Reference System Model Predictive Control. 1. Continuous Time Formulation and Case Studies on Performance

Lokesh Kalra,† Christos Georgakis,*,‡ and Luis C. Oliveira-Lopes§

Chemical Process Modeling and Control Research Center and Department of Chemical Engineering, Lehigh University, Bethlehem, Pennsylvania 18015

Reference system model predictive control (RS-MPC) aims to fill a void in predictive control technology that is caused mainly by the lack of intuitiveness of tuning of currently popular MPC schemes. It addresses the two main drawbacks of weight-based approaches: (1) the absence of clear easy-to-follow guidelines for the selection of such weights to obtain the desired closed-loop response; (2) the lack of a uniform closed-loop response over a wide operating range without the need to retune the controller. Because the determination of tuning parameter values is based on a specification of the desired closed-loop response, RS-MPC offers significant advantages, particularly in the context of nonlinear systems. Presented herein is the development of the RS-MPC algorithm as well as some representative example cases demonstrating its tuning advantages. It is shown herein that RS-MPC consistently delivers the asked performance without needing retuning at different operating conditions. This is in sharp constrast to weight-based techniques which we show often fail to provide a uniform closed-loop response as the underlying open-loop dynamics change due to the nonlinear character of a process.

1. Introduction

Reference system synthesis (RSS)[1] and related techniques (generic model control or GMC[2]) have been around for some time and have been used with considerable success.[3−5] Some of the features that make them attractive are their simple philosophy, their ability to handle nonlinear processes, and their tuning transparency. With a straightforward requirement that the process follow a reference system, a large variety of processes can be handled. However, their inability to handle nonminimum-phase (NMP) systems and process constraints severely limits their applicability to the increasingly tightly designed and integrated plants of today. The fact that it addresses these issues in a relatively straightforward manner is a major reason for the increasing use of model predictive control (MPC) schemes.[6,7] However, the tuning of the available conventional MPC schemes is nontrivial even for simple multivariable linear systems.[8] Shridhar and Cooper in their papers[9,10] on the tuning of unconstrained MPC review a comprehensive list of more than 15 papers published mostly in the 1990s on this active topic of research. For nonlinear systems, the problem of tuning MPC becomes more difficult, with retuning required when the dynamics of the plant are different from those used to tune the controller. This could occur as the process moves around in its operating range in response to disturbances or to bring about grade changes. The lack of simple tuning guidelines and the high maintenance that might be required could be a major argument against the use of conventional MPC techniques for highly nonlinear processes.

Reference system MPC (RS-MPC) is a methodology that was developed to address the special needs of NMP plants and processes with constraints. It is a framework that is based on predictive control concepts but one that has none of the tuning drawbacks of conventional MPC. The purpose of the present paper is to introduce RS-MPC and to demonstrate some of its features and tuning advantages with the help of simulation examples. To begin with, in section 2, the RSS philosophy to controller design is reviewed. The problems posed by NMP plants and a new RSS structure that allows for the handling of these plants is discussed in section 3. The RS-MPC algorithm is then described in section 4 along with guidelines for tuning parameter selection and some implementation details. Section 5 contains simulation examples for a single-input single-output (SISO) system to illustrate some features of RS-MPC. Some of the most significant advantages of RS-MPC are in the context of multivariable and nonlinear systems. This is pointed out in section 6 with the help of simulation examples from the Amoco model IV FCCU. Sections 7 and 8 contain some closing remarks.

2. Reference System Control (RSC)

In traditional controller design, a process model is used to design a controller, the controller is applied to the system, and then the closed-loop characteristics are examined to determine the controller tuning parameters. The philosophy of RSC reverses this order; the desired closed-loop characteristics are first specified through the RS, and the controller law is then obtained. The idea behind RSC[1] then is to design a controller that forces the closed-loop process to follow a prescribed trajectory, the RS. In this sense, the RS methodology is related to Richalet's et al.'s seminal 1978 contribution[11]

* To whom all correspondence should be addressed.

† Present address: Air Products & Chemicals Inc., Allentown, PA 18105.

‡ Present address: Department of Chemistry, Chemical Engineering, and Material Science, 728 Rogers Hall, Polytechnic University, Six Metrotech Center, Brooklyn, NY 11201. Phone: (718) 260-3579. Fax: (718) 260-3125. E-mail: C.Georgakis@Poly.edu. URL: http://chem.poly.edu.

§ Present address: School of Chemical Engineering, Federal University of Uberlandia, Avenida Joao Naves de Avila, 2160 38400-902 Uberlandia/MG, Brazil.

10.1021/ie000891j CCC: $22.00
Published on Web 06/04/2002

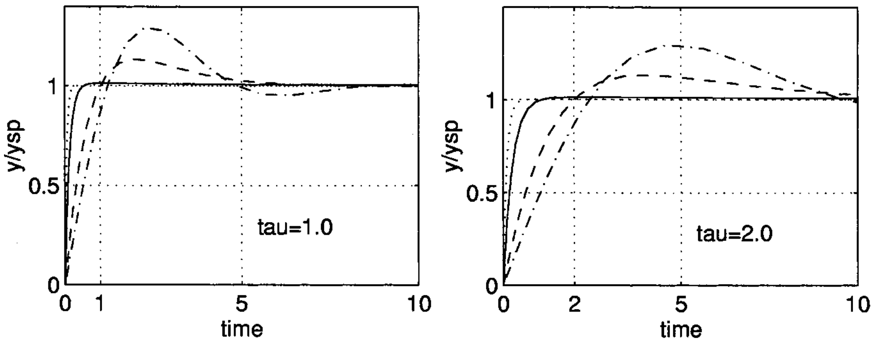

Figure 1. Shape of the closed-loop response with RS tuning for ξ values of 0.5 (dash-dotted), 1.0 (dashed), 4.0 (solid), and 10.0 (dotted) with $\tau = 1.0$ (left) and 2.0 (right).

which introduced the concept of a reference trajectory and led to the IDCOM technology. However, as will be seen in the next few paragraphs, in section 3.1, and in the appendix, RSC is a significant extension because of the tuning improvements resulting from the way the reference trajectory is applied to the process.

The RSC methodology consists of three main steps. First, the process to be controlled is described by a general state-space model of the following form:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \sum_{i=1}^{m} \mathbf{g}_i(\mathbf{x})\, u_i, \quad \mathbf{y} = \mathbf{h}(\mathbf{x}) \tag{1}$$

where the inputs $\mathbf{u}$, and the outputs $\mathbf{y}$, belong to $\mathbf{R}^m$ and the states $\mathbf{x} = [x_1, x_2, ..., x_n]^{\mathrm{T}}$ belong to $\mathbf{R}^n$. $\mathbf{f}(\mathbf{x})$ and $\mathbf{g}_i(\mathbf{x})$, $i = 1, ..., m$, are n-dimensional and $\mathbf{h}(\mathbf{x})$ is an m-dimensional field of $\mathbf{x}$. The reference system methodology assumes a system that is affine in the inputs. However, this assumption is not overly restrictive because Henson and Seborg (1990)[12] have introduced a way of reducing general nonlinear systems to a form that is affine in the input.

Next, the desired closed-loop response of the system is specified through a reference system represented as $\mathrm{d}\mathbf{y}_{\mathrm{r}}/\mathrm{d}t = \mathbf{f}_{\mathrm{r}}(\mathbf{y}^{\mathrm{set}},\mathbf{y}_{\mathrm{r}},\mathbf{u},\mathbf{d})$. Then, in the third step, the desired control law is obtained by minimizing the difference between the system equation and the reference system equation, min $\{\mathrm{d}\mathbf{y}/\mathrm{d}t - \mathrm{d}\mathbf{y}_{\mathrm{r}}/\mathrm{d}t\}$. In general cases, numerical methods must be used to solve for the inputs. However, in the case of square multiple-input multiple-output (MIMO) systems with relative order 1, the equations $\dot{\mathbf{y}} - \dot{\mathbf{y}}_{\mathrm{r}} = \mathbf{0}$ can be solved to obtain an expression for the inputs.

All of the tuning here is through the choice of the reference system and its parameters. One form of the reference system that has been used with success in the past[1,3] is the PI reference system written as $\dot{\mathbf{y}}_{\mathrm{r}} = \mathbf{K}_1(\mathbf{y}^{\mathrm{sp}} - \mathbf{y}_{\mathrm{r}}) + \mathbf{K}_2\int_0^t(\mathbf{y}^{\mathrm{sp}} - \mathbf{y}_{\mathrm{r}})\,\mathrm{d}t$. A major advantage of this reference system is a particularly simple form of the closed-loop transfer function that results if the $\mathbf{K}$ matrices are chosen to be diagonal, with their elements being functions of parameters ξ and τ such as $k_{1ii} = 2\xi_i/\tau_i$, $i = 1, 2, ..., m$, and $k_{2ii} = 1/\tau_i^2$, $i = 1, 2, ..., m$, where i refers to the ith output. In the event of no plant–model mismatch and disturbances, the relationship between the outputs and the setpoints can be written for the ith output as

$$\frac{y_i(s)}{y_i^{\mathrm{sp}}(s)} = \frac{2\tau_i\xi_i s + 1}{\tau_i^2 s^2 + 2\tau_i\xi_i s + 1} \tag{2}$$

The parameter ξ_i here is similar to the damping coefficient for linear second-order systems, and τ_i is akin to the time constant. These two parameters can be used to shape the response of the closed-loop system at the control system design stage, giving the designer an early feel for the desired closed-loop performance.

Figure 1 shows the effect that the ξ and τ tuning parameters have on the closed-loop response. It is basically the response of the output y_i to a step change in its setpoint y_i^{sp}, described by eq 2. The plot on the left shows, for four values, how ξ affects the performance for a fixed τ value of 1.0. It can be seen that $\xi = 0.5$ yields a response that is the slowest and also has the largest overshoot. The plot on the right-hand side in Figure 1 shows the system's responses with the same four ξ values and a larger τ value of 2.0. It can be seen that the responses are self-similar, and each one would be identical with the corresponding one on the left if the time axis was made dimensionless with the time constant, τ, of the reference system. These observations form the basis of how the two parameters can be used in setting the desired closed-loop response. A value of ξ is picked based on a preference for the shape of the response of the output to a setpoint change. Once this is done, a τ value can be set based on the desired speed of the closed-loop response. If a fast response is wanted, a small τ value is used, while larger values of τ give slower responses. Thus, with the help of these two reference system parameters, considerable flexibility is possible under RSC.

