Proceedings of the 1997 IEEE International
Conference on Control Applications
Hartford, CT • October 5-7, 1997

MP05 4:40

A Numerical Method for Design of PI Controllers

H. Panagopoulos, K. J. Åström, and T. Hägglund

Department of Automatic Control
Lund Institute of Technology
Box 118
S-221 00 Lund, Sweden
Phone: +46-46 2228789, Fax: +46-46 138118
Email: hp@control.lth.se

Abstract This paper presents an efficient numerical method to design PI controllers. The design is based on optimization of the load disturbance rejection, with constraints on the sensitivity and weighting of the set point response. Thus, the formulation of the design problem captures many aspects of industrial control problems. It leads to a non-convex optimization problem. Efficient ways to solve the problem are presented.

Keywords PI control. Design. Optimization. Specifications. Load disturbance rejection. Set point response. Robustness. Sensitivity.

1. Introduction

The PI controller is unquestionably the most commonly used control algorithm. In spite of its wide spread use there exist no generally accepted design method for the controller. The empirical method described in Ziegler and Nichols (1942) is probably the most commonly used technique even if it inherently gives closed loop systems with poor robustness to process variations, see Åström and Hägglund (1995). Most traditional techniques are based on little information about the process and simple computational procedures.

With the computational power that is available today it is possible to use systematic design methods. In order to do so we start by giving specifications and then proceed by finding controllers which satisfy these specifications if this is at all possible. Such a systematic approach has many advantages: it is invaluable when comparing different controllers and it is essential in devices for automatic tuning.

There are many requirements on an efficient design method. It should be applicable to a wide range of systems and it should have the possibility to introduce specifications that capture the essence of real control problems. Furthermore, the method should be robust in the sense that it provides controller parameters if they exist, or if the specifications can not be met an appropriate diagnosis should be presented. We believe these requirements are satisfied by the method presented in this paper.

The method presented assumes a linear process whose dynamics is characterized in terms of a transfer function, which does not have to be rational. This means that the method can be applied to systems described by partial differential equations. Consequently, for a system with a given transfer function the method will give a PI controller that satisfies the specifications if such a controller exists.

This work builds on previous work based on the idea to optimize load disturbance response with constraints on sensitivity, see Shinskey (1990), Persson (1992), Persson and Åström (1992), and Schei (1994). The main contributions of this paper are the efficient numerical scheme and a systematic method for finding the set point weighting.

2. Formulation of the Design Problem

The formulation of a design problem includes a characterization of the process and its environment, the controller structure, and specifications on the performance of the closed-loop system.

The Process is assumed to be linear, and specified by a transfer function $G(s)$ which is analytical with finite poles and possibly an essential singularity at infinity. The description covers finite dimensional systems with time delays and infinite dimensional systems described by linear partial differential equations. This is very useful because it admits design of controllers for a very wide range of processes.

The Controller is assumed to be a PI controller described by

$$u(t) = k\big(by_{sp}(t) - y(t)\big) + k_i \int_0^t \big(y_{sp}(\tau) - y(\tau)\big)d\tau \quad (1)$$

where $u(t)$ is the control signal, $y(t)$ is the process output, y_{sp} is the set point, and k, k_i, and b are controller parameters. It is industry practice to use integration time, defined as $T_i = k/k_i$, instead of parameter k_i. However, for the computations it is more convenient to use k_i. Industrial controllers typically use either $b = 0$ or $b = 1$, but lately it has been recognized that it is advantageous to use full range of b-values, that is $0 \le b \le 1$. The controller given by Equation (1) is said to have two degrees of freedom when $b \neq 1$. The advantage of such structures has been pointed out by Horowitz (1963) and their use in PID controllers is discussed in Shigemasa *et al.* (1987) and Åström and Hägglund (1995).

Specifications express requirements on:

- Load disturbance response
- Set point response
- Robustness with respect to model uncertainties

In process control applications efficient rejection of load disturbances is of primary concern whereas set point responses are typically of secondary importance. However, set point response may be of primary importance for example in motion control systems. Although it has been frequently pointed out by engineers that load disturbances is of primary concern, it is interesting to note that papers on PI control traditionally focus on set point response, see e.g. Shinskey (1990).

The sensitivity to model uncertainty is of primary significance. The poor sensitivity is one of the major drawbacks of the classical Ziegler-Nichols tuning scheme.

