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

Model-free PID controller optimization for loop shaping

Masami Saeki *

* *Hiroshima University, Higahsi-Hiroshima, 739-8527 Japan (Tel: 81-82-424-7589; e-mail: saeki@ hiroshima-u.ac.jp).*

Abstract: In this paper, a data-driven model-free design method of PID controller is proposed for a single-input single-output linear time-invariant plant for a loop shaping problem where the integral gain is maximized under the maximum sensitivity constraint. This design problem is reduced to a linear programming problem. The constraints on the PID gains are given from many fictitious data, which are obtained by applying the wavelet transform to a step response. Our design method does not require the iteration of modeling and control design and the performance index refinement. Numerical examples show the robustness against measurement noise.

1. INTRODUCTION

Since PID control is widely used, such design methods that do not require trial and error for tuning and that depends on a plant response that can be easily measured would be desirable. Although mathematical model plays key roles in the analysis and the synthesis of control systems, modeling error causes the intrinsic difficulty such as the iteration of modeling and control design(Albertos et al [2002]). In order to avoid modeling errors, data-driven designs that do not rely on plant models would be useful. Further, in order to avoid weight refinement in the design, a problem setting whose performance index does not use weighting function would be useful. Such a design problem is given by Åström et al [1998], where the integral gain is maximized subject to the maximum sensitivity constraint.

There are a few model-free methods; for example, unfalsified control(Safonov et al. [1997]), iterative feedback tuning (Hjalmarsson et al. [2003]), and virtual reference feedback tuning (Cabral et al. [2003]). The unfalsified control is a method of falsifying the controllers that do not satisfy the performance criterion by the input-output response data. In the iterative feedback tuning, the performance index of LQ theory is minimized by a descent method by iterating experiment and design. The virtual reference feedback tuning reduces a matching problem of control design to the identification problem of the controller.

Unfalsified control has the features that no model is required and the performance criterion of the robust control can be treated. Robust adaptive control has been studied (Jun et al. [1999]Cabral et al. [2003],Wang et al. [2005]). The author applied it to the off-line control design, and proposed a method of drawing the set of PID gains falsified by the criterion of the mixed sensitivity control on the parameter plane(Saeki. [2004]). Further, the class of input output responses that effectively falsify the controllers is clarified(Saeki et al. [2004]), and a method of generating fictitious signals from a single response data is proposed (Saeki et al. [2006]). It is useful to visualize the shape of the unfalsified regions on the parameter plane. However, it is not easy when the number of parameters is more than two.

In this paper, we will give a loop shaping design method of PID gains by developing an optimization method for the problem of Åström et al [1998] with the additional condition that no mathematical model but a plant response is given.

For signals $w(t), v(t), t \in [0, \infty)$, we use the following notations. $\|w\|_2 = \sqrt{\int_0^\infty w^2(\tau)d\tau}$, $\|w\|_{2T} = \sqrt{\int_0^T w^2(\tau)d\tau}$, $\langle w, v\rangle = \int_0^\infty w(\tau)v(\tau)d\tau$, and $\langle w, v\rangle_T = \int_0^T w(\tau)v(\tau)d\tau$.

2. A LOOP SHAPING PROBLEM

Let us consider the feedback system described by

$$y = Pe \tag{1}$$

$$e = w - u \tag{2}$$

$$u = Ky \tag{3}$$

The plant P is linear time-invariant and single-input and single output. K is a PID controller given by

$$K(s) = K_P + K_I \frac{1}{s} + K_D s \tag{4}$$

The maximum sensitivity is defined as

$$M_s = \max_{0 \le \omega < \infty} \left| \frac{1}{1 + P(j\omega)K(j\omega)} \right| \tag{5}$$

$$= \max_{0 \le \omega < \infty} |S(j\omega)| \tag{6}$$

where S is the sensitivity function. M_s is a good index for the sensitivity and the stability margin, and typical values of M_s are in the range of 1.2 to 2 (Åström et al [1994]).