A more quantitative feel for the closed-loop speed of response is also possible. An expression for the rise time as a function of τ and ξ can easily be derived with the help of the closed-loop transfer function equation (2). This expression was found to be (eq 3 is good for $\xi \neq 1$; for $\xi = 1$, it can easily be verified that $t_{\mathrm{r}} = \tau$)

$$t_{\mathrm{r}} = \tau\frac{\ln(\xi_i + \sqrt{\xi_i^2 - 1})}{\sqrt{\xi_i^2 - 1}} \tag{3}$$

This equation can be used as an indicator of the expected speed of the response of the system under RSC, although in many cases plant–model mismatch will cause the closed-loop response to deviate from that predicted by the reference system. The method thus allows the control system designer to quickly arrive at values of the τ and ξ parameters that would give the desired speed and shape of response.

3. Handling NMP Plants

NMP plants are known to pose problems and to limit the achievable closed-loop performance. The extra phase lag in systems that contain time delays and/or unstable *zero dynamics* is known to contribute to instability[13] by making the system closer to being unstable at lower frequencies. Thus, controllers cannot be tuned as aggressively as would be possible for minimum-phase (MP) plants. In model-based strategies, the controller transfer function includes the inverse of the process model.[14] The presence of time delays or RHP zeros makes the controller either unstable or noncausal.[15,16] Hence, special attention needs to be paid to control the design for such processes.

In this section a form of RSC is proposed that handles NMP plants. This development is crucial because the RS-MPC algorithm that handles NMP plants is based on it. First, in section 3.1, the structure of RSC is exposed to (i) show the source of the methodology's tuning advantages and (ii) show why the current formulation is unable to handle NMP plants. The proposed approach for implementing RSC on NMP plants is discussed in section 3.2. It will be seen there that the approach is based on the factorization of the plant model into a MP and an AP factor. Section 3.3 contains a note on the factorization and how it can be carried out for systems of interest. The tuning of RSC for NMP plants is somewhat more involved from that presented in section 2, and it is discussed in section 3.4.

3.1. Structure of RSC. Insights into RSC tuning and the source of the diagonal closed-loop transfer matrix are presented in detail in work by Kalra.[17] The analysis is based on state feedback for input−output decoupling and linearization. Here only the major result is presented to illustrate the structure of RSC as applied to a linear system. Consider the square multivariable system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \quad \mathbf{y} = \mathbf{C}\mathbf{x} \tag{4}$$

where $\mathbf{x}$, $\mathbf{u}$, and $\mathbf{y}$ are vectors as in eq 1 and $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are matrices in $\mathbf{R}^{n\times n}$, $\mathbf{R}^{n\times m}$, and $\mathbf{R}^{m\times n}$, respectively. A linear state feedback of the form $\mathbf{u} = -\mathbf{K}\mathbf{x} + \mathbf{H}\mathbf{v}$ can be applied with $\mathbf{K}$ in $\mathbf{R}^{m\times n}$, $\mathbf{H}$ in $\mathbf{R}^{m\times m}$, and $\mathbf{v}$ a vector of m reference inputs. For a system of relative order (the relative order for the ith output y_i is defined as the smallest integer r_i such that $\mathbf{c}_i\mathbf{A}^{r_i-1}\mathbf{B} \neq 0$; the more general relative order case is presented in ref 8) unity in all of the outputs, Chen[18] shows that if we set $\mathbf{K} = [\mathbf{CB}]^{-1}\mathbf{CA}$ and $\mathbf{H} = [\mathbf{CB}]^{-1}$, each output depends only on the corresponding reference input via the relationship $\mathrm{d}y_i/\mathrm{d}t = v_i$. In RSC this result is extended further with the imposition of output feedback in an outer loop by defining v_i to be $(2\xi_i/\tau_i)(y_i^{\mathrm{sp}} - y_i) + (1/\tau_i^2)\int_0^t(y_i^{\mathrm{sp}} - y_i)\,\mathrm{d}t$ for each i. Then, in the Laplace domain, the closed-loop relationship between any output y_i and the corresponding setpoint y_i^{sp} is exactly the transfer function in eq 2. In other words, a particular choice of state + output feedback has decoupled the closed-loop relationship and imposed a desired dynamics between the outputs and the corresponding setpoints.

For the case of nonlinear systems, a state feedback can be found that decouples *and* linearizes the closed-loop relationships. In the thesis[17] we build on the results of Kravaris and co-workers in the context of global linearizing control[19−21] and demonstrate that the closed-loop transfer function matrix is once again diagonal with each element represented by eq 2. This decoupling, linearizing state + output feedback is what gives RSC its tuning advantages. Initial tuning is easier because of the freedom to specify without trial and error a desired response through the reference system. The decoupling makes any necessary fine tuning easier, and for significantly nonlinear systems, the linearization leads to a uniform closed-loop response over a wide operating range.

Let us now look deeper into the RSC law as proposed above to expose a limitation. For the system in eq 4 and with the choice of $\mathbf{K}$ and $\mathbf{H}$ as above, the control law is $\mathbf{u} = [\mathbf{CB}]^{-1}[-\mathbf{CAx} + \mathbf{v}]$. In the Laplace domain now, we can use in the above equation the simplest form of the state estimator, an open-loop observer given by $\mathbf{x}(s) = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{Bu}(s)$. With the help of algebra, it can then be shown that the control law reduces to $\mathbf{u}(s) = (1/s)[\mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}]^{-1}\mathbf{v}(s)$. It is easy to see that the controller transfer function incorporates the inverse of the process model $\mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}$. This perspective of RSC is useful because it provides a way of studying the robustness properties of RS-MPC.[17] A more relevant point that can be drawn from the controller transfer function relates to systems with unstable zeros. When the full process model is used in RSC, the inverse will have a pole in the RHP. This will make the controller unstable, clearly indicating the problems that NMP plants pose for RSC. In the general nonlinear case, if the open-loop system were to have unstable (RHP) zero dynamics or time delays, the resulting state feedback based RSC would be unstable and/or noncausal. For this reason, the original RSC formulation cannot handle NMP plants. This limitation is addressed in the next few sections.

3.2. RSC for NMP Plants. The frequency with which NMP processes are encountered[22] makes their handling quite important. The simplest approach around some of the problems caused by RHP zeros and time delays is to not invert the process model exactly. In a predictive implementation of model-based control such as MPC, an inversion of the process model is done implicitly and in an approximate manner. This is an outcome of the minimization of the objective function to determine the control moves to be implemented on the plant. As a result of the approximate inversion performed, the resulting controller will be stable for many cases. Rawlings and Muske[23] introduced the concept of terminal constraints to force stability. In general though, the selection of the tuning parameter values must be done carefully in order to avoid stability problems (exact model inverses could result if a control horizon equal to the prediction horizon and a zero input weighting matrix are used[24]). This, when combined with the fact that the determination of tuning parameters of MPC is not intuitive to begin with, results in a very difficult tuning problem.

A better approach to handling NMP plants is one where only the part of the model that can be inverted without causing difficulties is used in controller synthesis. In this manner, the NMP behavior is explicitly accounted for rather than implicitly via an optimization as was done above. Such a method has been used in the IMC framework[15] as well in the context of generalized Smith predictors.[25] The basic idea is to use a model that has been factorized into two parts in parallel with the plant. The first part is the so-called *outer factor* and is the MP or invertible part of the model. The second

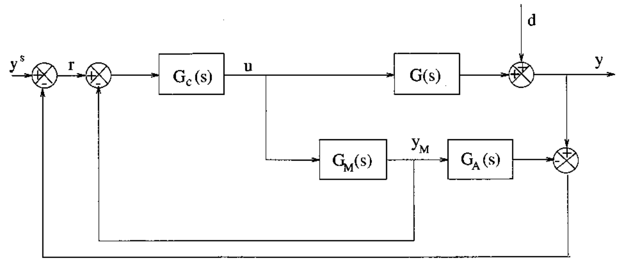

Figure 2. Block diagram for RSS for NMP processes.

part is the *inner factor*, which in the SISO case reduces to an all-pass (AP) system. Thus, for a linear system $\mathbf{G}(s) = \mathbf{G}_A(s)\ \mathbf{G}_M(s)$, with $\mathbf{G}_A(s)$ representing the inner factor and $\mathbf{G}_M(s)$ the outer MP factor. Then, instead of the reference system being specified on the plant output, it is specified at the output of an inner loop that contains the outer MP factor of the model, as shown in Figure 2 for the case of a linear system. Because RSC is implemented on this inner loop, which only contains the invertible part of the model, the resulting closed-loop controller $\mathbf{G}_c(s)$ will be causal and stable. The outer loop provides feedback action so that plant−model mismatch and/or unmeasured disturbances can be handled.

In the context of RSC for linear systems, the strategy for taking the NMP characteristics of a plant into account can be further described with the help of the *proper* state-space system:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \quad \mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} \tag{5}$$

This system can be written in terms of two factors (see ref 17 for details). The outer factor is *strictly proper* and MP, with the state equation the same as above. By a proper choice of the output map $\mathbf{C}_M$ of the MP factor, the output $\mathbf{y}_M$ can be made to be such that it has the same steady-state gain as $\mathbf{y}$ but does not exhibit any NMP dynamics. Hence, this factor is suitable for use in RSC, as is illustrated in Figure 2.

The second factor, the inner factor, contains all of the noninvertible dynamics of the process. With the subscript A denoting that the system is an AP, it can be written as

$$\dot{\mathbf{x}}_A = \mathbf{A}_A\mathbf{x}_A + \mathbf{B}_A\mathbf{y}_M, \quad \mathbf{y} = \mathbf{C}_A\mathbf{x}_A + \mathbf{D}_A\mathbf{y}_M \tag{6}$$

Thus, its input is the output of the MP factor and its output the original model's output $\mathbf{y}$, a result that has already been indicated through the equation for $\mathbf{G}(s)$ above. This output can be compared with the incoming measurement from the plant, as is shown in Figure 2, to determine any inaccuracies in the model.

For a restrictive class of nonlinear systems, there are some general results available on factorization. The interested reader is referred to papers by Ball and Helton,[26] van der Schaft and Ball,[27] Kravaris and Daoutidis,[28] Wright and Kravaris,[16] and Doyle et al.[29] In all of the methods mentioned in the literature, the MP factor is different from the full model only in its output map, leaving the state equations intact. For these cases, exactly as is shown in Figure 2, the reference system requirement can be specified on this MP factor with the inner or AP factor left open loop for RSC. For other more general nonlinear problems, locally linear models can be factorized for use online and the vast theory on linear factorization can be utilized. In this manner then, RSC for NMP nonlinear processes is possible.

