Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Non-Parametric Tuning of PID Controllers via Modified Second-Order Sliding Mode Algorithms

Igor M. Boiko*

**University of Calgary, Calgary, AB, T2W 3R8*
Canada (Tel: 403-281-4427; e-mail: i.boiko@ieee.org).

Abstract: An application of the modified second-order sliding mode (SOSM) algorithms to a PID controller tuning is presented. The proposed method utilizes the opportunity of exciting test oscillations in the controller-process loop at the frequency corresponding to a desired phase lag of the process. This allows for this frequency to become the phase crossover frequency in the closed-loop system with a PID controller if the tuning rules are formulated as non-parametric rules in terms of the "ultimate frequency" and "ultimate gain". The use of those properties results in designing a simple tuning method that provides the desired stability of the system. A simple test that involves measurements of the amplitude and of the frequency of the self-excited oscillations and non-parametric tuning rules that provide desired gain margins exactly are presented.

1. INTRODUCTION

PID control is the main type of control used in the process industries. PID controllers are usually implemented as configurable software modules within the distributed control systems (DCS). The DCS software is constantly evolving providing the developers and process control engineers with a number of new features. One of most useful features is the controller autotuning functionality. This trend can be seen in the new releases of such popular DCS as Honeywell Experion PKS® and Emerson DeltaV®. The practice of the use of a number of autotuning algorithms shows that many of them do not provide a satisfactory performance if the process is subject to noise, variable external disturbance or nonlinear. On the other hand the simplest algorithms such as closed-loop tuning method proposed by Ziegler and Nichols (1942) and the relay feedback test (Astrom & Hagglund, 1984) provide a satisfactory performance in those conditions despite the inherent relatively low accuracy of those methods. The explanation of this phenomenon lies in the area of analysis of the parametric versus non-parametric tuning. Apparently, the use of an underlying model of the process in a parametric method (not fully matching to the actual process dynamics) that usually has three or higher number of parameters may result in the significant deterioration of the identification-tuning accuracy if the test conditions are affected by noise, disturbances or nonlinearities. Only the most basic characteristics of the system, such as the ultimate gain and ultimate frequency (Ziegler & Nichols, 1942), remain almost unaffected in those conditions.

However, the use of only two measurements cannot ensure sufficient accuracy of tuning. For that reason, in practice, tuning almost always includes the steps of trial end error that follow the step of autotuning. Therefore, a trade-off between the accuracy and reliability of tuning (which also translates into accuracy) is apparent. The cause of the relatively low accuracy of the referenced (and other) non-parametric methods is well known. This is the use of only two measurements of the test over the process. Yet, it is also known that a satisfactory accuracy of identification for most processes can be achieved if at least a three-parameter model is used (Astrom & Hagglund, 1995).

There is one more factor that also contributes to the issue of accuracy. This is the widely accepted statement that *the most important test point for the subsequent tuning is the phase cross-over frequency*, i.e. the point in which the phase characteristic of the process is -180° (frequency ω_π). Yet, if the controller that is used to control this process is of PI-type (the most common option) then in the open-loop dynamics containing the PI controller the frequency ω_π is lower than the frequency ω_π of the closed-loop tests (Ziegler & Nichols, 1942) or (Astrom & Hagglund, 1984). The parametric methods of tuning that utilize the relay feedback test are also based on the measurements of the process characteristics at this frequency and do not account for the change of this frequency due to the controller introduction. Therefore, the choice of the test frequency and the means of generating this frequency of oscillations are considered in the paper.

The paper is organized as follows. At first the problem of selection of the test point on the frequency response of the process is analyzed. After that a modified relay feedback test that provides generation of the oscillations at a given point of the phase response of the process is proposed. Finally, tuning rules that provide a higher accuracy of non-parametric tuning in comparison with (Ziegler & Nichols, 1942) and (Astrom & Hagglund, 1984) are derived.

2. EFFECT OF TEST POINT SELECTION ON STABILITY OF CLOSED-LOOP SYSTEM

Consider the following motivating example, which illustrates how the introduction of the controller may affect the results of identification and tuning.