For $\gamma \in [1.2, 2]$, this condition is represented by

$$\|S(s)\|_\infty < \gamma \tag{7}$$

or it is represented by the time domain condition:

$$\|e\|_2 < \gamma \|w\|_2 \tag{8}$$

is satisfied for all $w \in L_2$ where $e = Sw$.

In Åström et al [1998], an optimization problem where the integral gain of the PID controller is maximized under the maximum sensitivity constraint is given based on the property that the integrated error IE is the reciprocal of the integral gain. This problem is also interpreted from loop shaping.

When $K_I P(0) \neq 0$, the next approximation is satisfied for low frequencies.

$$|S(j\omega)| \approx \frac{\omega}{K_I P(0)} \tag{9}$$

Thus, when the condition (7) is satisfied and $K_I P(0) \neq 0$, the gain plot of the sensitivity function has the shape illustrated in Fig. 1.

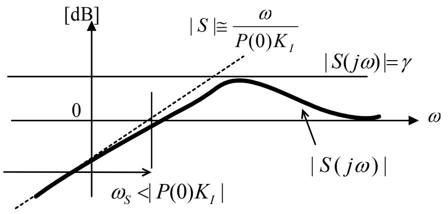

Fig. 1. Loop shaping for the sensitivity function

By substituting the approximation (9) into $|S(j\omega)| < 1$,

$$\omega < K_I P(0) \tag{10}$$

It is expected that the disturbance attenuation is improved by feedback in the frequency range $[0, K_I P(0)]$, and that the frequency range can be made wider by making K_I larger.

Remark 1 Let us compare this with the criterion $\|V(s)S(s)\|_\infty < \gamma$ that uses a weighting function $V(s)$. From this,

$$|S(j\omega)| < \frac{\gamma}{|V(j\omega)|}, \omega \in R \tag{11}$$

Roughly speaking, an appropriate function V is determined as the reciprocal of the optimal S. However, the optimal S cannot be known before giving the appropriate V, which results in the iteration of the control design for the weight selection. In our problem setting, the controller gains are the only parameters to be determined, because γ can be easily selected. Another drawback is that the gain shape cannot be specified freely, because $V(s)$ is usually a low order rational function. For example, suppose that a designer wants to set the bound shown by the lines of Fig. 2. Then, the designer probably chooses it without confidence, say $V(s) = 0.7(s + 10)/(s + 1)$, whose gain plot is shown by the curve. The bound given by the curve is more restrictive than that given by the lines.

3. PROBLEM SETTING AND DERIVATION OF CONSTRAINT

3.1 Problem setting

Assume that an input-output response of the plant, $e(t), y(t)$, is given on the finite interval $t \in [0, T]$ where the

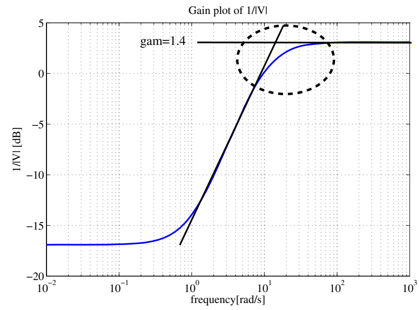

Fig. 2. Gain plot of $1/|V(j\omega)|$

plant is at the steady state for $t < 0$, i.e. $e(t) = 0, y(t) = 0$ for $t < 0$. Further assume that the plant is linear time-invariant (This assumption will not be used in this section, but it will be used in the next section for the generation of the fictitious data by filtering e and y).

Our problem is to construct a data driven design method of PID controller for the next loop shaping problem on the above assumptions.

Loop shaping problem : For the feedback system (1)-(3), find a PID controller that maximizes the integral gain K_I subject to

$$\|e\|_2 < \gamma \|w\|_2 \tag{12}$$

for all $w \in L_2$.

Note that the performance criterion (12) is represented in the time domain instead of (7), because it is suitable for the following discussions.

3.2 Derivation of constraint

Lemma 1 Suppose that the system satisfies causality and it is in the steady state at $t = 0$. If (12) is satisfied, then

$$\| e \|_{2T} < \gamma \| w \|_{2T} \tag{13}$$

for any $T > 0$.

