Design of Multivariable Linear-Quadratic Controllers Using Transfer Functions

Optimal linear-quadratic (LQ) controller design is usually associated with state space techniques. However, when one has measurements of the outputs to be controlled, there are many advantages to designing these LQ controllers using input-output transfer function models. The design procedure leads to a discrete equivalent of the Wiener-Hopf equation, which can be solved using a spectral factorization approach.

In this paper the design procedure is presented and various interpretations of the resulting controllers are discussed. In particular, the controllers are shown to be of the internal model controller (IMC) form, and the Wiener-Hopf procedure is shown to be a powerful way of selecting approximate model inverses and filters that yield good performance and robustness characteristics. The approach treats the problem of simultaneous disturbance rejection and set-point tracking, and it easily handles nonsquare systems.

The design approach, its performance/robustness trade-offs, and the structure of the resulting controllers are demonstrated using models for several processes, including a two-input/one-output sheet forming process, a (3 x 3) multivariable level control problem, and a (2 x 2) multivariable catalytic reactor.

T. J. Harris
Department of Chemical Engineering
Queen's University
Kingston, Ontario, Canada

J. F. MacGregor
Department of Chemical Engineering
McMaster University
Hamilton, Canada L8S 4L7

Introduction

The objectives of process control schemes can vary considerably. In many material and energy integrated processes the control system must maintain stable operation. There may be little economic incentive for tighter control. Tighter control can reduce stability if material and energy balances are upset by large or rapid changes in flows.

However, there exists a broad class of processes for which tight control is desired. These include quality variables in polymer processing, sheet forming, and fiber and other "no blend" type processes. In other instances, there can be economic incentives for moving process set points closer to process or quality constraints. To achieve this one must usually reduce variability.

For situations where tight control is desirable, a reasonable control objective in the single-input/single-output (SISO) case is to minimize the variance of the process output about its set point. Although such a strategy has an appealing motivation, the resulting controllers, known as minimum variance controllers or deadbeat controllers (if the disturbances are deterministic), may have undesirable properties. These include excessive and erratic control action, intersample ripple, and performance and stability characteristics that are sensitive to the accuracy of the process model.

To reduce the magnitude of the control action and improve the robustness of the controller to modeling errors, the variance of the output can be minimized subject to a constraint on the variance of the manipulated variable; that is, one can minimize the quadratic objective function

$$J_1 = \lim_{N \to \infty} \frac{1}{N} E\left\{\sum_{t=1}^{N} e_t^2 + q\,U_t^2\right\}$$
$$= \text{Var}\,(e_t) + q\,\text{Var}\,(U_t) \tag{1}$$

where $E\{\cdot\}$ is the expectation operator, e_t is the deviation of the output Y_t from its set point, and U_t is the deviation of the input from its steady state. This constrained minimum variance (CMV) controller, or linear-quadratic (LQ) output controller, is a subset of the more general class of linear-quadratic state variable controllers. These optimal controllers have been a major topic in the control literature for years. Design procedures for univariate processes were first presented in the frequency domain for continuous-time LQ output controllers by Newton et al. (1957) and Chang (1961). The more general LQ state variable control theory was introduced by Kalman (1960), and has today become so well established (texts by Astrom, 1970; Kwakernaak and Sivan, 1972) that the mere mention of LQ control conjures up the vision of the state space models and Riccati equations of this time domain approach. However, recently there has been a reemergence of the frequency domain approaches (Youla et al. 1976; Youla and Bongiorno, 1985; Kucera, 1979, 1985; Grimble, 1979). These use transfer function models and lead to a Wiener-Hopf equation that is solved using spectral factorization methods. Also in recent years, the idea of internal model control (IMC) has been introduced as a means of both analyzing and designing control systems (Zames, 1981; Brosilow, 1979; Garcia and Morari, 1982, 1985).

In this paper we consider the design of discrete multivariable linear-quadratic controllers using pulse transfer function models to represent the behavior of multivariable processes and their disturbances. The design procedures are presented and related to those of IMC. In particular, these LQ output controllers are shown to be internal model controllers capable of accommodating any arbitrary form of disturbance or set-point variation, and the design procedure is shown to provide a way of obtaining model inverses and robust filters for complex multivariable processes. Furthermore, even though these LQ controllers are designed to optimize a performance criterion in time, they are shown to have some very desirable frequency domain robustness properties. Finally, in the last section of the paper the behavior and the structure of these LQ controllers are illustrated with several examples.

Discrete Multivariable Process and Disturbance Models

Process transfer functions

The input/output behavior of many processes about some operating point can be adequately modeled by discrete transfer functions of the form

$$Y_t = G_M(z^{-1})U_t = \frac{B(z^{-1})}{A(z^{-1})} z^{-b} U_t \tag{2}$$

where Y_t and U_t are deviation variables at discrete time t, and $\{A(z^{-1}), B(z^{-1})\}$ are polynomials in the backward shift operator z^{-1}. Models of the form of Eq. 2 can be obtained by taking the modified z transform of a continuous model, or by identifying the structure and parameters from process data collected from a designed experiment.

The process in Eq. 2 is stable if the poles of $G_M(z^{-1})$, i.e., the roots of $A(z^{-1})$, lie inside the unit circle in the z plane. The process is invertible if the roots of $B(z^{-1})$ lie inside the unit circle in the z plane. The term nonminimum phase is frequently used to describe a process whose zeros lie outside the unit circle. Discrete processes can have zeros outside the unit circle when the underlying continuous process is minimum phase. The choice of control interval then determines whether a discrete process is invertible (Astrom and Wittenmark, 1984).

Multivariable processes with n outputs and m inputs are described by

$$\underline{Y}_t = G_M(z^{-1})\, \underline{U}_t \tag{3}$$

The transfer function between the ith input and jth output is of the form of Eq. 2. To facilitate manipulation of the transfer function matrix, Eq. 3, it is convenient to express it in a right matrix fraction form

$$G_M(z^{-1}) = L(z^{-1}) \cdot [R(z^{-1})]^{-1} \tag{4}$$

Matrix fractions are not unique; two convenient choices are the following. $R(z^{-1})$ can be chosen as the identity matrix multiplied by a scalar polynomial $r(z^{-1})$, which is the greatest common factor of the denominators of $G_M(z^{-1})$. Alternatively, the ith diagonal element of $R(z^{-1})$ can be chosen as the greatest common factor of the denominators of the ith column of $G_M(z^{-1})$. With both these choices, $R(z^{-1})$ is a diagonal polynomial matrix whose leading coefficient is I. $L(z^{-1})$ is a matrix polynomial of the form

$$L(z^{-1}) = L_1 z^{-1} + \cdots L_\ell z^{-\ell}$$

Matrix fractions can be constructed easily by inspection from a transfer function.

In general, $L(z^{-1})$ and $R(z^{-1})$ may be expressed as

$$L(z^{-1}) = \tilde{L}(z^{-1})\, T(z^{-1}) \tag{5}$$

and

$$R(z^{-1}) = \tilde{R}(z^{-1})\, T(z^{-1}) \tag{6}$$

where the pair $\{\tilde{L}, \tilde{R}\}$ have no common factor whose determinant is a function of z^{-1}. The pair $\{\tilde{L}, \tilde{R}\}$ are then said to be coprime, and they form a minimal order realization of the process. The poles of $G_M(z^{-1})$ are defined as the roots of the equation

$$|\tilde{R}(z^{-1})| = 0 \tag{7}$$

(Callier and Desoer, 1982). For square systems the zeroes of $G_M(z^{-1})$ are the roots of the equation

$$|\tilde{L}(z^{-1})| = 0 \tag{8}$$

The zeros located at infinity are associated with the process time delays. The remaining zeros are referred to as the finite zeros of the process (Wolovich and Elliot, 1983). If these zeros are less than one in magnitude the process is said to be invertible. Unlike a SISO process, multivariable processes can have coincident poles and finite zeros that do not cancel.

To avoid unstable pole/zero cancellations in multivariable controller designs, coprime factorizations are required if the process is unstable. For open-loop stable processes coprime representations are not required since at worst they will only lead to stable pole/zero cancellations. Therefore, for such systems right matrix fractions representations obtained by inspection are usually adequate.

Stochastic disturbances

The process description in Eq. 2 is incomplete. If the process is stable and the manipulated variable held at its equilibrium value, Eq. 2 indicates that the process variable will return to its equilibrium value. Industrial processes tend to drift away from their equilibrium values when the manipulated variable is held constant, due to the presence of process disturbances. These disturbances may be deterministic in nature, i.e., steps, ramps, or an exponential rise to a new level. Many times they are of a more random or stochastic nature, and can be adequately modeled by autoregressive integrated moving average (ARIMA) time series models (Box and Jenkins, 1970). Many industrial disturbances, both stochastic and deteministic, can be modeled by simple or low-order models.

Consider the disturbance defined by the difference equation

$$D_t = (1 + \Phi) D_{t-1} - \Phi D_{t-2} + a_t \tag{9}$$

D_t is the current value of the disturbance. Φ is the autoregressive parameter $\{-1 < \Phi < 1\}$. $\{a_t\}$ is a sequence of independently distributed random variables with mean 0 and variance σ_a^2. Typical realizations of Eq. 9 for $\Phi = 0$ (known as a random walk) and for $\Phi = 0.8$ are shown in Figures 1a and 1b. We notice in both cases that the disturbance is not centered about zero. This is referred to as nonstationarity. The transfer function representation of Eq. 9 is

$$D_t = \frac{1}{\nabla(1 - \Phi z^{-1})} a_t \tag{10}$$

where ∇ is defined as the backward difference operator $(1 - z^{-1})$. The presence of a single root on the unit circle allows for nonstationarity behavior in the mean. Equation 10 is known as an integrated autoregressive (IAR) model.

