PROCEEDINGS OF THE

Volume I

1980

Joint Automatic Control Conference

An ASME Century 2 Emerging Technology Conference

August 13-15, San Francisco, California

Sponsoring Organizations
American Institute of Chemical Engineers
American Society of Mechanical Engineers
Institute of Electrical and Electronics Engineers
Instrument Society of America

Participating Organizations
American Institute of Aeronautics and Astronautics
Society of Manufacturing Engineers

Published on behalf of the American Automatic Control Council

IEEE: 80CH1580-0

AN IDENTIFICATION APPROACH TO OPTIMIZING CONTROL OF INTERCONNECTED PROCESSING SYSTEMS

Manfred Morari

Carlos E. Garcia
Chemical Engineering Department
University of Wisconsin
Madison, Wisconsin

ABSTRACT

The economically optimal setpoint of a process will shift when unknown disturbances in the form of changes of the feed properties or the operating characteristics and intentional changes of the production specifications occur. Currently utilized techniques for optimizing control are either model based and suffer consequently from the unavoidable model inaccuracies or use a steady state gradient search on the real system which makes them very slow. A new method is presented where a linear dynamic system model is continuously updated on line using the instrumental variable technique and forms the basis of the optimizing control strategy. The complexity of the algorithm is reduced by a new decomposition approach suited to the natural structure of industrial plants (interconnected subsystems).

INTRODUCTION

Increasing energy and raw material 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 in 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 optimum despite changing environmental conditions it is possible to achieve an important performance improvement.

Converter catalyst deactivation, heat exchanger equipment fouling, ambient temperature drifts and change 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.

REQUIREMENTS OF AN OPTIMIZING CONTROLLER

To be applicable in an industrial environment in a non-trivial situation any optimizing control scheme has to satisfy the following requirements:

1. It should operate on-line:
   i.e. it should not be model dependent but determine the best operating conditions through direct experiments on the plant.

2. It should be fast:
   An optimizing controller whose adaptation time constant is slower than or comparable to the period of important process disturbances will be continuously hunting after an optimum without ever reaching it.

3. It should be able to deal with noisy measurements:
   The desire to keep the required process perturbations at a minimum to avoid unnecessary plant upsets results in small signal to noise ratios.

4. It should be able to deal with multivariate, large scale systems:
   Given the great complexity, large dimensionality and inherent interaction between units that characterize modern chemical plants, a unit by unit optimizing controller will prove insufficient for our purposes. Hence, adequate handling of multivariable systems while at the same time providing for division and coordination of tasks is a mandatory requirement of an industrially successful scheme.

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

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 inaccurate (Webb, et al., 1978, Prett & Gillette, 1979).

On-line Methods
1. Methods using a continuously varying perturbation signal: Probably the oldest method is due to Draper and Li (1951) using a ramp shaped perturbation signal to determine the gradient of the objective function. More recent techniques use sinusoidal inputs (Box and Chanmugan, 1962; Kotnour et al. 1966). These methods are generally noise sensitive and essentially limited to one dimensional searches. Extensions even to two dimensions encounter complex stability problems (White, et al., 1968) and no applications have been reported.

2. Direct search methods: Multivariate systems have been handled by implementing either gradient or pattern-type search methods from optimization theory directly on the plant. Being essentially steady-state procedures, measurements should be taken only when the

[handwritten margin note: comment: Draper and Li (1951) also consider sinusoidal perturbations]

process is allowed to settle after each change in the manipulated variables which makes the controller very slow.

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 in an off-line optimization routine which determines how the manipulated variables should be varied 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 the speed and the noise insensitivity typical for most dynamic identification schemes.

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. After the identification step the model is used in an optimization routine. It is natural to carry out the identification in a decentralized and the optimization in a centralized fashion as shown in Fig. 1. With the necessary coordination each subsystem is identified separately. The optimizer is supplied with the subsystem models and chooses changes in the inputs to the system such that the economic objective is improved.

