

Article

Improved Autotuning Using the Shape Factor from Relay Feedback

T. Thyagarajan, and Cheng-Ching Yu

Ind. Eng. Chem. Res., **2003**, 42 (20), 4425-4440 • DOI: 10.1021/ie011006f

Downloaded from <http://pubs.acs.org> on November 28, 2008

More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 2 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

[View the Full Text HTML](#)



ACS Publications
High quality. High impact.

Improved Autotuning Using the Shape Factor from Relay Feedback

T. Thyagarajan and Cheng-Ching Yu*

Department of Chemical Engineering, National Taiwan University, Taipei 106-17, Taiwan

Since the introduction of the relay feedback test by Åström and Hägglund (*Automatica* **1984**, *20*, 645–651), autotuning of PID controller has received much attention, and many commercial autotuners have also been designed accordingly. Without knowledge of the model structure, most of these relay feedback autotuners use Ziegler–Nichols-type tuning rules to set controller parameters. This can lead to poor performance in some cases, because no single tuning rule can work well for all model structures over the entire range of parameter values. Luyben points out that additional information can be obtained from relay feedback tests, namely, the shape of the response (*Ind. Eng. Chem. Res.* **2001**, *40* (20), 4391–4402). In this work, relay feedback tests are conducted on processes with different orders and a wide range of dead-time-to-time-constant ratios. On the basis of the shape of the response from the relay feedback tests, these processes can be broadly classified into three major categories (model structures). Procedures are given to find parameters for the corresponding model structures, and then different tuning rules are employed to find appropriate PI controller settings. The procedures are tested against linear systems with and without noise. Simulation results clearly indicate that, by incorporating the shape information, improved autotuning can be achieved in a straightforward manner. Moreover, possible dead-time compensation and higher-order compensation can also be devised when necessary. It should be emphasized that the improvement is obtained from the conventional relay feedback test and no additional testing is required.

1. Introduction

The relay feedback test proposed by Åström and Hägglund¹ has received much attention from process control practitioners. Luyben² is among the first to employ the relay feedback test for system identification. The autotune variation (ATV) identification method has become a standard practice in chemical process control. Chang et al.³ derived transfer functions from relay feedback tests with increased accuracy. Several groups^{4–6} have developed autotuning procedures for PID controllers for open-loop unstable process having dead time. Methods to obtain exact parameter estimations from asymmetrical limit cycle data have also been proposed.^{7,8} The progress in relay feedback autotuning up to 1999 is comprehensively documented in the book by Yu.⁹ Ultimate-gain- and ultimate-frequency-based autotuning might produce poor performance in some instances, because information on model structure is lacking. Typical examples include first-order systems with large dead-time-to-time-constant ratios and second-order systems with small dead-time-to-time-constant ratios. Many authors have proposed modifications to the relay feedback test, and some works recommend additional tests to extract extra information about the process dynamics. Luyben¹⁰ suggested that additional testing might not be necessary if the extra information available from a single test is used, namely, the *shape* of the relay feedback response.

The purpose of this work is to utilize the shape information from the relay feedback test to identify the correct model structure of the process and to determine appropriate PID controller settings. The additional shape information is also useful for devising dead-time

compensation and higher-order compensation when necessary. This paper is organized as follows. A review of relay feedback, the shapes of relay feedback responses of various processes, observations related to process dynamics made from these shapes, and the classification of different processes based on the shape information are presented in section 2. The identification procedure is detailed in section 3. The implications of proposed method in process control, dead-time compensation, and higher-order compensation devised and the results are discussed in section 4. Concluding remarks are presented in the final section.

2. Shapes of Relay Feedback Response

2.1. Relay Feedback. The Åström and Hägglund relay feedback test is based on the observation that, when the output lags behind the input by π radians, the closed-loop system can oscillate with a period of P_u . A relay of magnitude h is inserted in the feedback loop. Initially, the input $u(t)$ is increased by h . Once the output $y(t)$ starts increasing after a time delay (D), the relay switches to the opposite direction, $u(t) = -h$. Because there is a phase lag of $-\pi$, a limit cycle of amplitude a is generated, as shown in Figure 1. The period of the limit cycle is the ultimate period, P_u . The approximate ultimate gain, K_u , and the ultimate frequency, ω_u , are

$$K_u = \frac{4h}{\pi a} \quad (1)$$

$$\omega_u = \frac{2\pi}{P_u} \quad (2)$$

* Corresponding author. Tel.: +886-2-3365-1759. Fax: +886-2-2362-3040. E-mail: ccyu@ccms.ntu.edu.tw.

The relay feedback test is a useful tool for system identification because it identifies two important pa-

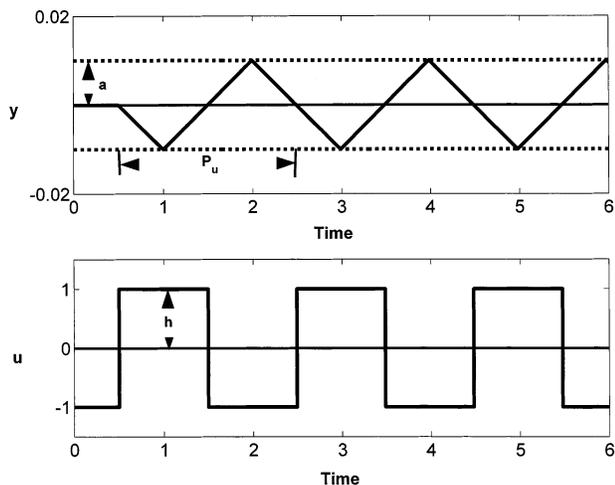


Figure 1. Typical relay feedback response.

parameters for controller tuning, the ultimate gain and the ultimate frequency. Typically, Ziegler–Nichols-type tuning rules are applied because K_u and P_u are the quantities required to set PID controller parameters. Unfortunately, satisfactory performance is not always guaranteed because no single tuning rule works well for the entire range of dead time (D) to time constant (τ) ratios (D/τ), even for first-order plus dead time (FOPDT) processes.^{10,11} Luyben¹⁰ demonstrates that, for first-order plus dead time processes, different D/τ ratios give different shapes in relay feedback tests (Figure 2) and these shape factors can be utilized to find the corresponding D/τ values, so that different tuning rules can be applied accordingly. This approach represents significant progress in relay feedback identification, and much reliable autotuning has resulted, as shown by Luyben.¹⁰ Figure 2 shows the transition from a triangular to an almost rectangular shape as D/τ changes from 0.1 to 10. Similar figures have also been provided by Friman and Waller.¹² In Luyben's work, the time to the midpoint of the amplitude (a) is used to characterize D/τ .

2.2. Shapes. To characterize model structures and parameter values (e.g., D/τ), processes with different orders (i.e., 1st, 2nd, 3rd, 8th, 15th, and 20th) and different dead-time-to-time-constant ratios (i.e., $D/\tau = 0.01, 1, \text{ and } 10$) were studied. In this work, only overdamped processes are considered (underdamped processes and systems with inverse responses are not included). Figure 3 shows the relay feedback responses for the higher-order processes. Note that all process gains were assumed to be 1, and a relay height of 1 (i.e., $h = 1$) was used to generate sustained oscillations.

From the shapes of the curves in Figures 2 and 3, several observations can immediately be made.

1. *First-Order Plus Dead Time (FOPDT) Process.* If the response curves have sharp edge (discontinuities) at the peak amplitudes (i.e., $y = \pm a$), the process can be considered as a FOPDT system, as shown in Figure 2.

2. *Effect of D/τ for FOPDT Process.* If the relay feedback gives a triangular wave, the process can be treated as a time-constant-dominant process (i.e., small D/τ for FOPDT). Specifically, the time required to reach the peak amplitude is equal to the dead time, as will be shown later. If the dead-time-to-time-constant ratio becomes larger, curvature begins to appear (e.g., Figure 2), and this implies gradual development toward a step

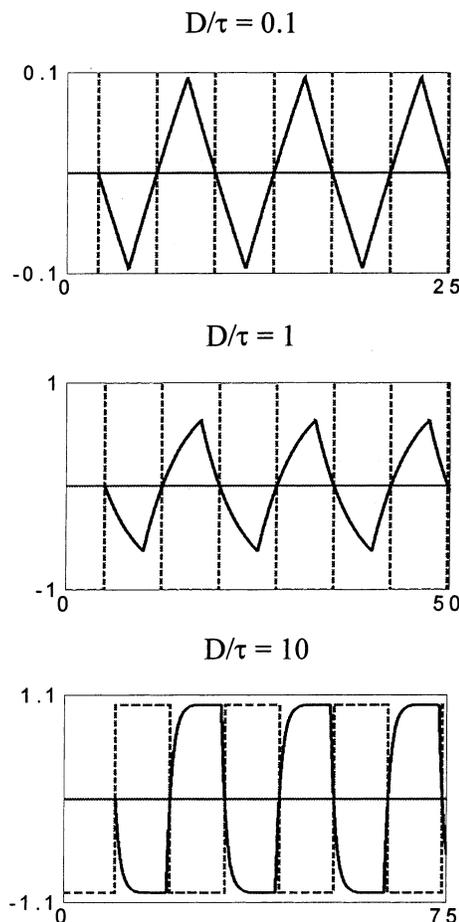


Figure 2. Relay feedback responses of first-order plus dead time (FOPDT) processes with different D/τ values (controlled variable, solid lines; manipulated variable, dashed lines).

response. As D/τ approaches infinity ($D/\tau \rightarrow \infty$), the response resembles a symmetrical rectangular wave. Actually, FOPDT processes represent a very unique class in terms of relay feedback responses.

3. *Effect of Order.* If the order of the process increases to 2 and beyond (e.g., $n = 2, 3, 8, 15, \text{ and } 20$ in Figure 3), the sharp edges disappear, and the responses resemble sinusoidal oscillations. Generally, sustained oscillations develop in all cycles, except for second-order processes with small D/τ values (Figure 3). Again, when the dead-time-to-time-constant ratios become large, the responses approach rectangular waves.

4. *Exponentially Developed Cycling.* If the response consists of sinusoidal oscillations with exponentially increasing magnitude that reach steady state after many cycles, the process can be considered as a second-order process with a small D/τ value. This again represents a special class in relay feedback responses.

2.3. Model Structures. The observations presented above are useful in identifying different model structures. The basic principle in classification is to use the fewest classes while capturing all possible curve shapes in the relay feedback responses. From the responses in Figures 2 and 3, three distinct classes can be identified.