Another frequently encountered disturbance model is the integrated moving average (IMA) model. The difference equation representation is

$$D_t = D_{t-1} + a_t - \theta a_{t-1} \tag{11}$$

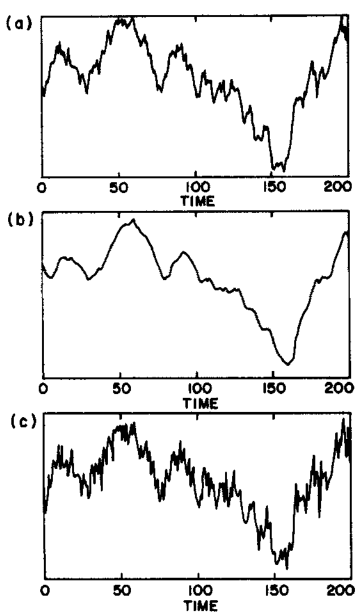

Figure 1. Simulated stochastic disturbances.
a. Random walk $\nabla D_t = a_t$
b. IAR $(1 - 0.8 Z^{-1})\nabla D_t = a_t$
c. IMA $\nabla D_t = (1 - 0.5 z^{-1})a_t$

For $\theta = 0$ this again reduces to a random walk, Figure 1a, while for $\theta > 0$ it can be shown to be equivalent to a random walk plus superimposed white noise. Larger values of θ denote larger measurement noise (Box and Jenkins, 1970). A realization of Eq. 11 for $\theta = 0.5$ is shown in Figure 1c. More general ARIMA models of the form

$$\Phi(z^{-1}) \nabla^d D_t = \theta(z^{-1}) a_t \tag{12}$$

are discussed by Box and Jenkins.

Multivariable ARIMA disturbance models can be defined in a similar fashion. It is also often convenient to express these models in a right matrix fraction form

$$\underline{D}_t = \theta(z^{-1}) \Phi^{-1}(z^{-1}) \nabla^{-d} \underline{a}_t \tag{13}$$

where $\Phi(z^{-1})$ is a diagonal polynomial matrix and $\nabla^d = (1 - z^{-1})^d I$.

Deterministic disturbances

In Eqs. 12 and 13 it was assumed that the driving force, the a_t's, were always present. Instead, nonzero values may occur only infrequently. Figures 2a and 2b show the behavior of the previously studied IAR disturbances in these circumstances. The disturbances are now deterministic in nature, but occur at random times. For the IAR disturbances, the time constant τ of the exponential rise is related to the autoregressive parameter by $\Phi = \exp(-T/\tau)$ where T is the sampling interval. MacGregor et al. (1984) refer to these as randomly occurring deterministic disturbances, Astrom and Wittenmark (1984) call them piecewise deterministic signals, and Johnson (1972) calls them waveform descriptions.

The essential difference between models for stochastic and deterministic disturbances is in the frequency or the probability distribution of the a_t sequence. The difference equations modeling the disturbances are of the identical structure. Therefore, as discussed by MacGregor et al. (1984), since LQ controllers are

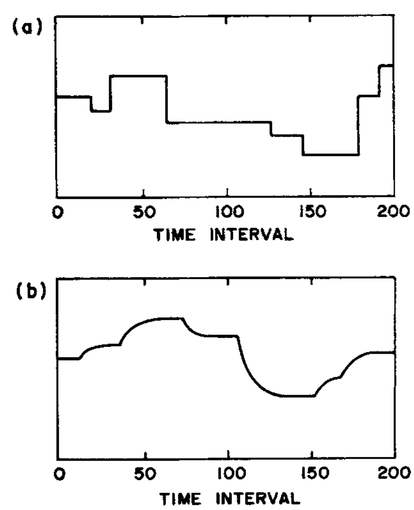

Figure 2. Simulated randomly occurring deterministic disturbances.
a. Random steps $\nabla D_t = a_t$
b. Random exponential $(1 - 0.8 z^{-1})\nabla D_t = a_t$

independent of the distribution of the a_t's, the controller designed for a given difference equation structure, e.g., Eq. 9, will be optimal for both the stochastic and the deterministic disturbances described by that model. Thus, no real distinction should be made between optimal stochastic control and optimal control for deterministic disturbances occurring randomly in time. The important point is that in designing the controller one must use appropriate models to characterize the major features of the disturbances that are present in the system.

Role of disturbances

The complete process model is the superposition of the process dynamics plus the disturbances

$$Y_t = G_M(z^{-1})U_t + D_t \tag{14}$$

The disturbances, deterministic or stochastic, have a dual role. Frequently the model, Eq. 14, must be identified from input/output data. The parameters of the disturbance model need to be jointly estimated with those of the transfer function G_M if the correct structure (order) and parameters of the transfer function are to be identified efficiently. The disturbance model used for controller design might differ from that determined from the process identification stage if the nature of the process disturbances or set-point changes expected in the future is different from that present during identification.

All optimal controllers (i.e., those optimizing a performance index) must be based on a specified disturbance (or set-point change) model. Often in the design of such controllers the disturbance model is not explicitly stated, but rather indirectly implied in the problem formulation. A common example is the state variable LQ design approach commonly presented in text books, in which the implied disturbance is simply an impulse or sudden deviation from steady state at $t = 0$. Such unrealistic disturbance assumptions often lead to very poor controllers, such as controllers with no integral action. If one specifies, directly or indirectly, in an optimal controller design that there are no nonstationary disturbances that might cause the process to drift away from set point, then obviously integral action will not be needed. Ignoring such basic ideas about the nature of process disturbances in designing controllers has often led to poor results and totally unwarranted criticisms of the design procedure.

Although it is essential to model the nature of the process disturbances or set-point changes, detailed models and precise estimates of their parameters are not usually essential. It is usually sufficient that the model capture only the major features of the disturbance, such as the degree of nonstationarity and the dominant frequency content of its power spectrum. (For example, is it a smooth drifting load type of disturbance or one with a considerable amount of high-frequency noise?) One reason for this is that mismatch between the nature of the disturbances and that predicted by the model will not directly affect the stability of the closed-loop system. However, the structure of the disturbance model assumed will have a strong effect on the performance of the controller, and it will affect the sensitivity of the controller to mismatch between the process and its transfer function model (Bergh and MacGregor, 1987a).

In some sense all model-based controllers must make predictions of both the effect that past control actions will have on future outputs, and the effect that disturbances will have on these future outputs. The transfer function provides the prediction of the former, and the disturbance model provides the prediction of the latter. For the IMA(1, 1) model, Eq. 11, it can be shown that the best least-squares predictor at time t is given by the exponentially weighted moving average filter acting on past estimates of the disturbance, i.e.,

$$\begin{aligned}\hat{D}_{t+b|t} &= (1-\theta)(D_t + \theta D_{t-1} + \theta^2 D_{t-2} + \cdots) \\ &= \frac{(1-\theta)}{(1-\theta z^{-1})} D_t \end{aligned} \tag{15}$$

Internal model control (IMC) and dynamic matrix control (DMC) are based on the assumption of step disturbances ($\theta = 0$). As can be seen from Eq. 15, this implies that the prediction for all future time is given by $\hat{D}_{t+b/t} = D_t$, the present disturbance.

These approaches can perform poorly in the presence of much smoother load type disturbances such as steps passing through the process (Bergh and MacGregor, 1987a). Therefore, in such circumstances it may be worthwhile modifying them to account for this type of disturbance by employing integrated autoregressive models such as Eq. 9 with $\Phi > 0$; see the $\Phi = 0.8$ case in Figures 1b and 2b. For this disturbance the prediction is given by the complementary solution to difference Eq. 9, i.e.,

$$\hat{D}_{t+b|t} = \left[\frac{1-\Phi^{b+1}}{1-\Phi}\right] D_t - \left[\frac{\Phi(1-\Phi^b)}{1-\Phi}\right] D_{t-1} \tag{16}$$

The generalized analytical predictor of Wellons and Edgar (1986) uses such a model for the disturbance prediction.

Controller Design

In this section, we give an overview of the control theory, the controller design, and some properties of the controller. Most of the theoretical derivations and computational discussions have been placed in the appendices available as supplementary material.

Quadratic objective function

As discussed in the introduction, we are interested in the design of linear controllers that optimize a quadratic performance index involving the process outputs and inputs. One such performance index for a SISO system was given in Eq. 1. However, when the disturbance models exhibit nonstationary behavior, it is impossible to stabilize the variance of Y from set point when the manipulated variable is constrained to its steady state value. The manipulated variable must be allowed to float. This is accomplished by constraining the change in control action $\nabla U_t = U_t - U_{t-1}$. The objective function to be minimized is then of the form

$$\begin{aligned} J_2 &= \lim_{N\to\infty} \frac{1}{N} E\left\{\sum_{t=1}^{N} [(e_t)^2 + q(\nabla U_t)^2]\right\} \\ &= \mathrm{Var}\,(e) + q\,\mathrm{Var}\,(\nabla U) \end{aligned} \tag{17}$$

When the process disturbances are modeled with a root at unity and the objective function, Eq. 17 is minimized, the resulting controller will always contain integral action. A common misconception is that integral action is introduced by simply penalizing the incremental control action, i.e., ∇U_t. In fact, the disturbance model must be nonstationary for integral action to be present (Box and Jenkins, 1970; MacGregor, 1973; Palmor and Shinnar, 1979). The consequences of this for adaptive control are discussed by Harris et al. (1978, 1980).

The multivariate version of Eq. 17 is to minimize

$$J_3 = \lim_{N\to\infty} \frac{1}{N} E\left\{\sum_{t=1}^{N} \underline{e}_t^T Q_1 \underline{e}_t + \nabla \underline{U}_t^T Q_2 \nabla \underline{U}_t\right\} \qquad (18)$$

where $\{Q_1, Q_2\}$ are positive semidefinite weighting matrices.

The choice of Q_1 is usually straightforward. It is frequently chosen to be diagonal, with the ith diagonal element, q_{ii}, directly related to the relative importance of the ith output. Typically, q_{ii} is inversely proportional to the square of the desired control range for the ith output. Off-diagonal elements could be used to penalize covariance between variables. For example, in reactor control one might want to avoid simultaneous high concentrations of the limiting reactant and high temperatures in order to minimize the potential for runaway. A positive off-diagonal element in the appropriate position of Q_1 would tend to prevent this situation.

