SRC TR 89-88

Robust Model Predictive Control of Processes with Hard Constraints

By

E. Zafiriou

Robust Model Predictive Control of Processes with Hard Constraints *

Evanghelos Zafiriou

Chemical Engineering
and Systems Research Center
University of Maryland
College Park, MD 20742

Accepted for publication in
Computers and Chemical Engineering

October 16, 1989

Abstract

A significant number of Model Predictive Control algorithms solve on-line an appropriate optimization problem and do so at every sampling point. The major attraction of such algorithms, like the Quadratic Dynamic Matrix Control, lies in the fact that they can handle hard constraints on the inputs (manipulated variables) and outputs of a process. The presence of such constraints results in an on-line optimization problem that produces a nonlinear controller, even when the plant and model dynamics are assumed linear. This paper provides a theoretical framework within which the stability and performance properties of such algorithms can be studied. Necessary and/or sufficient conditions for nominal and robust stability are derived and two simple examples are used to demonstrate their effectiveness in capturing the nonlinear characteristics of the system. These conditions are also used to analyze simulation results of a 2×2 subsystem of the Shell Standard Control Problem.

1 Introduction

The problem of input saturation is present in almost every chemical system, even when the process dynamics can be assumed linear. In addition to the input constraints, safety and cer-

*Supported in part by the National Science Foundation's Engineering Research Centers Program: NSFD CDR 8803012. Additional support was provided by Shell Development Co. through an unrestricted research grant.

tain performance specifications also require the presence of hard constraints on some output and state variables. The urgency of rigorous theoretical work in this area has been pointed out by the industry (Garcia and Prett, 1986). An approach that has been tried industrially during the past few years is to on-line solve an appropriate optimization problem and to do so at every sampling point. The repeated application of such methods on industrial problems with considerable success (Garcia and Morshedi, 1986; Ricker *et al*, 1989) indicates that sufficient degrees of freedom exist in these formulations and has resulted in their incorporation in academic curricula (see, e.g., Arkun *et al*, 1988). A drawback that has prohibited their widespread use is the fact that no exact tuning procedure for the optimization parameters exist and such tuning often has to be carried out on-line by experienced designers.

The presence of hard constraints in the on-line optimization problem produces a nonlinear controller even when the plant and model dynamics are assumed linear. The fact that the overall control system (plant + controller) is nonlinear makes the study of its properties quite involved, especially since no analytic expression is available for the controller. The problems are compounded when robustness with respect to model-plant mismatch is also considered, because no straighforward extension of the results of the Robust Linear Control Theory to this particular problem exists, even though the plant and model dynamics are assumed linear. Some efforts have been made (Campo and Morari, 1987; Garcia and Prett, 1986) to achieve robustness by modifying the "min" optimization problem that is solved on-line to a "min max" problem that minimizes the objective function over all possible plants. One of the problems of this approach is that either the computations for solving the optimization problem are too time consuming to be carried out on-line at every sample point or to simplify the computations one has to use simplistic model uncertainty descriptions that are unrealistic. Another, quite serious problem is the fact that these methods inherently assume that by solving the "min max" problem to obtain a sequence of future inputs (manipulated variables) and then implementing the first one and repeating the computation at the next sample point, one is guaranteed robust stability and performance, provided that a sufficiently long horizon is used in the objective function. However, feedback from an uncertain plant exists in reality and it is not taken into account in the formulation of the optimization problem, which is an open-loop minimization of the objective function over all possible plants. This fact can lead to performance deterioration and instability. Note that the situation is quite different from studying (and guaranteeing) a stabilizing control algorithm when no model error is present, in which case the assumption is reasonable, although not proven for the general case.

The problems discussed just above, cannot possibly be satisfactorily addressed without considering the problem in its proper nonlinear framework. It is the author's opinion that instead of augmenting the objective functions to add robustness, an action that dramatically increases the computational load and at the same time produces no rigorous robustness guarantees, one should study the problem in its nonlinear nature, obtain conditions that guarantee nominal and robust stability and performance and tune the parameters of the original optimization problems (e.g., Quadratic Dynamic Matrix Control (QDMC)) to satisfy them.

2 Preliminaries

Although Model Predictive Control (MPC) algorithms have been applied to systems with nonlinear dynamic models (Garcia, 1984; Eaton *et al*, 1989), it is usually assumed that the dynamics are linear, the nonlinearity of the problem arising from the hard constraints. The properties of the controller are independent of the type of model description used for the plant (see, e.g., Morari *et al*, 1989). The impulse response description is a convenient one:

$$y(k+1) = H_1 u(k) + H_2 u(k-1) + \ldots + H_N u(k-N+1) \tag{1}$$

where y is the output vector, u is the input vector and N is an integer sufficiently large for the effect of inputs more than N sample points in the past on y to be negligible. The plant is assumed to be open-loop stable, but it may be non-square.

The QDMC-type algorithms (Garcia and Morshedi, 1986; Garcia and Morari, 1985) use a quadratic objective function that includes the square of the weighted norm of the predicted error (setpoint minus predicted output) over a finite horizon in the future (sample points $\bar{k}+1,\ldots,\bar{k}+P$, where $\bar{k}$ is the current sample point) as well as penalty terms on u or Δu:

$$\begin{aligned}\min_{u(\bar{k}),\ldots,u(\bar{k}+M-1)} \sum_{l=1}^{P} [e(\bar{k}+l)^T \Gamma^2 e(\bar{k}+l) + u(\bar{k}+l-1)^T B^2 u(\bar{k}+l-1) \\ + \Delta u(\bar{k}+l-1)^T D^2 \Delta u(\bar{k}+l-1)]\end{aligned} \tag{2}$$

The minimization of the objective function is carried out over the values of $\Delta u(\bar{k})$, $\Delta u(\bar{k}+1)$,..., $\Delta u(\bar{k}+M-1)$, where M is a specified parameter. The minimization is subject to possible hard constraints on the inputs u, their rate of change Δu, the outputs y and other process variables usually referred to as associated variables. The details on the formulation of the optimization problem can be found in Prett and Garcia (1988). After the problem is solved on-line at $\bar{k}$, only the optimal value for the first input vector $\Delta u(\bar{k})$ is implemented and the problem is solved again at $\bar{k}+1$. The optimal $u(\bar{k})$ depends on the tuning parameters of the optimization problem, the current output measurement $y(\bar{k})$ and the past inputs $u(\bar{k}-1)$,..., $u(\bar{k}-N)$ that are involved in the model output prediction. Let f describe the result of the optimization:

$$u(k) = f(y(k), u(k-1), \ldots, u(k-N), r_P(k)) \tag{3}$$

where $r_P(k)$ includes all the values of the reference signal (setpoint) during the prediction horizon from $k+1$ to $k+P$.

The optimization problem of the QDMC algorithm can be written as a standard Quadratic Programming problem:

$$\min_v q(v) = \frac{1}{2} v^T G v + g^T v \tag{4}$$

subject to

$$A^T v \geq b \tag{5}$$

where

$$v = [\ \Delta u(\bar{k}) \ \ \ldots \ \ \Delta u(\bar{k}+M-1)\ ]^T \tag{6}$$

and the matrices G, A, and vectors g, b are functions of the tuning parameters (weights, horizon P, M, some of the hard constraints). The vectors g, b are also linear functions of $y(\bar{k})$, $u(\bar{k}-1)$,..., $u(\bar{k}-N)$. For the optimal solution v^* we have (Fletcher, 1981):

$$\begin{bmatrix} G & -\hat{A} \\ -\hat{A}^T & 0 \end{bmatrix} \begin{bmatrix} v^* \\ \lambda^* \end{bmatrix} = - \begin{bmatrix} g \\ \hat{b} \end{bmatrix} \tag{7}$$

where $\hat{A}^T$, $\hat{b}$ consist of the rows of A^T, b that correspond to the constraints that are active at the optimum and λ^* is the vector of the Lagrange multipliers corresponding to these constraints. The optimal $\Delta u(\bar{k})$ corresponds to the first m elements of the v^* that solves (7), where m is the dimension of u.

The special form of the LHS matrix in (7) allows the numerically efficient computation of its inverse in a partitioned form (Fletcher, 1981):

$$\begin{bmatrix} G & -\hat{A} \\ -\hat{A}^T & 0 \end{bmatrix}^{-1} = \begin{bmatrix} H & -T \\ -T^T & U \end{bmatrix} \tag{8}$$

Then

$$v^* = -Hg + T\hat{b} \tag{9}$$

$$\lambda^* = T^T g - U\hat{b} \tag{10}$$

and

$$u(\bar{k}) = u(\bar{k}-1) + \underbrace{\begin{bmatrix} I & 0 & \dots & 0 \end{bmatrix}}_{M} v^* \stackrel{\text{def}}{=} f(y(\bar{k}), u(\bar{k}-1), \dots, u(\bar{k}-N), r_P(k)) \tag{11}$$