This lemma means that if (13) is not satisfied, (12) is not satisfied neither. This basic lemma appears in the input-output stability theory, and Safonov used it first for the unfalsified control(Safonov et al. [1997]). The condition (12) requires the infinite time interval data, whereas (13) requires the data in the finite time-interval $[0, T]$.

From (13)

$$\langle w, w \rangle_T > \frac{1}{\gamma^2} \langle e, e \rangle_T \tag{14}$$

The disturbance w that gives e, y is given by

$$w(t) = e(t) + K(s)y(t) \tag{15}$$
$$= e(t) + \hat{y}(t)^T \hat{K} \tag{16}$$

where

$$y_I(t) = \int_0^t y(\tau) d\tau \tag{17}$$

$$y_D(t) = \frac{dy}{dt}(t) \tag{18}$$

$$\hat{K} = [\, K_P \;\; K_I \;\; K_D \,]^T \tag{19}$$

$$\hat{y}(t) = [\, y(t) \;\; y_I(t) \;\; y_D(t) \,]^T \tag{20}$$

w given by (16) is called a fictitious disturbance in the model-free control methods.

Then, by substituting (16) into (14), we obtain

$$\langle e+v, e+v \rangle_T > \frac{1}{\gamma^2} \langle e, e \rangle_T \tag{21}$$

where

$$v(t) = \hat{y}(t)^T \hat{K} \tag{22}$$

This inequality gives a concave constraint on $\hat{K}$ for fixed e and y. Since it is difficult to treat non-convex constraints in optimization problems generally, in the following, we will derive a linear constraint from (21) as a sufficient condition.

For any $v_0(t)$,

$$\langle v - v_0, v - v_0 \rangle_T \geq 0 \tag{23}$$

Equivalently,

$$\langle v, v \rangle_T \geq 2 \langle v_0, v \rangle_T - \langle v_0, v_0 \rangle_T \tag{24}$$

From (21),

$$\langle e, e \rangle_T + 2 \langle e, v \rangle_T + \langle v, v \rangle_T > \frac{1}{\gamma^2} \langle e, e \rangle_T \tag{25}$$

From (24) and (25),

$$\langle e + v_0, v \rangle_T \geq \frac{1}{2} \left\{ \left(\frac{1}{\gamma^2} - 1 \right) \langle e, e \rangle_T + \langle v_0, v_0 \rangle_T \right\} \tag{26}$$

This inequality is expressed as

$$a_1 K_P + a_2 K_I + a_3 K_D > b \tag{27}$$

Thus, the next constraint is obtained.

$$a\hat{K} > b \tag{28}$$

where $a = [\, a_1 \;\; a_2 \;\; a_3 \,]$ and

$$a_1 = \langle e + v_0, y \rangle_T \tag{29}$$

$$a_2 = \langle e + v_0, y_I \rangle_T \tag{30}$$

$$a_3 = \langle e + v_0, y_D \rangle_T \tag{31}$$

$$b = \frac{1}{2} \left\{ \left(\frac{1}{\gamma^2} - 1 \right) \langle e, e \rangle_T + \langle v_0, v_0 \rangle_T \right\} \tag{32}$$

3.3 Selection of v_0

We will consider the case that a PID gain $\hat{K} = \hat{K}_a$ that stabilizes the closed-loop system is given. Denote $v_a(t) = \hat{y}(t)^T \hat{K}_a$ and assume that (21) is satisfied for $v(t) = v_a(t)$.

The set of v that satisfies (21) corresponds to the outside region of the sphere with center $-e$ and radius $\|e\|_2/\gamma$ (Fig. 3). This set is concave and v_a lies outside the sphere by assumption. Let v_0 be the intersection of the segment that connects $-e$ and v_a and the sphere. We consider that the sphere is appropriately approximated by the plane which touches the sphere at the point v_0 as illustrated in Fig. 3. Note that this convex set determined by the plane is described by (28) with this v_0.

