Algebraic Solution to H_2 Control Problems. II. The Multivariable Decoupling Case

Weidong Zhang,*,† Chen Lin,‡ and Linlin Ou†

Department of Automation, Shanghai Jiaotong University, Shanghai 200030, People's Republic of China, and Department of Space Science, Lulea University of Technology, 97187 Lulea, Sweden

The control of multivariable systems with time delays is an important problem in linear control systems. The goal of this paper is to develop an algebraic design method for multivariable systems with multiple time delays. In the method, the algebra-based input-output design technique is adopted, because it is easy to understand and use, and it requires no state variables. We focus mainly on decoupled response. First, all stabilizing decoupling controllers are parametrized. Second, the optimal decoupling controller is derived by minimizing the sensitivity function. Finally, the performance and robustness are analyzed and a simple tuning rule is developed for quantitative performance and robustness. The most important feature of the proposed method is that the controller is optimal, analytical, and can provide decoupling response. Typical examples are provided to illustrate the proposed method.

1. Introduction

Since J. C. Maxwell formulated his mathematical theory that was related to automatic control in 1868,[1] the area of linear control system design has advanced rapidly. Early works focused mainly on single-input/single-output (SISO) control problems. Inspired by the need to solve multi-input/multi-output (MIMO) control problems that were arising from aerospace and defense industries, some new design theories were developed from 1950 to 1960. Among those theories, the Linear Quadratic Gaussian (LQG) design method was particularly successful.[2,3]

In the 1980s, it was recognized that LQG controllers could exhibit poor robustness properties, although, with hindsight, it is now known that much improved robustness can be obtained by following proper design procedures.[4,5] The recognition of these potential problems led to the development of methods to improve LQG design, and, more importantly, it stimulated the development of a different optimal control theory called H_∞ control.

Zames[23] noted in 1981 that minimizing a new measure of system performance, namely, the H_∞ norm, could provide significant advantages, particularly for systems that were uncertain. Although sensitivity and robustness problems had been considered by other researchers, it was the work of Zames that led to an explosion of interest in this special problem. The H_∞ optimal control problem is originally studied in the framework of frequency domain theory and algebra theory[6] and is solved using the state space method.[5] Nevertheless, it is not easy to use the H_∞ method in industrial control systems. On one hand, the design procedure to ensure desired specifications still involves too much trial and error and requires expert understandings, and it is a difficult task to translate the quantitative design requirement such as overshoot into the H_∞ performance criterion and related weighting functions. On the other hand, it almost always provides high-order controllers, although one can reduce the orders by model reduction methods.

There are also many other multivariable design methods, of which the most well-known may be the Internal Model Control (IMC) and model predictive control (MPC).[7,8] The two internally related methods provide an alternative for linear systems and have been applied to some industrial control systems.

Despite all the advances and improvements, multivariable design problem still need further research. For example, the LQG, H_∞, and IMC procedures mainly involve delay-free systems. Many industrial control systems contain time delays. The control problem of multivariable systems with time delays is particularly difficult. Several papers have been devoted to this problem (see, for example, the work of Ogunnaike and Ray,[9] Wang,[10] Wang,[11] Gu et al.,[12] Niculescu and Gu,[13] and Liu et al.[14]).

The objective of this paper is to develop an algebraic design method for multivariable systems with multiple time delays. In the method, the algebra-based input−output design technique is adopted, because it is easy to understand and use, and it requires no state variables. Different from those developed methods (see, for example, the work of Morari and Zafiriou[7] and Kwakernaak[15]), this method focuses mainly on decoupled response. First, all stabilizing decoupling controllers are parametrized. Second, the optimal decoupling controller is derived in one step by minimizing the sensitivity function, instead of designing the decoupler first and then the controller. Finally, the performance and robustness are analyzed and a simple tuning rule is developed for quantitative performance and robustness. Compared with classical decoupling design methods, the most important feature of the proposed method is that the controller is optimal and analytical, and applicable to general plants. In addition, in the proposed method, the performance and robustness can be directly analyzed by those developed techniques.

The paper is organized as follows: In Section 2, the IMC design procedure is briefly reviewed and some important background is introduced. In Section 3, all stabilizing controllers for plants with time delay are parametrized. A heuristic procedure is proposed in Section 4 for exploring the required factorization and a new factorization is defined for plants with multiple time delays. The H_2 optimal controller is derived in Section 5. The filter design problem is discussed in Section 6.

* To whom correspondence should be addressed. Tel.: +86.21.34204019. Fax: +86.21.54260762. E-mail address:wdzhang@sjtu.edu.cn.

† Shanghai Jiaotong University.

‡ Lulea University of Technology.

10.1021/ie0602310 CCC: $33.50
Published on Web 09/19/2006

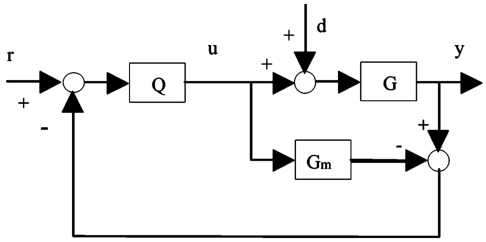

Figure 1. Schematic of the internal model control (IMC) control structure.

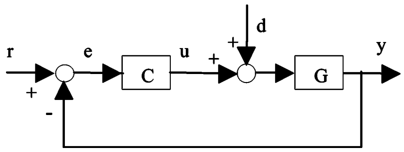

Figure 2. Schematic of the unity feedback control loop.

In Section 7, the performance and robustness are analyzed and a simple tuning rule is developed for quantitative performance and robustness. In Section 8, the proposed method is illustrated by several typical plants. Finally, some conclusions are given in Section 9.

2. Internal Model Control of Multi-input−Multi-output (MIMO) Processes

Consider the IMC structure shown in Figure 1, where $G(s)$ is a square plant of dimension p, $G_m(s)$ is the model, and $Q(s)$ is the IMC controller. The IMC structure can be related to the unity feedback loop through

$$C(s) = Q(s)[I - G(s)Q(s)]^{-1} \tag{1}$$

(see Figure 2).

Definition 2.1. A system G(s) is proper if all its elements are proper, and it is strictly proper if all its elements are strictly proper. Especially, G(s) is semi-proper if G(s) is proper but not strictly proper. All systems that are not proper are improper.

Definition 2.2. The pole polynomial $\pi(s)$ is the least common denominator of all nonidentically zero minors of all orders of G(s). The poles of G(s) are the roots of the characteristic equation $\pi(s) = 0$.

Definition 2.3. A point is a zero of G(s) if the rank of G(s) at this point is less than the normal rank of G(s).

Definition 2.4. G(s) is a non-minimum phase (NMP) if its transfer function contains zeros in the closed right half-plane or there exists a time delay term that can be factored out.

The task of how to factor a "time delay term" will be discussed in Section 5.

In the case that the model is exact, that is, $G(s) = G_m(s)$, the sensitivity transfer matrix (i.e., the transfer matrix from setpoint $r(s)$ to the error $e(s)$) is given by

$$S(s) = I - G(s)Q(s) \tag{2}$$

and the complementary transfer matrix (i.e., the transfer matrix from the setpoint $r(s)$ to the system output $y(s)$) is given by

$$T(s) = G(s)Q(s) \tag{3}$$

Let $W_1(s)$ and $W_2(s)$ be the performance weighting functions. The performance index of the MIMO IMC design is

$$\min \|W_2(s)S(s)W_1(s)\|_2 \tag{4}$$

which implies that the sum of the integral of square errors (ISEs) that each of the inputs would cause when applied to the system separately is minimized.

Assume that the plant is rational and can be factored into a stable all-pass portion $G_A(s)$ and a minimum phase (MP) portion $G_M(s)$, such that

$$G(s) = G_A(s)G_M(s) \tag{5}$$

Hence, $G_A(s)$ and $G_M^{-1}(s)$ are stable and $G_A^*(s)G_A(s) = I$. Here, the asterisk (*) denotes the complex conjugate transpose of a matrix.

The procedure for performing this inner−outer factorization was performed by the state space method and used for special plants. Morari and Zafiriou[7] have shown that, for a unit step input, the optimal controller is given by

$$Q_{opt}(s) = G_M^{-1}(s) \tag{6}$$

$Q_{opt}(s)$ is generally improper. This is acceptable because a diagonal filter $J(s)$ will be introduced to make it proper: $Q(s) = Q_{opt}(s)J(s)$.

The IMC design provides an excellent result, but only nondecoupling design is developed for plants without time delay. If there are time delays in the plant, they must be replaced by their rational approximations. In this paper, we will develop a decoupling design procedure for general linear systems with multiple time delays. The first step is to parametrize all stabilizing controllers. The second step is to define a factorization. The parametrization and the factorization will be used in the third step to derive the optimal solution. Finally, an efficient tuning method will be developed.

3. Parametrization of All Stabilizing Controllers

Throughout this paper, we consider square plants that are linear time-invariant and causal. For control system design, the following is always assumed:

(1) There is no closed right half-plane zero-pole cancellation in $G(s)$.

(2) $G(s)$ has no finite zeros on the imaginary axis.

(3) $G(s)$ is not identically singular; that is, $\det[G(s)]$ is not identically zero.

(4) $I - G(s)Q(s)$ is not identically singular if the plant is semi-proper.

When mentioning zero-pole cancellation in this paper, we mean that there are zeros and poles at the same point. If the first condition is not satisfied, the internal stability of the closed-loop system cannot be guaranteed. The second condition has to be made for the H_2 optimal decoupling control, because zeros on the imaginary axis will cause internal instability. The third condition is equivalent to that $G(s)$ is of full normal rank; that is, the plant is of full rank for every s in the set of complex numbers, except for a finite number of elements of complex numbers. (This condition is made so that the optimal control problem is solvable.) The fourth condition ensures that the transfer matrixes from any point of the control system to any other point exist and are proper. These assumptions are not very restrictive, because they will not be destroyed after a slight perturbation in the coefficients of the plant is introduced.