In practically all applications it is useful to have a tuning parameter that permits adjustment of the trade off of aggressiveness versus robustness. The effects of the tuning parameter should be transparent to the user. In the method presented we have a tuning parameter to specify the sensitivity to model uncertainties.

A Formal Description

In order to use a formal design method it is necessary to capture specifications in mathematical form which is extensively discussed in Åström and Hägglund (1995). Load disturbance rejection can be expressed very neatly in terms of the integrated error due to a load disturbance response at the process input. In Åström and Hägglund (1995) it is shown that the integrated error is given by

$$IE = \int_0^\infty e(t)dt = \frac{1}{k_i}, \quad (2)$$

where the input is a unit step at the process input. However, integrated error by itself is not a satisfactory criterion of controller performance, as it would not penalize a loop oscillating uniformly. In our case this is ensured by the sensitivity constraints, which can be expressed in two ways.

Sensitivity to modeling errors can be expressed in terms of the largest value of the sensitivity function. Let the loop transfer function be $L(s) = G(s)G_c(s)$ where G_c is the controller transfer function, and let the sensitivity function be $S(s) = 1/(1 + L(s))$. The maximum sensitivity is then characterized by $M_s = \max|S(i\omega)|$. Keep in mind that the quantity M_s is simply the inverse of the shortest distance from the Nyquist curve of the loop transfer function to the critical point -1. Typical values of M_s are in the range of 1 to 2.

Let $T(s) = 1 - S(s) = L(s)/(1 + L(s))$ be the complementary sensitivity function. The sensitivity can also be expressed by the largest value of the complementary sensitivity function, i.e. $M_p = \max|T(i\omega)|$. The value M_p is the size of the resonance peak of the closed loop system obtained with $b = 1$, see Equation (1). Typical values of M_p are in the range of 1.0 to 1.5.

The design has so far focused on the response to load disturbances, which is of primary concern. However, it is also important to have a good response to set point changes. The transfer function from set point to process output is characterized by,

$$\frac{Y(s)}{Y_{sp}(s)} = \frac{k_i + bks}{k_i + ks}\frac{L(s)}{1 + L(s)} = G_{sp}(s). \quad (3)$$

One way to give specifications on the set point response is to specify the resonance peak of the transfer function $G_{sp}(s)$, i.e.

$$M_{sp} = \max|G_{sp}(i\omega)|. \quad (4)$$

Consequently, the b-value is determined such it fulfills Equation (4). Notice that $M_{sp} < M_p$ if $b < 1$.

Design Parameters

The tradeoff between performance and robustness varies between different control problems. Therefore, it is desirable to have a design parameter to change the properties of the closed-loop system. Ideally, the parameter should be directly related to the performance of the system, it should not be process oriented. There should be good default values so a user is not forced to select some value. This is of special importance when the design procedure is used for automatic tuning. The design parameter should also have a good physical interpretation and natural limits to simplify its adjustment.

The variables M_s and M_p are both possible candidates for design variables. In Åström *et al.* (1997) it is shown that M_s is more suitable as a design parameter. On the other hand, it is important that the resulting M_p value is not too large. We will therefore also calculate M_p when the design is completed. If M_p is too large there are several possibilities. One is to repeat the design with a smaller M_s value. Another is to use constraints on both M_s and M_p, see Åström *et al.* (1997).

3. The Optimization Problem

The design problem discussed in the previous section can be formulated as an optimization problem: Find controller parameters that maximize k_i subject to the constraints that the closed loop system is stable and that the Nyquist curve of the loop transfer function is outside a circle with center at $s = -C$ and radius R.

Introduce $L(s) = (k + k_i/s)G(s)$ and the function

$$f(k, k_i, \omega) = |C + (k - i\frac{k_i}{\omega})G(i\omega)|^2. \tag{5}$$

The sensitivity constraint can then be expressed as

$$f(k, k_i, \omega) \geq R^2 \tag{6}$$

and the optimization problem is to maximize k_i subject to the sensitivity constraint (6).

Let $a(\omega)$ and $b(\omega)$ be the real and imaginary parts of the process transfer function. Hence

$$G(i\omega) = a(\omega) + ib(\omega) = r(\omega)e^{i\varphi(\omega)}$$

where

$$\begin{aligned} a(\omega) &= r(\omega)\cos\varphi(\omega),\\ b(\omega) &= r(\omega)\sin\varphi(\omega). \end{aligned}$$

The function f can then be written as

$$f(k, k_i, \omega) = C^2 + 2Ca(\omega)k + 2C\frac{b(\omega)}{\omega}k_i + r^2(\omega)k^2 + \frac{r^2(\omega)}{\omega^2}k_i^2. \tag{7}$$

In the following, we will occasionally drop the argument ω in a, b, r, and φ in order to simplify the writing.