3.3. Factorization. For linear systems, the separation into an invertible part and an inner or AP part is straightforward.[17] The method that was available in the literature[15] was used for the most part. However, a difficulty is encountered for strictly proper systems. An approach is proposed by Kalra (1997) for producing low-order inner and outer factors that do not contain any artificial poles, and zeros are only summarized here. It involves introducing a $\mathbf{D}$ matrix of the form $\epsilon\mathbf{I}_m$, with ϵ a small number, before applying the inner−outer algorithm. Postprocessing is required on the resulting inner factor to cancel out common poles and transmission zeros associated with the $\mathbf{D}$ matrix. This is done via (i) a minimal realization of the inner factor to cancel out common poles and transmission zeros and (ii) the elimination of states that correspond to large modes (poles) through the use of a Jordan or modal form[30] of the inner factor. An example of this approach applied to a second-order SISO system can be found in work by Kalra.[17]

3.4. Tuning of RSC for NMP Plants. Assuming a perfect model and in the absence of disturbances, there is no feedback from the outer loop in Figure 2. Then, the closed-loop transfer function between the output and the setpoint is simply that between the output of the MP factor and the setpoint, modified by that of the inner factor, or $\mathbf{G}^{CL}(s) = \mathbf{G}_A(s)\ \mathbf{G}_M^{CL}(s)$. If a PI reference system is used, for SISO cases, the overall closed-loop transfer function $\mathbf{G}^{CL}(s)$ can be written as

$$\mathbf{G}^{CL}(s) = \mathbf{G}_A(s)\,\frac{2\xi\tau s + 1}{\tau^2 s^2 + 2\xi\tau s + 1} \tag{7}$$

For this NMP case tuning is still straightforward because of the known effect of ξ and τ but would require simple simulations. Trial values of ξ and τ are chosen, as for MP systems, based on the desired shape of the response and a feel for the achievable closed-loop speed. Then, via eq 7 the actual response that will be obtained is determined using (for example) MATLAB. The presence of the RHP zero will skew the response from one that would have been obtained for a MP plant. Using this response from eq 7 as the starting point, a new set of ξ and τ values can be chosen if the initial set was not found to be satisfactory. Thus, the presence of NMP characteristics could result in a few offline trials before a preferred closed-loop dynamics is reached. However, the availability of an expression for the expected closed-loop response eliminates the need for extensive searching for the most appropriate parameter values, such as is needed for the input and output weights in MPC. An example of a NMP SISO system is presented later in section 5.

For the multivariable case, we utilize the example of a 2 × 2 system because higher dimensional systems introduce no additional challenge or complications. The closed-loop transfer matrix around the MP part can be approximated by the diagonal matrix $\mathbf{G}_M^{CL}(s)$ below. The lowest order inner factor obtained from the algorithm in section 3.3 for the case of no time delay and one RHP transmission zero z is $\mathbf{G}_A(s)$ (note that for consistency with the SISO case the inner and outer factors are scaled such that the steady-state gain matrix of the inner factor is the identity matrix, i.e., $\mathbf{G}_A(0) = \mathbf{I}$).

$$\mathbf{G}_{\mathrm{M}}^{\mathrm{CL}}(s) = \begin{bmatrix} \dfrac{2\xi_1\tau_1 s + 1}{\tau^2 s^2 + 2\xi\tau s + 1} & 0 \\ 0 & \dfrac{2\xi_2\tau_2 s + 1}{\tau_2{}^2 s^2 + 2\xi_2\tau_2 s + 1} \end{bmatrix},$$

$$\mathbf{G}_{\mathrm{A}}(s) = \begin{bmatrix} \dfrac{d_{11}s + z}{s + z} & \dfrac{d_{12}s + z}{s + z} \\ \dfrac{d_{12}s}{s + z} & \dfrac{d_{22}s + z}{s + z} \end{bmatrix} \quad (8)$$

Here the d_{ij} are the elements of the symmetric $\mathbf{D}_{\mathrm{A}}$ matrix of the inner (AP) factor. Thus, the overall closed-loop transfer matrix $\mathbf{G}^{\mathrm{CL}}(s) = \mathbf{G}_{\mathrm{A}}(s)\ \mathbf{G}_{\mathrm{M}}^{\mathrm{CL}}(s)$ is

$$\begin{bmatrix} y_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} \dfrac{d_{11}s + z}{s + z}\ \dfrac{2\xi_1\tau_1 s + 1}{\tau_1{}^2 s^2 + 2\xi\tau s + 1} & \dfrac{d_{12}s}{s + z}\ \dfrac{2\xi_2\tau_2 s + 1}{\tau_2{}^2 s^2 + 2\xi_2\tau_2 s + 1} \\ \dfrac{d_{12}s}{s + z}\ \dfrac{2\xi_1\tau_1 s + 1}{\tau_1{}^2 s^2 + 2\xi_1\tau_1 s + 1} & \dfrac{d_{22}s + z}{s + z}\ \dfrac{2\xi_2\tau_2 s + 1}{\tau_2{}^2 s^2 + 2\xi_2\tau_2 s + 1} \end{bmatrix} \times \begin{bmatrix} y_1{}^{\mathrm{sp}} \\ y_2{}^{\mathrm{sp}} \end{bmatrix} \quad (9)$$

Thus, instead of the diagonal closed-loop transfer matrix $\mathbf{G}_{\mathrm{M}}^{\mathrm{CL}}(s)$ that would be obtained for a MP plant, a full 2×2 transfer matrix is obtained here. Each output is affected by both of the setpoints. Even though at steady state the outputs will naturally only be affected by their own setpoints, during the transient period the effect of the magnitude of the other setpoint change will be seen via the off-diagonal transfer function elements of the matrix in eq 9.

As in the SISO case above, simulations will have to be performed to determine the tuning for the NMP case. Values for ξ and τ corresponding to each of the outputs should be available if the designer has a feel for the desired shape and the speed of the response based on the open-loop dynamics. Then, with the help of eq 9, the expected response of the outputs for the chosen values of the tuning parameters and the desired setpoint changes can be calculated. This provides the designer with an offline look at whether the desired speeds and shapes of the two responses were obtained. If the obtained response is not satisfactory, the ξ and τ parameters can be changed until the desired response is arrived at. This final tuning set would then be used online. The impact of the ξ and τ parameters is explored in more detail with the help of a NMP MIMO system in section 6.

4. RS-Based MPC

RS-MPC is a technique that is directly based on the RSC formulations outlined in sections 2 and 3. Thus, its tuning is based on the same philosophy that is used in RSC, namely, the selection of a reference system or a desired closed-loop behavior. It handles NMP plants in the straightforward manner presented in the previous section. Being a model-based predictive control strategy, it incorporates constraint handling, a feature that is increasingly a requirement of modern process control. Also, because it can utilize nonlinear models, it offers a transparent and low-maintenance method for controlling the process over a wide operating range.

RS-MPC is conceptually close to an algorithm put forward by Brown et al..[31] However, there are algorithmic differences. First, Brown's algorithm uses rate of change constraints instead of the more explicit upper and lower limits on the input and output variables involved in the optimization. The desired rate of change is indicated to the controller via *specification curves*. This formulation and the use of specification curves leads to problems (see ref 17): (i) the controller requires action because of the constraint variable even if the variable is within its constraints and (ii) the number of tuning parameters required increases greatly because of the constants in the specification curves. Second, the one-step NLP formulation used to bring about computational savings could compromise constraint compliance when compared to multistep formulations. Third, a vital difference between Brown's strategy and the one proposed here is the fact that in the former no special consideration is given to NMP plants. Hence, whether the method leads to good results for a given NMP system will depend to a greater degree than desirable on the tuning parameter values. RS-MPC, on the other hand, explicitly accounts for RHP zeros and time delays. Because these noninvertible parts of the process model are explicitly accounted for in controller synthesis, the stability and tuning performance will be superior to those in Brown's approach.

In the next section, the RS-MPC algorithm is outlined in detail for the case of a general nonlinear NMP plant. This is followed by section 4.2, where some guidelines are provided for the selection of tuning parameter values.

4.1. RS-MPC Algorithm. The overall algorithm consists of two main steps: (i) determining the effect on the internal model of the last control move implemented on the plant and (ii) solving the optimization problem in order to determine future control moves. Each of these warrants a discussion of the motivation and the approach taken.

1. *Effect of the last move*: The model that is used by the controller to predict the future behavior of the process should approximate closely the present condition of the plant. This can be achieved at every control instant by using a state observer. In the present study, an "open-loop" state observer similar to the one in conventional MPC[32,33] was used. The predicted output values are compared to the incoming plant measurements to yield the current output prediction *bias* of the observer. This bias is assumed to be constant over the prediction horizon in the absence of a knowledge of future disturbances.[34,35] Such a choice of the observer has been found in the past to be satisfactory for stable processes.[36,37] For future implementations of the RS-MPC controller, a closed-loop observer such as a Kalman filter[38] will be considered.

2. *Determination of control moves*: The aim of this step is to find a set of future control moves that minimizes the deviations of the process from the reference trajectory. Following the receding horizon approach in traditional MPC,[39] only the first move is implemented on the plant and the entire two-step procedure repeated at the next sampling instant with updated measurements.

The full set of nonlinear equations incorporating the process model and the structure of RS-MPC is mostly a continuous implementation with a discrete approximation to the calculated control variable profile to be sent to the plant for implementation. The motivation for formulating the optimization problem in continuous time is related to the availability of convenient solution techniques. Approximation of the continuous-time problem by orthogonal collocation[36,40] yields a much smaller number of nonlinear algebraic equations compared to a discrete formulation of the problem.

The tasks that need to be accomplished at each control instant in determining the control moves can now be listed:

1. *Initialization*: The controller must be initialized with the current state, inputs, and outputs of the process as best as possible.

2. *One-step ahead prediction*: Take into account the effect of the last control move made on the states by integrating the observer (the open-loop model as in eq 1) over one sampling instant. Compare the model output $\mathbf{y}$ with the measured value $\hat{\mathbf{y}}$ to determine the error at the current time t_0 for every output, i.e., $\mathbf{b}(t_0) = \hat{\mathbf{y}}(t_0) - \mathbf{y}(t_0)$. This will form the bias term which brings feedback correction into the algorithm. Next, determine the output error integral for each output using the value $\mathcal{I}(t_{-1})$ from the previous control instant. That is, $\mathcal{I}(t_0) = \mathcal{I}(t_{-1}) + \int_{t_-1}^{t_0}[\mathbf{y}^{sp}(\theta) - \mathbf{y}(\theta)]\,d\theta$, where $\mathbf{b}$ and $\mathcal{I}$ are n_y-dimensional vectors.

