Automatica, Vol. 14, pp. 413–428
Pergamon Press Ltd. 1978. Printed in Great Britain

0005-1098/78/0901-0413 $02.00/0

Model Predictive Heuristic Control: Applications to Industrial Processes*

J. RICHALET,† A. RAULT,† J. L. TESTUD† and J. PAPON†

Different industrial processes are being computer controlled using a new dual algorithm which identifies input–output impulse responses and computes control inputs as needed to realize desired output trajectories even when system noise, disturbances, and parameter variations occur.

Key Word Index—Chemical industry; computer control; control theory; dual control; identification; Lyapunov methods; multivariable control systems; oil refining; predictive control; steam generators.

Summary—A new method of digital process control is described. It relies on three principles:

(a) The multivariable plant is represented by its impulse responses which will be used on line by the control computer for long range prediction;

(b) The behavior of the closed-loop system is prescribed by means of reference trajectories initiated on the actual outputs;

(c) The control variables are computed in a heuristic way with the same procedure used in identification, which appears as a dual of the control under this formulation.

This method has been continuously and successfully applied to a dozen large scale industrial processes for more than a year's time. Its effectiveness is due to the ease of its implementation (e.g. constraints on the control variables) and to its amazing robustness as concerns structural perturbations.

The economics of this control scheme is eloquent and figures can be put forward to demonstrate its efficiency. Optimality does not come from extraneous criteria on the control actions but from minimization of the error variance which permits computation of the set points of the dynamic control in a hierarchical way.

INTRODUCTION: USE OF DIGITAL COMPUTERS IN INDUSTRIAL PROCESS CONTROL

THE GROWTH of digital technology in the last few years has represented a challenge to automatic control research workers.

With the availability of much more powerful computers, should not the basic approaches to control systems application be reconsidered?

The theory of feedback systems originated from work on continuous electrical networks. Despite the progress of technology, the fundamentals of control theory remain unchanged. It is now natural to ask: can we conceive a novel type of control system that would use to the full the capabilities of currently available computers:

—storage of information—fast access memory

—fast computation, choice of a solution among several possibilities according to a criterion?

The achievements of modern control theory are well-known. Successful applications to aerospace guidance problems are remarkable. However the implementation of such techniques to industrial control has not been so successful.

Industrial processes are quite different, they are highly multivariable systems, perturbations affect the plant structure more often than the measurable variables. Industrial processes have their own performance criteria and reliability requirements. The economic and psychological environment required for a successful implementation is often not met in practice so that many constraints prevent the implementation of on-line control schemes on production plants.

We tried to conceive a control theory where digital computation and modeling play a major role and support new concepts.

Before publishing the principles of this new control scheme in order to give evidence of its efficiency we wanted to have a significant number of complex industrial applications working continuously for more than a year's time.

The organization of the paper is as follows:

—In Section 1, the general scheme is given of this new Model Predictive Heuristic Control whose software is called IDCOM (Identification–Command). An overall presentation of the method is given in the simplest way; we emphasize the implementation aspects. No comparison of this method with the classical approach is presented. In this section, applications are stressed.

*Received January 13, 1977; revised August 26, 1977; revised March 29, 1978. The original version of this paper was presented at the 4th IFAC Symposium on Identification and System Parameter Estimation which was held in Tbilisi, Georgian Republic USSR during September 1976. The published Proceedings of this IFAC Meeting may be ordered from: North Holland Publishing Company, P.O.B. 103, Amsterdam, West Netherlands. This paper was recommended for publication in revised form by associate editor K. J. Åström.

†ADERSA/GERBIOS, 53 avenue de l'Europe, F78140 Vélizy, France.

—In Section 2, the outlines of the most significant present industrial applications of the method are given, with a brief description of the systems and an economic appraisal of the results.

—In Appendix A, the identification algorithm is presented and its convergence analysed.

—In Appendix B, stability and robustness of the control scheme are analysed.

More details are given in [22, 23].

Since new concepts are put forward, new expressions have had to be coined. However the basic ideas are straightforward and can easily be understood. Interpreting or "naturalizing" in terms of classical control does not help to comprehend the method. No block diagram or black box representation can be used.

SECTION 1

1.1. *Model predictive heuristic control*

The MPHC strategy relies on three principles:

(a) The multivariable process to be controlled is represented by its impulse-responses which constitute the *internal model*. This model is used on-line for prediction, its inputs and outputs are updated according to the actual state of the process. Though it could be identified on-line, this internal model is most of the time computed off-line as explained later in the self-adaptation Section 1.4.

(b) The strategy is fixed by means of a *reference trajectory* which defines the closed-loop behavior of the plant. This trajectory is initiated on the actual output of the process and tends to the desired set-point.

(c) Controls are not computed by a one-shot operator or controller but through a procedure which is heuristic in the general case. Future inputs are computed in such a way that, when applied to the fast time internal predictive model, they induce outputs as close as possible to the desired reference trajectory.

1.1.1 *Impulse response representation*

(a) *Why*? Use of impulse responses to represent systems is a controversial technique. It is not new (Frechet 1910–Volterra 1930) [25], but it is not regarded as convenient by the classical or "modern" control theory. Differential equations, transfer functions and state representation are very well suited to analyse the type of processes that were considered to be at the origin of circuit analysis and control theory. All these processes can be considered as analogous to a system composed of an inertial mass moving in a field force. The physical phenomena are basically *incremental*: increase of the speed of a mass, decrease of the voltage across a capacitor, etc.

In a large category of processes encountered in industrial plants, the physical phenomenon involved can best be described as that of the *influence* of a source, through channels of different lengths, thus of different losses and time delays, on a localized sensor. One can interpret the impulse response as the probability density function of the occurring effect of the source on the sensed point. The phenomena involved can be: heat, physical transformation, propagated chemical effects, etc. Time delays and non-minimal phase effects are often encountered. Since most industrial processes are interconnectable, their "order" would necessarily have been large (e.g. 20), in such a way that if N is the number of parameters defining the impulse response, with $N=40$ the redundancy factor is generally close to 2 and does not exceed 3.

The main criticisms that were traditionally put forward against this representation dealt with the non-minimality of this modeling. The ill-posed nature of the identification involved—if any—can be avoided by a proper conditioning of the relaxation factor or a proper selection of the structural distance to be minimized, as discussed in Appendix A, and in ref. [16].

Now, why object to a procedure which saves time? What is the use of a lengthy analysis where one has to characterize the process, test several hypotheses, look for a minimal representation? The goal of modeling is clearly in this case to predict the behavior of the plant when it is submitted to a known input. The accuracy of the prediction depends largely on the adequacy of the test signals and on the identification technique.

As we will see, the one used gives easily the unbiased parameters of a multivariable process submitted to unknown perturbations. Since these parameters appear linearly in a non-matrix form, 12 impulse responses with 40 parameters each, are commonly identified.

Lastly we have to point out that there are no "initial conditions" with this representation, which makes the identification of multivariable systems easier, and that "structure" and "variables" play a symmetrical role, which will be the base of the control algorithm as described in Section 1.1.3.

The characteristics of this representation are listed below.

(b) *Universality*. Each output $s_j(n)$ of a multivariable system is a weighted sum of the N past values $e_k(n-i)$ of the NE inputs

$$s_j(n)=\sum_{k=1}^{NE}\sum_{i=1}^{N} a(i)_{k,j}e_k(n-i)$$

or in a vector form

$$s_j(n) = \mathbf{a}_j^{\mathrm{T}} \mathbf{e}(n)$$

$$\mathbf{a}_j = \{a(i)_{k,j}\} \qquad \mathbf{e}(n) = \{e_k(n-i)\}.$$

The information used to represent the system with NE inputs and one output is thus carried by a vector of dimension $NE \times N$, if we assume for simplicity that all impulse responses have the same time memory NT (T = sampling period).

The choice of N and T should be such that

$$NT > TR$$

where TR is the time response of the system.

The system is assumed asymptotically stable, which implies that it is possible to find N such that the truncation error can be arbitrarily small. If one integration is present in the transfer function, such as in level control, one can represent the system by an equation of the following type

$$s(n) = s(n-1) + \mathbf{b}^{\mathrm{T}} \mathbf{e}.$$

—Truncation errors should have the same order of magnitude as the other errors: representation, sampling, computation...

—If the systems were to be naturally unstable—which rarely happens—they could be stabilized by some standard procedure with no claim on performance while IDCOM would optimize their performance in a supervisory way as discussed later in the section on Implementation.

(c) *Linearity*. The major feature of this representation is linearity with respect to the parameters a_j. It is then permitted to model systems by equations of the following type

$$s(n) = \mathbf{a}^{\mathrm{T}} \mathbf{u}(n)$$

where $u(n) = f(e(n))$, $f(\cdot)$ being any function provided $u(n)$ can be measured or estimated. Nonlinearities induced by the actuators can easily be taken into account by this procedure.

(d) *Identification*. Linearity with respect to the unknown parameters facilitates identification procedures. The identification algorithm and proof of convergence under various conditions are given in Appendix A. In most industrial cases, perturbations affect the outputs only and not the measured inputs of the process so that if identification is performed in open-loop conditions, unbiased estimates of the parameters are easily obtained in realistic conditions where poor signal to noise ratio exists.

No matrix computation being involved, this identification technique can be implemented easily on a small on-line digital computer. However it will not be used on-line continuously.

(e) *Test signals*. The only critical phases are experiment planning and data collecting. A proper hierarchical approach should have first defined the control and feed-forward variables and the outputs. This part turns out to be difficult on large systems. Active test signals should perturb normal operating conditions as little as possible and should give the best information on the system structure. These demands cannot be satisfied without some previous computation. It is necessary to select carefully the nature of test signals that will be added to the actual controls.

Pseudo Random Binary Signal, or deterministic dedicated signals, have proved to be convenient provided that the spectrum and the cross-correlation of inputs were appropriate.

The amplitude of inputs should be above the significant threshold of the sensors. The variance of the estimated parameters depends on the energy (amplitude × time) of the test signals and on the perturbation variance. The time duration of tests is limited by practical considerations and by the non-stationary characteristics of the process.