978-1-1234-7890-2/08/$20.00 © 2008 IFAC

10.3182/20080706-5-KR-1001.0276

Example 1. Let us assume that the certain process is given by the following transfer function (Note: this process model is used in a number of works as a test model for the subsequent approximation; see (Kaya & Atherton, 2001) for example):

$$W_p(s) = \mathrm{e}^{-2s} \frac{1}{(2s+1)^5}, \tag{1}$$

Find the first order plus dead time (FOPDT) approximating model to the process (1):

$$\hat{W}_p(s) = \frac{K_p e^{-\tau s}}{T_p s + 1}, \tag{2}$$

where K_p is the process static gain, T_p is the time constant, and τ is the dead time, so that both (1) and (2) produce the same ultimate gain and ultimate frequency in the closed-loop test (Ziegler & Nichols, 1942) or the same values of the amplitude and the ultimate frequency in the relay feedback test (Astrom & Hagglund, 1984). (Note: strictly speaking, the values of the ultimate frequency in tests (Ziegler & Nichols, 1942) and (Astrom & Hagglund, 1984) are slightly different. The frequency of the oscillations generated in the relay feedback test does not exactly correspond to the phase characteristic of the process -180°, which follows from the relay systems theory (Tzypkin, 1984), (Boiko, 2005)). Obviously, this problem has infinite number of solutions, as there are three unknown parameters of (2) and only two measurements obtained from the test. Assume for simplicity, that the value of the process static gain is known: K_p=1, and determine T_p and τ. Those parameters can be found from equation $\hat{W}_p(j\omega_\pi) = W_p(j\omega_\pi)$, where ω_π is the frequency that corresponds to the phase characteristic value -180°: $\arg W_p(j\omega_\pi) = -\pi$. The value of ω_π is 0.283, which gives $W_p(j\omega_\pi) = (-0.498, j0)$, and the FOPDT approximation is:

$$\hat{W}_p(s) = \frac{e^{-7.393s}}{6.153s + 1}, \tag{3}$$

The Nyquist plots of the process (1) and its approximation (3) are depicted in Fig. 1. The point of intersection of the two plots (denoted as Ω_0) is also the point of intersection with the real axis. Also $\Omega_0 = \omega_\pi$ for both process dynamics (1) and (3), and therefore $\hat{W}_p(j\Omega_0) = W_p(j\Omega_0)$.

If the designed controller is of proportional type then the stability margins (gain margins) for processes (1) and (3) are the same. However, if the controller is of PI type then the stability margins for (1) and (2) are different. Let us illustrate that. Design the PI controller given by the following transfer function:

$$W_c(s) = K_c\left(1 + \frac{1}{T_c s}\right), \tag{4}$$

using the tuning rules (Ziegler & Nichols, 1942). This results in the following transfer function of the controller:

$$W_c(s) = 0.803\left(1 + \frac{1}{17.76s}\right), \tag{5}$$

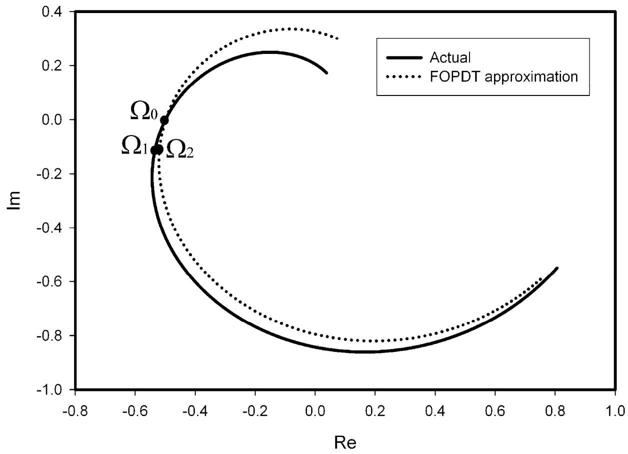


Fig. 1. Nyquist plots for process (1) and FOPDT approximation (3)

