A Tuning Strategy for Unconstrained SISO Model Predictive Control

Rahul Shridhar and Douglas J. Cooper*

Chemical Engineering Department, University of Connecticut, U-222, Storrs, Connecticut 06269-3222

This paper presents an easy-to-use and reliable tuning strategy for unconstrained SISO dynamic matrix control (DMC) and lays a foundation for extension to multivariable systems. The tuning strategy achieves set point tracking with minimal overshoot and modest manipulated input move sizes and is applicable to a broad class of open loop stable processes. The derivation of an analytical expression for the move suppression coefficient, λ, and its demonstration in a DMC tuning strategy is one of the significant contributions of this work. The compact form for the analytical expression for λ is achieved by employing a first order plus dead time (FOPDT) model approximation of the process dynamics. With tuning parameters computed, DMC is then implemented in the classical fashion using a dynamic matrix formulated from step response coefficients of the actual process. Just as the FOPDT model approximation has proved a valuable tool in tuning rules such as Cohen–Coon, ITAE, and IAE for PID implementations, the tuning strategy presented here is significant because it offers an analogous approach for DMC.

Introduction

This paper details and demonstrates a tuning strategy for single-input single-output (SISO), unconstrained dynamic matrix control (DMC) (Cutler and Ramaker, 1980, 1983) that is applicable to a wide range of open-loop stable processes. The DMC control law is given by

$$\Delta \bar{u} = (A^{\mathrm{T}}A + \lambda \mathbf{I})^{-1} A^{\mathrm{T}} \bar{e} \qquad (1)$$

where A is the dynamic matrix, $\bar{e}$ is the vector of predicted errors over the next P sampling instants (prediction horizon), λ is the move suppression coefficient, and $\Delta \bar{u}$ is the manipulated input profile computed for the next M sampling instants, also called the control horizon. The $A^{\mathrm{T}}A$ matrix, to be inverted in the evaluation of the DMC control law, is referred to in this work as the system matrix.

Implementation of DMC with a control horizon greater than one manipulated input move necessitates the inclusion of a move suppression coefficient, λ (Maurath et al., 1988a). This coefficient serves a dual purpose of conditioning the system matrix before inversion and suppressing otherwise aggressive control action. It is often used as the primary adjustable parameter to fine-tune DMC to desirable performance. Though the dual effect of λ has been discussed by others (e.g., Ogunnaike (1986)), past researchers have focused most attention on the latter effect in the selection of λ.

This paper designs an analytical expression that computes an appropriate λ by recognizing and exploiting the strong correlation between the condition number of the system matrix, $A^{\mathrm{T}}A$, and resultant manipulated input move sizes. The move suppression coefficient, λ, which effectively modifies the condition number of the system matrix, is computed such that the condition number is always bounded by a fixed low value. By holding the condition number at a low value, desirable DMC closed-loop performance is achieved while the manipulated input move sizes are prevented from becoming excessive.

The derivation of an analytical expression that computes λ and its demonstration in a DMC tuning strategy is one of the significant contributions of this work. Table 1 illustrates the step-by-step procedure to compute the DMC tuning parameters, including the analytical expression that computes λ, based on a user-specified control horizon and a given sample time.

Derivation of the analytical expression for λ considers a first order plus dead time (FOPDT) approximation of the process dynamics. It must be emphasized that the FOPDT approximation is employed only in the derivation of the analytical expression for λ. The examples presented later in this work all use the traditional DMC step response matrix of the actual process upon implementation.

The primary benefit of a FOPDT model approximation is that it permits derivation of a compact analytical expression for computing λ. Although a FOPDT model approximation does not capture all the features of some higher-order processes, it often reasonably describes the process gain, overall time constant, and effective dead time of such processes (Cohen and Coon, 1953). In the past, tuning strategies based on a FOPDT model such as Cohen–Coon, IAE, and ITAE have proved useful for PID implementations. The tuning strategy presented here is significant because it offers an analogous approach for DMC.

The theoretical details in this paper are organized as follows: (i) the DMC transfer function form is derived and from it the existence of a gain-scaled move suppression coefficient is established; (ii) an approximation of the scaled system matrix, $\bar{A}^{\mathrm{T}}\bar{A}$, is obtained from the DMC control law; (iii) formulas for the maximum and minimum eigenvalues of the gain-scaled $(A^{\mathrm{T}}A + \lambda \mathbf{I})$ matrix are developed and an expression for its condition number is obtained; (iv) based on this expression for the condition number, an analytical expression for computing the move suppression coefficient is derived and an overall DMC tuning strategy is formulated; (v) a straightforward extension of this tuning strategy to multivariable systems is proposed; and (vi) issues important to the implementation of the tuning strategy are discussed with guidelines for selection of the sample time and the control horizon.

Through demonstrations of several simulation ex-

---

* Author to whom correspondence should be addressed. Phone: (860) 486-4092. Fax: (860) 486-2959. Email: cooper@eng2.uconn.edu.

Table 1. DMC Tuning Strategy

1. Approximate the process dynamics with a first order plus dead time (FOPDT) model:

$$\frac{y(s)}{u(s)} = \frac{K_p e^{-\theta_p s}}{\tau_p s + 1}$$

2. It is desirable but not necessary to select a value for the sampling interval, T. If possible, select T as the largest value that satisfies

$$T \le 0.1\tau_p \quad \text{and} \quad T \le 0.5\theta_p$$

3. Calculate the discrete dead time (rounded to the next integer):

$$k = \theta_p/T + 1$$

4. Calculate the prediction horizon and the model horizon as the process settling time in samples (rounded to the next integer):

$$P = N = 5\tau_p/T + k$$

5. Select the control horizon, M (integer, usually from 1 to 6) and calculate the move suppression coefficient:

$$f = \begin{cases} 0 & M = 1 \\ \dfrac{M}{500}\left(\dfrac{3.5\tau_p}{T} + 2 - \dfrac{(M-1)}{2}\right) & M > 1 \end{cases}$$

$$\lambda = fK_p^2$$

6. Implement DMC using the traditional step response matrix of the actual process and the following parameters computed in steps 1−5:

> sample time, T
> model horizon (process settling time in samples), N
> prediction horizon (optimization horizon), P
> control horizon (number of moves), M
> move suppression coefficient, λ

amples that encompass a range of process characteristics, the tuning strategy is validated for set point tracking performance using different choices of the control horizon and sampling rates. The tuning strategy is further validated for disturbance rejection. An example application that validates an extension to the tuning strategy applied to multivariable systems is also presented.

Background

Model predictive control (MPC) has established itself over the past decade as an industrially important form of advanced control. Since the seminal publication of Model Predictive Heuristic Control (later Model Algorithmic Control) (Richalet et al., 1976, 1978) and Dynamic Matrix Control (Cutler and Ramaker, 1980, 1983), MPC has gained widespread acceptance in academia and in industry. Several excellent technical reviews of MPC recount the significant contributions in the past decade and detail the role of MPC from an academic perspective (García et al., 1989; Morari and Lee, 1991; Ricker, 1991; Muske and Rawlings, 1993) and from an industrial perspective (Richalet, 1993; Clarke, 1994; Froisy, 1994; Camacho and Bordons, 1995; Qin and Badgwell, 1996).

MPC refers to a family of controllers that employs a distinctly identifiable model of the process to predict its future behavior over an extended prediction horizon. A performance objective to be minimized is defined over the prediction horizon, usually as a sum of quadratic set point tracking error and control effort terms. This cost function is minimized by evaluating a profile of manipulated input moves to be implemented at successive sampling instants over the control horizon. Closed-loop optimal feedback is achieved by implementing only the first manipulated input move and repeating the complete sequence of steps at the subsequent sample time. This "moving horizon" concept of MPC, where the controller looks a finite time into the future, is illustrated in Figure 1.
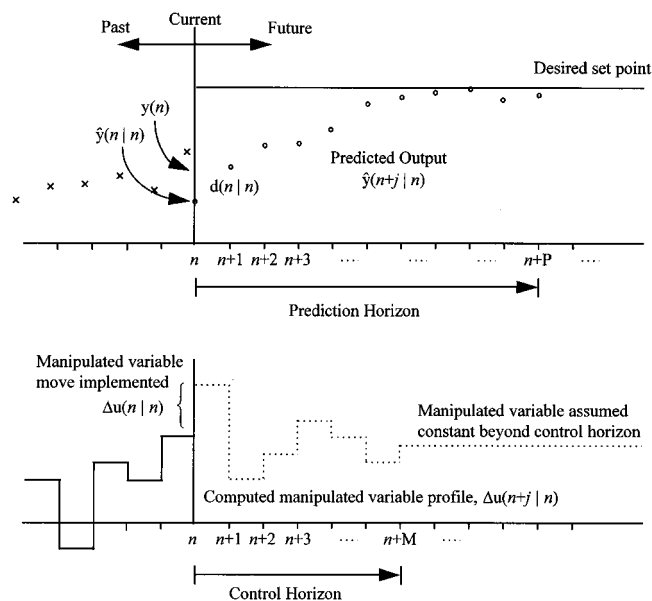
Figure 1. Moving horizon concept of model predictive control.

Dynamic matrix control is arguably the most popular MPC algorithm currently used in the chemical process industry. Qin and Badgwell (1996) reported about 600 successful applications of DMC. It is not surprising why DMC, one of the earliest formulations of MPC, represents the industry's standard today. A large part of DMC's appeal is drawn from an intuitive use of a finite step response (or convolution) model of the process, a quadratic performance objective over a finite prediction horizon, and optimal manipulated input moves computed as the solution to a least squares problem. Because of its popularity, this work focuses on an overall tuning strategy for DMC.

Another form of MPC that has rapidly gained acceptance in the control community is Generalized Predictive Control (GPC) (Clarke et al., 1987a,b). It differs from DMC in that it employs a controlled autoregressive and integrated moving average (CA-

RIMA) model of the process which allows a rigorous mathematical treatment of the predictive control paradigm. The GPC performance objective is very similar to that of DMC but is minimized via recursion on the diophantine identity (Clarke et al., 1987a,b; Lalonde and Cooper, 1989). Nevertheless, GPC reduces to the DMC algorithm when the weighting polynomial that modifies the predicted output trajectory is assumed to be unity (McIntosh et al., 1991). Therefore, without any loss of generality, the tuning strategy proposed in this paper is directly applicable to GPC.

Unconstrained SISO DMC. DMC does not always compete with, but sometimes complements, classical three-term PID (proportional, integral, derivative) controllers. That is, it is often implemented in advanced industrial control applications embedded in a hierarchy of control functions above a set of traditional PID loops (Prett and García, 1988; Qin and Badgwell, 1996).

The unconstrained SISO DMC formulation considered in this work (eq 1) does not unleash the full power of MPC. This restricted form of DMC does not allow multivariable control while satisfying multiple process and performance objectives. However, the analysis presented here provides a foundation upon which more advanced tuning strategies may be developed, and this is illustrated later in the work with an extension to a tuning strategy for multivariable DMC.

In any event, unconstrained SISO DMC does offer some useful capabilities. For example, past comparison studies between unconstrained DMC and traditional PI control (e.g., Farrell and Polli, 1990) show that DMC provides superior performance when disturbance tuning differs significantly from servo tuning. DMC has also demonstrated superior performance in the case of plant-model mismatch, except for process gain mismatch.

Additionally, incorporation of process knowledge in the controller architecture provides DMC with anticipatory capabilities and facilitates control of processes with nonminimum phase behavior and large dead times. The form of the performance objective provides a convenient way to balance set point tracking with control effort, leading to an intuitive tradeoff between performance and robustness. Also, the DMC form considered in this work allows the introduction of feed-forward control in a natural way to compensate for measurable disturbances.

Challenges in SISO DMC Tuning. Tuning of unconstrained and constrained DMC for SISO and multivariable systems has been addressed by an array of researchers. In the past, systematic trial-and-error tuning procedures have been proposed (e.g., Cutler, 1983; Ricker, 1991). Marchetti et al. (1983) presented a detailed sensitivity analysis of adjustable parameters and their effects on DMC performance. The method of principal component selection was presented by Maurath et al. (1985, 1988b) as a method to compute an appropriate prediction horizon and a move suppression coefficient (Callaghan and Lee, 1988). To simplify DMC tuning, Maurath et al. (1988a) also proposed the "$M = 1$" controller configuration of DMC.

Other tuning strategies for DMC have concentrated on specific aspects such as tuning for stability (García and Morari, 1982; McIntosh, 1988; Clarke and Scattolini, 1991; Rawlings and Muske, 1993), robustness (Ohshima et al., 1991; Lee and Yu, 1994), and performance (McIntosh et al., 1992; Hinde and Cooper, 1993, 1995). Although some of the above methods provide a complete design of DMC, they also require fairly so-

phisticated analysis tools and an advanced knowledge of control concepts for their implementation. Hence, there still exists a need for easy-to-use tuning strategies for DMC.

Tuning of unconstrained SISO DMC is challenging because of the number of adjustable parameters that affect closed-loop performance. These include the following: a finite prediction horizon, P; a control horizon, M; a move suppression coefficient, λ; a model horizon, N; and a sample time, T.

The first problem that needs to be addressed is the selection of an appropriate set of tuning parameters from among those available for DMC. Practical limitations often restrict the availability of sample time, T, as a tuning parameter (Franklin and Powell, 1980; Åström and Wittenmark, 1984). The model horizon is also not an appropriate tuning parameter since truncation of the model horizon, N, misrepresents the effect of past moves in the predicted output and leads to unpredictable closed-loop performance (Lundström et al., 1995).