Thus it follows that test signals optimisation can be done only when the process model is known. To break this vicious circle, an iterative procedure is used. Intermediate models can be used, one or two trials usually being sufficient. Identification can be performed by either of the following procedures:

—The system is open-loop. A weak manual control is tolerated on some outputs, inducing a negligible bias on the parameters.

—The system is controlled in a "transparent" or supervisory way, as described later in the Applications section, where the analog controllers are still working on-line. An overall model will then be derived. Test signals are added to the set-point values of the PID controllers. These operating conditions are much in favor. Since risks are minimized, operators are unstressed and more tolerant.

1.1.2 *Reference trajectory*

For sake of simplicity let us consider a single input, single-output system. Let C be the constant value of the prescribed output and $s_0(n)$ the actual output value at time n, as illustrated in Fig. 1.

From the last sampled value $s_0(n)$ a trajectory $s_{MR}(n+i)$ is initialized which reaches C according to some criterion (e.g. no overshoot, fixed time response). These desired values of the future output can be obtained from stored data or

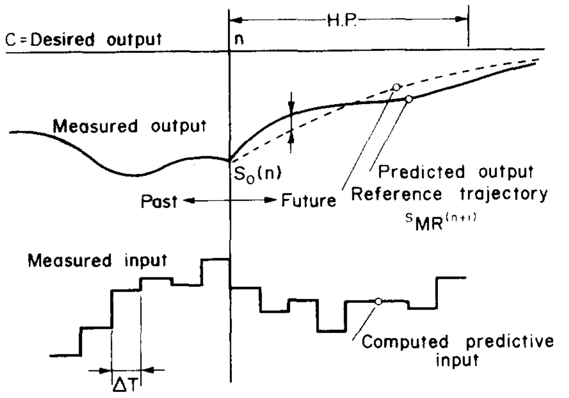


FIG. 1. Model predictive heuristic control: reference trajectory updating.

computed by a recursive equation. One of the simplest (1st order) is

$$s_{MR}(n+i)=\alpha s_{MR}(n+i-1)+(1-\alpha)C \quad (i>0)$$
$$s_{MR}(n)=s_0(n).$$

The main characteristic of the reference trajectory is its time duration: TMR. This parameter is one of the few to be specified in the program and must be accessible to users.

The control algorithm has to find a set of future control variables such that the future outputs of the internal model will be as close as possible to the reference trajectory. The whole procedure is to be repeated at every sampling period. Due to state and structural perturbations and to computational errors, the trajectory of the actual process does not usually fit the one that was prescribed in the past. The control problem is then reduced to the computation of control variables acting on a known and non-perturbed system: the internal model, so that its future outputs from n to infinity follow the reference trajectory.

The future trajectories that will be followed by the actual process and the internal model may be different, mainly due to the internal model mismatch which may induce instability if too large. This will be analysed in Appendix B.

1.1.3 *Control algorithm*

Given a model of the process, and fast time computer facilities, if a solution exists, then by any heuristic means the control variables will be computed in such a way that the output of the simulated process follows the reference trajectory.

In the particular case of industrial processes, the impulse-response representation will be used to advantage. In this case, the systems are "state-linear" with respect to inputs $e_j(n-i)$ and "structure-linear" with respect to parameters $\mathbf{a}_j$. Commutativity of the inner product is used as schematized on Fig. 2

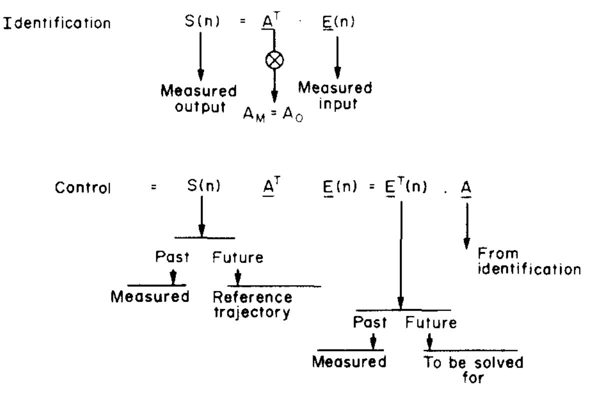


FIG. 2. Duality of identification and control.

$$s(n)=\mathbf{a}^{\mathrm{T}}\mathbf{e}(n)=\mathbf{e}^{\mathrm{T}}(n)\mathbf{a}.$$

In the identification scheme $s(n)$, $\mathbf{e}(n)$ are given; the problem is to find $\mathbf{a}$.

In the control problem:

$\mathbf{a}$ is given by the previous identification;

$s(n)$ is known, in the past from the collected data, in the future by the reference model trajectory;

$\mathbf{e}(n)$ is given in the past from the controls actually applied; the problem is to find $\mathbf{e}$ in the future. The same type of algorithm as used for identification will be used, since it appears that *control and identification are dual.*

Constraints on the controls should be introduced. Limitations imposed by actuators are described by

$$m_e<e(n)<M_e$$
$$|e(n)-e(n-1)|<V_M.$$

Constraints on internal variables or secondary outputs can be introduced

$$u(n)=\mathbf{h}^{\mathrm{T}}\mathbf{e} \qquad u(n)\in D.$$

Theoretically, to take into account properly the constraints, all the future controls have to be computed. Practically, the predicted time span can be shortened to a few points with little loss of performance: $HP\simeq 10$.

The different steps of the control algorithm are brought up in the diagram given in Fig. 3.

1.2 *Program specifications*

The control being heuristic in the sense that no structure or control law is used, the tuning of "parameters" to plant modification or operating point variations is meaningless here. Indeed, the *process model* is used in its initial form, the impulse response, for prediction. The only argu-

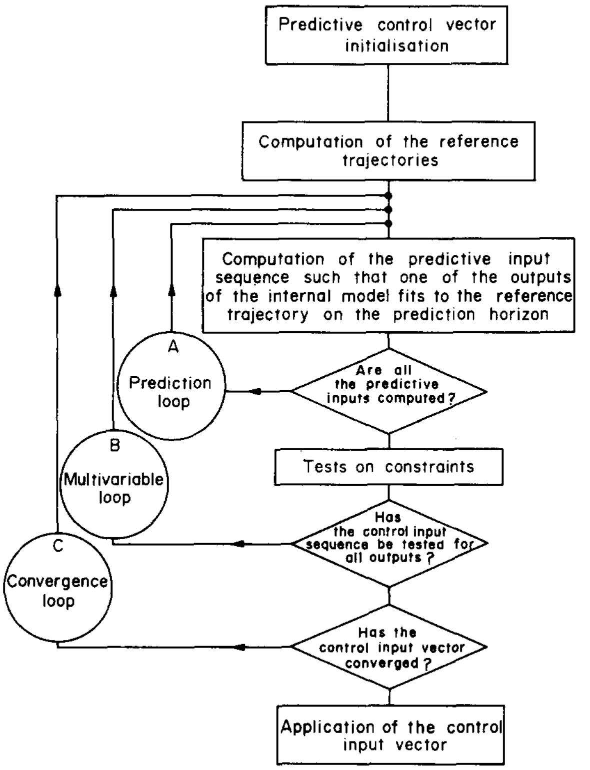


FIG. 3. Flow-chart of the control algorithm. The (A) loop iterates to compute each predicted input vector sequence needed to obtain a fit between the internal model output and the reference trajectory for a number of sample times in the future. Computation is based on equation (15) of Appendix B using the dual form of the algorithm defined in Appendix A. The number of future points HP to be fitted is defined in Fig. 1. Tests on constraints being taken into account, loop (B) iterates on the number of outputs to be controlled. Loop (C) iterates on the whole predicted input control vector sequence ensuring convergence. Once this is satisfied only the first input control vector of the HP predicted sequences is applied. The whole procedure is repeated at the next sampling time.

ments then to be specified have a direct meaning readily understood by users.

—Internal model: **a**, where $s=\mathbf{a}^{T}\cdot\mathbf{e}$, where control variables will be distinguished from measurable perturbations used as feedforward prediction.

—Reference trajectory: for a given class, its time-response T_{MR}.

—Constraints on controls: max, min, velocity max on internal variables if necessary.

The last parameters should be left at the disposal of the process operator and chosen on-line. In particular, when the control is implemented for the first time, time-response and constraints can be strictly set and optimized afterwards.

1.3. *Performances*

Taking into account the fact that some investment, either intellectual or financial, is necessary to obtain a model of the process, computed off-line or on-line on a digital computer, noticeable improvements are to be expected on the quality of the control if such effort is to be justified.

—In fact, performances achieved by this control scheme are up to expectations and this is clearly demonstrated by figures given in Section 2.

—However, other schemes, on particular applications, can give similar results. Our claim must then rest on the remarkable robustness of the method which is essential if it is to survive in an industrial environment. These requirements are:

—ease of implementation of software;

—program parameters must have a clear control significance, e.g.: time response—constraints;

—large tolerance to variations of the process structure due to errors of identification: natural alteration of the gains and dynamics of subsystems, change of transfer characteristics according to set-points (nonlinear effects), weak sensitivity to noise... All these conditions should be satisfied while control of multivariable processes (time delays—non minimum phase) is guaranteed, and constraints satisfied.

In Appendix B, results are given on the stability limit which show that one can cope with large gain variations of the process (>4) at the cost of quite small changes in performance. Thanks to this approach, stability appears to be not so critical. The reference trajectory intermediate goal and the internal model prediction being responsible for it. Controlling a "second order" system is as easy as to control a "tenth order" one, if the internal models are correct. Robustness and, at the limit, stability will then depend on the fitness of the internal model.

1.4. *Necessity of adaptation*

It is quite obvious that if the internal model is far from reality, and if the reference trajectory is much faster than the process, the MPHC strategy will not be efficient. The better the internal model is, the more we can demand from the control. Adaptation of this model thus appears to be necessary if we want to maintain the optimal operating conditions.

However, in the general case, auto-adaptation will not be the solution required and continuous identification will not be performed on a controlled industrial process. There are several reasons for such an assertion:

—if the set-points are constant, for a large category of problems, the natural robustness of IDCOM can cope passively with the natural variations of the plant's structure.