Q_2 is the main tuning parameter. It is most often selected as a diagonal matrix and the individual elements are adjusted so that the variances of the inputs and outputs are jointly acceptable (MacGregor and Wong, 1980). Larger elements in the ith diagonal position will penalize the variance of the ith input variable more heavily.

The objective function, Eq. 17, can be interpreted as the Lagrangian multiplier solution to minimizing the variance of the deviations of Y from set point, subject to the constraint that the variance of ∇U not exceed some prescribed value (Chang, 1961). The multivariate interpretation of Eq. 18 is then to simultaneously minimize the weighted variance of each Y_i, subject to constraints on the variances of the inputs. Hence, a very important property of constrained minimum variance (CMV) controllers is that, when the process model is correct:

1. For a given variance of Y, no other linear controller will have less control effort as measured by the variance of ∇U
2. For a given control effort, as measured by the variance of ∇U, no other linear controller has a smaller variance of Y for the specified disturbance

For a specific disturbance structure, the best performance, as measured by the variance, is the solution to Eq. 18 with $Q_1 = I$ and $Q_2 = 0$. This controller then defines the best performance against which all other designs, (IMC, PID, etc.) can be compared. This is an important property. It may happen that the current level of control is perceived as being unacceptable. In some instances this level of performance may not differ appreciably from the best achievable performance. In these cases, reduction of the dead time via process and/or sampling procedures, or feedforward control may be the only means of improving performance.

The variance of the inputs and outputs is closely linked to frequency domain properties. Consider the closed-loop transfer function between the process output and the disturbance.

$$Y_t = (1 + GC)^{-1} D_t \qquad (19)$$

where C is the feedback controller transfer function. The variance of Y for a SISO system is then given by Parseval's theorem as

$$\mathrm{Var}\,(Y) = \frac{1}{\pi} \int_0^{\pi} [|1 + GC|^{-1} |D|]^2 \, d\omega \qquad (20)$$

where $z^{-1} = e^{-i\omega T}$. For MIMO systems this is discussed further in a later section. If the magnitude of the sensitivity function $(1 + GC)$ for a controller is less, for all frequencies, than the magnitude of that for an alternative controller, then the variance of Y for the first controller will be less than for the second for any arbitrary disturbance.

Development of the control equations

The controller that minimizes the quadratic criterion, Eq. 18, can be obtained by different approaches depending upon the structure of the models used. If the state space representation of the models in Eqs. 3 and 13 is used, then the calculus of variations (Astrom, 1970) can be used to obtain the controller

$$\underline{U}_t = -L_\infty \hat{x}_{t+b-1/t} \qquad (21)$$

where L_∞ is the steady state ($N \to \infty$) solution of a matrix Riccati equation, and $\hat{x}_{t+b-1/t}$ is the state vector estimate for time $t + b - 1$, given information up to and including time t obtained from a Kalman filter. This approach is well understood and computer programs for solving it are readily available. When the disturbances are modeled as being nonstationary these optimal LQ controllers will contain integral action (MacGregor, 1973; MacGregor and Wong, 1980). However, when one is interested in output control, as in Eq. 18, it can be argued that by using the above general state variable approach, physical insight into the controller problem is obscured for the sake of the mathematical optimization procedures.

In the past decade there has been a reemergence of transfer function approaches for solving the LQ problem (Youla et al., 1976; Youla and Bongiorno, 1985; Kucera, 1979; Grimble, 1979). Many of these solutions appear quite formidable, requiring coprime factorizations and solutions of simultaneous Diophantine equations. However, when the processes are stable, or at most involve roots on the unit circle, the solutions can be greatly simplified (Youla et al.; Youla and Bongiorno). Although SISO unstable processes are easily handled, MIMO unstable systems are more difficult. Since nearly all chemical processes fit into the category of being open-loop stable, we make this assumption throughout the remainder of this paper. The transfer function solution to the LQ problem for MIMO processes is outlined in the next section, and comparisons are made to other design procedures.

The algorithm

Consider the general process transfer function model, Eq. 3, together with the disturbance and set point models of the form of Eq. 13. It is shown in appendix A1 of the supplementary material that the controller minimizing the quadratic objective function, Eq. 18, is of the form

$$\underline{U}_t = -H_1 [\underline{Y}_t - G_M \underline{U}_t] + H_2 \underline{Y}_{sp,t} \qquad (22)$$

where $\underline{Y}_{sp,t}$ is the mean corrected set point and G_M is the process model transfer function. H_1 and H_2 are rational polynomial matrices that may be conveniently expressed as the product of two other rational matrices:

$$H_1 = \tilde{G}_M^{-1} \cdot F_1 \tag{23}$$

$$H_2 = \tilde{G}_M^{-1} \cdot F_2 \tag{24}$$

The operator z^{-1} has been omitted where obvious for conciseness. $\tilde{G}_M^{-1}$ is an approximate transfer function model inverse that is independent of the disturbance model. F_1 and F_2 are filters of dimension $m \times n$ that depend on the nature of the set-point changes and disturbances, respectively, as well as on $\tilde{G}_M^{-1}$. The block diagram of the controller, Figure 3, shows it to be of the internal model form (Zames, 1981; Garcia and Morari, 1985).

The control is not expressed in terms of only error feedback as is common with many controllers. The reason for this is that the disturbance model may differ from that of the set-point model. If the models for the set-point changes and the disturbances are identical then the filters F_1 and F_2 will be identical, and the controller can be expressed only in terms of error feedback, that is as

$$\underline{U}_t = -H_1[(\underline{Y}_t - \underline{Y}_{sp,t}) - G_M\underline{U}_t] \tag{25}$$

The approximate model inverse can be expressed as

$$\tilde{G}_M^{-1} = R\Gamma^{-1} \tag{26}$$

where R is the right matrix fraction term in G_M, Eq. 4, and Γ is an $m \times m$ matrix polynomial of the form

$$\Gamma = \Gamma_0 + \Gamma_1 z^{-1} + \cdots + \Gamma_k z^{-k}$$

which is the unique solution to the spectral factorization equation

$$\Gamma^*\Gamma = L^*Q_1L + \nabla^{*d}R^*Q_2R\nabla^d \tag{27}$$

having all its roots inside the unit circle. Γ^* denotes the complex conjugate of Γ. Γ^{-1} is always stable and causal. For the case of minimum variance control ($Q_2 = 0$) in a SISO system with finite zeroes inside the unit circle, $\tilde{G}_M^{-1}$ is simply the inverse of the process transfer function excluding the dead-time term. For noninvertible models the solution to Eq. 27 will always give a stable inverse. In the minimum variance SISO case it simply reflects the finite zeros outside the unit circle (Z_i) to the position ($1/Z_i$) inside the unit circle.

Using the above relationships, Eqs. 23, 24, and 26, the controller, Eq. 22, can be conveniently expressed as a function of past inputs and outputs as

$$(\Gamma - F_1L)R^{-1}\underline{U}_t = -F_1\underline{Y}_t + F_2\underline{Y}_{sp,t} \tag{28}$$

The filters $\{F_1, F_2\}$ depend on the set-point variation and disturbance models. They do not depend upon the distribution of the driving shock sequences $\{\boldsymbol{a}_t\}$ or on their covariance structure, but only on the structure of the difference equation model. For disturbance or set-point models in the right matrix fraction form, Eq. 13, the filters can be expressed as

$$F = T\theta^{-1} \tag{29}$$

where the polynomial $T(z^{-1})$ comes from the solution of the bilateral or Diophantine equation

$$L^*Q_1\theta = \Gamma^*T + zP^*\nabla^d\Phi \tag{30}$$

In the minimum variance SISO case, F is simply the minimum variance prediction filter for the disturbance b steps into the future ($\hat{D}_{t+b|t} = FD_t$). In the MIMO case, the interpretation is not as clear unless the minimum delays occur along the diagonal elements of the process transfer function. In these instances, F also has the interpretation of a disturbance prediction over a finite horizon. For example, for independent step changes in a set point, F will simply equal the identity matrix. For the more general constrained minimum variance (CMV) problem (i.e., $Q_2 \neq 0$) the filters F_1 and F_2 will also depend upon Γ in the approximate inverse, Eq. 26.

The controller matrices H_1 and H_2 obtained by this spectral factorization solution to the Wiener-Hopf equations (appendix A1) are guaranteed to be stable. For any controller parameterized in the form of Eq. 22, the stability of the controller matrices $\{H_1, H_2\}$ and the open-loop process G guarantee stability of the closed-loop system when there is no process model mismatch (Zames, 1981; Desoer and Chen, 1981; Garcia and Morari, 1985). (However, even though these conditions may be satisfied, the controller, Eq. 28, may still have poles outside the unit circle.) When there exists process/model mismatch then closed-loop stability cannot be guaranteed. However, as shown later the closed-loop system can be made more robust to modeling errors by increasing Q_2 in the objective function, that is, by constraining the size of the input manipulations. Furthermore, for an open-loop stable process there exists some value of Q_2 sufficiently large than it can guarantee robustness to arbitrarily large modeling errors, provided that $Re\{\lambda_j[G(1)\tilde{G}_M^{-1}(1)]\} > 0$; $j = 1, 2, \ldots n$, where $\lambda_j[A]$ denotes the jth eigenvalue of A, and G is the true process transfer function (Grosdidier et al., 1985).

Very efficient algorithms with guaranteed convergence properties exist for the computation of the spectral factor Γ (Wilson, 1970; Davis and Dickinson, 1983; Jezek and Kucera, 1985); these are discussed in appendix A2 of the supplementary material. The solution of the Diophantine equations for the filters is straightforward for simple disturbance models. For example, for the IMA disturbance, Eq. 11, the filter is given by $F_1 = Q_1^{1/2} \cdot (I - \theta)(I - \theta z^{-1})^{-1}$ regardless of the approximate model inverse $\tilde{G}_M^{-1}$. For more general structures see appendix A3, supplementary material.