3 Formulation of the Problem as a Contraction Mapping

The framework selected for the study of the properties of the overall nonlinear system is that of the Operator Control Theory (Economou, 1985). In this approach, the stability and performance of the nonlinear system can be studied by applying the contraction mapping principle on the operator F that maps the "state" of the system (plant + controller) at sample point k to that at sample point $k+1$. The fact that the plant dynamics are assumed linear allows us to obtain results and carry out computations that are not yet feasible in the general case. We can define as the "state" of the system at sample point k the following vector

$$x(k) = \begin{bmatrix} x_1(k) \\ \vdots \\ x_N(k) \end{bmatrix} \tag{12}$$

where

$$
\begin{array}{llcl}
x_1(k+1) & \stackrel{\text{def}}{=} u(k) & = & f(y(k), u(k-1), \ldots, u(k-N)) \\
 & & = & f(H_1 u(k-1) + \ldots + H_N u(k-N) + d(k), \\
 & & & u(k-1), \ldots, u(k-N), r_P(k)) \\
 & & \stackrel{\text{def}}{=} & \Psi(u(k-1), \ldots, u(k-N); r_P(k), d(k)) \\
 & & = & \Psi(x(k); r_P(k), d(k)) \\
x_2(k+1) & \stackrel{\text{def}}{=} u(k-1) & = & x_1(k) \\
\vdots & & \vdots & \\
x_N(k+1) & \stackrel{\text{def}}{=} u(k-N+1) & = & x_{N-1}(k)
\end{array}
\tag{13}
$$

and $d(k)$ represents the disturbance effect at k on the plant output. The "state" vector $x(k)$ is defined so that knowledge of it and of the external inputs $r_P(k)$ and $d(k)$ allows the computation of $x(k+1)$ by applying the plant and controller equations on it. Indeed the operator F that maps $x(k)$ to $x(k+1)$ is given by

$$
x(k+1) = F(x(k); r_P(k), d(k)) = \begin{bmatrix} \Psi(x(k); r_P(k), d(k)) \\ x_1(k) \\ \vdots \\ x_{N-1}(k) \end{bmatrix} \tag{14}
$$

Note, however, that although $f(.)$ can be computed, since it describes the on-line optimizing control algorithm and it involves only the process model, Ψ is not exactly known, because it involves the "true" plant impulse response coefficients H_1,..., H_N.

Convergence of the successive substitution $x(k+1) = F(x(k))$ implies stability of the overall nonlinear system; fast convergence implies good performance. The use of the contraction mapping principle allows the development of conditions for robust stability and performance in terms of some consistent matrix norm [1] of the derivative F' of the above operator F.

4 Stability Conditions

We shall now proceed to obtain stability conditions for the overall nonlinear system by obtaining conditions under which the mapping described by F is a contraction. The terms stability and instability of the control system are used in the global asymptotical sense over the domain of F under consideration.

Let us first examine the differentiability of F. From (14) it follows that this is equivalent to differentiability of $\Psi(x)$ and from (13) to differentiability of f. Let us assume that for some point x in the domain of F, an infinitesimal change in x (which results in a change of g, b in (4), (5)) does not change $\hat{A}$, i.e., the set of active constraints at the optimum does not change (note that A is independent of x). Then from (7) it follows that the derivative of Ψ exists and it has a constant value in a neighbourhood of that x.

[1] A consistent matrix norm $||.||$ is a norm for which there exists a vector norm $|.|$ such that $|Ax| \leq ||A|| |x|$. Such a norm has the property $||AB|| \leq ||A|| ||B||$.

Let J_i be a set of indices for the active constraints of (4) and $J_1,\ldots, J_n$ correspond to all possible active sets of constraints when all xs in the domain of F are considered. Every such J_i corresponds to an $\hat{A}_i$ and a $\hat{b}_i$. Then from the above discussion, it is evident that for all xs that correspond to the same J_i and for which an infinitesimal change in their value does not change the set of active constraints, the derivative of Ψ and therefore of F exist and it has the same value that depends on the particular set J_i:

$$F'_{J_i} = \begin{bmatrix} (\nabla_{x_1}\Psi)_{J_i} & (\nabla_{x_2}\Psi)_{J_i} & \cdots & (\nabla_{x_{N-1}}\Psi)_{J_i} & (\nabla_{x_N}\Psi)_{J_i} \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix} \tag{15}$$

where from (13) it follows that

$$(\nabla_{x_j}\Psi)_{J_i} = (\nabla_{x_j}f)_{J_i} + (\nabla_y f)_{J_i} H_j \tag{16}$$

The derivatives of f can be computed easily from (11):

$$(\nabla_{x_j}f)_{J_i} = \begin{bmatrix} I & 0 & \ldots & 0 \end{bmatrix} (-H_{J_i}\nabla_{x_j}g + T_{J_i}\nabla_{x_j}\hat{b}_i) \tag{17}$$

where the derivatives of g, b_i are constant since g, b are linear functions of $y(\bar{k})$, $u(\bar{k}-1),\ldots,$ $u(\bar{k}-N)$. The same expression as in (17) is also true for the derivative with respect to $y(\bar{k})$, the current measurement. Also note that in the case of x_1, the identity matrix I should be added to the RHS of (17).

It is clear from the above discussion that $F(x)$ is piecewise linear and that it is differentiable everywhere except the points where an infinitesimal change will change the set of active constraints at the optimum of (4). It follows then that for F to be a contraction, it is necessary that

$$||F'_{J_i}|| \leq \theta < 1, \quad i = 1,\ldots,n \tag{18}$$

where $||.||$ is some induced matrix norm[2], the same for all i. Since an induced norm is a consistent norm as well, it is necessary that there exists some consistent matrix norm for which (18) is satisfied. Such a condition, however, can be shown to be sufficient as well. Consider two points x^a, x^b and let the straight path connecting them in the domain of F be broken into the succesive segments $x^a \to x^1$, $x^1 \to x^2,\ldots,$ $x^l \to x^b$, the points of each of which correspond to the same J_i: J_{k_0}, $J_{k_1},\ldots,$ J_{k_l}, respectively. Then, if $|.|$ is the vector norm that is consistent with $||.||$, we have

$$\begin{aligned} |F(x^a) \; - \; & F(x^b)| \\ = \; & |(F(x^a) - F(x^1)) + (F(x^1) - F(x^2)) + \ldots + (F(x^l) - F(x^b))| \\ = \; & |F'_{J_{k_0}}(x^a - x^1) + F'_{J_{k_1}}(x^1 - x^2) + \ldots + F'_{J_{k_l}}(x^l - x^b)| \\ = \; & |(a_0 F'_{J_{k_0}} + a_1 F'_{J_{k_1}} + \ldots + a_l F'_{J_{k_l}})(x^a - x^b)| \end{aligned}$$

[2]An induced matrix norm is a norm for which there exists a vector norm $|.|$ such that $||A|| = \sup_{x\neq 0}|Ax|/|x|$. An induced norm is also a consistent norm.

$$
\begin{aligned}
&\leq a_0|F'_{J_{k_0}}(x^a - x^b)| + a_1|F'_{J_{k_1}}(x^a - x^b)| + \ldots + a_l|F'_{J_{k_l}}(x^a - x^b)| \\
&\leq (a_0||F'_{J_{k_0}}|| + a_1||F'_{J_{k_1}}|| + \ldots + a_l||F'_{J_{k_l}}||)|(x^a - x^b)| \\
&\leq (a_0 + a_1 + \ldots + a_l)\theta|x^a - x^b| \\
&= \theta|x^a - x^b| \qquad (19)
\end{aligned}
$$

where a_j is the relative length of the respective segment as compared to $x^a \to x^b$. From (19) it follows that F is a contraction. The fact that there is only a finite number of J_is allows us to drop the θ from (18) to obtain:

Theorem 1 *F is a contraction if and only if there exists a consistent matrix norm $||.||$, for which*

$$||F'_{J_i}|| < 1, \quad i = 1, \ldots, n \qquad (20)$$

The practical use of (20) is limited by the fact that finding an appropriate consistent norm is not a trivial task. The following two subsections provide conditions which are more readily computable. The third subsection formulates the respective robustness conditions.

4.1 Sufficient Condition

By selecting one particular consistent matrix norm and stating (20) for that norm, one can get a sufficient only condition.

Let us select the following norm, which can be shown to be a consistent one on $\mathbf{R}^{mN \times mN}$ (Stewart, 1973), where m is the plant dimension:

$$||A|| = ||DAD^{-1}||_\infty \qquad (21)$$

where

$$
\begin{aligned}
&||B||_\infty = \max_i \sum_{j=1}^{N} |b_{ij}| && (22)\\
&D = \mathrm{diag}(I, \eta I, \eta^2 I, \ldots, \eta^{N-1} I) && (23)\\
&0 < \eta < 1 && (24)
\end{aligned}
$$

Then

$$
DF'_{J_i}D^{-1} =
\begin{bmatrix}
(\nabla_{x_1}\Psi)_{J_i} & (\nabla_{x_2}\Psi)_{J_i}\eta^{-1} & \ldots & (\nabla_{x_{N-1}}\Psi)_{J_i}\eta^{-(N-2)} & (\nabla_{x_N}\Psi)_{J_i}\eta^{-(N-1)} \\
\eta I & 0 & \ldots & 0 & 0 \\
0 & \eta I & \ldots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \ldots & \eta I & 0
\end{bmatrix}
\qquad (25)
$$

From (21)-(25) we get

$$
\begin{aligned}
& ||DF'_{J_i}D^{-1}||_\infty < 1 \\
\Leftrightarrow \; & || \; (\nabla_{x_1}\Psi)_{J_i} \;\; (\nabla_{x_2}\Psi)_{J_i}\eta^{-1} \;\; \ldots \;\; (\nabla_{x_N}\Psi)_{J_i}\eta^{-(N-1)} \; ||_\infty < 1 \\
\Leftarrow \; & || \; (\nabla_{x_1}\Psi)_{J_i} \;\; (\nabla_{x_2}\Psi)_{J_i} \;\; \ldots \;\; (\nabla_{x_N}\Psi)_{J_i} \; ||_\infty < \eta^{N-1}
\end{aligned} \tag{26}
$$

Since any η in $(0,1)$ will do and there is only a finite number of J_is, from (26) we can obtain:

Theorem 2 *The control system is asymptotically stable if*

$$|| \; (\nabla_{x_1}\Psi)_{J_i} \;\; (\nabla_{x_2}\Psi)_{J_i} \;\; \ldots \;\; (\nabla_{x_N}\Psi)_{J_i} \; ||_\infty < 1, \quad i = 1, \ldots, n \tag{27}$$

Note that for single-input single-output plants (27) becomes

$$\sum_{j=1}^{N} \left| \frac{\partial \Psi_{J_i}}{\partial x_j} \right| < 1, \quad i = 1, \ldots, n \tag{28}$$

which for the unconstrained case is simply a sufficient condition for the closed-loop poles to lie inside the Unit Circle.

