Shaping Time-Domain Responses with Discrete Controllers

Leonardo L. Giovanini and Jacinto L. Marchetti*

Institute of Technological Development for the Chemical Industry (INTEC), CONICET—Universidad Nacional del Litoral, Güemes 3450, 3000 Santa Fe, Argentina

A model-based design and tuning method for discrete controllers for single-input—single-output systems is presented in this article. It stresses the possibility of specifying the desired closed-loop behavior by using a set of time-domain conditions defining the performance that is desirable in the controlled output. The problem formulation is developed in the time domain since this is the domain where many important and different design conditions can be visualized and written in an easy mathematical form, particularly for most continuous chemical processes. This approach allows shaping the time-domain closed-loop response by one or more constraints representing the limits of minimum performance desired when specific changes on either the setpoint or the load disturbance are expected. Model uncertainties caused by slow time-variant or nonlinear systems can also be accounted for, guaranteeing robust performance and stability of the closed-loop system.

1. Introduction

The problem of tuning either continuous or discrete controllers has been widely studied for decades. During all this time many methods for tuning the classical PID controller have been applied to very different process systems, surely the most well-known are the Ziegler–Nichols25 and Cohen–Coon2 rules. More recently, however, Rivera et al.19 introduced a H∞ design approach based on internal model control (IMC) that, among others things, has the advantage of yielding a PI or PID controller with only one tuning parameter. Then, the results are combined with an H∞ adjustment for robust stability and performance.15

For discrete controllers, Isermann8 indicates two main alternative approaches for the synthesis: (i) determining structure-optimized controllers and (ii) determining parameter-optimized controllers. In the first case both the controller structure and the parameters are adapted optimally to the structure and parameters of the process model; cancellation and state controllers are typical examples. In the second case the form and the order of the controller equation are given or arbitrarily adopted, and the parameter values are adjusted to the process model using an optimizing criterion. Common parameter optimized controllers are P, PI, or PID, where if small sampling times are used, the settings obtained for continuous controllers (tuning rules) can be used to determine the discrete controller parameters, slightly modified to consider the effective time delay caused by the zero-order hold.23

Discrete feedback controllers can also be designed using pole placement in the z domain;6 however, little is specified in the time domain about the behavior of the system variables.

We cannot avoid mentioning the remarkable importance given also to the linear quadratic problem (LQP).9 A consequence of the interest in this approach is that many of its properties have already been revealed; the most important is that it yields a linear controller that stabilizes the closed-loop system. However, when incorporating new restrictions in the optimization problem, most of its properties disappear. Furthermore, the presence of restrictions makes difficult the stability analysis for the closed-loop system.13,22

Many discussions about performance specifications focus on mathematical approaches stemming from space vector norms and frequency domain, leaving in a second place very familiar and intuitive time-domain alternatives. In this article we show that it is not only possible but it also easy to design and adjust a discrete controller using design conditions that cannot be defined by a single performance function. The proposed technique stresses the possibility of specifying the desired closed-loop behavior by using a set of time-domain conditions defining limits of minimum performance that are desirable in the controlled output.

The problem formulation is developed in the time domain since this is the domain where many important and different design conditions can be visualized and written in an easy mathematical form, particularly for most continuous chemical processes. This approach allows shaping the time-domain closed-loop response by one or more constraints representing the limits of minimum performance desired when specific changes on either the setpoint or the load disturbance are expected. Furthermore, since model uncertainties can be considered using a multimodel representation, lately referred to as the polytopic paradigm,5,10 a simple extension of the method can also guarantee the robust stability and performance for the closed-loop system.

The organization of this paper is as follows: In section 2, the linear discrete models are recalled and the extension to multimodels or polytopic models is commented on. Section 3 refers to discrete controllers and their main structural features. In section 4, the basic mathematical problem formulation is presented and the meanings of some design parameters are discussed. Furthermore, different performance specifications including the closed-loop stability condition are analyzed and comments about the initialization of the optimizing algorithm are made. Section 5 shows the results obtained from the application of the proposed method to selected examples, including a nonlinear continuously
2. Discrete Process Modeling

Typically, controllers are adjusted to meet performance requirements during setpoint changes or disturbance rejection. The adjustment must be made for a specific process system; i.e., each application requires a particular design or at least a particular tuning. This fact turns our attention to the availability or the development of a model of the involved process to describe its dynamics and consequently for controller design or controller adjustment. For the time being, model-based process control has become an almost exclusive strategy for a rational approach for controller designing or controller tuning.

As we restrict ourselves to sampled-data systems, the possibilities of representing process dynamics are still important. Difference equations, discrete convolutions, and discrete state space representations are different but, most of the times, equivalent forms of representing discrete linear systems. In this article, we mainly deal with difference equations of the form

\[ y(k) + \sum_{j=1}^{n} a_j y(k-j) = \sum_{i=1}^{m} b_i u(k-i-k_d) \]  

(2.1)

where \( y \) and \( u \) are the output and input variables, respectively, and \( k_d \) stands for the time delay.

The design and tuning method discussed in this article can be supported by today's available computational capabilities and software packages oriented to general optimization problems. The approach is not dependent on the process model structure, which is a very frequent important condition in most design and adjusting procedures. Linear or nonlinear process representations in the form of discrete or continuous sampled process models might be used for determining an appropriate controller order and satisfactory parameter values. Though the idea can be extended to general dynamic simulators for representing the process, we do not address this issue in this article.

If the available model is linear, in any of the possible representation forms, the formulation of the optimization problem is straightforward. If we have to deal with a nonlinear system, the use of a multiple-model representation (see Appendix I) combined with concepts coming from robust control theory have shown to be computationally efficient and to yield appropriate solutions for controller adjustment.

3. Controller Structure

Designing a controller implies giving an answer to the following two questions: (i) how to determine the controller structure and (ii) how to determine the controller parameter values. One approach to solve the problem is by optimizing both the structure and parameters, simultaneously. This leads, for instance, to state controllers and to cancellation controllers.  

Other alternatives suggest the adoption of the general input–output structure, in this case, for the discrete linear controller:

\[ C(z) = \frac{u(z)}{e(z)} = \frac{Q(z^{-1})}{P(z^{-1})} = \frac{\sum_{i=0}^{w} q_i z^{-i}}{1 + \sum_{j=1}^{v} p_j z^{-j}} \]  

(3.1)

The selection of \( w \) and \( v \) may follow some practical criterion like the simple requirement of looking for low-order implementations, the most typical case being when adopting PI or PID controllers. In this article we show a convenient way for defining the number of parameters to be used and their adjustment to reach a particular performance objective.