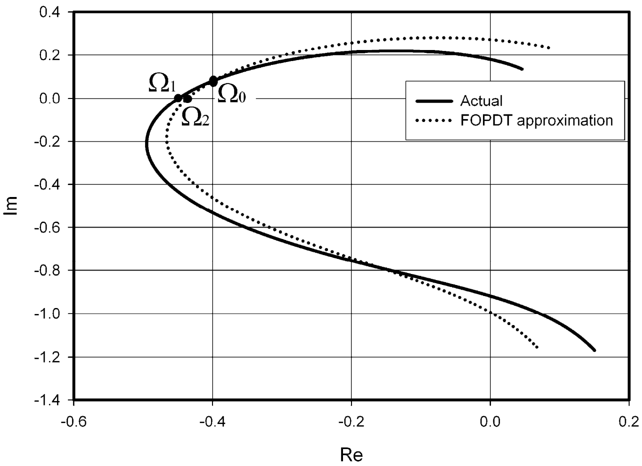


Fig. 2. Nyquist plots for open-loop system with PI controller and process

The Nyquist plots of the open-loop systems containing the process (1) or its approximation (3) and the controller (5) are depicted in Fig. 2. It follows from the Ziegler-Nichols tuning rules that the mapping of point Ω_0 in Fig. 1 into point Ω_0 in Fig. 2 is done via clockwise rotation of vector $\vec{W}_p(j\Omega_0)$ by the angle $\psi = \arctan\frac{1}{0.8 \cdot 2\pi} = 11.25^\circ$ and multiplication of its length by such value, so that its length becomes equal to 0.408. However, for the system containing the PI controller, the points of intersection of the Nyquist plots of the open-loop system and of the real axis are different for the system with process (1) and with process approximation (3). They are shown as points Ω_1 and Ω_2 in Fig. 2. The mapping of those points to the Nyquist plots of the process and its approximation is shown in Fig. 1. The stability margins of the systems containing a PI controller are not the same any more. It is revealed as different point of intersection of the plots and of the real axis in Fig. 2. In fact the position of vector

$\vec{W}_{ol}(j\Omega_0) = \vec{W}_c(j\Omega_0)\vec{W}_p(j\Omega_0)$ is fixed, but this vector does not reflect on the stability of the system, and as we can see in Fig. 2, the gain margin of the system containing the FOPDT approximation of the process is higher than the one of the system with the original process.

The considered example enlightens a fundamental problem of all methods of identification-tuning based on the measurements of process response in the critical point (Ω_0). This problem is the shift of the critical point due to the introduction of the controller. The question that naturally follows from the given analysis is whether the test point can be selected in a different way, so that the introduction of the controller would shift this point to the real axis. And if this is possible then what the test should be to ensure the measurements in the desired test point.

Address the first question first. Assume that we can design a certain test, so that we can assign the test point at the desired phase lag of the process $\arg W_p(j\Omega_0) = \varphi$, where φ is a given quantity, and measure $W_p(j\Omega_0)$ in this point. Consider the following example.

Example 2. Let the plant be the same as in Example 1. Assume that the introduction of the controller will be equivalent to the mapping similar to the mapping described above – the vector of the frequency response of the open-loop system in the point Ω_0 will be a result of clockwise rotation of the vector $\vec{W}_p(j\Omega_0)$ by a known angle and multiplication by a certain known factor: $\vec{W}_{ol}(j\Omega_0) = \vec{W}_c(j\Omega_0)\vec{W}_p(j\Omega_0)$. Also assume that the controller will be the same as in Example 1 (for illustrative purpose - because the tuning rules are not formulated yet). Therefore, let us find the values of T_p and τ for the transfer function (2) (we still assume K_p=1) that ensure that the equality $\hat{W}_p(j\Omega_0) = W_p(j\Omega_0)$ holds, where $\arg W_p(j\Omega_0) = -180^\circ + 11.25^\circ = -168.75^\circ$ (the angle is selected considering the subsequent clockwise rotation by 11.25°). Therefore, $\Omega_0 = 0.263$, and $W_p(j\Omega_0) = (-0.532, -j0.103)$. The matching FOPDT approximation of the process is

$$\hat{W}_p(s) = \frac{e^{-7.293s}}{5.897s + 1}, \qquad (6)$$

One can notice that both the time constant and the dead time in (6) are smaller than in (3). Application of controller (5) shifts the point Ω_0 of intersection of $W_p(j\Omega_0)$ and $\hat{W}_p(j\Omega_0)$ to the real axis. This point remains the point of intersection of the two Nyquist plots. Therefore, the gain margin of both systems: with the original process and with the approximated process are the same.