A fundamental requirement of any feedback control system design is the stability of the closed-loop system. Therefore, the design problem may be considered as a search or optimization over the class of all stabilizing controllers. To conduct the proposed design, we must develop a parametrization of all stabilizing controllers. Zhang et al.[16] proposed a new parametrization for nonsquare plants. The controller parametrization here is, in fact, a special case of that given by Zhang et al.[16] and can be regarded as a natural extension on the parametrization of SISO plants.

Consider the IMC structure shown in Figure 1 again. The structure cannot be used to control unstable plants. Nevertheless, even for unstable plants, we will exploit the features of the IMC structure for control system design and then implement the controller in the unity feedback control loop shown in Figure 2.

The closed-loop system is internally stable if and only if all elements of the following transfer matrix are stable:[7]

$$H(s) = \begin{bmatrix} G(s)Q(s) & [I - G(s)Q(s)]G(s) \\ Q(s) & -Q(s)G(s) \end{bmatrix} \tag{7}$$

This implies that the unity feedback control system is internally stable if and only if

(1) $Q(s)$ is stable.

(2) $I - G(s)Q(s)$ has zeros wherever $G(s)$ has unstable poles.

(3) There is no closed right half-plane zero-pole cancellation in $[I - G(s)Q(s)]G(s)$ and $C(s)$.

When the system performance is considered, it is always desirable that the system has a zero steady-state error. Assume that $G(s)$ has r_p unstable poles and the unstable pole p_j ($j = 1, 2, ..., r_p$), $\mathrm{Re}(p_j) \geq 0$, and is a multiple of l_j, and l_{ij} is the largest multiple of p_j in any element of the ith row of $G(s)$; z_j ($j = 1, 2, ..., r_z$), $\mathrm{Re}(z_j) > 0$, is an open right half-plane zero of $G(s)$ and k_{ij} is the maximum multiple of the unstable pole in the ith column of $G^{-1}(s)$. Let $L(s) = G(s)C(s)$ and m be the largest integer for which

$$\mathrm{rank}[\lim_{s\to 0} s^m L(s)] = p \tag{8}$$

The system $L(s)$ then is said to be of Type m. Note that $L(s)$ has at least $p \times m$ poles at the origin. For a Type m system, the sensitivity transfer function satisfies

$$\lim_{s\to 0} s^{-k} S(s) = 0 \qquad (\text{for } 0 \leq k < m) \tag{9}$$

If the closed-loop system is stable, as $t \to \infty$, the system can perfectly track inputs of the form $\sum_{k=0}^{m} a_k s^{-k}$, where a_k are real constant vectors. In particular, a Type 1 system requires that

$$\lim_{s\to 0} G(s)Q(s) = I \tag{10}$$

Theorem 3.1. Assume that G(s) is a plant with time delays. All controllers that make the unity feedback control system internally stable and have a zero steady-state error for a step reference can be parametrized as

$$C(s) = Q(s)[I - G(s)Q(s)]^{-1}$$

where

$$Q(s) = G^{-1}(0)[I + sQ_1(s)]$$

$Q_1(s)$ is any stable transfer matrix that makes Q(s) proper and satisfies

$$\lim_{s\to p_j} \frac{d^k}{ds^k} \det[I - G(s)G^{-1}(0) - sG(s)G^{-1}(0)Q_1(s)] = 0 \qquad (\text{for } j = 1, 2, ..., r_p, 0 \leq k < l_j)$$

and there is no closed right half-plane zero-pole cancellation in [I − G(s)Q(s)]G(s) and C(s).

Proof: If

$$\lim_{s\to 0} [I - G(s)Q(s)] = 0$$

The closed-loop system possesses the asymptotic property. Then,

$$Q(0) = G^{-1}(0)$$

All transfer matrixes that satisfy the condition can be written as

$$Q(s) = G^{-1}(0)[I + sQ_1(s)]$$

To guarantee the internal stability of the closed-loop system, first, Q(s) should be stable. This implies that Q(s) should be proper.

Second, [I − G(s)Q(s)]G(s) should be stable. Thus, I − G(s)Q(s) must cancel all closed right half-plane poles of G(s). To achieve this, Q(s) must satisfy the condition

$$\lim_{s\to p_j} \frac{d^k}{ds^k} \det[I - G(s)G^{-1}(0) - sG(s)G^{-1}(0)Q_1(s)] = 0 \qquad (\text{for } j = 1, 2, ..., r_p, 0 \leq k < l_j)$$

The condition cannot guarantee the stability of [I − G(s)Q(s)]G(s) unless there is no closed right half-plane zero-pole cancellation in [I − G(s)Q(s)]G(s) and C(s).

Write the closed-loop transfer matrix as

$$T(s) = G(s)Q(s) = G(s)G^{-1}(0)[I + sQ_1(s)] \tag{11}$$

In practice, it is often desirable that the closed-loop response is decoupled; that is, the closed-loop transfer matrix is diagonal.

Corollary 3.2. Assume that the closed-loop response is decoupled. Let the ith element of T(s) be $T_i(s)$. All controllers that make the unity feedback control system internally stable and have a zero steady-state error for a step reference can be parametrized as

$$C(s) = Q(s)[I - G(s)Q(s)]^{-1}$$

where

$$Q(s) = G^{-1}(0)[I + sQ_1(s)]$$

$Q_1(s)$ is any stable transfer matrix that makes Q(s) proper and T(s) diagonal, and satisfies

$$\lim_{s\to p_j} \frac{d^k}{ds^k} [1 - T_i(s)] = 0 \qquad (\text{for } j = 1, 2, ..., r_p; 0 \leq k < l_{ij})$$

and there is no closed right half-plane zero-pole cancellation in [I − G(s)Q(s)]G(s) and C(s).

Corollary 3.2 implies that more right half-plane zeros in $I - G(s)Q(s)$ may have to be introduced to guarantee the internal stability. A question of interest is when we have to do this. We must do this if the multiples of the closed right half-plane poles of at least one row of $G(s)$ are not the same.

4. Factorization of Plants

In this section, a new factorization will be defined. To motivate the procedure to follow, let us see how to choose a reasonable factorization. Assume that the plant is written in the form

$$G(s) = \begin{bmatrix} G_{11}(s)e^{-\theta_{11}s} & \cdots & G_{1p}(s)e^{-\theta_{1p}s} \\ \vdots & \ddots & \vdots \\ G_{p1}(s)e^{-\theta_{n1}s} & \cdots & G_{pp}(s)e^{-\theta_{pp}s} \end{bmatrix} \quad (12)$$

where $G_{ij}(s)$ are scalar rational transfer functions and $\theta_{ij} > 0$ are time delays. Recall the multiloop controller design, only those diagonal elements of the plant are considered and nondiagonal elements are regarded as uncertainties. Thus, it seems that the plant can be factored into the form

$$G(s) = G_D(s)G_O(s) \quad (13)$$

where

$$G_D(s) = \mathrm{diag}\{e^{-\theta_1 s}, \cdots, e^{-\theta_p s}\}$$

$$G_O(s) = \begin{bmatrix} G_{11}(s) & \cdots & G_{1p}(s)e^{-(\theta_{1p}-\theta_{11})s} \\ \vdots & \ddots & \vdots \\ G_{p1}(s)e^{-(\theta_{p1}-\theta_{pp})s} & \cdots & G_{pp}(s) \end{bmatrix}$$

Evidently, $G_D(s)$ is all-pass and $G_D(0) = I$. Unfortunately, such a factorization may not be feasible, prediction may exist in the controller.

To explore how to choose the factorization, let us review the SISO design first. Given a plant, the ideal solution is

$$Q_{opt}(s) = G^{-1}(s) \quad (14)$$

However, the controller is not realizable, because of the prediction. Therefore, we factored the plant $G(s)$ into two parts: $G(s) = G_D(s)G_O(s)$, where $G_D(s)$ is the time delay and $G_O(s)$ the delay-free component. If $G_O(s)$ is MP, the controller is taken to be

$$Q_{opt}(s) = G_O^{-1}(s)$$

It can be proven that such a design is H_2 optimal.[7,18] We can give an alternative explanation for the design. As the time delay is unavoided, we must remove the prediction. Factor the plant $G(s)$ into two parts: $G(s) = G_D(s)G_O(s)$, where $G_D(s)$ should be chosen such that it just counteracts the prediction in the controller $Q_{opt}(s) = G^{-1}(s)$. The controller is then taken to be

$$Q_{opt}(s) = G^{-1}(s)G_D(s) = G_O^{-1}(s)$$

In other words, the inverse that is closest to $G^{-1}(s)$ is adopted.

We now return to the MIMO case. For the MIMO case, it is desirable to take

$$Q_{opt}(s) = G^{-1}(s) \quad (15)$$

Let $G^{ji}(s)e^{-\theta^{ji}s}$ be the element of $G^{-1}(s)$; then,

$$G^{-1}(s) = \begin{bmatrix} G^{11}(s)e^{-\theta^{11}s} & \cdots & G^{p1}(s)e^{-\theta^{p1}s} \\ \vdots & \ddots & \vdots \\ G^{1p}(s)e^{-\theta^{1p}s} & \cdots & G^{pp}(s)e^{-\theta^{pp}s} \end{bmatrix}$$

Some elements of $G^{-1}(s)$ may contain predictions that are not physically realizable (that is, θ^{ji} is negative). To avoid this, we have to remove these predictions. This can be reached by post-multiplying it by $G_D(s)$ and then the controller is

$$Q_{opt}(s) = G^{-1}(s)G_D(s) = G_O^{-1}(s)$$

$G_D(s)$ should be chosen such that it just counteract those predictions.

Let θ_i be the maximum prediction of the ith column of $G^{-1}(s)$; that is,

$$\theta_i = \max_j \theta^{ji}$$

Definition 4.1. The factorization for the time delay is

$$G_D(s) = \mathrm{diag}\{e^{-\theta_1 s}, \ldots, e^{-\theta_p s}\} \quad (16)$$

The factorization for time delays is not conservative. On one hand, any time delays smaller than that in Definition 4.1 will result in an internally unstable closed-loop system. On the other hand, it will be proven in Section 5 that the aforementioned conjecture on the optimal solution is correct and the factorization developed in this section will result in the unique optimal solution for decoupling responses. It is easy to verify that $G_D^*(s)G_D(s) = I$ and $G_D(0) = I$.