Effect of dimensions of input and output vectors

The Wiener-Hopf design procedure outlined in the preceding section is perfectly general and is capable of handling cases where the dimensions of the input and output vectors are different. However, a solution for the existence of a stable inverse exists if and only if the righthand side of the spectral factorization equation, Eq. 27, is nonsingular. This is equivalent to the condition that the $[(n + m) \times m]$ matrix

$$\begin{bmatrix} Q_1 L(z^{-1}) \\ (1 - z^{-1})^d Q_2 R(z^{-1}) \end{bmatrix} \tag{31}$$

has rank m for all z on the unit circle (Wilson, 1972; Kucera, 1979). This rank condition can be checked by letting $z^{-1} = e^{-i\omega T}$, and testing the rank condition at a number of equispaced points between $[0, \pi]$.

The properties of the approximate inverses and filters arising from the LQ controller design depend upon the dimensions of the input and output vectors. Various cases are examined in turn.

1. Number of Outputs Equals Number of Inputs ($n = m$). When the rank condition holds, and when the disturbances are nonstationary, the approximate inverse and filter have the property that

$$\tilde{G}_M^{-1}(1) \cdot F(1) = [G_M(1)]^{-1}$$

where $G_M(1)$ is the steady state gain of the process model. The controller then has integral action in each channel since the lefthand side of Eq. 28 vanishes at $z = 1$. For square systems an alternative scaling for spectral factorization Eq. 27 is to factorize the lefthand side as $\Gamma^* Q_1 \Gamma$. This scaling ensures that the model inverse and filter gains are given by $\tilde{G}_M^{-1}(1) = [G_M(1)]^{-1}$ and $F(1) = I$, as is usual in most IMC designs.

2. More Inputs than Outputs ($m > n$). To satisfy the rank condition, Eq. 30, at $\omega = 0$ it is necessary that $(m - n)$ manipulated variables not have reset action, but rather be constrained about some steady state value. This implies that the second term in the quadratic objective function, Eq. 18, be of the form

$$\nabla^d \underline{U}_t^T Q_2 \nabla^d \underline{U}_t \tag{32}$$

where

$$\nabla^d = \text{diag}\,\{(1 - z^{-1})^{d_1}, \ldots, (1 - z^{-1})^{d_m}\}$$

where $(n - m)$ of the d_i's are zero. An example in a later section will illustrate this situation.

3. More Outputs than Inputs ($n > m$). In this case the objective function need not be modified. However, there will not be enough manipulated variables to eliminate offsets in all of the outputs resulting from set-point changes or nonstationary disturbances. The n outputs can now be controlled about their set points only in a weighted least squares sense.

Robustness and sensitivity norm interpretations

These multivariable LQ or constrained minimum variance (CMV) controllers are optimal controllers in the sense that they are designed to optimize the quadratic criterion, Eq. 18, for a particular class of disturbances. For these controllers to be useful they must also satisfy other criteria such as stability robustness to process/model mismatch, and performance robustness to disturbance structures other than that used in the design. From an intuitive viewpoint one would expect that, as the input variables are more heavily constrained via increasing Q_2 in the controller design, the closed-loop system would exhibit greater robustness to process modeling errors because much less demand is being placed upon the model to predict the effect of large changes in the inputs. The robustness of SISO controllers for changes in the model parameters has been investigated by Palmor and Shinnar (1979) and by Bergh and MacGregor (1987a). In this section we try to shed further light on these properties by examining frequency domain interpretations of these LQ controllers.

Consider the block diagram of Figure 3 with $F_1 = F_2$. Under conditions of no process/model mismatch the closed-loop relationship between the output and the disturbance is given by

$$\begin{aligned} \underline{Y}_t &= (I - G_M H_1)\underline{D}_t \\ &= S\,\underline{D}_t \end{aligned} \tag{33}$$

where $S(z^{-1})$ is the sensitivity matrix of the closed-loop system. The relationship between the output and set point is

$$\begin{aligned} \underline{Y}_t &= G_M H_1 \underline{Y}_{sp,t} \\ &= (I - S)\underline{Y}_{sp,t} \end{aligned} \tag{34}$$

where $I - S(z^{-1})$ is the complementary senstivity matrix.

For good performance one would like to minimize some norm of the sensitivity matrix S over the frequency ranges for which the power spectra of the disturbances and set-point changes are large. One could choose the weighted H^2 norm (Francis, 1982), i.e.,

$$\text{Min}\ \|W_D S\|^2 \tag{35}$$

where the weighting matrix $W_D(z^{-1})$ might be chosen such that $\|W_D\|^2$ is the power spectrum of the disturbance Φ_D.

When process/model mismatch is present, the robustness of the system to this mismatch is an important issue. Robustness theorems based on the Nyquist stability criterion (Doyle and Stein, 1981; Lehtomanki et al., 1985; Kwakernaak, 1985) show that to achieve robustness to uncertainties in the model the "gain" of the complementary sensitivity matrix should be small at frequencies where the mismatch is large. Typically, this mismatch is largest at high frequencies, and therefore, this "small gain theorem" would require that the gain (usually expressed in terms of singular values) of $I - S(z^{-1})$ should roll off at high frequencies. However, if a continuous time system is treated appropriately with anti-aliasing filters and is sampled slowly enough, the resulting discrete time system will contain much less high frequency uncertainty (Rohrs et al., 1985). This increased robustness resulting from the sampling process is a generally unrecognized advantage of discrete model-based controllers.

From the above discussion, it would appear that to improve the robustness of the closed-loop system one should minimize some norm of the complementary sensitivity matrix over the frequency range where the mismatch is large (usually high frequencies). Again one might consider a quadratic norm, that is,

$$\text{Min}\ \|W_E(I - S)\|^2 \tag{36}$$

Finally, to achieve a controller design that would yield a good compromise between performance and robustness, one could combine Eqs. 35 and 36 (Francis, 1982), that is,

$$\text{Min}\ \{\|W_D S\|^2 + \|W_E(I - S)\|^2\} \tag{37}$$

To compare this intuitive design criterion with the LQ criterion, Eq. 18, we can express the LQ criterion in the frequency domain as (Kwakernaak and Sivan, 1972)

$$\text{Min}\ \frac{1}{2\pi j} \oint_{|z|=1} \text{tr}\,\{Q_1 \Phi_Y + Q_2 \Phi_{\nabla U}\} \frac{dz}{z} \tag{38}$$

where Φ_y and $\Phi_{\nabla U}$ are the spectra of Y and ∇U, respectively.

Using Eqs. 33 and 34 this can be expressed as

$$\text{Min } \frac{1}{2\pi j} \oint_{|z|=1} \text{tr } \{Q_1 \Phi_D SS^* + Q_2 G^{-1} G^{*-1} \Phi_{\nabla D}(I - S)(I - S)^*\} \frac{dz}{z} \quad (39)$$

$$= \text{Min } \{\|W_1 S\|^2 + \|W_2(I - S)\|^2\} \quad (40)$$

where $|W_1|^2 = Q_1 \Phi_D$ is the weighted spectrum of the disturbance, and $|W_2|^2 = Q_2 G^{-1} G^{*-1} \Phi_{\nabla D}$ where $\Phi_{\nabla D}$ is the spectrum of the differenced disturbance.

Hence, we see that the standard LQ or CMV controller is of the same form as that obtained by the H^2-norm minimization criterion, Eq. 37.

In most process control situations the magnitude of the W_1 weight on the sensitivity function will be large at low frequencies and small at high frequencies, since this is the nature of most common disturbance spectra. On the other hand, the magnitude of the weight W_2 on the complementary sensitivity function will be small at low frequencies and large at high frequencies. This is exactly the type of weighting compromise that one would normally choose for a design to have good performance and good robustness (Morari and Doyle, 1986). In general, by changing the magnitude of W_2 (or equivalently the input constraint matrix Q_2) we can influence the relative weight we are giving in the design to performance as opposed to robustness. However, it should be mentioned that LQ designs only treat robustness in the above general manner. They do not provide designs with guaranteed robustness for a specific process/model mismatch. H^∞ sensitivity norm designs are usually required for such guarantees (Kwakernaak, 1985). However, these Wiener-Hopf design techniques using spectral factorization can be applied to much more general quadratic criteria than given in Eq. 18 or equivalently in Eq. 39. In particular, any quadratic terms that relate to other performance or robustness measures can be included to give what the designer feels to be a better overall objective (Youla and Bongiorno, 1985). It has also been shown by Grimble (1986) that H^∞-norm minimizations can also be solved by Wiener-Hopf design techniques.

Discussion

In this section we briefly discuss further relationships between the LQ or CMV controllers and some other approaches to multivariate controller design.

Internal model structure

As shown earlier, these LQ controllers are of the internal model form, Figure 3. The spectral factorization design procedures provide a natural way of obtaining stable approximate multivariable model inverses, and of obtaining filters for any given type of disturbance or set point change. The combination of model inverses and filters arising from this design will give the optimal performance in terms of output variances for any specified bounds on the input variances. Furthermore, as one increases the input constraining matrix Q_2 these LQ controllers become very robust to process/model mismatch.

Efficient numerical algorithms for spectral factorization with guaranteed convergence properties exist for computing the spectral factors (appendix A2, supplementary material). The frequency domain algorithms are independent of the complexity of the transfer function models. That is, models having elements with different orders, different periods of delay, noninvertible zeroes, etc., are just as easily inverted as simple models. The computation difficulty depends only upon the dimension of the process (n). These algorithm have been used to find approximate inverses for problems as large as $m = n = 12$. Furthermore, the algorithm always guarantees a stable inverse having the correct gains to ensure the presence of zero steady state offsets.