2.3.1. Categories 1a and 1b: First-Order Plus Dead Time (FOPDT). As pointed out in the previous section, two distinct features signify FOPDT systems: (1) response curves showing sharp edges and (2) responses reaching stationary oscillations in the first cycle (see Figure 1). Therefore, category 1 is represented by the FOPDT process

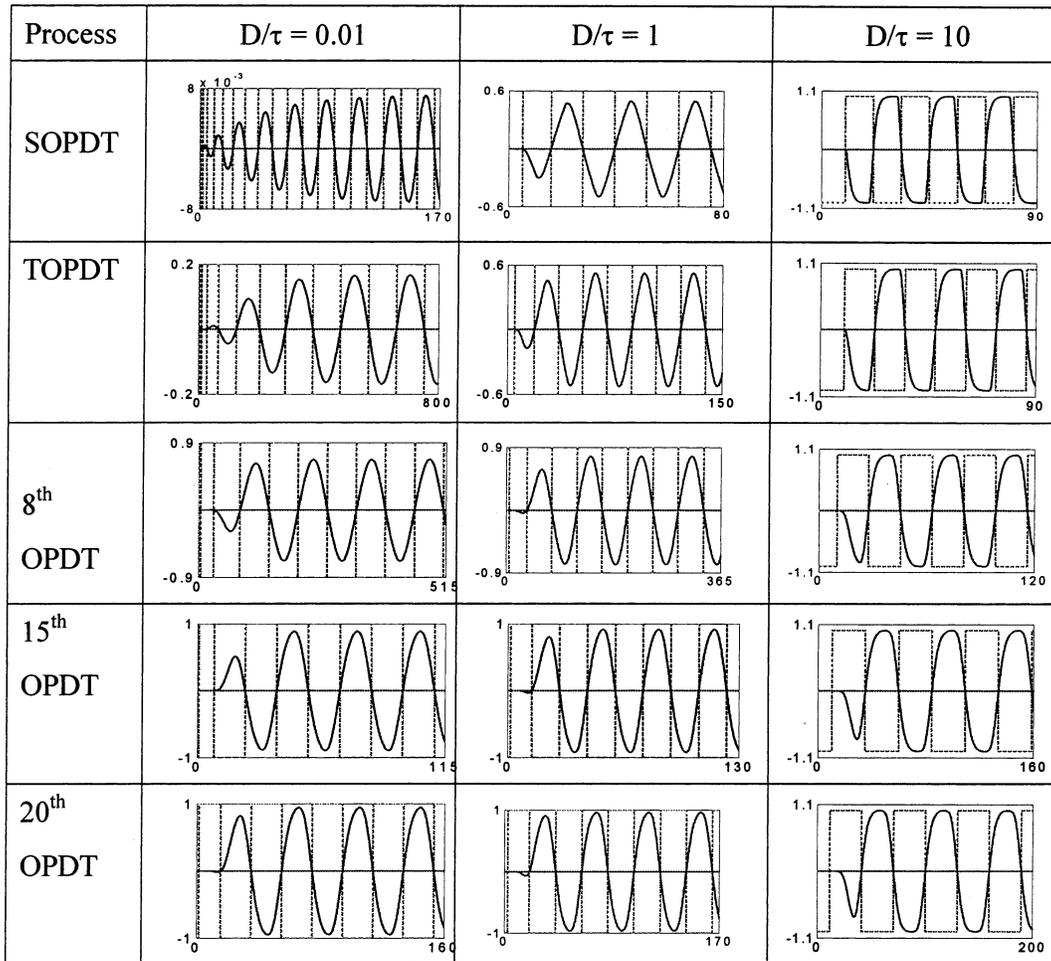


Figure 3. Relay feedback response for processes with different orders and various D/τ ratios (controlled variable, solid lines; manipulated variable, dashed lines).

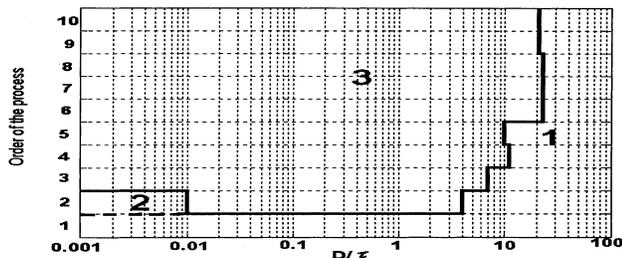


Figure 4. Quantitative classification of different model structures (categories 1–3) based on integrated absolute error from frequency response.

$$G(s) = \frac{K_p e^{-Ds}}{\tau s + 1} \quad (3)$$

where K_p is the steady-state gain, D is the dead time, and τ denotes the time constant. Figure 4 shows that, if the process is truly first-order plus dead time, it certainly falls into category 1. However, a higher-order process with a large D/τ value can also be classified into this category, as suggested by inspection of the relay feedback responses shown in the top two rows of Figure 3 with $D/\tau = 10$. A quantitative comparison of the multiplicative error (defined by eq 30 and reported in Table 3) also reveals that higher-order systems with D/τ greater than 10 are better represented by FOPDT system.

2.3.2. Category 2: Second-Order Plus Small Dead Time (SOPSDT). Inspection of Figure 3 indicates that a model structure to describe exponentially developed cycling (top row with $D/\tau = 0.01$ in Figure 3) is needed. An ideal candidate is a second-order plus small dead time (SOPSDT) process, i.e.

$$G(s) = \frac{K_p e^{-Ds}}{(\tau s + 1)^2} \quad (4)$$

Typically, if the ratio $\epsilon = D/\tau$ is less than 0.01, the oscillation develops slowly. In this work, the ratio ϵ is set to 0.001. Thus, the transfer function can be expressed as

$$G(s) = \frac{K_p e^{-\epsilon \tau s}}{(\tau s + 1)^2} \quad (5)$$

Again, a quantitative assessment (Figure 4) confirms this category as SOPSDT.

2.3.3. Category 3: Higher-Order (HO). In addition to the two above-mentioned categories, Figure 3 indicates that the rest of the responses show sinusoidal oscillations with stationary cycling reached in 1–2 cycles. This behavior can be described by a higher-order (HO) process without dead time. A typical transfer function for such a process is

$$G(s) = \frac{K_p}{(\tau s + 1)^n} \quad (6)$$

As long as $n \geq 3$, similar relay feedback responses can be seen. In this work, a default value of $n \geq 5$ is used, which leads to

$$G(s) = \frac{K_p}{(\tau s + 1)^5} \quad (7)$$

Figure 4 shows that this category, category 3, covers the largest parameter space for the systems studied.

The model structures presented above (eqs 3, 5, and 7) show that there are only two unknown parameters for categories 2 and 3 and that, for the FOPDT model, there are three unknown parameters (i.e., K_p , D , and τ). The parameters chosen for categories 2 and 3 (i.e., ϵ and n in eqs 5 and 6) can affect the distribution of model structures in the parameter space of Figure 4. Nonetheless, most of the curve shapes are well-represented using these three classes.

3. System Identification

The detailed procedures for system identification of various processes under different categories are presented below.

3.1. Category 1: FOPDT. This category includes two types of systems. One is the true first-order plus dead time process, as shown in the last row of Figure 4, and this is denoted category 1a. The second one is higher-order systems with large D/τ values, as shown on the right-hand side of Figure 4, which is called category 1b.

3.1.1. Category 1a. This category has two important characteristics: (1) sharp edges at the peak amplitude and (2) stationary oscillations that develop in the first cycle. The relay feedback response of an FOPDT system can actually be described analytically. Figure 5A shows the original response curve where the output starts to increase after the dead time D . If the output (y) is aligned with the input (u) by being shifted to the left (Figure 5B), it becomes clear that the increasing part of the output response is the result of a step increase in u of magnitude h . After a delay time D , the relay switches to $-h$, and the second step change becomes effective immediately, which results in the decreasing portion of the half-cycle, as shown in Figure 5B. Therefore, the analytical expression for the first half-cycle becomes

$$y(t) = K_p h(1 - e^{-t/\tau}) \quad \text{for } 0 < t < D \quad (8)$$

$$y(t) = K_p h(1 - e^{-t/\tau}) - 2K_p h(1 - e^{-(t-D)/\tau}) \quad \text{for } D < t < P_u/2 \quad (9)$$

The continuous step change repeats itself, and a sustained oscillation results. Note that similar derivations were proposed by Wang et al.⁷ for stable first-order plus dead time systems and Tan et al.⁴ and Huang and Chen⁵ for unstable first-order open-loop systems. Equations 8 and 9 clearly indicate that the time required to reach the peak amplitude is exactly the dead time (D) for the FOPDT process and that this value can be validated repeatedly at each half-cycle ($P_u/2$). Given the boundary conditions $y(D) = a$ and $y(P_u/2) = 0$, one can solve for the other two model parameters, K_p and τ .

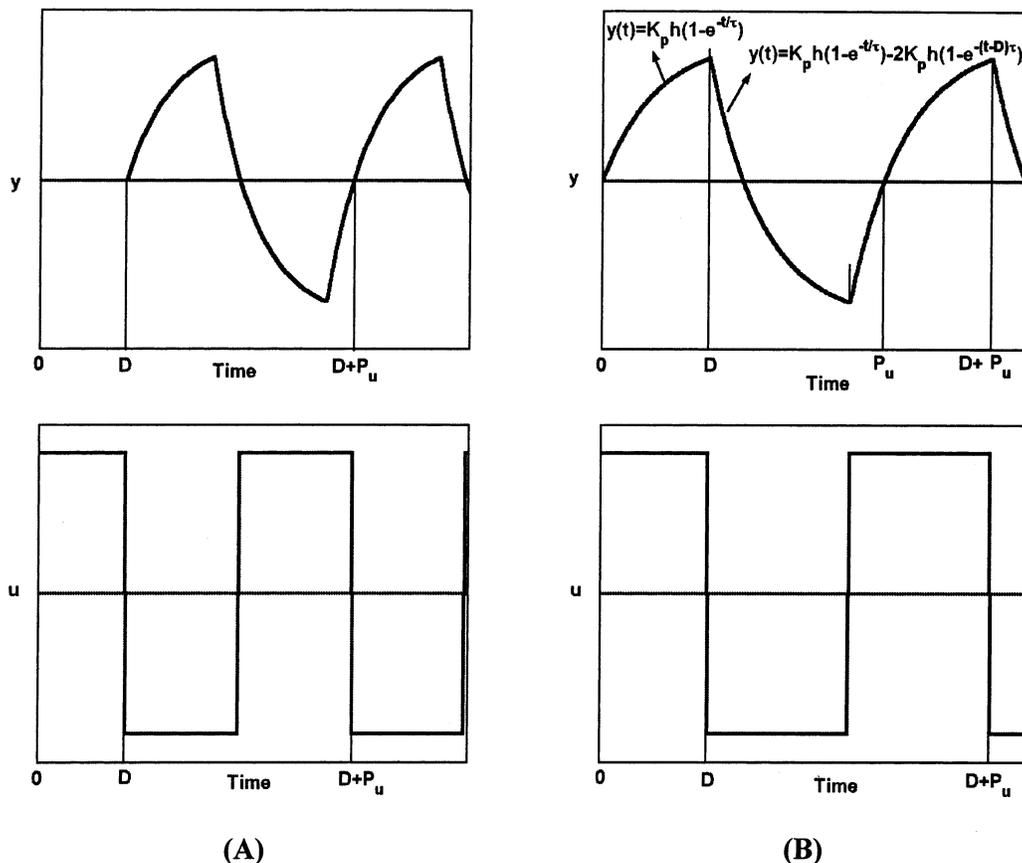


Figure 5. Analytical expressions of relay feedback response for FOPDT systems: (A) original response and (B) shifted version.