4.2 Instability Conditions

For every consistent matrix norm we have

$$\rho(A) \leq ||A|| \tag{29}$$

where $\rho(A)$ is the spectral radius of A, defined as $\rho(A) = \max_j |\lambda_j(A)|$, $\lambda_j(A)$, being the eigenvalues of A. Then from (20) and (29) we get

Theorem 3 *F can be a contraction only if*

$$\rho(F'_{J_i}) < 1, \quad i = 1, \ldots, n \tag{30}$$

Note that if the optimization (4) is not subject to (5), then $n = 1$ and (30) becomes sufficient as well, because, given a matrix one can always find a consistent norm arbitrarily close to its spectral radius (Stewart, 1973). The reason that (30) is not sufficient in general is that such a consistent norm is in general a different one for two different matrices (different J_is), while (20) requires the same norm for all i. In the case of $n = 1$, (30) translates to the requirement that the closed-loop poles of the system are located inside the Unit Circle.

If (30) is not true, then F is not a contraction. This however does not necessarily imply that the control system is unstable. The following theorem provides a condition that is sufficient for instability.

Theorem 4 *The control system is unstable if*

$$\rho(F'_{J_i}) > 1, \quad i = 1, \ldots, n \tag{31}$$

The proof follows the argument that if a stable local equilibrium point existed, then for the J_i corresponding to that point we would have $\rho(F'_{J_i}) < 1$.

Theorem 4 can be used to predict instability of the overall nonlinear system. Theorem 3 on the other hand does not seem at a first glance to be of much use, since violation of (30) does not necessarily imply instability. From a practical point of view, however, violation of that condition for some i, should be taken as a very serious warning that the control system parameters should be modified. The reason is that when in the region of the domain of F that corresponds to that i, the system will behave as a virtually unstable system, the only hope for stability being to move to a region with $\rho(F'_{J_i}) < 1$. It might be the case that for a particular system in question this will always happen, making this system a stable one. But even in this case, a temporary unstable-like behavior might occur, thus making the control algorithm practically unacceptable. The examples in Section 8 demonstrate situations where violation of (30) is enough to produce an unstable system although (31) is not satisfied.

4.3 Robustness Conditions

From (16) we see that F'_{J_i} depends on the impulse response coefficient matrices $H_1, \ldots, H_N$ of the actual plant. These matrices are never known exactly and so in order to guarantee stability for the actual plant, one has to compute the conditions of Sections 4.2, 4.1, not just for the model, but for all possible plants. To do so, one needs to have some information on the possible modeling error associated with the H_is. Let $\mathcal{H}$ be the set of possible values for these coefficients. Then we can write the following conditions:

Theorem 5 *The control system is asymptotically stable for all plants with coefficients in $\mathcal{H}$ if*

$$\sup_{\mathcal{H}} \| \; (\nabla_{x_1}\Psi)_{J_i} \quad (\nabla_{x_2}\Psi)_{J_i} \quad \ldots \quad (\nabla_{x_N}\Psi)_{J_i} \; \|_\infty < 1, \quad i = 1, \ldots, n \tag{32}$$

Theorem 6 *F can be a contraction for all plants with coefficients in $\mathcal{H}$ only if*

$$\sup_{\mathcal{H}} \rho(F'_{J_i}) < 1, \quad i = 1, \ldots, n \tag{33}$$

5 A Robust Linear Control Stabilization Interpretation of the Necessary Conditions

The following re-formulation of the necessary conditions of the previous section, allows us to bypass the problem of dealing with uncertainty in the Hs directly, and use the tools that were developed for Robust Linear Control (e.g., the structured singular value (Doyle, 1982)) to treat any of the types of model error that can be handled by that theory. Consider a standard feedback controller $C(z)$. Then

$$u(z) = C(z)(r(z) - y(z)) \tag{34}$$

where r is the setpoint vector. Define

$$C_{J_i}(z) \stackrel{\text{def}}{=} -\left[I - (\nabla_{x_1} f)_{J_i} z^{-1} - \ldots - (\nabla_{x_N} f)_{J_i} z^{-N}\right]^{-1} (\nabla_y f)_{J_i} \tag{35}$$

Since the plant is assumed to be open-loop stable, for stability of this linear control system we need that the closed-loop transfer function between u and r or d (disturbance at the plant output) be stable. From (34), (35) we get by using (1)

$$u(z) = -\left[I - (\nabla_{x_1}\Psi)_{J_i}z^{-1} - \ldots - (\nabla_{x_N}\Psi)_{J_i}z^{-N}\right]^{-1}(\nabla_y f)_{J_i}r(z) \tag{36}$$

where $(\nabla_{x_j}\Psi)_{J_i}$ is given by (16). Hence, stability of the linear unconstrained system under feedback control $C_{J_i}(z)$ is equivalent to stability of the transfer matrix in (36), which, assuming that $(\nabla_y f)_{J_i}$ is full rank, is itself equivalent to (30) since F'_{J_i} is the companion matrix of the denominator of (36). For the case where $(\nabla_y f)_{J_i}$ is not full rank, stabilty under $C_{J_i}(z)$ is still necessary for satisfaction of (30). For J_is, however, for which $(\nabla_y f)_{J_i} = 0$, (30) is equivalent to requiring that the transfer matrix

$$Q_{J_i}(z) \stackrel{\text{def}}{=} \left[I - (\nabla_{x_1}f)_{J_i}z^{-1} - \ldots - (\nabla_{x_N}f)_{J_i}z^{-N}\right]^{-1} \tag{37}$$

be stable. Note that $Q_{J_i}(z)$ is independent of the "uncertain" plant coefficients H_1, ..., H_N.

Hence, from the above discussion we have

Theorem 7 *F can be a contraction only if all feedback controllers $C_{J_i}(z)$, $i \ni (\nabla_y f)_{J_i} \neq 0$, produce a stable system when applied to the unconstrained process and all transfer matrices $Q_{J_i}(z)$, $i \ni (\nabla_y f)_{J_i} = 0$, are stable.*

Theorem 8 *F can be a contraction for all plants in a set Π, only if all feedback controllers $C_{J_i}(z)$, $i \ni (\nabla_y f)_{J_i} \neq 0$, stabilize all plants in the set Π and all transfer matrices $Q_{J_i}(z)$, $i \ni (\nabla_y f)_{J_i} = 0$, are stable.*

The advantage of Thm. 8 over Thm. 6 lies in the fact that through Thm. 8 we can handle any set Π that Robust Linear Control theory can (for a discussion of the possible Πs see Morari and Zafiriou, 1989). This new interpretation of the conditions also indicates that robust performance conditions can be formulated for the same set of feedback controllers. For the sufficient conditions a similar formulation may be possible but it would probably involve some conservativeness.

6 Practical Relevance of a Condition Violation

Conditions (30), (27) can be used to examine the nominal stability of the system for a particular set of tuning parameters. An important question is what are the implications if for a particular $\hat{A}_i$ the conditions are not satisfied. This would only be relevant if the particular combination of active constraints at the optimum can actually occur during the operation of the control system. The following is a procedure that can decide if a certain set of active constraints at the optimum is relevant.

Let $\check{A}^T$, $\check{b}$ consist of the rows of A^T, b that correspond to the inactive constraints at the optimum. Then by using (9), (10) we see that in order for such a combination to be possible at the optimum we need to have

$$\check{A}^T(-Hg + T\hat{b}) \geq \check{b} \tag{38}$$

$$T^T g - U\hat{b} \geq 0 \tag{39}$$

Since g, b are linear combinations of the past manipulated variables and the current measurement, (38), (39) can be combined with the hard constraints on the past us, the past Δus and the output $y(\bar{k})$ to constitute a system of linear inequalities that have to have a feasible solution over the values of the past inputs and the current measurement. Note that depending on the estimate of expected disturbances, one may wish to modify the bounds on $y(\bar{k})$ that are used in the above problem. If the problem has no feasible solution, then the fact that for that particular $\hat{A}$ the stability conditions are not satisfied, is of no practical importance.

Note that the above procedure can also serve to construct a sequence of possible past inputs that can lead to a situation during the operation of the control system where the stability conditions are not satisfied.

7 Analysis of Simulation Results

The computation of the stability conditions at all possible combinations of active constraints at the optimum of the on-line optimization problem can be extremely time-consuming and therefore a systematic method that does not have to check all possibilities is needed. Since no such method for checking the conditions is currently available, the following procedure for providing the designer with insights on tuning the controller parameters can be used.

For a given set of values for the tuning parameters, the designer can simulate the overall system for certain disturbances and/or setpoints that he considers of practical relevance. Such simulations can show instability or simply bad behavior at certain points during the simulation. This behavior which stops short of instability might be captured as a violation of condition (30) which is necessary for F to be a contraction. By computing these conditions at every sampling point during the simulation and by studying the robustness properties of the C_{J_i}s that correspond to the points where the conditions were violated, the designer may be able to improve the tuning parameters. The above procedure is used in the study of a realistic process in section 8.3.