—on-line identification—if a non-academic case is to be considered—imposes practically the addition, in one way or another, of some independent extraneous signals to the controls[19, 20, 24]. Not to perturb the natural operating conditions, these test signals should be small and identification should be performed with a poor signal-to-noise ratio. Such identification techniques exist but large artefacts are always probable (e.g. instrumentation defects...), they are easily detected by the inspection of a human operator but very difficult to analyse in all cases by an algorithmic procedure. Auto-adaptation then appears not to be so necessary and perhaps too risky to users. On the contrary, adaptation or tuning of the internal models from *a priori* information is well accepted and used satisfactorily. This information is derived from measured variables, e.g. the sampling rate is continuously fixed by the power, in the boiler example described in Section 2.

In practice, depending on the local conditions, a general overhaul of the internal model can be performed once a year for instance. An optimal test-signal procedure will be used on-line, but the results should be inspected by a human operator before being fed back into the control program.

1.5. *Implementation*

The poor reliability of digital computers initially hindered the progress of digital process control. Nowadays, the central unit Mean Time Between Failure is large enough to provide a good service. If software and control architecture do not demand magnetic disks or tapes, pure solid electronic devices will ensure a reliable control. Nevertheless, it is useful for different reasons to keep a hierarchy of controllers and analog PID mixed with digital computers. Two main possibilities are offered:

Direct digital control (*DDC*). As indicated in Fig. 4, computers control the process control variables, which are often set-points of cascaded "level 0" PID controllers as discussed below in Section 1.6. The internal model is in this case the process model. Outputs and set-points are fed through the appropriate peripherals. In this implementation, constraints on the actuators are easily formulated. On the other hand, identification procedure and supervisory software (watch dog) require careful attention.

Transparent control (*TC*). As shown in Fig. 5, the outputs of the control computer are the set-points of conventional analog controllers operating at "level 1". In the classical SPC mode (Set-Point Control), the set-points are constant. The

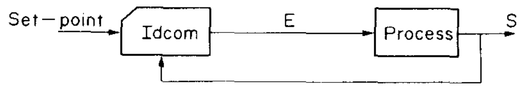


FIG. 4. Direct digital control.

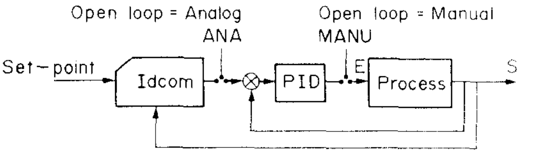


FIG. 5. Transparent control.

TC mode can be considered either as a "dynamic SPC" or a DDC procedure applied to a closed-loop system with sluggish—thus robust—analog controls. If the controllers are PI, unity static gains are ensured for diagonal transfers and zero gain for cross-transfer. A progressive set of solutions: manual—analog, digital can be used in a permanent way. Some outputs can be controlled by an analog controller, some by IDCOM which, in a direct or supervisory way, ensures optimal control of the process.

Security insured by the analog back-up and progressivity of the implementation make this type of control more attractive to users[22].

1.6 *Necessity of a sophisticated control algorithm?*

Nowadays effective control schemes used in practical industrial applications are digital transcriptions of analog control laws[21]. What is the purpose of an industrial controller? What are the criteria? Strict dynamic control must be imbedded in a larger problem which can be divided into 4 hierarchical levels.

Level 0	Control of ancillary systems (e.g. servo-valves) where PID controllers are quite efficient.
Level 1	Dynamic control of the plant—multivariable process perturbed by state and structural non-measured perturbations.
Level 2	Optimization of the set-points with minimization of cost-functions ensuring quality and quantity of production.
Level 3	Time and space scheduling of production (planning—operation research).

The economic benefits induced by levels 0 and 1 are in practice usually negligible. In contrast, level 2 optimization can bring valuable improvements in the economics of the systems. However, a necessary but not sufficient condition for satisfactory level 2 optimization is first to have levels 0 and 1 optimized. If a regulator is operating around a fixed set-point, no significant gain in energy and raw material consumption can be obtained by a sophisticated dynamic control. On the contrary, an optimized level 2 setting needs a

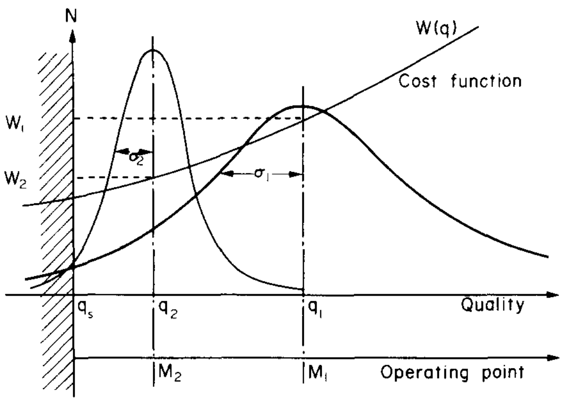


FIG. 6. Level 1: optimisation.

good quality of control around the level 1 set-points. Reducing the variance of the actual output around their prescribed mean values allows level 2 to be set in a better way, closer to the specified quality variables.

A diagram of great generality is given by Fig. 6 where the distribution is plotted of the measured quality of the product which is specified to be $q \geq q_s$, see Application 1 for example.

With a non-optimized control scheme, the set-point q_1 should be such that in "all cases" (95%) $q > q_s$. With an optimized scheme $q_s < q_2 < q_1$, the operating conditions will be closer to the prescribed limit q_s. The cost function $W(q)$ is such that $W(q_1) > W(q_2)$.

Thus it is clear that improvements on the dynamic control laws, provided stability is ensured, will not be concerned with "energy" of the controls. Constraints have to be respected and the variance of the outputs has to be minimized.

From the important, though classical, scheme of Fig. 6 [23], we maintain that in this case no quadratic criterion on the control variables is of any use. Optimality, in an industrial process, comes from level 2.

For a given frequency-spectrum of perturbations, the time-response of the system (TMR) is then the main parameter to be optimized.

To fulfill such a goal, a robust method of control is needed. A global approach is necessary if it is to be applied to any system. The classical method based on the processing of the error, a set-point minus actual value, which leads to the well-known trade-off between stability and precision appears to be fragile since the parameters of the controller depend too much on the structure of the process.

SECTION 2: EXAMPLES OF APPLICATIONS

Three different types of applications which have been implemented for at least one year and are now used on a routine basis, will be presented. Two of them come from the chemical industry:

—a complete PVC plant where most of the processes are controlled through the MPHC strategy;

—a distillation column in an oil refinery.

The third example is a power plant in which the algorithm ensures the control of the whole steam generator.

Because of the generality of approach of the control problem through the MPHC procedure, the three examples will be analysed at the same time.

2.1. *Description of the processes*

In the oil refinery, the problem is to control the qualities respectively of the heavy and light products. It is a "level 2" type of control; however the qualities of the outflowing products are correlated to definite temperatures. Therefore the first problem (level 1) is to control the respective temperatures. The column is schematically described by Fig. 7.

The temperatures to be controlled are TI 62 and TI 63, representative of the light and heavy products respectively. The control variables are the product outflow RD208 and RD209. The system is perturbed by several interactions from the rest of the plant. These are fairly well represented by the top and bottom temperatures TI 18 and TI 61.

A physical constraint which depends very much on the outflows of the products is the "pan" level; it should be kept within a fixed

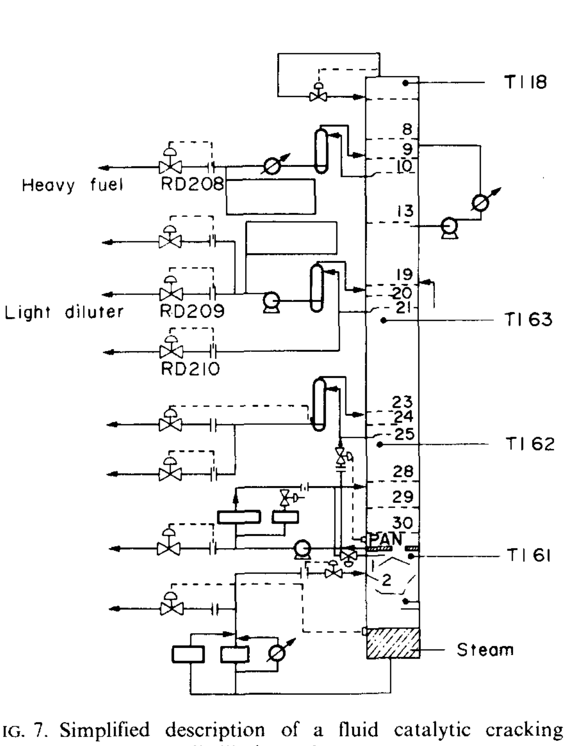


FIG. 7. Simplified description of a fluid catalytic cracking distillation column.

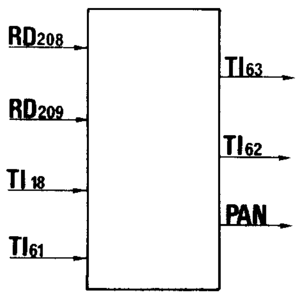


FIG. 8. Control diagram of the process of a fluid catalytic cracking.

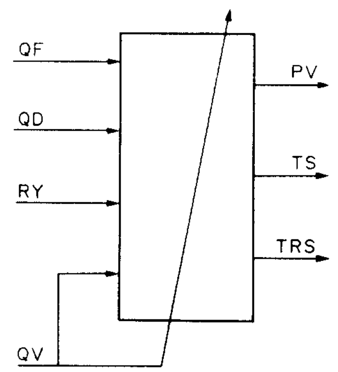


FIG. 10. Control diagram of the 250 MW steam generator.

interval to ensure correct working conditions of the distillation column. The mathematical model representative of the problem is given by the block diagram of Fig. 8 on which each input variable is related to each output through an impulse response to be identified.

Two variables are to be controlled with a physical constraint on a third one, using two control variables. The procedure which was finally adopted was to check on the "pan" level and each time it goes beyond its allotted interval, to progressively change the set-points of TI 62 and TI 63 in order to satisfy the "pan" level. This is only possible with an algorithmic type of control. The control is a direct digital one with a sampling period of 3 minutes.