Consider now the problem of the design of the test that can ensure matching the actual and approximating processes in the point corresponding to a specified phase lag.

3. CLOSED-LOOP TEST VIA MODIFIED SUB-OPTIMAL ALGORITHM

Consider the following discontinuous control:

$$u(t) = \begin{cases} h & \text{if } \sigma(t) \ge \Delta_1 \text{ or } (\sigma(t) > -\Delta_2 \text{ and } u(t-) = h) \\ -h & \text{if } \sigma(t) \le \Delta_2 \text{ or } (\sigma(t) < \Delta_1 \text{ and } u(t-) = -h) \end{cases} \qquad (7)$$

where $\Delta_1 = \beta\sigma_{\max}$, $\Delta_2 = -\beta\sigma_{\min}$, $\sigma_{\max}$ and $\sigma_{\min}$ are last "singular" points of the error signal (Fig. 3) corresponding to last maximum and minimum values of $\sigma(t)$ after crossing the zero level, β is a positive constant parameter.

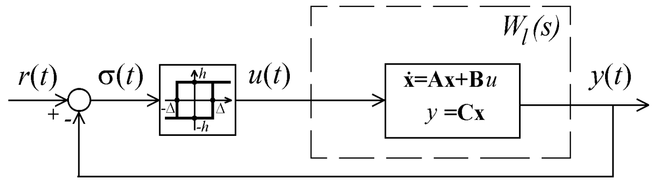


Fig. 3. Relay feedback test

The algorithm (7) is similar to the so-called "generalized sub-optimal" algorithm used for generating a second-order sliding mode in systems of relative degree two (Bartolini, 1997), (Bartolini, 2003). The difference between the two is that the generalized sub-optimal algorithm involves an advance switching but the proposed algorithm involves a delayed switching of the relay in comparison with the ideal relay. Let the reference signal $r(t)$ be zero in Fig. 3. Let us show that in the steady mode, the motions in the system Fig. 3, where the control is given by (7) are periodic.

Apply the describing function (DF) method (Atherton, 1975) to the analysis of motions in Fig. 3. Assume that the steady mode is periodic, and after that prove that this is a valid assumption by finding the parameters of this periodic motion. If the motions in the system are periodic then $\sigma_{\max}$ and $\sigma_{\min}$ represent the amplitude of the oscillations: $a = \sigma_{\max} = -\sigma_{\min}$, which can be represented by the equivalent hysteresis value of the relay $\Delta = \Delta_1 = \Delta_2 = \beta\sigma_{\max} = -\beta\sigma_{\min}$. The DF of the hysteretic relay is given as follows

$$N(a) = \frac{4h}{\pi a}\sqrt{1 - \left(\frac{\Delta}{a}\right)^2} - j\frac{4h\Delta}{\pi a^2}, \; a > \Delta \qquad (8)$$

However, system Fig.3 with control (7) is not a conventional relay system. This system has the hysteresis value that is unknown *a-priori* and depends on the amplitude value: $\Delta = \beta a$. Therefore, (8) can be rewritten as follows:

$$N(a) = \frac{4h}{\pi a}\left(\sqrt{1 - \beta^2} - j\beta\right), \qquad (9)$$

The relay feedback test will generate oscillations in the system under control (7). We shall refer to that test as to "modified relay feedback test". Parameters of the oscillations can be found from the harmonic balance equation:

$$W_p(j\Omega_0) = -\frac{1}{N(a_0)}, \tag{10}$$

where a_0 is the amplitude of the periodic motions, and the negative reciprocal of the DF is given as follows:

$$-\frac{1}{N(a)} = -\frac{\pi a}{4h}\left(\sqrt{1-\beta^2} + j\beta\right) \tag{11}$$

Finding a periodic solution in system Fig.3 with control (7) has a simple graphic interpretation (Fig. 4) as finding the point of intersection of the Nyquist plot of the process and of the negative reciprocal of the DF, which is a straight line that begins in the origin and makes a counterclockwise angle $\psi = \arcsin\beta$ with the negative part of the real axis.