Model Description

Each subsystem; $(i=1,...s)$ will be described by the discrete input-output model

$$A_i(z^{-1})y_i(k) = z^{-\tau_i}\{B_i(z^{-1})m_i(k) + C_i(z^{-1})u_i(k)\}$$
$$+ v_i(k) \qquad (1)$$

where y_i R^{r_i} is the measured subsystem output, m_i R^{m_i} the noise free measured subsystem input and u_i the noise corrupted interconnection input arising from neighboring subsystems. A_i, B_i and C_i are polynomial matrices of order p_i in the backward shift operator z^{-1}. p_i and τ_i are assumed known, v_i is a vector of stationary noise in general correlated with other sub-

system noise vectors v_j, $j \neq i$. We obtain the steady state part of (1) for $z=1$ as

$$A_i y_i = B_i m_i + C_i u_i \qquad i=1,...s \qquad (2)$$

or in integrated form by making the appropriate substitutions for u_i

$$Ay = Bm \qquad (3)$$

Optimization Schemes

In its general form the optimization problem can be stated as

$$\min_{m} P(y(m),m) \qquad (4)$$

where all the equality constraints are included implicitly in $y(m)$ and the absence of inequality constraints is assumed in this work. There are numerous techniques available to solve problem (4). Considering that the model of the process is known through identification the gradient can be computed as easily as the steady-state objective. Therefore gradient methods are prefered over direct or pattern searches since they exhibit faster convergence rates and more efficient algorithms. At iteration i any direction

$$q^i = -S^i \nabla_m P\big|_i \qquad (5)$$

is a direction of descent if S^i is positive definite. Because the current gradient is always available it is best to choose a fixed step size μ such that the new improved operating point is

$$m^{i+1} = m^i + \mu q^i \qquad (6)$$

For $S^i=I$ we have the familiar steepest descent search. A variable metric update for which S^i approximates the inverse Hessian can also be obtained (Avriel, 1976).

Identification of Interconnected Systems

Many reviews of identification techniques for discrete systems have been presented, e.g. Eykhoff and Åstrom (1971) and Eykhoff's book (1974) and comparative studies have evaluated the merits of the different methods (Isermann et al., 1974, Söderstrom et al., 1978). Except in very special circumstances the linear least squares (LS) approach is known to yield biased estimates. The instrumental variable (IV) method corrects this flaw while preserving the simplicity. This was the main reason for choosing it here, though it is known that the estimates are not efficient. However, in simulation studies of IV the observed parameter variances were found not to be significantly different from their minimum achievable values, the Cramer-Rao lower bonds (Rowe, 1970). Assume for simplicity in notation (1) to describe a SISO system (neglecting the subscripts) and define

$$Y_N(k) = (y(k),y(k+1),...y(N+k-1))^T$$

$$V_N(k) = (v(k),v(k+1),...v(N+k-1))^T$$

$$M_N(k) = (m(k),m(k+1),...m(N+k-1))^T$$

$$\Phi_N(y,m) = (Y_N(0),Y_N(-1)...Y_N(1-n_a),M_N(0),...M_N(1-n_b))$$

$$\Theta = (a_1.... a_{n_a}, b_1... b_{n_b})^T$$

Then the system model (1) can be rewritten as

$$Y_{N(1)} = \Phi_N(y,m)\theta + V_{N(1)} \qquad (7)$$

and the IV estimate is

$$\hat{\theta}_N = (Z_N^T \Phi_N(y,m))^{-1} Z_N^T Y_{N(1)} \qquad (8)$$

Wong and Polak (1967) suggest different choices for the IV matrix Z_N in order to guarantee unbiased estimates and Smets (1970) discusses their implementation. An IV selection that reduces estimation variance requires a knowledge of the noise statistics. If this is avoided the "bootstrap method" has been found to yield good results:

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

where $\hat{x}(k) = \hat{a}_1\hat{x}(k-1)+\ldots+\hat{a}_{n_a}\hat{x}(k-n_a)+\hat{b}_1 m(k-1)+\ldots$

$$\ldots+\hat{b}_{n_b} m(k-n_b) \qquad (9)$$

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.

Only a few results on the identification of composite systems have been reported in the literature (e.g. Wismer, et al., 1970, Arafeh & Sage, 1974). In them the estimation problem is converted to a deterministic optimization problem without due regard to the statistical properties of the estimates. In this paper a simple scheme within the IV framework is presented requiring a minimum of coordination effort among subsystems.

The main difficulty in applying the IV technique directly to individual subsystem identification consists in having a noisy signal as an input to the model. Under these circumstances the IV method was shown to yield biased estimated (Garcia, Morari, 1980). The bias can be eliminated by introducing a coordinator which supplies the local estimator with the instrumental variables of the neighboring subsystems. This suggests a coordination structure as indicated in Fig. 2 for the special case of two subsystems. Each individual estimator uses all its available measurements plus the information from the coordinator which is basically a transfer of instrumental variables.