8 Illustrations

The first two examples demonstrate the effectiveness of the nominal and robust stability conditions in capturing the nonlinear behavior of the control system. These two examples are simple so that the effect of incuding hard constraints in the on-line optimization problem is clear. They are not meant as difficult to control systems.

The third example is based on a system described by industrial practitioners. It is intended to demonstrate that even at this early stage in the development of the theoretical framework, the available conditions can be useful in analyzing the behavior of a rather complex process.

8.1 Nominal Stability of a 2×2 process

Let us consider a system with the following transfer function:

$$P(s) = \begin{bmatrix} \frac{1}{s+1} & 0 \\ \frac{-2e^{-5s}}{s+1} & \frac{-s+2}{(s+2)(s+1)} \end{bmatrix} \tag{40}$$

A sampling time $T = 0.5$ is used and the following objective function is minimized on-line:

$$\min_{u(\bar{k}),\ldots,u(\bar{k}+M-1)} \sum_{l=1}^{P} \left[e(\bar{k}+l)^T \Gamma^2 e(\bar{k}+l) + u(\bar{k}+l-1)^T B^2 u(\bar{k}+l-1) \right] \tag{41}$$

where $\bar{k}$ is the current sample point, e is the predicted difference between the setpoints and the plant outputs and Γ, B, are weights.

$$\Gamma = \begin{bmatrix} 1 & 0 \\ 0 & 0.5 \end{bmatrix} \tag{42}$$

is selected signifying that the first output is more important than the second.

Let us first consider the unconstrained problem. First we select $P = M = 2$, which is a selection that is expected (Garcia and Morari, 1982, 1985) to produce an unstable control system if $B = 0$. The reason is the right-half plane (RHP) zero of $P(s)$. Indeed, one can easily check that for these values of the tuning parameters, we have $\rho(F'_{J_1}) > 1$, where J_1 corresponds to the case where no constraints are active at the optimum. Hence the necessary condition (30) predicts the instability. From theory (Garcia and Morari, 1985) we know that by making B sufficiently large, we can stabilize the system. Indeed by making

$$B = \begin{bmatrix} 0 & 0 \\ 0 & 0.1 \end{bmatrix} \tag{43}$$

the system is stabilized ($\rho(F'_{J_1}) < 1$, which is sufficient for $n = 1$). The fact that the RHP zero is pinned to the second plant output, made it unnecessary to increase the 1,1 element of B. The response to a unit step change in setpoint 1 is shown in Fig. 1. The steady-state offset in output 2 is expected from theory and can be avoided by modifying the control algorithm, but we will not do so to avoid the unnecessary complication of the example.

Let us now assume that after looking at the response, the designer decides that a slight tightening of the specifications is in order, namely the addition in the optimization problem of a lower bound on output 2 at the value -0.9. Since output 2 only slightly violated this bound when the unconstrained algorithm was used, one might think that the response for the constrained algorithm should be almost the same as that in Fig. 1. This is not so, however. The response for the same setpoint change is shown in Fig. 2. The system is unstable! An instability warning was issued by the necessary condition for F to be a contraction (30), since $\rho(F'_{J_2}) > 1$, where J_2 corresponds to the case where the low constraint on output 2 is active at the optimum. Indeed by looking at a close-up of Fig. 2 in Fig. 3, we see that the system went unstable as soon as output 2 reached the low bound to which the on-line minimization was subject. The constraint remained active at the subsequent sample points and the system never stabilized.

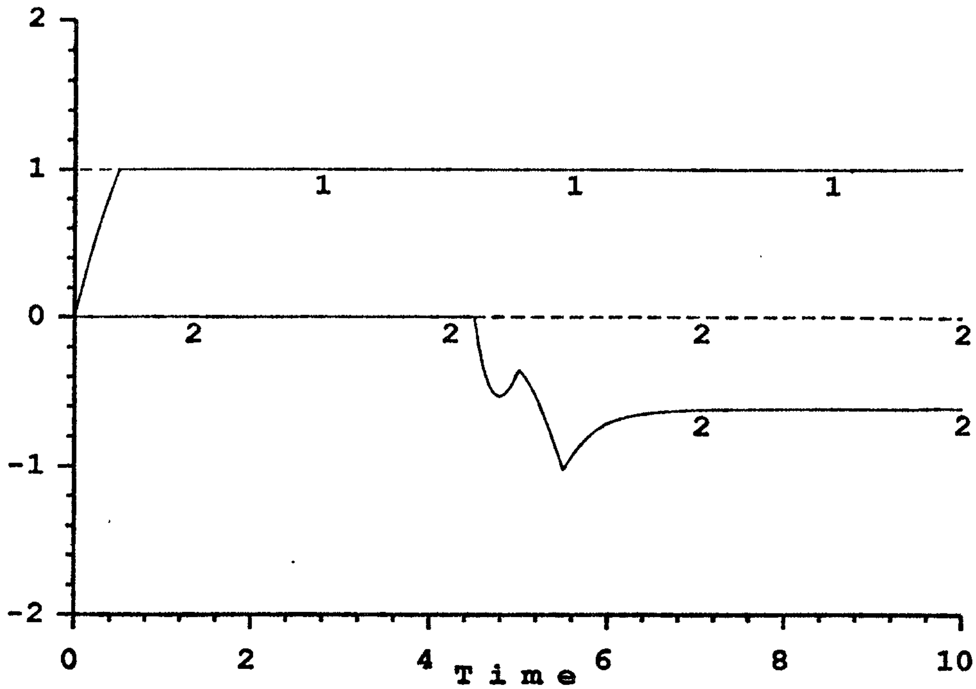

Figure 1: Example 1: Unconstrained minimization.

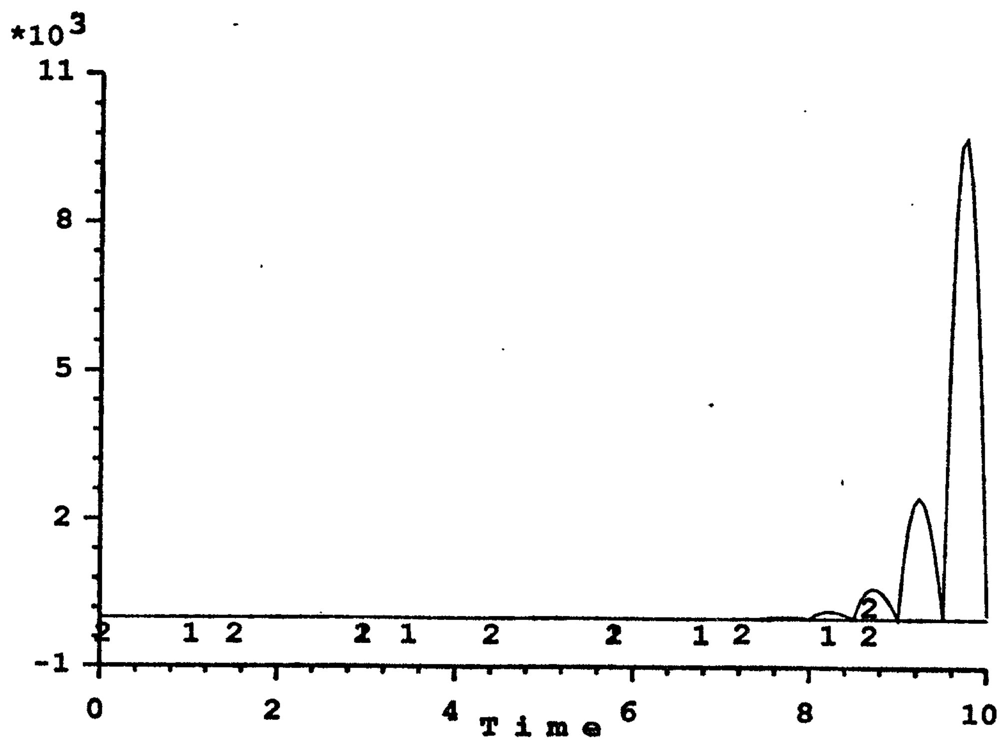

Figure 2: Example 1: Minimization subject to lower bound constraint on output 2.

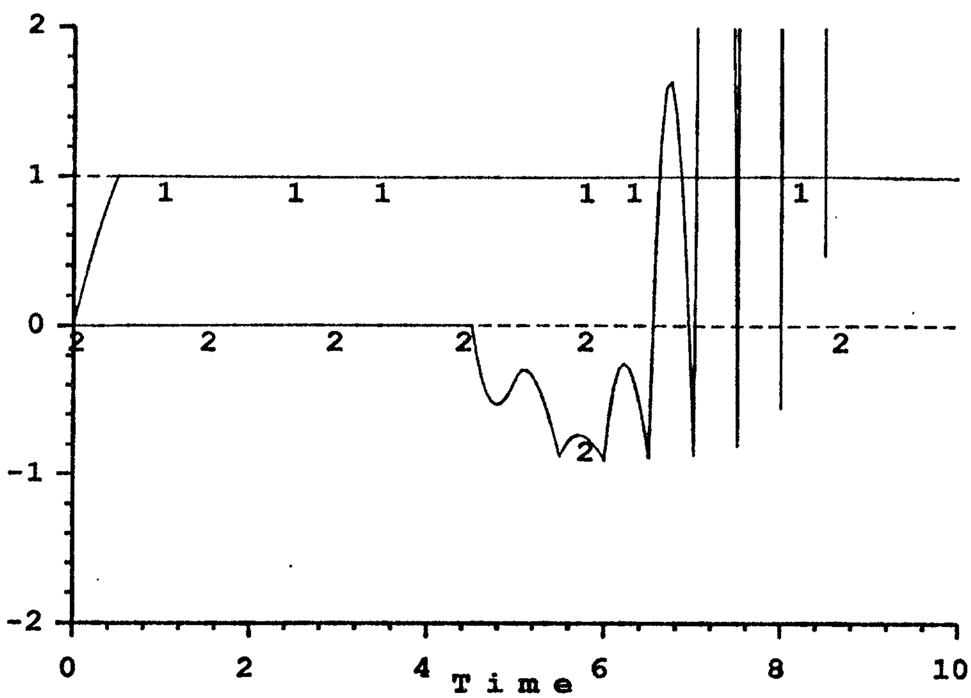

Figure 3: Close-up of Fig. 2.