Because $G_D(s)$ is diagonal and elements of $G_D(s)$ are time delays, no right half-plane zeros and poles are canceled when forming $G_O(s) = G_D^{-1}(s)G(s)$. If $G_O(s)$ is MP, we can directly take $Q_{opt}(s) = G_O^{-1}(s)$. In cases where the plant has right half-plane zeros except for the time delays, we must factor it further.

Let

$$Q_{opt}(s) = G_O^{-1}(s) = \frac{\mathrm{adj}\{G_O(s)\}}{\det\{G_O(s)\}} \quad (17)$$

Some elements of $G_O^{-1}(s)$ may be unstable. To make it stable, we must remove the unstable poles in each element. Similarly, this can be reached by post-multiplying it by $G_N(s)$:

$$Q_{opt}(s) = G_O^{-1}(s)G_N(s) = G_M^{-1}(s) \quad (18)$$

As z_j ($j = 1, 2, ..., r_z$) is a right half-plane zero of $G(s)$, it is an unstable pole of $G_O^{-1}(s)$. $G_N(s)$ can be constructed as follows.

Definition 4.2. The factorization for closed right half-plane zeros is

$$G_N(s) = \mathrm{diag}\left\{\prod_{j=1}^{r_z}\left(\frac{-s+z_j}{s+z_j^*}\right)^{k_{1j}}, \ldots, \prod_{j=1}^{r_z}\left(\frac{-s+z_j}{s+z_j^*}\right)^{k_{pj}}\right\} \quad (19)$$

The factorization for right half-plane zeros is also not conservative. It is easy to verify that $G_N^*(s)G_N(s) = I$ and $G_N(0) = I$. The plant may have infinite right half-plane zeros. Such a case is illustrated in Example 7.1.

Now we can let

$$G_A(s) = G_D(s)G_N(s) \quad (20)$$

Especially, for rational plants $G_D(s) = I$,

$$G_A(s) = G_N(s) \quad (21)$$

An important property of an all-pass transfer matrix is that it will not affect the value of 2-norm, that is, $\|G_A(s) \times \text{some transfer function}\|_2 = \|\text{some transfer function}\|_2$. This result can be directly obtained by the definition of 2-norm. In Section 5, the property will be used to derive the optimal solution.

The factorization that has been defined in this section is unique. For stable plants, it will result in findings similar to those in Wang.[11] Compared to the inner−outer factorization, the feature of the proposed factorization is that it is analytical and can give decoupled responses. For details, see Zhang and Lin.[19]

From the definition of the factorization, it is determined that more time delays will be introduced if the predictions of at least one column of $G^{-1}(s)$ are not the same, and more right half-plane zeros will be introduced if multiples of any unstable poles in at least one column of $G_O{}^{-1}(s)$ are not the same. This is the price that is paid for decoupling.

It is also observed that some time delays or right half-plane zeros may influence other channels. In some current literature, the problem is analyzed by zero directions. The zero directions are not intuitive; therefore, we will analyze the problem by the time delay or the zero itself.

Definition 4.3. The time delay is canonical if at least one element of $G_D(s)$ contains a time delay, provided that the greatest common time delay of all elements of $G_D(s)$ has been removed.

Definition 4.4. A right half-plane zero is canonical if at least one element of $G_N(s)$ has the zero, provided that the greatest common factor of all elements of $G_N(s)$ has been removed.

A canonical time delay or right half-plane zero will not spread its influence over all channels, whereas a noncanonical time delay or right half-plane zero will affect all channels to the same extent.

Example 4.1. Consider the plant with the following transfer matrix:

$$G(s) = \frac{1}{(s+3)(s-1)}\begin{bmatrix} s-2 & 2(s-2) \\ 1 & s-1 \end{bmatrix}$$

The plant has a nude zero, at $s = 2$, and a nonnude zero, at $s = 3$. Because

$$G_N(s) = \begin{bmatrix} \dfrac{-(s-2)}{s+2} & 0 \\ 0 & 1 \end{bmatrix} \frac{-(s-3)}{s+3}$$

the zero *$s = 2$ is canonical and the zero $s = 3$ is noncanonical.*

5. Optimal Controller Design for Processes with Time Delays

The idea of the H_2 optimal control is to find a controller that stabilizes the system and minimizes a given quadratic cost function. We now derive the optimal controller for plants with time delays. Our derivation is based on input−output techniques. The design procedure is different from but internally related to those developed (see, for example, the works of Morari and Zafiriou,[7] Zhou et al.,[5] and Skogestad and Postlethwaite[17]).

The H_2 optimal control problem is defined by

$$\min \|W_2(s)S(s)W_1(s)\|_2 \tag{22}$$

where $W_1(s)$ and $W_2(s)$ are performance weighting functions. To design the controller, the two weighting functions must be determined first.

Let us determine what the weighting functions imply. We excite the system in separate experiments with p different linearly independent inputs $r_i(s)$ ($i = 1, 2, ..., p$). For one experiment, the error is $e_i(s) = S(s)r_i(s)$. We define $W_1(s) = [r_1(s), r_2(s), ..., r_p(s)]$. The columns of $S(s)W_1(s)$ are the errors from the p experiments. Thus, $W_1(s)$ is the input weighting function. In the H_2 controller design, the weighting function is usually taken to be the input so that the input is normalized into impulses. In practice, step inputs or inputs similar to steps (that is, the input that has only one pole at the origin) are of primary importance, then $W_1(s) = s^{-1}I$ is a reasonable weight. In this paper, we consider only step inputs. For other inputs, one can choose the weighting function similarly. For example, for ramp inputs or inputs similar to ramps, one can take $W_1(s) = s^{-2}I$.

Sometimes, the design specification is expressed in terms of the shape of $S(s)$. This implies that the input has the spectrum $S^{-1}(s)$. If $S^{-1}(s)$ has only one pole at the origin, we can also take $W_1(s) = s^{-1}I$, while the shape of $S(s)$ can be obtained by choosing an appropriate filter. In the case where the input is stochastic but with a known spectrum, the treatment is similar.

Consider now premultiplication by the output weight $W_2(s)$ to generate $W_2(s)S(s)W_1(s)$, a matrix whose columns are the weighted errors. The output weight is used because some errors may have to be made small over different frequency ranges. In our treatment, a filter can be used to achieve this. Compared to using a complex weight, the process of penalizing errors through the use of a filter is simple and easy to understand. We then can simply take $W_2(s) = I$.

Therefore, the H_2 optimal control problem can be rewritten as

$$\min \|S(s)W_1(s)\|_2 \tag{23}$$

The following procedure is general enough to allow general weighting functions. However, it is only performed for pre-specified weighting functions, to avoid meaningless complexity.

Introduce the symbol L_2 for the family of all strictly proper functions with no poles on the imaginary axis. Let H_2 denote the subset of L_2, $H_2^\perp$ the set of strictly proper transfer functions analytical in Re $s \le 0$, and $H_2 + H_2^\perp$ the set of all sums. Thus, $H_2 + H_2^\perp$ consists precisely of all strictly proper transfer functions with no poles on the imaginary axis. Furthermore, every function $F(s)$ in L_2 can be uniquely expressed as

$$F = F_1 + F_2 \qquad (F_1 \in H_2, F_2 \in H_2^\perp) \tag{24}$$

Lemma 5.1. If $F_1 \in H_2$ and $F_2 \in H_2^\perp$, then

$$\|F_1 + F_2\|_2^2 = \|F_1\|_2^2 + \|F_2\|_2^2$$

Proof: Omitted.

The control system design problem can be considered as a search or an optimization over the class of all stabilizing controllers with asymptotic properties. For H_2 optimal performance criterion, we have

$$\|S(s)W_1(s)\|_2^2 = \|s^{-1}[I - G(s)Q(s)]\|_2^2 = \|s^{-1}[I - G(s)G^{-1}(0) - sG(s)G^{-1}(0)Q_1(s)]\|_2^2 \tag{25}$$

Theorem 5.2. Assume that the plant with time delay can be factored into two portions,

$$G(s) = G_A(s)G_M(s)$$

such that

(1) *$G_A(s)$ and $G_M{}^{-1}(s)$ are stable.*

(2) *$G_A(s)$ is diagonal, $G_A^*(s)G_A(s) = I$, and $G_A(0) = I$.*

(3) *$G_M(s)$ has $\sum_{j=1}^{r_p} l_j + \sum_{i=1}^{p}\sum_{j=1}^{r_z} k_{ij}$ closed right half-plane poles, including those removed poles in forming $G_M(s)$ by $G_A{}^{-1}(s)G(s)$.*

(4) There are no zero-pole cancellations in G_A*(s)* and G_M*(s).*

The unique optimal solution for decoupled response then is

$$Q_{\mathrm{opt}}(s) = G_{\mathrm{M}}^{-1}(s)$$

Proof: The H_2 optimal performance criterion can be written as

$$\|S(s)W_1(s)\|_2^2 = \|s^{-1}[G_{\mathrm{A}}^{-1}(s) - I] + s^{-1}\{ I - G_{\mathrm{M}}(s)G^{-1}(0)[I + sQ_1(s)]\}\|_2^2$$

Because $G_A(0) = I$, $G_M(0)G^{-1}(0) = I$, s must be a factor of $G_A^{-1}(s) - I$ and $I - G_M(s)G^{-1}(0)$, and the expression $s^{-1}[G_A^{-1}(s) - I]$ is strictly proper, $s^{-1}\{I - G_M(s)G^{-1}(0)[I + sQ_1(s)]\}$ is also strictly proper if $Q(s)$ is proper. The assumption the closed right half-plane poles of $G_M(s)$ are $\prod_{j=1}^{r_p} (-p_j^{-1}s + 1)^{m_{ij}}$ implies that $G_A(s)$ has only right half-plane zeros or time delay, and, thus, $s^{-1}[G_A^{-1}(s) - I]$ is unstable. According to the parametrization, $s^{-1}\{I - G_M(s)G^{-1}(0)[I + sQ_1(s)]\}$ is stable.