The optimization problem is nontrivial because the constraint, which is infinite dimensional, defines a set in parameter space which is not convex. There are also other subtleties which may cause problems. For a specific problem it is not difficult to solve the problem numerically with standard optimization routines because the search range can often be limited and if the optimization fails it is possible to interfere manually. Since PI controllers are very common it is, however, worthwhile to make special algorithms which are tailored for the problem. Such procedures are also required for automatic tuning where manual interaction is very inconvenient.

The constraint given by Equation (6) has a nice geometric interpretation. For fixed ω Equation (6) represents the exterior of an ellipse in the k-k_i plane. The ellipse has its axes parallel to the coordinate axes. For $0 \leq \omega < \infty$ the ellipses generate the envelope defined by

$$\begin{aligned} f(k, k_i, \omega) &= R^2,\\ \frac{\partial f}{\partial \omega}(k, k_i, \omega) &= 0. \end{aligned} \tag{8}$$

Since the function f is quadratic in k_i the envelope has two branches. Only one branch corresponds to stable closed loop systems.

The admissible gains that give a maximum sensitivity less than M_s and a stable closed loop system are characterized by a set in the k-k_i plane. This set which is bounded by one of the envelopes and the k-axis is not convex.

It is a major computational effort to compute the envelope and we will therefore look for methods that require less computations. We will achieve this by reducing the problem to solving nonlinear equations.

The envelope given by Equation (8) defines implicitly k_i as a function of k. To find the maximum of this function we observe that

$$df = \frac{\partial f}{\partial k}dk + \frac{\partial f}{\partial k_i}dk_i + \frac{\partial f}{\partial \omega}d\omega = 0. \tag{9}$$

It follows from Equation (8) that $\partial f/\partial\omega = 0$ on the envelope. At a local extrema we have $dk_i = 0$. For arbitrary variations of dk we must thus require that $\partial f/\partial k = 0$. Combining this with the envelope conditions (8) we get

$$\begin{aligned} \frac{\partial f}{\partial k}(k, k_i, \omega) &= 0,\\ \frac{\partial f}{\partial \omega}(k, k_i, \omega) &= 0,\\ f(k, k_i, \omega) &= R^2, \end{aligned} \tag{10}$$

which corresponds to the situation when the maximum occurs at a point on the envelope where it has continuous derivatives. In the Nyquist diagram this corresponds to the case when the loop transfer function has a tangency with the circle at one point.

The extremum may, however, also occur when the envelope has a discontinuous derivative. In the following the details will be worked out for the case when the maximum occurs at a point on the envelope where it has continuous derivatives. See Åström *et al.* (1997) for the details when the envelope has a discontinuous derivatives.

A Simplification

Equation (10) is a nonlinear equation in three variables, k, k_i and ω. It is possible to solve this equation directly, but a much more efficient algorithm will be obtained by eliminating some variables.

Inserting expression (7) into Equation (10), gives

$$\begin{aligned}\frac{\partial f}{\partial k} &= 2Ca + 2r^2k = 0,\\ \frac{\partial f}{\partial \omega} &= 2C(\frac{b}{\omega})'k_i + 2Ca'k + \left(\frac{r^2}{\omega^2}\right)' k_i^2 + 2rr'k^2 = 0,\\ f &= R^2.\end{aligned} \tag{11}$$

Solving k and k_i from the first and last equations gives

$$\begin{aligned}k &= -C\frac{a}{r^2} = -C\frac{1}{r}\cos\varphi,\\ k_i &= -\frac{\omega bC}{r^2} - \frac{\omega R}{r} = -\frac{\omega}{r}(C\sin\varphi + R),\end{aligned} \tag{12}$$

where the positive sign is chosen to satisfy the encirclement criterion. The following condition is obtained by inserting the expressions of k and k_i in the second equation of (11),

$$\begin{aligned}h(\omega) &= 2R\left((C\frac{b}{r} + R)(\frac{r'}{r} - \frac{1}{\omega}) - C(\frac{b}{r})'\right)\\ &= 2R\left((R + C\sin\varphi)(\frac{r'}{r} - \frac{1}{\omega}) - C\varphi'\cos\varphi\right) = 0.\end{aligned} \tag{13}$$

Thus, the solution to Equation (11) is reduced to a single algebraic equation (13) in ω. Solving it gives the frequency ω_0 for which we can compute the controller gains k and k_i given by Equation (12).