A question that one may ask at this point is whether the use of a

$$B = \begin{bmatrix} 0 & 0 \\ 0 & \beta \end{bmatrix} \tag{44}$$

with a β larger than the previously used value of 0.1, will stabilize the system. We know that this would indeed be the case for the unconstrained problem; however, for the constrained case that does not happen. By examining the analytic expression for F'_{J_2} one sees that β does not even appear in it and can therefore in no way influence the stability of the system when the constraint becomes active. When the constraint is reached, the algorithm puts as its higher priority keeping output 2 above the lower bound and to do so it inverts the 2,2 element of $P(s)$ and causes instability.

Finally, one should note that this is a very simple example, used to demonstrate the nonlinear behavior of the constrained controller. It is not a difficult system to control. Putting a weight on the first element of B will stabilize the system. This, however is not needed for the unconstrained controller, since the RHP zero is pinned in the 2,2 element of $P(s)$.

8.2 Robust Stability of a SISO process

Consider the process model

$$\tilde{p}(s) = \frac{1}{s+1} \tag{45}$$

A sampling time $T = 0.1$ will be used and the control algorithm will minimize on-line the objective function

$$\min_{u(\bar{k}),\ldots,u(\bar{k}+M-1)} \sum_{l=1}^{P} \left[e(\bar{k}+l)^T \Gamma^2 e(\bar{k}+l) + \Delta u(\bar{k}+l-1)^T D^2 \Delta u(\bar{k}+l-1) \right] \tag{46}$$

To allow the analytic study of the properties of the control system we shall choose the parameters to be $P = M = \Gamma = 1$. A choice of $D = 0$, when there are no hard constraints, will result in an IMC controller that inverts the model (Garcia and Morari, 1982).

Let us now consider a model-plant mismatch caused by a delay term in the plant:

$$p(s) = \frac{e^{-0.15s}}{s+1} \tag{47}$$

For this plant, robust linear control theory can easily show that the control system will be unstable for $D = 0$. D has to be increased to over $D = 0.2$ to stabilize it. The choice $D = 0.4$ results in reasonable performance.

Our interest in this example has to do with the effect of hard constraints on its output. Let us specify a lower bound of -1 and an upper bound of $+1$ for y and include these constraints in the on-line optimization problem. Since the horizon $P = 1$, it is not possible for both to be active at the optimum. In this case $n = 3$, corresponding to (i) no active constraints, (ii) upper constraint active, (iii) lower constraint active. Analytic computaion of $c_{J_i}(z)$, $i = 1, 2, 3$, results in the expressions

$$\begin{aligned} c_{J_1}(z) = H_1/[\ & (D^2 + H_1^2) + (H_1H_2 - H_1^2 - D^2)z^{-1} \\ & + H_1(H_3 - H_2)z^{-2} + \ldots + H_1(H_N - H_{N-1})z^{-N+1} - H_1H_Nz^{-N}] \end{aligned} \tag{48}$$

$$\begin{aligned} c_{J_2}(z) = c_{J_3}(z) = 1/[\ & H_1 + (H_2 - H_1)z^{-1} \\ & + (H_3 - H_2)z^{-2} + \ldots + (H_N - H_{N-1})z^{N+1} - H_Nz^{-N}] \end{aligned} \tag{49}$$

One can easily see from these exressions that c_{J_2} and c_{J_3} correspond to an IMC controller that inverts the process model, the same as c_{J_1} for $D = 0$. The difference is that D does not appear in (49) and therefore this controller will be unstable when the model-plant mismatch is present. The question that arises now, is the one discussed in Section 6. For the case of the upper constraint and for a setpoint equal to zero, (39) predicts that if the system is at equilibrium, a disturbance of magnitude greater than 1.06 will result in an on-line optimization where the upper constraint is active. The system could however manage to return to the contraction region of no active constraints. Indeed for a disturbance of 1.70, as Fig. 4 shows, the system is still stable, although at the edge of instability. An increase of the disturbance to 1.75 however results in an unstable system as Fig. 5 shows. Note that $D = 0.4$ is being used; although D does not appear in (49), it does play a role on whether the constraints are active at the optimum. Both simulations use the plant of (47).

Let us now remove the constraints from the optimization problem and repeat the simulation for the same $d = 1.75$ and $D = 0.4$. The result is shown in Fig. 6. The response is reasonable and the constraints are virtually satisfied, although they were removed from the optimization problem. This example is not meant to suggest that output constraints should never be included in the optimization, but merely to point out that their effect should be studied carefully before their inclusion and to demonstrate that the stability conditions that were provided in this paper can predict this effect successfuly.

8.3 Robust Stability of a Heavy Oil Fractionator

In this section we will use the conditions of section 5 to analyze simulation results for the Shell Standard Control Problem (SSCP) (Prett and Garcia, 1988). Let us consider the top