Let us assume for the moment that the process is represented by (2.1), or equivalently

\[ G_p(z) = \frac{y(z)}{u(z)} = \frac{B(z^{-1})}{A(z^{-1})} = \frac{\sum_{i=1}^{m} b_i z^{-i}}{1 + \sum_{j=1}^{n} a_j z^{-j}} \]  

(3.2)

Then, \( w = n \), or \( v = m + k_d \), are the lowest order values that enable the complete placement of the closed-loop poles (see Appendix II). If these equalities hold, the coefficients of the closed-loop characteristic equation give a system of linear equations with a unique solution, where the controller parameters are the unknown variables, and all the closed-loop poles can be located as desired.

Moreover, a default condition is frequently required for (3.1) to avoid steady-state offsets, i.e.,

\[ P(1) = 1 + \sum_{j=1}^{v} p_j = 0 \]  

(3.3)

If arbitrarily adopting \( v = 1 \), eq 3.3 asks for \( p_1 = -1 \), and (3.1) gives a controller having a single integrating pole at \( z = 1 \), and \( w - 1 \) poles at the origin,

\[ C(z) = \frac{Q(z^{-1})}{(1 - z^{-1})(1 - z^{-1})} = \frac{\sum_{i=0}^{w} q_i z^{-i}}{(1 - z^{-1})(1 - z^{-1})} \]  

(3.4)

This structure includes the discrete version of the very familiar controllers PI and PID, which are obtained by restricting the order \( w \) to 1 or 2, respectively, and the parameters values to the following conditions, for positive gain:

\[ -q_0 < 0 \]
\[ q_1 < -q_0 \]
\[ q_2 < q_0 \]
\[ -(q_0 + q_1) < q_2 \]  

(3.5)

However, approaches involving optimization procedures might give results not necessarily constrained by the
above conditions, thus allowing performances otherwise unreachable by the PID controller.

4. Design-and-Tuning Problem Formulation

4.1. Basic Case. Our proposal in this article is to determine the minimum controller orders, w and v, that are necessary to meet a desired performance, and to adjust all the parameters in $Q(z^{-1})$ and $P(z^{-1})$ by a pseudo-optimization procedure until a set of time-domain specifications are met.

Let us assume a discrete model of the system is available and that there is no model mismatch, for the moment. There are at least two controller design conditions almost always required: (i) closed-loop stability and (ii) the minimization of some performance or objective function $J$.

The general basic formulation of a discrete-controller optimization problem could be described by a mixed-integer-nonlinear problem (MINLP) where the integer variables to be optimized are the polynomials orders $v$ and $w$. However, since these two parameters are determined among a reduced set of values, a direct search is recommendable. This approach decomposes the original problem in few NLPs that can be sequentially solved until a satisfactory result is obtained. Hence, for each $(v, w)$ pair, we have the following NLP formulation:

$$\min_{q, b} J[r(k), y(k), u(k)], \quad \forall k \in [1, N], \quad i \in [0, w], \quad j \in [1, v] \; (4.1)$$

s.t.

$$y(k) = F[y(k - \zeta), u(k - \gamma - k_d), d(k)], \quad \forall k \in [1, N]$$

$$u(k) = \sum_{j=1}^{v} p_j u(k - j) + \sum_{i=0}^{w} q_i [r(k - i) - y(k - i)], \quad \forall k \in [1, N]$$

$$u(N) = u(N - 1)$$

$$y(0) = 0$$

$$u(0) = 0 \; (4.2)$$

where $N$ is the overall number of sampling instants being considered.

Rather than defining the final desired performance, like in more standard tuning methods, in this approach $J(\cdot)$ stands for any convenient objective function capable of guiding the parameter adjustment. The desired closed-loop performance is defined through additional restrictions as discussed in a following subsection.

The first equation in (4.2) is a general expression of the discrete process model which remarks that any type of representation may be used at this point; $y$ and $u$ are the controlled and the control variables, respectively; $\zeta$ and $\gamma$ implies that historical values might be included, and the displacement $k_d$ is the system time delay. Besides, $d$ is a disturbance, which may be defined through an analytical expression or by a time series representing any practical or realistic change.

The second equation represents a linear controller with parameters $q$, $i \in [0, w]$ and $p$, $j \in [1, v]$, where $r$ is the setpoint. Finally, an additional constraint is suggested in this formulation through the third equation in (4.2), which deserves specific comments in the next subsection. The initial conditions for $y$ and $u$ complete the basic formulation.

4.2. Final Control Condition. The third equation in (4.2) asks for null or negligible control movement at the end of the time interval being considered. We show in Appendix III that this condition forces the stability of the closed loop and that it is equivalent to requiring both $y$ and $u$ to remain constant after the time instant $k = N$. This constraint may also be used as a design condition affecting the closed-loop performance: defining the value of $N$ is an alternative way for determining the closed-loop settling time. The larger this parameter value is, the lesser the controller sensitivity will be and slower the closed-loop dynamics will look.

However, the effect of this constraint on the controller adjustment is better visualized recalling concepts related to complete output controlable systems. Any complete output controlable time-invariant discrete system, with unbounded input, can be transferred from any initial output to any final output in a finite number of steps. In particular, if this number of steps is equal to the dimension of the system and we require both $y$ and $u$ to remain constant thereafter, then the output variable exhibits deadbeat response. It also implies that all the closed-loop poles are at the origin of the unit circle.

From Appendix III, the final control condition forces a deadbeat response for the lower bound $N = n + v + k_d$, and when this condition is numerically relaxed with a small $\varepsilon$ gap, the closed-loop poles move away from the origin as $N$ increases. In other words, if $N = n + v + k_d$ is adopted, a deadbeat type of response is obtained as long as the control input remains unbounded, and no other design condition is imposed. As $N$ increases, the room for combining special design specifications increases too, but still requiring the closed-loop poles to stay inside the unit circle. Issues related to internal stability are discussed in the next subsection.

Hence, the final control condition applied at a time instant larger than the closed-loop system dimension helps to select feasible solutions with bounded input/output trajectories and consequently it speeds up the numerical convergence. Furthermore, it avoids oscillations or ripples between sampling points, a problem frequently arising from discrete cancellation controllers, particularly when applied to high-order systems.