Sensitivity Study and Final Selection of DMC Tuning Parameters. Based on the above discussion, candidate parameters for developing a systematic tuning strategy for DMC include the prediction horizon, P, the control horizon, M, and the move suppression coefficient, λ. Though this simplifies the task of sensitivity analysis, the appropriate choice of these parameters is strongly dependent on the sample time and the nature of the process.

Over the past decade, detailed studies of DMC parameters have provided a wealth of information about their qualitative effects on closed-loop performance (Marchetti et al., 1983; García and Morshedi, 1986; Downs et al., 1988; Maurath et al., 1988a). In this section, a brief sensitivity study investigates the extent to which various parameters affect DMC performance. This study is targeted toward selection of appropriate tuning parameters for developing a DMC tuning strategy.

A base case process is employed to illustrate the effect of adjustable parameters on DMC response for a step change in set point (Figures 2–4). The base case process has the transfer function

process 1

$$G_p(s) = \frac{e^{-50s}}{(150s + 1)(25s + 1)} \tag{2}$$

Figure 2 illustrates the importance of λ in tuning DMC with a control horizon greater than one manipulated input move. The sample time is selected for this base case such that the T/τ_p ratio is 0.1. This sample time to overall time constant ratio lies within the recommended range for digital controllers (Seborg, 1986; Ogunnaike, 1994). A FOPDT model approximation of the process described by eq 2 yields a process gain, $K_p = 1$, an overall time constant, $\tau_p = 157$, and effective dead time, $\theta_p = 70$. The prediction horizon and model horizon are fixed at 54 and represent the complete settling time of process 1 in samples.

Graph a of Figure 2 illustrates the response to a step change in set point when the control horizon is 1 ($M = 1$) and no move suppression is used ($\lambda = 0$). Note that for $M = 1$, the set point step response is sluggish and a $\lambda > 0$ will only further slow the process response.

$$G_p(s) = \frac{e^{-50s}}{(150s+1)(25s+1)}$$

FOPDT fit :
$K_p = 1$
$\tau_p = 157$
$\theta_p = 70$
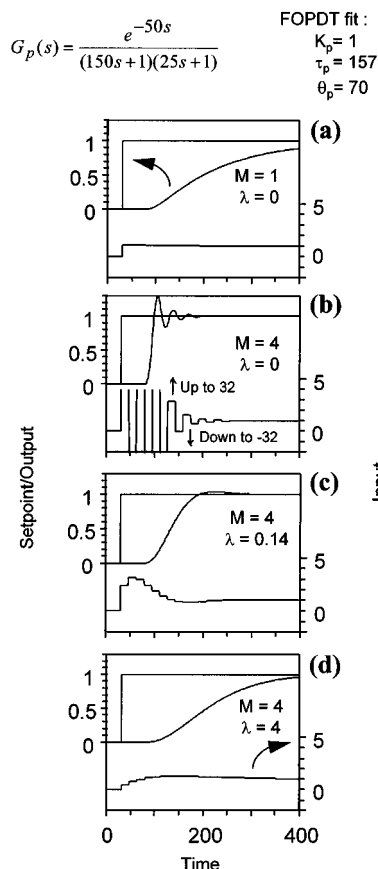
**Figure 2.** Importance of the move suppression coefficient, λ, in tuning DMC with the control horizon, M, greater than one manipulated input move ($T = 16$, $T/\tau_p = 0.1$, $P = 54$).

Graphs b–d show the impact of λ on performance for a control horizon of $M = 4$. Graph b demonstrates that, with $M > 1$, the lack of move suppression results in dramatically aggressive control effort and a significantly underdamped measured output response. Graph c shows that an intermediate response can be achieved by an appropriate choice of λ. Graph d shows that a larger move suppression coefficient ($\lambda = 4.0$) results in a slower response. Further increasing λ can lead to an undesirable sluggish response for most applications. Consequently, this study shows that the choice of λ is critical to the performance achieved by DMC.

Figure 3 demonstrates the interdependence of prediction horizon, P, and sample time, T. A matrix of closed-loop response results is generated for different choices of P and T while maintaining the control horizon, M, and move suppression coefficient, λ, constant. When the prediction horizon is altered, the model horizon, N, is always kept long enough to ensure that the DMC step response model correctly predicts the steady state. By eliminating the truncation errors that result when the model horizon does not account for the process steady state, Figure 3 isolates the effect of prediction horizon, P, and sample time, T, on DMC performance.

The trend in the Figure 3 graphs h → e → b shows that by reducing prediction horizon while maintaining a constant sample time ($T/\tau_p = 0.1$), the output response becomes increasingly underdamped. Also, the trend in graphs f → e → d illustrates that by decreasing the sample time for a constant prediction horizon ($P = 9$), the output response becomes increasingly underdamped. Thus, Figure 3 shows that the choice of P and T are interrelated.

Graphs f, h, and i in Figure 3 show that a large
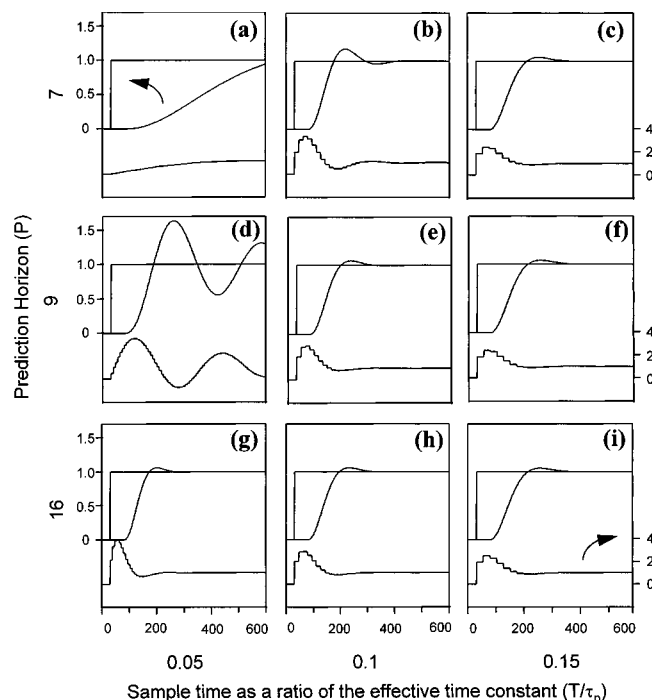
Figure 3. Effect of prediction horizon, P, and sample time, T, on DMC performance for process 1 ($M = 4$, $\lambda = 0.14$).

prediction horizon ($P = 16$ with $T/\tau_p = 0.1$) or a large sample time ($T/\tau_p = 0.15$ with $P = 9$) does not improve closed-loop performance significantly. However, it has been shown by past researchers (García and Morari, 1982; Clarke and Scattolini, 1991; Rawlings and Muske, 1993; Scokaert and Clarke, 1994) that a larger prediction horizon improves nominal stability of the closed loop. For this reason, the prediction horizon should be selected such that it includes the steady state effect of all past computed manipulated input moves. Therefore, the prediction horizon cannot be used as the primary DMC tuning parameter.

Figure 4 illustrates the effect of control horizon, M, for fixed $P = 54$, $T/\tau_p = 0.15$, and $\lambda = 0.14$. Graphs a–c show that increasing the control horizon from 2 to 6 does not alter the performance significantly. Actually, the only noticeable effect is a slight increase in overshoot for a larger control horizon. This is due to the additional degree of freedom from a larger control horizon. This allows more aggressive initial moves that are later compensated for by the extra moves available.

Rawlings and Muske (1993) have shown that a necessary condition to ensure nominal stability of infinite horizon MPC is that the control horizon should be greater than or equal to the number of unstable modes of the system. Nevertheless, it is clear from this study that control horizon, due to its negligible effect on closed-loop performance, is not well suited as the primary DMC tuning parameter.

A conclusion from the above discussion is that the choice of prediction horizon, P, cannot be made independent of the sample time, T. For stability reasons, P must be selected such that it includes the steady state effect of all past manipulated input moves; i.e., it should equal the open-loop settling time of the process in samples. Stability considerations also restrict the choice of control horizon, M. Besides, the control horizon does not have a significant impact on closed-loop performance in the presence of move suppression. Therefore, this brief study supports the opinions of other researchers

$$G_p(s) = \frac{e^{-50s}}{(150s+1)(25s+1)}$$

FOPDT fit :
$K_p = 1$
$\tau_p = 157$
$\theta_p = 70$
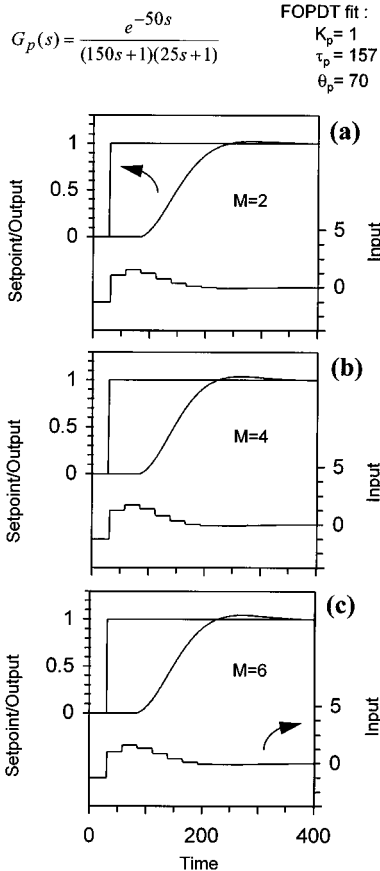
Figure 4. Effect of the control horizon, M, on DMC performance for process 1 ($T = 16$, $T/\tau_p = 0.1$, $P = 54$, $\lambda = 0.14$).

that the one candidate best suited as the final DMC tuning parameter is the move suppression coefficient, λ.

Analytical Framework

In this section, the form of the DMC transfer function is derived from the DMC control law. The difficulty in using this form for development of an overall DMC tuning strategy is highlighted. It is then shown how some useful information relevant to the selection of move suppression coefficient, λ, can be extracted from the form of the DMC transfer function.

DMC Control Law. The cornerstone of the DMC algorithm is a discrete convolution or step response model of the process that predicts the output ($\hat{y}(n + j)$) j sampling instants ahead of the current time instant, n:

$$\hat{y}(n + j) = y_0 + \underbrace{\sum_{i=1}^{j} a_i \Delta u(n + j - i)}_{\substack{\text{effect of current and} \\ \text{future moves}}} + \underbrace{\sum_{i=j+1}^{N-1} a_i \Delta u(n + j - i)}_{\text{effect of past moves}} \quad (3)$$

In eq 3, y_0 is the initial condition of the measured output, $\Delta u_i = u_i - u_{i-1}$ is the change in the manipulated input at the ith sampling instant, a_i is the ith unit step

response coefficient of the process, and N is the process settling time in samples.

The current and future manipulated input moves are yet undetermined and are not used in the computation of the predicted output profile. Therefore, eq 3 reduces to

$$\hat{y}(n + j) = y_0 + \sum_{i=j+1}^{N-1} (a_i \Delta u(n + j - i)) + d(n + j) \quad (4)$$

where the term $d(n + j)$ lumps together possible unmeasured disturbances and inaccuracies due to plant-model mismatch. From eq 4, the predicted output at the current instant ($j = 0$) is

$$\hat{y}(n) = y_0 + \sum_{i=1}^{N-1} (a_i \Delta u(n - i)) + d(n) \quad (5)$$

Since the future values of $d(n + j)$ are not available, an estimate is used. In the absence of any additional knowledge of $d(n + j)$ over future sampling instants, the predicted disturbance is assumed to be equal to that estimated at the current time instant. Therefore

$$d(n + j) = d(n) = y(n) - y_0 - \sum_{i=1}^{N-1} (a_i \Delta u(n - i)) \quad (6)$$

where $y(n)$ is the measurement of the current output.

A more accurate estimate of the $d(n + j)$ is possible, provided the load disturbance is measured and a reliable load disturbance to measured output model is available. Using the unit step response coefficients from this model, an equation similar to eq 5 above can be used to predict the future disturbance (Asbjornsen, 1984; Muske and Rawlings, 1993). A DMC configuration that uses a process model augmented with a disturbance to output model is known as feed-forward DMC.

Now, if the predicted output is to follow the set point in the wake of a set point change or unmeasured disturbance, then the current and future manipulated input moves in eq 3 have to be determined such that

$$y_{sp}(n + j) - \hat{y}(n + j) = 0 \qquad j = 1, 2, ..., P \quad (7)$$

Substituting eq 3 in eq 7 gives

$$\underbrace{y_{sp}(n + j) - y_0 - \sum_{i=j+1}^{N-1} a_i \Delta u(n + j - i) - d(n)}_{\text{predicted error based on past moves, } e(n + j)} = \underbrace{\sum_{i=1}^{j} a_i \Delta u(n + j - i)}_{\substack{\text{effect of current and future} \\ \text{moves to be determined}}} \qquad j = 1, 2, ..., P \quad (8)$$

The terms on the left in eq 8 represent the error between the predicted output and the set point that will exist over the next P sampling instants provided no further manipulated input moves are made by the controller. The term on the right represents the effect of the yet undetermined current and future manipulated input moves.