By Lemma 5.1, we have

$$\|S(s)W_1(s)\|_2^2 = \|s^{-1}[G_A^{-1}(s) - I]\|_2^2 + \|s^{-1}\{I - G_{\mathrm{M}}(s)G^{-1}(0)[I + sQ_1(s)]\}\|_2^2$$

Minimizing the right-hand side of the equation, one gets

$$Q_{1\mathrm{opt}}(s) = s^{-1}G(0)G_{\mathrm{M}}^{-1}(s)[I - G_{\mathrm{M}}(s)G^{-1}(0)]$$

The unique diagonal optimal solution is

$$Q_{\mathrm{opt}}(s) = G_{\mathrm{M}}^{-1}(s)$$

The result is thoroughly different from the IMC design in Morari and Zafiriou,[7] because of its property of decoupling. Sometimes, the decoupled response can also achieve the global optimal performance, rather than diagonal optimal performance. If the time delays in each row of $G(s)$ are the same, and all closed right half-plane zeros have the same multiples in each row of $G(s)$, the diagonal optimal solution achieves the global optimal solution.

The proposed design procedure is also applicable to MIMO processes with state delays. The state delay is seldom considered in input−output methods, because the plant with state delays usually contains an infinite number of unstable poles. A plant with output and state time delays can be described by state equations of the form

$$\dot{x}(t) = A_0x(t) + \sum_{i=1}^{n_i}A_ix(t - \vartheta_i) + B_0u(t) + \sum_{j=1}^{n_j}B_ju(t - \tau_j) \tag{26a}$$

$$y(t) = C_0x(t) + \sum_{r=1}^{n_r}C_rx(t - \varsigma_r) \tag{26b}$$

where ϑ_i, τ_j, and ς_r are lumped delays. The corresponding transfer matrix can then be written as

$$G(s) = (C_0 + \sum_{r=1}^{n_r}C_r\mathrm{e}^{-\varsigma_rs})(sI - A_0 - \sum_{i=1}^{n_i}A_i\mathrm{e}^{-\vartheta_is})^{-1} \times (B_0 + \sum_{j=1}^{n_j}B_j\mathrm{e}^{-\tau_js}) \tag{27}$$

We can rewrite it into the form of

$$G(s) = \begin{bmatrix} G_{11}(s)\mathrm{e}^{-\theta_{11}s} & \cdots & G_{1p}(s)\mathrm{e}^{-\theta_{1n}s} \\ \vdots & \ddots & \vdots \\ G_{p1}(s)\mathrm{e}^{-\theta_{n1}s} & \cdots & G_{pp}(s)\mathrm{e}^{-\theta_{pp}s} \end{bmatrix}$$

The numerators and denominators of $G_{ij}(s)$ are quasi-polynomials. If the plant has an infinite number of unstable poles, to use the proposed approach, these quasi-polynomials should be expanded into polynomials by model reduction techniques. This implies that a rigorously analytical solution cannot be obtained for general MIMO processes with state delays. For some special cases, however, an analytical solution is possible.

6. Filter Design

The function of the filter is 2-fold:

(1) The optimal controller $Q_{\mathrm{opt}}(s) = G_{\mathrm{M}}^{-1}(s)$ is usually improper. A filter $J(s)$ must be introduced to make it proper: $Q(s) = Q_{\mathrm{opt}}(s)J(s)$.

(2) The filter will be used to tune the shape of the closed-loop response and satisfy the performance and robustness specifications.

The filter should satisfy the following requirements:

(1) The controller $Q(s) = Q_{\mathrm{opt}}(s)J(s)$ should be proper.

(2) The closed-loop system should be internally stable.

(3) Asymptotic tracking should be met.

It is not difficult to satisfy the first condition. The third condition can be satisfied by choosing

$$J(0) = I \tag{28}$$

If there are no constraints on $G_{\mathrm{A}}(s)$, the second condition is equivalent to

$$\lim_{s\to p_j} \frac{d^k}{ds^k}[I - G(s)Q_{\mathrm{opt}}(s)J(s)] = 0 \quad (\text{for } j = 1, 2, ..., r_p;\ 0 \le k < l_j) \tag{29}$$

and there is no right half-plane zero-pole cancellation in $[I - G(s)Q(s)]G(s)$ and $C(s)$.

Now let us determine how to choose the filter for the proposed method. In principle, the structure of $J(s)$ can be as complex as the designer wishes. However, for decoupling response, a simple structure such as a diagonal matrix is enough:

$$J(s) = \mathrm{diag}\{J_1(s), ..., J_p(s)\} \tag{30}$$

Similar to that observed in the SISO case, the following form for elements of the filter is reasonable:

$$J_i(s) = \frac{N_{xi}(s)}{(\lambda_is + 1)^{n_i}} \tag{31}$$

where $N_{xi}(s)$ is a polynomial with all roots in the left half-plane, $N_{xi}(0) = 1$, and λ_i is a positive real constant (called the performance degree).[18,20] As clearly observed, the nominal stability can always be guaranteed by positive performance degrees.

Assume that the largest relative degree in any element of the ith column of $Q_{\mathrm{opt}}(s)$ is α_i. The first condition can be achieved by choosing $n_i = \deg\{N_{xi}(s)\} - \alpha_i$ for strictly proper column (that is, at least one element of the column is strictly proper) and $n_i = \deg\{N_{xi}(s)\} + 1$ for semi-proper column (that is, every element of the column is semi-proper).

The third condition can be satisfied by choosing

$$J(0) = I \tag{32}$$

Seemingly, the second condition can be satisfied by choosing a filter $J(s)$ such that

$$\lim_{s \to p_j} \frac{d^k}{ds^k} [I - G(s)Q_{\mathrm{opt}}(s)J(s)] = 0 \qquad (\text{for } j = 1, 2, ..., r_p;\ 0 \le k < l_j) \tag{33}$$

and there is no right half-plane zero-pole cancellation in $[I - G(s)Q(s)]G(s)$ and $C(s)$. However, as we discussed in Section 3, the internal stability may not be guaranteed, because $G_A(s)$ is resticted to be diagonal. Let the ith element of $G_A(s)$ is $G_{Ai}(s)$. The filter that satisfies the following condition then can guarantee the internal stability:

$$\lim_{s \to p_j} \frac{d^k}{ds^k} [I - G_{Ai}(s)J_i(s)] = 0 \qquad (\text{for } j = 1, 2, ..., r_p;\ 0 \le k < l_{ij}) \tag{34}$$

Here, $\deg\{N_{xi}(s)\} = \sum_{j=1}^{r_p} l_{ij}$. If the multiples of closed right half-plane poles in each row of $G(s)$ are the same, the condition reduces to

$$\lim_{s \to p_j} \frac{d^k}{ds^k} [I - G(s)Q_{\mathrm{opt}}(s)J(s)] = 0 \qquad (\text{for } j = 1, 2, ..., r_p;\ 0 \le k < l_j)$$

7. Analysis and Discussion

In the preceding sections, we proposed an optimal analytical design method for the design of MIMO control systems with multiple time delays. Strictly speaking, the method is not analytical for all plants; we may have to use rational approximations for some complex MIMO plants.

Assume that $G(s)$ is a plant with only output time delays. According to the method developed in Section 5, the plant can be factored into

$$G(s) = G_D(s)G_O(s)$$

If $G_O(s)$ still contains time delays, the resulted optimal controller is of infinite dimension. Although the design may also work, we must use a complex feedback structure to implement the rigorously analytical controller. To implement it using a simple rational transfer function, it is recommended to expand the time delay first and then design an appropriate controller. This is illustrated by the following example.

Example 7.1. Consider the plant described by the following transfer matrix:

$$G(s) = \frac{1}{(s+2)(s-3)} \begin{bmatrix} (s-2)e^{-4s} & (s-2)e^{-s} \\ e^{-2}e^{-3s} & e^{-2s} \end{bmatrix}$$

The plant contains both closed right half-plane poles and time delays. With the proposed factorization, we get

$$G(s) = \begin{bmatrix} e^{-2s} & 0 \\ 0 & e^{-3s} \end{bmatrix} \frac{1}{(s+2)(s-3)} \begin{bmatrix} (s-2)e^{-2s} & (s-2)e^{s} \\ e^{-2} & e^{s} \end{bmatrix}$$

and

$$G_O(s) = \frac{1}{(s+2)(s-3)} \begin{bmatrix} (s-2)e^{-2s} & (s-2)e^{s} \\ e^{-2} & e^{s} \end{bmatrix}$$

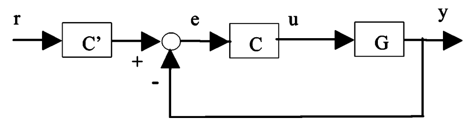

Figure 3. Schematic of the two-degrees-of-freedom control system.

It is observed that $G_O(s)$ still contains time delays. If we do not use rational approximation, the following factorization will be obtained:

$$G_O(s) = \frac{e^{-2s} - e^{-2}}{e^{-2s} - e^{2}} \begin{bmatrix} \left(\frac{-s+2}{s+2}\right) & 0 \\ 0 & 1 \end{bmatrix} \times \frac{e^{-2s} - e^{2}}{(s+2)(s-3)(e^{-2s} - e^{-2})} \begin{bmatrix} -(s+2)e^{-2s} & -(s+2)e^{s} \\ e^{-2} & e^{s} \end{bmatrix}$$

and

$$G_M(s) = \frac{e^{-2s} - e^{2}}{(s+2)(s-3)(e^{-2s} - e^{-2})} \times \begin{bmatrix} -(s+2)e^{-2s} & -(s+2)e^{s} \\ e^{-2} & e^{s} \end{bmatrix}$$

It follows that

$$Q_{\mathrm{opt}}(s) = \frac{\begin{bmatrix} 1 & (s+2) \\ -e^{-2}e^{-s} & -(s+2)e^{-3s} \end{bmatrix}(s-3)}{-(e^{-2s} - e^{2})/(s+2)}$$

The controller is of infinite dimension. Because the plant is unstable, the controller must be implemented in the unity feedback loop. The design procedure is as follows: First, a filter is introduced. Second, the controller is derived, through

$$C(s) = Q(s)[I - G(s)Q(s)]^{-1} \tag{35}$$

Finally, the right half-plane zero-pole cancellations in *$C(s)$ are removed by rational approximations. The procedure is very tedious and the resulting $C(s)$ will be of high order. If we approximated $G_o(s)$ or $Q_{opt}(s)$ by a rational transfer matrix, the design effort could be significantly reduced*.