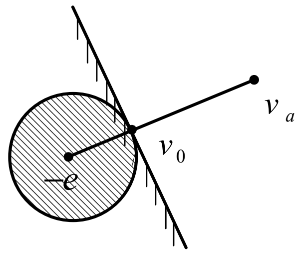

Fig. 3. Approximation of the concave region by plane

Now, calculate the point v_0. The segment is described by

$$v = qv_a + (1-q)(-e), \quad 0 \leq q \leq 1 \tag{33}$$

Substituting this into (21),

$$q^2 \langle e + v_a, e + v_a \rangle_T \geq \frac{1}{\gamma^2} \langle e, e \rangle_T \tag{34}$$

Then, the minimum value of q is found to be

$$q_0 = \frac{1}{\gamma} \frac{\|e\|_{2T}}{\|v_a + e\|_{2T}} \tag{35}$$

, and v_0 is given by

$$v_0 = q_0 v_a - (1 - q_0) e \tag{36}$$

Note that $0 < q_0 < 1$ from the assumption that (21) is satisfied for $v = v_a$. If $q_0 \geq 1$, this implies that v_a lies inside the sphere and the linear approximation does not hold.

4. FICTITIOUS DATA GENERATION BY FILTERING

From the discussion in the previous section, one linear constraint (28) is derived from an input-output response $e(t), y(t), t \in [0, T]$ by the equations (35),(36). In order to determine the PID gains, many linear inequalities are necessary. One method is to experiment many times to obtain many response data. But it is usually difficult. Therefore, we will generate many fictitious data $e_i(t), y_i(t)$ by

$$e_i(t) = F_i(s) e(t) \tag{37}$$

$$y_i(t) = F_i(s) y(t), t \in [0, T] \tag{38}$$

where $F_i(s)$ is a filter transfer function.

Since P is assumed to be linear time-invariant,

$$y_i(t) = P(s) e_i(t) \tag{39}$$

is obviously satisfied. This means that the data $e_i(t), y_i(t)$ is the input-output response of the plant.

One of the important representation form of the system dynamics is the frequency response. It is shown by numerical examples that the many frequency response data can falsify the gains effectively in (Saeki. [2004]), and that the above filtering method is proposed in (Saeki et al. [2006]). In this section, we will give a new $F_i(s)$ and explain the filtering from the viewpoint of wavelet transform(C.K. Chui [1992]).

The input-output response treated here is a non-stationary signal, and the important information about the frequency property lies locally in the time-domain. In these ten years, wavelet transform is becoming popular as the time-frequency analysis tool. It is also known that wavelet transform is a signal processing with many band-pass filters.

In order to extract the frequency components of $\omega_i, i = 1, 2, \dots, n_F$, we will use the next bandpass filters $F_i(s)$. The gain plot of $\psi(s)$ is shown in Fig. 4.

$$F_i(s) = \hat{\psi}(s/\omega_i) \tag{40}$$

$$\hat{\psi}(s) = \left(\frac{s}{(s+\alpha)^2 + 1}\right)^4 \tag{41}$$

Exactly speaking, the peak gain is taken not at ω_i but at $\omega_i\sqrt{1+\alpha^2}$. Since this is not important, we do not take the error into consideration. ω_i are usually logarithmically equally spaced frequencies. The impulse response $\psi(t)$ of $\hat{\psi}(s)$ with $\alpha = 0.5$ is shown in Fig. 5.

In (Saeki et al. [2006]), the bandpass-filter made from the butterworth filter is used. Though the difference by the filters is small, $F_i(s)$ is more suitable for the analysis because the transfer function is simple.