Equation 8 is a system of linear equations which can be represented as a matrix equation of the form

$$
\begin{bmatrix}
e(n+1) \\
e(n+2) \\
e(n+3) \\
\vdots \\
e(n+M) \\
\vdots \\
e(n+P)
\end{bmatrix}_{P \times 1}
=
$$

$$
\begin{bmatrix}
a_1 & 0 & 0 & \cdots & 0 \\
a_2 & a_1 & 0 & & 0 \\
a_3 & a_2 & a_1 & \ddots & 0 \\
\vdots & \vdots & \vdots & & 0 \\
a_M & a_{M-1} & a_{M-2} & & a_1 \\
\vdots & \vdots & \vdots & & \vdots \\
a_P & a_{P-1} & a_{P-2} & \cdots & a_{P-M+1}
\end{bmatrix}_{P \times M}
\times
$$

$$
\begin{bmatrix}
\Delta u(n) \\
\Delta u(n+1) \\
\Delta u(n+2) \\
\vdots \\
\Delta u(n+M-1)
\end{bmatrix}_{M \times 1}
\tag{9}
$$

or in a compact matrix notation as

$$
\bar{e} = A \Delta \bar{u} \tag{10}
$$

where $\bar{e}$ is the vector of predicted errors over the next P sampling instants, A is the dynamic matrix, and $\Delta \bar{u}$ is the vector of manipulated input moves to be determined.

An exact solution to eq 10 is not possible since the number of equations exceeds the degrees of freedom ($P > M$). Hence, the control objective is alternatively posed as a least-squares optimization problem with a quadratic performance objective (cost function) of the form

$$
\underset{\Delta \bar{u}}{\text{Min}} \; J = [\bar{e} - A\Delta \bar{u}]^{\mathrm{T}} [\bar{e} - A\Delta \bar{u}] \tag{11}
$$

In the unconstrained case, this minimization problem has a closed-form solution which represents the DMC control law

$$
\Delta \bar{u} = (A^{\mathrm{T}} A)^{-1} A^{\mathrm{T}} \bar{e} \tag{12}
$$

Implementation of DMC with the control law in eq 12 results in excessive control action, especially when the control horizon is greater than 1. Hence, a quadratic penalty on the size of manipulated input moves is introduced into the DMC performance objective. The modified performance objective has the form

$$
\underset{\Delta \bar{u}}{\text{Min}} \; J = [\bar{e} - A\Delta \bar{u}]^{\mathrm{T}} [\bar{e} - A\Delta \bar{u}] + [\Delta \bar{u}]^{\mathrm{T}} \lambda [\Delta \bar{u}] \tag{13}
$$

where λ is the move suppression coefficient. For the modified performance objective the closed form solution takes the form (Marchetti et al., 1983; Ogunnaike, 1986):

$$
\Delta \bar{u} = (A^{\mathrm{T}} A + \lambda \mathbf{I})^{-1} A^{\mathrm{T}} \bar{e} \tag{14}
$$

DMC Transfer Function. The DMC transfer function is developed by building upon prior work by Gupta (1987, 1993). Let c_i denote the ith first row element of the pseudoinverse matrix, $(A^{\mathrm{T}} A + \lambda \mathbf{I})^{-1} A^{\mathrm{T}}$. Then, $\Delta u(n)$,

the first element of $\Delta \bar{u}$ in eq 14 to be implemented at the current instant, can be written as

$$
\Delta u(n) = \sum_{i=1}^{P} c_i e(n+i) \tag{15}
$$

The expression for the predicted error from eq 8 can be modified to eliminate y_0 using eq 5:

$$
e(n+j) = y_{\mathrm{sp}}(n+j) - y_0 - \sum_{i=j+1}^{N-1} a_i \Delta u(n+j-i) - d(n)
$$

$$
= y_{\mathrm{sp}}(n+j) - \hat{y}(n) - \sum_{j=1}^{N-1} (a_{i+j} - a_j) \Delta u(n-j) \tag{16}
$$

Substituting eq 16 in eq 15 gives

$$
\Delta u(n) = \sum_{i=1}^{P} c_i \{ y_{\mathrm{sp}}(n+i) - \hat{y}(n) - \sum_{j=1}^{N-1} (a_{i+j} - a_j) \Delta u(n-j) \} \tag{17}
$$

Let $y_{\mathrm{sp}}(n)$ be a weighted average of the desired set point profile over P future sampling instants:

$$
y_{\mathrm{sp}}(n) = \frac{\displaystyle\sum_{i=1}^{P} c_i y_{\mathrm{sp}}(n+i)}{\displaystyle\sum_{i=1}^{P} c_i} \tag{18}
$$

Using eq 18 in eq 17, the latter becomes

$$
\Delta u(n) = \sum_{i=1}^{P} c_i \{ y_{\mathrm{sp}}(n) - \hat{y}(n) \} - \sum_{i=1}^{P} c_i \{ \sum_{j=1}^{N-1} (a_{i+j} - a_j) \Delta u(n-j) \} \tag{19}
$$

Replacing $y_{\mathrm{sp}}(n) - y(n) = e(n)$, the error at the current sample time, eq 19 becomes

$$
\Delta u(n) + \sum_{i=1}^{P} c_i \{ \sum_{j=1}^{N-1} (a_{i+j} - a_j) \Delta u(n-j) \} = \sum_{i=1}^{P} c_i e(n) \tag{20}
$$

Transformation of eq 20 to the z-domain gives the DMC transfer function form:

$$
D(z) = \frac{u(z)}{e(z)} = \frac{1}{(1 - z^{-1}) \dfrac{1}{\displaystyle\sum_{i=1}^{P} c_i} + \displaystyle\sum_{j=1}^{N-1} \left(\dfrac{\displaystyle\sum_{i=1}^{P} c_i a_{i+j}}{\displaystyle\sum_{i=1}^{P} c_i} - a_j \right) z^{-j}} \tag{21}
$$

Although eq 21 provides insight into the form of the DMC transfer function, theoretical analysis of the closed-loop system is very complicated (even after the assumption of a FOPDT model approximation of the

process). This is primarily because it is difficult to represent the elements, c_i, of the pseudoinverse matrix analytically in terms of the step response coefficients.

Consequently, eq 21 does not serve as a convenient starting point for devising the analytical expression for computing λ. However, as shown in the next section, some useful information regarding the form of the move suppression coefficient rule can be extracted from the DMC transfer function.

Gain Scaling of the Move Suppression Coefficient. Previous work has proposed that the choice of the move suppression coefficient can be made independent of the process gain (Lalonde and Cooper, 1989; McIntosh et al., 1991). Gain-scaling is a term coined to represent a transformation where a mathematical expression is stripped of the effects of process gain for analysis independent of gain. For example, gain-scaling of the move suppression coefficient can be done by expressing it as a product of a scaled move suppression coefficient, f, and the square of the process gain, K_p^2. Note that this approach is restricted to DMC applied to SISO systems and cannot be directly extended to multivariable systems.

Consider a form of the move suppression coefficient given by

$$\lambda = f K_p^2 \tag{22}$$

The step response coefficient of any linear system can be written as

$$a_i = K_p \tilde{a}_i \tag{23}$$

where $\tilde{a}_i$ represents the part of the unit step response coefficient that is independent of the process gain, K_p. Using eq 22 and eq 23, it is possible to separate the gain-related effects from the first row elements, c_i, of the pseudoinverse matrix:

$$c_i = i\text{th first row element of } \{(A^T A + \lambda I)^{-1} A^T\}$$
$$= i\text{th first row element of } \{(K_p^2 \tilde{A}^T \tilde{A} + f K_p^2 I)^{-1} K_p \tilde{A}^T\}$$
$$= 1/K_p \times i\text{th first row element of } \{(\tilde{A}^T \tilde{A} + f I)^{-1} \tilde{A}^T\}$$
$$= \tilde{c}_i / K_p \tag{24}$$

Here, $\tilde{A}$ is the gain-scaled dynamic matrix, $\tilde{A}^T \tilde{A}$ is the gain-scaled system matrix, and $\tilde{c}_i$ is the ith first row element of the gain-scaled pseudo-inverse matrix.

Substituting eq 23 and eq 24 in eq 21 shows that the gain dependence of the SISO DMC transfer function is separable from the dependence on remaining process and controller parameters:

$$D(z) = u(z)/e(z) =$$

$$\cfrac{1}{K_p(1 - z^{-1})\left(\cfrac{1}{\sum\limits_{i=1}^{P} \tilde{c}_i} + \sum\limits_{j=1}^{N-1} \left(\cfrac{\sum\limits_{i=1}^{P} \tilde{c}_i \tilde{a}_{i+j}}{\sum\limits_{i=1}^{P} \tilde{c}_i} - \tilde{a}_j \right) z^{-j} \right)} = \tilde{D}(z)/K_p \tag{25}$$

Similarly, the gain dependence of a linear process transfer function is separable from the remaining process parameters:

$$G(z) = y(z)/u(z) = K_p \tilde{G}(z) \tag{26}$$

Using eq 25 and eq 26, the open-loop transfer function is of the form

$$D(z)G(z) = y(z)/e(z) = (\tilde{D}(z)/K_p)K_p \tilde{G}(z) = \tilde{D}(z)\tilde{G}(z) \tag{27}$$

and the closed-loop transfer function is of the form

$$\frac{y(z)}{y_{sp}(z)} = \frac{D(z)G(z)}{1 + D(z)G(z)} = \frac{\tilde{D}(z)\tilde{G}(z)}{1 + \tilde{D}(z)\tilde{G}(z)} \tag{28}$$

By casting the move suppression, λ, as a scaled coefficient, f, times the square of the process gain, K_p^2, the SISO DMC response looks similar for all first-order systems when the time dimension is scaled appropriately. In other words, eq 28 shows that, by gain-scaling λ as in eq 22, the closed-loop performance is rendered independent of the process gain. As a result, derivation of an analytical expression for λ yields a scaled coefficient, f, that is a function of parameters other than the process gain.

Derivation of an Analytical Expression for λ

The analysis of the gain-scaled system matrix, $\tilde{A}^T \tilde{A}$, involves the development of an approximate form of the gain-scaled system matrix. Such an approximation is made possible with the use of a FOPDT model approximation of the process. It must be emphasized that the use of this FOPDT approximation is employed only in the derivation of the analytical expression for λ. The examples presented later in this work all use the traditional DMC step response matrix of the actual process upon implementation.

An Approximate Form of the $\tilde{A}^T \tilde{A}$ Matrix. A FOPDT model with zero-order hold is represented by a discrete transfer function as

$$H_0 G_p(z) = \frac{K_p(1 - e^{-(T/\tau_p)})z^{-k}}{(1 - e^{-(T/\tau_p)}z^{-1})} \tag{29}$$

where K_p is the process gain, τ_p is the overall process time constant, T is the discrete sample time, and k is the effective discrete dead time given by

$$k = \theta_p/T + 1 \tag{30}$$

In eq 30, θ_p is the effective dead time in the process. From eq 29, the step response coefficients of a FOPDT process are given by

$$\tilde{a}_i = \begin{cases} 0 & 0 \le j \le k - 1 \\ (1 - e^{-(T/\tau_p)(j-k+1)}) & k \le j \end{cases} \tag{31}$$

Using a FOPDT model approximation of the process and the gain-scaled step response coefficients from eq 31, the dynamic matrix from eq 9 has the form

$$\tilde{A} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \\ (1-e^{-\frac{T}{\tau_p}}) & 0 & 0 & \cdots & 0 \\ (1-e^{-\frac{2T}{\tau_p}}) & (1-e^{-\frac{T}{\tau_p}}) & 0 & \cdots & 0 \\ (1-e^{-\frac{3T}{\tau_p}}) & (1-e^{-\frac{2T}{\tau_p}}) & (1-e^{-\frac{T}{\tau_p}}) & \ddots & 0 \\ \vdots & \vdots & \vdots & & 0 \\ (1-e^{-(M)\frac{T}{\tau_p}}) & (1-e^{-(M-1)\frac{T}{\tau_p}}) & (1-e^{-(M-2)\frac{T}{\tau_p}}) & \cdots & (1-e^{-\frac{T}{\tau_p}}) \\ \vdots & \vdots & \vdots & & \vdots \\ (1-e^{-(P-k+1)\frac{T}{\tau_p}}) & (1-e^{-(P-k)\frac{T}{\tau_p}}) & (1-e^{-(P-k-1)\frac{T}{\tau_p}}) & \cdots & (1-e^{-(P-k-M+1)\frac{T}{\tau_p}}) \end{bmatrix}_{P \times M} \quad (32)$$

For this form of the dynamic matrix, the gain-scaled system matrix, $\tilde{A}^T\tilde{A}$, in the DMC control law has the form

$$\tilde{A}^T\tilde{A} = \begin{bmatrix} \sum_{i=1}^{P-k+1}(1-e^{-\frac{T}{\tau_p}})^2 & \sum_{i=2}^{P-k+1}(1-e^{-\frac{T}{\tau_p}})(1-e^{-(i-1)\frac{T}{\tau_p}}) & \sum_{i=3}^{P-k+1}(1-e^{-\frac{T}{\tau_p}})(1-e^{-(i-2)\frac{T}{\tau_p}}) & \cdots \\ \sum_{i=2}^{P-k+1}(1-e^{-\frac{T}{\tau_p}})(1-e^{-(i-1)\frac{T}{\tau_p}}) & \sum_{i=2}^{P-k+1}(1-e^{-(i-1)\frac{T}{\tau_p}})^2 & \sum_{i=3}^{P-k+1}(1-e^{-\frac{T}{\tau_p}})(1-e^{-(i-2)\frac{T}{\tau_p}}) & \cdots \\ \sum_{i=3}^{P-k+1}(1-e^{-\frac{T}{\tau_p}})(1-e^{-(i-2)\frac{T}{\tau_p}}) & \sum_{i=3}^{P-k+1}(1-e^{-\frac{T}{\tau_p}})(1-e^{-(i-2)\frac{T}{\tau_p}}) & \sum_{i=3}^{P-k+1}(1-e^{-(i-2)\frac{T}{\tau_p}})^2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}_{M \times M} \quad (33)$$

An approximate form of the gain-scaled system matrix can be obtained by approximating individual terms of the matrix in eq 33 for large values of the prediction horizon, P, and small sample times, T (small T/τ_p). Let $\tilde{\alpha}_{ij}$ ($i,j = 1, 2, ..., M$) be the term in the ith row and jth column of the gain-scaled system matrix. The approximation of one such term, $\tilde{\alpha}_{11}$, is shown in eq 34.