4.3. Internal Stability Condition. Internal stability is a basic requirement for a practical feedback system discussed by Morari and Zafiriou and lately by Zhou and Doyle among others. The convenience of considering concepts associated with internal stability during the tuning problem formulation is quite obvious. In this regard, observe that the objective function typically requires the controlled output, $y(k)$ to follow a bounded sequence, $r(k)$, or alternatively, a specific design might ask for the rejection of a bounded disturbance signal, $d(k)$. Besides, the third constraint forces the control signal $u(k)$ to also be bounded. However, the resulting closed-loop will be internally stable only if there is no cancellation of process poles or zeros located outside the unit circle. If the process model is given by (3.2), the possibility of a cancellation like this during the numerical tuning can be easily detected by a preliminary inspection of $A(z) = 0$ and $B(z) = 0$.

If $A(z) = 0$ shows an unstable root at $z = \pi$ and the above basic tuning formulation is applied for a setpoint
change, the possibility of cancelling this pole with a controller zero exists. In this case, a small external disturbance in the control variable \( u(k) \) will reveal the unstable condition, exposing an unbounded \( y(k) \). To prevent this undesired design, we suggest the additional constraint

\[
\sum_{i=0}^{w} q_i(T)^{w-i} > \delta \quad (4.3)
\]

where \( \delta \) is a positive value defining a prohibit zone for the controller zeros around the location \( z = \xi \). Eventually, the magnitude of \( \delta \) could be related to the parametric uncertainty affecting this particular root.

Similarly, if \( B(z) = 0 \) shows a root \( z = \xi \) outside the unit circle or very close to it, and the above basic tuning formulation is applied for rejecting a load change on \( u(k) \), the possibility of cancelling the process zero with a controller pole exists. A small change in the setpoint will expose the problem, showing an unbounded \( u(k) \). This can be prevented by including the constraint

\[
\sum_{j=1}^{v} p_j(\xi)^{v-j} > \delta \quad (4.4)
\]

where \( \delta \) deserves the same interpretation given above.

Hence, if the process model presents either a zero or a pole outside the unit circle, then the problem formulation should include the constraint \((4.4)\) or \((4.3)\), respectively, to guarantee internal stability.

### 4.4. Performance Specification

The choice of the performance function, for measuring the quality of the closed-loop response to a given reference change or load disturbance, has played a central role in controller design. Very familiar are the definitions of the integral criteria ISE (integral of the square error), ITAE (integral of the absolute value of the error), and ITAE (integral of the two-weighted absolute error), which become sum criteria for discrete signals. Also, very frequently the ISE index is extended to include the energy delivered by the control variable through an arbitrary weight coefficient. Many resulting tuning formulas presented in the literature are based on one of these criteria, assuming a simple process model and a step change in the setpoint or the load disturbance.20 However, most of them have shown some limitations for including explicitly others design conditions like rise time, settling time, decay ratio, maximum or no overshoot, allowed undershoot, etc. Notice that using a single performance index only gives no clear information about the output response or the controller performance until the system is simulated or the controller is commissioned.

It is possible and many times desirable that the system response achieve more than one single requirement. To this end, a technique has been presented on the basis of the use of a single comprehensive index of performance.7 Though the technique seems to have low sensitivity to the starting values used for the optimization, the final approach includes a sort of trial and error until the most favorable values are selected. An alternative as been the use of a multiobjective programming approach1 aimed at finding Pareto optimal solutions characterized as a solution set in which any improvement in one performance criterion necessarily implies the degradation in some other. This last technique has to also solve a number of optimization problems, in an iterative fashion, to adjust a controller having a fixed predetermined structure.

Most of the times the meaning of “good control performance” in process control involves several desirable features different from the absolute minimum of the integral of some function of the error. From a practical point of view, the “additional” design conditions could be more important than reaching the absolute minimum of the objective function.

Going back to the optimization problem formulated above, observe that a sort of closed-loop simulation takes place each time the NLP tries new parameter values, and therefore, the closed-loop time-domain response is directly evaluated as well as the evolution of all the internal variables. Hence, additional design specifications different from those included in the objective function can also be included as well as performance constraints on variables other than the controlled output. Furthermore, during the numerical adjustment these constraints are met before finding the minimum of the objective function—the opposite case implies there is not feasible solution. Then, the NLP optimization may be stopped once the design conditions are satisfied. This does not mean to reduce the importance of the global minimum; it simply provides a practical criterium for stopping the numerical tuning.

Under this framework we certainly can use the ISE (or \( I_2 \)) for guiding the tuning process,

\[
J = ||e||^2 = \sum_{k=1}^{N} e(k)^2
\]

Alternatives such as the sum of the absolute error or the sum of the time multiplied by the absolute error might result in similar effectiveness. The possibility of a direct influence for damping the control variable by adding the quadratic deviation of the manipulated variable is also well-known. The use of the discrete version of the infinite norm of the error (\( l_{\infty} \)),

\[
J = ||e||_{\infty} = \sup_{0 \leq |k| \leq N} |e(k)|
\]

might be interesting when adjusting the controller to load disturbances rather than to set point changes.

Less conventional—but many times more important—objective functions can be used as additional constraints or shaping constraints. These functions are typically defined to measure certain attributes of the temporal response such as the rise time \( t_r \) and settling time \( t_s \) and the overshoot \( y_{os} \) or undershoot \( y_{us} \) and to account for a physical control constraint \((|u|_{max})\); in particular, for the positive-gain case they can be written as follows:

\[
t_r = \inf_{k} t_r(y(k) \geq y(N); 1 \leq k < N)
\]

\[
t_s = \inf_{k} t_s(|e(k)| \leq \varepsilon; 1 \leq k \leq N)
\]

\[
y_{os} = \sup_{k} y(k); 1 \leq k < N
\]

\[
y_{us} = \sup_{k} (-y(k); 1 \leq k < N)
\]

\[
|u|_{max} = \sup_{k} |u(k)|; 1 \leq k < N
\]

Similar functions have been suggested in connection with a multiobjective design,1 for optimal design,4 and pole placement.6,18 In this article we show that they can
be used as shaping constraints for reaching the desired performance in the tuning problem formulation.