Summary of New Technique

After specifying both the optimization and identification schemes to be used by our controller, we proceed to analyze potential difficulties to be encountered when implemented on a process. Specifically, measures to prevent divergence of the IV algorithm are discussed together with the proper selection of tuning parameters for the procedure.

Tuning parameters.

1. Sampling time (TSAM)

Since the convergence rate of the identification actually depends on the amount of data processed per unit time it is desirable to use small sampling times. Foster and Hansen (1979) in an application paper suggest a value of less than one-third the time constant of the system (usually we have a rough estimate of the settling times of the process). However, we suggest a much shorter sampling period since faster parameter convergence is desired between input moves (i.e. TSAM < $\tau/80$).

2. Initial LS iterations (ILSQ)

When there is no previous information about parameter values, the LS algorithm should be used initially. Once fair estimates are obtained the IV recursion can be started and used subsequently for all iterations. Smets (1970) suggests an ILSQ of 10m where m is the total number of parameters, although proper selection really depends on the parameter rate of convergence.

3. Iterations between input moves (IMOVE)

The input should be changed as soon as good estimates are obtained, although too frequent perturbations have no effect on the optimization. A value for this parameter actually depends on practical considerations and thus on the specific problem. In any case, enough iterations should be allowed for convergence of the estimates.

4. Forgetting factor (XLAM)

For continuous adaptation of the parameters recent data points have to be weighted more than old ones. A simple way of achieving this is by defining at the Nth iteration

$$Z_{N+1}^T \Phi_{N+1} = \lambda Z_N^T \Phi_N + z_{N+1}\phi_{N+1}^T \qquad (10)$$

$$Z_{N+1}^T Y_{N+1} = \lambda Z_N^T Y_N + z_{N+1} Y(N+1) \qquad (11)$$

where $0 < \lambda \le 1$ is the forgetting factor and z_{N+1}, ϕ_{N+1}, Y_{N+1} are the vectors containing the latest samples. In this work the recursive form of the IV algorithm (Eykhoff, 1974) with $\lambda \le 1$ was used.

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.
Thus there should be an optimal value for the forgetting factor but it can only be determined experimentally for each different problem. A reasonable initial guess is XLAM = 0.95

Divergence of the Recursive Algorithm

Through our simulation studies we have found that certain conditions make the IV algorithm yield unbounded estimates. It is certain that if the information matrix $Z_N^T \Phi$ is singular the estimate $\hat{\theta}_N$ will diverge. According to the analysis in Söderström (1974) the following equivalence holds for the noise free bootstrap estimator:

$$Z_N^T \Phi_N = \hat{D} M D^T \qquad (12)$$

where: M is a square matrix of delayed inputs which is pos. def. iff the $m_j(k)$ sequences are persistently exciting (Eykhoff, 1974), 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. Then it follows that either

(i) M singular and $\hat{D}$, D of full rank, or
(ii) M non-singular, and $\hat{D}$ or D of defficient rank imply singularity of the information matrix.

Case (i) can occur close to the optimum where the decrease in the gradient brings about a decrease in the input moves. An increase in $\mu(6)$ was found to have a positive effect without causing significant oscillations at the optimum.

Case (ii) can occur when some parameters vanish. This can happen when the objective function is equal to one of the output variables. Then at the optimum the output is at its extremum and the input gain becomes zero. By choosing the initial covariance small enough this can be avoided (Albert and Sittler, 1965) but it slows down the identification significantly. We found it preferable to switch off the identification of the autoregressive part of the model close to the optimum.

EXAMPLE

In order to demonstrate the basic features of the proposed optimizing control algorithm, a simple reactor system shown in Fig. 3 was simulated on the digital computer. It is assumed that the concentration transients are dominant and that instant temperature response is obtained. The reaction is

$$A + B \underset{k_-}{\overset{k_+}{\rightleftharpoons}} 2B$$
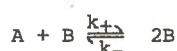
with rate $A = -k_+AB + k_-B^2$

The physical and design constants of the system are

$$\alpha = 0.5$$

$$\tau_1 = 30 \text{ min}$$

$$\tau_2 = 25 \text{ min}$$

$$E_+/R = 17786 \quad E_-/R = 23523 \text{ °K}$$

$$A_+ = 9.73 \times 10^{22}, \quad A_- = 3.1 \times 10^{30} \quad \ell/\text{gmoles-sec}$$

with an overall time constant of approximately 40 min.