Recognizing that the summation terms in $\tilde{\alpha}_{11}$ are in geometric progression results in the exact expression

$$\tilde{\alpha}_{11} = \sum_{i=1}^{P-k+1}(1-e^{-i(T/\tau_p)})^2$$

$$= \sum_{i=1}^{P-k+1}(1-2e^{-i(T/\tau_p)}+e^{-2i(T/\tau_p)})$$

$$= (P-k+1) - \frac{2e^{-(T/\tau_p)}(1-e^{-(P-k+1)(T/\tau_p)})}{(1-e^{-(T/\tau_p)})} + \frac{e^{-(2T/\tau_p)}(1-e^{-2(P-k+1)(T/\tau_p)})}{(1-e^{-(2T/\tau_p)})} \quad (34)$$

With a FOPDT model approximation available, the prediction horizon, P, can be computed as the open-loop process settling time in samples as $P = (5\tau_p/T) + k$. This supports the findings of past researchers (García and Morshedi, 1986; Maurath et al., 1988a; Lundström et al., 1995) that P should be large enough to include the steady state effect of all past input moves.

For large values of prediction horizon, the term in eq 34 simplifies to

$$\tilde{\alpha}_{11} \cong (P-k+1) - \frac{2e^{-(T/\tau_p)}}{(1-e^{-(T/\tau_p)})} + \frac{e^{-(2T/\tau_p)}}{(1-e^{-(2T/\tau_p)})} \quad (35)$$

Notice that the approximation in eq 35 becomes increasingly accurate as P increases and is exactly true for an infinite horizon implementation of DMC. A first-order Taylor series expansion, $e^{-(T/\tau_p)} = 1 - T/\tau_p$, that is valid for small sample times ($T/\tau_p \le 0.1$), can be applied to eq 35 to yield

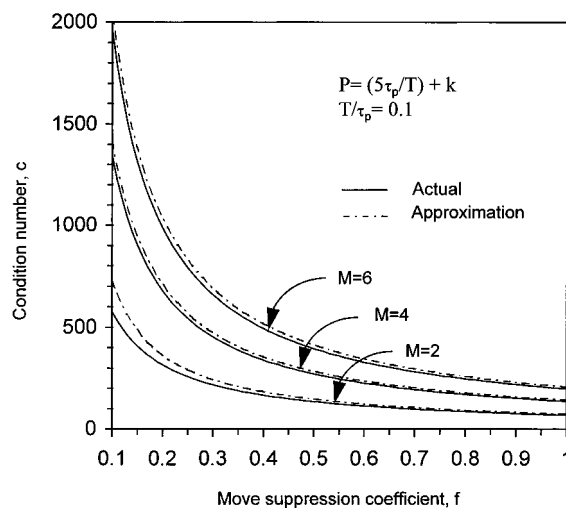$$\tilde{\alpha}_{11} \cong P - k - \frac{3}{2}\frac{T}{\tau_p} + 2 \quad (36)$$

Figure 5. Comparison of the true and approximate condition numbers of the $\tilde{A}^T\tilde{A}$ matrix for large values of the prediction horizon, P.

The validity of the approximation in eq 36 is explored later in this section.

The other terms of the $\tilde{A}^T\tilde{A}$ matrix can be approximated using a similar procedure. The final approximate form of the matrix that results is

$$\tilde{A}^T\tilde{A} \cong \begin{bmatrix} P-k-\frac{3}{2}\frac{T}{\tau_p}+2 & P-k-\frac{3}{2}\frac{T}{\tau_p}+\frac{3}{2} & P-k-\frac{3}{2}\frac{T}{\tau_p}+1 & \cdots \\ P-k-\frac{3}{2}\frac{T}{\tau_p}+\frac{3}{2} & P-k-\frac{3}{2}\frac{T}{\tau_p}+1 & P-k-\frac{3}{2}\frac{T}{\tau_p}+\frac{1}{2} & \cdots \\ P-k-\frac{3}{2}\frac{T}{\tau_p}+1 & P-k-\frac{3}{2}\frac{T}{\tau_p}+\frac{1}{2} & P-k-\frac{3}{2}\frac{T}{\tau_p} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}_{M \times M} \quad (37)$$

Let $\beta = \tilde{\alpha}_{11} \cong P - k - (3/2)(T/\tau_p) + 2$, then

$$\tilde{A}^T\tilde{A} = \begin{bmatrix} \beta & \beta-\frac{1}{2} & \beta-1 & \cdots \\ \beta-\frac{1}{2} & \beta-1 & \beta-\frac{3}{2} & \cdots \\ \beta-1 & \beta-\frac{3}{2} & \beta-2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}_{M \times M} \quad (38)$$

Note that the approximate $\tilde{A}^T\tilde{A}$ matrix has a Hankel matrix form with the added feature that the elements of every row successively decrease by 0.5 from left to right. The approximate matrix of eq 38 is singular when $M \ge 3$. This supports the observation made by prior researchers (Marchetti et al., 1983) that the A^TA matrix (hence, the $\tilde{A}^T\tilde{A}$ matrix) becomes increasingly singular for large values of the prediction horizon, P, and control horizon, M.

Figure 5 provides evidence that the approximate gain-scaled system matrix is a good approximation of the true system matrix, especially for the analysis of the $\tilde{A}^T\tilde{A} + f\tilde{I}$ matrix in the DMC control law. Specifically, Figure 5 shows the condition number for the exact and approximate $\tilde{A}^T\tilde{A} + f\tilde{I}$ matrix as a function of the scaled move suppression coefficient, f, for different choices of the control horizon. The prediction horizon employed in Figure 5 is $(5\tau_p)/T + k$, which is the open-loop settling time of the process based on a FOPDT model ap-

proximation of the process. This value of prediction horizon is recommended in the proposed tuning strategy based on closed-loop stability considerations (García and Morari, 1982; Clarke and Scattolini, 1991; Rawlings and Muske, 1993; Scokaert and Clarke, 1994). From Figure 5 it can be seen that, for large values of prediction horizon, the condition number of the approximate system matrix in eq 38 closely follows the true condition number.

QR Factorization of the Approximate $\tilde{A}^T\tilde{A} + fI$ Matrix. Since the approximate form of the system matrix (eq 38) was shown above to be a reasonable approximation of the true system matrix (eq 33), the following treatment of the $\tilde{A}^T\tilde{A} + fI$ matrix is based on this approximate form.

Consider two linearly independent M-vectors:

$$\bar{h}_1^T = (1 \quad 1 \quad 1 \quad \cdots \quad 1)_{1\times M}$$

$$\bar{h}_2^T = (0 \quad 1 \quad 2 \quad \cdots \quad M-1)_{1\times M} \tag{39}$$

The approximate $\tilde{A}^T\tilde{A}$ matrix (eq 38) can be written in terms of these vectors as

$$\tilde{A}^T\tilde{A} = \bar{v}\cdot\bar{h}_1^T + \bar{h}_1\cdot\bar{v}^T \tag{40}$$

where

$$\bar{v} = \frac{\beta}{2}\bar{h}_1 - \frac{1}{2}\bar{h}_2 \tag{41}$$

Hence, $\bar{h}_1$ and $\bar{h}_2$ form a basis for the approximate $\tilde{A}^T\tilde{A}$ matrix.

A Gram–Schmidt orthonormalization of h_1 and h_2 yields the orthonormal basis for $\tilde{A}^T\tilde{A}$:

$$\bar{e}_1^T = \frac{1}{\sqrt{M}}(1 \quad 1 \quad 1 \quad \cdots \quad 1)_{1\times M}$$

$$\bar{e}_2^T = \sqrt{\frac{12}{M(M-1)(M+1)}}(0 \quad 1 \quad 2 \quad \cdots \quad M-1)_{1\times M} - \frac{(M-1)}{2}(1 \quad 1 \quad 1 \quad \cdots \quad 1)_{1\times M} \tag{42}$$

Therefore, $\bar{h}_1$ and $\bar{h}_2$ can be QR factored as

$$[\bar{h}_1 \quad \bar{h}_2]_{M\times 2} = [e_1 \quad \bar{e}_2]_{M\times 2}R_{2\times 2} \tag{43}$$

where R is upper triangular and invertible. Using eq 43, eq 40 can be transformed to

$$\tilde{A}^T\tilde{A} = \bar{v}\cdot\bar{h}_1^T + \bar{h}_1\cdot\bar{v}^T =$$

$$[\bar{e}_1 \quad \bar{e}_2]_{M\times 2}\begin{bmatrix} a & b \\ b & 0 \end{bmatrix}_{2\times 2}[\bar{e}_1 \quad \bar{e}_2]_{2\times M}^T \tag{44}$$

where a and b are simple linear functions of β. Equation 44 can also be written as

$$\tilde{A}^T\tilde{A} = [\bar{e}_1 \quad \bar{e}_2 \quad \ldots \quad 0]_{M\times M} \times$$

$$\begin{bmatrix} a & b & \vdots & 0 \\ b & 0 & & \\ \cdots & & \cdots & \cdots \\ 0 & & \vdots & 0 \end{bmatrix}_{M\times M} [\bar{e}_1 \quad \bar{e}_2 \quad \ldots \quad 0]_{M\times M}^T \tag{45}$$

Analytical expressions for a and b can be obtained from

eq 45 by solving the system of equations that result from the top left partitioned block using eq 38 and eq 42. Thus, a and b are given by

$$a = M\left\{\beta - \frac{(M-1)}{2}\right\}$$

$$b = -\frac{M\sqrt{(M-1)(M+1)}}{2\sqrt{12}} \tag{46}$$

Now, the $\tilde{A}^T\tilde{A} + fI$ matrix, to be inverted in DMC control law, has the form

$$\tilde{A}^T\tilde{A} + fI =$$

$$\begin{bmatrix} \beta + f & \beta - \frac{1}{2} & \beta - 1 & \cdots \\ \beta - \frac{1}{2} & \beta - 1 + f & \beta - \frac{3}{2} & \cdots \\ \beta - 1 & \beta - \frac{3}{2} & \beta - 2 + f & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}_{M\times M} \tag{47}$$

Adopting the factored form in eq 45 and eq 46, eq 47 is written as

$$\tilde{A}^T\tilde{A} + fI = [\bar{e}_1 \quad \bar{e}_2 \quad \ldots \quad 0]_{M\times M} \times$$

$$\begin{bmatrix} a+f & b & \vdots & 0 \\ b & f & \vdots & \\ \cdots & & \cdots & \cdots \\ 0 & & \vdots & fI \end{bmatrix}_{M\times M} [\bar{e}_1 \quad \bar{e}_2 \quad \ldots \quad 0]_{M\times M}^T \tag{48}$$

Equation 48 can now be used to determine explicit analytical expressions for the eigenvalues of $\tilde{A}^T\tilde{A} + fI$. From eq 47 it is clear that the approximate form of $\tilde{A}^T\tilde{A} + fI$ has eigenvalues $\mu = f$ with multiplicity $(M-2)$. Therefore

$$\mu_{min} = f \tag{49}$$

The remaining two eigenvalues are found from the top left partitioned block as the μ that satisfies

$$\begin{vmatrix} a+f-\mu & b \\ b & f-\mu \end{vmatrix} = 0 \tag{50}$$

A solution to eq 50 (using eq 46) yields the larger of the two eigenvalues as

$$\mu_{max} =$$

$$\frac{M\beta + 2f - \frac{M(M-1)}{2} + \sqrt{M^2\beta^2 - M^2(M-1)\beta + \frac{(M-1)M^2(2M-1)}{6}}}{2} \tag{51}$$

Alternatively, the minimum and maximum eigenvalues of $\tilde{A}^T\tilde{A} + fI$ can be determined by reasoning. Note that the $\tilde{A}^T\tilde{A}$ matrix (eq 38) is nearly singular for $M = 2$ and is perfectly singular for $M > 2$. Therefore, the minimum absolute eigenvalue of $\tilde{A}^T\tilde{A}$ for $M \geq 2$ is close to or exactly zero. When a constant quantity, f, is added to the leading diagonal of such a matrix, all its eigenvalues shift by that quantity (Hoerl and Kennard, 1970; Ogunnaike, 1986). Hence, the minimum absolute eigenvalue of the resultant $\tilde{A}^T\tilde{A} + fI$ matrix, μ_{min}, is equal to f (eq 49).

Analytical expressions for the maximum eigenvalue can be derived for the square matrix, $\tilde{A}^T\tilde{A} + f\mathbf{I}$, with successively increasing dimensions ($M \times M$). By recognizing that the various coefficients in the analytical expressions follow a series pattern that is a function of the matrix dimension ($M \times M$), a general formula for the maximum eigenvalue can be obtained as a function of M. This general analytical expression is identical to the analytical expression obtained in eq 51.

Formulation of the Analytical Expression for λ. Past researchers (e.g., Ogunnaike, 1986) have indicated that the move suppression coefficient, λ, serves a dual purpose in the DMC control law. Its primary role in DMC is to suppress otherwise aggressive controller action when $M > 1$. Additionally, λ serves to improve the conditioning of the system matrix by rendering it more positive definite.

One premise underlying this work is that both these effects are interrelated. When λ is increased, the manipulated input move sizes decrease, as does the condition number. Clearly, if the effect of λ on the condition number of the system matrix can be analytically expressed, then this condition number can be maintained within specified bounds by an appropriate choice of λ. An upper bound on the condition number would, in turn, prevent the manipulated input move sizes from becoming excessive.