The steam generation process in the power plant is represented by Fig. 9. The variables to be controlled are the steam pressure P_v delivered to the turbine and the steam temperatures at the superheater T_s and at the resuperheater T_{RS} using as control variables:

the deheater flow: Q_d
the recycling air flow: R_y
the fuel inflow: Q_f submitted to load variations of the plant, measured by the steam outflow Q_v.

This can be reduced to the schematic representation given by Fig. 10.

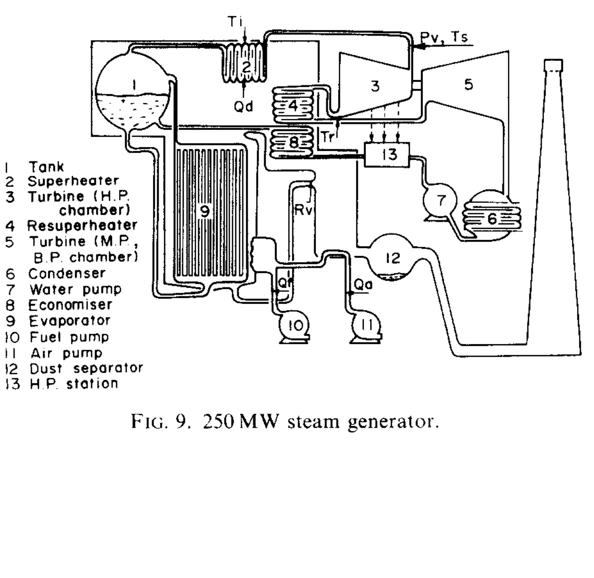


FIG. 9. 250 MW steam generator.

The structure of control in this example is a mixed one, P_v is controlled in a transparent mode while T_s and T_{RS} are controlled in a DDC mode as schematized on Fig. 11.

The third implementation presented is a whole chemical plant synthetizing vinyl chloride (PVC); a description of the process is given in Fig. 12. It is composed of two types of processes: separation processes, which actually are distillation columns, and transformation processes, which are cracking furnaces. Four distillation columns and three furnaces are controlled in a transparent mode according to the MPHC procedure.

This implementation is particularly interesting because it provides evidence of the generality of the method: five processes are controlled and within each process several types of controls are performed (temperatures, levels, impurities). The particular analysis of each process is quite similar to the preceding ones and thus will not be repeated.

2.2. *Process modeling and identification*

It has been shown in the previous descriptions of the industrial processes that they generally are multivariable systems. Under the hypothesis of linearity in the neighborhood of an operating point, the most appropriate mathematical model is the impulse-response representation.

Thus, the model of the refinery distillation shown in Fig. 7 is constituted of a set of 12 impulse responses.

Identification has been realized on-line and off-

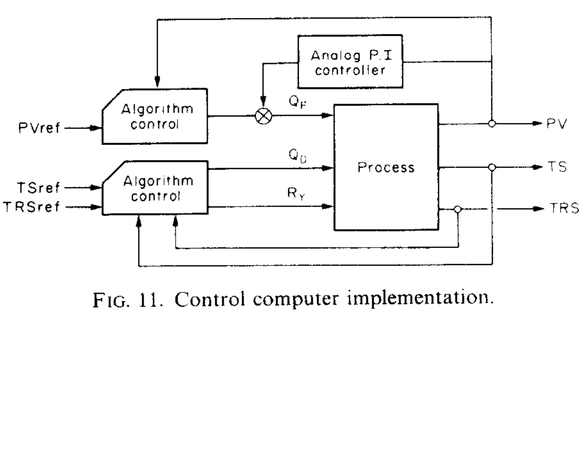


FIG. 11. Control computer implementation.

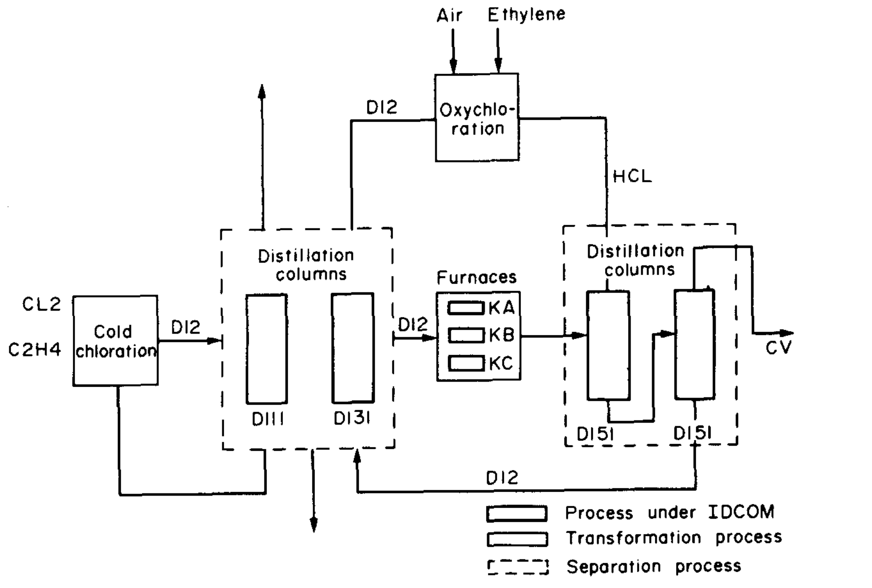


FIG. 12. P.V.C. plant description.

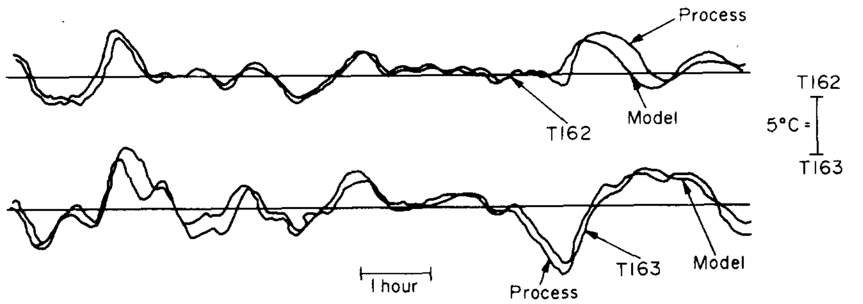


FIG. 13. Identification of the fluid catalytic cracking distillation column.

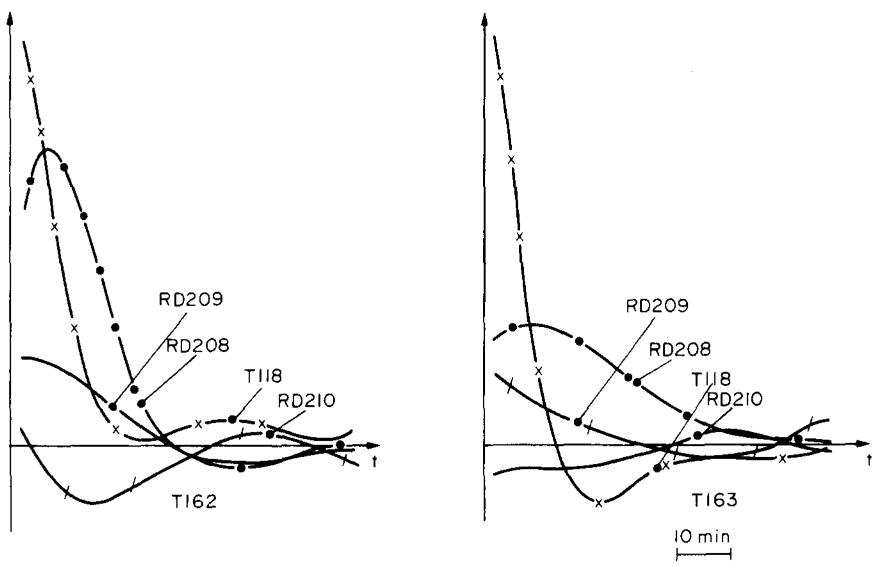


FIG. 14. Identified impulse responses of the fluid catalytic cracking distillation column (ordinate: arbitrary units).

line and results are shown in Fig. 14. Each impulse response is composed of 30 points. Comparison of the behavior of the plant and its mathematical model is shown on Fig. 13.

Results of the identification of the pressure of the steam generator are given on Fig. 15. These identifications have been performed on open loop systems.

In most applications, the control algorithm used thereafter being sufficiently robust, it was not necessary to have an on-line identification scheme, because generally the processes work around the same operating point.

In the case where the load of the process is liable to change, as for example in the power plant application, on-line adaptation becomes

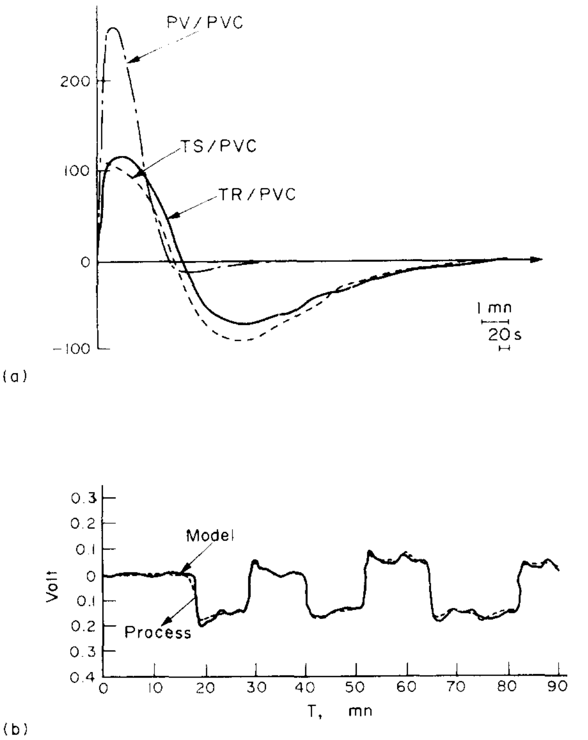


FIG. 15. Identification of a steam generator (a) impulse responses (ordinate: arbitrary units); (b) process-model outputs (1 volt = 20 bars).

necessary. This problem requires a more sophisticated solution and in general we are faced with a dual control situation. In the power plant case it was found after identification at different load levels (100%, 80%, 50%) that the time response of the transfers related to the temperatures varied as an inverse function of the load Q_v.

