Relay-Based Autotuning of PID Controller for Improved Load Disturbance Rejection

Tao Liu, Furong Gao*

*Department of Chemical Engineering, Hong Kong University of Science & Technology, Kowloon, Hong Kong (Tel: +852-2358-7139; e-mail addresses: liu@router.ieee.org; kfeng@ust.hk)

Abstract: To overcome sluggish load disturbance response for industrial/chemical processes with slow time constant(s), an improved design for on-line autotuning of proportional-integral-derivative (PID) controller is proposed in this paper, based on relay identification of the widely used first-order-plus-dead-time (FOPDT) process model. Using the fitting conditions established for process response at the oscillation frequency under a relay test, the identification algorithm is transparently developed. An analytical controller tuning method is then developed using an asymptotic constraint established thereby for reducing the influence of the slow process time constant on load disturbance rejection. Illustrative examples are given to show the effectiveness and merits of the proposed algorithms.

1. INTRODUCTION

Sluggish load disturbance response is usually resulted for processes with slow time constant(s), as widely recognized in the process industry. To deal with this problem, model-based controller tuning methods have been effectively developed (Morari and Zafiriou, 1989; Seborg, Edgar, and Mellichamp, 2003). Recently improved methods for tuning the proportional-integral-derivative (PID) controller, which is most commonly used in engineering practice, can be found in the literature (Piazzi and Visioli, 2006; Leva, Bascetta and Schiavo, 2005; Sree, Srinivas and Chidambaram, 2004; Skogestad, 2003; Ho et al., 2003; Sung, Lee and Park, 2002; Lee and Edgar, 2002; Hwang and Hsiao, 2002; Tan, Lee and Jiang, 2001). For such a controller design, a low-order process model is needed, of which the first-order-plus-dead-time (FOPDT) model structure is mostly used since it can effectively reflect the fundamental characteristics of process response, in particular for the low frequency range primarily referred to controller tuning (Åström and Hägglund, 1995; Yu, 2006). Relay identification for obtaining low-order process models has received increasing attention in the process control community (Atherton, 2006; Hang, Åström and Wang, 2001), owing to that such an identification test can be performed online while preventing process response from drifting too far away from the operational level required therein. Based on a single run of unbiased relay feedback, Luyben (2001) proposed a FOPDT modeling method by defining curvature factors for the relay response shapes of stable and unstable processes; Vivek and Chidambaram (2005) reported another FOPDT identification algorithm using the Fourier analysis of the process relay response; Huang, Jeng and Luo (2005) developed a simple formulation of FOPDT model for on-line tuning of PI/PID controllers. Based on a single run of biased relay test, Shen, Wu and Yu (1996) gave a FOPDT modeling method according to the sustained oscillation conditions from the describing function analysis; Wang, Hang and Zou (1997) derived a FOPDT algorithm using the algebra properties of periodic oscillation; Kaya and Atherton (2001) developed a FOPDT identification method based on the so-called A-locus analysis.

In this paper, two identification algorithms are respectively derived according to whether an unbiased or biased relay test is used. Based on the internal model control (IMC) theory, (Morari and Zafiriou, 1989), a modified IMC filter design is proposed to derive the PID controller within the framework of a unity feedback control structure. As a result, apparently improved disturbance rejection performance can be obtained.

2. RELAY IDENTIFICATION ALGORITHMS

In a relay feedback test for identification, the relay function is usually specified as

\[
    u(t) = \begin{cases} 
    u_+ & \text{for } \{e(t) > e_+ \text{ and } \dot{e}(t) = 0\} \\
    u_- & \text{for } \{e(t) < -e_- \text{ and } \dot{e}(t) = 0\} 
    \end{cases}
\]

where \( u_+ = \Delta \mu + \mu_0 \) and \( u_- = \Delta \mu - \mu_0 \) denote, respectively, the positive and negative relay magnitudes; \( e_+ \) and \( e_- \) denote, respectively, the positive and negative relay switch hysteresis. Note that letting \( u_+ = -u_- \) and \( e_+ = -e_- \) leads to an unbiased relay function.