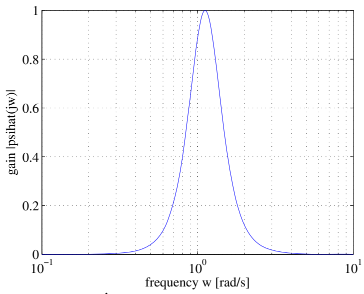

Fig. 4. Gain plot of $\hat{\psi}(j\omega)$

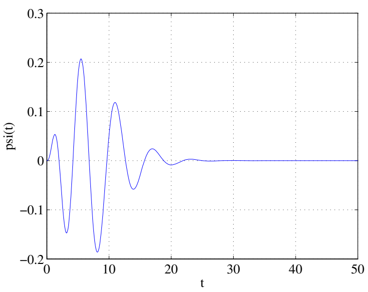

Fig. 5. Impulse response $\psi(t)$ for $\alpha = 0.5$

The impulse response of $F_i(s) = \hat{\psi}(s/\omega_i)$ is represented as

$$L^{-1}\{F_i(s)\} = \omega_i \psi(\omega_i t) \tag{42}$$

By the convolution formula,

$$y_i(t) = F_i(s)y(t) \tag{43}$$

$$= \omega_i \int_0^t \psi(\omega_i(t-\tau))y(\tau)d\tau \tag{44}$$

On the other hand, the integral wavelet transform W_ϕ by the mother wavelet $\phi(t)$ is defined by

$$(W_\phi y)(b,a) = |a|^{-1} \int_{-\infty}^{\infty} \phi\left(\frac{\tau - b}{a}\right) y(\tau)d\tau \tag{45}$$

By comparing (44) and (45), we can immediately find the next correspondense.

$$a \leftrightarrow \frac{1}{\omega_i}, \quad b \leftrightarrow t, \quad -\phi(-t) \leftrightarrow \psi(t) \tag{46}$$

Since $-\phi(-t)$ corresponds to $\psi(t)$, the graph of $-\phi_{db10}(-t)$ is shown in Fig. 6 for the Daubechies wavelet "db10". Note that $\alpha = 0.5$ is the value so that $\psi(t)$ may be close to $-\phi_{db10}(-t)$. From the uncertainty principle in the wavelet analysis, there is a trade-off between the time window and the frequency window. The time-frequency window can be tuned by the parameter α.

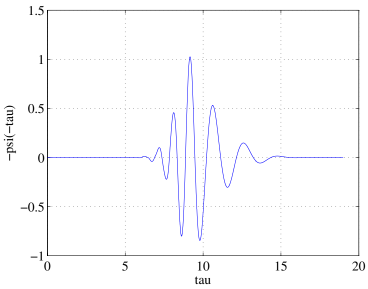

Fig. 6. Mother wavelet db10 $y = -\phi_{db10}(-\tau)$

The next admissible condition is required for the Fourier transform $\hat{\phi}(\omega)$ of the mother wavelet for the existence of the inverse transform.

$$\int_{-\infty}^{\infty} \frac{|\hat{\phi}(\omega)|^2}{|\omega|} d\omega < \infty \tag{47}$$

Since $\hat{\psi}(s)$ has four zeros at $s = 0$ and the difference between the numerator and the denominator degrees is four, the admissible condition is satisfied.

Thus, the filtering is considered as the wavelet transform and therefore we can expect that this filtering will be useful for extracting the local information from the non-stationary response.

5. ALGORITHM

In this section, we summarize the previous results as an algorithm for control design.

In the first step, the input output response $e(t), y(t), t \in [0, T]$ is measured by exciting the system at the steady state. If $y(0), e(0)$ are nonzero at the steady state, they should be unbiased as $y(t) - y(0), e(t) - e(0), t \in [0, T]$ before computation.

In the second step, set $\omega_i, i = 1, 2, \dots, n_F$ as logarithmically equally spaced n_F points of an important frequency range for control. Then, the next fictitious responses are generated by (38) from $e(t), y(t), t \in [0, T]$.

$$e_i(t), y_i(t), t \in [0, T], i = 1, 2, \dots, n_F \tag{48}$$

In the third step, give a PID gain $\hat{K}_a$ that stabilizes the closed-loop system and that satisfies (12) for a given γ (For stable plants, this condition is satisfied for $K(s) = 0$ or $K(s) = K_P$ where K_P is sufficiently small. Therefore, $\hat{K}_a$ is easily found for stable plants). Then, compute linear constraints for n_F set of responses $e_i(t), y_i(t)$, and represent them as $a_i \hat{K} < b_i$ for $i = 1,\ 2,\ \dots,\ n_F$.