The condition (11) does not tell if the extremum is a minimum, a maximum or saddle point, but constraint (6) implies that the function should be a minimum with respect to ω. A necessary condition for this is that the curvature of the Nyquist curve of the loop transfer function is greater than R at ω_0. This gives the following local condition

$$\frac{d^2 f}{d\omega^2}(\omega_0) > 0. \tag{14}$$

Equation (13) can be solved iteratively with the Newton-Raphson method which converges very fast if suitable initial conditions are given. Notice, however, that in general there may be several solutions which can be found by starting the iteration from different initial conditions.

For special classes of systems, for example systems with monotone transfer function, it is possible to provide good initial conditions, see Åström *et al.* (1997).

Set Point Weighting

The set point response is governed by the transfer function G_{sp} given by Equation (3). In order to have a small overshoot in set point response, set point weighting b will be determined so that $M_{sp} = \max|G_{sp}(i\omega)|$ is close to one. It follows from Equations (3) that $M_{sp} \le M_p$ when $0 \le b \le 1$. A bound of M_{sp} is thus given indirectly through M_p.

We will make the approximation that maximum of $|G_{sp}(i\omega)|$ occurs for ω_{mp}, where $\omega = \omega_{mp}$ is the frequency where the maximum of $|L(i\omega)/(1 + L(i\omega))|$ occurs. Parameter b will be determined so that

$$|G_{sp}(i\omega_{mp})| = 1 \tag{15}$$

with the constraint $0 \le b \le 1$. Using Equation (3) and Equation (15) this implies that

$$b = \begin{cases}\dfrac{\sqrt{k^2\omega_{mp}^2 - k_i^2(M_p^2 - 1)}}{k\omega_{mp}M_p} & \text{if } (\omega_{mp}k/k_i)^2 \ge M_p^2 - 1,\\ 0 & \text{if } (\omega_{mp}k/k_i)^2 < M_p^2 - 1.\end{cases} \tag{16}$$

If $b = 0$, it is not sure that the design objective (15) will be obtained. If the set point response is important and the M_p value is large, the design can be repeated, see Åström *et al.* (1997).

4. Examples

We will consider some representative systems which are normally encountered in process control. They have the following transfer functions,

$$\begin{aligned}G_1(s) &= \frac{1}{(s+1)(1+0.2s)(1+0.04s)(1+0.008s)},\\ G_2(s) &= \frac{1}{(s+1)^3}, & G_3(s) &= \frac{e^{-5s}}{(s+1)^3},\\ G_4(s) &= \frac{1}{s(s+1)^2}, & G_5(s) &= \frac{1-2s}{(s+1)^3},\\ G_6(s) &= \frac{9}{(s+1)(s^2+2s+9)}.\end{aligned}$$

The first four models capture typical dynamics encountered in the process industry. Models G_1 and G_2 represent processes that are relatively easy to control. Model G_3 has a long dead time, and G_4 models an integrating process. Model G_5 has a zero in the right half plane, and model G_6 has complex poles with relative damping 0.33. These processes are uncommon in process control, but they have been included to demonstrate the wide applicability of the design procedure.

Figure 1 shows the Nyquist curves of the loop transfer functions obtained for two values of the design parameter M_s. The responses to changes in set point and load

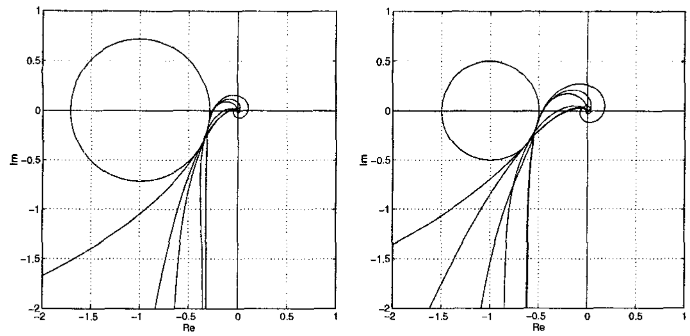


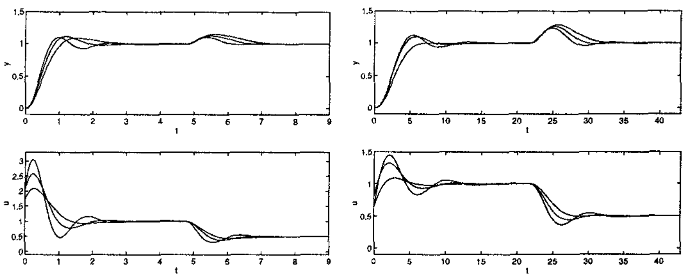
Figure 1 The Nyquist plots of the open loop frequency response for the process models $G_j(s)$, $j = 1 \dots 6$ for $M_s = 1.4$ to the left and $M_s = 2.0$ to the right.