In the problem of analysis, frequency Ω_0 and amplitude a_0 are unknown variables and are found from the complex equation (10). In the problems of identification and tuning, frequency Ω_0 and amplitude a_0 are measured from the modified relay feedback test, and on the basis of the measurements obtained either parameters of the underlying model are calculated (for parametric tuning) or tuning parameters are calculated immediately from Ω_0 and a_0 (for non-parametric tuning).

Reviewing again Example 2, we can note that if, for example, Ziegler-Nichols tuning rules are supposed to be applied, and the subsequent transformation via introduction of the PI controller involving clockwise rotation by angle $\psi = \arctan\frac{1}{0.8\cdot 2\pi} = 11.25^\circ$ is going to be applied, then parameter β of the controller for the modified relay feedback test should be $\beta = \sin 11.25^\circ = 0.195$.

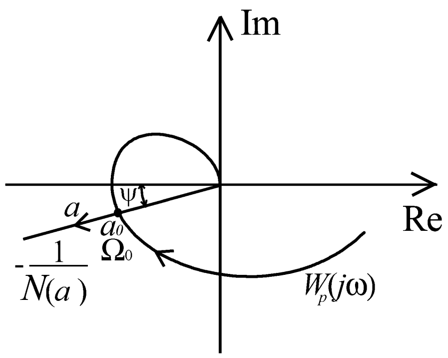


Fig. 4. Finding a periodic solution

The modified relay feedback test also allows for the exact design of the gain margin. Since the amplitude of the oscillations a_0 is measured from the test, the process gain at frequency Ω_0 can be obtained as follows: $|W_p(j\Omega_0)| = \frac{\pi a_0}{4h}$, which after introduction of the controller will become the process gain at the critical frequency. Therefore, if the tuning rules are given in the format:

$$K_c = c_1\frac{4h}{\pi a_0}, \quad T_c = c_2\frac{2\pi}{\Omega_0}, \tag{12}$$

where c_1 and c_2 are parameters that define the tuning rule, then the frequency response of the controller at Ω_0 becomes

$$W_c(j\Omega_0) = c_1\frac{4h}{\pi a_0}\left(1 - j\frac{1}{2\pi c_2}\right), \tag{13}$$

and for obtaining the gain margin γ (γ >1) parameter c_1 should be selected as $c_1 = 1\Big/\left(\gamma\sqrt{1+\frac{1}{4\pi^2 c_2^{\,2}}}\right)$. In the considered example, if we keep parameter c_2 the same as in (Ziegler & Nichols, 1942): c_2=0.8, then to obtain, for example, gain margin γ=2 tuning parameter c_1 for the modified relay feedback test should be selected c_1=0.49. Any process regardless of the actual dynamics will have gain margin γ=2 (6dB) exactly (within the framework of the filtering hypothesis of the DF method).

4. NON-PARAMETRIC TUNING RULES

Given a large variety of possible process dynamics, it is very difficult to formulate certain universal rules for tuning. In practice of process control, tuning rules that provide a less aggressive response than the one provided by IAE, ITAE criteria or Ziegler-Nichols formulas (or other rules) are widely used. This trend is reflected in the review of the modern PID control (Astrom & Hagglund, 2006). Let us consider the PI controller only, and only the rules given in the format of the proportional dependence of the controller gain on the ultimate gain and of the integral time constant on the period of the oscillations in the modified relay feedback test – as given by formula (12). Considering the fact that the frequency-domain characteristics of all loops tuned via the modified relay feedback test will be very consistent (the gain margin is the same), let us analyze the time-domain characteristics of the loops with different process dynamics and generate the tuning rules that provide the best consistency of the time-domain characteristics.