When the process response moves into the limit cycle under relay feedback, the process output becomes a periodic signal with the oscillation angular frequency, \( \alpha_\omega = 2\pi / P_\omega \). By using the idea of time shift, we may view it as a periodic signal from the very beginning, so its Fourier transform can be derived as

\[
    Y(j\alpha_\omega) = \lim_{N \to \infty} N \int_0^N y_m(t)e^{-j\alpha_\omega t}dt = \lim_{N \to \infty} \int_{-\infty}^{\infty} y(t)e^{-j\alpha_\omega t}dt
\]

where \( y_m(t) = y(t) \) for \( t \in [t_m, \infty) \) and \( t_m \) may be taken as any relay switch point in steady oscillation, such that the influence from the initial response can be excluded.
Similarly, it follows that
\[ U(j\omega) = \lim_{N \to \infty} \int_{t_m}^{t_m+N} u(t)e^{-j\omega t}dt \] (3)

Thereby, the process frequency response at \( \omega \) can be obtained as
\[ G(j\omega) = \frac{Y(j\omega)}{U(j\omega)} = \int_{t_m}^{t_m+N} y(t)e^{-j\omega t}dt \] (4)

Note that the Laplace transform for the relay response can be naturally decomposed as
\[ Y(s) = \int_{t_m}^{\infty} y(t)e^{-st}dt + \int_{t_m}^{\infty} y(t)e^{-st}dt \] (5)

where \( t_{m+1} \) denotes the moment after which \( y(t) \) becomes a periodic signal. The second integral in (5) can be derived as
\[ \int_{t_m}^{\infty} y(t)e^{-st}dt = \lim_{M \to \infty} \int_{t_m}^{t_{m+M}} y(t)e^{-st}dt = \lim_{M \to \infty} \frac{1-e^{-st}}{st} \int_{t_m}^{t_{m+M}} y(t)e^{-st}dt \]

For \( \text{Re}(s) > 0 \), there exists \( e^{-st} \to 0 \) as \( n \to \infty \). Hence, we can obtain for \( \text{Re}(s) > 0 \) that
\[ Y(s) = \int_{t_m}^{\infty} y(t)e^{-st}dt + \frac{1}{1-e^{-st}} \int_{t_m}^{t_{m+P}} y(t)e^{-st}dt \] (6)

Likewise, the Laplace transform of the relay output for \( \text{Re}(s) > 0 \) can be derived as
\[ U(s) = \int_{t_m}^{\infty} u(t)e^{-st}dt + \frac{1}{1-e^{-st}} \int_{t_m}^{t_{m+P}} u(t)e^{-st}dt \] (7)

Therefore, the process transfer function for \( \text{Re}(s) > 0 \) can be obtained as
\[ G(s) = \frac{Y(s)}{U(s)} = \frac{1-e^{-st}}{1-e^{-st}} \int_{t_m}^{t_{m+P}} y(t)e^{-st}dt \] (8)

By substituting \( s = \alpha + j\omega \) into (8) we can obtain
\[ G(j\omega + \alpha) = \frac{1-e^{-st}}{1-e^{-st}} \int_{t_m}^{t_{m+P}} y(t)e^{-st}dt \] (9)

where \( \alpha \in (0, \infty) \) may be viewed as a constant shift operator of Laplace transform. Note that \( G(j\omega + \alpha) \to 0 / 0 \) as \( \alpha \to \infty \). It is therefore suggested to choose \( \alpha \) with a numerical constraint of \( \min[\text{Re}(t_{m+P} + e^{-st})] < \text{Re}(t_{m+P} + e^{-st})] > 10^4 \), such that both the initial and steady responses of \( y(t) \) and \( u(t) \) under relay feedback can be effectively included in the computation of (9).

To identify a FOPDT model generally in the form of
\[ G_m = \frac{k_p e^{-\tau_d}}{\tau s + 1} \] (10)

we can establish two fitting conditions by substituting (10) into (4), i.e.,
\[ \frac{k_p}{\sqrt{\dot{\tau}^2 + \tau^2 \omega^2}} = A_0 \]
\[ -\theta_0 \omega - \arctan(\tau \omega) = \varphi_0 \]

By substituting (10) into (9) and letting \( G(j\omega + \alpha) = A_0 e^{j\varphi_0} \), we obtain another fitting condition,
\[ \frac{k_p e^{-\alpha\tau}}{\sqrt{(\alpha^2 + \tau^2 \omega^2)}} = A_0 \]

For the case that a biased relay test is used, the process gain can be derived as
\[ k_p = G(0) = \frac{\int_{t_m}^{t_{m+P}} y(t)dt}{\int_{t_m}^{t_{m+P}} u(t)dt} \] (14)

Accordingly, the other two model parameters, \( \tau \) and \( \theta \), can then be derived from (11) and (12), respectively.