3. *Computing the control move*: The optimization to be performed at each control instant solves for the unknown functions $\mathbf{u}(t)$, $\mathbf{x}(t)$, and $\mathbf{y}(t)$ and two slack variables per output λ_P and λ_N from the present to the prediction horizon [$t \in (t_0, t_0 + t_p)$; note that only the inputs $\mathbf{u}(t)$ are the independent variables because $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are constrained through the system equations]. In vector form

$$\min_{\mathbf{u}(t)} \int_{t_0+t_p}^{t_0} [\mathbf{w}_P^{\,T}(t)\, \lambda_P(t) + \mathbf{w}_N^{\,T}(t)\, \lambda_N(t)]\, dt \qquad (10)$$

such that

$$\mathbf{H}_M(\mathbf{x}(t))[\mathbf{f}(\mathbf{x}(t)) + \sum_{i=1}^{n_u} \mathbf{g}_i(\mathbf{x}(t))\, u_i(t)] + \lambda_P(t) - \lambda_N(t) = \mathbf{K}_1[\mathbf{y}^{sp}(t) - (\mathbf{h}_M(\mathbf{x}(t)) + \mathbf{b}(t_0))] + \mathbf{K}_2[\mathcal{I}(t_0) + \int_{t_0+t_p}^{t_0} [\mathbf{y}^{sp}(t) - (\mathbf{h}_M(\mathbf{x}(t)) + \mathbf{b}(t_0))]\, dt] \qquad (11a)$$

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}(t)) + \sum_{i=1}^{n_u} \mathbf{g}_i(\mathbf{x}(t))\, u_i(t), \quad \mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t)) \qquad (11b)$$

$$\mathbf{u}_{min} \le \mathbf{u}(t) \le \mathbf{u}_{max}, \quad \mathbf{y}_{min} \le \mathbf{y}(t) \le \mathbf{y}_{max}, \quad \mathbf{x}_{min} \le \mathbf{x}(t) \le \mathbf{x}_{max} \qquad (11c)$$

$$\left(\frac{d\mathbf{u}}{dt}\right)_{min} \le \frac{d\mathbf{u}(t)}{dt} \le \left(\frac{d\mathbf{u}}{dt}\right)_{max} \qquad (11d)$$

$$\lambda_P(t) \ge 0, \quad \lambda_N(t) \ge 0 \qquad (11e)$$

4. Approximate the control move to be sent to the plant based on the value of the optimum control variable $\mathbf{u}(t)$ between t_0 and $t_0 + \Delta T$.

The optimization problem to be solved at each instant is further explained here. Equation 10 is the objective function, which is simply a weighted sum of the slack variables over all of the outputs and time instants in the prediction horizon. λ_P and λ_N are positive slack variable functions. They are a measure of the deviation of the process model from the PI reference system constraint, eq 11a. Also in eq 11a, $\mathbf{H}_M(\mathbf{x}(t))$ is the Jacobian of the output map $\mathbf{h}_M(\mathbf{x}(t))$ of the MP factor of the process model with respect to the state vector. The reference system gains $\mathbf{K}_1$ and $\mathbf{K}_2$ are the diagonal tuning matrices referred to in section 2. They and the weighting functions $\mathbf{w}_P$ and $\mathbf{w}_N$ are important tuning parameters, and guidelines for their selection are provided in section 4.2.

The solution of the above algorithm results in continuous profiles for the state and manipulated variables obtained by interpolation of the values at the orthogonal collocation points. The sample interval value for $\Delta\mathbf{u}$ to be sent to the plant is calculated as an average of the values of the continuous profile $\mathbf{u}(t)$ at the start and end of the next sample interval. In all simulations reported here, three finite elements with three collocation points per finite element were used.

Developing a closed-loop relationship will be very difficult for a constrained, mixed continuous−discrete implementation such as RS-MPC. For this reason we do not present a closed-loop relationship such as eq 2 obtained for RSC. Instead, we argue intuitively that a very similiar relationship should hold for this predictive reference system formulation as well. Just as in RSC, RS-MPC utilizes state feedback for input−output decoupling and then imposes a desired closed-loop response via the PI reference system. Consequently, RS-MPC has a significant advantage over conventional MPC formulations in that its tuning is much more intuitive and directly linked to the desired closed-loop behavior. This is clearly demonstrated in the following sections of this paper by case studies.

4.2. Tuning of RS-MPC. The major tuning parameters in the RS-MPC methodology are (1) the prediction horizon (t_p), (2) the control horizon t_m, (3) the gains $\mathbf{K}_1$ and $\mathbf{K}_2$ that contain the ξ and τ parameters of the reference system for each output, and (4) the weights on the slack variables in the objective function. As will be shown in the following paragraphs, all of the tuning parameters can be easily determined via a knowledge of the process and its dynamics.

The first two tuning parameters, the prediction and control horizons, are chosen with respect to the natural dynamics of the process. A simple choice for the prediction horizon is for it to be related to the settling time of the process. The prediction horizon is then divided into three finite elements with three collocation points each. For the control horizon, the possible choices are up to the end of the first, second, or third finite element.[17] For a closed-loop performance in RS-MPC that is closest to the continuous-time RSC formulation, a control horizon equal to the prediction horizon is recommended.[41] If this leads to a large number of unknowns to be solved for, the control horizon can be shortened to the end of the first finite element.

One set of reference system parameters have to be chosen corresponding to each CV. This is best done with the help of eqs 2 and 3 and Figure 1. The damping coefficient ξ can be chosen based on the amount of overshoot that can be tolerated in that particular CV. A value of 4.0 was used in all simulations performed herein and yielded a fairly good performance. As was pointed out earlier, a ξ of 4.0 leads to a closed-loop rise time that is approximately half that of the τ value used.

Table 1. Parameter Values Used for the SISO Nonlinear System (Time in min)

OP I		OP II		common	
u_{ss}	0.0125	u_{ss}	0.015	a	1
z_1	0.185	z_2	0.032	b	0.21
k_1	0.05	k_2	0.29	c	0.0064
				α_1	−0.392
				α_2	0.951

Table 2. Summary of the Tuning Comparisons between RS-MPC and MPC for Nonlinear Systems

	good local tuning		good tuning over range	
process characteristic	RS-MPC	MPC	RS-MPC	MPC
steady-state gain nonlinearity	yes	yes	yes	no
time constant nonlinearity	yes	yes	yes	maybe
inverse response nonlinearity	yes	no	yes	no
dead-time nonlinearity	yes	yes	yes	yes
FCCU gain + time constant nonlinearity	yes	yes	yes	no

Based on the desired closed-loop dynamics of the process, an approximate value of the closed-loop rise time can be guessed. This can then be used to determine a value for τ. A good initial guess at the speed of response is the open-loop response of the process. If the obtained performance is not satisfactory, fine tuning of the desired closed-loop τ can be easily done because of the decoupling nature of RS-MPC.

Two separate λ functions give the designer the extra freedom to penalize positive and negative deviations differently through the appropriate choice of the weighting functions $\mathbf{w}_P$ and $\mathbf{w}_N$ (under the orthogonal collocation implementation, these functions are approximated by vectors that contain the values of the function at the collocation points; hence, these weight functions will often be referred to as *weighting vectors* in the remaining text). In addition to directional penalties, the weight functions can also be used to prioritize the importance among the different control variables (CVs). The effect of different units can also be accounted for via the weights. It must be stressed, however, that the reference system gains $\mathbf{K}_1$ and $\mathbf{K}_2$, defined through ξ_i and τ_i, are the primary tuning parameters; all other weight functions (i.e., $\mathbf{w}_P$ and $\mathbf{w}_N$) can be set equal to 1.

5. Tuning Advantages for SISO Plants

In this section the tuning features of RS-MPC in the context of SISO systems are explored. The studies conducted as a part of this project show that when the local linear dynamics of the underlying process change because of nonlinearities, RS-MPC tuning consistently leads to a closed-loop performance that is quite uniform, while this is often not the case for weight-based tuning methods. This was done with the help of four case studies. These cases examine the effect of a nonlinearity in (i) the steady-state gain, (ii) the dominant dynamics, (iii) the inverse response, and (iv) the dead time, on the closed-loop performance of RS-MPC and compare it to that of conventional MPC. In section 5.1, we present only one of these cases: a SISO system with nonlinear inverse response. This example highlights the major features of RS-MPC tuning and demonstrates the advantages of the RS-MPC tuning methodology over the weight-based tuning methods of conventional MPC. We summarize the other SISO cases examined in work by Kalra (1997) in Table 2.

5.1. Example of a System with Nonlinear Inverse Response. A SISO system is examined that exhibits the following nonlinearity: the extent of its inverse response changes significantly when moving from one operating point to another. With the help of this system, some tuning properties of RS-MPC were explored. In addition, how MPC and RS-MPC perform with regard to tuning over a broad operating range was also examined. The second-order nonlinear process that was used can be represented with the help of the following equations:

$$\dot{x}_1 = -\frac{b}{a}x_1 - \frac{c}{a}x_2 + u + d$$

$$\dot{x}_2 = x_1$$

$$y = k_1 z_1\left\{\frac{-x_1}{(\alpha_1 x_2 + \alpha_2)a} + \frac{x_2}{a}\right\} \tag{12}$$

The state equations $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},u,d)$ of this model are linear, but the output map $y = h(\mathbf{x})$ is nonlinear. The output equation can be linearized around a steady state (u_{ss}, $\mathbf{x}_{ss}$, y_{ss}) [here the steady-state values of the state and output can be expressed as $x_{1ss} = 0$, $x_{2ss} = (a/c)u_{ss}$, $y_{ss} = (k_1z_1/a)x_{2ss}$]. The transfer function at any given steady state will be

$$\frac{y(s)}{u(s)} = \frac{k_1 z_1}{\alpha_1 x_{2ss} + \alpha_2}\left\{\frac{-s + (\alpha_1 x_{2ss} + \alpha_2)}{as^2 + bs + c}\right\} \tag{13}$$

It can be seen that the location of the RHP zero at any steady state is $\alpha_1 x_{2ss} + \alpha_2$. The inverse response nonlinearity is due to the dependence of the RHP zero on the process' operating point through x_{2ss}. Further, the pole locations and the steady-state gain (k_1z_1/c) are completely independent of the operating point. These were the characteristics that motivated the form of the nonlinear model chosen in eq 12.

The values for the various parameters used here at the two operating points are summarized in Table 1. With these choices, the RHP zero locations at OP I and OP II are z_1 and z_2, respectively. The transfer functions at the two operating points are $k_1(-s + z_1)/(as^2 + bs + c)$ and $k_2(-s + z_2)/(as^2 + bs + c)$, where $k_2 = k_1z_1/z_2$. The step responses obtained from these transfer functions are plotted in Figure 3. As Table 1 shows, the location of the RHP zero is much closer to the origin at OP II as compared to OP I. This results in the significant difference in the magnitude of the inverse responses at the two operating points seen in Figure 3.