Let us use the FOPDT model as the implied process dynamics for the purpose of optimal selection of the coefficients c_1 and c_2. Analysis of the time-domain performance of FOPDT processes with different ratios between the dead time and the time constant (subject to the same value of the gain margins) would allow us to find the optimal tuning rules. Within the time domain, it would be difficult to compare such characteristics as settling time or other measures of speed of response – due to the difference in time constants of different processes. The only parameter that can be used as a "universal" characteristic (in that sense) is the value of the overshoot in the step response. Therefore, let us find the overshoot values of the step responses of a series of FOPDT dynamics with dead time to time constant ratio ranging as follows $\tau / T_p = [0.3;1.5]$, subject to equal gain margins in those loops, by varying gain margin and parameter c_2 values. The noted dependence is presented in Fig. 5, where gain margin $\gamma \in [2;4]$ and parameter $c_2 \in [0.3;3.3]$.

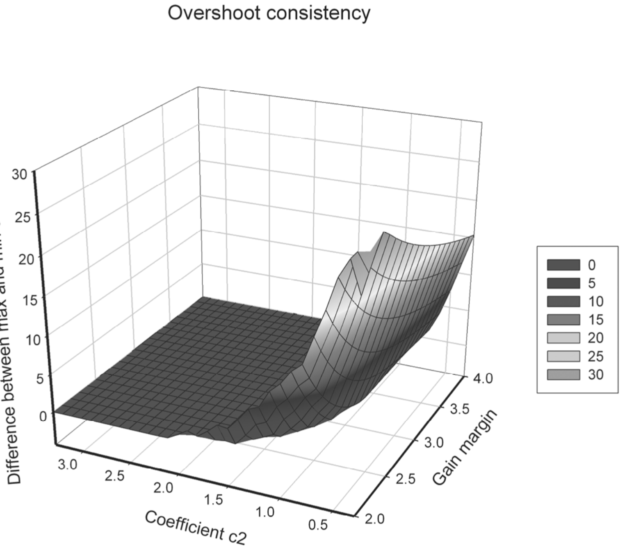


Fig. 5. Difference between maximum and minimum overshoot [%] for proposed closed-loop test

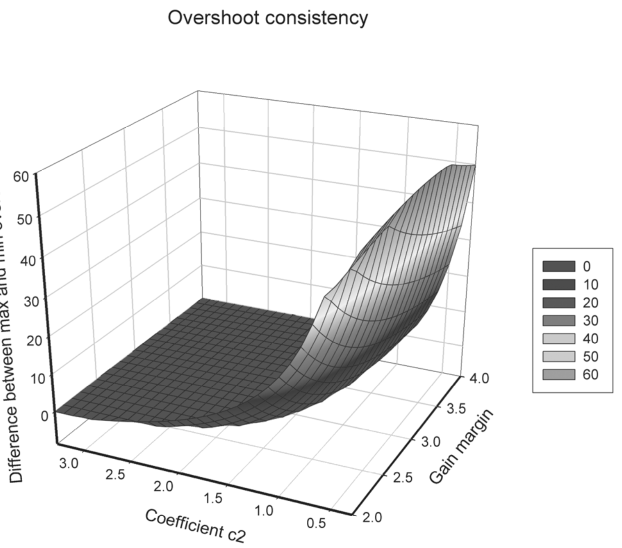


Fig. 6. Difference between maximum and minimum overshoot [%] for conventional relay feedback test (note different from Fig. 5 scale)

One can see that at lower values of parameter c_2 the difference between the maximum and minimum overshoots can be high, which does not allow using too low values. On the other hand, higher values of parameter c_2 result in the decrease of the integral action of the controller, which may not be acceptable. With respect to the dependence on the gain margin, higher values of the gain margin lead to a smaller overshoot, and to a higher consistency of the step response for various τ / T_p ratios.

A similar plot for the case of the conventional relay feedback test is presented in Fig. 6 for comparison. In that case the gain margins are not equalized by respective selection of parameter β, and the difference between the maximum and the minimum overshoots is about three times of the former. Analysis of the data presented in Fig. 5 shows that for satisfactory consistency of the step response (difference between maximum and minimum overshoots is lower than 10%) the gain margin and the value of c_2 should not be smaller than certain values. In particular, for γ=2 $c_2 \geq 1.1$; γ=2.5 $c_2 \geq 0.7$; γ=3 $c_2 \geq 0.6$; γ=3.5 $c_2 \geq 0.5$, and γ=4 $c_2 \geq 0.5$. Therefore, the recommended settings for non-aggressive tuning with expected overshoot 0-3.3% might be γ=3 c_2=0.7, which results in the following tuning rules:

$$K_c = 0.33 \frac{4h}{\pi a}, \quad T_c = 0.7 \frac{2\pi}{\Omega_0}, \tag{14}$$

As follows from formula (13) the PI controller would introduce the lag at the frequency Ω_0 equal to $\arctan \frac{1}{2\pi c_2} \approx 12.81^\circ$. Therefore, the parameter β of the modified relay feedback test should be $\beta = \sin\left(\arctan \frac{1}{2\pi c_2}\right) = 0.222$.