We desire to find the temperatures T1 and T2 which maximize the concentration of B at the second reactor. The important disturbance for optimization is C_{AO}, the feed concentration of A which has a significant effect on the optimum as shown in the following table.

$C_{AO}\left(\dfrac{\text{gmoles}}{\ell}\right)$	$C_{B2}\left(\dfrac{\text{gmoles}}{\ell}\right)$	T_1(°K)	T_2(°K)
0.1	0.0731	315.	309.
0.06	0.0420	317.	312.
0.20	0.1536	312.	306.

For the identification each tank is modelled as a first order system with two inputs. A 0.5 minute sampling time is used throughout. For the optimization a maximum step size (STMX) of 1 degree is specified. The LS algorithm is run during the first 40 iterations and the search is started after 80. During these, PRBS of 1° of amplitude in each temperature are employed. All the following runs start at (T1,T2) = (308,299)K for C_{AO} = 0.10. Figs. 4 and 5 show the results when the measurements are noise free, XLAM=0.9, IMOVE=15. Convergence is achieved after about four system time constants which is about the time necessary for a system to reach a new steady state after a step change. Because of divergence, the identification of the autoregressive part of the model was interrupted in the vicinity of the optimum. An increase of the measurement noise requires XLAM to be increased. When XLAM is increased the parameter adaptation is slowed down and IMOVE has to be increased. It was observed, for example, that in the noise free case with IMOVE=15 and XLAM=0.95 the optimum was not found and the optimization algorithm seemed to be lost. The excellent performance in the presence of significant measurement noise can be seen in Fig. 6. XLAM and IMOVE had to be increased which brought about an increase in the search time. Fig. 7 shows the tracking of a shifting optimum caused by changes in the feed concentration at times of 600, 1500 and 2700 minutes. Again the response of the optimizing controller is seen to be very fast. In all examples a steepest descent search was used. The variable metric techniques were found to perform very poorly in the presence of noise. For comparison with Fig. 4 and 5, the performance of a SIMPLEX pattern search was tested under the same conditions with IMOVE=15 (Fig. 8). The convergence is much slower and persistent oscillations at the optimum appear. For noisy observations very poor results were obtained.

ACKNOWLEDGMENT

REFERENCES

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

Arafeh, S.A., Sage, A.P., "Multilevel discrete time system identification in large scale systems," Int. J. Syst. Sci., 5, 753, 1974.

Åström, K.J. and Eykhoff, P., "System Identification - A Survey," Automatica, 7, 123, 1971.

Avriel, M., "Nonlinear Programming Analysis and Methods," Prentice Hall, Englewood Cliffs, N.J., 1976.

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

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

Draper, C.S. and Li, Y.T., "Principles of Optimizing Control Systems and an Application to the Internal Combustion Engine," Amer. Soc. of Mech. Engr., Sept. 1951.

Eykhoff, P., System Identification, Wiley, London, 1974.

Foster, R.D. and Hansen, J.L., "Modelling Dynamic Chemical Processes Using Time Series," AIChE 72nd Annual Mtg. Nov. 1979.

Garcia, C.E and Morari, M., 1980, submitted for publication.

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

Kotnour, K.D., Bos, G.E.P. and Altpeter, R.J., "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.

Prett, D.M. and Gillette, R.D., "Optimization and Constrained Multivariable Control of a Catalytic Cracking Unit," Paper No. 51C, AIChE 86th Nat. Mtg., 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.

Sittler, Albert A. and Sittler, R.W., "A Method for Computing Least Squares Estimators that Keep up With the Data," SIAM J. Control, 3, 384, 1965.

Smets, A.J., "The Instrumental Variable Method and Related Identification Schemes," Eindhoven Univ. 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., Ljung, L., and Gustavsson, I., "A Theoretical Analysis of Recursive Identification Methods," Automatica 14, 231, 1978.

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

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

Wismer, D.A., Perrine, R.L., and Haimes, Y.Y., "Modelling and Identification of Aquifier Systems of High Dimension," Automatica, 6, 77, 1970.