In the last step, solve the next linear programming problem.

Minimize

$$J = c^T \hat{K} \tag{49}$$

subject to

$$A\hat{K} > B \tag{50}$$

where

$$A = [a_1^T, a_2^T, \cdots, a_{n_F}^T]^T \in R^{n_F \times 3}$$
$$B = [b_1, b_2, \cdots, b_{n_F}]^T \in R^{n_F \times 1}$$
$$c = [0 \ \ 1 \ \ 0]^T$$

In the programming, the integral and the derivative calculations are approximated by discretization.

Remark: When $y(0), e(0)$ are nonzero at the steady state and $y(t), e(t)$ are given to the filter whose initial condition is zero, the filtered outputs do not satisfy the input-output relation of the plant. Therefore, the operation is essential in the first step.

6. NUMERICAL EXAMPLES

Suppose that $e(t), y(t), t \in [0, 30]$ are given as shown in Fig. 7. This data is obtained for the feedback system

$$y = Pe \tag{51}$$
$$e = K(r - y) \tag{52}$$

for $r(t) = 1, t \in [0, 30]$ where $K(s) = 0.1$ and

$$P(s) = \frac{-s+1}{(s+1)^2} e^{-s} \tag{53}$$

Since the steady state error $r(\infty) - y(\infty)$ is very large, this response is not good. Note that this transfer function and the related information that the plant is non-minimum phase with time delay will not be utilized for design.

6.1 Design of PI controller

Let us design a PI controller for $\gamma = 1.7$. ω_i's are logarithmically equally spaced $n_F = 90$ points in the interval $[0.1,\ \ 10]$. For $\hat{K}_a = [K_P,\ K_I]^T = [0.1,\ 0]^T$, the linear constraints are derived. The solution is obtained as

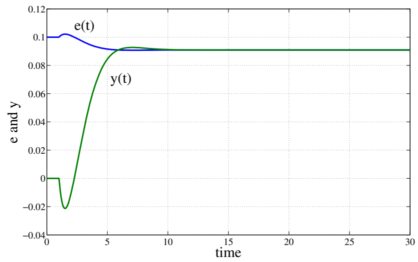

Fig. 7. Step response of the closed-loop system for $K = 0.1$

$$K(s) = 0.3756 + 0.2092/s \tag{54}$$

by solving the linear programming problem once.

The gains of the sensitivity function S and the complementary sensitivity function T are shown in Fig. 8. M_s can be larger than γ, because the basic condition (13) is a necessary condition, and γ is a little bit larger than the specified value 1.7. A good loop shaping is attained and the step response for $r = 1$ is shown in Fig. 9.

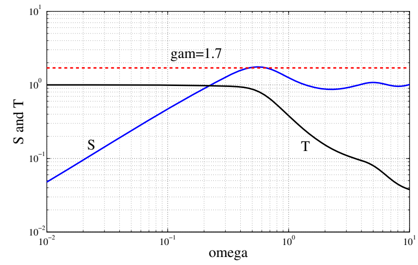

Fig. 8. S and T

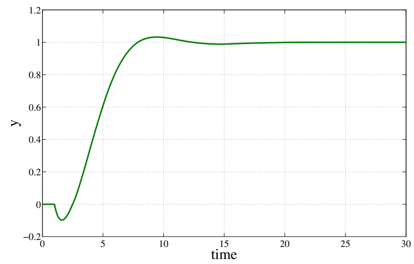

Fig. 9. Step response of the closed-loop system

The linear inequality has been obtained by approximating the concave inequality as a sufficient condition. In order to make this certain visually, the concave regions are drawn on the $[K_P, K_I]$ plane in Fig. 10. In this figure, each curve is the part of an oval, and the outside of the oval is the region defined by (21). By the way, the stability condition requires $K_I > 0$ theoretically. Therefore, the shaded portion is the region unfalsified by (21) and the stability condition.