$$\tau = \frac{P_u/2}{\ln(2e^{D/\tau} - 1)} \quad (10)$$

$$K_p = \frac{a}{h(1 - e^{-D/\tau})} \quad (11)$$

Thus, all three model parameters can be determined from the relay feedback response for true FOPDT processes. Therefore, the identification consists the following steps: (0) Record the time to the peak amplitude (D), the peak amplitude (a), and the period of oscillation (P_u). (1) Set the dead time (D) as the time to the peak value (Figure 5). (2) Compute the time constant τ from eq 10. Note that eq 10 is an implicit equation for τ that requires an iterative solution. One can use the relationship between the ultimate frequency and τ given below to obtain a first guess for the value of τ

$$\tau = \frac{\tan(\pi - D\omega_u)}{\omega_u} \quad (12)$$

(3) Compute K_p from eq 11.

Similarly, for an unstable first-order plus dead time system, all three model parameters can be obtained directly from the relay feedback responses

$$G(s) = \frac{K_p e^{-Ds}}{\tau s - 1} \quad (13)$$

First, the output responses can be represented analytically as shown in Figure 6. From the first half-period, the expression becomes

$$y(t) = K_p h(e^{t/\tau} - 1) \quad \text{for } 0 < t < D \quad (14)$$

$$y(t) = K_p h(e^{t/\tau} - 1) - 2K_p h(e^{(t-D)/\tau} - 1) \quad \text{for } D < t < P_u/2 \quad (15)$$

Substituting the boundary conditions $y(D) = a$ and $y(P_u/2) = 0$ into eqs 14 and 15, the time constant (τ) and steady-state gain (K_p) can be computed directly via

$$\tau = \frac{P_u/2}{\ln[1/(2e^{-D/\tau} - 1)]} \quad (16)$$

$$K_p = \frac{a}{h}(e^{D/\tau} - 1) \quad (17)$$

The identification procedure for the unstable process is exactly the same as that for the stable process except that the equations for computing τ and K_p are eqs 16 and 17, respectively, rather than eqs 10 and 11. Equation 16 also reveals that the condition for the existence of a limit cycle requires $D/\tau < \ln 2$ (when $D/\tau = \ln 2$, the ultimate period becomes infinite).

3.1.2. Category 1b: Approximated FOPDT. This category also has the FOPDT model structure. As in category 1a, stationary cycling develops in the first cycle, but the sharp edge around the peak amplitude is not quite obvious as it is for category 1a. Second- and third-order systems with $D/\tau = 10$ fall into this category. Because the true process is not exactly an FOPDT system, determining the dead time directly from the response (e.g., Figure 5B) can be erroneous. As in Luyben's approach, we first define the time required to reach the peak amplitude (a) as t_a and the time required to reach one-half of the peak amplitude ($a/2$) as $t_{a/2}$.

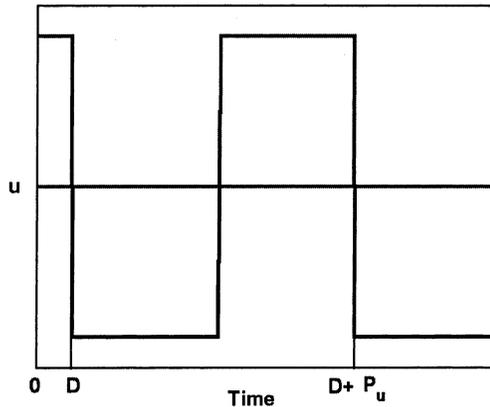
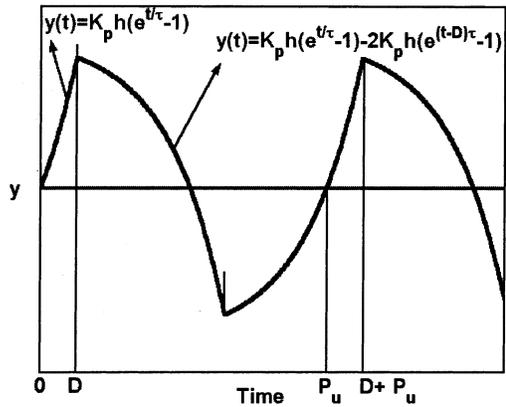


Figure 6. Analytical expressions of relay feedback response for unstable FOPDT systems (shifted version).

Following the analytical expressions in eqs 8 and 9, we have

$$a = K_p h(1 - e^{-t_a/\tau}) \quad (18)$$

$$\frac{a}{2} = K_p h(1 - e^{-t_{a/2}/\tau}) \quad (19)$$

Dividing eq 18 by eq 19, we can solve for τ using

$$2e^{-t_{a/2}/\tau} - e^{-t_a/\tau} = 1 \quad (20)$$

Once τ is known, we can solve for the other two model parameters, K_p and D , from the ultimate properties

$$D = \frac{\pi - \tan^{-1}(\tau\omega_u)}{\omega_u} \quad (21)$$

$$K_p = \frac{\sqrt{1 + (\tau\omega_u)^2}}{K_u} \quad (22)$$

Therefore, the identification consists of the following steps: (0) Record the time to the peak amplitude (t_a), the time to one-half of the peak amplitude ($t_{a/2}$), the peak amplitude (a), and the period of oscillation (P_u). (1) Compute the time constant τ from eq 20. (2) Calculate the dead time from eq 21. (3) Calculate the steady-state gain (K_p) from eq 22.

This procedure enables us to find the approximate FOPDT model.

3.2. Category 2: SOPSDT. As pointed out earlier, if a small dead-time-to-time-constant ratio ($\epsilon = D/\tau$) is

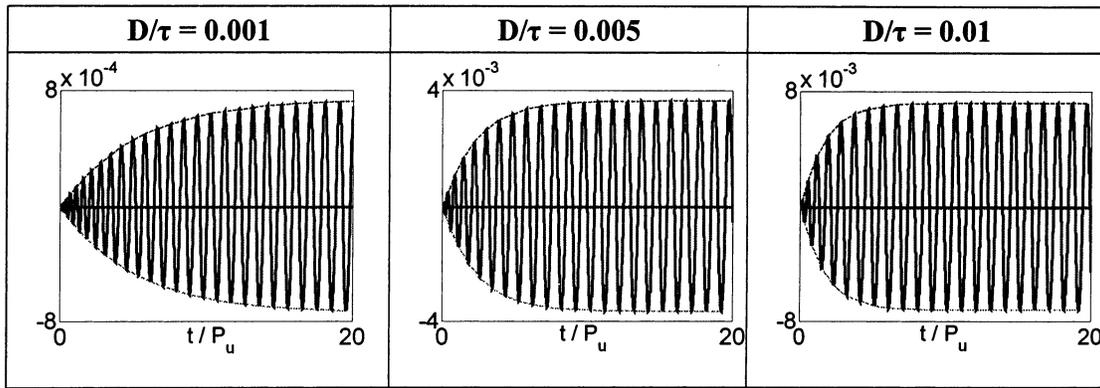


Figure 7. Relay feedback responses and global responses (defined by peaks) of SOPSDT processes with different D/τ ratios.

specified, then there are only two unknown parameters, as shown in eq 5. After several numerical simulations, the default value of $\epsilon = 0.001$ was found to work well for a wide range of parameter values. The time constant can be obtained from phase-angle information

$$-\pi = -\epsilon\tau\omega_u - 2 \tan^{-1}(\tau\omega_u) \quad (23)$$

and the steady-state gain can be computed according to

$$K_p = \frac{1 + (\tau\omega_u)^2}{K_u} \quad (24)$$

Therefore, the procedure consists of the following steps: (0) Record the values of the peak amplitude (a) and the period of oscillation (P_u). (1) Compute the time constant τ from eq 23. (2) Calculate the steady-state gain (K_p) from eq 24.

This procedure enables us to find the SOPSDT model with the default setting of $\epsilon = 0.001$.

However, for certain cases, the model parameters obtained using the default value of $\epsilon = 0.001$ might not be satisfactory. This is because the ratio ϵ depends on the rate at which the oscillations develop. Figure 7 presents the relay feedback responses obtained for SOPSDT processes with $\epsilon = 0.001, 0.005,$ and 0.01 . Figure 7 indicates that, as ϵ increases, the normalized time constant [ratio of the time constant (τ_G) defined by the peaks of the oscillations to the period of oscillation (P_u)] decreases. In other words, as D/τ increases, fewer cycles are required to reach static oscillations. Figure 8 displays a plot showing the dependence of the ratio D/τ on the normalized time constant. A linear model is used to relate $\log(\epsilon)$ to the normalized time constant (τ_G/P_u)

$$\log(\epsilon) = -0.3031 \left(\frac{\tau_G}{P_u} \right) - 1.4767 \quad (25)$$

Thus, with the value of τ_G/P_u from the relay feedback tests, we are able to calculate the ratio of D to τ , ϵ .

Therefore, this more elaborate procedure consists of the following steps: (0) Record the values of the peak amplitude (a), the period of oscillation (P_u), and the time constant (τ_G) from the global response (Figure 7). (1) Compute the value of the normalized time constant τ_G/P_u . (2) Compute the value of ϵ using eq 25. (3) Compute the time constant τ from eq 23. (4) Calculate the steady-state gain (K_p) from eq 24.

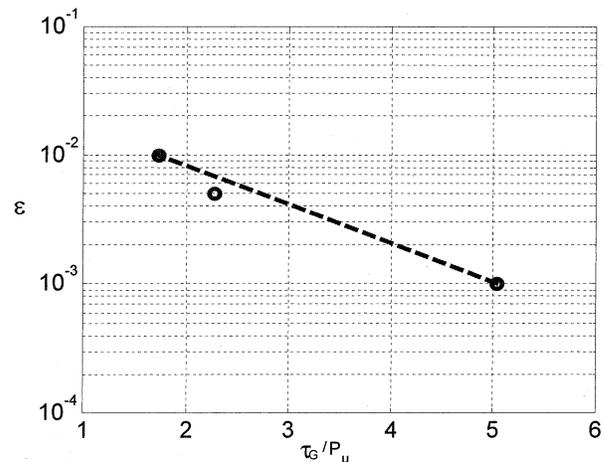


Figure 8. Relationship between the normalized global time constant (τ_G/P_u) and the dead time/time constant ratio ($\epsilon = D/\tau$) for SOPSDT processes.