The condition number is defined for a square matrix as

$$c = \frac{|\mu_{\max}|}{|\mu_{\min}|} \tag{52}$$

where $\mu_{\max}$ and $\mu_{\min}$ represent the maximum and minimum eigenvalues of the matrix. From eqs 49, 51, and 52, the condition number for the $\tilde{A}^T\tilde{A} + f\mathbf{I}$ matrix is

$$c = \frac{1}{2f}\left(M\beta + 2f - \frac{M(M-1)}{2} + M\sqrt{\beta^2 - (M-1)\beta + \frac{(M-1)(2M-1)}{6}}\right) \tag{53}$$

An interesting approximate relation from eq 53 is that the condition number of the DMC gain-scaled system matrix, $c \propto M(P-k)/f$.

Equation 53 is rearranged to give an expression for the scaled move suppression coefficient as

$$f = \frac{M}{2c}\left(\beta - \frac{(M-1)}{2} + \sqrt{\beta^2 - (M-1)\beta + \frac{(M-1)(2M-1)}{6}}\right) \tag{54}$$

where $\beta = P - k - (3/2)(T/\tau_p) + 2$. Equation 54 expresses the gain-scaled coefficient as a function of the condition number of $\tilde{A}^T\tilde{A} + f\mathbf{I}$, the control horizon, and FOPDT model parameters other than the process gain.

Evaluation of f in eq 54 can be simplified by recognizing the contribution of each term to the value of f. The last term within the square root in eq 54 can be modified to ease evaluation:

$$\beta^2 - (M-1)\beta + \frac{(M-1)(2M-1)}{6} \cong$$
$$\beta^2 - (M-1)\beta + \frac{(M-1)^2}{4} = \left(\beta - \frac{(M-1)}{2}\right)^2 \tag{55}$$

With this simplification, the expression for the gain scaled coefficient becomes

$$f = \frac{M}{c}\left(\beta - \frac{(M-1)}{2}\right) \tag{56}$$

Past researchers (Maurath et al., 1988a,b; Callaghan and Lee, 1988; Farrell and Polli, 1990) have indicated typical condition numbers for a moderately ill-conditioned DMC system matrix. Condition numbers reported range from about 100 for single-input single-output systems to about 2000 for multivariable systems. Apart from the variety of systems considered, the approximate relation, $c \propto M(P-k)/f$, indicates that some of the differences are due to the different prediction and control horizons used by the individual designers. In this work, a condition number of 500 is selected to represent the upper allowable limit of ill-conditioning in the system matrix (corresponding to modest control effort).

The choice of the condition number, and hence the upper allowable limit of ill-conditioning in the system matrix, lies with the individual designer. Depending on the criterion for good closed-loop performance for a given application, the condition number can be conveniently fixed. The choice of a condition number of 500 was motivated by the rule of thumb that the manipulated input move sizes for a change in set point should not exceed 2−3 times the final change in manipulated input (Maurath et al., 1985; Callaghan and Lee, 1988). However, if a faster or slower closed-loop response is more desirable, a larger or smaller condition number respectively, can be used instead.

Substituting the expression for β in eq 56, with $P = (5\tau_p)/T + k$, the analytical expression for the move suppression coefficient, λ, is given by

$$f = \frac{M}{500}\left(\frac{7}{2}\frac{\tau_p}{T} + 2 - \frac{(M-1)}{2}\right)$$
$$\lambda = fK_p^2 \tag{57}$$

The analytical expression for λ in eq 57, applied in conjunction with the recommended values for the other tuning parameters, gives the tuning strategy for DMC with $M > 1$.

Note that eq 57, which computes the move suppression coefficient, does not contain a dead time term. The reason is not intuitively obvious but is an interesting one. The elements, and the condition number, of the $(A^TA + \lambda\mathbf{I})$ matrix depend on the number of non-zero terms in the dynamic matrix, A. This is clearer from eq 33, where the first term of the $(A^TA + \lambda\mathbf{I})$ matrix is a series summation performed over the $P - k + 1$ non-zero terms in the dynamic matrix A. (Equation 37 also conveys the dependence of condition number of $(A^TA + \lambda\mathbf{I})$ on $P - k$, rather than P alone). Additionally, the choice of prediction horizon as $P = (5\tau_p/T) + k$ causes the elements of $(A^TA + \lambda\mathbf{I})$ to depend on $P - k = (5\tau_p/T)$. Hence, the condition number of $(A^TA + \lambda\mathbf{I})$ and the move suppression from eq 57, are dependent on $(5\tau_p/T)$ and are independent of dead time. Of course, this is

true only if the prediction horizon is chosen as $P = (5\tau_{\mathrm{p}}/T) + k$ (settling time of the process).

Extension of Tuning Strategy to Multivariable Systems

The selection of tuning parameters for multiple-input multiple-output (MIMO) DMC is significantly more challenging for the practitioner. The analytical expression that computes the move suppression coefficient (eq 57), developed in the previous section for SISO DMC, provides the foundation upon which a similar analytical expression can be developed for MIMO DMC. This is possible in a straightforward fashion even though the performance objective for MIMO DMC is defined over several manipulated inputs and measured outputs and results in a more complex MIMO DMC control law.

For a system with S manipulated inputs and R measured outputs, the MIMO DMC performance objective (cost function) has the form (García and Morshedi, 1986)

$$\min_{\Delta\bar{\boldsymbol{u}}} J = [\bar{\boldsymbol{e}} - \boldsymbol{A}\Delta\bar{\boldsymbol{u}}]^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}[\bar{\boldsymbol{e}} - \boldsymbol{A}\Delta\bar{\boldsymbol{u}}] + [\Delta\bar{\boldsymbol{u}}]^{\mathrm{T}}\boldsymbol{\Lambda}^{\mathrm{T}}\boldsymbol{\Lambda}[\Delta\bar{\boldsymbol{u}}]$$

(58)

In eq 58, $\bar{\boldsymbol{e}}$ is the vector of predicted errors for the R measured outputs over the next P sampling instants, $\boldsymbol{A}$ is the multivariable dynamic matrix, and $\Delta\bar{\boldsymbol{u}}$ is the vector of M moves to be determined for each of the S manipulated inputs. $\boldsymbol{\Lambda}^{\mathrm{T}}\boldsymbol{\Lambda}$ is a square diagonal matrix of move suppression coefficients with dimensions ($M{\cdot}S \times M{\cdot}S$). Similarly, $\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}$ is a square diagonal matrix that contains the controlled variable weights (equal concern factors) with dimensions ($P{\cdot}R \times P{\cdot}R$).

A closed-form solution to the MIMO DMC performance objective (eq 58) results in the MIMO DMC control law (García and Morshedi, 1986):

$$\Delta\bar{\boldsymbol{u}} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{A} + \boldsymbol{\Lambda}^{\mathrm{T}}\boldsymbol{\Lambda})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}\bar{\boldsymbol{e}}$$

(59)

The MIMO DMC control law (eq 59) is similar to the SISO DMC control law (eq 14), except for the form of the move suppression coefficients, $\boldsymbol{\Lambda}^{\mathrm{T}}\boldsymbol{\Lambda}$, and the introduction of controlled variable weights, $\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}$.

The move suppression coefficients in multivariable DMC follow a notation slightly different from SISO DMC. $\boldsymbol{\Lambda}$ is a square diagonal matrix with dimensions ($M{\cdot}S \times M{\cdot}S$). This matrix can be divided into S^2 square blocks, each with dimensions ($M \times M$). The leading diagonal elements of the first ($M \times M$) matrix block along the diagonal of $\boldsymbol{\Lambda}$ are λ_1, the leading diagonal elements of the next ($M \times M$) matrix block along the diagonal of $\boldsymbol{\Lambda}$ are λ_2, and so on. All off-diagonal elements of the matrix $\boldsymbol{\Lambda}$ are zero. Hence, the matrix of move suppression coefficients, $\boldsymbol{\Lambda}^{\mathrm{T}}\boldsymbol{\Lambda}$, has the form

$$\boldsymbol{\Lambda}^{\mathrm{T}}\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1^2\mathbf{I}_{M\times M} & \vdots & 0 & \vdots & 0 \\ \cdots & \vdots & \cdots & \vdots & \cdots \\ 0 & \vdots & \lambda_2^2\mathbf{I}_{M\times M} & \vdots & 0 \\ \cdots & \vdots & \cdots & \vdots & \cdots \\ 0 & \vdots & 0 & \vdots & \ddots \end{bmatrix}_{M\cdot S \times M\cdot S}$$

(60)

Thus, in the MIMO DMC control law (eq 59), the move suppression coefficients that are added to the leading diagonal of the multivariable system matrix, ($\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{A}$), are λ_i^2 ($i = 1, 2, ..., S$). Similarly, the matrix of

controlled variable weights, $\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}$, has the form

$$\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma} = \begin{bmatrix} \gamma_1^2\mathbf{I}_{P\times P} & \vdots & 0 & \vdots & 0 \\ \cdots & \vdots & \cdots & \vdots & \cdots \\ 0 & \vdots & \gamma_2^2\mathbf{I}_{P\times P} & \vdots & 0 \\ \cdots & \vdots & \cdots & \vdots & \cdots \\ 0 & \vdots & 0 & \vdots & \ddots \end{bmatrix}_{P\cdot R \times P\cdot R}$$

(61)

Hence, the controlled variable weights are γ_i^2 ($i = 1, 2, ..., R$).

Similar to SISO DMC, adjustable parameters that can be used to alter closed-loop performance for MIMO DMC include the prediction horizon, P, the control horizon, M, the model horizon, N, the sample time, T, and the move suppression coefficients, λ_i^2. In addition, MIMO DMC has yet another set of adjustable parameters in the form of the controlled variable weights, γ_i^2.

Just as with the SISO case, practical limitations often restrict the availability of sample time, T, as a tuning parameter for MIMO DMC (Franklin and Powell, 1980; Åström and Wittenmark, 1984). The model horizon, N, is also not an appropriate tuning parameter since truncation of the model horizon can lead to unpredictable closed-loop performance (Lundström et al., 1995). As demonstrated earlier for SISO DMC, the control horizon, M, does not have a significant effect on closed-loop performance, especially in the presence of move suppression.

The controlled variable weights, γ_i^2, serve a dual purpose in MIMO DMC. These weights can be appropriately chosen by the user to scale measurements of the R measured outputs to comparable units. Also, it is possible to achieve tighter control of a particular measured output by selectively increasing the relative weight of the corresponding least-square residual. Hence, controlled variable weights are usually specified by the user for a certain application and cannot be employed as the primary tuning parameters for MIMO DMC.

Analogous to SISO DMC, the move suppression coefficients can be conveniently used to fine tune MIMO DMC for desired closed-loop performance. Since the dual benefit of the move suppression coefficients, λ_i^2, is again to improve the conditioning of the MIMO DMC system matrix ($\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{A}$) and move size suppression for the S inputs, a strategy analogous to SISO DMC can be used to extend eq 57 to compute the move suppression coefficients for MIMO DMC.

Building upon the analogy, an approximation of the MIMO DMC system matrix, ($\boldsymbol{A}^{\mathrm{T}}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{A}$), has the form of a mosaic Hankel matrix (not shown here) comprised of S^2 Hankel matrix blocks. The S^2 Hankel matrix blocks, each with dimensions ($M \times M$), have a form identical to that obtained earlier (eq 47) from a similar approximation of the scaled SISO DMC system matrix, $\tilde{\boldsymbol{A}}^{\mathrm{T}}\tilde{\boldsymbol{A}}$.

The impact of a change in the ith manipulated input on all measured outputs is reflected in the ith diagonal Hankel matrix block. Hence, it is possible to select the ith move suppression coefficient, λ_i^2, such that the condition number of the ith diagonal matrix block is always bounded by a fixed low value. By holding the condition number of the ith diagonal matrix block at a low value, desirable closed-loop performance is achieved while preventing the ith manipulated input move size from becoming excessive.

With this understanding, an analytical expression that computes the move suppression coefficients for MIMO DMC can be obtained as

$$\lambda_i^2 = \frac{M}{500} \sum_{j=1}^{R} \left[\gamma_j^2 K_{p_{ij}}^2 \left(P - k_{ij} - \frac{3}{2} \frac{\tau_{p_{ij}}}{T} + 2 - \frac{(M-1)}{2} \right) \right]$$

$$(i = 1, 2, ..., S) \quad (62)$$

Using eq 62, the S move suppression coefficients, one for each input, can be computed for a given sampling time, T, control horizon, M, and controlled variable weights, γ_i.

In eq 63, it is not possible to substitute $P = (5\tau_p)/T + k$ as was done for the SISO case since the MIMO DMC architecture requires a single value of the prediction horizon to be selected for all manipulated input and measured output pairs. Nevertheless, the analytical expression for MIMO DMC is similar in form to that obtained for SISO DMC (eq 57). In fact, for a single-input single-output process, the analytical expression for MIMO DMC (eq 62) reduces to the analytical expression for SISO DMC (eq 57) when $P = (5\tau_p)/T + k$.

Implementation of the DMC Tuning Strategy

The proposed DMC tuning strategy, which includes the analytical expression for the move suppression coefficient, λ, is presented in Table 1. This tuning strategy can be applied to unconstrained DMC in closed loop with a broad class of SISO processes that are open-loop stable, including those with challenging control characteristics such as high process order, large dead time, and noniminimum phase behavior.