If one observes the process with a sampling period varying in the same way it appears as a nearly stationary system. This has been implemented and gives full satisfaction; the only identification left to be done on-line deals with the gain of each impulse response.

It should also be mentioned that the non-linearities of the actuators have been included in the internal model of the process as for example in the recycling air inflow R_y; such cases should be identified independently.

2.3. *Parameters of the control algorithm*

The main parameters of the control algorithm are the time-constants of the reference trajectories. They define the desired behavior and stability of the controlled variables.

Usually in industrial processes, the goal is to accelerate the natural response of the system within the limitations of the constraints on the actuators.

The constraints are of the (max, min) type on the amplitude and speed of variation of the control. For example the water injection flow Q_d in the power plant as well as its speed of variation are both limited.

Models of the processes were not—except for the gain of the steam generator—modified online; no adaptation was necessary.

For a "4 inputs × 2 outputs" system the whole software needs less than 2 K words of 16 bits, including data and programs.

2.4. *Results*

With such a general procedure, results should be given at different levels. The numerous simulations necessary to verify the different principles of this MPHC strategy will not be mentioned; on the contrary, results will only be concerned with industrial applications. Although industry may be a constrained experimental field, it is a source of truly unsuspected actual problems from whose analysis great benefit can be derived.

(a) *Recording.* For the oil refinery distillation column, the effects of a shift from analog control to IDCOM are presented in Fig. 16.

In the steam generator example, a standard triangular 10 MW/minute power perturbation was applied, with regular analog control and with IDCOM as shown in Fig. 17.

(b) *Variance.* A more objective way which yields numerical results is to compute on-line the mean value and variances of the variables to be controlled. This has been done on the distillation column (refinery).

Variances on the two temperatures are divided by a factor 4 which permits the shift of the set-points to better level 2 operating conditions as revealed in Fig. 18.

(c) *Economics.* As described above the level 1 optimization is significant only on a level 2 management of the process. A few examples are given:

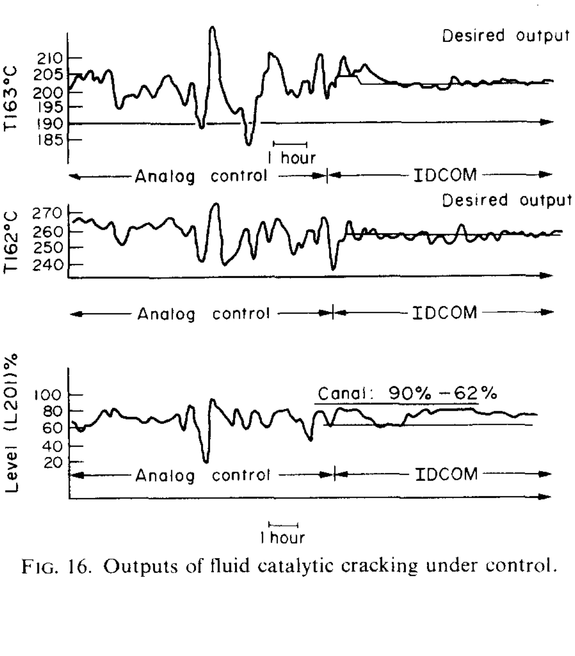


FIG. 16. Outputs of fluid catalytic cracking under control.

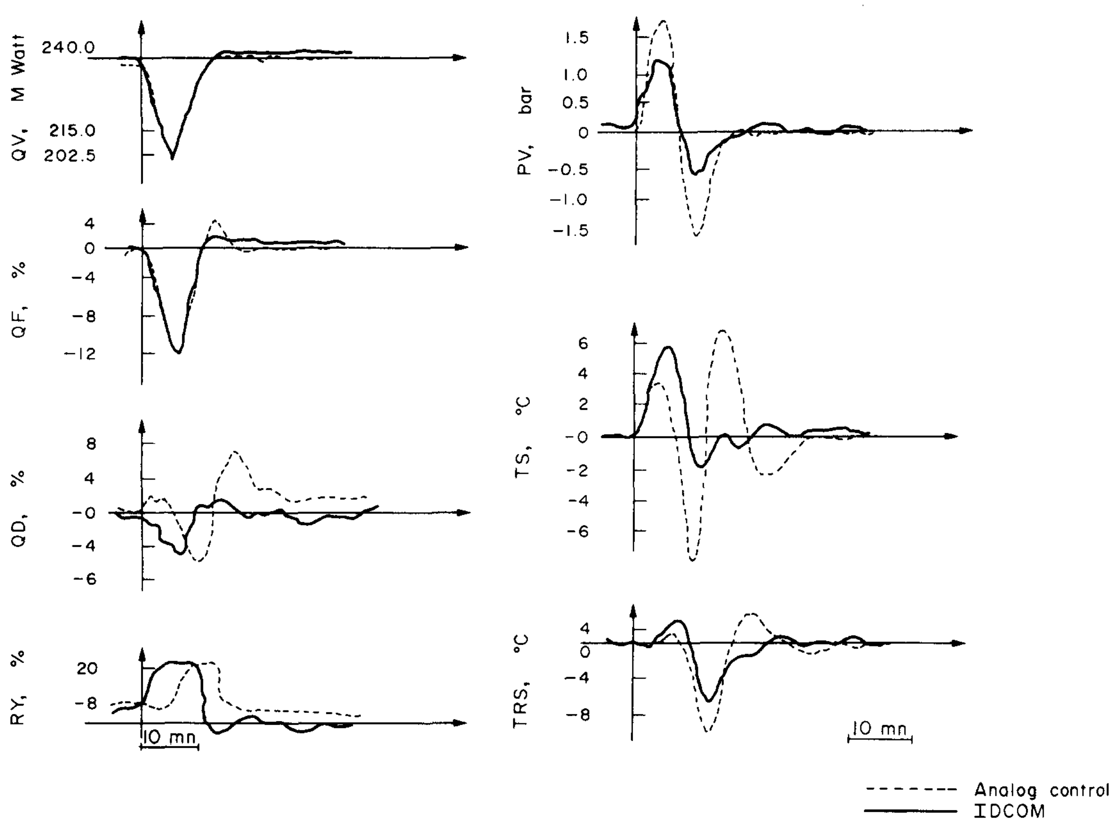
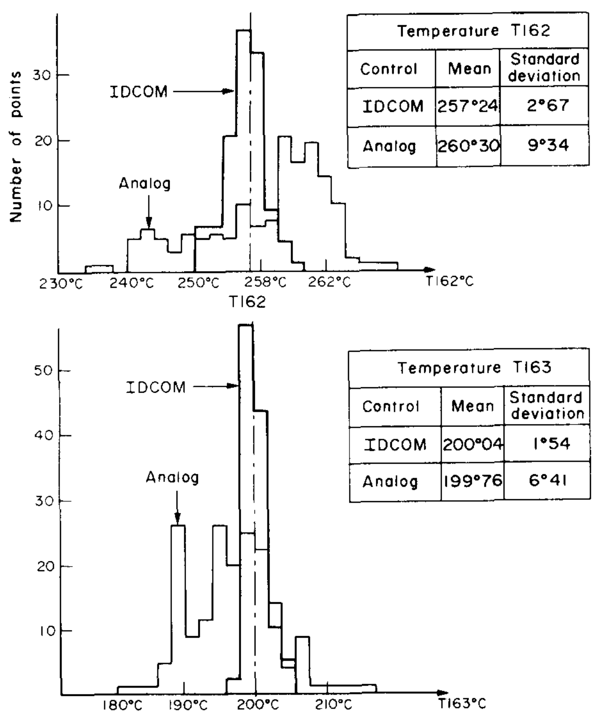


FIG. 17. Analog and IDCOM control with a load perturbation (Qv).



Temperature T162		
Control	Mean	Standard deviation
IDCOM	257°24	2°67
Analog	260°30	9°34

Temperature T163		
Control	Mean	Standard deviation
IDCOM	200°04	1°54
Analog	199°76	6°41

FIG. 18. Histograms of temperature control (FCC).

—In the PVC plant on a distillation column (D 131 Purification), a small variance of the outputs and security feeling induced by the new control, allowed to reconsider, after a few months of satisfactory continuous operating conditions, the specifications of level 2. Step by step in a year's time, specifications were almost doubled and the minimal constraint on the outflow was lowered from 45 T/H to 32 T/H. That induced an economy of energy on this column of about 1.5 T/H of steam ($\simeq 15\%$) which on an 8000 hours/year basis represents about \$120 000 a year.

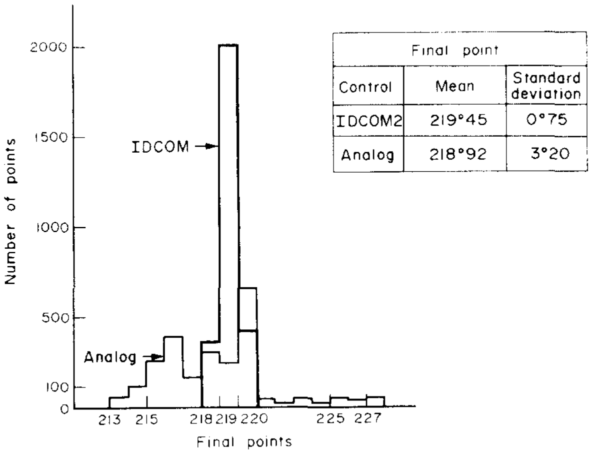
—On distillation column D 111, under similar conditions the minimal constraint taken into account by IDCOM was lowered from 10 T/H to 9,2 T/H with an estimated gain to \$100 000 a year.

—On the cracking tower in the refinery, lowering of the variance of the temperatures, while ensuring the prescribed quality permits an increase of output of the most valuable product (light gas) and a gain estimated at \$150 000 a year.

Moreover IDCOM has been applied to "level 2" control of the quality of the "Final Point" product as shown in the schematic diagram on Fig. 19. Improvement of the variance of the final point, the quality index of the gas oil, is by an histogram.

This is the economic aspect of the results; it is to be kept in mind. With such figures the pay-back time of a digital control system is very short and profitability can no longer be invoked to disregard such techniques.