This procedure enables us to find the dead-time-to-time-constant ratio (ϵ) and the model parameters for the SOPSDT model.

3.3. Category 3: HO. After several numerical simulations, the default value of $n = 5$ was found to work well for a wide range of parameter values. Thus, if the order n is chosen for category 3 (e.g., $n = 5$ in eq 6), we are left with two unknown parameters. They can be determined from the ultimate properties via

$$\tau = \frac{\tan(\pi/n)}{\omega_u} \quad (26)$$

$$K_p = \frac{[1 + (\tau\omega_u)^2]^{n/2}}{K_u} \quad (27)$$

Similarly, the identification procedure becomes: (0) Record the values of the peak amplitude (a) and the period of oscillation (P_u). (1) Compute the time constant τ from eq 26 (with $n = 5$). (2) Calculate the steady-state gain (K_p) from eq 27 (with $n = 5$).

This procedure enables us to find the parameters for the 5th-order process (HO model).

However, for certain cases, the model parameters obtained using the default value of $n = 5$ might not be satisfactory. This is because the value of n depends on the rate at which the oscillations develop. Similar to discussion in section 2.3.2, we can relate the normalized time constant (τ_G/P_u) to the order n by eq 28. A linear

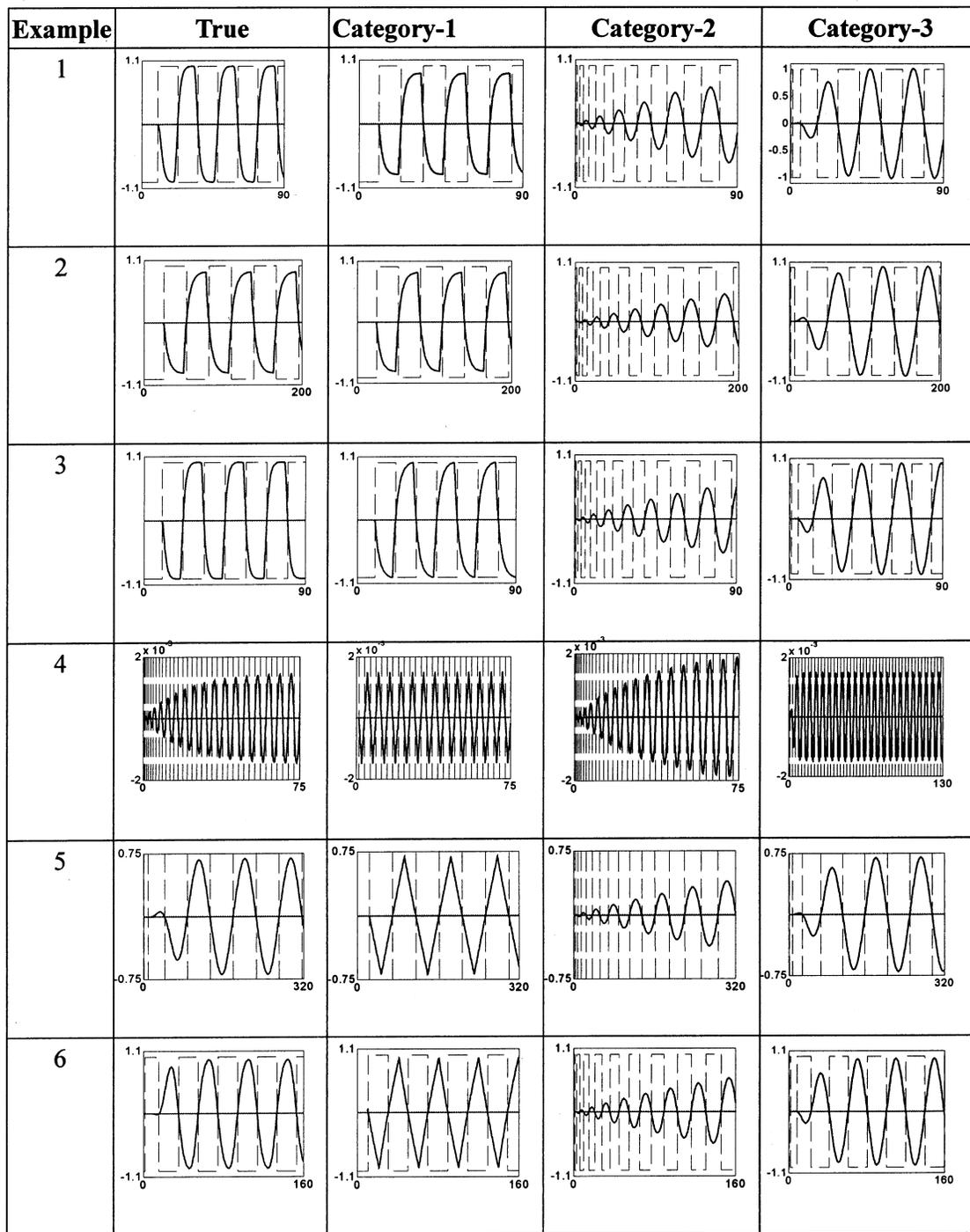


Figure 9. Reproduction of relay feedback responses (controlled variable, solid lines; manipulated variable, dashed lines) for six examples with the assumption of different model structures (from left to right: true process, FOPDT, SOPSDT, and HO).

model is used to interpolate for n and τ_C/P_u between $n = 3$ and $n = 10$

$$\log(n) = 1.9040 - 1.3736 \left(\frac{\tau_C}{P_u} \right) \quad (28)$$

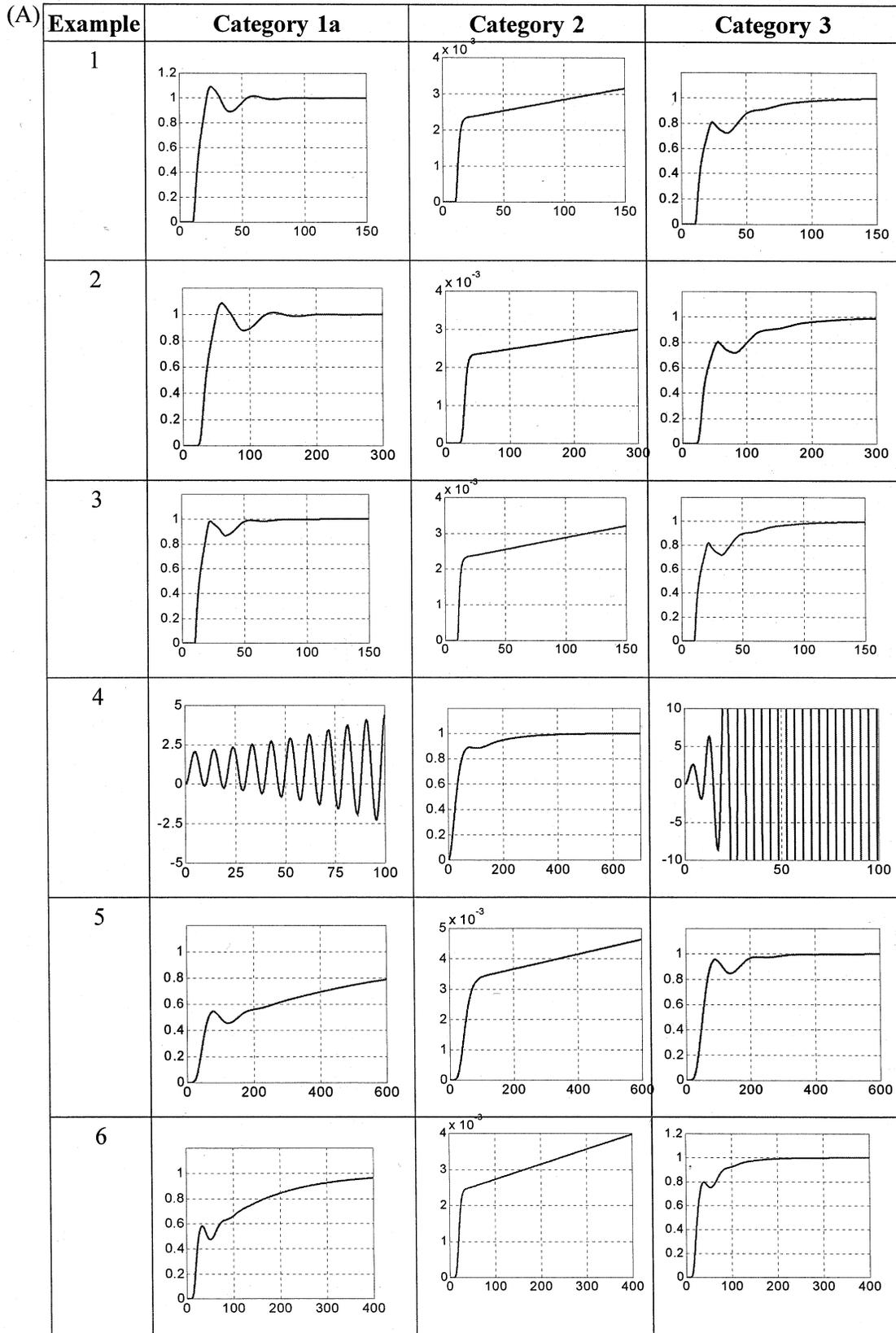
Thus, with the value of τ_C/P_u from the relay feedback tests, we are able to calculate the order n .

Therefore, the more elaborate procedure consists of the following steps: (0) Record the values of the peak amplitude (a), the period of oscillation (P_u), and the time constant of the curvature (τ_C). (1) Compute the value of the curvature factor $C = \tau_C/P_u$. (2) Compute the value

of n using eq 28. (3) Compute the time constant τ from eq 26. (4) Calculate the steady-state gain (K_p) from eq 27.

This procedure enables us to find the order as well as model parameters for the HO model.

3.4. Validation. To illustrate the appropriateness of the proposed classification, six typical examples representing the various categories are considered (Table 1). Relay feedback tests were conducted on all of the above examples, and the relay feedback responses thus obtained are shown in Figure 9 under the heading "true". Time-domain responses clearly indicate that examples 1–3 can be classified as approximated FOPDT systems



(category 1b), example 4 is an SOPSDT system (category 2), and examples 5 and 6 can be classified as HO processes (category 3). The ultimate properties computed from the relay experiments are also presented in Table 1.