5.1.1. Tuning Advantages. A sample time of 1 min was used based on the open-loop dynamics of the process and the closed-loop bandwidth required. The prediction and control horizons were both picked to be 67 min, which is close to the 95% settling time of the open-loop process. The weight vectors in the objective function were picked to be unity for both of the slack variables corresponding to the positive as well as negative deviations from the reference trajectory. For the reference system parameters, $\xi = 4.0$ and $\tau = 133$ min were used, giving a response that was a little faster than the open-loop response. With local linear models instead of a nonlinear model used in the RS-MPC algorithm as constraints, controller calculations involved solving a linearly constrained optimization problem with a linear objective function (see eqs 10 and 11). This optimization

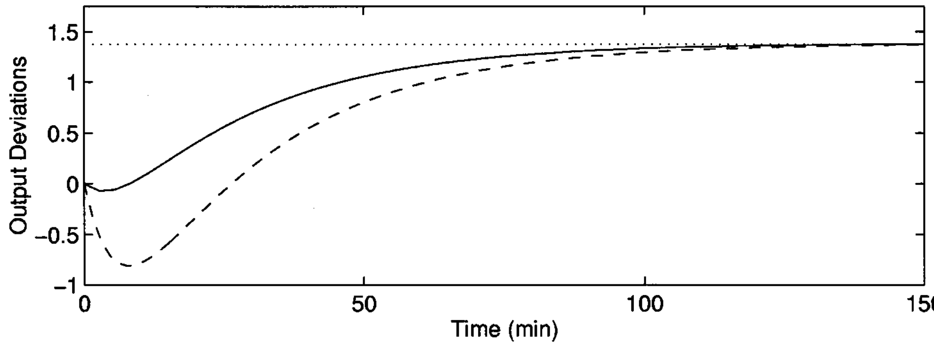

Figure 3. Comparison of the step responses of the models at OP I (solid) and OP II (dashed). In each case it is the deviation of the output from steady state that is plotted.

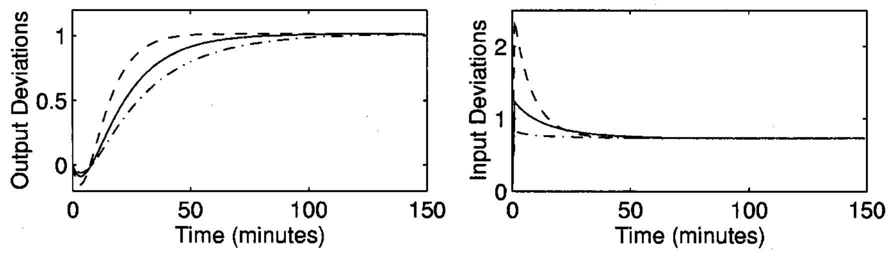

Figure 4. Closed-loop response of the system under RS-MPC to a setpoint change at OP I for $\tau = 67$ (dashed), 133 (solid), and 200 (dash-dotted).

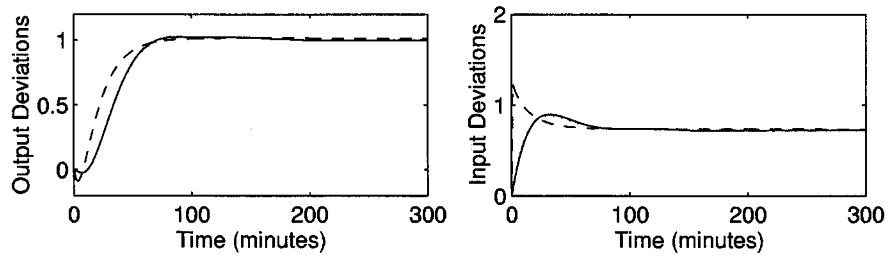

Figure 5. Closed-loop response of the system under RS-MPC (dashed) and MPC (solid) to a unit setpoint change at OP I. The original tuning was used for both of the controllers. The dotted line (behind the solid line and barely visible) corresponds to the case when a smaller control horizon $m = 20$ was used.

problem was solved at each sample instant using IMSL's[42] linear program solver (DLPRS).

With the tuning parameters as selected (referred to henceforth as the original tuning for RS-MPC), the servo control performance of RS-MPC was examined. The output response to a unit setpoint change is shown as the solid line in Figure 4. The presence of the RHP zero is what causes the inverse response in the closed-loop response. The response coincides with the closed-loop transfer function (see section 3) given by $G^{CL}(s) = [(-s + z_1)/(s + z_1)][(2\xi\tau s + 1)/(\tau^2 s^2 + 2\xi\tau s + 1)]$. Three different values of τ were tried out with a ξ of 4.0. These runs are shown in Figure 4 as well. It can be seen that the change in τ affects the closed-loop response in a direct and predictable way.

For examining the response of conventional MPC at this operating point, the nominal tuning values selected were as follows: control and prediction horizons $m = p = 67$ (a model horizon $n = 100$ was used; note that identical control and prediction horizons are being used for consistency with RS-MPC tuning; for nonzero $\mathbf{\Gamma}^u$, it was observed that changing the control horizon had no discernible effect on the response) and weighting matrices $\mathbf{\Gamma}^y = 1.0$ and $\mathbf{\Gamma}^u = 13.0$. A trial and error effort was involved in arriving at the weighting factor $\mathbf{\Gamma}^u$ that gave the desired response. This tuning set does extremely well in closed loop, as the solid lines in Figure 5 show. Here, the MPC response and the RS-MPC response are shown on the same figure so as to compare the two at OP I. Note that, in the RS-MPC response, a larger initial input movement is seen. A rate-constrained version of RS-MPC was also tested[17] that restricted the rate of change in the controlled variable without significantly affecting the dynamics of the RS-MPC controlled process.

Next, the performance of RS-MPC is compared with that of conventional MPC at the second operating point. The tuning is kept the same as it was at OP I, but the local linear models used in the two controllers are updated to correspond to the ones at the new operating point OP II. Figure 6 shows the two closed-loop setpoint step responses at OP II. The RS-MPC response (dashed line) appears to be a bit different from the one at OP I. This is because the NMP character of the open-loop process has changed because of the greater effect of the RHP zero at this operating point. The closed-loop behavior obtained is exactly what was expected from the reference system and is described by the corresponding closed-loop transfer function $[(-s + z_2)/(s + z_2)][(2\xi\tau s + 1)/(\tau^2 s^2 + 2\xi\tau s + 1)]$, which is different from that at OP I because of the location of the RHP zero.

It can be seen in Figure 6 that while RS-MPC gives a response identical with the one predicted for it at OP

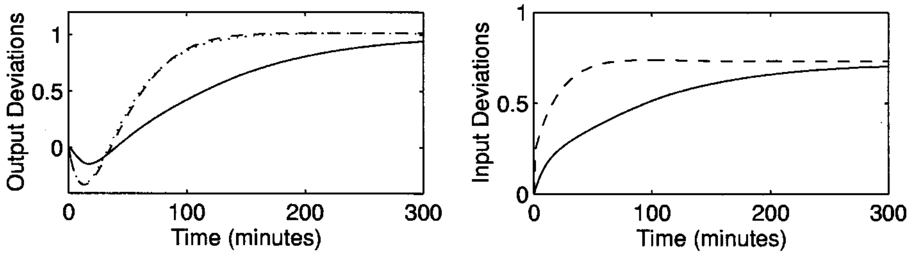

Figure 6. Closed-loop responses of the system under RS-MPC (dashed) and MPC (solid) to a unit setpoint change at OP II with the original tuning. The dotted line (coincident with the dashed line) corresponds to the response predicted by the closed-loop transfer function.

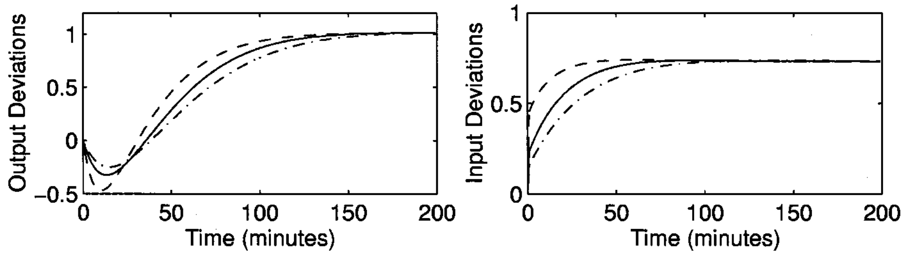

Figure 7. Closed-loop response of the system under RS-MPC to a setpoint change at OP II for $\tau = 67$ (dashed), 133 (solid), and 200 (dash-dotted).

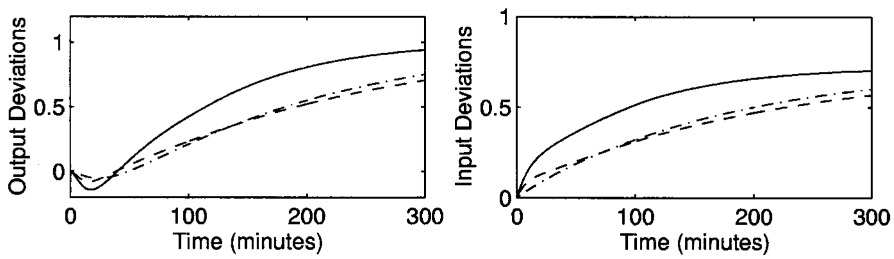

Figure 8. Closed-loop response of the system under MPC to a setpoint change at OP II for $\mathbf{\Gamma}^u = 9$ (dashed), 13 (solid), and 52 (dash-dotted).

II, the performance of MPC has degraded considerably. The output response is very sluggish and has not reached the new setpoint even at the end of 300 min. Thus, the presence of the NMP nonlinearity, which significantly alters the local inverse response at the two operating points, has a considerable impact on conventional MPC tuning characteristics. The values of the prediction and control horizons and the weight values that work well at one operating point can no longer be relied upon to give a good performance at other operating points. This illustrates the major drawback of conventional MPC tuning in handling systems with nonlinearities.

The local tuning capabilities of the two controllers at OP II were also examined. For RS-MPC it was found that the closed-loop behavior has the same dependence on τ and ξ as at OP I (Figure 7). For conventional MPC, results of the simulations carried out with three different values of the input weight led to some surprising observations.[17] The large inverse response in the process caused the simulations with $\mathbf{\Gamma}^u$ values of less than 9 to go unstable. Because NMP characteristics were not explicitly taken into account in the MPC formulation, tuning values had to be chosen carefully to prevent unstable closed-loop dynamics. Another fallout of this fact was the nonintuitiveness of the effect of the input weight at the present operating point. From the nominal value, both decreasing as well as increasing $\mathbf{\Gamma}^u$ led to a slower closed-loop response (Figure 8).

The deterioration in MPC performance that was seen in servo control simulations is evident in disturbance rejection as well. In ref 17, the rejection of a unit step in the disturbance d by MPC and RS-MPC is compared at OP I and OP II. While the overshoot for MPC was slightly larger, the controllers did equally well in rejecting the disturbance with the original tuning. However, at OP II, the MPC response was slower and the controller took much longer to bring the process output back to setpoint. Thus, in both servo as well as regulatory control simulations, it is seen that RS-MPC with a fixed set of tuning parameter values performs better at different operating points than conventional MPC does.