However, conflicting conditions could be simultaneously required during the controller design. A set of individually admissible constraints might define an empty solution space if the final problem formulation is not consistent with dynamic process limitations, which means that the process system will admit realistic performance requirements only.\textsuperscript{14,21}

4.5. Computing Approach. In the introduction to this section we characterize the general basic formulation as a MINLP since v and w-two integer variables—are also design parameters to be optimally determined. Hence, a standard NLP solver is used to find the controller parameters values for a given pair (v, w), while a simple direct and guided search is used to determine the lowest v and w combination that meets the design conditions. The NLP readjusts the controller until all the design conditions are simultaneously satisfied and stops even though the objective function J(\cdot) does not achieve the absolute extreme. If any restriction is violated, the direct search changes the controller structure, changing v and w by 1 until the minimum required values are reached or the improvement is meaningless. In this way, the result gives the lowest order controller capable to meet the required performance. This controller might not be optimum from a rigorous or more traditional point of view, but it will certainly satisfy the designer specifications.

Furthermore, something must be said regarding numerical initialization. If the controller is given by (3.1) and the plant dynamics are represented by (3.2), without losing generality we assume \( n = m \), then the characteristic equation takes the general form

\[
P(z)z^{\frac{n}{2}} - v^{\frac{n}{2}} A(z) + Q(z)B(z) = 0 \quad (4.8)
\]

Since a stable closed-loop solution is desired, the convenient initial parameter values must locate all the closed-loop poles into the unit circle. In particular, for stable plants, v and w may be lower than n, and the easiest and safe initialization is \( p_i = -1/v, \forall v \in [1, v]; q_i = 0, \forall i \in [0, w]. \) These parameter values locate the roots of \( P(z) = 0 \) inside the unit circle,\textsuperscript{3} meet the (3.3) condition, and locate \( k_d + w - v \) closed-loop poles at the origin; the remaining poles are the stable open-loop process poles. This initialization helps to provide stability to the numerical search, which is equivalent to a sequence of dynamic simulations, and leads to rapid convergence.

For unstable systems the initial values must be carefully chosen, since an inadequate selection might lead to an unfeasible solution space. If we desire the capability to locate stable closed-loop poles arbitrarily, the initial controller structure is determined by \( w = n \) and \( v = m + k_d \). However, stable solutions can be found for a lower number of controller parameters, mainly due to the final control condition plus the internal stability condition discussed above. In any case, we have to propose an initial stable pole configuration and determine the first trial coefficients \( q_i \) and \( p_j \) by solving the system of linear equations resulting from (4.8).

5. Applications

The full potential of the tuning method proposed in this article comes from the possibility of including as many time-domain design conditions as desired, as long as they are consistent with the system limitations and they do not create conflicting requirements. In this section we show this feature through several application examples, first by working on simple linear systems and then controlling the operation of a continuously stirred tank reactor.

5.1. Linear Systems. To introduce the application example, let us see a very typical problem where a discrete linear model of the process (3.2) is available and a controller like (3.1) is adopted. If the design requires (i) offset elimination for setpoints changes and (ii) closed-loop stability assuming there is no model mismatch, then the basic NLP formulation of an off-line parameter search guided by the ISE function can be described as follows:

\[
\min \left\{ \sum_{k=1}^{N} [r(k) - y(k)]^2 \right\}, \quad \forall k \in [1, N], \quad i \in [0, w],
\]

\[
\text{s.t.}
\]

\[
y(k) = -\sum_{j=1}^{u} a_j y(k - j) + \sum_{i=1}^{m} b_i u(k - i - k_d) + d(k),
\]

\[
\quad \forall k \in [1, N]
\]

\[
u(k) = -\sum_{j=1}^{v} p_j u(k - j) + \sum_{i=0}^{w} q_i [r(k - i) - y(k - i)]
\]

\[
\quad \forall k \in [1, N]
\]

\[
u(N) = u(N - 1)
\]

\[
\sum_{j=1}^{v} p_j = -1
\]

\[
y(0) = 0
\]

\[
u(0) = 0
\]

Though \( v \) and \( w \) are fixed for this formulation, they are also unknown variables in the general problem. Note that \( w = n \) and \( v = m + k_d \) are the reasonable values to be used in the first trial (see Appendix II); however, there are many cases where lower values meet the needs.

Example 1. Let us consider a continuous process showing inverse response to input changes and whose dynamics are represented by the following transfer function:

\[
G(s) = \frac{1 - 5s}{(10s + 1)^2}
\]

The problem is to find a controller such that the following performance can be achieved: (i) the error be lower than \( \Delta = 0.05 \) about 30 s after a step change in setpoint is made; (ii) the overshoot of the system output, \( y_{os} \) must be null; (iii) its undershoot, \( y_{us} \) must be smaller than 0.25; and (iv) the final offset must be null.

Since the system has a real right-half plane open-loop zero at \( \zeta_2 = 0.2 \) and we require a settling time of \( T_{sett} = 30 \) s, the minimum undershoot that could be expected...
if there is no constraint on the manipulated variable\(^{14}\) is

\[
y_{us} \geq \frac{\Delta}{\exp(s_T \cdot T_{set}) - 1} = 0.0024 \tag{5.4}
\]

This value is smaller than 0.25, the maximum undershoot specified, which means that the first and the third design conditions are consistent with the process dynamics. As expected, however, the discrete process transfer function shows a zero outside the unit circle, \(\xi = 1.22\), whose cancellation must be avoided during the numerical synthesis by including the constraint (4.4).

Hence, the specification for the design can be written as follows:

\[
y(k) \leq r(k) - 0.05, \quad N_0 + 30 \leq k \leq N
\]

\[
y(k) \leq r(k), \quad N_0 \leq k \leq N
\]

\[
y(k) \geq -0.25, \quad N_0 \leq k \leq N_0 + 30
\]

\[
|v(1.22)^v + \sum_{j=2}^{v} p_j (1.22)^{v-j}| > 0.1
\]

which are added to the basic tuning problem (5.1–5.2), where \(N_0\) is the sampling instant at which the setpoint change is made. The bound 0.1 in the fourth condition of (5.5) is arbitrary at this point.

For completeness, this problem still requires the following definitions: (i) the adopted sampling time is \(T_s = 1\) s, and assuming a zero-order hold for converting the transfer function (5.3) to the \(z\) domain, the resulting difference equation has \(n = m = 2\), and \(k_d = 0\); (ii) the time length considered is \(N = 180\) sampling intervals (common sense indicates this should be at least as long as the open-loop process settling time). The input change is a unit step at \(N_0 = 20\), and no perturbation, \(d(k) = 0\), \(\forall k\), is included.