3.4.1. Time-Domain Analysis. The equivalent models in the 3 categories for the above six examples were formulated using the identification procedures described

in the previous section. The 18 equivalent models thus obtained are presented in Table 2. Note that each equivalent model has the same values of K_u and ω_u but very different model structures. Relay experiments were conducted on all three equivalent models, as shown in Figure 9. The results show that the correct model structure reproduces the relay feedback response while the mismatched model structures give completely dif-

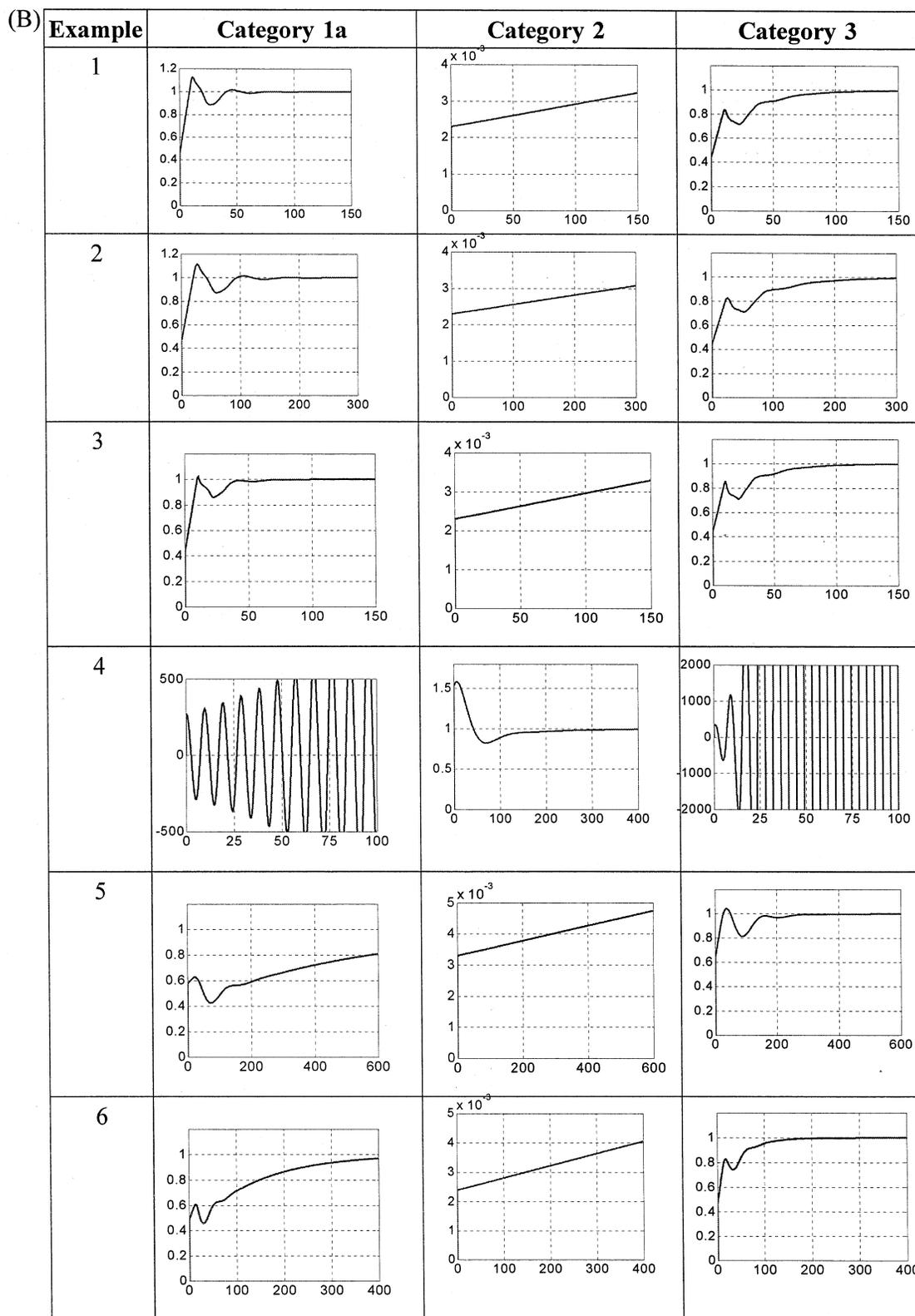


Figure 10. Comparison of set-point responses for six examples with PI controllers: (A) controlled variables, (B) manipulated variables.

ferent curve shapes, despite having the same K_u and P_u . The time domain responses confirm that examples 1–3 belong to category 1, example 4 belongs to category 2, and examples 5 and 6 belong to category 3.

The identified models in Table 2 also reveal the importance of applying the appropriate model structure. In example 1, using the same values of K_u and ω_u , the SOPSDT model gives a steady-state gain 1600 times the

true value, and in example 4, the FOPDT model structure results in an unstable system. The results clearly indicate the need to extract model structure information from relay feedback tests.

3.4.2. Frequency-Domain Analysis. The above categorization can be validated in the frequency domain by evaluating the integrated absolute error (IAE) for each of the equivalent models. Multiplicative error is

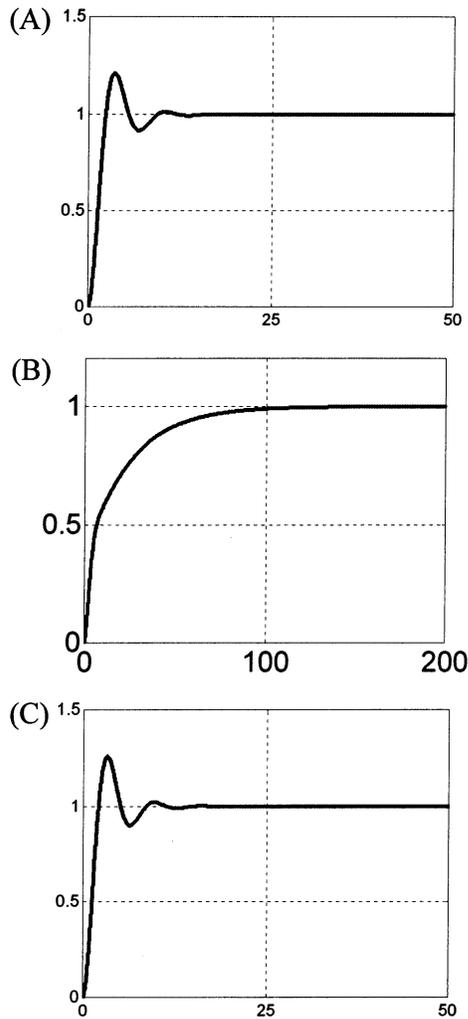


Figure 11. Comparison of closed-loop responses of an SOPSDT process (special case) for set-point tracking: (A) controller settings obtained using the true process, (B) controller settings obtained using the (SOPSDT) model with the default value of $\epsilon = 0.001$, and (C) controller settings obtained using the SOPSDT model with ϵ obtained from eq 25 ($\epsilon = 0.00685$).

employed here

$$e_{(\omega)} = \frac{G_m(j\omega) - G(j\omega)}{G(j\omega)} \quad (29)$$

where G is the true process and G_m is the derived model. The integrated absolute error is evaluated between $0.1\omega_u$ and $10\omega_u$

$$\text{IAE}_{\omega} = \int_{0.1\omega_u}^{10\omega_u} |e_{(\omega)}| d\omega \quad (30)$$

The numerical values of IAE_{ω} are presented in Table 3. By comparing the values of IAE_{ω} of the models under categories 1, 2, and 3 of a particular true process, one can easily identify the category to which the true process belongs. The category to which the true process belongs offers the lowest integrated absolute error, thereby validating the proposed categorization. Even though all three model structures give small errors at the ultimate frequency, the correct category results in the lowest overall IAE_{ω} . With the help of IAE_{ω} values computed for different model structures for various processes over

Table 1. Processes Studied and Corresponding Ultimate Properties

example	true process	K_u	P_u
1	$e^{-10s}/(s+1)^3$	1.274	25.348
2	$e^{-15s}/(s+1)^{15}$	1.2736	59.338
3	$e^{-10s}/(s+1)^2$	1.2735	23.358
4	$e^{-0.05s}/(25s+1)^2$	862.8633	5.434
5	$e^{-0.6s}/(6s+1)^8$	1.8424	91.94
6	$e^{-0.001s}/(s+1)^{20}$	1.3424	39.332

a wide range of D/τ values as well as different orders, a quantitative classification can be made as shown in Figure 4.

4. Implications for Control

After identifying the appropriate model structure and associated model parameters, different tuning rules can be designed to achieve improved performance.

4.1. PID Control. 4.1.1. Category 1: FOPDT. Following Luyben's approach,¹⁰ different tuning formulas can be applied for different D/τ values. PI controllers are used here, but the approach can be extended to PID controllers with little difficulty.

1. $D/\tau < 0.1$. For processes in category 1 having D/τ ratios less than 0.1, the Tyreus–Luyben tuning rule is found to be suitable.

The Tyreus–Luyben tuning equations¹³ for PI controllers are

$$K_c = \frac{K_u}{3.2} \quad (31)$$

$$\tau_I = 2.2P_u \quad (32)$$

2. $0.1 \leq D/\tau \leq 1$. The minimum ITAE tuning rule developed by Rovira¹¹ is found to be suitable for FOPDT processes in category 1 with D/τ ratios ranging from 0.1 to 1.

The ITAE tuning equations¹¹ for a PI controller are

$$K_c = \frac{0.586}{K_p} \left(\frac{\tau}{D} \right)^{0.916} \quad (33)$$

$$\tau_I = \frac{\tau}{1.03 - 0.165 \left(\frac{D}{\tau} \right)} \quad (34)$$

3. $D/\tau > 1$. For processes in category 1 having D/τ ratios greater than 1, the PI controller with the IMC tuning rule is found to be suitable.

The IMC tuning equations¹⁴ for a PI controller are

$$\lambda = \max(1.7D, 0.2\tau) \quad (35)$$

$$K_c = \frac{\tau + \frac{D}{2}}{K_p \lambda} \quad (36)$$

$$\tau_I = \tau + \frac{D}{2} \quad (37)$$

The various tuning rules for FOPDT processes with different D/τ ratios are summarized in Table 4. For unstable FOPDT systems, the tuning rules given by Tan et al.,⁴ Huang and Chen,⁵ Marchetti et al.,⁶ and Jacob and Chidambaram¹⁵ can be used.