Human aspects. To dispel the ever-present fear of innovation in industrial control systems, the acceptability of any new product should be



Final point		
Control	Mean	Standard deviation
IDCOM2	219°45	0°75
Analog	218°92	3°20

FIG. 19. Histogram of quality control.

guaranteed. The people concerned range from managers to operators and it is the latter who, in the long run, will accept or reject any modification.

To be accepted, innovation should not in itself be a source of troubles (premium non nocere); hardware and software should be reliable. The amazing robustness of the MPHC procedure makes it applicable in tough and changing conditions. Any new method should be clearly understandable: with IDCOM there is no gap between the operator's capability and plant complexity since all control parameters have a straightforward physical significance (time response—constraints...). Last, but not least, innovation should make control easier. For example, the oxychloration furnaces were difficult to drive, taking too much of the operator's attention—the "PAN" level constraint was so hard on operators that most of their time was occupied by this non-challenging but dangerous task: after a few months' of application of IDCOM, as confidence gradually pervaded, the operators were relieved from that preoccupation to such an extent that they would never accept a return to the previous situation.

Being no longer concerned with what is, in fact, a minor problem: dynamic control at "level 1", more significant improvements—"level 2" setting—can be looked for on a sound basis. The way is then opened for true optimization.

CONCLUSION

The conclusion is two-fold. Two types of results can be put forward.

1. *Digital control of multivariable industrial processes*

Several points are to be mentioned:

(a) Compared with the classical transfer-function or state-equation representations of systems, the impulse-response can be used to advantage, since characterization and identification become universal and easier. A parametric model of a system with a minimal order need not be hypothesized, derived and identified.

(b) Detailed knowledge of the process and search for the minimum order of the system appear from practice to be superfluous for control purposes. On the contrary, redundancy of the impulse-response representation is appreciated.

(c) A large number of parameters and variables can be dealt with by modern digital computers and iterative non-matrix calculus.

(d) Identification is generally not performed on-line; it could be mixed with control in an auto-adaptive scheme, but:

—it must be used with great caution because of sensor reliability and non-parametric structural changes of the process.

—the results are biased or may be biased if no external test signals and/or debiasing procedures are applied.

—it is not necessary in most cases because of the natural passive robustness of IDCOM.

—on the contrary, passive adaptation is most often necessary and efficient (e.g. adaptive sampling period in the steam generator case).

(e) In controlling industrial plants the main objective is to reduce the deviation of the variables from their set-points which can be optimized thereupon, through an appropriate hierarchical approach.

(f) The taking-into-account of practical constraints on actions and internal variables and the robustness of control are the important features of industrial dynamic control. The economics depends more on the higher hierarchical levels where optimality may come from than on the dynamic control level.

(g) If we can obtain the black box model of an industrial multivariable process then the latter can be controlled in most cases by an MPHC procedure. The ease of implementation and its amazing robustness make it a convenient and reliable method, capable of standing the adverse industrial conditions.

The economics is clear and the optimal "level 1" dynamic control of a complex multivariable plant does pay back if imbedded in the above-mentioned hierarchical approach. Considering the demands on the conservation of energy and the still-decreasing cost of industrial computers, such dependable implementation may now develop rapidly.

2. *Methodology*

Efficient modeling and fast-computing are new tools; their use broadens the fields of control. If one gets the mathematical input–output descriptive model of a process—whatever this system could be—then through a fast-time heuristic computation of the present and future control variables, if they exist, more complex control problems can be tackled.

The basic ideas underlying this approach are related to the "scenario technique" used in Predictive Economics. To some extent it is also similar to what the human operator is assumed to do with his internal model of the external world.

The control variables are no longer a combination, algebraic or differential, of the measured or observed variables defining the state of a system. MPHC relies on far less restrictive assumptions. Modern computers permit designs which skip unnecessary problems that showed up artificially but necessarily in the past, such as designing a dedicated controller which is of the same functional nature as the process.

Since actions are not computed through a fixed operator from limited local observations but through a totally predictive scheme, stability is not critical and blends into the *robustness analysis*.

Since many tentative future control-variables can be tested and selected according to a criterion, constrained optimization is feasible.

The stress is then on modeling and updating of the identification to avoid the internal model mismatch. Various numerical-analysis problems may arise when computing the optimal controls. That field is open to research, but the control of really complex systems seems more reachable with this fast-time-heuristic-scenario technique.

Acknowledgement—The authors wish to thank Dr. Raman Mehra, the editor, and the referees for their constructive comments in the revision of this paper.

REFERENCES

[1] S. Kaczmarz: Angenäherte Auflösung von Systemen Linearer Gleichungen. *Bull. Intern. Acad. Pol. Sci. Lett. Cl. Sci. Math. Natur.* (1937).

[2] J. Nagumo and A. Noda: A learning method for system identification. *IEEE PGAC* **AC-12**, 281–287 (1967).

[3] F. Lecamus and J. Richalet: Identification des systèmes discrets linéaires multivariables par minimisation d'une distance de structure. *Elect. Lett.* **4**, (24), (1968).

[4] J. Richalet and B. Gimonet: Identification des systèmes discrets linéaires monovariables par minimisation d'une distance de structure. *Elect. Lett.* **4**, (24), (1968).

[5] P. R. Belanger: Comment on a learning method for system identification. *IEEE PGAC* **AC-13**, 207–208 (1968).

[6] I. D. Landau: A hyperstability criterion for model reference adaptive control systems. *IEEE Trans. Aut. Control* **AC-14**, 552–555 (1969).

[7] J. Richalet, F. Lecamus and P. Hummel: New trends in identification minimization of a structural distance: weak topology. *2nd IFAC Symposium on Identification*, Prague (1970).

[8] Y. Z. Tsypkin: *Adaptation and Learning*. Academic Press (1971).

[9] K. J. Aström, P. Eykhoff: System identification, a survey. *IFAC Symposium on Identification and Process Parameter Estimation*, Prague (1970); *Automatica* **7**, 123–162 (1971).

[10] J. Richalet, A. Rault and R. Pouliquen: *Identification des Processus par la Méthode du Modèle*. Gordon & Breach (1971).

[11] D. Turtle and P. Phillipson: Simultaneous identification and control. *Automatica* **7**, 445–453 (1971).

[12] E. Tse and Y. Bar Shalom: Application of adaptive tuning of filters of exo-atmospheric target tracking. *3rd Symposium Nonlinear Estimation Theory*, San Diego (1972).

[13] D. Graupe: *Identification of Systems*. Van Nostrand Reinhold (1972).

[14] A. Feldbaum: *Principles Théoriques des Systémes Asservis Optimaux*. Ed. MIR Moscou (1973).

[15] J. M. Mendel: *Discrete Techniques of Parameter Estimation*. M. Dekker, New York (1973).

[16] D. R. Audley and D. A. Lee: Ill-posed and well-posed problems in system identification. *IEEE Trans. Aut. Control* **AC-19**, 738–748 (1974).

[17] P. Kudva and K. Narendra: An identification procedure for discrete multivariable systems. *IEEE Trans. Aut. Control* **AC-19**, 549–552 (1974).

[18] H. Kurz and R. Isermann: Methods for on-line process identification in closed loop. *6th IFAC World Congress* (1975).

[19] T. Söderström, L. Ljung and I. Gustavson: On the accuracy problem in identification. *6th IFAC World Congress* (1975).

[20] B. Wittenmark: Stochastic adaptive control methods: a survey. *Int. J. Control* (21), (1975).

[21] L. Donaghey: Microcomputer systems for chemical process control. *Proc. IEEE* **64** (1976).

[22] Adersa/Gerbios: IDCOM—Conduite algorithmique des processus industriels de fabrication. *Proc. Conf. on the Algorithmic Control of Industrial Systems*, Paris (1976).

[23] J. Richalet, A. Rault, J. L. Testud, J. Papon: Algorithmic control of industrial processes. *4th IFAC Symposium on Identification and System Parameter Estimation*, Tbilisi (USSR), (1976).

[24] I. Gustavson, L. Ljung and T. Söderström: Identification of processes in closed loops. Identificability and accuracy aspects. *Automatica* **13**, (1). 50 (1977).

[25] G. Hung and L. Stark: Introductory Review—The Kernel identification (1910–1977—Review of theory, calculation, application and interpretation). *Math. Biosci.* (37) 135–190 (1977).

[26] P. R. Latour: The hidden benefits from better process control. *ISA-76*, Houston, Texas (1976).

APPENDIX A: IDENTIFICATION ALGORITHM

1. *Introduction*

On-line identification in a closed-loop system leads to a biased result due to the fact that inputs and outputs are correlated. It has been shown, [9], that without external signals (known set-point variation or perturbation) the identification scheme tends to yield the inverse regulation law.

To get out of this non-identificability condition it is necessary to introduce an external perturbation signal, [23, 24]. It can be proved that the bias depends on the ratio of the energy of the external signal introduced to the energy of the perturbation. The higher this ratio, the smaller is the bias.

Simultaneous identification and control (i.e. complete self-adaptation) not being necessary in industrial processes, the usual conditions under which identification is performed are: either open loop or very weak feedback and high level of external test signal.

The main points of the identification algorithm are presented hereafter. The objective is not to give an exhaustive coverage of an on-line algorithm but rather to emphasize points which are thought to be more original.

2. *Minimization of a structural distance*

The systems to be considered are represented under the following bilinear form. Indices O and M correspond respectively to the actual physical system and its mathematical model

$$s_O(n) = \mathbf{a}_O^{\mathrm{T}}\mathbf{e}(n)$$

$$s_M(n) = \mathbf{a}_M^{\mathrm{T}}\mathbf{e}(n)$$

where $s(n)$ is the observed output at instant n, $\mathbf{a}$ is the *structural vector* whose components are the unknown parameters a_i, $\mathbf{e}(n)$ being the *information vector* at instant n.

In the case of a one-input one-output system, the structural vector $\mathbf{a}$ is formed with the impulse-response parameters $a(i)$ $(i=0,\ldots,N)$ and the information vector is formed with the past inputs $e(n-i)$ $(i=1,\ldots,N)$.