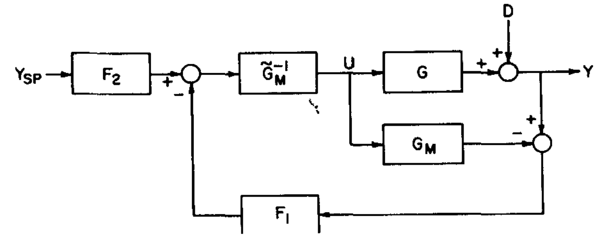

Figure 3. Internal model structure of LQ controllers.

The filters obtained from the Diophantine Eq. 30 depend upon the nature of the disturbance or set-point changes in the system. They also depend, in general, upon the model inverse obtained via the spectral factorization algorithm, Eq. 27; that is, they are matched to the inverse. For set-point changes the filters are usually diagonal, but for load disturbances they are usually full matrices. A necessary, but not sufficient, condition for diagonal filters is that the disturbances are decoupled, that is, all disturbances in the system affect one and only one output.

Decoupling is rarely optimal for the quadratic performance index, Eq. 18. As shown by the LQ designs discussed in this paper, decoupling will only be optimal when $G_M H = L\Gamma^{-1}F$ is diagonal, that is when both the filter F and $L\Gamma^{-1}$ are diagonal. The first condition rarely holds for process disturbances encountered in practice, and the latter conditions would almost never hold when $Q_2 \neq 0$ (see Eq. 27), or when the dynamics have complex delay patterns or noninvertible characteristics.

In other IMC design approaches the filters are chosen to be diagonal with elements of a fixed but arbitrary form (usually first-order low pass filters) in order to make the design simple, and to allow for on-line tuning of the filter parameters. To retune the LQ controllers presented here the spectral factorization algorithms would have to be resolved for a new choice of Q_2 or Q_1. That is, these controllers do not provide a set of arbitrary parameters that can be explicity tuned on-line. However, it should be recognized that the obvious advantage of using the former simplification cannot usually be achieved without some loss in performance and/or robustness.

Relationship to prediction error methods

Dynamic matrix control or DMC (Cutler and Ramaker, 1976; Prett and Gillette, 1979) and model algorithmic control or MAC (Richalet et al., 1978) are also *linear-quadratic control* algorithms. If in DMC the prediction horizon and the number of future control actions (U_t) to be calculated both are allowed to tend to infinity, then, for the same weighting matrices in the quadratic performance index, DMC would be equivalent to the LQ or CMV controller discussed here. However, since in DMC the matrix size is directly related to the number of future controls (U) to be solved for, this number is usually kept small, but only the first control action is implemented; then the control actions are recomputed at the next interval. Hard constraints on the inputs and outputs can be accommodated within DMC by using an on-line numerical optimization (quadratic programming) at each interval on a finite number of future inputs and over a finite prediction horizon. The analytical LQ solutions presented here do not directly accommodate such hard constraints. However, they can be modified in a simple manner to provide for any form of input saturation (linear, nonlinear, or time-varying) in a one-step optimal sense (Segall et al., 1987). Although the DMC algorithm is designed for step disturbances in the output, it can easily be modified to handle other disturbance structures.

Examples

In this section we examine the behavior and structure of these LQ controllers in four example systems.

SISO example

We start by considering a simple SISO process, and examine the behavior of the controller for a range of values of the constraining parameter q. The transfer function model

$$Y_t = \frac{0.22\, z^{-3}}{1 - 0.78\, z^{-1}} u_t + D_t \tag{41}$$

is the discrete representation of the continuous first-order plus delay process

$$Y(s) = \frac{e^{-0.5s}}{s+1} u(s)$$

with sampling interval $T = 0.25$. We wish to design a controller to guard against random step-type load disturbances that pass through the process. The disturbance model is therefore of the form

$$D_t = \frac{1}{(1 - 0.78\, z^{-1})\nabla} a_t \tag{42}$$

For the LQ controller a plot of the expected integral square error (ISE) of the process output Y plotted against that of the input changes ∇U for a range of values of the constraining parameter q in Eq. 17 is given in Figure 4. The behavior of the process under minimum variance control ($q = 0$) is shown in Figure 5 for a unit step load disturbance. The open-loop response that would have been followed under no control is also shown. As can be seen, because of the delay the controller cannot affect the output until three periods have passed, and then a dead-beat response is achieved. Although no controller can have a lower ISE (0.48), the excessive response to the manipulated variable is undesirable and the controller is not at all robust to modeling errors. If we increase the constaining factor q, the variance of the manipulated variable is rapidly reduced, Figure 4. The LQ controller is of the form of Eq. 28, that is,

$$(\Gamma - F_1 L) R^{-1} U_t = -F_1 Y_t + F_2 Y_{SP,t}$$

Allowing for the possibility of step changes in set point (i.e., $\nabla Y_{SP,t} = a_t$) and taking $q = 0.5$, this controller is given by

$$[(0.97 - 1.15\, z^{-1} + 0.4\, z^{-2}) - (3.38 - 2.38\, z^{-1})(0.22\, z^{-3})] \cdot (1 - 0.78\, z^{-1})^{-1} U_t = -(3.38 - 2.38 z^{-1}) Y_t - Y_{sp,t}$$

The term in the square bracket on the lefthand side contains $\nabla = (1 - z^{-1})$ and $(1 - 0.78\, z^{-1})$ as factors, and so the final control equation becomes

$$(0.97 + 0.58\, z^{-1} + 0.67\, z^{-2})\nabla U_t = -(3.38 - 2.38\, z^{-1}) Y_t + Y_{sp,t} \tag{43}$$

In Figure 5 the closed-loop response with this controller is also shown for the single load disturbance. Although the ISE has increased to 0.85, the manipulated variable changes much more smoothly and shows very little oscillation.

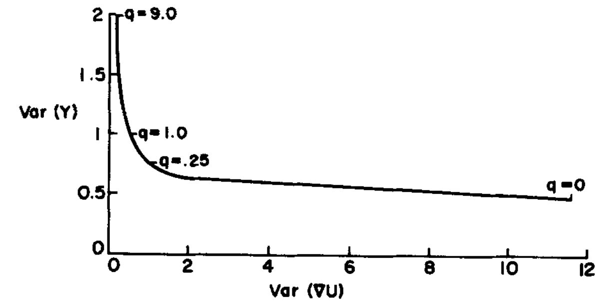

Figure 4. Trade-off between variances or ISE of output and input for LQ control using different constraint parameter values q, SISO example.

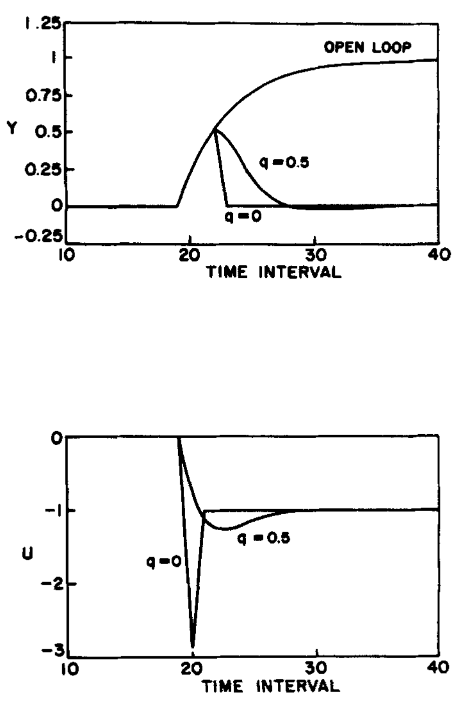

Figure 5. LQ control for a load disturbance, SISO example.

Two-input/one-output example: A sheet forming process

In this section we consider the case where there exist two input variables for the control of one output. Such situations often arise in practice. By making use of a second input variable, the control system designer has much greater flexibility in achieving secondary objectives.

We consider a degenerate example of the control of sheet and film forming processes. A treatment of the full spatial control problem for these processes is given by Bergh and MacGregor (1987b). Consider the special case of a paper-making machine in which it is desired to control the scan average moisture content of the sheet (Y) by manipulating the steam flow to the drying rolls ($-2 \le u_2 \le 2$), and the fractional power input to an infrared heater ($0. \le u_1 \le 1.0$). The process is described by

$$Y_t = -1.0\, u_{1,t-1} - \frac{0.1}{(1 - 0.9\, z^{-1})} u_{2,t-3} + \frac{(1 - 0.7z^{-1})}{\nabla} a_t \qquad (44)$$

Note that the IR heating element is located much closer to the final moisture sensor than the steam rolls. Optimal output performance could be achieved by manipulating u_1 only. However, IR heater power is an expensive variable, and it is desirable to maintain it at as low a value as possible while still having some room for control. Therefore, we consider the problem of minimizing the variance of Y by manipulating u_1 and u_2 while trying to maintain u_1 about a value of 0.25. We accomplish this by minimizing the objective

$$J = \mathrm{Var}\,(Y) + q_1\, \mathrm{Var}\,(u_1 - 0.25) + q_2\, \mathrm{Var}\,(\nabla u_2) \qquad (45)$$

This is accomplished by taking $\bar{u}_1 = (u_1 - 0.25)$ and by choosing $d_1 = 0$, $d_2 = 1$, and $Q_2 = \mathrm{diag}\,(q_1, q_2)$ in the objective function term, Eq. 32. The last term in Eq. 45 penalizes the rate of change of steam flow to the drier rolls.

The model can be expressed in the right matrix fraction form as

$$Y_t = LR^{-1}\underline{U}_t - \theta\nabla^{-1}a_t$$

where

$$L = L_1 z^{-1} + L_3 z^{-3}$$
$$= [-1.0\; 0.]z^{-1} + [0.\; -0.1]z^{-3}$$
$$R = \mathrm{diag}\,[1.0, (1 - 0.9z^{-1})]$$

The controller is given by

$$(\Gamma_0 + \Gamma_1 z^{-1} + \Gamma_2 z^{-2})R^{-1}\underline{U}_t = -\frac{0.3/\sqrt{2}}{1 - 0.7\, z^{-1}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} (Y_t - LR^{-1}\,\underline{U}_t) \qquad (46)$$

We consider two cases.