4.1.2. Category 2: SOPSDT. For second-order plus small dead time system, Ziegler–Nichols tuning gives

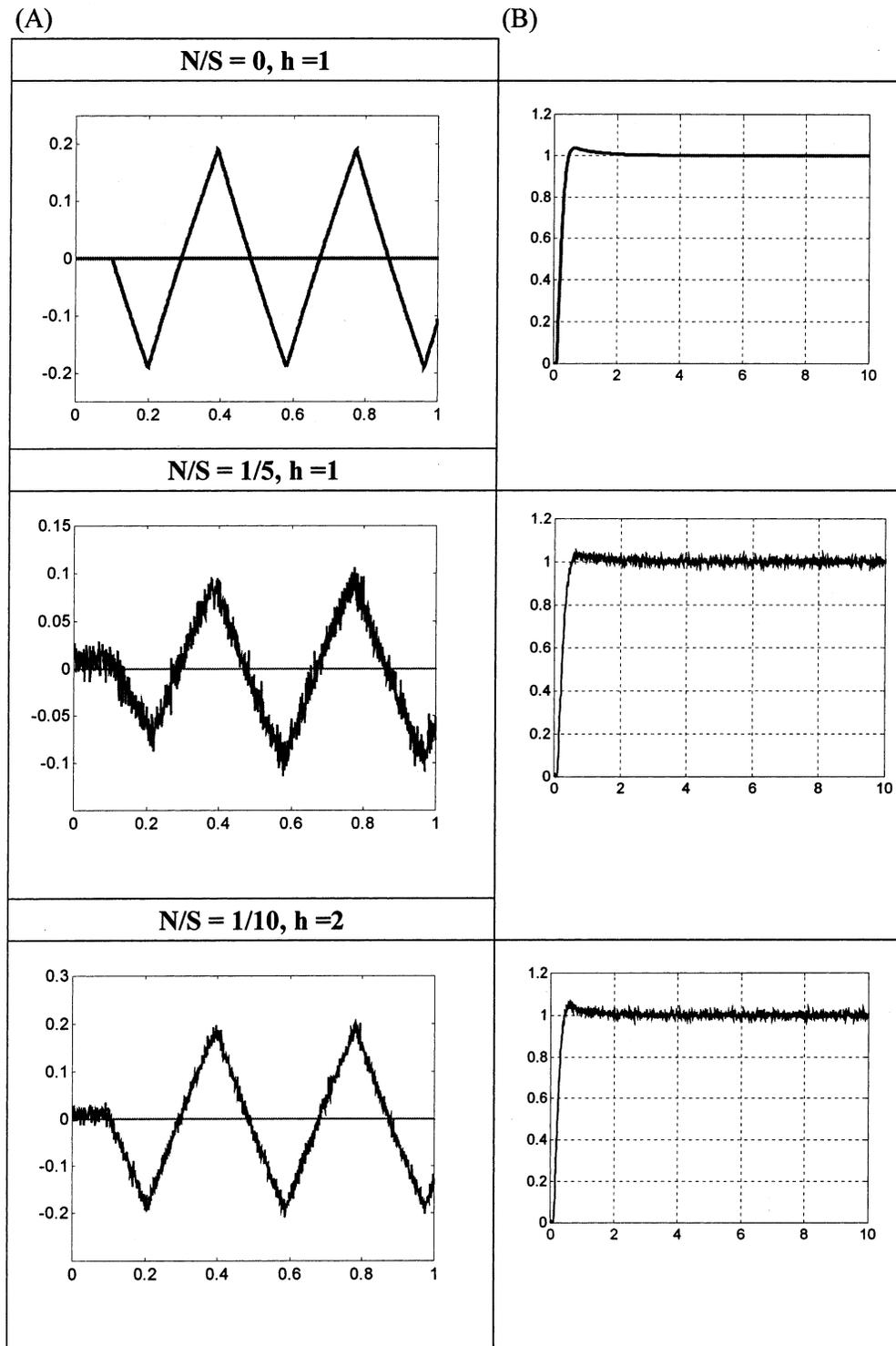


Figure 12. Performance of an FOPDT process ($D/\tau = 0.1$) with different levels of measurement noise (NSR = 0, $1/5$, and $1/10$): (A) relay feedback responses, (B) set-point responses with a PI controller.

poor performance. It should also be emphasized that this category is just an approximate modeling of the true process, so modeling error should be expected. For SOPSDT processes with small values of D/τ (i.e., $D/\tau = 0.01-0.001$), we find the following rules appropriate. For a PI controller, first set

$$\tau_I = 2\tau \quad (38)$$

Then, find that value of K_c that gives a 45° phase margin. For $D/\tau = 0.001$, this gives

$$K_c = \frac{2[(0.0432\tau\omega_u)^3 + 0.0432\tau\omega_u]}{K_p[\sqrt{4(0.0432\tau\omega_u)^2 + 1}]} \quad (39)$$

4.1.3. Category 3: HO. Again, this is just an approximate model of a large variety of processes (Figure 4), and therefore, a conservative tuning rule should be devised. For HO processes with orders ranging from 3 to 10 (i.e., $n = 3-10$), first, set the reset time to

$$\tau_I = (n - 1)\tau \quad (40)$$

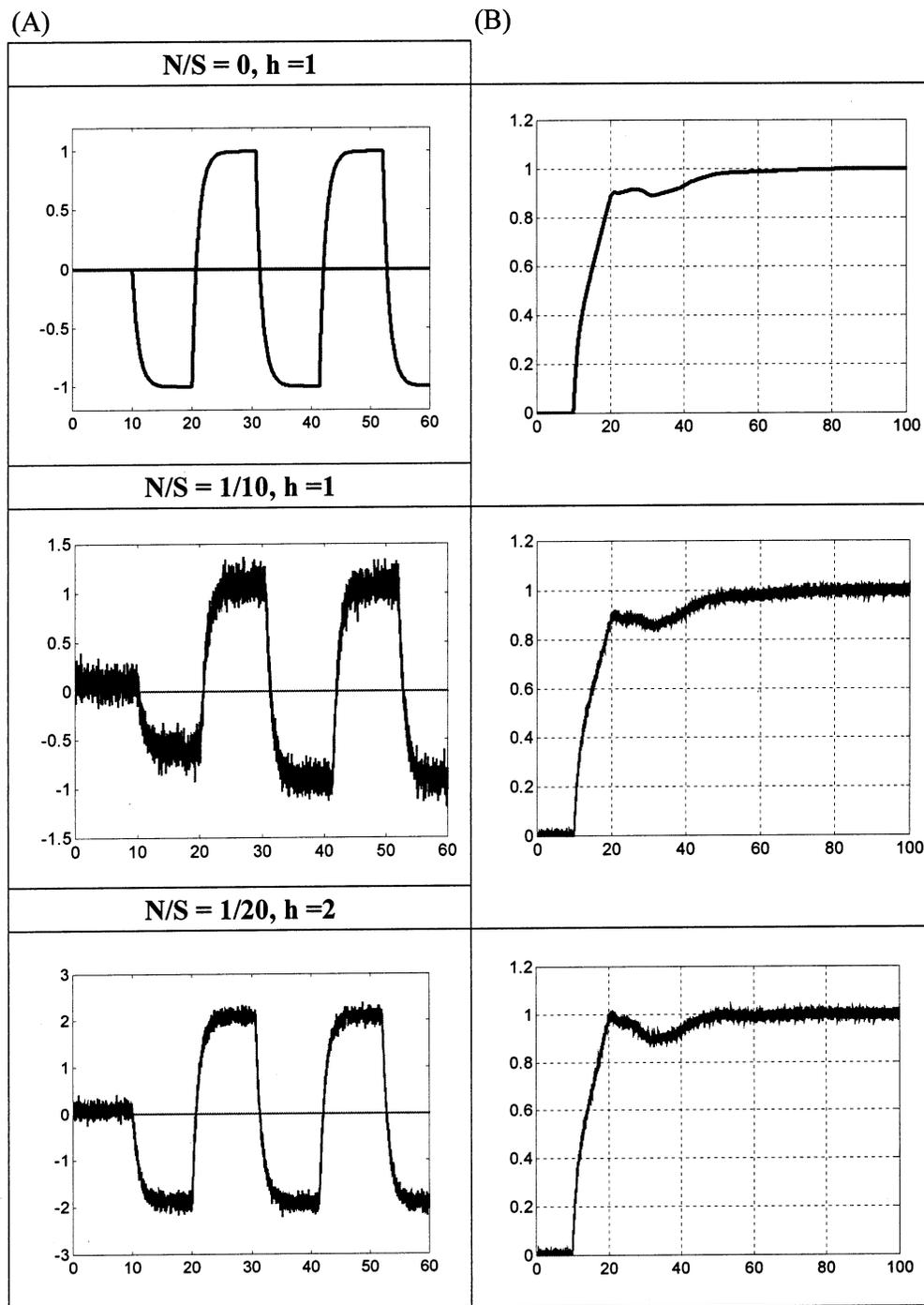


Figure 13. Performance of an FOPDT process ($D/\tau = 10$) with different levels of measurement noise ($NSR = 0, 1/5,$ and $1/10$): (A) relay feedback responses, (B) set-point responses with a PI controller.

Next, adjust K_c to give a closed-loop logarithmic modulus (L_c^{\max}) of 3 dB. For the default value of $n = 5$, this gives a very simple expression for K_c

$$K_c = \frac{1}{K_p} \quad (41)$$

4.2. Results. 4.2.1. Noise-Free Systems. After the exact category of the true process and the appropriate control strategy had been identified, closed-loop studies were carried out on all six examples listed in Table 1.