For the case that an unbiased relay test is used, the three model parameters, \( k_p \), \( \tau \), and \( \theta \), can be derived by solving (11-13) together. To perform numerical computation possibly involved thereby, iterative algorithms such as the Newton-Raphson method may be used, and correspondingly, a relay response fitting constraint can be adopted to determine a suitable solution, i.e.,
\[ \sum_{k=1}^{N} [y(kT_s + t_m) - y(kT_s + t_{m+1})]^2 / N < \epsilon \] (15)

where \( y(kT_s + t_m) \) and \( y(kT_s + t_{m+1}) \) denotes respectively the model and process responses, \( T_s \) is the sampling period and \( N = P_s / T_s \), and \( \epsilon \) is a user-specified fitting threshold that may be set between 0.01% - 1%.

2. PID TUNING METHOD

Consider the unity feedback control structure shown in Fig.1.

![Fig. 1. Unity feedback control structure](image)

where \( C \) is a PID controller as mostly used in practice, \( d_i \) and \( d_d \) denote, respectively, load disturbances injected at the process input and output sides, and \( \dot{d}_d \) indicates the load disturbance with a transfer function of \( G_d \). In many industrial cases, the influence of \( \dot{d}_d \) may be transformed into \( d_i \) to be treated (Seborg, Edgar and Mellichamp, 2003).

It is well known that the IMC theory has been successfully applied to PID tuning within the framework of a unity
feedback control structure (Skogestad and Postlethwaite, 2005; Braatz, 1995). The key to the use of IMC theory for the closed-loop controller design lies with the choice of a suitable IMC filter to construct the desired closed-loop complementary sensitivity function. Given a process model $G_m = G_c G_m$, where $G_c$ is an all-pass portion and $G_m$ a minimum-phase portion, the IMC-based complementary sensitivity function can be ascertained as

$$T = G_\alpha f$$

where $f$ denotes the IMC filter. A conventional IMC filter of type I, $f(s) = 1/(\lambda s + 1)$, is generally chosen for step changes in set-point and load disturbance, and type II, $f(s) = (n\lambda s + 1)/(\lambda s + 1)$, is for ramp changes where $n$ is an integer large enough to make $f/G_m$ proper. Most of existing IMC-based tuning methods are based on the filter type I, since a step change in set-point or load disturbance can be physically regarded as a summation of sinusoidal signals of different frequencies. The key feature of IMC filter type I is that it can lead to the $H_\infty$ optimal performance objective for step change in set-point and the load disturbance acting on the process output side (i.e., $d_i$ shown in Fig.1). However, for a load disturbance that seeps into the process, denoted as $d_i$ in Fig.1, the corresponding transfer function is in the form of

$$H_{di} = GSD_i$$

where $S = 1 - T$ is the sensitivity function. It can be seen that the time constant(s) of $G$ is enclosed in the characteristic equation of $H_{di}$, and therefore, affects the achievable disturbance rejection, no matter how the IMC filter is tuned in $T$. This can be used to explain why the recovery trajectory of the disturbance response is subject to ‘a long tail’, i.e., sluggish load disturbance suppression, for a process with slow time constant(s).

To reduce the influence arising from the slow process time constant(s) to the load disturbance response, a good idea is to eliminate the corresponding pole(s) from the characteristic equation of the above load disturbance transfer function, as illustrated by Horn et al. (1996) for cancelling the slowest pole with rational approximation. It is thus expected that $1-T$ (i.e., $S$), rather than $T$, has the corresponding zero(s) to cancel the slow pole(s) of $G$, such that the load disturbance response is governed only by the time constant of $T$ (i.e., an adjustable parameter in the IMC filter). The numerator of $1-T$, however, is unavoidably involved with time delay factor(s) for a process with time delay, so it cannot be factorized to make exact zero-pole cancellation with the denominator of $G$. The following asymptotic constraint is therefore proposed to realize the above idea based on the identified FOPDT model of (10),

$$\lim_{s \to 0} (1-T) = 0$$

Correspondingly, the conventional IMC filter is rectified as

$$f_{RMC} = \frac{\alpha s + 1}{(\lambda_s s + 1)^2}$$

where $\alpha$ is an additional parameter used for satisfying the above asymptotic constraint. It follows from (10), (16) and (19) that

$$T_{RMC} = \frac{(\lambda s + 1)e^{-\theta}}{(\lambda_s s + 1)^2}$$

Substituting (20) into (18) yields

$$\alpha = \tau \left[1 - \frac{\lambda}{\tau} e^{-\frac{\theta}{\tau}}\right]$$

It is seen that $\alpha$ is a function of $\lambda$. Hence, there is essentially a single adjustable parameter, $\lambda$, in the proposed IMC filter.