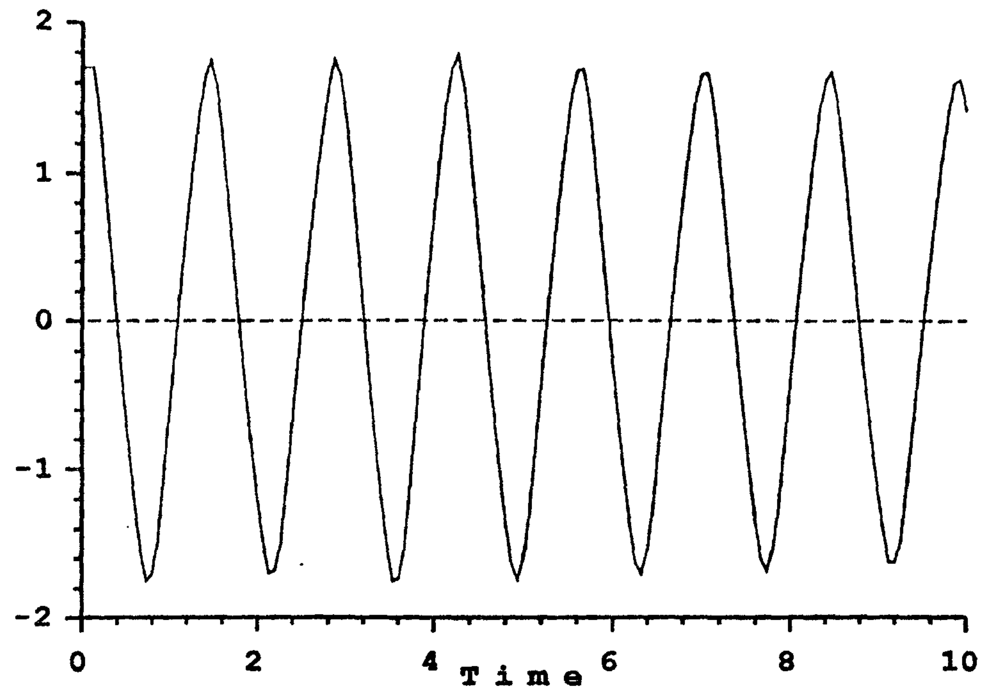

Figure 4: Example 2: Constrained; D=0.4 and d=1.70

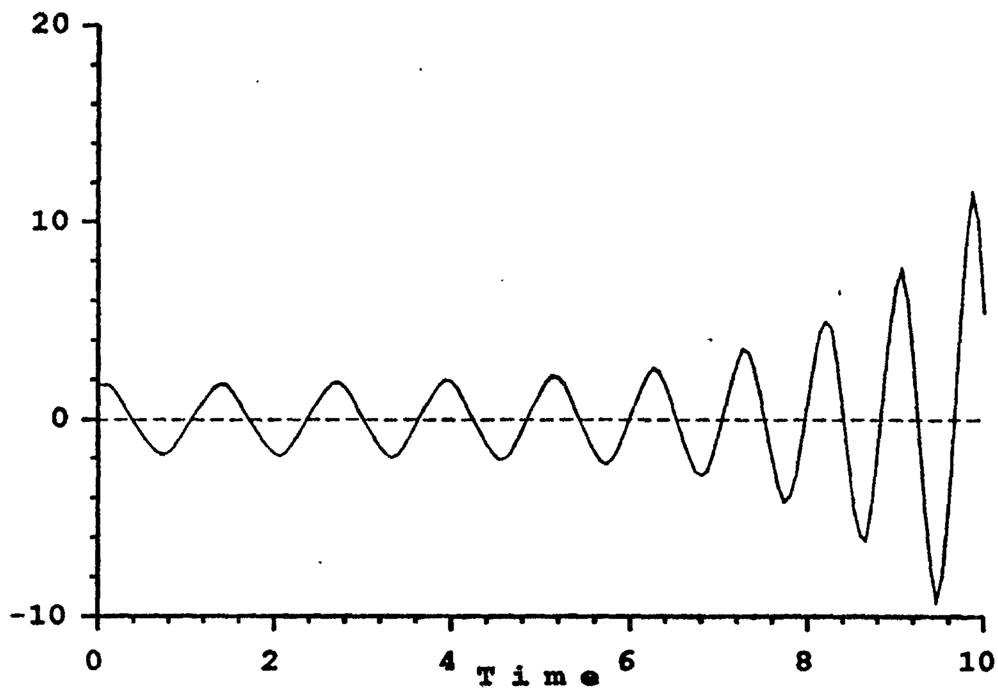

Figure 5: Example 2: Constrained; D=0.4 and d=1.75

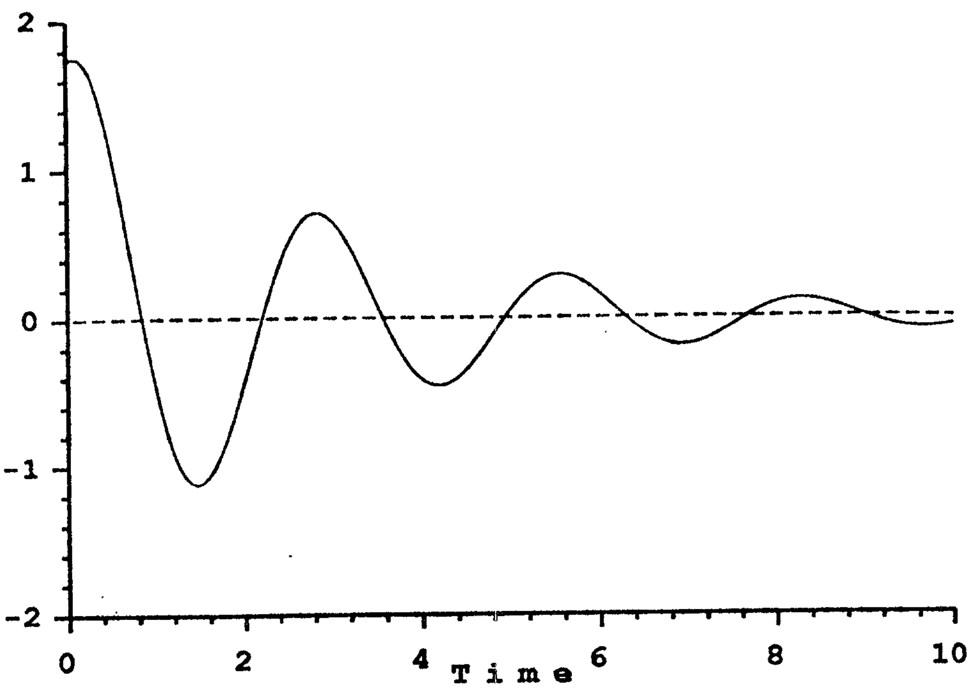

Figure 6: Example 2: Unconstrained; D=0.4 and d=1.75

2×2 part of the Heavy Oil Fractionator of the SSCP. This system has as outputs 1 and 2, the Top End Point and the Side End Point correspondingly. The inputs are the Top Draw and the Side Draw. The transfer function of this subsystem is

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{(4.05+2.11\epsilon_1)e^{-27s}}{50s+1} & \frac{(1.77+0.39\epsilon_2)e^{-28s}}{60s+1} \\ \frac{(5.39+3.29\epsilon_1)e^{-18s}}{50s+1} & \frac{(5.72+0.57\epsilon_2)e^{-14s}}{60s+1} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{50}$$

where ϵ_1, ϵ_2 represent the model uncertainty and they can take any value between -1 and $+1$, 0 corresponding to the nominal model. A sampling time of $T = 6$ *min* is selected which results in lower and upper constraints of -0.3 and 0.3 for the changes in the inputs from one sampling point to the next. Lower and upper constraints of -0.5 and 0.5 exist for all the inputs and outputs.

Our goal is to see how the stability conditions can be used to analyze simulation results. In the objective function of (2) we select $P = 6$, $M = 2$, $B = D = 0$. The minimization is carried out subject to the above described hard constraints. The Constraint Window for the outputs includes future points 5-6 for the Top End Point and 3-4 for the Side End Point. Beginning the windows at earlier times may result in infeasibilities because of the longer time delays. It should be noted that this selection of parameters is meant as a simple one rather than an "optimal" one.

The simulation for no model-plant mismatch is shown in Fig. 7, where a disturbance in the form of simultaneous step changes of 0.5 in the Upper and the Intermediate Reflux Duties is used. The same disturbance is used in all simulations in this section. Use of the disturbance transfer function models yields the following output disturbance vector:

$$d(s) = \begin{bmatrix} \frac{1.20e^{-27s}}{45s+1} & \frac{1.44e^{-27s}}{40s+1} \\ \frac{1.52e^{-15s}}{25s+1} & \frac{1.83e^{-15s}}{20s+1} \end{bmatrix} \begin{bmatrix} 0.5/s \\ 0.5/s \end{bmatrix} \tag{51}$$

Note that the plot of ρ is the value of the necessary condition for the particular J_i occuring at the sample points during the simulation. When a model-plant mismatch is present, as in

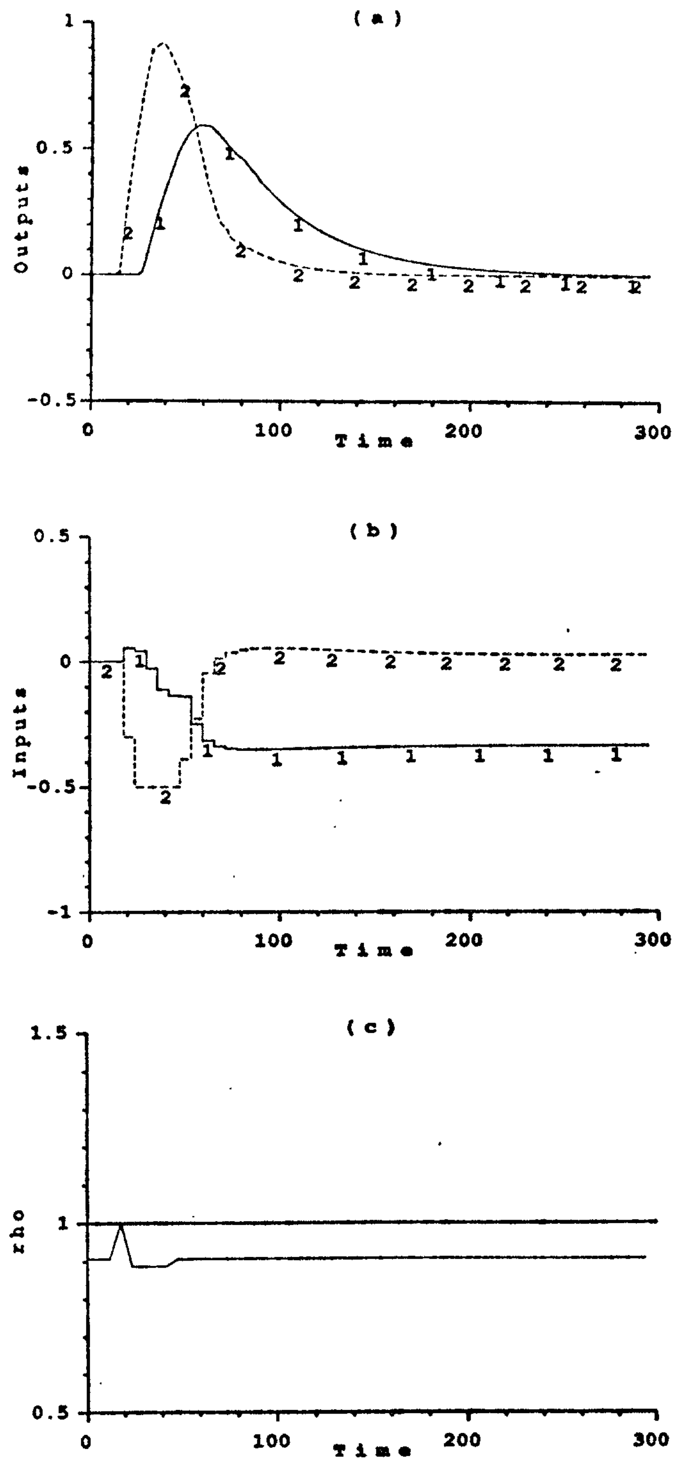

Figure 7: Example 3: Nominal; lower input constraint at -0.5. (a) Outputs; (b) Inputs; (c) $\rho(F')$

the following simulations, it is computed for the coefficients of the actual plant used in the simulation.

Next, a mismatch between the model and the plant is assumed, corresponding to $\epsilon_1 = -1$ and $\epsilon_2 = 1$. The simulation is shown in Fig. 8. By looking just at the outputs and inputs there is no indication of a potential problem. However by looking at the plot of ρ we see that the necessary condition is close to being violated during part of the simulation. It is simple to check that this part of the simulation corresponds to the case where at the optimum of the on-line optimization no constraint is active. The problem is not significant in this simulation because eventually, the lower constraint for the Top Draw becomes active at the optimum and we move to a well-behaved region. Let us now repeat the simulation of Fig. 8 but with the lower constraints for the inputs at -1 rather than -0.5. The simulation is shown in Fig. 9 and this time the system sufers from persistent oscillations because the constraint does not become active early on. Figure 10 repeats the simulation of Fig. 9 but with a larger mismatch. We are using $\epsilon_1 = -1.2$ and $\epsilon_2 = 1.2$. This time we are in the instability region as the plots show. The question of interest at this point is how to use the plot of ρ in Fig. 10 to make a parameter change so that the system is stabilized. From the previous simulations it is clear that one way would be to simply increase the value of the lower input constraint, i.e., use this constraint as a tuning parameter. What is important to note however is the following:

Tuning Observation. *The values of the hard constraints do not appear in the expressions of the $C_{J_i}s$; hence they can influence stability only by keeping a destabilizing J_i from occuring. They cannot change a C_{J_i} into a stabilizing controller; this can be accomplished only by the parameters of the objective function.*

Hence it seems that is safer to actually try to find values for the parameters of the objective function that make C_{J_1} stabilizing (where J_1 is defined to correspond to the case of no active constraints at the optimum), without changing the values of the hard constraints. But this is a problem that can be addressed through Robust Linear Control Theory. Use of the Structured Singular Value shows that a $B = 0.2$ stabilizes the system. The simulation is given in Fig. 11. Note that if the problematic C_{J_i} corresponded to some active constraints, the situation would still be treated through the same tools.