The resulting tuning problem is solved using a gradient-based optimization algorithm. The structure that allows a complete closed-loop pole placement (\(w = 2, v = 2\)) has no difficulty satisfying the required conditions. The parameters values obtained for this controller (3.1) are

\[
w = 2, \quad q_0 = 4.368, \quad q_1 = -7.877, \quad q_2 = 3.548
\]

\[
v = 2, \quad p_0 = 1, \quad p_1 = -1.451, \quad p_2 = 0.451 \tag{5.6}
\]

However, it becomes instructive exploring the sensitivity of the controller structure, in particular, the sensitivity to a reduction in the parameter \(v\). The case \(v = 1\) yields a PID controller despite the constraints in (3.5) not being included; this can be readily verified from the following values:

\[
w = 2, \quad q_0 = 5.740, \quad q_1 = -10.021, \quad q_2 = 4.345
\]

\[
v = 1, \quad p_0 = 1, \quad p_1 = -1 \tag{5.7}
\]

Figure 1 shows the closed-loop responses obtained using the above controllers; it is apparent from the plots in Figure 2. This result is important to exemplify that an optimal ISE tuning, for instance, does not necessarily yield the desired output response.

**Example 2.** Now, let us consider a continuous process showing unstable open-loop behavior whose dynamics are represented by the following transfer function model:

\[
G_p(s) = \frac{1}{(10s - 1)^2} \tag{5.8}
\]

The problem is to find a discrete stabilizing controller such that the closed-loop response to a unit step change in the setpoint satisfies the following characteristics: (i) the system output overshoot, \(y_{os}\), be lower than 25%; (ii) the error magnitude be lower than 0.05 just 30 s after a step change in setpoint is made, and (iii) the steady-state offset be null.
Because the open-loop continuous system has a real right-half-plane pole at \( s_p = 0.1 \) and the discrete implementation introduces a time delay of one sampling interval \( T_s = 1 \text{ s} \), the system response must have an overshoot. Middleton\(^{14}\) shows that the overshoot expected using a stable unity feedback satisfies

\[
y_{os} \geq s_p T_s = 0.1 \tag{5.9}
\]

which is lower than the maximum specified \( y_{os} \leq 0.25 \).

Furthermore, the discrete process transfer function shows in this case an unstable pole at \( \pi = 1.10 \). Thus, the convenience of preventing a possible cancellation is apparent, and this is done by including the constraint \((4.3)\). Hence, the additional design conditions are written as follows:

\[
y(k) \leq 1.25, \quad N_0 \leq k \leq N
\]

\[
y(k) \geq r(k) - 0.05, \quad N_0 + 30 \leq k \leq N
\]

\[
y(k) \leq r(k) + 0.05, \quad N_0 + 30 \leq k \leq N \tag{5.10}
\]

\[
|\sum_{i=0}^{w} q_i(1.10)^{w-i}| > 0.1
\]

For this case \( N, N_0, \) and \( d(k) \) have the same values as those adopted for Example 1.

Since \( n = m = 2 \), then \( w = v = 2 \) defines the first-trial structure for the controller. Again, using a gradient-based optimization algorithm to solve the problem formulated above, a feasible solution space is found and the ISE index is minimized to 1.11.

\[
w = 2, \quad q_0 = 236.39, \quad q_1 = -432.98, \quad q_2 = 197.88
\]

\[
v = 2, \quad p_0 = 1.0, \quad p_1 = -0.025, \quad p_2 = -0.975 \tag{5.11}
\]

Though these parameter values yield a closed-loop response satisfying all the constraints defined by the design conditions, undesired oscillations are observed as a consequence of excessive movements in the control variable. Since minimum ISE is not a main feature required to the adjustment, an additional condition is included to limit the control energy:

\[
\sum_{k=1}^{N} \Delta u^2(k) \leq 250 \tag{5.12}
\]

In this case the parameter values obtained are

\[
w = 2, \quad q_0 = 110.95, \quad q_1 = -206.42, \quad q_2 = 96.18
\]

\[
v = 2, \quad p_0 = 1.0, \quad p_1 = -0.536, \quad p_2 = -0.464 \tag{5.13}
\]

and the ISE index goes up to 1.33. Figure 3 shows the output variable in the upper part of the response to the setpoint change and Figure 4 shows the correspondent manipulated movements obtained using controllers \((5.11)\) and \((5.13)\).

An additional comment in regard to the control variable is that if neither the final control condition nor uncertainty function is used, the perturbation \( d \) in the basic formulation \((5.1-5.2)\) can be determined by \( d(\tau) = \tilde{m}(\tau)\tilde{G}_d(\tau)u(\tau) + \tilde{d}(\tau) \), where \( \tilde{G}_d(\tau) \) is the nominal model and \( \tilde{d}(\tau) \) is the load disturbance to be rejected. If this additional information is used, the resulting controller would provide robustness to the required closed-loop performance.\(^{15}\) Alternatively and more simply, a set of models determined in the neighborhood of the nominal operating point can be directly used.

\[
\text{Figure 3. Closed-loop responses of the open-loop unstable linear system (5.8) to a unit step change in setpoint showing the effect of using an energy constraint on the control variable under shaping constraints in (5.10).}
\]

\[
\text{Figure 4. Control variable correspondent to the closed-loop responses shown in Figure 3.}
\]

\[(5.12)\] are included in the formulation, then the result gives

\[
w = 2, \quad q_0 = 240.1, \quad q_1 = -438.4, \quad q_2 = 200.1
\]

\[
v = 2, \quad p_0 = 1.0, \quad p_1 = 0.0, \quad p_2 = -1.0 \tag{5.14}
\]

i.e., the controller shows a rippling pole at \( \pi = -1.0 \).