6. Tuning Advantages for MIMO Plants

Some of the most significant advantages expected from RS-MPC over conventional MPC are in the context of MIMO systems. The decoupling implicit in RS-MPC and the tuning parameters related in a simple manner to the desired closed-loop response makes initial selection of tuning values straightforward and fine tuning easier. The model-based input−output linearization brought about via state feedback ensures a uniform closed-loop performance in the presence of nonlinear characteristics.

The tuning features and capabilities of the RS-MPC approach for MIMO systems are explored in detail in ref 17. Here just one of the cases is presented. RS-MPC is applied on the Amoco FCCU simulation, a realistic process that requires the identification of linear black box models for process control. As was the case in

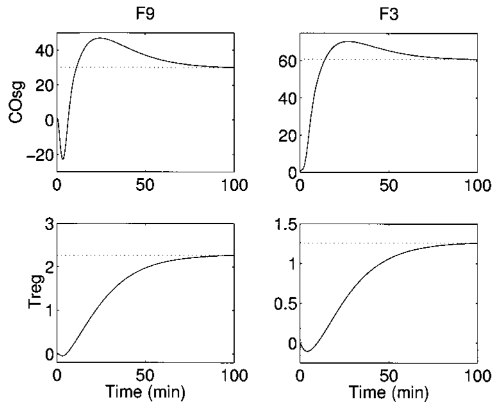

Figure 9. Open-loop model of the FCCU 2 × 2 subsystem at OP A.

section 5, only RS-MPC utilizing local linear models is examined. The transition problem from one steady state to another using nonlinear models will be treated in a separate publication.

6.1. Amoco Model IV FCCU. With a detailed model and simulation already available, Amoco FCCU was the natural choice of a process to demonstrate the tuning advantages of RS-MPC. It is a very realistic simulation of a complex and important industrial process that affords a number of opportunities and potential benefits from the application of improved advanced control techniques.[43] For the present case, a 2 × 2 subsystem of the FCCU is used. For this multivariable system, the outputs are the carbon monoxide concentration in the regenerator stack gas (CO_{sg}) and regenerator temperature (T_{reg}), and they are controlled by the inputs regenerator lift air flow rate (F_9) and feed flow rate (F_3). The details of how the black box input–output models were identified can be found in ref 17.

To begin with, the advantages of the RS methodology in FCCU operation at a given steady state (OP A) is examined. Because there is little evidence of local nonlinearity, linear models as identified above were used. Tuning was once again done with the help of the open-loop dynamics of the process. The step responses of the four input–output pairs of the MIMO linear model identified at OP A are shown in Figure 9. A sample time of 1 min was considered adequate based on the open-loop response and the bandwidth required. A settling time in the 60–70 min range led to the choice of a prediction horizon of 67 min (4000 s). To achieve the closest possible response to RSC, the control horizon was set to 67 min as well.

It was observed from the process that stack gas carbon monoxide variation to unit step changes in the two inputs is approximately an order of magnitude greater than that for the regenerator temperature. With a view to normalizing the relative importance of these two reference systems in the objective function, weights for the CO reference system that are one-tenth those of the regenerator temperature were used. Thus, the $\mathbf{w}_N$ and $\mathbf{w}_P$ vectors were each assigned as $[0.1, 1]^T$. A ξ value of 4.0 was picked for the reference systems of both of the outputs because it strikes a good balance in terms of the overshoot and the speed of the response (Figure 1). The value of τ was chosen based on a feel for the open-loop behavior of the plant and the desired closed-loop dynamics. As shown in Figure 9, the carbon monoxide response to the two inputs is faster than that of the regenerator temperature. With that in mind, rise times of 17 min for carbon monoxide and 33 min for the regenerator temperature were decided upon because they are reflective of the open-loop speed of response of these outputs. With the chosen ξ values, a simple calculation based on eq 3 shows that these rise times correspond to τ values of 33 and 67 min for CO_{sg} and T_{reg}, respectively [roundoff in converting these values from seconds to minutes is the reason for the τ values not being exactly double the chosen rise times; for example, for CO_{sg} the exact rise time desired was 1000 s (16.67 min), giving a τ of 2000 s (33.33 min)].

With the tuning parameters as selected above, RS-MPC was implemented on the FCCU simulation at OP A. The solid line in Figure 10 shows the servo response of the controller to setpoint changes of 2.5 ppm in the carbon monoxide concentration and 0.5 °F in the regenerator temperature. It can be seen that the response is satisfactory; approximately halfway into the run the outputs are already close to their new setpoints. Thus, unlike the weight-based approach of conventional MPC, the RS-MPC methodology led to a set of tuning parameters that gave an acceptable closed-loop performance.

It must be pointed out that although a rise time of about 17 min was asked for carbon monoxide the actual value obtained was about 9 min. One reason for this difference is that while the expected values of the rise time are being calculated for a continuous-time reference system, the actual controller is a discrete implementation sampling at the finite rate of once every 1 min. However, the major cause of the disparity is the mismatch between the real process and the low-order model identified via black box system identification techniques. When RS-MPC was implemented on a

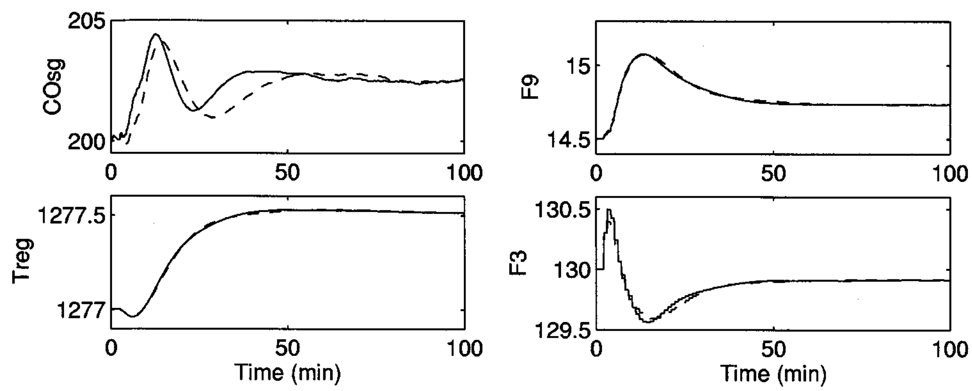

Figure 10. Effect of the reference system on the carbon monoxide concentration on the closed-loop performance: $\tau = 33$ (solid) and 150 (dashed) with $\xi = 4.0$.

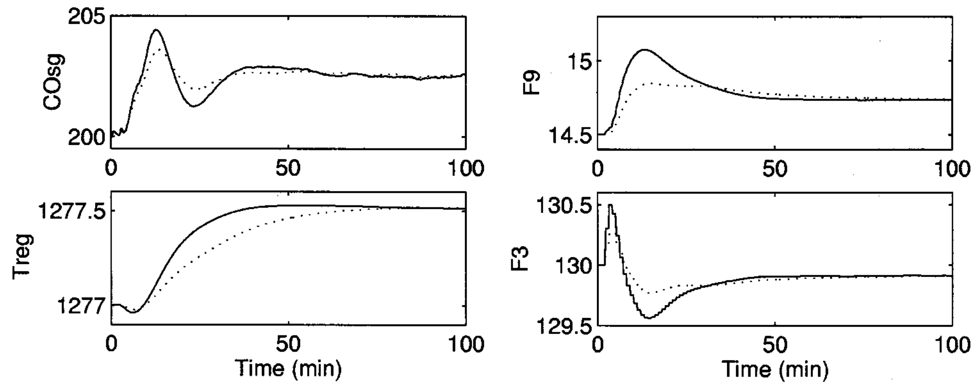

Figure 11. Effect of the reference system on the regenerator temperature on the closed-loop performance: $\tau = 67$ (solid) and 132 (dotted) with $\xi = 4.0$.

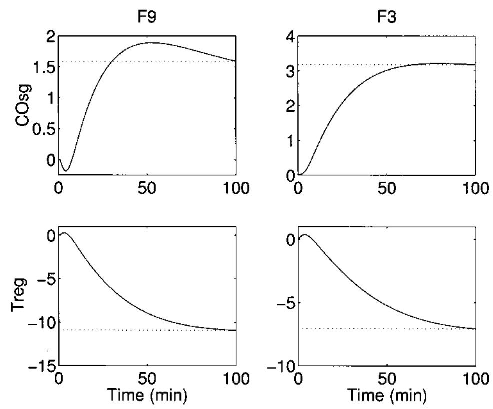

Figure 12. Open-loop model of the FCCU 2×2 subsystem at OP B.

system where the process is represented by the model used in the controller, the rise times obtained were identical with the expected values.

The decoupling property of RS-MPC provides a way of quickly getting the desired performance if, because of plant–model mismatch, one does not obtain it on the first pass. If one desires a slower stack gas carbon monoxide response to the setpoint change, one can simply increase τ in the reference system corresponding to it. In Figure 10 τ for carbon monoxide was increased from 33 to 150 min, leaving τ for the regenerator temperature unchanged. It can clearly be seen from Figure 10 that the carbon monoxide response is slower than it was before while the regenerator temperature was unchanged. Similarly, if it was the regenerator temperature response that was required to be slower, its reference system would be changed. This is shown in Figure 11. These results reemphasize the point that only the variable whose reference system is altered via one of its parameters is mostly affected in its closed-loop performance, leaving the response of the other variable as it was.

The tuning advantages of RS-MPC over conventional MPC are more marked if the underlying process dynamics changes significantly. Operating point OP B (Figure 12) has a slower dynamics and a smaller steady-state gain for carbon monoxide. Furthermore, the regenerator temperature process gain is now negative and 4 times larger than what it was at OP A.[17]

To begin with, the performance of MPC and RS-MPC using the local linear models was compared at the old operating point OP A. Figure 13 shows the responses in carrying out a setpoint change of 2.5 ppm and 0.5 °F in CO_{sg} and T_{reg}, respectively. The tuning for RS-MPC was the nominal set that was used in Figure 10. For MPC, the tuning was one that is arrived at after considerable trial and error.[8,17] Values were $n = 120$, $m = p = 60$, $\boldsymbol{\Gamma}^y = \mathrm{diag}(0.4, 4.0)$, and $\boldsymbol{\Gamma}^u = 5\ \mathrm{diag}(2, 0.5)$. It can be seen from Figure 13 that both of the controllers perform satisfactorily at this operating point.