As we known, a control system with unstable plant usually exhibits excessive overshoot. This problem can be solved well by a two-degrees-of-freedom structure, as shown in Figure 3. On the other hand, the disturbance rejection is always a primary objective of control system design. The two-degrees-of-freedom controller can isolate the disturbance from the setpoint and, thus, make better control possible. It is very easy to design the two-degrees-of-freedom controller in the proposed framework.

The controller for the disturbance loop is just the controller of the unity feedback loop:

$$C(s) = Q(s)[I - G(s)Q(s)]^{-1} \tag{36}$$

If the plant is stable, we can directly implement the controller in the IMC structure. Otherwise, we must implement it in the unity feedback loop. The closed-loop transfer matrix of the unity feedback loop is

$$T(s) = G(s)Q(s) = G_D(s)G_N(s)J(s) \tag{37}$$

By regarding $T(s)$ as a plant and repeating the proposed design, one can obtain the following optimal controller for the setpoint loop:

$$C'_{\mathrm{opt}}(s) = J^{-1}(s){G_{\mathrm{N}}}^{-1}(s)$$

The introduction of a diagonal filter $J'(s)$ for the optimal controller yields

$$C'(s) = J^{-1}(s){G_{\mathrm{N}}}^{-1}(s)J'(s) \tag{38}$$

$J'(s)$ has a structure that is similar to that of $J(s)$. Obviously, the controller is diagonal.

Each element of the filter $J(s)$ (the case is similar for $J'(s)$) has adjustable performance degrees, λ_i. The performance degrees are determined by the design specifications. Practical design specifications are usually given in terms of time-domain responses (such as overshoot and rise time) or frequency domain responses (such as stability margin or bandwidth). Because the closed-loop response is decoupled, each channel can be tuned independently. The ith element of the closed-loop transfer matrix is

$$T_i(s) = \frac{N_{xi}(s)}{(\lambda_i s + 1)^{n_i}} \prod_{j=1}^{r_z} \left(\frac{-s + z_j}{s + z_j^*}\right)^{k_{ij}} \mathrm{e}^{-\theta_i s} \tag{39}$$

Thus, the performance degrees are related to the nominal response monotonically. This implies that the tuning procedure can be significantly simplified.[20]

Fortunately, most practical processes are stable. For stable processes, we can implement the proposed controller by the IMC structure. The most important merit of doing this is that the decoupling property can be thoroughly preserved, although rational approximations may be used for $Q(s)$. For unstable processes, we must implement the proposed design in the unity feedback control structure.

No mathematical model can exactly model a physical system. In practice, uncertainties always exist. Often, the requirement on robustness has been internally included in design specifications. In cases where the requirement is independently proposed, the robustness of the closed-loop system can be calculated according to some rigorous criteria.[7]

Many methods can be taken to describe the uncertainty. Assume that the uncertainty is expressed in terms of unstructured uncertainty. Let $\tilde{G}(s)$ be the uncertain plant. We then have the additive uncertainty

$$\tilde{G}(s) = G(s) + W_{\mathrm{u1}}(s)\Delta(s)W_{\mathrm{u2}}(s) \tag{40}$$

where $\bar{\sigma}[\Delta(j\omega)] < 1$, and $W_{\mathrm{u1}}(s)$ and $W_{\mathrm{u2}}(s)$ are stable transfer matrixes that characterize the spatial and frequency structure of the uncertainty, confining the matrix $\tilde{G}(s)$ to a neighborhood of the nominal model $G(s)$. An alternative statement to the additive unstructured uncertainty is the so-called multiplicative output form:

$$\tilde{G}(s) = [I + W_{\mathrm{u_1}}(s)\Delta(s)W_{\mathrm{u_2}}(s)]G(s) \tag{41}$$

This statement confines the matrix $\tilde{G}(s)$ to a normalized neighborhood of the nominal model $G(s)$. For unstructured uncertainties $W_{\mathrm{u_1}}(s)$ is a scalar function and $W_{\mathrm{u_2}}(s) = I$.

With regard to the aforementioned uncertainty description, the closed-loop system is robust stable for all stable $\Delta(s)$ if and only if, for every frequency,[7]

$$\bar{\sigma}[T(j\omega)W_{\mathrm{u_1}}(j\omega)] < 1$$

and the robust performance is guaranteed if

$$\bar{\sigma}[W_1(j\omega)S(j\omega)] + \bar{\sigma}[W_{\mathrm{u1}}(j\omega)T(j\omega)] < 1$$

Although the robustness can be tested theoretically by rigorous criteria, they are seldom adopted in practice, because of their complexity. A simple engineering tuning procedure developed by Zhang's group[18,20] is recommended here. The tuning procedure is quantitative and very simple: Increase the performance degrees monotonically until the required response is obtained. The method can be used for both nominal performance and robustness. It is observed that the proposed design method provides a possible transition between the classical design specification, such as overshoot, and the new design theory, based on a widely accepted optimal performance criterion.

The performance degrees can also be used to tune the amplitude of the control variables (as well as the shape of $S(s)$, stability margin, etc.). The tuning procedure is the same as that for tuning performance and robustness. In this way, controllers with unrealistic gains and bandwidths can be avoided and anti-windup can be easily achieved. Compared to using weighted control variables in the performance criteria, the design is simplified.

As we know, to obtain a solution and limit the amplitude of the control variables, one must add weighted control variables to the performance criteria in classical optimal design and, thus, the performance criterion is in the form of

$$\min \int_0^\infty (e^{\mathrm{T}}Qe + u^{\mathrm{T}}Ru)\,\mathrm{d}t \tag{42}$$

On one hand, this makes the design very complex; the designer cannot intuitively grasp what the intrinsic performance capabilities and limitations of the plant are. On the other hand, this introduces a fictitious factor in a rigorous method. Because of the fact that no rigorous rules can be used to choose the weighting functions, even if the plant model is exactly known, one hundred designers will obtain one hundred different results.

Also note that, unlike many other developed methods, the proposed design procedure is not dependent on a predescribed uncertainty profile. When the uncertainty profile is known, we can tune the closed-loop response for the worst case to obtain a tight tradeoff between the performance and the robustness. If the uncertainty profile is not known, we can also perform the design and approximately tune the closed-loop response.

8. Illustrative Examples

Example 8.1. In this example, we must construct a plant, because we failed to find a proper practical plant to illustrate the proposed design. The plant is described by the following transfer matrix:

$$G(s) = \frac{1}{(s+3)(s-1)}\begin{bmatrix} s-2 & 2(s-2) \\ 1 & s-1 \end{bmatrix}$$

which is an NMP and unstable. The plant has two two multiple poles at $s = -3$ and $s = 1$, one zero at $s = 2$, one zero at $s = 3$, and two zeros at infinity.

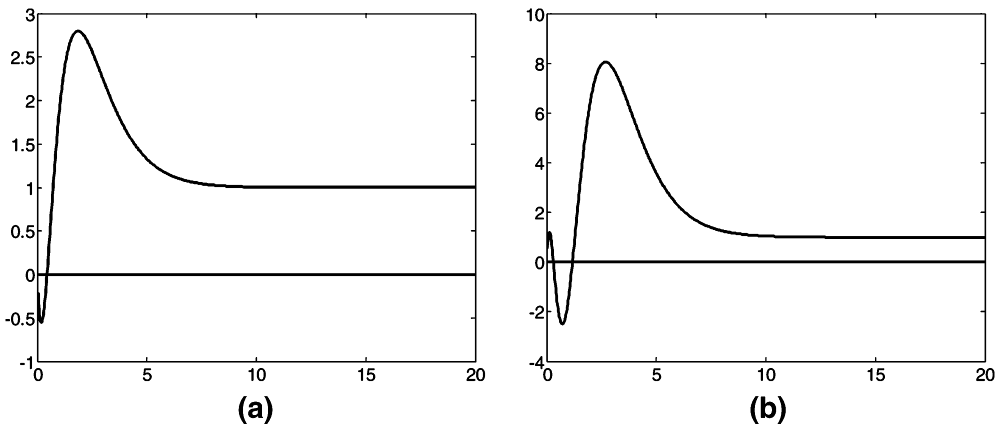

Figure 4. Graphical depiction of the closed-loop responses of the one-degree-of-freedom system: (a) $r_1 = 1/s$ and $r_2 = 0$; (b) $r_1 = 0$ and $r_2 = 1/s$.