9 Conclusions

This paper has provided a theoretical framework for the study of the properties of control algorithms that are based on the on-line minimization of some objective function, subject to certain hard constraints. The selected framework seems to be very promising since it allowed the derivation of necessary and/or sufficient conditions for nominal and robust stability of the overall nonlinear system. These conditions can be formulated in a way that allows the treatment of the kinds of model-plant mismatch that robust linear control theory can handle.

Simple examples were used to demonstrate in a clear way that one cannot afford to neglect the nonlinear phenomena caused by the hard constraints to which the on-line optimization is subject. These examples also indicate that inclusion of hard constraints on the plant outputs in the specifications can cause serious problems and that their effect should be carefully studied before they are used. The stability conditions of this paper can be used in

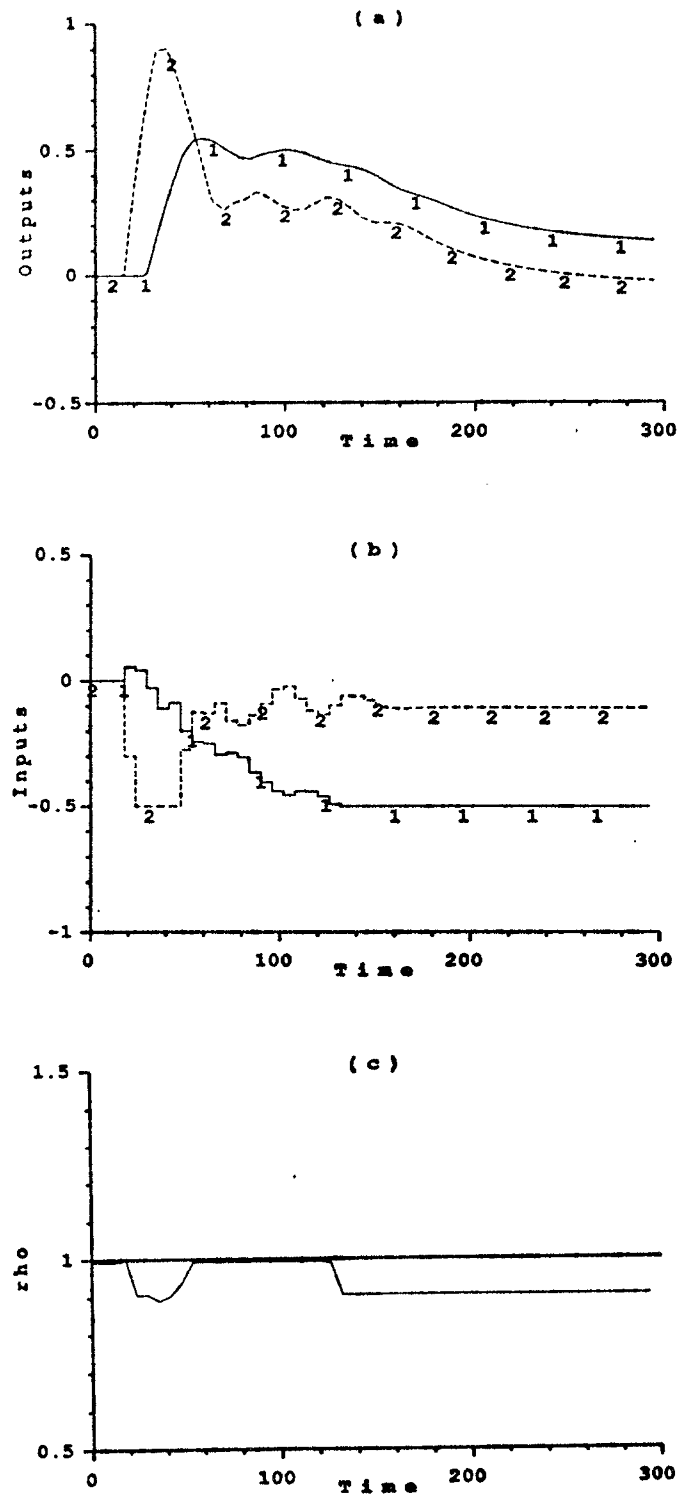

Figure 8: Example 3: $\epsilon_1 = -1$, $\epsilon_2 = 1$; lower input constraint at -0.5. (a) Outputs; (b) Inputs; (c) $\rho(F')$

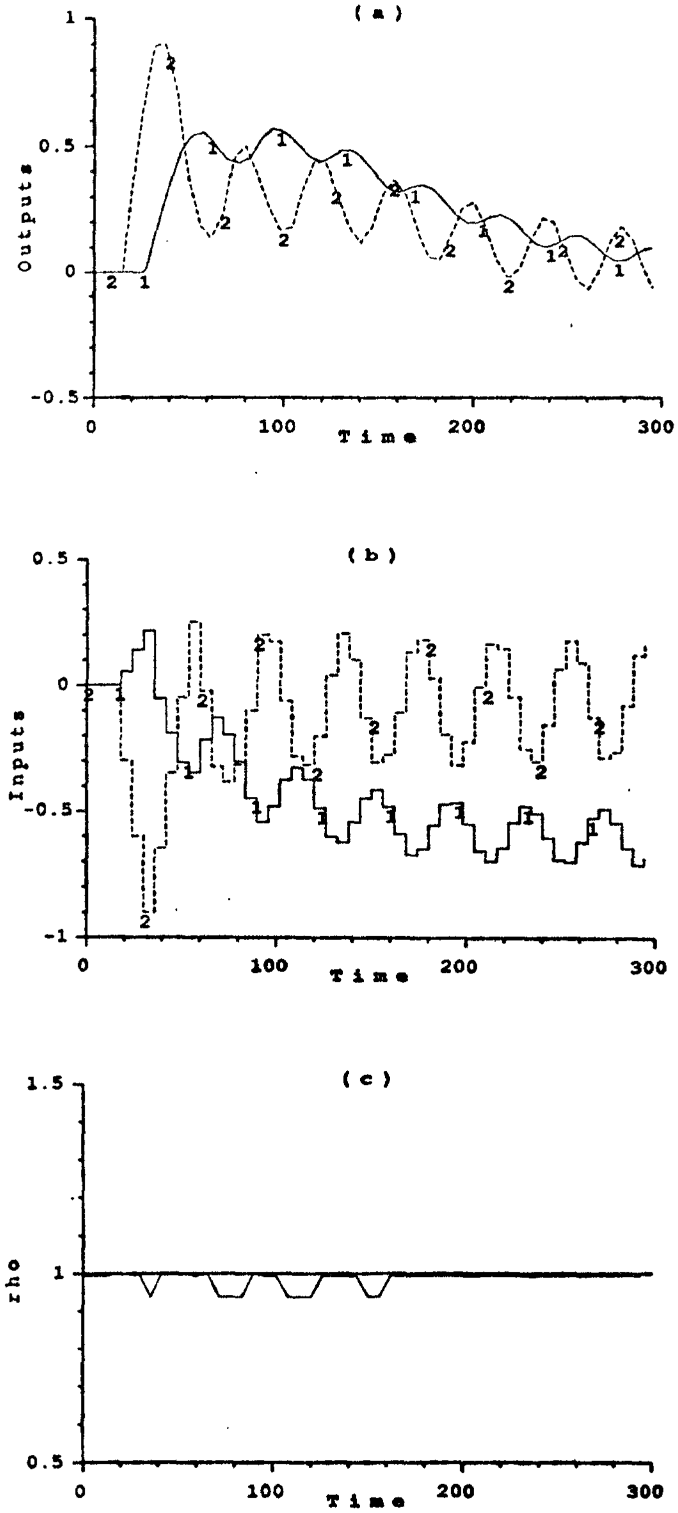

Figure 9: Example 3: $\epsilon_1 = -1$, $\epsilon_2 = 1$; lower input constraint at -1. (a) Outputs; (b) Inputs; (c) $\rho(F')$