(a) *No constraint on* ∇u_2 $(q_1 = 1.0,\ q_2 = 0.0)$

In this case the controller is

$$2.0\, \bar{u}_{1,t} + \frac{0.1}{1 - 0.9\, z^{-1}} u_{2,t-2} = \frac{0.3}{(1 - 0.7\, z^{-1})} (Y_t - LR^{-1}\underline{U}_t) \qquad (47)$$

$$\frac{0.1}{1 - 0.9\, z^{-1}} u_{2,t} = \frac{0.3}{(1 - 0.7\, z^{-1})} (Y_t - LR^{-1}\underline{U}_t) \qquad (48)$$

where $\bar{u}_{1,t} = (u_{1t} - 0.25)$

This controller structure is very appealing. The second control equation, Eq. 48, for $u_{2,t}$ is simply the minimum variance controller for controlling Y_t using u_{2t} and is independent of $u_{1,t}$. The first equation, Eq. 47, for $\bar{u}_{1t}$ involves the constrained minimum variance controller for controlling Y_t using $\bar{u}_{1,t}$ with a feedforward term for the effect of changes in $u_{2,t}$ made earlier. In effect, both u_1 and u_2 are being manipulated immediately upon an upset, but then u_1 is gradually reset back to its desired value of 0.25 as the action of $u_{2,t}$ becomes effective two periods later.

The responses of the process output and the input manipulations using this controller are shown in Figure 6a for the stochastic disturbance in Eq. 44. The responses to a unit step disturbance in Y are shown in Figure 6b. Note that in both cases u_1 (IR heater power) is being forced back to 0.25 while the nonstationary drift caused by the disturbance is being compensated for by an eventual shift in the level of u_2 (steam flow). The major problem with this controller is the severe minimum variance control action in u_2. In Figure 6b u_2 actually saturates at the upper limit (+2.0) for a few periods after the step disturbance occurs.

(b) *Constraint on Var* (∇u_2): $(q_1 = 1.0,\ q_2 = 1.0)$

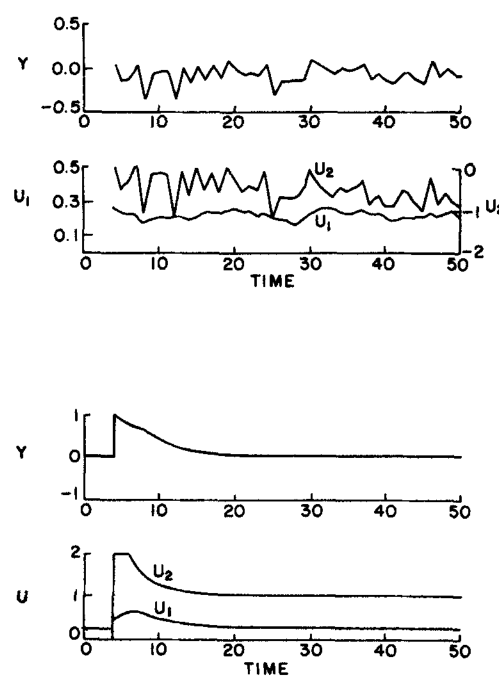

Figure 6. Sheet process: LQ controller ($q_1 = 1$, $q_2 = 0$).
a. Stochastic disturbance
b. Step disturbance

In this case the controller is

$$2.0\,\bar{u}_{1,t} + \frac{0.1}{(1-0.9\,z^{-1})}\,u_{2,t-2} = \frac{0.3}{(1-0.7\,z^{-1})}\,(Y_t - LR^{-1}\,\underline{U}_t) \quad (49)$$

$$\frac{1.64 - 2.64\,z^{-1} + 1.10\,z^{-1}}{1-0.9\,z^{-1}}\,u_{2,t} = \frac{0.3}{(1-0.7\,z^{-1})}\,(Y_t - LR^{-1}\,\underline{U}_t) \quad (50)$$

The first row of the controller, i.e., that for u_1, has remained unchanged and still contains the feedforward compensation for u_2. However, the control action for u_2 is now much more constrained. The responses to the stochastic disturbance and to the step disturbance are shown again for this controller in Figures 7a and 7b. Notice the much smoother behavior in u_2. The performance of the process output has not changed appreciably, but, particularly for the step disturbance, the IR heater variable u_2 has had to move away from 0.25 for a longer period in order to make up for the more constrained action in u_2. The convenient trade-off allowed between the manipulated variables is a feature of using more inputs than outputs.

Level control of tanks in series

Here we revisit an example, considered by Doss et al. (1983), of the control of liquid levels in three tanks in series, Figure 8. Doss et al. considered LQ control via a state space approach, but then resorted to other more intuitive approaches that they felt led to a simpler, more physically acceptable scheme. In this section we treat their problem using LQ control with transfer function models, and show that in the case of minimum variance control ($Q_2 = 0$) the resulting solution is identical to their intuitive solution.

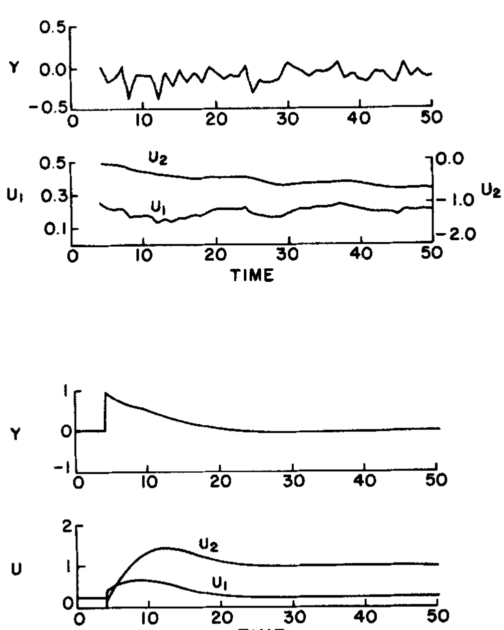

Figure 7. Sheet process: LQ controller ($q_1 = 1$, $q_2 = 1$).
a. Stochastic disturbance
b. Step disturbance

From a material balance the following differential equations can be shown to describe the tank levels Y_i as a function of the inlet and outlet flow rates u_i:

$$\frac{d}{dt}\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{A_t} & -\frac{1}{A_1} & 0 \\ 0 & \frac{1}{A_2} - \frac{1}{A_2} \\ 0 & 0 & \frac{1}{A_3} \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -u_4/A_3 \end{bmatrix} \quad (51)$$

Then the discrete transfer function representation of this system is given by

$$\underline{Y}_t = L\,R^{-1}\,\underline{U}_{t-1} + \underline{D}_t \quad (52)$$

where

$$L = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R = \nabla^{-1} = (1 - z^{-1}) \cdot I$$

and the Y_i's are deviations from set points. In Eq. 52 we have taken all the $\Delta T/A_i$ equal to unity where ΔT is the control interval. In order to design the controller we must also specify the nature of the disturbance D_t or set-point changes that are present. We shall consider first the case where

$$\underline{D}_t = \theta(z^{-1})\,\nabla^{-1}\,\underline{a}_t \quad (53)$$

where $\theta(z^{-1}) = \text{diag}\,\{(1-\theta_1 z^{-1}), (1-\theta_2 z^{-1}), (1-\theta_3 z^{-1})\}$. For $\theta(z^{-1}) = I$, this is equivalent to assuming either of the following:

1. Step set-point changes being called for in any of the levels
2. White noise disturbances in all the flow rates u_i

For $\theta_i > 0$ this allows for some measurement noise in the levels as well.

Assigning equal importance to all levels ($Q_1 = I$), the MIMO controller minimizing the quadratic performance index, Eq. 18, is given by

$$\Gamma\,\nabla^{-1}\,\underline{U}_t = -\,(I - \theta)[\theta(z^{-1})]^{-1}(\underline{Y}_t - L\nabla^{-1}\,\underline{U}_t) \quad (54)$$

For minimum variance control we take $Q_2 = 0$. The spectral factor Γ given by Eq. 27 is then equal to L. Substituting this into Eq. 54 and rearranging, the controller equation simplifies to

$$\underline{U}_t = -\,L^{-1}(I - \theta)\underline{Y}_t \quad (55)$$

Substituting for L and θ reveals that this is exactly the same feedforward structure arrived at by Doss et al. on the basis of intuition, i.e.,

$$u_{3,t} = -\,(1 - \theta_3)Y_{3t} \quad (56)$$

$$u_{2,t} = -\,(1-\theta_2)Y_{2t} - (1-\theta_3)Y_{3t} = -\,(1-\theta_2)Y_{2t} + u_{3t} \quad (57)$$

$$u_{1,t} = -(1-\theta_1)Y_{1t} - (1-\theta_2)Y_{2t} - (1-\theta_3)Y_{3t}$$
$$= -(1-\theta_1)Y_{1t} + u_{2t} \tag{58}$$

As illustrated in Figure 8 this is a feedforward strategy in which all upstream flows are adjusted when a downstream flow is changed. (Note that this minimum variance solution cannot be obtained for continuous time controllers because in continuous time a positive definite Q_2 is required.) The feedforward blocks are simply gains, and the controllers are all proportional in nature. This latter fact arises because, for the step type of level disturbance or set-point changes assumed, the integration in the process dynamics provides the integral action necessary to eliminate any offset. We discuss the consequences of different disturbance structures later. First consider the case where we constrain the variances of the flow rate manipulations by using a nonzero constraint matrix Q_2. Taking

$$Q_2 = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}$$

the spectral factor matrix becomes

$$\Gamma = \begin{bmatrix} 1.08 - 0.08z^{-1} & -1.01 + 0.01z^{-1} & -0.005 + 0.005\,z^{-1} \\ 0.073 - 0.073\,z^{-1} & 1.08 - 0.08\,z^{-1} & -1.01 + 0.01\,z^{-1} \\ 0.069 - 0.069\,z^{-1} & 0.073 - 0.073\,z^{-1} & 1.08 - 0.08\,z^{-1} \end{bmatrix} \tag{59}$$

and the controller, Eq. 54, will no longer have the above simple cascaded feedforward structure. Now all the flow rates will depend upon all of the upstream and downstream levels, as noted by Doss et al. in their state variable approach. However, by comparing the structure of Γ in this case, Eq. 59, with $\Gamma = L$ in the minimum variance case, we can see that the banded diagonal dominance remains, and the controller will still be dominated by the same feedback plus cascaded feedforward structure.