The instantaneous error between object and model is written

$$\varepsilon(n) = (\mathbf{a}_M(n) - \mathbf{a}_O)^{\mathrm{T}}\mathbf{e}(n).$$

A distance can be defined in the parametric space as

$$D_{O,M}(n) = \|\mathbf{a}_M(n) - \mathbf{a}_O\|_P = (\mathbf{a}_M - \mathbf{a}_O)^{\mathrm{T}}P(\mathbf{a}_M - \mathbf{a}_O)$$

where P is a positive definite matrix.

A converging identification algorithm will be one which ensures that

$$D(n+1) - D(n) < 0.$$

D being a Lyapunov function, stability is insured in the parametric space. However it is to be noticed that the state error ε^2 or any quadratic function of ε is not guaranteed to be minimum.

3. *Derivation in the deterministic case*

Let us note

$$\Delta\mathbf{a}_M(n) = \mathbf{a}_M(n) - \mathbf{a}_O$$

$$\Delta\mathbf{a}_M(n+1) = \mathbf{a}_M(n+1) - \mathbf{a}_O = \Delta\mathbf{a}_M(n) + \boldsymbol{\delta}\mathbf{a}(n+1).$$

Thus $D(n) = \Delta\mathbf{a}_M^{\mathrm{T}}(n)P\Delta\mathbf{a}(n)$ and $\varepsilon(n) = \Delta\mathbf{a}_M^{\mathrm{T}}(n)e(n)$ and it follows that

$$D(n+1) - D(n) = 2\,\Delta\mathbf{a}_M^{\mathrm{T}}(n)P\boldsymbol{\delta}\mathbf{a}(n+1) + \boldsymbol{\delta}\mathbf{a}^{\mathrm{T}}(n+1)P\,\boldsymbol{\delta}\mathbf{a}(n+1).$$

The algorithm is determined by the choice of $\boldsymbol{\delta}\mathbf{a}(n+1)$ as a function of the information vector; let it be

$$\boldsymbol{\delta}\mathbf{a}(n+1) = \mu P^{-1}\mathbf{e}(n), \qquad \mu \text{ scalar}.$$

It yields

$$D(n+1) - D(n) = \mu^2\mathbf{e}^{\mathrm{T}}(n)P^{-1}\mathbf{e}(n) + 2\mu\varepsilon(n).$$

Convergence of the algorithm will be ensured if

$$\mu = -\frac{\lambda\varepsilon(n)}{\mathbf{e}^{\mathrm{T}}(n)P^{-1}\mathbf{e}(n)} \qquad \textit{with } 0 < \lambda < 2.$$

The optimal value being $\lambda = 1$, if we look for an optimal minimisation from $D(n+1)$ to $D(n)$.

The on-line algorithm has the following expression

$$\mathbf{a}_M(n+1) = \mathbf{a}_M(n) - \lambda\frac{e(n)P^{-1}\mathbf{e}(n)}{\mathbf{e}^{\mathrm{T}}(n)P^{-1}\mathbf{e}(n)} \tag{1}$$

In the case of a metric distance $D_{O,M}(n) = \|\mathbf{a}_M(n) - \mathbf{a}_O\|$ it becomes

$$\mathbf{a}_M(n+1) = \mathbf{a}_M(n) - \lambda\frac{\varepsilon(n)\mathbf{e}(n)}{\mathbf{e}^{\mathrm{T}}(n)\mathbf{e}(n)} \tag{2}$$

with the corresponding distance variation

$$D(n+1) = D(n) + \lambda(\lambda - 2)\frac{\varepsilon^2(n)}{\mathbf{e}^{\mathrm{T}}(n)\mathbf{e}(n)}.$$

4. *Geometric interpretation*

Given a Euclidean distance $(P = I)$ and $\lambda = 1$, the algorithm can be interpreted as a projection in the parametric space. If $M(n)$ is the point representing the model, O the process, $\mathbf{e}(n)$ the input vector, equation (2) shows that the model at instant $(n+1)$ will be obtained by a fictitious projection of O onto $\mathbf{e}(n)$ as depicted in Fig. 20.

In an N dimensional parametric space, N orthogonal vectors $\mathbf{e}(n)$ would span the space and thus be sufficient to identify. Since $\mathbf{e}(n)$ does not completely span the parametric space, the identification stops in a subspace.

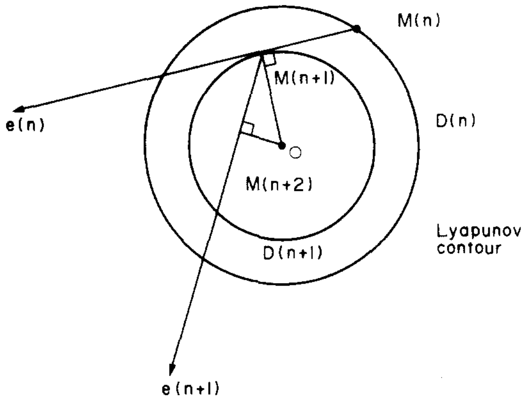


FIG. 20. Structural distance minimisation in the parameter space.

The relaxation factor λ and the P weighting matrix play a major role and may be optimized according to the information contained in the successive input vectors. This basic strategy can be made more sophisticated.

Equations (1) and (2) exhibit the structure of on-line identification algorithms which are fairly universal and found in various fields of applications. In fact, it seems that the original use of such a structure is due to Kaczmarz [1] in 1937. It has been used much as a learning method, [2], [5], [8], but also appears in on-line estimation methods [11], [13], [15], [17].

The interesting contribution in this presentation is that it shows that these types of algorithms minimize a *structural distance* which is the objective of any identification procedure.

5. *Identification in noisy environment*

5.1. *Noise in industrial environment.* Two types of noises are encountered in industrial systems: measurement noise and process noise.

—measurement noises are present on every observed variable (input, output, state variables). For these, the hypothesis of whiteness is generally acceptable.

—process noises correspond to state or structural perturbations. State perturbations take into account every perturbation on the state of the system and in particular secondary inputs not included in the model. Structural perturbations are those of the process itself due to the aging of elements or nonlinear effects not taken into account in the global representation model.

Classically the system is represented in the following form:

The observed process output is: $s_O^{\ddagger}(n) = s_O(n) + v(n)$

The observed past input vector is: $\mathbf{e}^{\ddagger}(n) = \mathbf{e}(n) + \mathbf{b}(n)$

where $\mathbf{b}(n)$ is measurement noise satisfying whiteness hypothesis and $v(n)$ takes into account measurement as well as process noise.

A first approach consists in applying an identical filter F on both measured input and output. Under the hypothesis of linearity and stationarity, process P is still identifiable between e_F and s_F; filters F eliminate noises outside the process bandwidth, mainly measurement noise, and non-zero mean structural noises, drifts.

This justifies that we will restrict the analysis to the case where noise affects only the output of the process.

It can be shown that noise on the input induces a bias in the identification whose computation is quite involved.

5.2. *Analysis of the identification algorithm in noisy environment.* Following the previous remarks, the processes are described by

$$s_O^{\ddagger}(n)=\mathbf{e}^{\mathrm{T}}(n)\mathbf{a}_O+v(n)$$

$$s_M(n)=\mathbf{e}^{\mathrm{T}}(n)\mathbf{a}_M\,.$$

Thus $\varepsilon(n)=\mathbf{e}^{\mathrm{T}}(n)\Delta\mathbf{a}_M(n)-v(n)$.
and the identification algorithm (2) takes the following expression

$$\mathbf{a}_M(n+1)=\mathbf{a}_M(n)-k\mathbf{e}(n)\varepsilon(n) \qquad \text{with } k=\frac{\lambda}{\mathbf{e}^{\mathrm{T}}(n)\mathbf{e}(n)}$$

Notation: In the foregoing the current index n will be omitted and index $n+1$ will be replaced by a superscript $+$

$$\mathbf{a}_M^+=\mathbf{a}_M-k\mathbf{e}[\mathbf{e}^{\mathrm{T}}\,\Delta\mathbf{a}_M-v]$$

$$\mathbf{a}_M^+=\mathbf{a}_M-k\mathbf{e}\mathbf{e}^{\mathrm{T}}\,\Delta\mathbf{a}_M+k\mathbf{e}v. \tag{3}$$

Hypotheses:

H 1 the process is assumed to be open loop

H 2 $E[v(n)]=0 \qquad E[v^2(n)]=\sigma_s^2$.

Subtracting $\mathbf{a}_O$ to each side of equation (3) and taking the mathematical expectation, yields

$$E[\Delta\mathbf{a}_M^+]=E[\Delta\mathbf{a}_M]-k\mathbf{e}\mathbf{e}^{\mathrm{T}}E[\Delta\mathbf{a}]+kE[\mathbf{e}v].$$

Assuming the process noise is uncorrelated with the input $E[\mathbf{e}v]=0$, one obtains

$$E[\Delta\mathbf{a}_M^+]=[I-k\mathbf{e}\mathbf{e}^{\mathrm{T}}]E[\Delta\mathbf{a}_M]. \tag{4}$$

Equation (4) represents the stochastic difference equation of the error term $\Delta\mathbf{a}_M$. Under the condition that the eigenvalues of the operator $[I-k\mathbf{e}\mathbf{e}^{\mathrm{T}}]$ be less than one, $E[\Delta\mathbf{a}_M]$ tends to zero.

Thus, as long as the noise on the output is not correlated with the past inputs, the identification algorithm is unbiased.

5.3. *Variance analysis.* Developing the expression of the error signal ε, the error equation is written as

$$\Delta\mathbf{a}_M^+=[I-k\mathbf{e}\mathbf{e}^{\mathrm{T}}]\Delta\mathbf{a}_M+k\mathbf{e}v.$$

Making use of the hypotheses that the covariance matrix of the estimation error term satisfies the following difference equation

$$E[\Delta\mathbf{a}_M\,\Delta\mathbf{a}_M^{\mathrm{T}}]=\Sigma \qquad E[\Delta\mathbf{a}_M^+\,\Delta\mathbf{a}_M^{+\mathrm{T}}]=\Sigma^+$$

$$\Sigma^+=\Sigma-2k\mathbf{e}\mathbf{e}^{\mathrm{T}}\Sigma+k^2\mathbf{e}\mathbf{e}^{\mathrm{T}}\Sigma\mathbf{e}\mathbf{e}^{\mathrm{T}}+k^2\sigma_s^2\mathbf{e}\mathbf{e}^{\mathrm{T}}. \tag{6}$$

Detailed computation is given in [22].