PI controllers were designed on the basis of the identified models. For example, for the third-order plus dead time process of example 1, three different PI

controllers were designed according to the FOPDT, SOPSDT, and HO model structures (i.e., first row of Table 2) using the tuning rules presented in Table 4, and eqs 38–41. This procedure was repeated for all six examples. Table 5 gives the controller settings for the examples studied. The effects of model structures on closed-loop performance can thus be compared. Closed-loop studies were carried out on the true processes. The set-point responses of three different controller settings (from different model structures) on the six examples are presented in Figure 10. A close look at the responses reveals the following:

For examples 1–3, the responses obtained from the controller designed using the FOPDT model structure

Table 2. True Process and Equivalent Models by Assuming Different Model Structures

example	true process	FOPDT	SOPSDT	HO
1	$e^{-10s/(s+1)^3}$	$0.8724e^{-10.8567s}/(1.9574s+1)$	$1615.9378e^{-0.1830s}/(182.95s+1)^2$	$2.2648/(2.9310s+1)^5$
2	$e^{-15s/(s+1)^{15}}$	$0.8970e^{-24.9161s}/(5.2159s+1)$	$1615.9378e^{-0.1830s}/(182.95s+1)^2$	$2.2656/(6.8613s+1)^5$
3	$e^{-10s/(s+1)^2}$	$1.0169e^{-10.0017s}/(2.4497s+1)$	$1615.9645e^{-0.1686s}/(168.6s+1)^2$	$2.2658/(2.7009s+1)^5$
4	$e^{-0.05s/(25s+1)^2}$	$0.0234e^{-1.3233s}/(21.2556s-1)$	$2.3822e^{-0.0392s}/(39.2s+1)^2$	$0.0033/(0.6283s+1)^5$
5	$e^{-0.6s/(6s+1)^8}$	$12.0599e^{-23.6439s}/(400.6714s+1)$	$1118.8082e^{-0.6637s}/(663.7s+1)^2$	$1.5661/(10.6313s+1)^5$
6	$e^{-0.001s/(s+1)^{20}}$	$4.6861e^{-10.835s}/(47.9111s+1)$	$1533.106e^{-0.2839s}/(283.91s+1)^2$	$2.1494/(4.5479s+1)^5$

Table 3. Comparison of Integrated Frequency Response Errors for Different Model Structures

example	true process	integrated absolute error		
		FOPDT	SOPSDT	HO
1	$e^{-10s/(s+1)^3}$	1.3083×10^{-4}	7.5365×10^{-3}	9.9724×10^{-4}
2	$e^{-15s/(s+1)^{15}}$	1.0958×10^{-4}	7.5468×10^{-3}	9.9986×10^{-4}
3	$e^{-10s/(s+1)^2}$	3.7721×10^{-5}	7.5233×10^{-3}	9.9337×10^{-4}
4	$e^{-0.05s/(25s+1)^2}$	9.9766×10^{-3}	2.3097×10^{-4}	5.1679×10^{-3}
5	$e^{-0.6s/(6s+1)^8}$	1.1144×10^{-3}	5.6718×10^{-3}	4.7446×10^{-4}
6	$e^{-0.001s/(s+1)^{20}}$	1.1952×10^{-3}	7.309×10^{-3}	9.3339×10^{-4}

Table 4. Tuning Rules for FOPDT Processes with Different D/τ Ratios

method	$D/\tau < 0.1$	$0.1 \leq D/\tau \leq 1$	$D/\tau > 1$
	TL	ITAE	IMC
formula	$K_c = \frac{K_u}{3.2}$	$K_c = \frac{0.586(\tau/D)^{0.916}}{K_p}$	$\lambda = \max(1.7D, 0.2\tau)$
	$\tau_1 = 2.2P_u$	$\tau_1 = \frac{\tau}{1.03 - 0.165(D/\tau)}$	$K_c = \frac{\tau + D}{K_p\lambda}$
			$\tau_1 = \tau + \frac{D}{2}$

Table 5. PI Controller Parameters for Different Examples

true process	FOPDT-IMC		SOPSDT-PM		HO- L_c^{\max}	
	K_c	τ_1	K_c	τ_1	K_c	τ_1
$e^{-10s/(s+1)^3}$	0.4587	7.3858	0.0023	365.9	0.4418	11.724
$e^{-15s/(s+1)^{15}}$	0.4652	17.6739	0.0023	856.64	0.4416	27.4452
$e^{-10s/(s+1)^2}$	0.4309	7.4506	0.0023	337.2	0.4415	10.8036
	FOPDT-TL					
$e^{-0.05s/(25s+1)^2}$	269.6448	11.9548	1.5279	78.4	303.15	2.5132
$e^{-0.6s/(6s+1)^8}$	0.5758	202.268	0.0033	1327.4	0.6387	42.5252
	FOPDT-ITAE					
$e^{-0.001s/(s+1)^{20}}$	0.4880	48.2641	0.0024	567.82	0.4654	18.1916

are superior to the other responses. This is because examples 1–3 fall into category 1, for which a PI controller with IMC tuning is the recommended controller ($D/\tau > 1$). The controller settings obtained using the SOPSDT model structure give very slow responses, whereas those obtained by assuming an HO process result in undershoot responses.

For example 4, the correct model structure (SOPSDT) gives a reasonable set-point response when compared to those of the other two model structures. The TL tuning (a conservative Ziegler–Nichols type of tuning) produces unstable responses for the second-order system with small dead-time-to-time-constant ratio, and the assumption of an HO model structure also fails to maintain stability. The reason is that this is almost a double integrator process, which is difficult to control.

For examples 5 and 6, the responses obtained from the controller designed using the equivalent HO models are superior to the other responses. The FOPDT model structure results in undershoot responses, whereas the SOPSDT equivalent model gives even more sluggish set-point responses. Note that the controller gain (K_c) of the SOPSDT model is almost 2 orders of magnitude smaller than that of the reasonable model.

Table 6. True Process, Equivalent Models, and Corresponding PI Controller Settings for Different Noise-to-Signal Ratios^a

	true	NSR = 0	NSR = $1/10$	NSR = $1/5$
G	$\frac{e^{-0.1s}}{(s+1)}$	$1.0091e^{-0.1s}/(1.0118s+1)$	$0.5526e^{-0.1036s}/(0.5414s+1)$	$0.3962e^{-0.1089s}/(0.3717s+1)$
K_c	4.1900	4.1900	4.383	3.9527
τ_1	0.8404	0.8404	0.8382	0.8492

^a FOPDT with $D/\tau = 0.1$.

Table 7. True Process, Equivalent Models, and Corresponding PI Controller Settings for Different Noise-to-Signal Ratios^a

	true	NSR = 0	NSR = $1/20$	NSR = $1/10$
G	$\frac{e^{-10s}}{(s+1)}$	$0.9998e^{-10.0010s}/(1.0258s+1)$	$1.0026e^{-9.1667s}/(1.6449s+1)$	$1.2527e^{-8.7135s}/(2.3101s+1)$
K_c	0.3529	0.3525	0.3986	0.3766
τ_1	6	6.0263	6.2282	6.6668

^a FOPDT with $D/\tau = 10$.

A special case under the SOPSDT category for which the default value of $\epsilon = 0.001$ does not work very well is also considered below. Consider the SOPSDT process represented by

$$G = \frac{e^{-0.01s}}{(2s+1)^2} \quad (42)$$

The relay feedback test output response indicates that the process belongs to category 2, i.e., SOPSDT process with a model structure given by eq 5. The model parameters obtained using the procedure described in section 3.2 with the default value of $\epsilon = 0.001$ is given by

$$G_m = \frac{6.027e^{-0.05s}}{(5s+1)^2} \quad (43)$$

The controller settings obtained ($K_c = 0.6038$ and $\tau_1 = 10$) using eqs 38 and 39 with the above model parameters results in a slow set-point response (Figure 11). Hence, there is a need to select an appropriate value of ϵ using the procedure described in section 3.2. The model parameters thus obtained with the modified value of $\epsilon = 0.00685$ are given in the equation

$$G_m = \frac{0.8709e^{-0.013s}}{(1.8978s+1)^2} \quad (44)$$

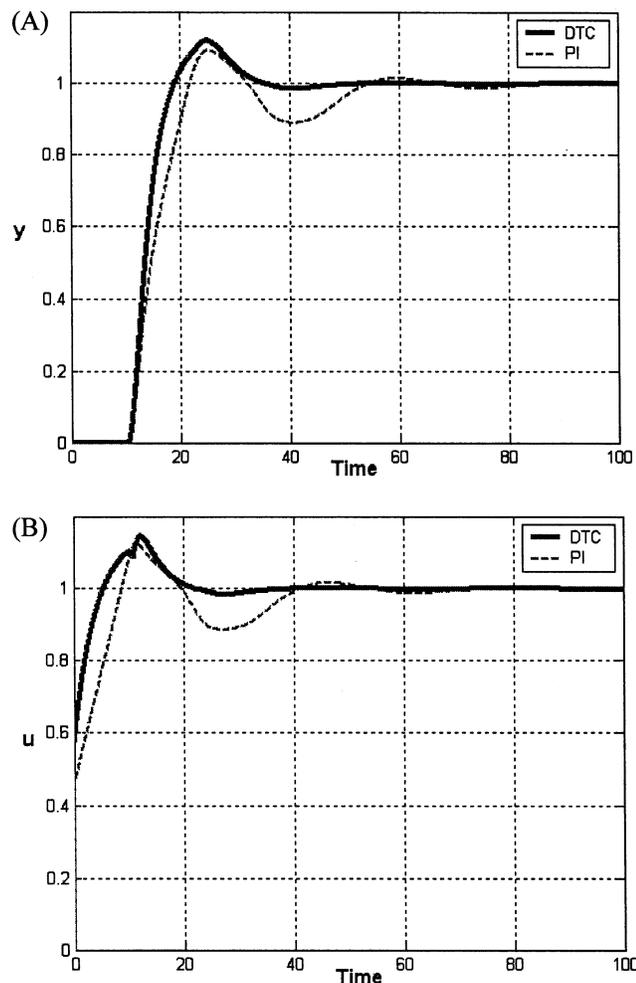


Figure 14. Set-point responses for a process with large dead time (category 1b) using a dead-time compensator (solid lines) and a PI controller (dashed lines).

The controller settings obtained ($K_c = 4.1794$ and $\tau_I = 3.7956$) using eqs 38 and 39 with the above model parameters result in a better set-point response (Figure 11), on a par with that obtained using the true process ($K_c = 3.6398$ and $\tau_I = 4$).

Thus, it is clear that the appropriate model structure with suitable tuning rules offers better closed-loop performance. More importantly, the improvement is achieved by taking the shape factor in the relay feedback response into account.

4.2.2. Systems with Noise. In a real process environment, measurement noise is unavoidable. The proposed method was tested against measurement noise. In the context of system identification, the noise-to-signal ratio (NSR) can be expressed as⁷

$$\text{NSR} = \frac{\text{mean}[\text{abs}(\text{noise})]}{\text{mean}[\text{abs}(\text{signal})]} \quad (45)$$

where $\text{abs}(\cdot)$ denotes the absolute value and $\text{mean}(\cdot)$ represents the mean value.