5. EXAMPLES

Example 3. Consider the following four transfer functions (Fig. 7) that were used as representative process models in (Kaya & Atherton, 2001). Apply the modified relay feedback test in the first case with parameter $\beta = 0.222$ and tuning rules (14), which correspond to γ=3, and in the second case with $K_c = 0.49 \frac{4h}{\pi a}$ and the same rule for T_c, which corresponds to γ=2, to those processes. The step responses of the tuned loops to the set point change (denoted as "s.p.") and to the disturbance application (denoted as "disturbance") are presented in Fig. 7. The graphs presented demonstrate a satisfactory loop performance in a conservative approach (γ=3) and in a more aggressive loop tuning (γ=2). The performance of the loops is in agreement with the design criteria selected: the desired gain margin.

6. CONCLUSIONS

A method of non-parametric tuning of a PI controller based on the modified relay feedback test and inspired by one of the second-order sliding mode algorithms is presented in the paper. The proposed method involves two parameter measurements from the modified relay feedback test, and the calculations similar to the ones as per Ziegler-Nichols formulas. It is proved that the proposed method ensures the desired value of the gain margin exactly (subject to the assumptions of the describing function method). It is shown that the proposed approach ensures an equalizing effect with respect to the time-domain characteristics of a variety of possible process dynamics – due to providing the same gain margin for all possible processes. Despite the consideration of only PI controller, the proposed method can easily be extended to the case of the PID control, and respective tuning rules can be obtained. The proposed method was implemented on a Honeywell TPS DCS as a module programmed in CL and successfully used for loop tuning.

REFERENCES

Astrom, K. J. and T. Hagglund (1984). Automatic tuning of simple regulators with specifications on phase and amplitude margins. *Automatica*, **20**, 645-651.

Astrom, K.J. and T. Hagglund (1995). *PID Controllers: Theory, Design and Tuning*, second ed. ISA, Research Triangle Park, USA.

Astrom, K.J. and T. Hagglund (2006). *Advanced PID Control*, ISA, Research Triangle Park, USA.

Atherton, D.P. (1975). *Nonlinear Control Engineering – Describing Function Analysis and Design*. Van Nostrand Company Limited, Workingham, Berks, UK.

Bartolini, G., A. Ferrara, and E. Usai (1997). Output tracking control of uncertain nonlinear second-order systems, *Automatica*, **33** (12), 2203-2212.

Bartolini, G., A. Pisano, E. Punta, and E. Usai (2003). A survey of applications of second-order sliding mode control to mechanical systems. *International Journal of Control*, **76** (9/10), 875-892.

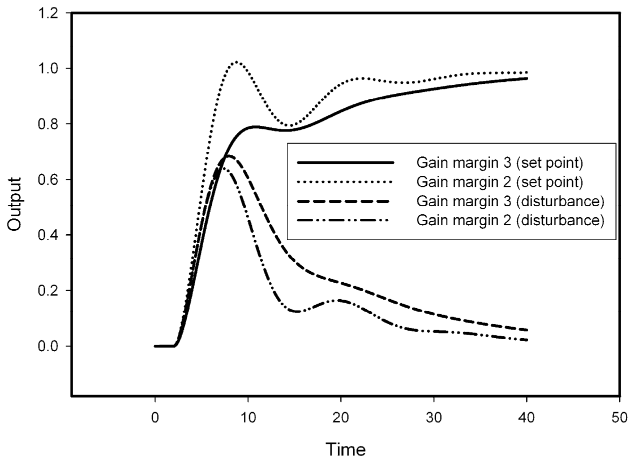
Boiko, I. (2005). Oscillations and transfer properties of relay servo systems – the locus of a perturbed relay system approach. Automatica, **41**, 677-683.