Returning to our comments on the disturbance model, we would like to note that the disturbance assumptions must depend upon the actual physical system. For instance, in this problem if the tank system were to supply process fluid to a downstream unit that had a changing demand, then the flow rate out of the last tank (u_4) would behave more like a random walk or randomly occurring steps. In terms of the disturbance seen by the liquid level in the last tank, Eq. 51, this would be of the nature of an integrated random walk type of disturbance, i.e.,

$$D_{3t} = \frac{(1-\theta z^{-1})}{(1-z^{-1})^2} a_t \tag{60}$$

The other disturbances in the levels of Y_1 and Y_2 would remain as before. The only major difference in the controller designs would be that the level controller for Y_3, Eq. 56, would be of a proportional plus integral form in this case.

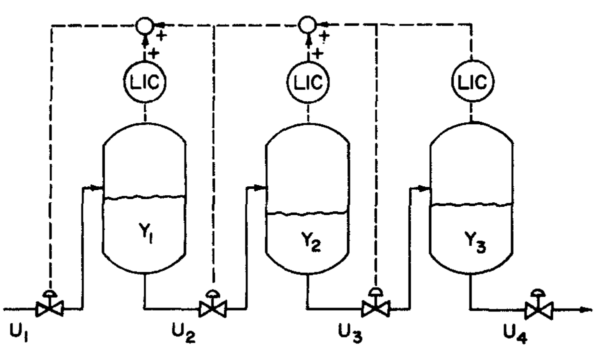

Figure 8. Liquid level process.

A catalytic packed-bed reactor

A tubular pilot plant reactor carrying out highly exothermic butane hydrogenolysis reactions over a nickel on silica gel catalyst is considered here. Previous LQ control studies on an earlier version of this system were performed by Jutan et al. (1977) using state space models derived from theoretical mass and energy balances, and by MacGregor and Wong (1980) using identified state variable models. In this example we show results obtained by Kozub (1986) for a modified reactor system. The reactor outputs to be controlled are the exit conversion of butane and the production rate of propane (an intermediate product). These outputs are measured every 3 min using an on-line gas chromatograph (GC). There is a dead time of one period introduced by the 3 min analysis period of the GC. The manipulated input variables are the hydrogen feed rate to the reactor, and the set point of the reactor temperature controller. This inner loop temperature controller manipulates the butane feed rate and serves to stabilize the reactor (Onderwater et al., 1987). An identification study (Kozub, 1986) led to the following model:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} \dfrac{1.8 - 6.64\,z^{-1} + 5.16\,z^{-2}}{1 - 1.04z^{-1} + 0.34\,z^{-2}} & \dfrac{0.538\,z^{-1}}{1 - 1.26\,z^{-1} + 0.52\,z^{-2}} \\ -0.74 - 3.40\,z^{-1} - 1.01\,z^{-2} & \dfrac{0.16\,z^{-2} + 0.11\,z^{-3}}{1 - 0.61\,z^{-1}} \end{bmatrix} \begin{bmatrix} u_{1,t-2} \\ u_{2,t-2} \end{bmatrix} + \begin{bmatrix} 1 + 0.13z^{-1} & -0.13 \\ 0.21\,z^{-1} & 1 - 0.39\,z^{-1} \end{bmatrix} \cdot \frac{1}{1-z^{-1}} \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} \tag{61}$$

where Y_{1t} = propane production rate − 24, mmol/s

Y_{2t} = butane conversion − 29.6, %

u_{1t} = (hydrogen flow rate − 75) × 10^{-1} mL/s

u_{2t} = hot-spot temperature rise set point − 7, °C

Note that the elements of the transfer function matrix have different periods of delay and some elements are noninvertible. The choice of an approximate model inverse ($\tilde{G}_M^{-1}$) for IMC controller design is far from trivial. However, the spectral factorization inverse, Eq. 26, provides a stable and causal inverse for any choice of the quadratic penalty matrices Q_1, Q_2. The disturbance model is a nonstationary IMA model. The LQ design filter F_1, Eq. 29, for this disturbance structure is given by the first-order filter:

$$F_1 = Q_1^{1/2}(I - \theta)(I - \theta z^{-1})^{-1} \quad (62)$$

where the filter parameter matrix θ is a full matrix because of the coupling between the disturbances in conversion and propane.

Details of the controller designs and the resulting model inverses and filters are given in Kozub (1986). Here we show simulation results for two controller designs. In both designs the matrix Q_1 in the quadratic objective function, Eq. 18, was taken as the identity matrix I in order to attach equal importance to both outputs. In the first simulation, Figure 9, the input penalty matrix Q_2 was chosen to be small, Q_2 = diag (10, 0.1). The regulation and servo control shown in Figure 9 are both good. A single degree of freedom controller, Eq. 25, was used because the models for the stochastic disturbances and the deterministic set-point changes were almost identical. Since the changes in the manipulated hot-spot temperature set point were considered to be too large in this case, the input constraint matrix Q_2 was increased to Q_2 = diag (100, 10). The simulation for this case, shown in Figure 10, reveals that the input manipulations are now much smoother and yet the output performance is almost identical. This controller would be much more robust than the first one.

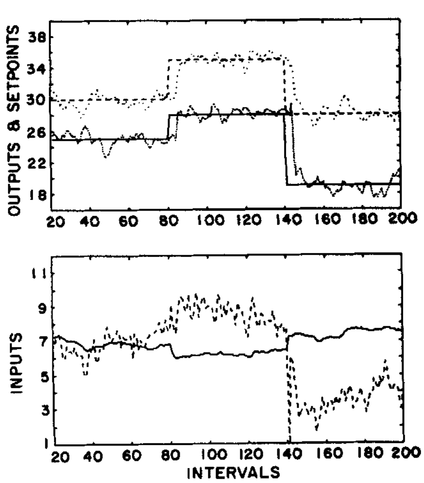

Figure 9. Catalytic reactor: LQ with Q_2 = diag (10, 0.1).
a. Outputs: — propane prod., --- % conversion
b. Inputs: — H_2 flow, --- hot spot set point

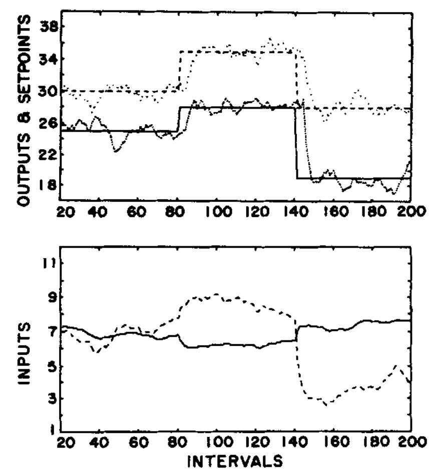

Figure 10. Catalytic reactor: LQ with Q_2 = diag (100, 10).
a. Outputs: — propane prod., --- % conversion
b. Inputs: — H_2 flow, --- hot spot set point

The reader is referred to Kozub et al. (1987) for details of the experimental results of applying these multivariable LQ controllers to the actual nonlinear pilot plant reactor over a range of operating conditions. The results of the actual implementations were very comparable to those of the simulations. Comparisons were also made with other multivariable IMC designs.

Summary

A design procedure for discrete multivariable LQ output controllers based on process transfer function and disturbance models has been presented. The method treats the problem both of disturbance rejection and of set-point tracking for either stochastic or deterministic disturbances, and it easily handles nonsquare systems. Since the design is carried out using input/output models directly, there is no need to construct state observers, and the structure of the controllers is much more evident. In particular, the controllers are shown to be internal model controllers (IMC) for specified disturbance structures. In the Wiener-Hopf design procedure, the solution of a spectral factorization algorithm is shown to lead easily to optimal multivariable inverses even for complex systems, and the solution of a Diophantine equation is shown to lead to stable, causal filters. The design shows that decoupling is rarely optimal (in the sense of the quadratic performance criterion), and when it is, it usually corresponds to a multivariable dead-beat controller. The role that the input constraining matrix Q_2 of the quadratic performance index has in effecting a trade-off between output performance and control effort, and in improving the robustness of these controllers, is clearly shown.

In summary, this optimal LQ design approach in the input/output space not only provides a very general and powerful multivariable controller design procedure, it also provides a means of understanding and unifying many of the concepts inherent in multivariable controller design.

Acknowledgment

The authors thank Derrick Kozub for granting permission to show some of his simulation results on the LQ control of a packed-bed reactor.

The authors also acknowledge financial support for this research from the Natural Sciences and Engineering Research Council of Canada (JFM), and from the Queen's University Advisory Research Council (TJH).