The following two FOPDT processes are used to illustrate the effect of process noises

$$G(s) = \frac{e^{-0.1s}}{(s+1)} \quad (46)$$

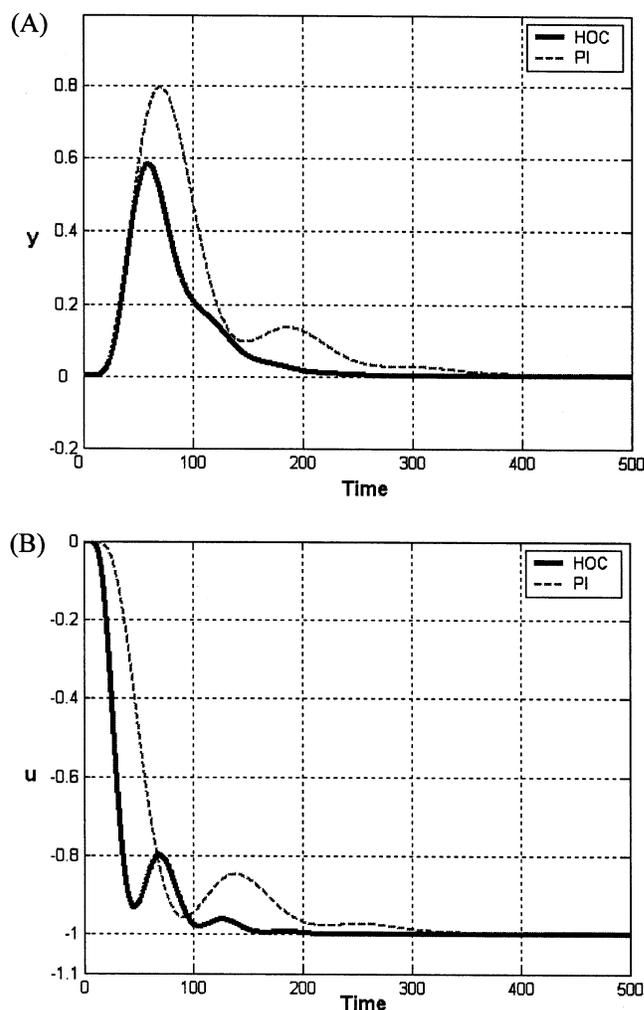


Figure 15. Load responses for a higher-order process (category 3) using a higher-order controller (solid lines) and a PI controller (dashed lines).

$$G(s) = \frac{e^{-10s}}{(s+1)} \quad (47)$$

In the case of the process represented by eq 46, relay feedback tests were performed with $\text{NSR} = 0$ and $1/5$ and with a relay height of 1. The relay feedback responses thus obtained are shown in Figure 12A. The limit cycle data was computed by taking the average of the fictitious peaks around the peak. Two cycles were employed to compute the average values of the limit cycle data. The FOPDT equivalent model (Table 6) was formulated using the procedure described in section 3.1.1. Tyreus–Luyben tuning rules (eqs 31 and 32) were used to tune the PI controller settings (Table 6). The set-point responses thus obtained are shown in Figure 12B and are on a par with the responses obtained for the true noise-free process. However, when the process was corrupted with noise, there was model mismatch in the FOPDT equivalent model derived. This mismatch can be alleviated by increasing the strength of the signal ($h = 2$, $\text{NSR} = 1/10$). Better FOPDT equivalent model parameters and controller settings were obtained (Table 6) by following a procedure similar to that described above. Figure 12A shows that the closed-loop performance results satisfactory set-point responses even under the influence of measurement noise.

Similarly, for the process represented by eq 47, relay feedback tests were performed with $NSR = 0$ and $1/10$ and with a relay height of 1. The corresponding relay feedback responses thus obtained are shown in Figure 13A. The limit cycle data were computed in a manner similar to that carried out for the process represented by eq 46. The FOPDT equivalent model was formulated using the procedure described in section 3.1.2. PI controller settings were obtained using the tuning rules given in eqs 35–37. Figure 13B shows the set-point responses obtained using the above controller settings. Even though the closed-loop response obtained for $NSR = 1/10$ is on a par with that obtained for the noise-free process, the FOPDT model formulated has a mismatch in the estimated model parameters (Table 7). An increase in the strength of the signal ($NSR = 1/20$, $h = 2$) can alleviate the model parameter mismatch. The PI controller settings derived using the above model parameters result in satisfactory set-point response, as shown in Figure 13B.

The observations from Figures 12B and 13B indicate that the proposed method works well even in the presence of measurement noise, resulting in satisfactory closed-loop performance.

4.3. Extension. Because the model structure and corresponding parameters are available, dead-time compensation or higher-order compensation can be provided whenever necessary. For the purpose of illustration, IMC¹⁵ was used to design controllers for an FOPDT system with a large D/τ value and for HO systems.

4.3.1. Dead-Time-Dominant Process. For processes in category 1b, a dead-time-compensating controller can be designed to improve the performance. Given the following model (G_m) and IMC filter F (a first-order filter)

$$G_m = \frac{K_p e^{-Ds}}{\tau s + 1} \quad (48)$$

$$F = \frac{1}{\lambda s + 1} \quad (49)$$

the conventional controller (K) using the IMC design becomes

$$K = \frac{G_c}{1 - G_c G_m} = \frac{1}{K_p} \frac{(\tau s + 1)}{[(\lambda s + 1) - e^{-Ds}]} \quad (50)$$

In this work, the filter time constant is set to

$$\lambda = 2\tau \quad (51)$$

Example 1 ($D/\tau = 5.5465$) is used to illustrate the potential improvement, Figure 14 shows that better set-point responses can be obtained for systems with large D/τ values.

4.3.2. Higher-Order Processes. For processes in category 3, a higher-order controller can be used to improve the control performance. Again, IMC control was employed with the following model (G_m) and filter (F)

$$G_m = \frac{K_p}{(\tau s + 1)^5} \quad (52)$$

$$F = \frac{1}{(\lambda s + 1)^5} \quad (53)$$

The recommended value for λ is 0.45τ . Thus, with the known values of G_m and F , the equivalent controller in the conventional feedback structure becomes

$$K = \frac{1}{5K_p\lambda s} \left[\frac{(\tau s + 1)^5}{(\lambda^4/5)s^4 + \lambda^3 s^3 + 2\lambda^2 s^2 + 2\lambda s + 1} \right] \quad (54)$$

Example 5 is used to illustrate the potential improvement. Figure 15 shows that a better load response is achieved using the higher-order controller without exciting the manipulated variable excessively.

5. Conclusion

The shapes of relay feedback responses are useful in extracting additional information about process dynamics. From a methodical analysis of the shape information, different processes can be broadly classified into three major categories. Analytical expressions (eqs 8, 9, 14, and 15) for the responses of stable and unstable FOPDT processes can be used to derive all three model parameters. From the insight gained, the identification procedures for different processes under various categories were evolved. Different tuning rules were employed to find appropriate PID controller settings. Procedures are tested against linear systems with and without noise. Further, dead-time compensation (for category 1b) and higher-order compensation (for category 3) can also be devised whenever necessary. The results show that the proposed method results in improved autotuning in a straightforward manner. Thus, shape information is useful in inferring the correct model structure of the process and also in selecting the appropriate control strategy to offer improved performance without the need for any additional testing.

Acknowledgment

We thank Prof. H. P. Huang for useful comments on shapes of relay feedback tests for second- and higher-order systems. This work was supported by the National Science Council of Taiwan.

Nomenclature

a = amplitude of output response
 D = dead time
DTC = dead time compensator
 F = filter transfer function
FOPDT = first-order plus dead time
 G = process transfer function
 G_m = model of the process
 h = relay magnitude
HO = higher-order
HOC = higher-order controller
IAE = integrated absolute error
IMC = internal model control
ITAE = integral of time-weighted absolute error
 K = controller transfer function
 K_c = controller gain
 K_p = steady-state gain
 K_u = ultimate gain
 L_c^{\max} = maximum closed-loop logarithmic modulus
 n = order of the process

P_u = ultimate period
 PI = proportional–integral controller
 PM = phase margin
 s = Laplace transform variable
 SOPDT = second-order plus dead time
 SOPSDT = second-order plus small dead time
 t_a = time to the peak amplitude
 $t_{a/2}$ = time to one-half of the peak amplitude
 TL = Tyreus–Luyben tuning
 u = manipulated variable
 y = controlled variable
 ϵ = constant with a value of 0.001
 λ = IMC tuning parameter
 τ = process open-loop time constant
 τ_G = global time constant characterized by the global relay feedback responses (defined by the peaks of cycling)
 τ_I = controller integral time
 ω_u = ultimate frequency

Literature Cited

- (1) Åström, K. J.; Hägglund, T. Automatic tuning of simple regulators with specifications on phase and amplitude margins. *Automatica* **1984**, *20*, 645–651.
- (2) Luyben, W. L. Derivation of transfer functions for highly nonlinear distillation columns. *Ind. Eng. Chem. Res.* **1987**, *26* (12), 2490–2495.
- (3) Chang, R. C.; Shen, S. H.; Yu, C. C. Derivation of transfer function from relay feedback systems. *Ind. Eng. Chem. Res.* **1992**, *31* (3), 855–860.
- (4) Tan, K. K.; Wang, Q. G.; Lee, T. H. Finite spectrum assignment control of unstable time delay processes with relay tuning. *Ind. Eng. Chem. Res.* **1998**, *37* (4), 1351–1357.
- (5) Huang, H. P.; Chen, C. C. Autotuning of PID controllers for second-order unstable process having deadtime. *J. Chem. Eng. Jpn.* **1999**, *32* (4), 486–497.
- (6) Marchetti, G.; Scali, C.; Lewin, D. R. Identification and control of open-loop unstable processes by relay methods. *Automatica* **2001**, *37*, 2049.
- (7) Wang, Q. G.; Hang, C. C.; Zou, B. Low-order modeling from relay feedback. *Ind. Eng. Chem. Res.* **1997**, *36* (2), 375–381.
- (8) Kaya, I.; Atherton, D. P. Parameter estimation from relay autotuning with asymmetric limit cycle data. *J. Process Control* **2001**, *11*, 429–439.
- (9) Yu, C. C. *Autotuning of PID Controllers*; Springer-Verlag: London, 1999.
- (10) Luyben, W. L. Getting more information from relay feedback tests. *Ind. Eng. Chem. Res.* **2001**, *40* (20), 4391–4402.
- (11) Smith, C. A.; Corripio, A. B. *Principles and Practice of Automatic Process Control*; John Wiley and Sons: New York, 1997.
- (12) Friman, M.; Waller, K. V. Autotuning of multiloop control systems. *Ind. Eng. Chem. Res.* **1994**, *33* (7), 1708–1717.
- (13) Tyreus, B. D.; Luyben, W. L. Tuning PI controllers for integrator–dead time processes. *Ind. Eng. Chem. Res.* **1992**, *35* (10), 3480–3483.
- (14) Morari, M.; Zaririou, E. *Robust Process Control*; Prentice Hall: Englewood Cliffs, NJ, 1989.
- (15) Jacob, E. F.; Chindambaram, M. Design of controllers for unstable first-order plus time delay systems. *Computers Chem. Eng.* **1996**, *20* (5), 579–584.
- (16) Luyben, W. L. Effect of derivative algorithm and tuning selection on the PID control of dead-time processes. *Ind. Eng. Chem. Res.* **2001**, *40* (16), 3605–3611.

Received for review December 13, 2001
 Revised manuscript received March 13, 2002
 Accepted March 15, 2002

IE011006F