Step 1 (Table 1) involves the identification of a first order plus dead time (FOPDT) model approximation of the process. An accurate identification of the FOPDT model parameters is essential to the success of this tuning strategy. Since a model is only as good as the data it fits, it is necessary that the input−output data used for model fitting be rich in dynamic information of the process. Typically, this is achieved by perturbing the process to obtain a signal-to-noise ratio greater than 10:1 (Seborg et al., 1986). Additionally, the data must be collected over a reasonable period of time to capture the complete process dynamics. An FOPDT fit thus obtained often reasonably describes the process gain, K_p, overall time constant, τ_p, and effective dead time, θ_p, of higher-order processes.

Step 2 involves the selection of an appropriate sample time, T. In most practical applications the user does not have the freedom to choose sample time. The tuning strategy does require that T be known. In cases where the designer can select the sample time, certain rules of thumb that guide in its selection are available. Rules that select T for a desired closed-loop bandwidth, ω_B, have been proposed in the past, e.g., $T \leq 2\pi/10\omega_B$ (Middleton, 1991). Since the choice of T is related to the overall time constant, τ_p, and the effective dead time of the process, θ_p (Seborg et al., 1986), the estimated FOPDT model parameters provide a convenient way to select T. A good rule of thumb is to select T such that the sampling rate is not less than 10 times per time constant. Hence, for DMC, the recommended sample time is

$$T \leq 0.1\tau_p \qquad (63)$$

An additional criterion for sample time selection is that the sampling rate should be high enough to sample two or three times per effective dead time of the process, $T \leq 0.5\theta_p$ (Seborg et al., 1986). This criterion is not a stringent requirement for DMC. However, if sample time can be chosen, the prudent approach is to pick the largest T that satisfies both criteria mentioned above. Once T is fixed, the discrete dead time, k, is calculated from the effective dead time of the process, θ_p in step 3.

Step 4 computes a model horizon, N, and a prediction horizon, P, from τ_p, θ_p, and T. Note that both N and P cannot be selected independent of the sample time, T. For DMC, it is imperative that N be equal to the open-loop process settling time in samples to avoid truncation error in the model prediction (Lundström et al., 1995). For computing P, a rigorous treatment by past researchers (García and Morari, 1982; Clarke and Scattolini, 1991; Rawlings and Muske, 1993; Scokaert and Clarke, 1994) has shown that a larger prediction horizon improves nominal stability of the closed loop. For practical applications this translates to using a reasonably large but finite P. Bearing in mind this important requirement for stability, P is also set equal to the open-loop process settling time in samples. With a FOPDT model approximation available, prediction horizon and model horizon can both be computed as $P = N = (5\tau_p/T) + k$.

In an industrial setting, there can be barriers to selecting different P and T values for different SISO DMC controllers. Also, the MIMO DMC architecture requires that a single P and a single T be selected for all manipulated input and measured output pairs. If the user has the freedom to select a sample time in such a scenario, a conservative choice would be the smallest value of T that satisfies eq 63 for all individual manipulated input and measured output pairs. Based on this sample time (or the fixed sample time), a single prediction horizon can be computed as $P = \max \{(5\tau_{p_{ij}})/T + k_{ij} (i = 1, 2, ..., S; j = 1, 2, ..., R)\}$ for a system with S manipulated inputs and R measured outputs.

Step 5 requires the specification of a control horizon, M, and the computation of an appropriate move suppression coefficient, λ. Recommended values of M range from 1 through 6. However, selecting $M > 1$ provides advance knowledge of the impending manipulated input moves by the controller and can be very useful to the practitioner (Ogunnaike, 1986). Also, Rawlings and Muske (1993) have shown that the stability of infinite horizon MPC can be guaranteed if M is greater than or equal to the number of unstable modes in the process.

The strategy for computing λ depends on the choice of M. As demonstrated earlier, with $M = 1$ the need for a move suppression is obviated and λ is set equal to zero. However, if $M > 1$, a scaled move suppression coefficient, f, is first computed from an analytical expression (eq 57). The product of the scaled move suppression coefficient, f, and the square of the process gain, K_p^2, determines an appropriate value for λ.

Step 6 summarizes the DMC tuning parameters. For certain applications, more specific or stringent performance criteria regarding the manipulated input move sizes or the nature of the measured output response may apply. For such cases, it may be necessary to fine tune DMC for desired performance by altering P and λ from the starting values given by the tuning strategy. The recommended approach is to increase λ for smaller move sizes and slower output response.

Validation of the DMC Tuning Strategy

All the simulation examples presented in this work use the traditional DMC step response matrix of the actual process upon implementation and a negligible plant-model mismatch is assumed. In the past, several researchers have investigated the effects of plant-model mismatch on controller performance (Morari and Zafiriou, 1989; Zafiriou, 1991; Ohshima et al., 1991; Gupta, 1993; Lee and Yu, 1994; Lundström et al., 1995). Hence, this work focuses strictly on the capabilities of the tuning strategy in providing desirable closed-loop performance.

Processes with a range of characteristics that pose specific challenges in process control are employed for the purpose of tuning strategy validation. Specifically, characteristics such as large dead time, minimum phase behavior, inverse response, and high process order and their effects on the performance of the tuning strategy are investigated. The impact of user-specified sample time, T, and control horizon, M, is also explored. Disturbance rejection performance of DMC tuned using the tuning strategy is evaluated. The proposed extension of the tuning strategy (eq 62) is also validated with an example application to binary distillation column control.

Tuning Strategy Validation for Challenging Processes.

For validation of the proposed tuning strategy, three example higher-order processes with different dynamic characteristics are considered in addition to process 1:

process 2

$$G_p(s) = \frac{(1 - 50s)e^{-10s}}{(100s + 1)^2} \qquad (64)$$

process 3

$$G_p(s) = \frac{(50s + 1)e^{-10s}}{(100s + 1)^2} \qquad (65)$$

process 4

$$G_p(s) = \frac{e^{-10s}}{(50s + 1)^4} \qquad (66)$$

Process 1 (eq 2) is a second-order process with a relatively large dead time, process 2 exhibits inverse response characteristics, process 3 has a minimum phase behavior, and process 4 is a fourth-order process with sluggish open-loop dynamics.

The results of the tuning strategy applied to DMC with process 1 are not explicitly shown since they are already demonstrated in Figures 2−4. For example, Figure 2c represents the results of the tuning strategy applied to process 1 (FOPDT fit: $K_p = 1.0$; $\tau_p = 157$; $\theta_p = 70$). Here, the sample time, T, is selected such that $T/\tau_p = 0.1$ and the control horizon $M = 4$. The effectiveness of the tuning strategy in selecting appropriate values for the tuning parameters for processes with a fairly large dead time is clear from these demonstrations.

Figures 6−8 are simulation results when the tuning strategy is applied to processes 2−4, respectively. These figures illustrate the response of the respective processes to a step change in set point. Specifically, each figure shows a matrix of four graphs, where each graph
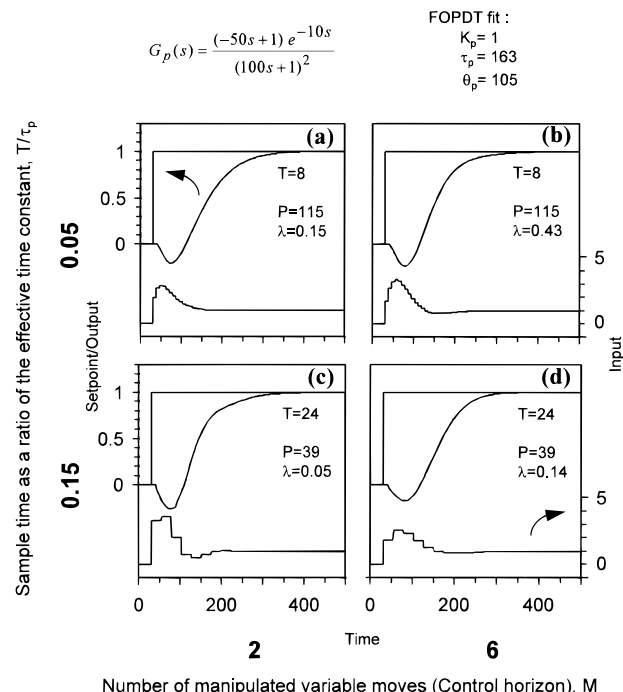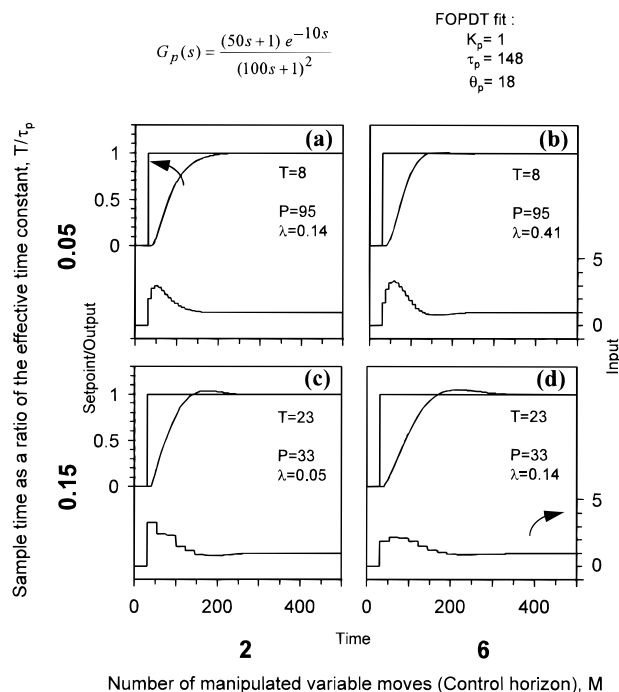

Figure 6. Effectiveness of the DMC tuning strategy demonstrated for process 2 for different choices of the control horizon, M, and sample time, T.

represents the performance of DMC tuned using the rules in Table 1, for a user-selected control horizon of 2 or 6 and sample time with a T/τ_p of 0.05 or 0.15.

Figure 6 shows the tuning results for process 2 (FOPDT fit: $K_p = 1.0$; $\tau_p = 163$; $\theta_p = 105$). The benefit of a FOPDT model approximation as a basis for the tuning strategy is evident here as the initial inverse response (due to the zero in the right half plane) of process 2 is approximated as an additional dead time in the FOPDT model fit. The additional dead time results in a larger P for this inverse response process (compared to the value of P recommended for the minimum phase process in Figure 7). Consequently, the DMC controller is able to look beyond the initial inverse response. Also, the initial inverse response prevents the output response from overshooting the final set point. Nevertheless, with the computed tuning parameters, a desirable closed-loop response results while also preventing the manipulated input move sizes per unit gain of the process from becoming excessive.

Figure 7 illustrates the tuning results for process 3 (FOPDT fit: $K_p = 1.0$; $\tau_p = 148$; $\theta_p = 18$). Despite the second-order characteristics of this process, the FOPDT model approximation results in a short dead time. This is because the process exhibits minimum phase behavior (due to the zero in the left half plane). The graphs in Figure 7 show that the measured output exhibits a modest overshoot and has a faster rise time than the process with the initial inverse response in Figure 6. However, the tuning parameters given by the tuning strategy result in acceptable closed-loop performance with manipulated input moves of a similar magnitude.

Figure 8 represents the results from the application of the tuning strategy to process 4 (FOPDT fit: $K_p = 1.0$; $\tau_p = 124$; $\theta_p = 99$). Due to the high order of the process, and the associated initial sigmoidal response, the FOPDT model approximation provides a larger dead time estimate. The graphs in Figure 8 show that the high order of the process results in a closed-loop response to a set point step that is more underdamped

$$G_p(s) = \frac{(50s+1)\,e^{-10s}}{(100s+1)^2}$$

FOPDT fit :
$K_p = 1$
$\tau_p = 148$
$\theta_p = 18$

Figure 7. Effectiveness of the DMC tuning strategy demonstrated for process 3 for different choices of the control horizon, M, and sample time, T.

$$G_p(s) = \frac{e^{-10s}}{(50s+1)^4}$$

FOPDT fit :
$K_p = 1$
$\tau_p = 124$
$\theta_p = 99$
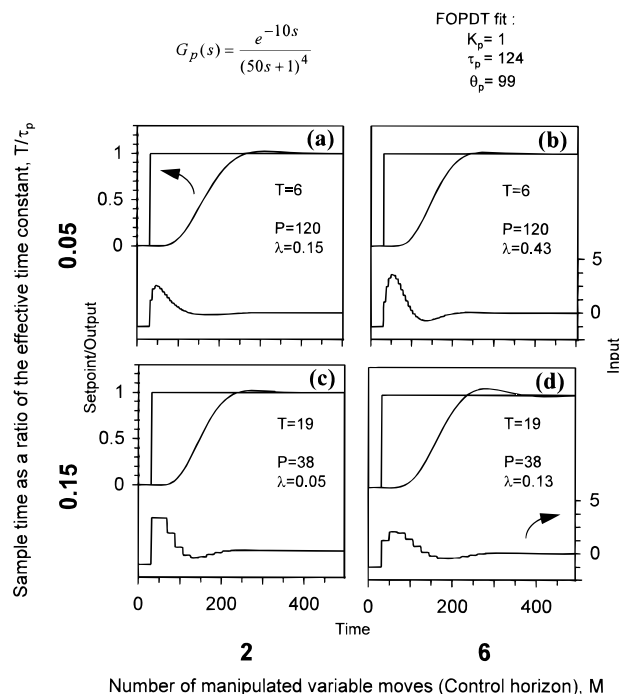
Figure 8. Effectiveness of the DMC tuning strategy demonstrated for process 4 for different choices of the control horizon, M, and sample time, T.

than low-order processes and exhibits a larger peak overshoot. However, even for this fourth-order process, the tuning strategy successfully implements DMC to provide desirable closed-loop performance.