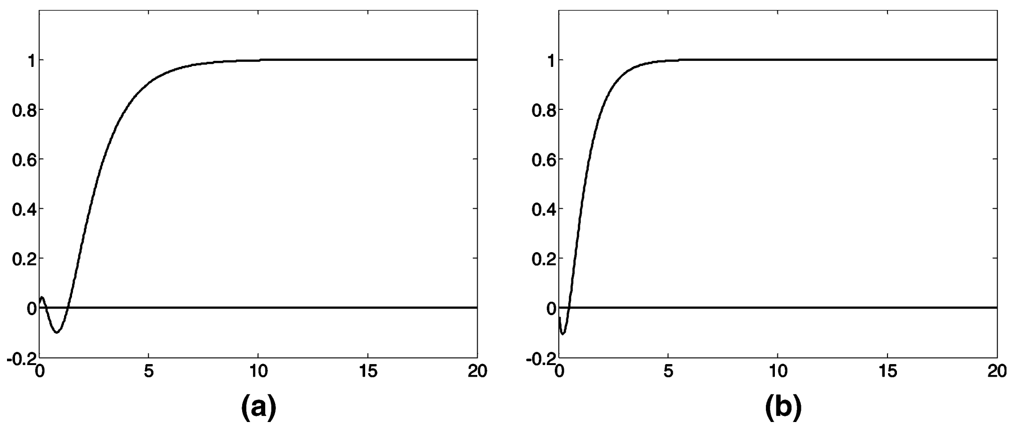

Figure 5. Graphical depiction of the closed-loop responses of the two-degrees-of-freedom system: (a) $r_1 = 1/s$ and $r_2 = 0$; (b) $r_1 = 0$ and $r_2 = 1/s$.

(1) $G(s)$ is factored into

$$G(s) = G_D(s)G_O(s)$$

and $G(s)$ contains no time delay, so $G_D(s) = 1$.

(2) $G_O(s)$ is factored into

$$G_O(s) = G_N(s)G_M(s)$$

Because the inverse of $G_O(s)$ is

$$G_O^{-1}(s) = \frac{\begin{bmatrix} s-1 & -2(s-2) \\ -1 & s-2 \end{bmatrix}}{[(s-2)(s-3)]/[(s-1)(s+3)]}$$

we have

$$G_N(s) = \begin{bmatrix} \left(\dfrac{-(s-2)}{s+2}\right) & 0 \\ 0 & 1 \end{bmatrix} \frac{-(s-3)}{s+3}$$

$$G_M(s) = \frac{-1}{(s-3)(s-1)} \begin{bmatrix} -(s+2) & -2(s+2) \\ 1 & s-1 \end{bmatrix}$$

(3) The optimal controller is

$$Q_{opt}(s) = G_M^{-1}(s) = \frac{\begin{bmatrix} s-1 & 2(s+2) \\ -1 & -(s+2) \end{bmatrix}}{(s+2)/(s-1)}$$

(4) The suboptimal controller is

$$Q(s) = Q_{opt}(s)J(s) = \begin{bmatrix} \dfrac{(s-1)(\beta_1 s+1)}{(\lambda_1 s+1)^2} & \dfrac{2(s+2)(\beta_2 s+1)}{(\lambda_2 s+1)^2} \\ \dfrac{-(\beta_1 s+1)}{(\lambda_1 s+1)^2} & \dfrac{-(s+2)(\beta_2 s+1)}{(\lambda_2 s+1)^2} \end{bmatrix} \frac{s-1}{s+2}$$

where

$$J(s) = \begin{bmatrix} \left(\dfrac{\beta_1 s+1}{(\lambda_1 s+1)^2}\right) & 0 \\ 0 & \left(\dfrac{\beta_2 s+1}{(\lambda_2 s+1)^2}\right) \end{bmatrix}$$

Because the plant has only one closed right half-plane pole at $s = 1$, and

$$S(s) = I - G(s)Q(s) = \begin{bmatrix} 1 - \left(\dfrac{(s-2)(s-3)(\beta_1 s+1)}{(s+2)(s+3)(\lambda_1 s+1)^2}\right) & 0 \\ 0 & 1 - \left(\dfrac{-(s-3)(\beta_2 s+1)}{(s+3)(\lambda_2 s+1)^2}\right) \end{bmatrix}$$

we get

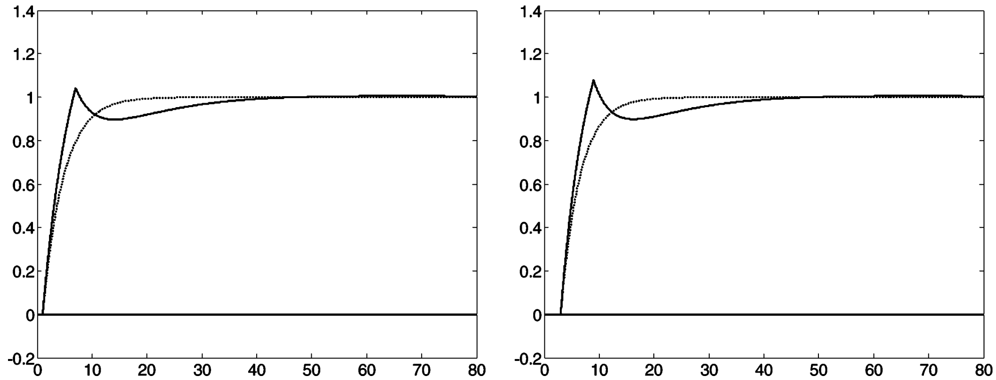

Figure 6. Graphical depiction of the closed-loop responses of the binary distillation column: (a) $r_1 = 1/s$ and $r_2 = 0$; (b) $r_1 = 0$ and $r_2 = 1/s$. (Solid line: proposed method, dotted line: IMC.)

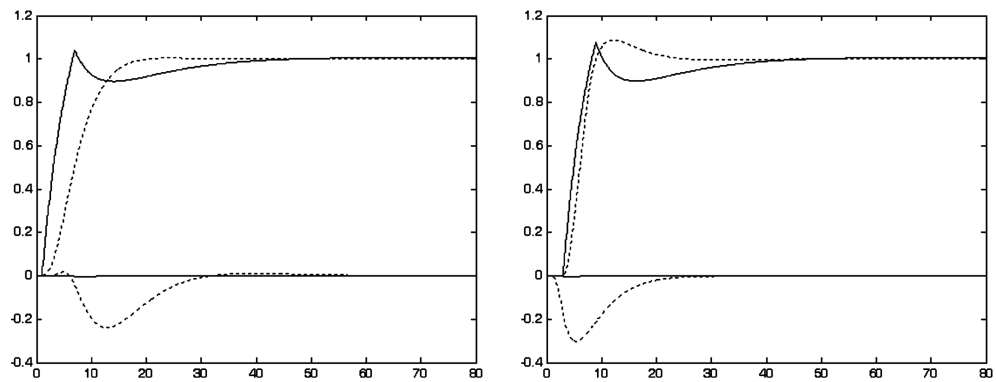

Figure 7. Graphical depiction of the closed-loop responses of the proposed method and the IMC: (a) $r_1 = 1/s$ and $r_2 = 0$; (b) $r_1 = 0$ and $r_2 = 1/s$. (Solid line, proposed method; dotted line, IMC.)

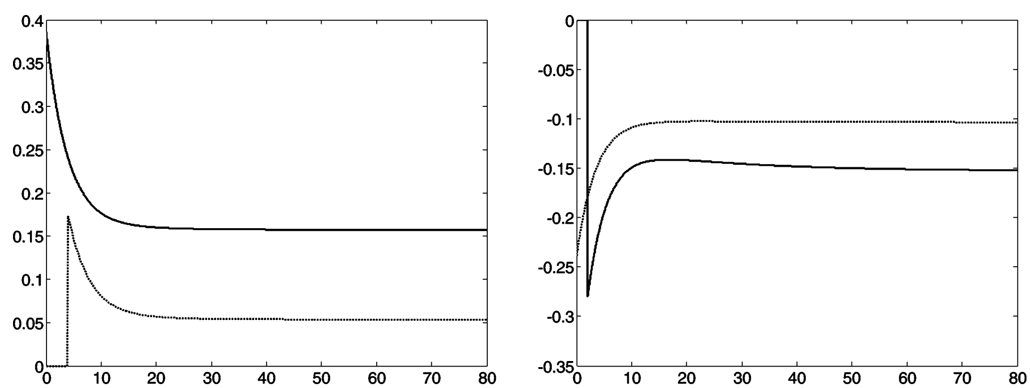

Figure 8. Graphical depiction of the responses of the control variables: (a) $r_1 = 1/s$ and $r_2 = 0$; (b) $r_1 = 0$ and $r_2 = 1/s$. (Solid line, u_1; dotted line, u_2.)

$$\beta_1 = 6(\lambda_1 + 1)^2 - 1$$

$$\beta_2 = 2(\lambda_2 + 1)^2 - 1$$

The unity feedback controller is

$$C(s) = \begin{bmatrix} \dfrac{(s-1)(s+3)\{[6(\lambda_1+1)^2-1]s+1\}}{s(\lambda_1{}^2s^2-(5+10\lambda_1)s+60\lambda_1+36\lambda_1{}^2+20)} & \dfrac{2(s+3)\{[2(\lambda_2+1)^2-1]s+1\}}{s(\lambda_2{}^2s+6\lambda_2{}^2+6\lambda_2+1)} \\ \dfrac{-(s+3)\{[6(\lambda_1+1)^2-1]s+1\}}{s(\lambda_1{}^2s^2-(5+10\lambda_1)s+60\lambda_1+36\lambda_1{}^2+20)} & \dfrac{-(s+3)\{[2(\lambda_2+1)^2-1]s+1\}}{s(\lambda_2{}^2s+6\lambda_2{}^2+6\lambda_2+1)} \end{bmatrix}$$

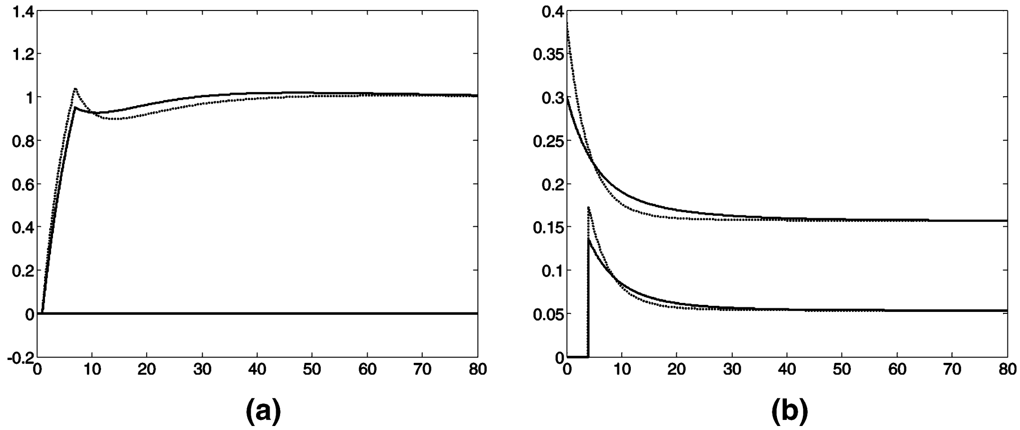

Figure 9. Graphical depiction of the responses of the unconstrained and constrained control variables: (a) output and (b) control variables. (Solid line, constrained; dotted line, unconstrained.)