Taking the trace of matrix equation (6) and making use of the following property $\operatorname{tr}[\mathbf{e}\mathbf{e}^{\mathrm{T}}\Sigma]=\mathbf{e}^{\mathrm{T}}\Sigma\mathbf{e}$ equation (6) becomes

$$\operatorname{tr}\Sigma^+=\operatorname{tr}\Sigma-2k\mathbf{e}^{\mathrm{T}}\Sigma\mathbf{e}+k^2[e^{\mathrm{T}}\Sigma e+\sigma_s^2]\mathbf{e}^{\mathrm{T}}\mathbf{e}. \tag{7}$$

Note that $\operatorname{tr}\Sigma$ is nothing else than the structural distance.

The optimum value of k minimising $E(\mathbf{a}_M-\mathbf{a}_O)^2$ is then

$$k_{\mathrm{opt}}=\frac{1}{\mathbf{e}^{\mathrm{T}}\mathbf{e}}\left[\frac{\mathbf{e}^{\mathrm{T}}\Sigma\mathbf{e}}{\mathbf{e}^{\mathrm{T}}\Sigma\mathbf{e}+\sigma_s^2}\right] \tag{8}$$

The variance $\bar{\varepsilon}^2$ of the error signal is easily computed as

$$\bar{\varepsilon}^2=\mathbf{e}^{\mathrm{T}}\Sigma\mathbf{e}+\sigma_s^2$$

and thus

$$k_{\mathrm{opt}}=\frac{1}{\mathbf{e}^{\mathrm{T}}\mathbf{e}}[1-\sigma_s^2/\bar{\varepsilon}^2].$$

The identification factor varies as the identification proceeds and tends to zero when there is only noise left in the error signal. This is a classical result of stochastic approximation. While σ_s^2 is easily estimated, estimation of $\bar{\varepsilon}^2$ must be made in a stationary state and an underestimation could yield a negative k factor and thus induce a divergence of the algorithm.

The deterministic approach had led to the definition of an identification factor $k=\lambda/(\mathbf{e}^{\mathrm{T}}\mathbf{e})$ with λ called the relaxation factor. The above result shows that λ should be smaller than one and even more so when the signal to noise ratio is bad.

The choice of a constant factor is appropriate in an adaptive situation where the identification algorithm should follow any stationary variations of the physical system.

5.4. *Adaptation of the deterministic algorithm.* If the identification factor is chosen as $k=\lambda/(\mathbf{e}^{\mathrm{T}}\mathbf{e})$, the covariance matrix equation (6) is written as

$$\Sigma^+=\Sigma-2\lambda\frac{\mathbf{e}\mathbf{e}^{\mathrm{T}}}{\mathbf{e}^{\mathrm{T}}\mathbf{e}}\Sigma+\frac{\lambda^2}{(\mathbf{e}^{\mathrm{T}}\mathbf{e})^2}\mathbf{e}\mathbf{e}^{\mathrm{T}}\Sigma\mathbf{e}\mathbf{e}^{\mathrm{T}}+\frac{\lambda^2}{(\mathbf{e}^{\mathrm{T}}\mathbf{e})^2}\sigma_s^2\mathbf{e}\mathbf{e}^{\mathrm{T}}.$$

Assuming the input sequences $\{\mathbf{e}\}$ have been such that the parametric space has been completely spanned, it has been shown (Section 5.2.) that the expected values of the parameters are unbiased.

Under the same conditions, from the previous equation, the final value of the parameter covariance matrix and consequently the expected value of the structural distance will be analysed.

The limit, if it exists, imposes $\Sigma^+=\Sigma$ and thus

$$2\lambda\frac{\mathbf{e}\mathbf{e}^{\mathrm{T}}}{\mathbf{e}^{\mathrm{T}}\mathbf{e}}\Sigma-\frac{\lambda^2}{(\mathbf{e}^{\mathrm{T}}\mathbf{e})^2}\mathbf{e}\mathbf{e}^{\mathrm{T}}\Sigma\mathbf{e}\mathbf{e}^{\mathrm{T}}=\frac{\lambda^2}{(\mathbf{e}^{\mathrm{T}}\mathbf{e})^2}\sigma_s^2\mathbf{e}\mathbf{e}^{\mathrm{T}}, \tag{10}$$

assuming the generated inputs are such that

$\mathbf{e}\mathbf{e}^{\mathrm{T}}=I\mathbf{e}^{\mathrm{T}}\mathbf{e}=I\bar{e}^2$ where I is the identity matrix.

Note that this hypothesis can be satisfied either in open loop or in closed-loop situation by designing convenient test signals.

Equation (10) yields the limit of the covariance matrix as

$$\Sigma=\frac{\lambda\sigma_s^2 I}{(2-\lambda)\bar{e}^2}. \tag{11}$$

Analysis of equation (11) shows that:

—The covariance matrix is diagonal which implies that every parameter has the same variance. Thus if a Euclidean structural distance is minimized, any impulse response is identified within a constant confidence channel.

—The value of the final variance, which is also the expected distance, depends on λ and on the ratio of the noise to signal energy. In particular, should the energy of the input signal tend to zero then the variance on the parameters becomes infinite though their expected value is still unbiased.

In order to cope with this configuration which could happen, in particular, in a self adaptive configuration an identification residual C is introduced. The identification factor k is thus redefined as

$$k = \frac{\lambda}{\mathbf{e}^{\mathrm{T}}\mathbf{e} + C}.$$

C is determined in such a way that when $e \to 0$, the identification being completed, the variance of the parameters is constrained within an uncertainty domain D_O.

This yields the following expression for C

$$C = \frac{\lambda N \sigma_s^2}{2D_O}. \tag{12}$$

Therefore, from the knowledge of the environment and of the admissible variance D_O, the identification residual is computed. It should however be noted that this term slows down the identification in normal operating ($\mathbf{e} \neq 0$) conditions.

APPENDIX B: STABILITY OF THE CONTROL ALGORITHM

In this appendix, for the sake of simplicity, the stability of the control algorithm will be studied in the case of a single output system. The multivariable case has been analyzed in [22].

A stability problem may arise if there is a mismatch between the actual process and its internal model. This could be due to several reasons:

—the identification is not perfect: model characterisation is not appropriate, estimation of the parameters is biased.

—the operating points may change and the system is nonlinear.

—the process is nonstationary: some elements keep aging, raw material and operating conditions keep changing with time.

It is thus important to evaluate the robustness of the algorithm with respect to variations of the process structure.

Let us restrict ourselves to the following case:

—the impulse response of the process $\mathbf{a}_O$ and of the model $\mathbf{a}_M$ are homothetic: $\mathbf{a}_O = q\mathbf{a}_M$ (q scalar).

—the reference trajectory is generated by a first order model.

These hypotheses may seem too restrictive or academic and chosen to ease the computations. In fact they are most commonly encountered in practice, and the stability margin thus derived is significant. Robustness is then analysed through the following steps.

(a) *Updating the reference trajectory.* The reference trajectory is obtained by equation (13)

$$s_{MR}(n+i) = \alpha s_{MR}(n+i-1) + (1-\alpha)C \tag{13}$$

where C is the desired value and α the parameter which fixes the time response of the controlled system.

The reference trajectory is computed through the following step:

$$\begin{aligned}
s_{MR}(n) &= s_O(n)\\
s_{MR}(n+1) &= \alpha s_O(n) + (1-\alpha)C\\
s_{MR}(n+i) &= \alpha s_{MR}(n+i-1) + (1-\alpha)C\\
i &= 1, 2 \ldots HP.
\end{aligned}$$

(b) *Updating the internal model output.* Let us call $s_{MI}(n)$ the output of the internal model. Due to several reasons (model mismatch, state perturbations...) $s_{MI}(n) \neq s_O(n)$: however the output of the actual process may be predicted by equation (14)

$$s_{MR}(n+1) = s_O(n) + \mathbf{a}_M^{\mathrm{T}}\mathbf{e}(n+1) - \mathbf{a}_M^{\mathrm{T}}\mathbf{e}(n). \tag{14}$$

In this way, model mismatch and other errors are essentially removed by using the difference in the model outputs at two different intervals to compute the corresponding change in the actual output rather than compute a new output value directly from the output of the model at a particular instant.

(c) *Computing the control variable.* To follow the reference trajectory $\mathbf{e}(n)$ should be such that $s_O(n+1) = s_{MR}(n+1)$ thus

$$s_{MR}(n+1) = s_O(n) + \mathbf{a}_M^{\mathrm{T}}\mathbf{e}(n+1) - \mathbf{a}_M^{\mathrm{T}}\mathbf{e}(n) \tag{15}$$

if $\mathbf{e}(n+1)$ is applied to the process it yields

$$s_O(n+1) = \mathbf{a}_O^{\mathrm{T}}\mathbf{e}(n+1) = q\mathbf{a}_M^{\mathrm{T}}\mathbf{e}(n+1).$$

Eliminating $s_{MR}(n+1)$ from equations (13) and (15).

$$\alpha s_O(n) + (1-\alpha)C = s_O(n) + \frac{s_O(n+1)}{q} - \frac{s_O(n)}{q}$$

$$s_O(n+1) = (1 - q(1-\alpha))s_O(n) + q(1-\alpha)C. \tag{16}$$

We note that no off-set is observed from equation (16)

$$s_O(n+1) = s_O(n) = C \quad \text{as} \quad n \to \infty.$$

The closed loop system behaves like a first order system with decrement

$$\alpha' = 1 - q(1-\alpha).$$

Stability is insured if the internal model mismatch q is such that

$$|1 - q(1-\alpha)| < 1.$$

If no oscillation response is wanted

$$0 < q < 1/1 - \alpha. \tag{17}$$

It is noticeable from (17) that under these assumptions robustness depends only on the decrement of the reference trajectory and on the internal model mismatch.