Wong, K.Y., Polak, E., "Identification of Linear Descrete Time Systems Using the Instrumental Variable Method," IEEE AC-12, 707, 1967.
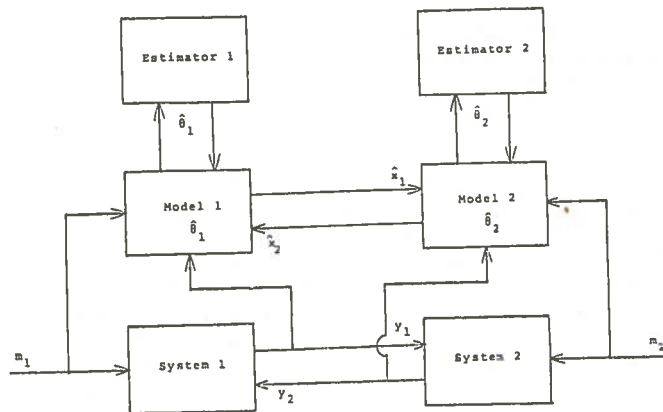
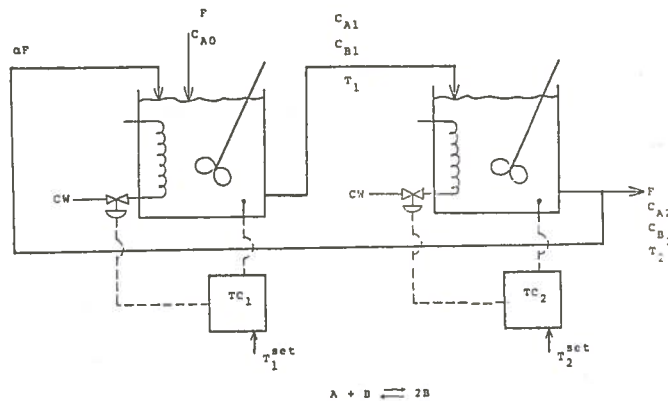Fig. 2. A decomposition approach to the IV identification of large interconnected systems.

Fig. 3. The structure of the physical system used in the example.
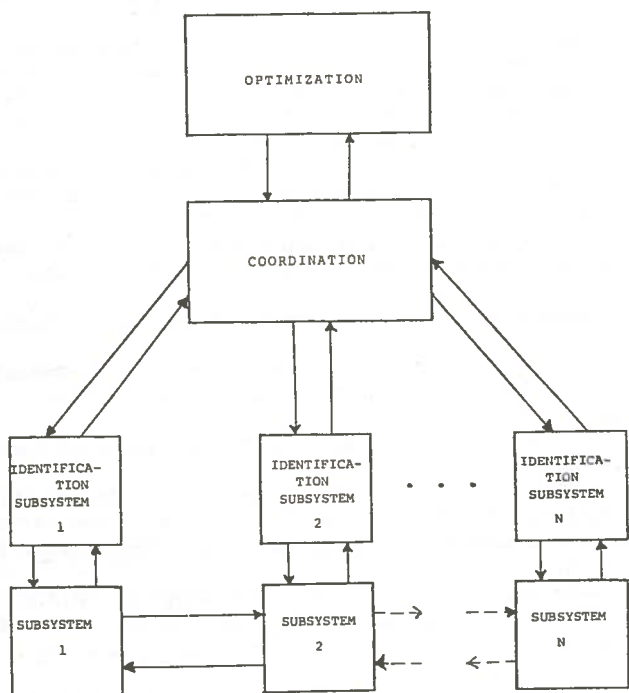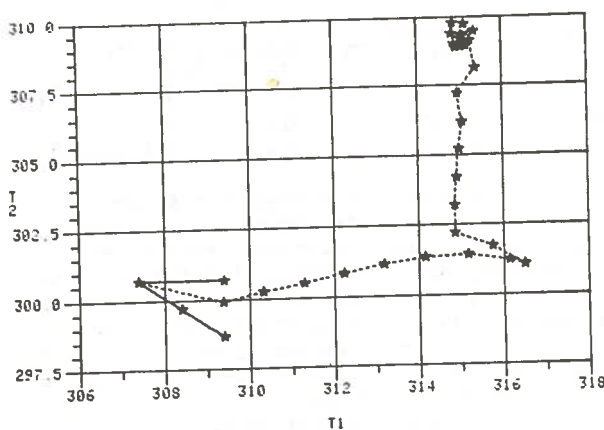

Fig. 4. Record of temperature setpoints during the search for the optimum (no noise, XLAM=0.9, IMOVE=15).