5.2. An Approach for Nonlinear Systems. The effectiveness of the method and philosophy being proposed for designing and tuning discrete controllers is quite clear from the above results. However, the always existing model-mismatch problem was not considered so far. This problem is certainly present any time the process has nonlinear characteristics and we attempt to use a linear representation. Assuming that a family of linear transfer functions is capable to capture a moderate nonlinearity and if a global multiplicative uncertainty function is used, the perturbation \( d \) in the basic formulation \((5.1-5.2)\) can be determined by \( d(\tau) = \tilde{m}(\tau)\tilde{G}_d(\tau)u(\tau) + \tilde{d}(\tau) \), where \( \tilde{G}_d(\tau) \) is the nominal model and \( \tilde{d}(\tau) \) is the load disturbance to be rejected. If this additional information is used, the resulting controller would provide robustness to the required closed-loop performance.\(^{15}\) Alternatively and more simply, a set of models determined in the neighborhood of the nominal operating point can be directly used.
Let us include model uncertainty in the basic formulation of section 5.1, for the case in which the process system is not linear. Suppose the uncertainty set $\Omega$ is defined by a polytope (see Appendix I) made of a set of difference equations. The basic problem formulation is now written as follows:

$$\min_{q,p} f(r(k), y_i(k), u_i(k)) \quad \forall k \in [1, N], \; i \in [0, w],$$

$$j \in [1, v] \quad (5.15)$$

s.t.

$$y_i(k) = -\sum_{j=1}^{n} a_{ij} y_i(k - j) + \sum_{i=1}^{w} b_{ij} u_i(k - i - k_d) + d(k),$$

$$\forall k \in [1, N]$$

$$u_i(k) = -\sum_{j=1}^{v} p_j u_i(k - j) + \sum_{i=1}^{w} q_{ij} [r(k - i) - y_i(k - i)],$$

$$\forall k \in [1, N]$$

$$u_i(N) = u_i(N - 1)$$

$$\sum_{j=1}^{v} p_j = -1$$

$$y_i(0) = 0$$

$$u_i(0) = 0 \quad (5.16)$$

where $l \in [1, M]$ stands for a vertex model and $M$ is the number of models being considered.

The above problem consists of the original restrictions (4.1) and (4.2) for each model, so that the mathematical program has $M(3N - K) + 1$ restrictions in which the controller parameters are the independent variables to optimize. In this case, the resulting controller stabilizes the closed loop and gives the best possible performance, satisfying all the models simultaneously, which is equivalent to saying that the solution to this problem gives a controller that provides robust performance to the closed-loop system.

**Example 3.** Let us consider the problem of controlling a continuously stirred tank reactor (CSTR) in which an irreversible exothermic reaction is carried out at constant volume (see Appendix III). This is a nonlinear system previously used by Morningred et al. for testing predictive control algorithms. Figure 5 shows the dynamic responses to the following sequence of changes in the manipulated variables $q_c$: +10, −10, −10, and +10 L min$^{-1}$, where the nonlinear nature of the system is apparent.

Four continuous linear models are determined using least-squared procedures to adjust the composition responses to the above four step changes in the manipulated variable. Notice that those changes imply three different operating points corresponding to the following stationary manipulated flow rates: 100, 110, and 90 L min$^{-1}$. Table 1 shows the process transfer functions obtained:

$$\frac{C_A(s)}{q(s)} = G_{p_l}(s), \; l = 1-4$$

They define the polytopic model associated with the nonlinear behavior in the operating region being considered. This representation should be associated with the $M$ vertex models in the above problem formulation.

Like in Morningred's work, the sampling time period was fixed in 0.1 min, which gives about four sampled-data points in the dominant time constant when the reactor is operating in the high-concentration region. Then, four fourth-order discrete linear models are used for representing the nonlinear reactor using the polytope idea. These models are obtained by $z$-transforming the continuous transfer functions and assuming a zero-order-hold device is included.

The discrete controller adjusted in this case is (3.4), i.e., $v = 1$ is arbitrarily adopted and since the zero-offset condition is included in the controller structure from the beginning ($p_0 = 1$, $p_1 = -1$), the fourth constraint in (5.16) becomes redundant. Since $n = 4$ and $m = 3$ in all the vertex models, an independent one-to-one closed-loop pole location cannot be made for them, but this is not mandatory from a practical point of view. In this application we stress the fact that the reactor operation becomes very sensitive once the manipulated variable exceeds 113 L min$^{-1}$. Hence, assuming a hard constraint is physically used on the coolant flow rate at 110 L min$^{-1}$, an additional restriction for the more sensitive model (model 1 in Table 1) must be considered for the deviation variable $u(k)$:

$$u_i(k) \leq 10 \quad (5.17)$$

This assumes that the nominal absolute value for the manipulated variable is around 100 L min$^{-1}$ and that the operation is kept inside the polytope whose vertices are defined by the linear models. The constraint (5.17) is then included in (5.16).

Notice in this case that it is the polytope or convex hull that must be shaped along the time being considered. Hence, the objective function necessary for driving the adjustment must consider all the linear models simultaneously. At a given time instant and operating point, there is not clear information about which model is the convenient one for representing the process. This is because it depends not only on the operating point but also on which direction the manipulated variable is going to move. The simpler way to solve the problem...
is by proposing the general form

\[ f = \sum_{i=1}^{M} \gamma_i f_i \]  

(5.18)

where the \( \gamma_i > 0 \) are arbitrary weights and \( f_i \) is the individual discrete ISE function for linear model \( i \). Observe that coefficients \( \gamma_i \) allow the user to emphasize or not the approach to the most sensitive model, for instance. Since in this application we found no reason to differentiate the models, we adopt \( \gamma_i = 1, \forall i \in [1, M] \). Hence, the objective function for the convex hull in this case becomes

\[ f = \sum_{i=1}^{M} \sum_{k=1}^{N} [r(k) - y_i(k)]^2 \]  

(5.19)

where the time span is defined by \( N = 100 \).

The problem described to this point has a rapid numerical solution using an algorithm based on the gradient method. Four \( q \) parameters \((w = 3)\) are needed basically to satisfy the combined design requirement expressed by (5.17), the no-offset condition and that sort of settling time imposed by the final control condition. The controller parameters obtained using the polytopic model in the basic-formulation problem are the following:

\[ q_0 = 272.98, \quad q_1 = -494.05, \quad q_2 = 222.94, \quad q_3 = 17.09 \]  

(5.20)

Morningred et al.\cite{16} have previously worked with this reactor model for testing different alternatives of predictive controllers and confronted the results with the responses obtained using a PI controller whose parameters were adjusted by the ITAE criterion; thus, we used the same settings: the gain value, 52 L\(^2\) mol\(^{-1}\) min\(^{-1}\) and the integration time constant, 0.46 min. The simulation tests are also similar to Morningred's work and consist of a sequence of step changes in the reference value and a sequence of load changes in the feed stream concentration and the refrigerant inlet temperature.