Simulations were now performed at the new steady state, OP B, with the same tuning but with the linear models now identified from this new operating point. Figure 14 shows the results of these runs. With the same setpoint changes and tuning as before, it can clearly be seen that conventional MPC response is poor because of evidence of oscillations. RS-MPC, on the other hand, is able to bring the carbon monoxide stack gas concentration and regenerator temperature to their new setpoints by the end of the simulation. It achieves this without the inputs oscillating, as was the case for

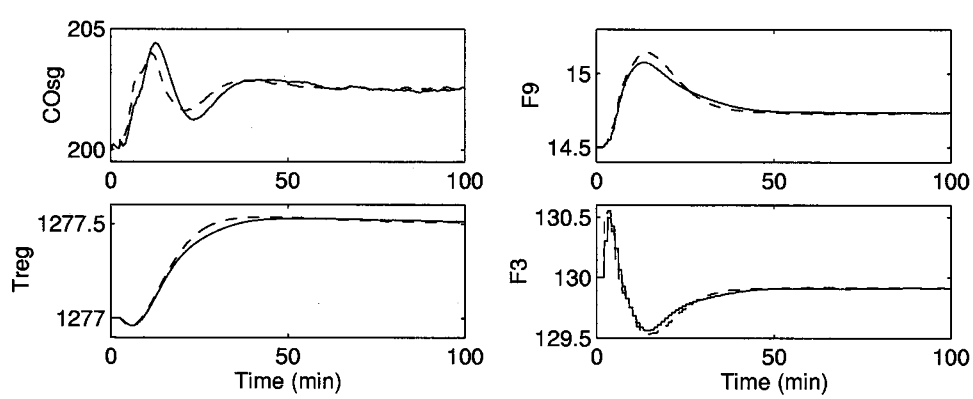

Figure 13. Comparison of conventional MPC (dashed) and RS-MPC (solid) at OP A. The nominal tuning set is used for both of the controllers.

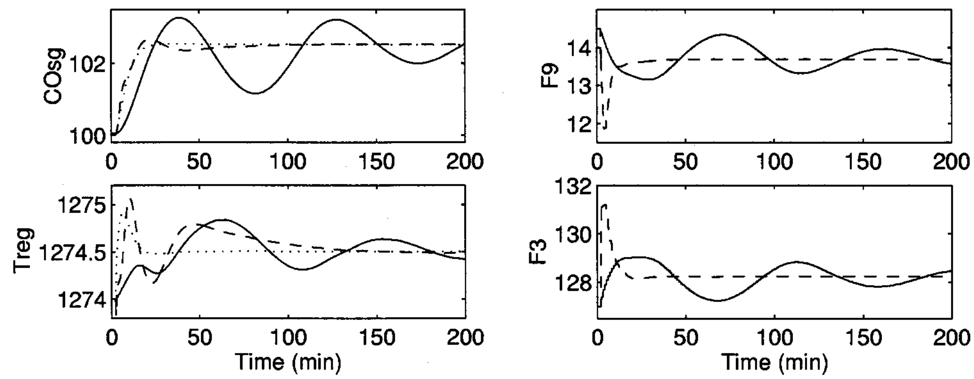

Figure 14. Comparison of conventional MPC (solid) and RS-MPC (dashed) at OP B. The nominal tuning set is used for both of the controllers.

conventional MPC. This illustrates the most important tuning advantage of the reference system approach over weight-based methods used in conventional MPC. Despite drastic changes in the underlying dynamics of the process, RS-MPC delivers a stable, acceptable closed-loop performance without requiring a retuning of the controller.

Note that the very different local characteristics at this operating point affected the closed-loop performance. This resulted in a closed-loop response that is not identical with that at OP A. Most notably, the regenerator temperature shows evidence of a large overshoot initially. While initially counterintuitive, these responses are entirely consistent with the NMP nature of this process. RS-MPC is expected to give the same closed loop at different operating points even as the underlying process' dynamics changes but only for MP systems. As is clearly indicated via eqs 7 and 9, any changes in the inner (AP) factor's dynamics will be passed through and shown in the closed-loop response.

The advantage of RS-MPC applied to these difficult processes was alluded to in section 3.4. It is that the methodology provides a means to predict this effect a priori with the help of the process model. For this NMP system, the model was factorized into AP and MP parts[17] for use in the controller. The predicted closed-loop transfer function, $\mathbf{G}_A(s)\,\mathbf{G}_M^{CL}(s)$, can be written as (see section 3.4)

$$\mathbf{G}^{CL} = \begin{bmatrix} \frac{0.22(s+0.57+2.13i)(s+0.57-2.13i)}{(s+3.6)(s+0.3)} & \frac{s(-s+3.78)}{(s+3.6)(s+0.3)} \\ \frac{s(s+3.78)}{(s+3.6)(s+0.3)} & \frac{0.22(s-0.57+2.13i)(s-0.57-2.13i)}{(s+3.6)(s+0.3)} \end{bmatrix} \times \begin{bmatrix} \frac{264s+1}{1089s^2+264s+1} & 0 \\ 0 & \frac{528s+1}{4356s^2+528s+1} \end{bmatrix} \quad (14)$$

The response of this closed-loop transfer function to the setpoint changes of 2.5 ppm and 0.5 °F in CO_{sg} and T_{reg}, respectively, is exactly the dotted line in Figure 14. Thus, clearly the differences in the RS-MPC responses between the two operating points are not due to controller performance issues. Rather, they can be attributed to the significantly larger impact of the AP factor $\mathbf{G}_A(s)$, i.e., the NMP nature of the plant, at OP B.

Thus, again it is seen that the RS-MPC tuning approach gives a performance which is consistent and quite predictable, given the local model and the tuning parameter values used. Conventional MPC, on the other hand, does not yield a good closed-loop response with the original tuning weights (see the appendix for a comparison between RSC and IDCOM). Once the underlying process characteristics change, it may require a retuning in order to give a good performance. This is a major advantage of RS-MPC in the context of nonlinear systems.

The value of this result cannot be underestimated. Using the RS-MPC approach, a controller is now possible that will perform satisfactorily over a wide operating range to the extent allowed by the NMP characteristics of the process. If a nonlinear model is used, this controller could seamlessly, and without operator intervention, handle the different steady states called for byproduct grade changes, feedstock variations, and any other kind of measured or unmeasured disturbances.

7. Summary of Case Studies

RS-MPC's advantages over the weight-based tuning of conventional MPC were examined as a part of this study with the help of five cases, two of which were presented here. Table 2 summarizes the conclusions made in the case studies. The controllers were evaluated based on the local tuning performance, i.e., good closed-loop response at a given operating point as well as an intuitive effect on the response when a tuning parameter change is made. The controllers were also assessed with regard to how well they perform when a wide operating range of the nonlinear process is considered. As can be seen from the table, RS-MPC does well in both of the categories for all of the cases studied. Conventional MPC, on the other hand, poses considerable difficulties in tuning at an operating point where inverse response characteristics are significant (section 5.1). More importantly, it was found that when a wide operating range is considered, conventional MPC does not perform well in retaining a uniform closed-loop response in three out of the five cases considered, with a fourth case being borderline. This reiterates the advantages of the RS-MPC approach in the context of nonlinear processes.

The tuning advantages of RS-MPC can be summarized as follows:

(i) The methodology allows the designer to specify a realistic closed-loop response through tuning parameter values arrived at without trial and error.

(ii) The closed-loop nominal response that will be obtained in the unconstrained case can be predicted a priori based on the tuning parameters used *and* a knowledge of the model of the process.

(iii) The effect of a change in the tuning parameter can be predicted using the expression for the expected closed-loop response of the system (such as eq 9 for a linear NMP system with one RHP zero).

(iv) For nonlinear processes, RS-MPC provides a uniform closed-loop response over the operating range of the process.

Even for NMP processes, the RS-MPC tuning approach retains the important advantage of a simple tuning methodology that is considerably less time-consuming than a trial and error approach such as the one used in MPC.

8. Conclusions

The major goal of this paper is the presentation of a new nonlinear MPC scheme and its tuning advantages. RS-MPC is a constraint handling, nonlinear control methodology which is based on RSS. It does not have the drawbacks of RSS and RSC in that it is suitable even for NMP plants and incorporates a methodology for the handling of constraints. RS-MPC's major advantage is that its tuning is considerably more intuitive than that of the weight-based methods currently used in predictive control schemes, leading to an easier determination of the tuning parameters. Furthermore, because its tuning is based on the specification by the designer of a desired closed-loop response, a uniform performance is obtained over a large operating range of highly nonlinear processes. This eliminates the need to constantly retune the controller, as is necessary in currently popular MPC techniques that use weights for tuning, when the process changes steady states in response to disturbances or product grade changes.

Appendix: Simulation-Based Comparison of RS-MPC with IDCOM

In this section a brief comparison of RSC with an IDCOM-like[11] setpoint trajectory MPC formulation is presented. Simulation results are presented at two operating points. A fixed set of tuning parameter values and local linear models are used in both of the controllers at each operating point. It is clear from this example that RSC, and by extension RS-MPC, provides a way of retaining tuning uniformity despite changes in the underlying process' response.

The MIMO models from section 6 were again used in this example. Details of these models which were obtained from the FCCU simulation using black box system identification techniques can be found in Appendix C of ref 17. The open-loop responses of these models are shown in Figures 9 and 12. There are two main differences from the simulations in section 6. The first and major one is that there is no plant−model mismatch in the simulations presented here. The models used to simulate the process are identical with the ones used in the controllers. Now, as shown in Figure 6 of section 5.1.1, the RSC response is nearly identical with that obtained by RS-MPC when there is no plant−model mismatch. Differences could exist because of the discrete time nature of RS-MPC, but with the sampling times chosen, these are insignificant. This brings up the second difference from the simulations in section 6: RSC rather than the discrete time RS-MPC was used for the simulations here.

Figure 15 shows the servo responses of the two controllers at the first operating point OP A. These responses are to a setpoint change of 2.5 ppm in the carbon monoxide concentration and 0.5 °F in the regen-

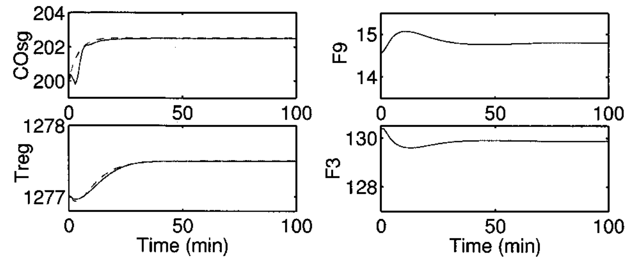

Figure 15. Comparison of RSC (dashed) and IDCOM (solid) at OP A.

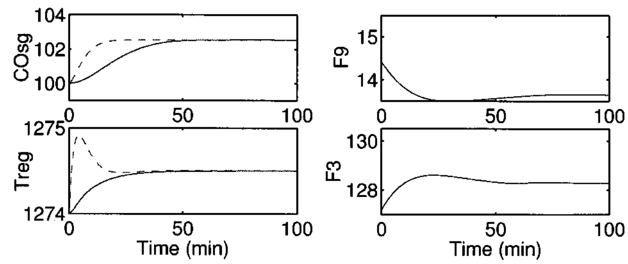

Figure 16. Comparison of RSC (dashed) and IDCOM (solid) at OP B. The tuning for both of the controllers is identical with the one used at OP A.

erator temperature. For RSC, τ values of 33 and 67 min for CO_{sg} and T_{reg}, respectively, and ξ values of 4.0 for both were used. This choice of tuning parameter values is identical with that used in section 6. The RSC response is shown by the dashed line in Figure 15.