Boiko, I. (2006). Method and apparatus for tuning a PID controller, *US Patent No. 7,035,695*.

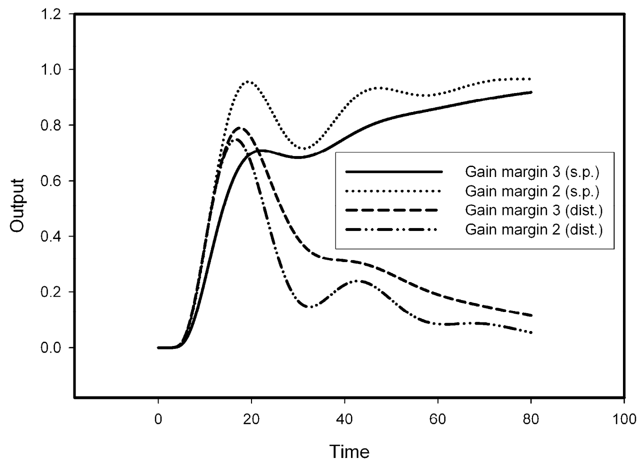
Kaya, I. and D.P. Atherton (2001). Parameter estimation from relay autotuning with asymmetric limit cycle data. *J. Process Control*, No. 11, 429-439.

Tsypkin, Ya. Z. (1984). *Relay Control Systems*, Cambridge, England.

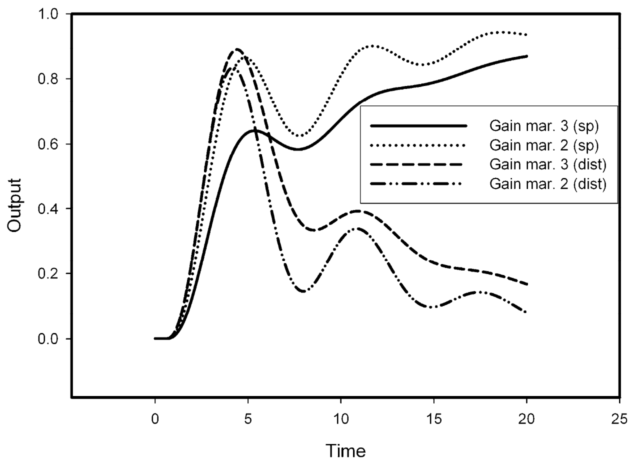
Ziegler, J.G. and N.B. Nichols (1942). Optimum settings for automatic controllers, *Trans. Amer. Soc. Mech. Eng.*, **64**, 759-768.



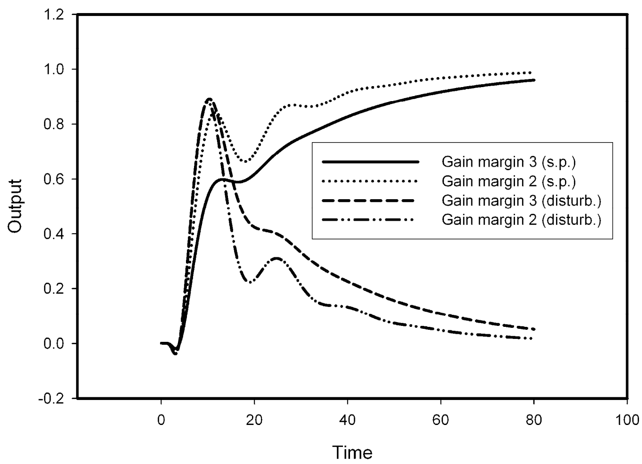
(a) $W(s) = \mathrm{e}^{-2s} \dfrac{1}{(2s+1)^2}$



(b) $W(s) = \mathrm{e}^{-2s} \dfrac{1}{(2s+1)^5}$



(c) $W(s) = \mathrm{e}^{-0.5s} \dfrac{1}{(s+1)(s^2+s+1)}$



(d) $W(s) = \mathrm{e}^{-s} \dfrac{-s+1}{(s+1)^5}$

Fig. 7. Step response of the processes of Example 3.