Figure 6 shows the results obtained when comparing the discrete controller with the mentioned PI. The setpoint was changed in intervals of 10 min from 0.09 mol L\(^{-1}\) to 0.125, returns to 0.09, then steps to 0.055, and returns to 0.09 mol L\(^{-1}\). The superior performance of the discrete controller is obtained through a vigorous initial movement in the manipulated variable, which however does not overcome the 110 L min\(^{-1}\) limit as shown in Figure 7.

Figures 8 and 9 show the results obtained when comparing the discrete controller with the mentioned PI under load changes. For testing the disturbance
1.05 mol L\(^{-1}\) first the feed stream concentration changes from 1 to 
rejection the following sequence of changes are made: first the feed stream concentration changes from 1 to 1.05 mol L\(^{-1}\) and 10 min later the refrigerant temperature goes down 10 °C; then, the feed concentration and refrigerant temperature returns to the original value, with a 10 min difference between them. A better disturbance rejection capability is observed in the discrete controller, adjusted accordingly to the procedure suggested in this paper.

6. Conclusions

An efficient method for designing and tuning discrete controllers for case-specific design conditions is presented in this article. It basically consists of formulating an optimization problem where the closed-loop desired behavior is represented by time-domain characteristics that are mathematically translated into limiting constraints. The design philosophy emphasizes the importance of performance characteristics different from those traditionally specified by objective functions like ISE, and consequently suboptimal solutions are accepted if those characteristics—specified by constraint equations—are satisfied. The basic formulation assures closed-loop stability for open-loop stable and minimal phase systems; those cases presenting an unstable pole or a zero outside the unit circle need an extra constraint to guarantee internal stability.

The full potential of this tuning method comes from the possibility of including as many controller design conditions as desired, as long as they do not create conflicting requirements. Furthermore, if modeling uncertainties due to slow time-variant or nonlinear systems can be described, a natural extension of the method allows the controller to be adjusted for providing robust closed-loop performance. The use of polytopic representations to deal with uncertainties have been preferred in this work.

The results obtained by simulating simple linear systems and a continuously stirred tank reactor (CSTR) with important nonlinearities show the effectiveness of the proposed design and tuning method.

Appendix I. Multiple Models for Bounding Nonlinear Systems

Different ways of describing the uncertainty associated with modeling real dynamic systems are being explored in robust control theory. A very frequent approach consists of defining this uncertainty in the frequency domain based on a family of linear time-invariant (LTI) models. An alternative time-domain realization for this paradigm consists of determining a set of linear time-invariant models in different operating points for a nonlinear system, or at different times for varying linear systems. These multimodels have received the name of polytopic models and permit the approximation of a nonlinear or varying system through a combined linear representation. For instance, in the case of a discrete state variable representation where \(u(k) \in \mathbb{R}\) is the control input, \(x(k) \in \mathbb{R}^n\) is the state of the plant and \(y(k) \in \mathbb{R}\) is the plant output,

\[
x(k + 1) = A(k)x(k) + B(k)u(k),
\]

\[
y(k) = Cx(k)
\]

such that

\[
[A(k), B(k)] \in \Omega
\]

where \(\Omega\) is the set,

\[
\Omega = \text{Co}([A_1, B_1], [A_2, B_2], \ldots, [A_M, B_M])
\]

and where “Co” refers to the convex hull defined by the linear models \([A_i, B_i]\), \(i = 1, 2, \ldots, M\). In other words, if \([A, B] \in \Omega\), then for some nonnegative values of \(\lambda_1, \lambda_2, \ldots, \lambda_M\), summing to 1, we can write

\[
[A(k), B(k)] = \sum_{i=1}^{M} \lambda_i(k)[A_i, B_i]
\]

It is possible to demonstrate that any trajectory of the original nonlinear system is also a trajectory of (6.1) for some linear time-invariant representation belonging to \(\Omega\). Therefore, it is reasonable to assume that any analysis and design method applied to the polytopic representation can be applied to the real system.

Since a complete equivalence exists among discrete state-space representations and difference equations,
the above multimodel can also be written as follows:

$$y_i(k) = -\sum_{j=1}^{n} a_{ij} y_i(k-j) + \sum_{i=1}^{m} b_{ij} u_i(k-i - k_d)$$  \hfill (6.5)

where \( l = 1, 2, ..., M \) identifies the vertices of the polytope.

**Appendix II. About the Number of Controller Parameters**

Let us assume the discrete model of the process is represented by (3.2) and the discrete controller is given by (3.1). Then, the closed-loop characteristic equation for the system is

$$\{1 + \sum_{j=1}^{n} a_{ij} z^{-j}\} \{1 + \sum_{j=1}^{v} p_{ij} z^{-j}\} + z^{-k_d} \sum_{j=0}^{w} q_{ij} z^{-j} \sum_{i=1}^{m} b_{ij} z^{-i} = \sum_{i=0}^{h} g_{i} z^{-i} = 0$$

where \( n \) is the degree of the closed-loop characteristic equation. Employing the formula of polynomials product, we have

$$1 + \sum_{i=1}^{n+v} h_{i} z^{-i} + z^{-k_d} \sum_{j=0}^{w} q_{ij} z^{-j} = 1 + \sum_{i=1}^{h} g_{i} z^{-i}$$ \hfill (6.6)

where the degree \( h \) is given by

$$h = \text{max}(n + v, w + m + k_d)$$

and \( a_{ij} \) and \( \beta_{ij} \) are given by the Cauchy formula:

$$a_{ij} = \sum_{i=1}^{1} a_{ij} p_{ij}, \quad i - j \leq n, \quad j \leq v$$ \hfill (6.7)

$$\beta_{ij} = \sum_{j=0}^{1} b_{ij} q_{ij}, \quad i - m \leq j \leq w$$ \hfill (6.8)

Replacing (6.7) and (6.8) in (6.6), we obtain the full expression of the closed-loop characteristic equation. The coefficients \( q \) determine a system of linear equations where there are \( w + v + 1 \) unknowns—the controller parameters—and if the no-offset condition (3.3) is included, there are \( h + 1 \) equations. Since the system equation matrix is made up by two Toeplitz submatrices (one for the \( q \) and the other for the \( p \)), the rank \( \rho \) is given by

$$\rho = \min(h + 1, w + v + 1)$$

Therefore, the equations system has a solution if and only if \( w, v \geq h \). If \( h = n + v \), it requires \( w \geq n \) and \( v \geq m + k_d \). If \( h = w + m + k_d \), the same result is obtained. In any case, the equality means that the controller provides just enough degrees of freedom as to arbitrarily locate all the closed-loop poles. Greater values than those indicated by the equalities might be necessary to achieve demanding performances.