It is observed that the right half-plane zero-pole cancellations in $C(s)$ have been removed. It might as well take $\lambda_1 = 1$ and $\lambda_2 = 1$ for disturbance loops. The controller then is

$$C(s) = \begin{bmatrix} \dfrac{23s^3 + 47s^2 - 67s - 3}{s(s^2 - 15s + 116)} & \dfrac{14s^2 + 44s + 6}{s(s + 13)} \\ \dfrac{-(23s^2 + 70s + 3)}{s(s^2 - 15s + 116)} & \dfrac{-(7s^2 + 22s + 3)}{s(s + 13)} \end{bmatrix}$$

Because the plant is unstable, there are usually large overshoots in the setpoint responses (Figure 4). We adopt the two-degrees-of-freedom structure. With a similar design procedure, we can obtain the controller for a setpoint loop:

$$C'(s) = \begin{bmatrix} \dfrac{(\lambda_1 s + 1)^2}{(\beta_1 s + 1)(\lambda_1' s + 1)} & 0 \\ 0 & \dfrac{(\lambda_2 s + 1)^2}{(\beta_2 s + 1)(\lambda_2' s + 1)} \end{bmatrix}$$

For 10% undershoot in each setpoint loop, we can take $\lambda_1 = 1.4$ and $\lambda_2 = 0.8$. The controller then is

$$C'(s) = \begin{bmatrix} \dfrac{s^2 + 2s + 1}{32.2s^2 + 24.4s + 1} & 0 \\ 0 & \dfrac{s^2 + 2s + 1}{5.6s^2 + 7.8s + 1} \end{bmatrix}$$

The closed-loop responses are shown in Figure 5. Now there are no overshoots in the setpoint responses. Because the plant has a right half-plane zero, every loop exhibits an inverse response.

Example 8.2. Consider a binary distillation column for separating a mixture of methanol and water (the feed) into a bottom product (mostly water) and a methanol-saturated distillate. In this application, the objective is to control the bottom and top methanol by manipulating the steam flow rate and the reflux flow rate, respectively. Because a change in either steam flow rate or reflux flow rate upsets both methanols, we have an interacting system. The model of the distillation column is described by[21]

$$G(s) = \begin{bmatrix} \dfrac{12.8e^{-s}}{16.7s + 1} & \dfrac{-18.9e^{-3s}}{21s + 1} \\ \dfrac{6.6e^{-7s}}{10.9s + 1} & \dfrac{-19.4e^{-3s}}{14.4s + 1} \end{bmatrix}$$

We use the model because it has been used in many papers on multivariable control and, in fact, the design problem can be regarded as a benchmark problem. The proposed design is conducted as follows:

(1) $G(s)$ is factored into

$$G(s) = G_D(s)G_O(s)$$

Because the inverse of the plant is

$$G^{-1}(s) = \frac{\mathrm{adj}\{G(s)\}}{\det\{G(s)\}}$$

where

$$\mathrm{adj}\{G(s)\} = \begin{bmatrix} \dfrac{-19.4e^{-3s}}{14.4s + 1} & \dfrac{18.9e^{-3s}}{21s + 1} \\ \dfrac{-6.6e^{-7s}}{10.9s + 1} & \dfrac{12.8e^{-s}}{16.7s + 1} \end{bmatrix}$$

and

$$\det\{G(s)\} = \left(\frac{-248.3}{240.5s^2 + 31.1s + 1} - \frac{-124.7e^{-6s}}{228.9s^2 + 31.9s + 1}\right)e^{-4s}$$

$G_D(s)$ is given by

$$G_D(s) = \begin{bmatrix} e^{-s} & 0 \\ 0 & e^{-3s} \end{bmatrix}$$

(2) $G_O(s)$ is an MP, thus $G_N(s) = I$.

(3) The optimal controller is

$$Q_{opt}(s) = \frac{\begin{bmatrix} \dfrac{-19.4}{14.4s + 1} & \dfrac{18.9e^{-2s}}{21s + 1} \\ \dfrac{-6.6e^{-4s}}{10.9s + 1} & \dfrac{12.8}{16.7s + 1} \end{bmatrix}}{\dfrac{-248.3}{240.5s^2 + 31.1s + 1} - \dfrac{-124.7e^{-6s}}{228.9s^2 + 31.9s + 1}}$$

In principle, we can perform the design exactly. By introducing the following filter,

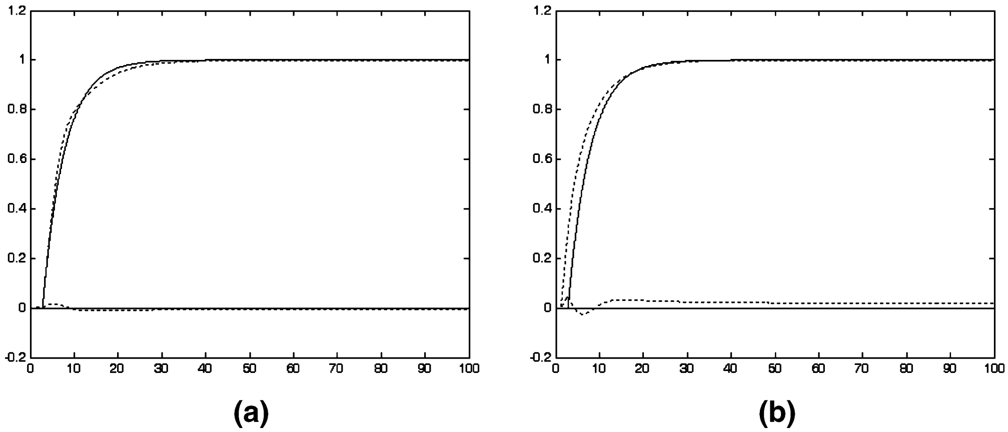

Figure 10. Graphical depiction of the closed-loop responses of the proposed method and the H_∞ method: (a) $r_1 = 1/s$ and $r_2 = 0$; (b) $r_1 = 0$ and $r_2 = 1/s$. (Solid line, proposed method; dotted line, H_∞ method.)

$$J(s) = \begin{bmatrix} \frac{1}{\lambda_1 s + 1} & 0 \\ 0 & \frac{1}{\lambda_2 s + 1} \end{bmatrix}$$

the rigorously analytical solution is

$$Q(s) = \frac{\begin{bmatrix} \frac{-19.4}{14.4s+1} & \frac{18.9e^{-2s}}{21s+1} \\ \frac{-6.6e^{-4s}}{10.9s+1} & \frac{12.8}{16.7s+1} \end{bmatrix}\begin{bmatrix} \frac{1}{\lambda_1 s + 1} & 0 \\ 0 & \frac{1}{\lambda_2 s + 1} \end{bmatrix}}{\frac{-248.3}{240.5s^2+31.1s+1} - \frac{-124.7e^{-6s}}{228.9s^2+31.9s+1}}$$

However, we must implement the controller by a complex feedback structure. As discussed in Section 7, model reduction will make the design simpler and easier without losing too much precision. Assume that we hope the maximum order of each element is two (the higher the order, the better the response and the more complicated the controller). Because

$$\frac{-248.32}{240.5s^2+31.1s+1} - \frac{-124.74e^{-6s}}{228.9s^2+31.9s+1} \approx \frac{-123.58}{134.63s^2+24.24s+1}$$

we get

$$Q_{\mathrm{opt}}(s) = \begin{bmatrix} \frac{-19.4}{14.4s+1} & \frac{18.9e^{-2s}}{21s+1} \\ \frac{-6.6e^{-4s}}{10.9s+1} & \frac{12.8}{16.7s+1} \end{bmatrix} \frac{134.63s^2+24.24s+1}{-123.58}$$

(4) The plant is stable. Let

$$J(s) = \begin{bmatrix} \frac{1}{\lambda_1 s + 1} & 0 \\ 0 & \frac{1}{\lambda_2 s + 1} \end{bmatrix}$$

The IMC controller is

$$Q(s) = \begin{bmatrix} \frac{-19.4}{(14.4s+1)(\lambda_1 s+1)} & \frac{18.9e^{-2s}}{(21s+1)(\lambda_2 s+1)} \\ \frac{-6.6e^{-4s}}{(10.9s+1)(\lambda_1 s+1)} & \frac{12.8}{(16.7s+1)(\lambda_2 s+1)} \end{bmatrix} \times \frac{134.63s^2+24.24s+1}{-123.58}$$

For an overshoot of <5% we can take $\lambda_1 = 3.8$ and $\lambda_2 = 3.5$. The closed-loop responses are shown in Figure 6. It is observed that the closed-loop responses are still thoroughly decoupled, with respect to the approximate $Q(s)$.

We now compare the proposed method with the IMC design, although the comparison is unfair, because of the different performance objectives and the use of rational approximations in IMC design. In the proposed method, the measure of good performance is to minimize the weighted sensitivity for the decoupling response of systems with time delays, whereas the objective of the IMC design is to obtain global optimal solution for the weighted sensitivity problem of rational plants.