are shown in Figure 2, and the details of the design calculations and simulations are summarized in Table 1. Although the models $G_1 - G_6$ represent processes with large variations in process dynamics, Figure 2 shows that the resulting closed loop responses are quite similar. This is important because it means that the proposed design procedure gives closed loop systems with desired and predictable properties. The fact that even integrating processes can be treated in the same way as stable processes is interesting. In many other design approaches, stable and integrating processes have to be treated separately, see Åström and Hägglund (1995).

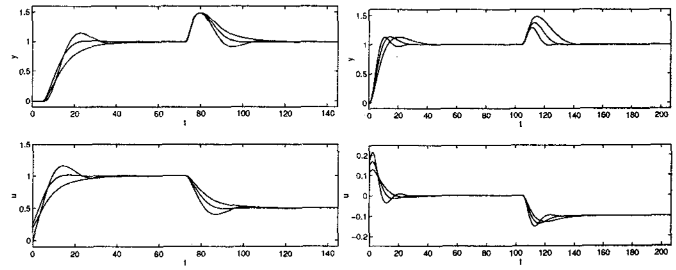
There is also a large similarity between the responses obtained with the different values of the tuning parameter M_s. This shows that the M_s-value is a suitable tuning parameter. Responses obtained with M_s=1.4 show little or no overshoot. This is normally desirable in process control. Responses obtained with M_s=2.0 gives faster responses. The settling time at load disturbances, t_s, is significantly shorter with this larger value of M_s. On the other hand, these responses are oscillatory with a larger overshoot. This can be seen from the comparison between IE and the integrated absolute error IAE in Table 1.

The controller gain K_c varies significantly with the design parameter M_s. However, integral time T_i is fairly constant for the stable processes, i.e., all processes except G_4. This means, that for PI control the different design specifications are mainly obtained by adjusting the gain only. This observation is made earlier, see Åström and Hägglund (1995).

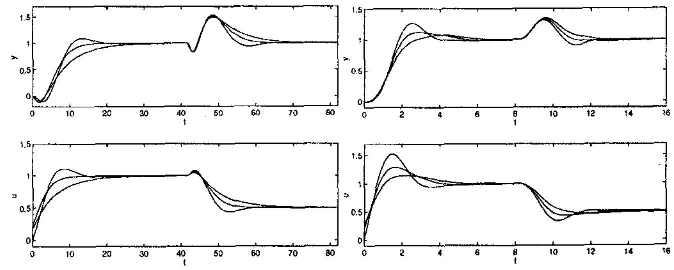
Except for the integrating process G_4, the M_p values obtained for M_s=1.4 are all close to one. Consequently, parameter b is also close to one. For M_s=2.0, the M_p values are, however, larger. This means that the overshoots would have been significant if the set point weighting were chosen to $b = 1$. However, acceptable set point responses are obtained by using small values of b. In some cases, $b = 0$. It means that the procedure has failed to obtain $M_{sp} = 1$. If set point responses are important and if the overshoots are unacceptable, a redesign may be done using smaller values of M_s or optimization with constraint on M_p, see Åström *et al.* (1997).



Control of the process model G_1 (left) and G_2 (right)



Control of the process model G_3 (left) and G_4 (right)



Control of the process model G_5 (left) and G_6 (right)

Figure 2 Comparison between the PI controllers for $M_s = 1.4$, 1.6 and 2.0. The graphs show a step response followed by a load disturbance. In all the diagrams, the fastest responses are obtained by the controllers designed for $M_s = 2.0$.