Fig. 1. Structure of hierarchical optimizing controller for large interconnected systems.
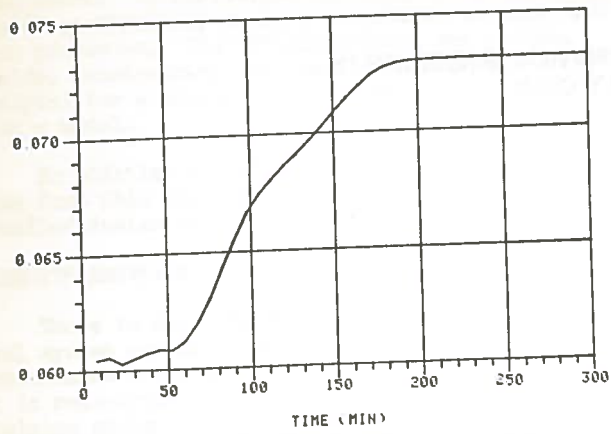
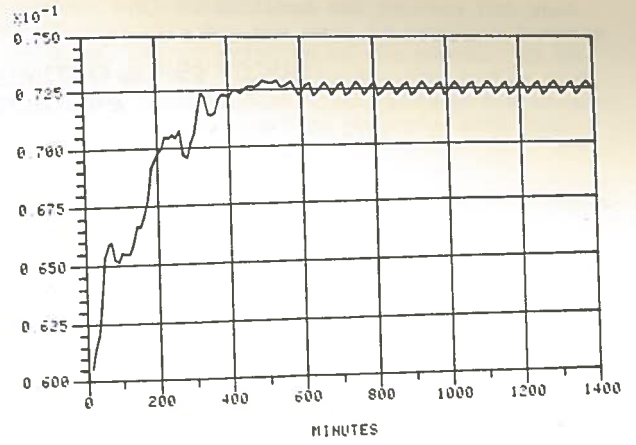Fig. 5.   Record of the objective function C_{B2} during the search for the optimum shown in Fig. 4.

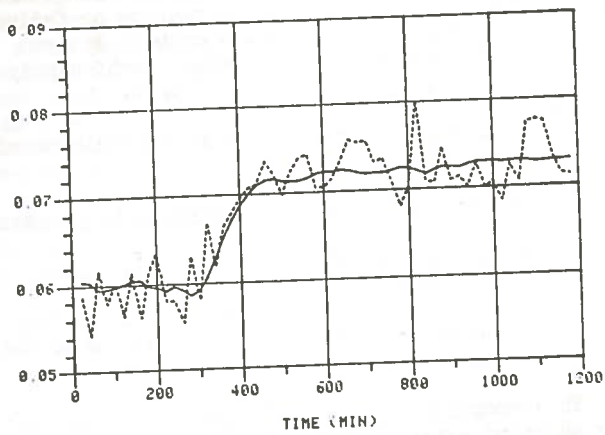Fig. 8.   Results for a SIMPLEX optimizing controller (IMOVE=15).

Fig. 6.   Search for the optimum under severe measurement noise (var=9 x 10^{-6}, XLAM=0.97, IMOVE=40).
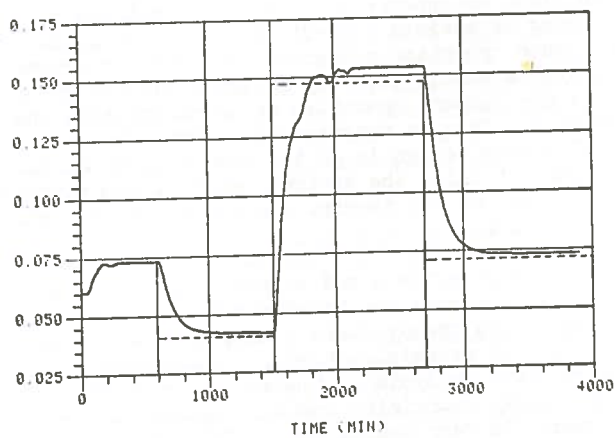

Fig. 7.   Tracking of a shifting optimum (--- steady state optimum without optimizing controller).

WP5-D