Because the IMC design is not applicable to systems with time delays, the time delays in the plant are expanded by the first-order Taylor series. The controller is

$$Q(s) = \frac{\begin{bmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{bmatrix}}{d}$$

where

$$n_{11} = 0.1261s^4 + 0.1931s^3 + 0.03477s^2 + 0.002122s + 4.223 \times 10^{-5}$$

$$n_{21} = -0.09838s^4 + 0.0625s^3 + 0.01268s^2 + 0.0007534s + 1.437 \times 10^{-5}$$

$$n_{12} = -0.5686s^4 - 0.2237s^3 - 0.03146s^2 - 0.001891s - 4.114 \times 10^{-5}$$

$$n_{22} = -0.5067s^4 - 0.1893s^3 - 0.02532s^2 - 0.001416s - 2.786 \times 10^{-5}$$

$$d = 2s^5 + 2.268s^4 + 0.8957s^3 + 0.1527s^2 + 0.0115s + 2.69 \times 10^{-4}$$

The closed-loop responses are shown in Figure 7.

The design specifications may impose a limitation on the control variables, because controllers with good performance may have unrealistic gains and bandwidths. For the aforementioned controllers, the responses of the control variables are shown in Figure 8. It is observed that, for the first output, the peaks of all of the first control variables are >0.36. Suppose that the design specification is that the peaks of the first control variables are <0.3. For the proposed controller, we monotonically increase λ_1 and find that a value of $\lambda_1 = 4.6$ satisfies the specification (Figure 9).

The plant may be uncertain and, thus, the design specification may be that the peaks of the first control variables are <0.3 for all possible plants. In this case, we first determine the worst plant and then monotonically increase λ_1 for the worst plant until the specification is satisfied.

Example 8.3. In this example, the proposed method is compared to the standard H_∞ design. The standard H_2 design can be regarded as a special case of the H_∞ design.[5]

Consider a basis weight and moisture content control system of Fourdrinier machines. When 78 g/m^2 of paper is produced, the plant model is given by[18,22]

$$G(s) = \begin{bmatrix} \dfrac{5.158e^{-2.8s}}{1.8s+1} & \dfrac{-0.2e^{-1.2s}}{2.23s+1} \\ \dfrac{0.44e^{-2.8s}}{1.8s+1} & \dfrac{-1.26e^{-1.2s}}{2.23s+1} \end{bmatrix}$$

The basic idea to compute a decoupling controller using the standard H_∞ method is to expand the time delays in the plant by rational approximations and then let the closed-loop transfer matrix $T(s)$ approximate a prescribed diagonal closed-loop transfer matrix $T_d(s)$, while preserving the stability margin, namely,

$$\min \left\| \begin{matrix} W_2[T(s) - T_d(s)] \\ W_1 C(s)[I + G(s)C(s)]^{-1} \end{matrix} \right\|_2$$

Take the weights

$$W_1 = \frac{0.1(s+0.1)}{s+1} I$$

and

$$W_2 = \frac{10(s+10)}{s+2} I$$

and choose the desired closed-loop transfer matrix as

$$T_d(s) = \begin{bmatrix} \dfrac{e^{-2.8s}}{5s+1} & 0 \\ 0 & \dfrac{e^{-1.2s}}{5s+1} \end{bmatrix}$$

This optimization problem can be solved by converting it to a standard H_∞ control problem. The obtained controller has 12 states. Using order reduction techniques, we readily obtain

$$Q(s) = \frac{\begin{bmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{bmatrix}}{d}$$

where

$$n_{11}(s) = 15.8436s^3 + 909.6482s^2 + 309.9057s + 0.0947$$

$$n_{21}(s) = 1.9703s^3 + 97.1894s^2 - 122.3365s - 1.0486$$

$$n_{12}(s) = 4.6130s^3 + 310.4975s^2 + 46.7891s - 1.0199$$

$$n_{22}(s) = 20.5378s^3 - 4515.6190s^2 - 1585.4215s + 0.1126$$

$$d(s) = s^4 + 272.2374s^3 + 12278.096s^2 + 4.6842s + 0.0268$$

The controller designed by the proposed method is

$$Q(s) = \frac{1}{-6.401} \times \begin{bmatrix} -1.26(1.8s+1) & 0.2(1.8s+1) \\ -0.44(2.23s+1)e^{-1.6s} & 5.15(2.23s+1)e^{-1.6s} \end{bmatrix} \times \begin{bmatrix} \dfrac{1}{\lambda_1 s+1} & 0 \\ 0 & \dfrac{1}{\lambda_2 s+1} \end{bmatrix}$$

To obtain a rise time that is similar to that of the H_∞ method, we take $\lambda_1 = \lambda_2 = 5$. The closed-loop responses are shown in Figure 10.

9. Conclusions

In this paper, we have proposed a simple design procedure for general multivariable systems with time delays. It provides a complete solution for decoupled closed-loop responses. Compared to the conventional decoupling control methods, the proposed method has several important merits:

(1) The conventional decoupling control methods cannot be used for unstable non-minimum phase (NMP) plants with multiple time delays, whereas the proposed method can be used.

(2) The optimal decoupling controller is derived in one step, instead of having the decoupler designed first and then the controller.

(3) It is very difficult to analyze the robustness in the conventional decoupling control system. The robustness issue can be naturally treated in the proposed method by those developed techniques.

Moreover, the proposed method exhibits some characteristics that may not directly be derived using other design methodologies:

(1) It can be used for the control of unstable NMP plants with time delays.

(2) The closed-loop response is decoupled and the design procedure is optimal and analytical.

(3) It provides a simple tuning rule for quantitative performance and robustness.

(4) No state variables and no observer are used.

(5) The controller is of lower order.

(6) The method can directly involve integrating plants.

Three typical examples have been provided to illustrate the proposed method with a step-by-step design procedure.

Our further research focuses on nondecoupled controller design. The decoupled optimal response is practically very important, whereas the optimal response that is not decoupled is theoretically more general, because a global minimum error can be obtained. The results obtained in this paper show a possible way toward the general solution.

Acknowledgment

The first author would like to thank Dr. A. Varga (Germany Aerospace Center) and Dr. Jochen M. Rieber (Institute of System Theory (IST)in Engineering, University of Stuttgart) for their discussions on the inner−outer factorization. This work was completed while W.Z. was an Humbold Research Fellow at the IST, University of Stuttgart. W.Z. would like to thank Prof. Frank Allgower for the good working environment and interesting discussions.

Supporting Information Available: This work is partly supported by the National Science Foundation of China (No. 60474031), NCET (No. 04-0383), SRSP (No. 04QMH1405), and the Alexander von Humboldt Research Fellowship.

Literature Cited

(1) Mayr, O. *The Origins of Feedback Control*; MIT Press: Cambridge, MA, 1970.

(2) Kalman, R. E. A new approach to linear filtering and prediction problems. *J. Basic Eng.* **1960**, *82*, 35−45.

(3) Anderson, B. D. O.; Moore, J. B. *Linear Optimal Control*; Prentice Hall: Englewood Cliffs, NJ, 1971.

(4) Maciejowski, J. M. *Multivariable Feedback Design*; Addison−Wesley: Wokingham, England, 1994.

(5) Zhou, K.; Doyle, J. C.; Glover, K. *Robust and Optimal Control*; Prentice Hall: Englewood Cliffs, NJ, 1996.

(6) Vidyasagar, M. *Control System Synthesis: A Coprime Factorization Approach*; MIT Press: Cambridge, MA, 1985.

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

(8) Cutler, C. R.; Ramaker, B. L. Dynamic matrix control—a computer control algorithm. *Proc. Am. Control Conf.* **1980**, WP5-B.

(9) Ogunnaike, B. A.; Ray, W. H. Multivariable controller design for linear systems having multiple time delays. *AIChE J.* **1979**, *25*, 1043−1057.

(10) Wang, W. *Robust Frequency Response Design Procedure for Multivariable Control Systems*; Shaker Verlag: Aachen, Germany, 2000.

(11) Wang, Q. G. *Decoupling Control*; Springer: Berlin, 2003.

(12) Gu, K.; Kharitonov, V. L.; Chen, J. *Stability of Time Delay Systems*; Birkhäuser: Berlin, 2003.

(13) Niculescu, S.-I., Gu, K.; Eds. *Advances in Time-Delay Systems*; LNCSE, Springer−Verlag: Berlin, 2004; Vol. 38.

(14) Liu, T.; Zhang, W. D.; Gu, D. Y. Analytical Decoupling Control Strategy Using a Unity Feedback Control Structure for MIMO Processes with Time Delays. *J. Process Control*, in press.

(15) Kwakernaak, H. H_2 optimization—Theory and applications to robust control design. In *Robust Control Design*; Kucera, V., Sebek, M., Eds.; Pergamon: Oxford, U.K., 2000; pp 437−448.

(16) Zhang, W. D.; Allgower, F.; Liu, T. Controller Parametrization for SISO and MIMO Plants with Time Delay. *Syst. Control Lett.* **2006**, 55, 794−802.

(17) Skogestad, S.; Postlethwaite, I. *Multivariable Feedback Control: Analysis and Design*; Wiley: Chichester, U.K., 2000.

(18) Zhang, W. D. Robust Control of Systems with Time Delay, Ph.D. Thesis, Zhejiang University, Hangzhou, PRC, 1996.

(19) Zhang, W. D.; Lin, C. Diagonal Inner−Outer and Spectral Factorizations. Unpublished work.

(20) Zhang, W. D.; Xi, Y. G.; Yang, G. K.; Xu, X. M. Design PID controllers for desired time- or frequency-domain response. *ISA Trans.* **2002**, *41*, 511−520.

(21) Wood, R. K.; Berry, M. W. Terminal composition control of a binary distillation column, *Chem. Eng. Sci.* **1973**, *28*, 1707−1717.

(22) Zhang, W. D.; Sun, Y. X.; Xu, X. M. Modeling of Fourdrinier and Cylinder Machines. *Dev. Chem. Eng., Miner. Process.* **2001**, *9*, 69−76.

(23) Zames, G. Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms, and Approximate Inverses. *IEEE Trans. Automatic Control* **1981**, *26* (4), 301−320.

Received for review February 22, 2006
Revised manuscript received July 31, 2006
Accepted August 1, 2006

IE0602310