A general rule of thumb for the manipulated input is that the move sizes should not exceed 2–3 times the steady state change in that manipulated input (Maurath et al., 1985; Callaghan and Lee, 1988). In addition to preventing the early wear of final control elements, small manipulated moves also contribute to an increase in robustness of the controller. Figures 6–8 show that the proposed tuning strategy results in manipulated

input sizes that lie within this recommended range. However, a stricter guideline imposed by the individual designer may be achieved by appropriately decreasing the upper bound of the system matrix condition number from the value of 500 used in this work.

Effect of Sample Time and Control Horizon. Figures 6–8, each comprise a matrix of closed-loop response results for different T and M. Results are presented for sample times such that the ratio T/τ_p is 0.05 and 0.15. M is selected to be either 2 or 6 manipulated input moves. The range of T and P explored corresponds to that recommended by the proposed tuning strategy.

The impact of T on DMC closed-loop performance when P is held constant was shown in Figure 3. Figures 6–8 show that the tuning strategy of Table 1 is effective in accounting for different sampling rates. For example, graphs a and c of Figure 6 show the effectiveness of the tuning strategy when T/τ_p for process 2 is increased from 0.05 to 0.15. As T increases, P is appropriately reduced to maintain similar closed-loop performance. Similar comparisons between other pairs of graphs vertically stacked in Figures 6–8 lead to the same conclusion.

Another interesting observation can be made about the effect of T on the analytical expression for λ. For example, graphs a and c of Figure 6 show that the λ computed from the analytical expression decreases when T is increased. This is due to a shorter P that results when T is increased. For a fixed M, as P decreases the system matrix becomes less singular (ill-conditioned) and the overall magnitude of its elements decreases. Hence, a smaller λ is sufficient to provide the same effect as a larger λ with a larger prediction horizon.

The effect of increasing M from 2 to 6 is rather insignificant for a second order plus dead time process (process 1), as shown earlier in Figure 4. However, for other higher-order processes considered in this study, an increase in control horizon causes the manipulated input moves to be more aggressive and the output to exhibit a greater peak overshoot. Figures 6–8 show that the control effort, and the closed-loop performance, can be reverted back to the original value by an increase in λ. For example, graphs a and b of Figure 6 show that λ computed from the analytical expression increases as the control horizon is increased from 2 to 6. Comparisons between other pairs of graphs placed side by side in Figures 6–8 lead to the same conclusion. From these demonstrations it is apparent that the proposed tuning strategy shows promise indeed for the initial tuning of unconstrained SISO DMC.

Effectiveness of Tuning Strategy for Disturbance Rejection. In addition to set point tracking, a controller should be capable of rejecting unexpected disturbances that cause the process to deviate from the desired operating conditions. Other researchers (e.g., Farrell and Polli (1990)) have established that an advantage of DMC over PID-type classical controllers for disturbance rejection is that it provides superior performance when disturbance tuning differs significantly from servo tuning. Hence, the study presented here investigates only the effect of move suppression on DMC performance for disturbance rejection and the effectiveness of the tuning strategy for selection of the appropriate move suppression coefficient.

It is important to emphasize that the form of DMC considered in this work allows the introduction of feedforward control in a natural way to compensate for

$$G_p(s) = \frac{e^{-50s}}{(150s+1)(25s+1)}$$

FOPDT fit :
$K_p = 1$
$\tau_p = 157$
$\theta_p = 70$
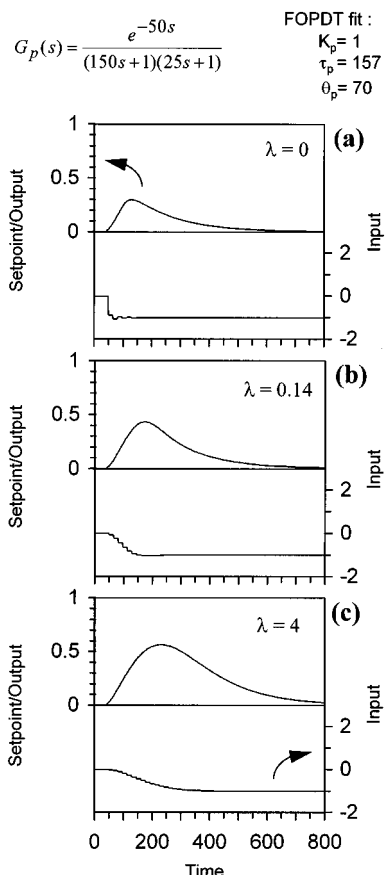
Figure 9. Effectiveness of the DMC tuning strategy in disturbance rejection demonstrated for process 1 ($T = 16$, $T/\tau_p = 0.1$, $P = 54$, $M = 4$).

measurable disturbances. Several researchers (Asbjornsen, 1984; Yocum and Zimmerman, 1988; Takamatsu et al., 1988; Glova et al., 1989; Hokanson et al., 1989; Muske and Rawlings, 1993; Prada et al., 1993; Soroush and Kravaris, 1994) have studied the performance of feed-forward DMC and demonstrated that it enhances the disturbance rejection capability of DMC, provided an accurate model of the disturbance to output model is available. However, the addition of a feed-forward feature in DMC does not affect the selection of tuning parameters, and the tuning strategy presented in this paper applies directly to a feed-forward configuration of DMC. For the same reason, this paper demonstrates the effectivness of the proposed tuning strategy only for DMC without feedforward.

Graphs a–c of Figure 9 demonstrate the effect of the move suppression coefficient, λ, on disturbance rejection for DMC without feed-forward and process 1 (eq 2). The disturbance to output dynamics considered also have a transfer function given by eq 2. In this study, all tuning parameters other than the move suppression coefficient are fixed at values computed using the DMC tuning strategy in Table 1.

Figure 9a shows that disturbance rejection is most effective when the move suppression is eliminated ($\lambda = 0$). Interestingly, with no move suppression, the manipulated input sizes are not excessive for disturbance rejection. However, $\lambda = 0$ is not an appropriate choice of the move suppression as it results in undesirable aggressive manipulated input moves for set point tracking. This was demonstrated earlier in Figure 2b.

Disturbance rejection performance for the move suppression coefficient computed using the analytical ex-

pression in eq 58 is illustrated in Figure 9b. In addition to effective disturbance rejection, the computed move suppression coefficient results in modest manipulated input move sizes. When the move suppression coefficient is increased to $\lambda = 4.0$, disturbance rejection performance worsens, as shown in Figure 9c. Such a large move suppression coefficient was also shown earlier in Figure 2d to be undesirable since it results in a relatively slower response for set point tracking. Hence, this study shows that the proposed tuning strategy results in tuning parameters that effectively reject unmeasured disturbances.

An Example Multivariable Application. The tuning strategy, presented and validated in this paper for SISO DMC, was extended to MIMO DMC as the analytical expression in eq 62. An example application of the tuning strategy extended to MIMO DMC is demonstrated here for a 2 × 2 distillation column model (Wood and Berry, 1973). This methanol–water column model has been used for several controller studies, including the application of model predictive controllers (Marchetti, 1982; Maurath, 1985, Lee and Yu, 1994) and can be written as

$$\begin{bmatrix} X_D(s) \\ X_B(s) \end{bmatrix} = \begin{bmatrix} \dfrac{12.8e^{-s}}{(16.7s+1)} & \dfrac{-18.9e^{-3s}}{(21.0s+1)} \\ \dfrac{6.6e^{-7s}}{(10.9s+1)} & \dfrac{-19.4e^{-3s}}{(14.4s+1)} \end{bmatrix} \begin{bmatrix} R_F(s) \\ Q_V(s) \end{bmatrix} \quad (67)$$

Here, X_D represents the mole fraction of methanol in the distillate, X_B the mole fraction of methanol in the bottoms, R_F the reflux flow, and Q_V the vapor boil-up rate. All the times are given in minutes.

A sample time, T, of 3 min was used in this study. The prediction horizon, P, and model horizon, N, were computed to include the steady state terms in the step response of all manipulated variable and measured output pairs as $P = N = \max\{(5\tau_{pij})/T + k_{ij} (i = 1, 2, ..., S; j = 1, 2, ..., R)\}$. Equal controlled variable weights were used for both measured outputs, i.e., $\mathbf{\Gamma}^T\mathbf{\Gamma} = \mathbf{I}$. The move suppression coefficients, $\mathbf{\Lambda}^T\mathbf{\Lambda}$, were computed from the analytical expression for MIMO DMC in eq 62 for two different choices of the control horizon, $M = 2$ and $M = 6$. For the control horizon, $M = 2$, eq 62 computes the diagonal elements of the matrix $\mathbf{\Lambda}$ as $\lambda_1 = 4.9$ and $\lambda_2 = 9.1$. For $M = 6$, the diagonal elements computed are $\lambda_1 = 8.3$ and $\lambda_2 = 15.2$.

Figure 10 illustrates the response of the two measured outputs and the corresponding manipulated input moves for a step change in the X_D set point. The results in Figure 10 show that, for a user-specified control horizon of $M = 2$ or $M = 6$, the analytical expression for MIMO DMC computes move suppression coefficients that result in desirable closed-loop performance. In either case, the output response tracks the set point rapidly without exhibiting overshoot and the corresponding manipulated input moves are modest in size.

Despite the interactions inherent in the moderately ill-conditioned Wood and Berry column, the effect of interactions on the measured output, X_B, are dampened effectively. This demonstration confirms that the tuning strategy presented for SISO DMC holds potential for a formal extension to the more challenging problem of tuning MIMO DMC.

Conclusions

A tuning strategy for DMC, with a novel analytical expression for the move suppression coefficient, λ, was
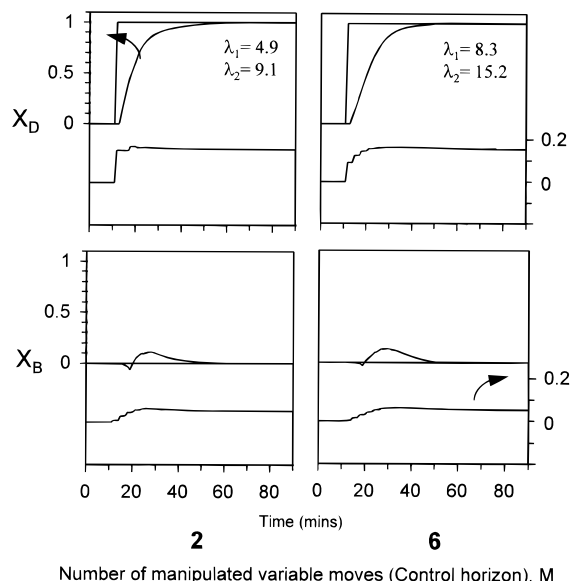
Figure 10. Effectiveness of the MIMO DMC tuning strategy demonstrated for the Wood and Berry column for different choices of the control horizon, M, and sample time, T.

presented. The application of this tuning strategy was demonstrated through simulations on a range of SISO, unconstrained, open-loop stable processes. Specifically, the effect of process order and process characteristics such as large dead time, inverse response, and minimum phase behavior were investigated. Also, the effect of control horizon, M, and sample time, T, on selection of λ was studied. The DMC controller tuned in this fashion exhibited good set point tracking capability with acceptable peak overshoot and modest control effort. Disturbance rejection performance of DMC tuned with the proposed tuning strategy was also explored. The DMC controller tuned using the proposed strategy effectively and smoothly rejected unexpected disturbances that impacted the process.

The elegance of the tuning strategy stems from the use of a simplistic FOPDT model approximation of the process. A FOPDT approximation made the mathematical treatment of the DMC control law more tractable. Also, such a model was shown to preserve certain characteristics of the process dynamics that helped compensate for the high order and inverse response behavior of processes. Specifically, in these cases, an increase in estimated dead time and time constant led to the recommendation of a larger prediction horizon. However, plant-model mismatch was avoided by using the traditional DMC step response matrix of the actual process upon implementation.

To allow tuning of SISO DMC independent of the process gain, existence of a scaled move suppression coefficient, f, observed by past researchers, was formalized. Subsequently, the systematic form of the approximate scaled system matrix obtained led to a better understanding of the effects of DMC and process parameters on the selection of the move suppression coefficient. Consequently, an analytical form for the move suppression coefficient was formulated. This analytical form directly addressed the issue of ill-conditioning of the system matrix and the resultant move sizes.

The form of the analytical expression for the condition number (eq 55) provided valuable insight into the extent of the impact that each of the DMC and process parameters have on the overall conditioning of the

system matrix. An important observation was that, as an approximate relationship, the condition number of the DMC system matrix is $c \propto M(P - k)/f$.

Tuning of MIMO DMC and the removal of ill-conditioning in such systems was addressed only in limited detail in this work. However, a mathematical framework upon which a formal tuning strategy can be built for MIMO DMC was presented. Toward obtaining a tuning strategy for MIMO DMC, the similarities between SISO and MIMO DMC were established. Building upon this similarity, an analytical expression that computes the move suppression coefficients for MIMO DMC was proposed. The effectiveness of this important extension was demonstrated by tuning MIMO DMC for control of a Wood and Berry distillation column.

Though not demonstrated here, the tuning strategy presented is also directly applicable to parameteric controllers like GPC, which employ an autoregressive moving average-based process model. Application to GPC should be fairly straightforward because the final control law and the tuning parameters of DMC and GPC are similar. For implementing the tuning strategy with GPC, the weighting polynomial that modifies the predicted output profile must additionally be set at unity. Hence, for those working with DMC or GPC this paper provides some useful insight into controller tuning and selection of λ.