According to the nominal closed-loop relationship

$$T = \frac{G_c C}{1 + G_c C}$$

we obtain by substituting (10) and (20) into (22) the desired closed-loop controller,

$$C = \frac{(\alpha s + 1)\tau (\tau s + 1)}{k_p((\lambda_s s + 1)^2 - (\alpha s + 1)e^{-\theta})}$$

It can be verified from (23) that

$$\lim_{s \to \infty} C = \infty$$

Hence, this controller has a property of integral for eliminating the steady-state error. To approximate it in a PID form for implementation, we hereby adopt the analytical approximation approach used in the recent literature (e.g., Liu et al., 2005a, b). Let $C = M/\tau$, it follows that

$$C = \frac{1}{\tau s}[M(0) + M'(0)s + \frac{M''(0)}{2!}s^2 + \cdots]$$

Accordingly, the first three terms in the above Maclaurin expansion constitute a PID controller, i.e.,

$$C_{PID} = k_c + \frac{1}{\tau_i s} + \frac{1}{\tau_d s}$$

where $k_c = M'(0)$ and $\tau_i = 1/M(0)$ and $\tau_d = M''(0)/2$. The pure derivative in (26) can be practically implemented by cascading with a first-order low-pass filter in which the time constant can be chosen as $(0.01 \sim 0.1)\tau_d$.

It should be noted that the rational high order approximation formula proposed in Liu et al. (2005a) in terms of the linear fractional Padé expansion may be used for obtaining further enhanced control performance.

Combining (21), (23) and (25), we can also see that the proposed PID controller is essentially tuned by the single adjustable parameter, $\lambda$, of the proposed IMC filter.

To explore the quantitative relationship between the disturbance response peak (DP) of a FOPDT process shown in (10) and the single tuning parameter $\lambda$, of the proposed PID controller, we may normalize the process model of (10) by scaling the Laplace variable as $s = \tau_s$, i.e.,
\[
\hat{G}_m = \frac{k}{s + 1}
\]

By substituting the scaled Laplace variable into (23), we can find that the tuning parameter is correspondingly scaled as \( \hat{\lambda} / \tau \) to obtain the same control effect. That is to say, using \( \hat{\lambda} / \tau \) for (27) can obtain the same DP with (10) in terms of \( \lambda \). Hence, we can study the quantitative relationship between DP / \( k_p \), \( \hat{\lambda} / \tau \) and \( \theta / \tau \), regardless of the variation of \( \tau \). Based on numerical computations and simulations, the result for a unity step change of the load disturbance added to the process (shown as \( d \) in Fig.1) is plotted in Fig.2.

Fig. 2. Disturbance response peak for FOPDT process

According to the small gain theorem (Zhou, Doyle and Glover, 1998), the closed-loop system shown in Fig.1 holds robust stability if and only if

\[
\|P\| \leq \frac{1}{|M|}
\]

where \( \Delta = (G - G_m)/G_m \) denotes the mismatch between the model and the real process. Given an upper bound of this mismatch, the admissible tuning range of \( \hat{\lambda} \) can be numerically determined by substituting (22) into (28). For instance, in the case that there exists the process time delay uncertainty \( \Delta \theta \), which may be converted to \( \Delta(s) = e^{-\Delta \theta} - 1 \), the robust stability constraint for tuning \( \hat{\lambda} \) of the proposed PID controller can be ascertained by substituting (22) and (26) into (28), i.e.,

\[
\left| \frac{G_m C_{PID}}{1 + G_m C_{PID}} \right| \leq \frac{1}{|e^{-\Delta \theta} - 1|}
\]

Based on the quantitative relationship given in Fig.2, it is suggested to tune \( \hat{\lambda} = \tau \) in the first place. Then by monotonously varying \( \hat{\lambda} \) online, the best trade-off between the nominal performance of the closed-loop system and its robust stability can be conveniently obtained.

4. ILLUSTRATION

Example 1. Consider the first-order process widely studied in the literature (see, e.g., Shen, Wu and Yu, 1996; Vivek and Chidambaram, 2005).

By using a biased relay test, Shen, Wu and Yu (1996) derived the process model, \( G_m = 0.999e^{2.0985s}/(8.118s + 1) \). Vivek and Chidambaram (2005) obtained the process model, \( G_m = 0.9467e^{2.1s}/(9.5028s + 1) \), from an unbiased relay feedback test. For comparison, unbiased \( (u = -u = 1) \) and biased \( (u = 1.3 \) and \( u = -0.7 \) ) relay tests with \( e = -e = 0.2 \) and \( \alpha = 0.1 \) are respectively performed, for which the intermediate values of the limit cycle for model identification are listed in Table 1.

Consequently, the proposed algorithm for unbiased relay test results in the model, \( G_m = 1.0048e^{2.0024s}/(10.049s + 1) \), and the proposed algorithm for biased relay test gives \( G_m = 1.0001e^{2.005s}/(10.001s + 1) \), both of which indicate good identification accuracy.

