE

This paper proposes a model and solution procedure for steady state continuous contact mass transfer problems as diverse as nonequilibrium operation of packed columns and the drying of porous media. Included are all the features listed as items (a) through (g). Such a combination represents an advance over previous work.

(a) Energy, mass and momentum are each required to be conserved in each phase.

(b) Equations of state are nontrivial in each phase.

(c) A multicomponent system is considered.

(d) Phase temperatures, velocities and concentrations are not assumed constant.

(e) Transfer coefficients for mass, energy or momentum are assumed neither constant nor equal for all components.

(f) One phase at some place in the unit is not assumed to be in equilibrium with the other phase at some other place in the unit.

(g) The problem of obtaining a numerical solution is addressed.

The model is sufficiently detailed and realistic to be representative of the best currently available.

Correspondence concerning this paper should be addressed to D. S. Billingsley.
Present address of Arun Chirachavala: 77/8 Soi Juladit, Petchburi Rd., Bangkok, Thailand.

0001-1541-81-2727-0968-$2.00.