Important stability issues have been addressed only indirectly. Based on rigorous stability studies by others, a large prediction horizon, P, and a control horizon, M, greater than 1 are recommended to ensure nominal closed-loop stability. Robustness of the final MPC design has also not been investigated in this work. Robustness becomes a critical design issue with greater uncertainty in the internal model of the controller. By specifying a bound on the condition number of the system matrix, $A^{\mathrm{T}}A$, some degree of robustness is assured by preventing errors in the formulation of the dynamic matrix to be pronounced in the computed manipulated input moves.

Nomenclature

a = constant
a_i = ith unit step response coefficient
$\tilde{a}_i$ = ith gain-scaled unit step response coefficient
A = dynamic matrix
$\tilde{A}$ = gain-scaled dynamic matrix
$A^{\mathrm{T}}A$ = DMC system matrix
$\tilde{A}^{\mathrm{T}}\tilde{A}$ = gain-scaled system matrix
b = constant
c = condition number
c_i = ith term of the pseudoinverse matrix
$\tilde{c}_i$ = ith term of the gain-scaled pseudoinverse matrix
d = lumped "disturbance"
$D(z)$ = discrete controller transfer function
$\tilde{D}(z)$ = gain-scaled discrete controller transfer function
e = predicted error
$\bar{e}$ = vector of predicted errors
$\bar{e}_j$ = jth orthonormal basis vector for the approximate $\tilde{A}^{\mathrm{T}}\tilde{A}$ matrix
f = gain-scaled move suppression coefficient
$G_{\mathrm{p}}(s)$ = continuous process transfer function
$G_{\mathrm{p}}(z)$ = discrete process transfer function
$\tilde{G}_{\mathrm{p}}(z)$ = gain-scaled discrete process transfer function
h_j = jth unit impulse response
h_j = jth basis vector for the approximate $\tilde{A}^{\mathrm{T}}\tilde{A}$ matrix
H_0 = zero-order hold
i = index

I = identity matrix
j = index of sampling instants
k = discrete dead time
K_p = process gain
M = control horizon (number of manipulated input moves)-
n = current sampling instant
N = model horizon (process settling time in samples)
P = prediction horizon
Q = equal concern factor (controlled variable weight)
Q_V = vapor boil-up rate
$\mathbf{R}$ = upper triangular and invertible QR factored matrix
$\mathcal{R}$ = number of measured outputs
R_F = reflux flow
s = Laplace domain operator
S = number of manipulated inputs
t = sample time
u = process input
X_B = mole fraction of methanol in bottoms
X_D = mole fraction of methanol in distillate
y = measured output
y_0 = initial condition of measured output
$\hat{y}$ = predicted output
y_{sp} = output set point
z = discrete time shift operator

Greek Symbols

Δu_i = change in manipulated input at the ith instant
θ_p = effective process dead time
λ = move suppression coefficient (manipulated input weight)
μ = eigenvalue
μ_{min} = minimum absolute eigenvalue of the approximate $\tilde{\mathbf{A}}^T\tilde{\mathbf{A}}$ matrix
μ_{max} = maximum absolute eigenvalue of the approximate $\tilde{\mathbf{A}}^T\tilde{\mathbf{A}}$ matrix
τ_c = overall closed-loop time constant
τ_n = natural period of oscillation
τ_p = overall process time constant
ω_B = closed-loop bandwidth
ζ = damping coefficient
γ_i^2 = controlled variable weight (equal concern factor) in MIMO DMC
λ_i^2 = move suppression coefficients in MIMO DMC
$\mathbf{\Gamma}^T\mathbf{\Gamma}$ = matrix of controlled variable weights
$\mathbf{\Lambda}^T\mathbf{\Lambda}$ = matrix of move suppression coefficients

Abbreviations

DMC = dynamic matrix control
GPC = generalized predictive control
ITAE = integral of time-weighted absolute error
IAE = integral of absolute error
MPC = model predictive control
PID = proportional integral derivative

Literature Cited

Asbjornsen, O. A. Feed Forward Predictive and Adaptive Control by the Dynamic Matrix. *Am. Control Conf.* **1984**, *3*, 1864.

Åström, K. J.; Wittenmark, B. In *Computer Controlled Systems, Theory and Design*; Kailath, T., Ed.; Prentice-Hall Information and System Science Series; Prentice-Hall: Englewood Cliffs, NJ, 1984.

Callaghan, P. J.; Lee, P. L. An Experimental Investigation of Predictive Controller Design by Principal Component Analysis. *Chem. Eng. Res. Des.* **1988**, *66*, 345.

Camacho, E. F.; Bordons, C. *Model Predictive Control in the Process Industry, Advances in Industrial Control*; Springer-Verlag: London, 1995.

Clarke, D. W. *Advances in Model-Based Predictive Control*; Oxford University Press: New York, 1994.

Clarke, D. W.; Scattolini, R. Constrained Receding Horizon Predictive Control. *IEE Proc.* **1991**, *138* (4), 347.

Clarke, D. W.; Mohtadi, C.; Tuffs, P. S. Generalized Predictive Control-I. The Basic Algorithm. *Automatica* **1987a**, *23*, 137.

Clarke, D. W.; Mohtadi, C.; Tuffs, P. S. Generalized Predictive Control-II. Extensions and Interpretations. *Automatica* **1987b**, *23*, 149.

Cohen, G. H.; Coon, G. A. Theoretical Considerations of Retarded Control. *Trans. ASME* **1953**, *75*, 827.

Cutler, C. R. Dynamic Matrix Control, An Optimal Multivariable Control Algorithm with Constraints. Ph.D. Dissertation, University of Houston, Houston, TX, 1983.

Cutler, C. R.; Ramaker, D. L. Dynamic Matrix Control-A Computer Control Algorithm. *Proc. JACC*; San Francisco, CA, 1980.

Downs, J. J.; Vogel, E. F.; Williams, V. J. Tuning and Sensitivity Study for Model Predictive Control. *Am. Control Conf.* **1988**, *3*, 2357.

Farrell, R. J.; Polli, A. Comparison of Unconstrained Dynamic Matrix Control to Conventional Feedback Control for a First Order Model. *Adv. Instrum. Control* **1990**, *45* (2), 1033.

Franklin, G. F.; Powell, J. D. *Digital Control of Dynamic Systems*; Addison-Wesley Publishing Co.: Reading, MA, 1980.

Froisy, J. B. Model Predictive Control: Past, Present and Future. *ISA Trans.* **1994**, *33*, 235.

García, C. E.; Morari, M. Internal Model Control. I. A Unifying Review and Some New Results. *Ind. Eng. Chem. Process Des. Dev.* **1982**, *21*, 308.

García, C. E.; Morshedi, A. M. Quadratic Programming Solution of Dynamic Matrix Control (QDMC). *Chem. Eng. Commun.* **1986**, *46*, 73.

García, C. E.; Prett, D. M.; Morari, M. Model Predictive Control: Theory and Practice-A Survey. *Automatica* **1989**, *25* (3), 335.

Glova, P. J.; Snowden, D. L.; Yokum, F. H. Multivariable Controllers Applied at Plastics Facility. *Adv. Instrum.* **1990**, *44* (2), 521.

Gupta, Y. P. A Simplified Model Predictive Control Approach. Dept. of Chem. Eng., TUNS: Halifax, Canada, 1987.

Gupta, Y. P. Characteristic Equations and Robust Stability of a Simplified Predictive Control Algorithm. *Can. J. Chem. Eng.* **1993**, *71*, 617.

Hinde, R. F.; Cooper, D. J. Using Pattern Recognition in Controller Adaptation and Performance Evaluation. *Proc. Am. Control Conf., IEEE Publ.* **1993**, 74.

Hinde, R. F.; Cooper, D. J. A Unified Excitation and Performance Diagnostic Adaptive Control Framework. *AIChE J.* **1995**, *41*, 110.

Hoerl, A. E.; Kennard, R. W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **1970**, *12* (1), 55.

Hokanson, D. A.; Houk, B. G.; Johnston, C. R. DMC Control of a Complex Refrigerated Fractionator. *ISA* **1989**, *44* (2), 541.

Lalonde, A. M.; Cooper, D. J. Automated Design, Implementation and Adaptation of the Generalized Predictive Controller. *Proc. Am. Control Conf., IEEE Publ.* **1989**, 1840.

Lee, J. H.; Yu, Z. H. Tuning of Model Predictive Controllers for Robust Performance. *Comput. Chem. Eng.* **1994**, *18* (1), 15.

Lundström, P.; Lee, J. H.; Morari, M.; Skogestad, S. Limitations of Dynamic Matrix Control. *Comput. Chem. Eng.* **1995**, *19* (4), 409.

Marchetti, J. L. Ph.D. Dissertation, University of California, Santa Barbara, CA, 1982.

Marchetti, J. L.; Mellichamp, D. A.; Seborg, D. E. Predictive Control Based on Discrete Convolution Models. *Ind. Eng. Chem. Process Des. Dev.* **1983**, *22*, 488.

Maurath, P. R.; Mellichamp, D. A.; Seborg, D. E. Predictive Controller Design by Principal Component Analysis. *Proc. Amer. Cont. Conf., IEEE Publications*, **1985**, 1059.

Maurath, P. R.; Mellichamp, D. A.; Seborg, D. E. Predictive Controller Design for Single Input/Single Output Systems. *Ind. Eng. Chem. Res.* **1988a**, *27*, 956.

Maurath, P. R.; Laub, A. J.; Mellichamp, D. A.; Seborg, D. E. Predictive Controller Design by Principal Components Analysis. *Ind. Eng. Chem. Res.* **1988b**, *27*, 1204.

McIntosh, A. R. Performance and Tuning of Adaptive Generalized Predictive Control. M.Sc. Thesis, University of Alberta, 1988.

McIntosh, A. R.; Shah, S. L.; Fisher, D. G. Analysis and Tuning of Adaptive Generalized Predictive Control. *Can. J. Chem. Eng.* **1991**, *69*, 97.

McIntosh, A. R.; Fisher, D. G.; Shah, S. L. Performance Adaptive Control: General Structure and a Case Study. *J. Process Control* **1992**, *2* (4), 213.

Middleton, R. H. Tradeoffs in Linear Control System Design. *Automatica* **1991**, *27*, 281.

Morari, M.; Zafiriou, E. *Robust Process Control*; Prentice-Hall, Inc.: Englewood Cliffs, NJ, 1989.

Morari, M.; Lee, J. H. Model Predictive Control: The Good, the Bad and the Ugly. In *Proc. Chem. Process Control Conf., CPC IV*, Fourth International Conference on Chemical Process Control; Arkun, Y., Ray, W. H., Eds.; Elsevier: Amsterdam, Holland, 1991.

Muske, K. R.; Rawlings, J. B. Model Predictive Control with Linear Models. *AIChE J.* **1993**, *39*, 262.

Ogunnaike, B. A. Dynamic Matrix Control: A Nonstochastic, Industrial Process Control Technique with Parallels in Applied Statistics. *Ind. Eng. Chem. Fundam.* **1986**, *25*, 712.

Ogunnaike, B. A.; Ray, W. H. Process Dynamics, Modeling and Control. In *Topics in Chemical Engineering*; Gubbins, K. E., Ed.; Oxford University Press: New York, 1994.

Ohshima, M.; Hashimoto, I.; Takamatsu, T.; Ohno, H. Robust Stability of Model Predictive Control (MPC). *Int. Chem. Eng.* **1991**, *31* (1), 119.

Prada, C.; Serrano, J.; Vega, P.; Piera, M. A. A comparative study of DMC and GPC controllers. In *Advances in Model-Based Predictive Control*; Clarke, D., Ed.; Oxford Science Publications: New York, 1993; p 38.

Prett, D. M.; García, C. E. *Fundamental Process Control*; Butterworths: Boston, MA, 1988.

Qin, J. S.; Badgwell, T. A. An Overview of Industrial Model Predictive Control Technology. *Prepr. Proc. Chem. Process Control Conf., CPC V*; Tahoe City, CA, 1996.

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

Richalet, J. Industrial Applications of Model Based Predictive Control. *Automatica* **1993**, *29* (5), 1251.

Richalet, J.; Rault, A.; Testud, J. L.; Papon, J. Algorithmic Control of Industrial Processes. 4th IFAC *Symposium on Identification and System Parameter Estimation*; Tbilisi, USSR, 1976.

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

Ricker, N. L. Model Predictive Control: State of the Art. In *Proc. Chem. Process Control Conf., CPC IV*, Fourth International Conference on Chemical Process Control; Arkun, Y., Ray, W. H., Eds.; Elsevier: Amsterdam, Holland, 1991.

Scokaert, P.; Clarke, D. W. Chapter: Stability and Feasibility in Constrained Predictive Control. *Advances in Model Based Predictive Control*; Oxford University Press: New York, 1994.

Seborg, D. E.; Edgar, T. F.; Shah, S. L. Adaptive Control Strategies for Process Control: A Survey. *AIChE J.* **1986**, *32* (6), 881.

Soroush, M.; Kravaris, C. Synthesis of Discrete-Time Nonlinear Feed-Forward/Feedback Controllers. *AIChE J.* **1994**, *40*, 473.

Takamatsu, T.; Hashimoto, I.; Ohshima, M.; Ohno, H. Study of model predictive control (MPC) in terms of its structure and degrees of freedom. *Int. Chem. Eng.* **1988**, *28* (4), 661.

Wood, R. K.; Berry, M. W. Terminal Composition Control of a Binary Distillation Column. *Chem. Eng. Sci.* **1973**, *28*, 1707.

Yocum, F. H.; Zimmerman, J. M. Real Time Control of Inverse Response using Dynamic Matrix Control. *Am. Control Conf.* **1988**, *1*, 266.

Zafiriou, E., The Robustness of Model Predictive Controllers. *Proc. CPC IV*, San Padre Island, TX, 1991.