To approximate IDCOM, a first-order setpoint trajectory was imposed onto a conventional MPC formulation. For MPC, the tuning was once again identical with the one used in section 6. Specifically, the parameter values were $n = 120$, $m = p = 60$, $\mathbf{\Gamma}^y = \mathrm{diag}(0.4, 4.0)$, and $\mathbf{\Gamma}^u = 5\ \mathrm{diag}(2, 0.5)$. The first-order setpoint filter time constants used were 5 and 10 for CO_{sg} and T_{reg}, respectively. These choices of the reference trajectory time constants were chosen so that at OP A the IDCOM responses are similar to RSC responses. Indeed, as the solid line curves in Figure 15 show, the IDCOM tuning used results in a response very close to that from RSC, thus setting up a fair baseline for comparison at OP B.

Next, the performance of the two controllers at the second operating point OP B is compared. As indicated earlier in this paper, the process' open-loop response at OP B (Figure 12) is quite different from that at OP A (Figures 9). Just as in section 6, local linear models were used in each of the controllers. The tuning parameter values that were used at OP A were once again utilized for the simulations at OP B.

Figure 16 shows the responses of the two controllers at OP B to the 2.5 ppm and 0.5 °F setpoint changes in the carbon monoxide concentration and the regenerator temperature, respectively. The RSC response (dashed line) is seen to have changed from that at OP A. While the CO_{sg} response is close to what it was before, the T_{reg} response is now significantly quicker and has a prominent overshoot. This change in RSC response at OP B is completely expected and predictable. As explained earlier in section 6.1, it is an outcome of the noninvertible, AP part of the process at OP B. In fact, the dashed line RSC response is identical with the dotted line in Figure 14 and is given by eq 14.

The solid line in Figure 16 shows the IDCOM response at OP B. It can be seen that the response, while good for T_{reg}, has deteriorated significantly for CO_{sg}. The controller now takes over 40 min to settle the process as compared to 10 min at OP A. In other words, when the underlying process' open-loop response changes, IDCOM's closed-loop performance changes. Further, as seen here, IDCOM's response could change in unpredictable ways. In this instance, it was not obvious that at OP B it would be the CO_{sg} response that would be significantly impacted. Clearly, a uniform and/or predictable closed-loop response over a wide operating range cannot be assured simply via the imposition of a setpoint reference trajectory such as the one used in IDCOM.

This simple example illustrates one of the key theses of this paper. RSC, and by extension RS-MPC, via state feedback used to effect input−output linearization and decoupling, provides a way of imposing a desired closed-loop response on the process. This same uniform response will then be obtained despite changes in the underlying dynamics of the process. For NMP plants, the response is affected by the dynamics of the noninvertible part of the process. However, the response will always be predictable, given the model of the process used in the controller. These are advantages not offered by weight-based tuning methodologies such as MPC. They are also not available in IDCOM-like methodologies where a setpoint trajectory is imposed on the outputs without state feedback for input−output linearization and decoupling.

Literature Cited

(1) Bartusiak, D.; Georgakis, C.; Reilly, M. Nonlinear feedback/feedforward control structures designed by reference system synthesis. *Chem. Eng. Sci.* **1989**, *44*, 1837.

(2) Lee, P.; Sullivan, G. Generic model control (GMC). *Comput. Chem. Eng.* **1989**, *12*, 573.

(3) Bartee, J. Identification and Control of Nonlinear Processes via Reference System Synthesis. Ph.D. Dissertation, Lehigh University, Bethlehem, PA, 1995.

(4) Adebekun, V.; Schork, F. Continuous solution polymerization reactor control. 1. Nonlinear reference control of methyl methacrylate polymerization. *Ind. Eng. Chem. Res.* **1989**, *28,* 1308.

(5) Adebekun, D.; Schork, F. Continuous solution polymerization reactor control. 2. estimation and nonlinear reference control during methyl methacrylate polymerization. *Ind. Eng. Chem. Res.* **1989**, *28*, 1846.

(6) Qin, S.; Badgwell, T. A. An overview of industrial model predictive control technology. *Preprints*−*Chemical Process Control*; CACHE: Austin, TX, 1996; Vol. V.

(7) Garcia, C. E.; Morari, M. Internal model control. 3. Multivariable control law computation and tuning guidelines. *Ind. Eng. Chem. Process Des. Dev.* **1985**, *24*, 484.

(8) Kalra, L.; Georgakis, C. Reference system based model predictive control. *Proceedings of the 13th Triennial IFAC World Congress,* San Francisco, CA, 1996; International Federation of Automatic Control: Laxenburg, Austria, 1996.

(9) Shridhar, R.; Cooper, D. J. A Tuning Strategy for Unconstrained SISO Model Predictive Control. *Ind. Eng. Chem. Res.* **1997**, *36*, 729.

(10) Shridhar, R.; Cooper, D. J. A Tuning Strategy for Unconstrained Multivariable Model Predictive Control. *Ind. Eng. Chem. Res.* **1998**, *37*, 4003.

(11) Richalet, J.; Rault, A.; Testud, J. L.; Papon, J. Model Predictive Heuristic Control: Applications to Industrial Processes. *Automatica* **1978**, *14*, 413.

(12) Henson, M.; Seborg, D. Input−Output Linearization of General Nonlinear Processes. *AIChE J.* **1990**, *37*, 1753.

(13) Holt, B. R.; Morari, M. Design of resilient processing plants−VI. The effect of right-half-plane zeros on dynamic resilience. *Chem. Eng. Sci.* **1985**, *40*, 59.

(14) Garcia, C.; Morari, M. Internal model control. 1. A unifying review and some new results. *Ind. Eng. Chem. Process Des. Dev.* **1982**, *21*, 308.

(15) Morari, M.; Zafiriou, E. *Robust Process Control*; Prentice Hall: Englewood Cliffs, NJ, 1989.

(16) Wright, R.; Kravaris, C. Nonminimum phase compensation for nonlinear processes. *AIChE J.* **1992**, *38*, 26.

(17) Kalra, L. Reference System Based Model Predictive Control of Nonlinear Processes. Ph.D. Dissertation, Lehigh University, Bethlehem, PA, 1997.

(18) Chen, C. *Linear System Theory and Design*; Oxford University Press, Inc.: New York, 1998.

(19) Kravaris, C.; Chung, C. Nonlinear state feedback synthesis by global input/output linearization. *AIChE J.* **1987**, *33*, 592.

(20) Kravaris, C. Input/output linearization: A nonlinear analogue of placing poles at process zeros. *AIChE J.* **1988**, *34*, 1803.

(21) Kravaris, C.; Soroush, M. Synthesis of multivariable nonlinear controllers by input/output linearization. *AIChE J.* **1990**, *35*, 1535.

(22) Jerome, N. F.; Ray, W. Model-predictive control of linear multivariable systems having time delays and right-half-plane zeros. *Chem. Eng. Sci.* **1992**, *47*, 763.

(23) Rawlings, J.; Muske, K. The Stability of Constrained Receding Horizon Control. *IEEE Trans. Autom. Control* **1993**, *38*, 1512.

(24) Kalra, L.; Georgakis, C. The effect of process nonlinearity on the performance of linear model predictive controllers for the environmentally safe operation of a Fluid Catalytic Cracking Unit. *Ind. Eng. Chem. Res.* **1994**, *33*, 3063.

(25) Ramanathan, S.; Curl, R.; Kravaris, C. Dynamics and control of quasirational systems. *AIChE J.* **1989**, *35*, 1017.

(26) Ball, J.; Helton, J. Inner−outer factorization of nonlinear operators. *J. Funct. Anal.* **1989**, *104*, 363.

(27) van der Schaft, A.; Ball, J. Nonlinear inner-outer factorization. Proceedings of the 33rd Conference on Decision and Control, 1994.

(28) Kravaris, C.; Daoutidis, P. Nonlinear state feedback control of second-order nonminimum-phase nonlinear systems. *Comput. Chem. Eng.* **1990**, *14*, 439.

(29) Doyle, F., III; Allgower, F.; Morari, M. A normal form approach to approximate input-output linearization for maximum phase nonlinear SISO systems. *IEEE Trans. Autom. Control* **1996**, *41*, 305.

(30) Friedland, B. *Control System Design*−*An introduction to state space methods*; McGraw-Hill: New York, 1986.

(31) Brown, M.; Lee, P.; Sullivan, G.; Zhou, W. A constrained nonlinear multivariable control algorithm. *Trans. Inst. Chem. Eng.* **1990**, *68*, 464.

(32) Lee, J.; Yu, Z. Tuning of model predictive controllers for robust performance. *Comput. Chem. Eng.* **1994**, *18*, 15.

(33) Li, S.; Lim, K.; Fisher, D. A state space formulation for model predictive control. *AIChE J.* **1989**, *35*, 241.

(34) Patwardhan, A.; Rawlings, J.; Edgar, T. Nonlinear model-predictive control. *Chem. Eng. Commun.* **1990**, *87*, 123.

(35) Sistu, P.; Gopinath, R.; Bequette, B. Computational issues in nonlinear predictive control. *Comput. Chem. Eng.* **1993**, *17*, 361.

(36) Eaton, J.; Rawlings, J. Feedback control of chemical processes using on-line optimization techinques. *Comput. Chem. Eng.* **1990**, *14*, 469.

(37) Daoutidis, P.; Kravaris, C. Dynamic output feedback control of minimum-phase multivariable nonlinear processes. *Chem. Eng. Sci.* **1994**, *49*, 433.

(38) Stengel, F. *Stochastic Optimal Control*, 1st ed.; John Wiley and Sons: New York, 1986; Chapter 3.

(39) Morari, M.; Prett, D.; Garcia, C.; Lee, J. Model Predictive Control. Draft prepublication copy, 1993.

(40) Cuthrell, J.; Beigler, L. Simultaneous optimization and solution methods for batch reactor control profiles. *Comput. Chem. Eng.* **1989**, *13*, 49.

(41) Kalra, L.; Georgakis, C. *Reference system based model predictive control*; Research Progress Report 22; Chemical Process Modeling and Control Research Center, Lehigh University: Bethlehem, PA, 1996.

(42) IMSL. *IMSL Problem Solving Software Systems*, 1.1 ed.; Visual Numerics: San Ramon, CA, 1989.

(43) McFarlane, R. C.; Reineman, R. C.; Bartee, J.; Georgakis, C. Dynamic simulator for a Model IV Fluid Catalytic Cracking Unit. *Comput. Chem. Eng.* **1993**, *17*, 275.

Received for review October 16, 2000
Revised manuscript received December 31, 2001
Accepted April 5, 2002

IE000891J