Suppose that a random noise of \( N(0, \sigma^2 = 0.045\%) \) is added to the process output measurement and feedback under the above biased relay test, causing the noise-to-signal ratio (NSR) to be 10%. Based on the statistical averaging of 10 steady oscillation periods for computation, the intermediate values of the limit cycle obtained thereby are also listed in Table 1. Correspondingly, the proposed algorithm for biased relay test results in \( G_m = 1.0236e^{1.9971s}/(10.233s + 1) \), indicating good identification robustness.

Example 2. Consider the high-order process studied in the recent literature (Wang et al., 1997; Kaya and Atherton, 2001)

\[
G = \frac{(-s + 1)e^{-s}}{(s + 3)^5}
\]

By using a biased relay test, Wang et al. (1997) derived a FOPDT model, \( G_m = 1.00e^{2.24s}/(2.99s + 1) \), and Kaya and Atherton (2001) gave the model, \( G_m = 1.00e^{-3.082s}/(2.292s + 1) \), both of which had shown their superiority over many other relay identification methods. For comparison, the biased relay test in example 1 is performed and correspondingly, the proposed algorithm results in the FOPDT model, \( G_m = 1.0001e^{4.8378s}/(2.3017s + 1) \). The Nyquist plots of these FOPDT models are shown in Fig.3. It can be seen that further improved fitting is captured by the proposed model. Note that the model response obtained hereby coincides with the real process at the oscillation frequency, i.e., (-0.686, -j0.1628), as shown in Fig.3.
Example 3. Consider the second-order process studied by Skogestad (2003)

\[ G_s = \frac{e^{-\lambda s}}{(20s+1)(2s+1)} \]

For improving disturbance rejection performance, a modified IMC design for tuning the closed-loop PID controller was developed in Skogestad (2003), of which the controller parameters are listed in Table 2. For illustration, an unbiased relay test as in example 1 is performed, of which the limit cycle data are listed in Table 1. Accordingly, the proposed identification algorithm results in the FOPDT model,

\[ G_m = 0.98e^{-2.7963s}/(21.8291s+1) \]

Based on this model, the proposed PID tuning method is used with \( \lambda = 0.9 \) to obtain the same DP with that of Skogestad (2003) for comparison. The corresponding controller parameters are listed in Table 2, together with the conventional IMC-based PID parameters obtained in terms of the exact process model, \( \lambda = 0.45 \) and the Maclaurin expansion.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Controller Parameters</th>
<th>( \lambda )</th>
<th>( \tau_i )</th>
<th>( \tau_d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed</td>
<td>13.6248</td>
<td>0.4129</td>
<td>16.263</td>
<td></td>
</tr>
<tr>
<td>Skogestad</td>
<td>12.5</td>
<td>0.8</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>IMC</td>
<td>11.6614</td>
<td>1.9</td>
<td>22.8324</td>
<td></td>
</tr>
</tbody>
</table>

The closed-loop system response for a step change of load disturbance injected into the process is shown in Fig.4. It can be seen that, to obtain the same DP, the proposed tuning method has reduced the recovery time by almost 80 percent in comparison with the conventional IMC filter, and by almost 50 percent in comparison with Skogestad (2003).

To demonstrate robust stability of the proposed tuning method, assume that there exists 20% error in modelling the original SOPDT parameters. The worst case is that the process time delay is actually 20% larger while the two time constants are actually 20% smaller. The corresponding load disturbance response is shown in Fig.5, which indicates that the proposed PID tuning method holds robust stability well in the presence of the severe process uncertainty.
5. CONCLUSIONS

Based on relay identification of the widely used FOPDT process model, an improved PID autotuning method has been developed. The proposed identification algorithms can be used for both unbiased and biased relay tests, and can result in improved fitting for the process frequency response compared to some existing FOPDT identification methods recently developed, in particular for the referred low frequency range for controller tuning. By proposing a modified IMC filter to reduce the influence of the slow process time constant on the load disturbance response, the corresponding PID controller in a unity feedback control structure has been analytically derived using the Maclaurin approximation. An important merit of the proposed PID design is that there is essentially a single adjustable parameter, which can be monotonously tuned on-line to meet the best compromise between the nominal closed-loop performance for load disturbance rejection and its robust stability. The quantitative tuning relationship between this adjustable parameter, DP, and the identified FOPDT model parameters has been given for practice reference.

ACKNOWLEDGEMENT

This work is supported in part by the Hong Kong Research Grants Council under Project No. 612906.

REFERENCES