Notation

A = denominator of process transfer function model
B = numerator of process transfer function model
b = periods of time delay in process transfer function model
C = feedback controller transfer function
D_t = disturbance in process output
d = degree of nonstationarity of disturbance, Eqs. 12, 13
E = expectation operator
e_t = output deviation from set point
F = filter transfer function, Eqs. 23, 24, 29
G_M = transfer function model for process
H_1, H_2 = rational polynomial matrices in LQ controller, Eq. 22
J = quadratic cost function
ℓ = order of the polynomial matrix L
L = polynomial matrix in right matrix fraction representation of G_M, Eq. 4
m = dimension of input vector
n = dimension of output vector
P = polynomial matrix in Diophantine equation
Q_1 = output penalty matrix in quadratic cost function
Q_2 = input penalty matrix in quadratic cost function
q = input penalty constant in quadratic cost function, Eq. 17
R = diagonal polynomial matrix in right matrix fraction representation of G_M, Eq. 4
S = sensitivity matrix, Eq. 33
T = polynomial matrix in filter F, Eq. 29
tr = trace of a matrix
U = vector of manipulated variables
u_i = elements of $\underline{U}$
W = weighting matrices in sensitivity norms
Y = vector of output or controlled variables
z^{-1} = backward shift operator or complex variable, depending upon the sense of the equation

Greek letters

Γ = matrix polynomial (solution to spectral factorization Eq. 27)
∇ = diagonal matrix $(1 - z^{-1}).I$
$\Phi(z^{-1})$ = diagonal polynomial matrix in right matrix fraction description of disturbance model, Eq. 13
Φ = parameter in $\Phi(z^{-1})$
$\theta(z^{-1})$ = polynomial matrix in right matrix fraction description of disturbance model, Eq. 13
θ = parameter in $\theta(z^{-1})$
$\oint_{|z|=1}$ = integral around unit circle in complex plane

Superscripts

$*$ = complex conjugate transpose
$\hat{}$ = prediction
T = transpose

Subscripts

M = model
sp = set point
t = discrete time period

Literature Cited

Astrom, K. J. *Introduction to Stochastic Control*, Academic Press, New York (1970).
Astrom, K. J., and B. Wittenmark, *Computer Controlled Systems*, Prentice-Hall, Englewood Cliffs, NJ.
Bergh, L. G., and J. F. MacGregor, "Spatial Control of Sheet and Film Forming Processes," Can. J. Chem. Eng., *65*, 148 (1987a).
———, "Constrained Minimum Variance Controllers: Internal Model Structure and Robustness Properties," *Ind. Eng. Chem.* (1987b).
Box, G. E. P., and G. M. Jenkins, *Times Series Analysis, Forecasting and Control*, Holden-Day, San Francisco (1970).
Brosilow, C. B., "The Structure and Design of Smith Predictors from the Viewpoint of Inferential Control," *Proc. Joint. Auto. Control Conf.*, Denver (1979).
Callier, and C. A. Desoer, *Multivariable Feedback Systems*, Springer-Verlag, New York (1982).
Chang, S. L., *Synthesis of Optimum Control Systems*, McGraw-Hill, New York (1961).
Cutler, C. R., and B. L. Ramaker, "Dynamic Matrix Control—A Computer Control Algorithm," 86th AIChE Meet. (Apr., 1976).
Davis, J. H., and R. G. Dickinson, "Spectral Factorization by Optimal Gain," *SIAM J. Appl. Math.*, **43,** 289 (1983).
Desoer, C. A., and M. J. Chen, "Design of Multivariable Feedback Systems with Stable Plant," IEEE **AC-26,** 408 (1981).
Doss, J. E., T. W. Doub, J. J. Downs, and E. F. Vogel, "New Directions for Process Control in the Eighties," AIChE Meet., Washington, DC (Nov., 1983).
Doyle, J. C., and G. Stein, "Multivariable Feedback Design: Concepts for a Classical/Modern Approach," IEEE, **AC-26,** 4(1981).
Francis, B. A., "On the Wiener-Hopf Approach to Optimal Feedback Design," *Syst. Contl. Lett.*, **2,** 197 (1982).
Garcia, C. E., and M. Morari, "Internal Model Control. 1: A Unifying Review and Some New Results," *Ind. Eng. Chem. Process Des. Dev.*, **21,** 308 (1982).
———, "Internal Model Control. 2: Design Procedure for Multivariable Systems," *Ind. Chem. Process Des. Dev.*, **24,** 472 (1985).
Grimble, M. J., "Solution of the Discrete Time Stochastic Optimal Control Problem in the Z-Domain," *Int. J. Syst. Sci.*, **10,** 1369 (1979).
———, "Optimal H_∞ Robustness and the Relationship to LQG Design Problems," *Int. J. Control*, **43,** 351 (1986).
Grosdidier, P., Morari, M., Holt, B. R., "Closed Loop Properties from Steady-State Gain Information," Ind. Eng. Chem. Fundam., *24*, 221–235 (1985).
Harris, T. J., J. F. MacGregor, and J. D. Wright, "Self-Tuning and Adaptive Controllers: An Application to Catalytic Reactor Control," *Proc. Joint. Auto. Control Conf., Philadelphia*, 41 (1978); *Technometrics*, **22,** 153 (1980).
Jezek, J., and V., Kucera, "Efficient Algorithm for Spectral Factorization," *Automatica*, **21,** 663 (1985).
Johnson, C. D., "Control of Dynamical Systems," *Stochastic Problems in Mechanics*, Univ. Waterloo, Ontario, Canada, 81 (1972).
Jutan, A., J. F. MacGregor, and J. D. Wright, "Multivariable Computer Control of a Butane Hydrogenolysis Reactor I, II, III," *AIChE J.*, **23,** 732 (1977).
Kalman, R. E., "Contributions to the Theory of Optimal Control," *Bol. Soc. Mat. Mexicana*, **5,** 102 (1960).
Kozub, D. J., "Advanced Multivariable Controller Design and Application to Packed-Bed Reactor Control," M. Eng. Thesis, Ch.E. Dept., McMaster Univ., Hamilton Ontario (1986).
Kozub, D. J., J. F. MacGregor, and J. D. Wright, "Application of LQ and IMC Controllers to Packed-Bed Reactor," *AIChE J.* (1987).
Kucera, V. *Discrete Linear Control: The Polynomial Equation Approach*, Wiley, New York (1979).
Kwakernaak, H., "Uncertainty Models and the Design of Robust Control Systems," *Proc. Int. Sem. Uncertainty and Control, Bonn, FDR, May, 1985*, J. Ackermann, ed., DFLVR, **70,** 84 (1985).
Kwakernaak, H., and R. Sivan, *Linear Optimal Control Systems*, Wiley, New York (1972).
Lehtomanki, N. A., D. A. Castanon, B. C. Leoy, C. Stein, N. R. Sandell, and M. Athans, "Robustness and Modeling Error Characterization," IEEE **AC-29,** 212 (1985).
MacGregor, J. F., "Optimal Discrete Stochastic Control Theory for Process Application," *Can. J. Chem. Eng.*, **51,** 468 (1973).
MacGregor, J. F. and A. K. L. Wong, "Multivariate Model Identification and Stochastic Control of a Chemical Reactor," *Technometrics*, **22,** 453 (1980).
MacGregor, J. F., T. J. Harris, and J. D. Wright, "Duality Between the Control of Processes Subject to Randomly Occurring Deterministic Disturbances and ARIMA Stochastic Disturbances," *Technometrics*, **26,** 389 (1984).
Morari, M., and J. C. Doyle, "A Unifying Framework for Control System Design Under Uncertainty and Its Applications for Chemical

Process Control," *Proc. 3rd. Chem. Proc. Control Conf.*, Asilomar (Jan., 1986).

Newton, G. C., L. A. Gould, and J. F. Kaiser, *Analytical Design of Linear Feedback Controls*, Wiley, New York (1957).

Onderwater, D., J. F. MacGregor, and J. D. Wright, "Temperature Control of a Tubular Reactor Using Self-Tuning Regulators with Nonlinear Transformations," *Can. J. Chem. Eng.*, **65** (1987).

Palmor, Z. J., and R. Shinnar, "Design of Sampled Data Controllers," *Ind. Eng. Chem. Process Des. Dev.*, **18**, 8 (1979).

Prett, D. M., and R. D. Gillette, "Optimization and Constrained Multivariable Control of a Catalytic Cracking Unit," AIChE 86th Ann. Meet., (Apr., 1979).

Richlalet, J. A., A. Rault, J. D. Testud, J. Papon, "Model Predictive Heuristic Control: Applications to Industrial Processes," *Automatica*, **14**, 413 (1978).

Rohrs, C., G. Stein, K. J. Astrom, "Uncertainty in Sampled Systems," *Proc. Am. Control Conf.*, Boston, 95 (1985).

Segall, N. L., J. F. MacGregor, and J. D. Wright, "One-Step Optimal Correction for Input Saturation in Discrete Model Based Controllers," Tech. Rept. Process Control Lab., Dept. Chem. Eng., McMaster University, Hamilton, Ontario, Canada (1987).

Wellons, M. C., and T. F. Edgar, "The Generalized Analytical Predictor. I: SISO Systems," Tech. Rep., Ch.E. Dept. Univ. of Texas (1986).

Wilson, G. T., "Modelling Linear Systems for Multivariable Control," Ph.D. Thesis, Univ. Lancaster, England (1970).

———, "The Factorization of Matricial Spectral Densities," *SIAM J. Appl. Math.*, **23**, 420 (1972).

Wolovich, W. A., and H. Elliot, "Discrete Models for Linear Multivariable Systems," *Int. J. Control*, **38**, 337 (1983).

Youla, D. C., and J. J. Bongiorno, "A Feedback Theory of Two-Degree-of-Freedom Optimal Wiener-Hopf Deisgn," IEEE, **AC-30**, 652 (1985).

Youla, D. C., H. A. Jabr, J. J. Bongiorno, "Modern Wiener-Hopf Design of Optimal Controllers. II: The Multivariable Case," *IEEE*, **AC-21**, 319 (1976).

Zames, C., "Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms and Approximate Inverses," IEEE, **AC-26**, 301 (1981).

Manuscript received June 26, 1986, and revision received Mar. 30, 1987.

See NAPS document no. 04532 for 8 pages of supplementary material. Order from NAPS c/o Microfiche Publications, P.O. Box 3513, Grand Central Station, New York, NY 10163. Remit in advance in U.S. funds only $7.75 for photocopies or $4.00 for microfiche. Outside the U.S. and Canada, add postage of $4.50 for the first 20 pages and $1.00 for each of 10 pages of material thereafter, $1.50 for microfiche postage.