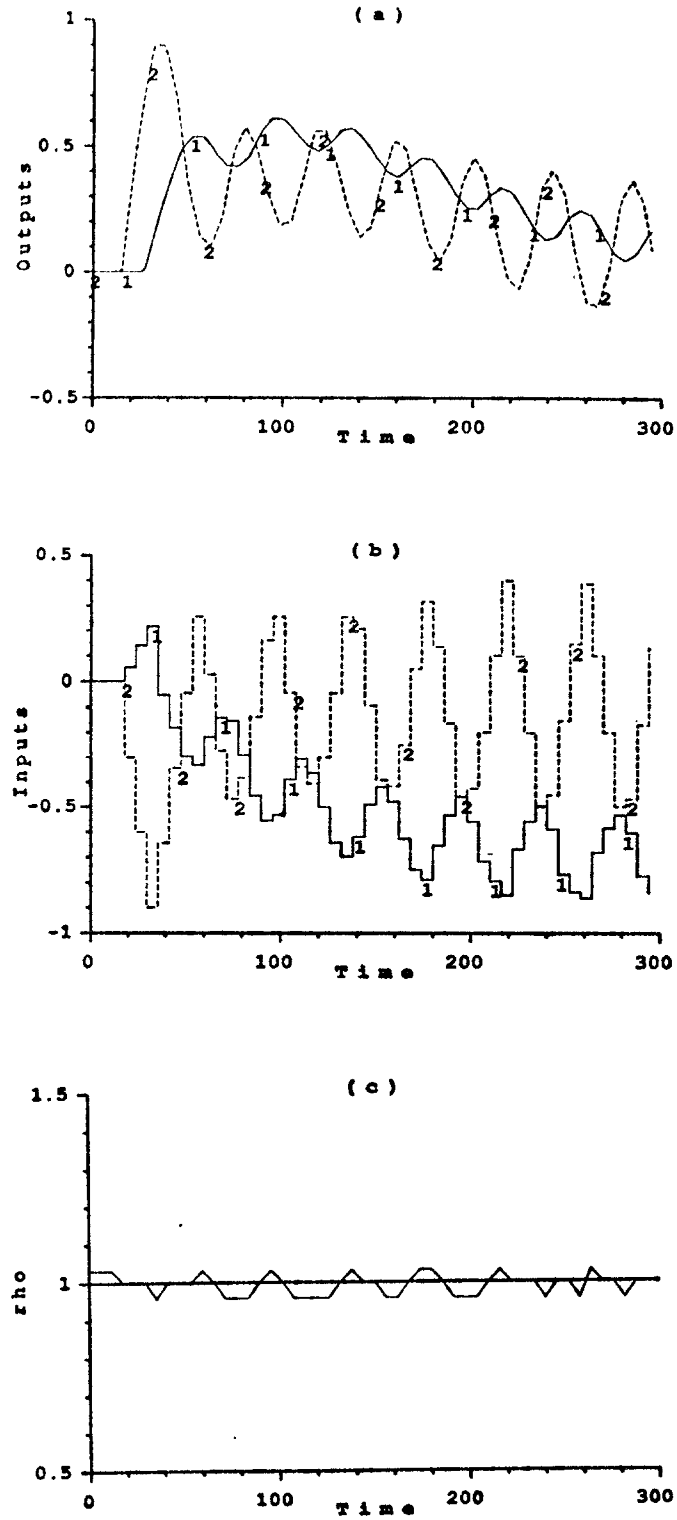

Figure 10: Example 3: $\epsilon_1 = -1.2$, $\epsilon_2 = 1.2$; lower input constraint at -1. (a) Outputs; (b) Inputs; (c) $\rho(F')$

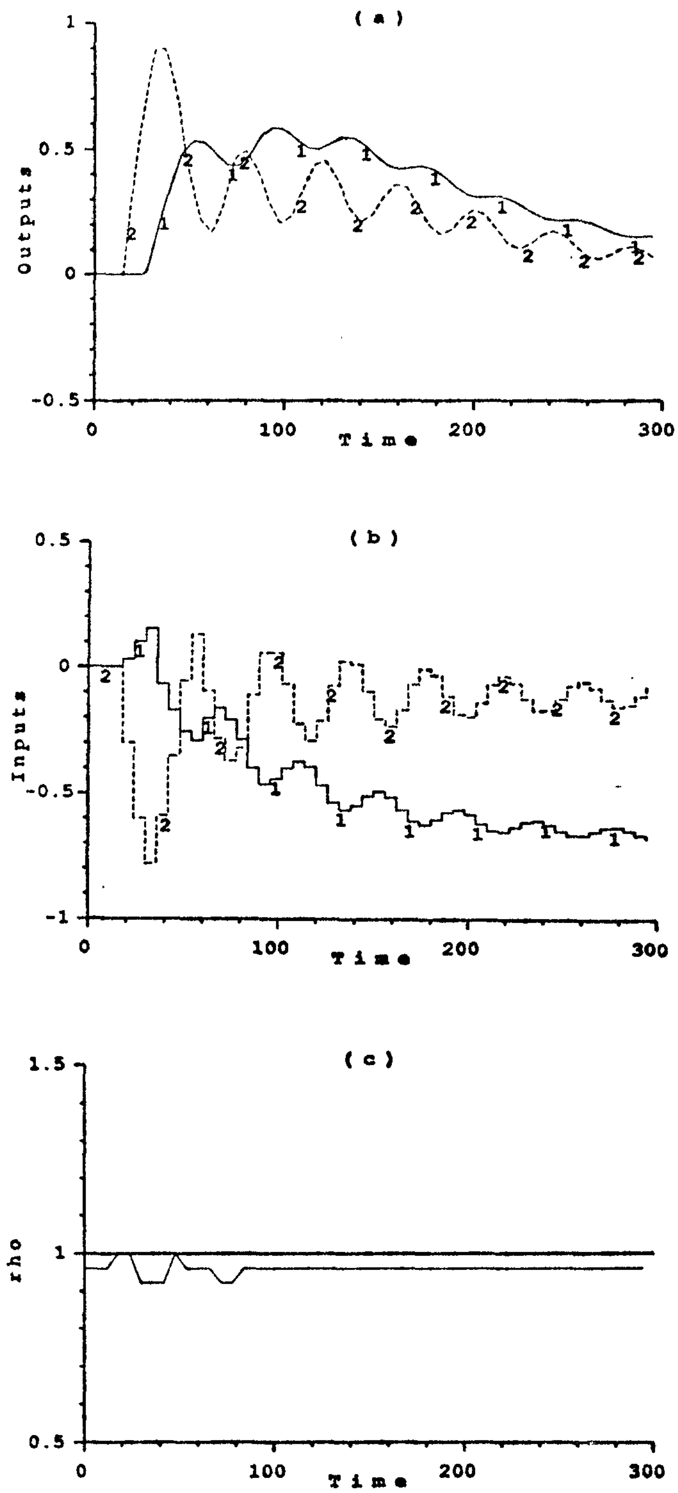

Figure 11: Example 3: $\epsilon_1 = -1.2$, $\epsilon_2 = 1.2$; lower input constraint at -1; $B = 0.2I$. (a) Outputs; (b) Inputs; (c) $\rho(F')$

this study.

The practical usefulness of the theory that was developed in this paper was demonstrated by the application of the stability conditions on the realistic model of the Shell Heavy Oil Fractionator. The conditions were able to capture characteristics not clearly visible in the plant outputs and inputs. In order to increase the applicability of the method on complex industrial systems, further work is needed on developing a computationally feasible search procedure for finding and testing all the J_is that can occur during the operation of the control system.

Acknowledgements

The control software package CONSYD, developed at Caltech (Dr. M. Morari's group) and the University of Wisconsin (Dr. W. H. Ray's group), was used in the simulations.

References

[] Y. Arkun, G. Charos and D. E. Reeves, "Model Predictive Control", *Chem. Eng. Education,* **22**, pp. 178-183, 1988.

[] P. J. Campo and M. Morari, "Robust Model Predictive Control", *Proc. Amer. Control Conf.,* pp. 1021-1026, Minneapolis MN, 1987.

[] J. C. Doyle, "Analysis of Feedback Systems with Structured Uncertainty", *I.E.E. Proc.,* Part D, **129**, pp. 242-250, 1982.

[] J. W. Eaton, J. B. Rawlings and T. F. Edgar, "Model-Predictive Control and Sensitivity Analysis for Constrained Nonlinear Processes", *Proc. IFAC Workshop on Model Based Process Control,* Pergammon Press, Oxford, 1989.

[] C. G. Economou, *An Operator Theory Approach to Nonlinear Controller Design,* Ph.D. Thesis, California Insitute of Technology, 1985.

[] R. Fletcher, *Practical Methods of Optimization; vol. 2: Constrained Optimization,* John Wiley and Sons, 1981.

[] C. E. Garcia, "Quadratic Dynamic Matrix Control of Nonlinear Processes: An Application to a Batch Reaction Process", *AIChE Ann. Mtg.,* San Francisco CA, 1984.

[] C. E. Garcia and M. Morari, "Internal Model Control. 1. A Unifying Review and Some New Results", *Ind. Eng. Chem. Proc. Des. Dev.,* **21**, pp. 308-323, 1982.

[] C. E. Garcia and M. Morari, "Internal Model Control. 3. Multivariable Control Law Computation and Tuning Guidelines", *Ind. Eng. Chem. Proc. Des. Dev.,* **24**, pp. 484-494, 1985.

[] C. E. Garcia and A. M. Morshedi, "Quadratic Programming Solution of Dynamic Matrix Control (QDMC)", *Chem. Eng. Commun.,* **46**, pp. 73-87, 1986.

[] C. E. Garcia and D. M. Prett, "Advances in Industrial Model Predictive Control", *Chem. Proc. Control Conf. III,* Asilomar CA, 1986.

[] M. Morari, C. E. Garcia and D. M. Prett,"Model Predictive Control: Theory and Practice", *Proc. IFAC Workshop on Model Based Process Control,* Pergammon Press, Oxford, 1989.

[] M. Morari and E. Zafiriou, *Robust Process Control*, Prentice–Hall, Englewood Cliffs, NJ, 1989.

[] D. M. Prett and C. E. Garcia, *Fundamental Process Control*, Butterworth Publishers, Stoneham, MA, 1988.

[] N. L. Ricker, T. Subrahmanian and T. Sim, "Case Studies of Model-Predictive Control in Pulp and Paper Production", *Proc. IFAC Workshop on Model Based Process Control,* Pergammon Press, Oxford, 1989.

[] G. C. Stewart, *Introduction to Matrix Computations*, Academic Press, New York, 1973.