**Appendix III. Analysis of the Final Control Condition**

The discrete controller in (3.1) may be equivalently written using the general state-space representation,
Hence, the control increment in (6.15) can be written now as

\[ \Delta u(k + 1) = C_u[(A_S - I)A_S^k x(0) + (A_S - I) \sum_{i=1}^{k} A_S^{i-1} B_S \tau(k - i) + B_S \tau(k)] + \]

\[ D_u[r(k + 1) - r(k)] \]  \hspace{1cm} (6.17)

Assuming a setpoint change from 0 to \( r \) at the time instant \( k = 0 \), and \( X(0) = 0 \) for simplicity, the last expression becomes

\[ \Delta u(k + 1) = C_u[(A_S - I) \sum_{i=1}^{k} A_S^{i-1} B_S \tau + B_S \tau] \]  \hspace{1cm} (6.18)

Recalling a property of geometric progressions, now we can write

\[ \sum_{i=1}^{k} A_S^{i-1} B_S \tau = (I - A_S)^{-1}(I - A_S^k) B_S \tau, \quad A_S \neq I \]  \hspace{1cm} (6.19)

Substituting (6.19) into (6.18) and rearranging gives

\[ \Delta u(k + 1) = C_u A_S^k B_S \tau \]  \hspace{1cm} (6.20)

Now, the constraint \( \Delta u(k + 1) = 0 \) included in the proposed tuning problem formulation may be analyzed considering an arbitrary accuracy \( \epsilon < 1 \), such that it is satisfied if

\[ |\Delta u(k + 1)| \leq \epsilon \]  \hspace{1cm} (6.21)

For a better visualization of the effect of (6.21) on the location of closed-loop characteristic values, let us take a conservative condition using a property of the 2-norm, i.e.,

\[ |C_u A_S^k B_S| \leq ||C_u|| ||A_S||^k ||B_S|| \leq \frac{1}{|\tau|^{\frac{1}{k}}} \]  \hspace{1cm} (6.22)

or, since \( C_u, B_S, \) and \( \tau \) are different from zero

\[ ||A_S||^k \leq M \epsilon \]  \hspace{1cm} (6.23a)

where \( M \) is a positive quantity. Observe that if (6.23a) is satisfied for the sampling instant \( k \), it will also verify for all subsequent sampling instants. This also implies

\[ \max_{i} |\lambda_i(A_S)| \leq ||A_S|| \leq (M \epsilon)^{1/k} \]  \hspace{1cm} (6.24)

Because \( \epsilon \) is an order lower than \( 10^{-6} \) in most optimization software packages, it is most probable that \( M \epsilon \ll 1 \) for most practical cases. Hence, the roots are enclosed by a circle whose diameter increases asymptotically up to 1 when \( k \rightarrow \infty \). Figure 10 shows the effect of \( k \) on the closed-loop pole locations for the case given in Example 1.

Furthermore, because \( A_S \) is \( R^{(n+v)(n+v)} \), assuming \( \epsilon = 0 \) requires \( A_S \) to be at least nilpotent of order \( k = (n + v) \) and from (6.24), it also means

\[ |\lambda_i(A_S)| = 0, \quad \forall i \in [1, n + v] \]  \hspace{1cm} (6.25)

This gives the lower bound \( n + v \) for parameter \( N \), where \( n + v \) is the dimension of the process-plus-controller system without pole-zero cancelation (we assume time delay \( k_d = 0 \) in this appendix).

**Appendix IV. The Nonlinear Reactor Model**

The model of a continuous reactor where an irreversible exothermic reaction takes place has been selected for testing the proposed design and tuning method.

The reaction is

\[ A \rightarrow B \]

and occurs in a constant volume reactor cooled by a single coolant stream. The operation is modeled by the following equations:

\[ \frac{dC_A(t + k_d)}{dt} = \frac{q(t)}{V} [C_A^0(t) - C_A(t + k_d)] - k_0 C_A(t + k_d) \exp\left[ -E \frac{1}{RT(t)} \right] \]

\[ \frac{dT(t)}{dt} = \frac{q(t)}{V} [T_i(t) - T(t)] - \frac{k_0 \Delta H}{\rho_c p_c} C_A(t + k_d) \times \exp\left[ -E \frac{1}{RT(t)} \right] + \frac{\rho_c p_c}{\rho_p} q_c(t) \left( 1 - \exp\left[ -\frac{-hA}{\rho_p c_p T_c(t) - T(t)} \right] \right) \]

The nominal parameter values for this model appear in Table 2. The objective is to control the measured concentration of reactive A at the outlet stream, \( C_A \), by

<table>
<thead>
<tr>
<th>Table 2. Nominal CSTR Parameters Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>parameter</td>
</tr>
<tr>
<td>measured concentration</td>
</tr>
<tr>
<td>reactor temperature</td>
</tr>
<tr>
<td>coolant flow rate</td>
</tr>
<tr>
<td>process flow rate</td>
</tr>
<tr>
<td>feed concentration</td>
</tr>
<tr>
<td>feed temperature</td>
</tr>
<tr>
<td>inlet coolant</td>
</tr>
<tr>
<td>temperature</td>
</tr>
<tr>
<td>CSTR volume</td>
</tr>
<tr>
<td>heat-transfer term</td>
</tr>
<tr>
<td>reaction rate</td>
</tr>
<tr>
<td>constant activation energy</td>
</tr>
<tr>
<td>heat of reaction</td>
</tr>
<tr>
<td>liquid densities</td>
</tr>
<tr>
<td>specific heats</td>
</tr>
</tbody>
</table>
manipulating the coolant flow rate $q_c$. The concentration has a measured time delay of $k_d = 0.5$ min. The nonlinear characteristics are clearly appreciated in Figure 5 where the responses to equal amplitude step changes are shown.

**Literature Cited**


Received for review January 15, 1999
Revised manuscript received August 2, 1999
Accepted August 16, 1999

IE990037W