Process	M_s	K_c	T_i	b	IE	IE/IAE	w_0	t_s	M_p
$G_1(s)$	1.4	1.93	0.74	0.89	0.19	1.00	3.33	2.25	1.10
	1.6	2.74	0.67	0.75	0.12	1.00	3.83	1.68	1.27
	1.8	3.47	0.62	0.63	0.09	0.96	4.25	1.39	1.46
	2.0	4.13	0.59	0.52	0.07	0.91	4.40	1.21	1.66
$G_2(s)$	1.4	0.63	1.95	1.00	1.54	1.00	0.74	10.3	1.00
	1.6	0.86	1.87	0.93	1.08	0.95	0.79	7.87	1.05
	1.8	1.06	1.82	0.70	0.86	0.86	0.82	6.77	1.24
	2.0	1.22	1.78	0.50	0.73	0.77	0.85	6.27	1.45
$G_3(s)$	1.4	0.19	2.99	1.00	7.84	1.00	0.22	38.2	1.00
	1.6	0.24	2.85	1.00	5.87	0.98	0.23	23.0	1.00
	1.8	0.28	2.75	0.88	4.89	0.84	0.23	19.1	1.03
	2.0	0.31	2.68	0.00	4.30	0.73	0.24	17.9	1.19
$G_4(s)$	1.4	0.17	14.00	0.70	8.40	0.97	0.29	35.3	1.40
	1.6	0.23	10.67	0.64	4.62	0.99	0.34	24.1	1.49
	1.8	0.29	9.00	0.57	3.15	0.99	0.38	18.5	1.62
	2.0	0.33	8.00	0.50	2.40	0.96	0.41	15.7	1.77
$G_5(s)$	1.4	0.18	1.78	1.00	4.94	0.90	0.38	28.6	1.00
	1.6	0.23	1.69	1.00	3.71	0.87	0.40	18.6	1.00
	1.8	0.27	1.64	0.87	3.09	0.80	0.41	14.7	1.04
	2.0	0.29	1.60	0.00	2.71	0.70	0.41	13.5	1.20
$G_6(s)$	1.4	0.31	0.37	0.88	0.59	0.87	1.98	4.13	1.04
	1.6	0.39	0.34	0.51	0.44	0.79	2.05	2.94	1.15
	1.8	0.44	0.33	0.00	0.37	0.70	2.05	2.69	1.26
	2.0	0.48	0.31	0.00	0.33	0.64	2.12	2.56	1.37

Table 1 Properties of controllers obtained for different systems and different values of the design parameter M_s.

5. Conclusions

This paper describes a design method for PI controllers that results in a constrained optimization problem. Robust and efficient numerical procedures to solve the problem are also presented.

The primary design goal is to obtain good load disturbance responses. This is done by minimizing the integrated control error IE. Robustness is guaranteed by requiring that the maximum sensitivity is less than a specified value M_s. The M_s value is shown to be a suitable tuning variable. Finally, set point responses are treated using set-point weighting. The set-point weight b is determined from the M_p value.

The main contributions of the paper are the efficient numerical procedure and the systematic method to find the set point weighting b. The design procedure has been applied to a variety of process models; stable and integrating, with short and long dead times, with real and complex poles, and with positive and negative zeros. The method provides suitable PI controller parameters that results in load- and set-point responses that correspond to the specifications.

6. References

ÅSTRÖM, K. J. and T. HÄGGLUND (1995): *PID Controllers: Theory, Design, and Tuning*, second edition. Instrument Society of America, Research Triangle Park, NC.

ÅSTRÖM, K. J., H. PANAGOPOULOS, and T. HÄGGLUND (1997): "Design of PI controllers." Submitted to Automatica.

HOROWITZ, I. M. (1963): *Synthesis of Feedback Systems.* Academic Press, New York.

PERSSON, P. (1992): *Towards Autonomous PID Control.* PhD thesis ISRN LUTFD2/TFRT--1037--SE, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

PERSSON, P. and K. J. ÅSTRÖM (1992): "Dominant pole design—A unified view of PID controller tuning." In *Preprints 4th IFAC Symposium on Adaptive Systems in Control and Signal Processing*, pp. 127–132. Grenoble, France.

SCHEI, T. S. (1994): "Automatic tuning of PID controllers based on transfer function estimation." *Automatica*, **30:12**, pp. 1983–1989.

SHIGEMASA, T., Y. IINO, and M. KANDA (1987): "Two degrees of freedom PID auto-tuning controller." In *Proceedings of ISA Annual Conference*, pp. 703–711.

SHINSKEY, F. G. (1990): "How good are our controllers in absolute performance and robustness?" *Measurement and Control*, **23**, May, pp. 114–121.

ZIEGLER, J. G. and N. B. NICHOLS (1942): "Optimum settings for automatic controllers." *Trans. ASME*, **64**, pp. 759–768.