Each oval is approximated by a half-plane as shown in Fig. 11 where the point $\hat{K}_a = [0.1,\ 0]$ is shown by the diamond. The polygon region determined by the linear inequalities is found to be a good approximation of the unfalsified region of Fig. 10. The solution of the linear programming problem is given by the point at which K_I is maximized in the polygon region.

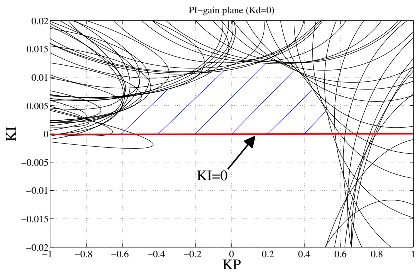

Fig. 10. Inside the ovals are falsified on the parameter plane

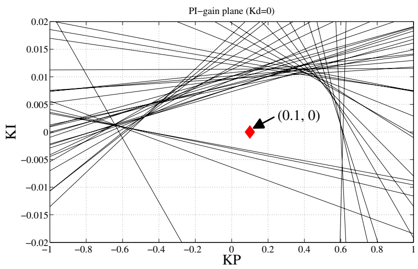

Fig. 11. Linearly approximated regions on the parameter plane

6.2 Robustness against observation noise

The response data affected by the observation noise is shown in Fig. 12. From this data,

$$K(s) = 0.3763 + 0.2165/s \tag{55}$$

, which shows the insensitiveness of our design method against noise. The step response for $r(t) = 1$ is shown in Fig. 13.

Our method also gives nice results for both PI and PID controllers for other typical industrial processes given in Chapter 5 of Åström et al [1994] where PID design is more affected by the noise. Since the filter can remove the constant disturbance effect from the response data, our method can be also applied to a certain case even if disturbance exists.

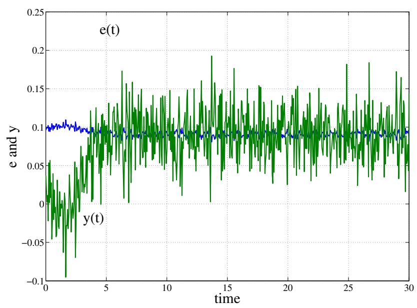

Fig. 12. Step response of the closed-loop system for $K = 0.1$ with observation noise

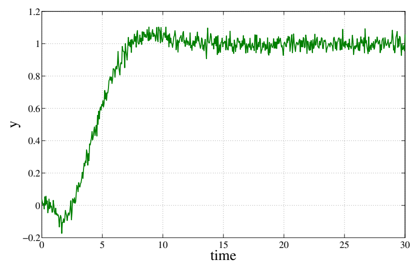

Fig. 13. Step response of the closed-loop system

7. CONCLUSION

In this paper, we proposed a data-driven PID controller design method for loop shaping. The constraints on the PID gains are derived from a step response by applying the wavelet transform to the response to generate many fictitious responses. Then, a PID gain that maximizes the integral gain subject to the constraints is obtained by solving a linear programming problem once. Extension to a multi-input multi-output case is given in Saeki. [2008].

REFERENCES

P. Albertos and A. Sala. *Iterative Identification and Control*, Springer-Verlag, London, 2002.

K. Åström and T. Hägglund. *PID controllers: Theory, design, and tuning*. Instrument Society of America, North Carolina, 1994.

K. Åström, H., Panagopoulous, and T. Hägglund. Design of PI controllers based on non-convex optimization. *Automatica*, volume 34, pages 585–601, 1998.

F.B. Cabral and M.G. Safonov. Unfalsified model reference adaptive control using the ellipsoid algorithm. *Proc. of the 42nd Conference on Decision & Control*, pages 3250–3255, 2003.

M.C. Campi, A. Lecchini, and S. M. Savaresi. Virtual reference feedback tuning: a direct method for the design of feedback controllers. *Automatica*, volume 38, pages 1337–1346. 2002.

C.K. Chui. *An introduction to wavelets*. Academic press, INC., Florida, 1992.

H. Hjalmarsson, M. Gevers, S. Gunnarsson, and O. Lequin. Iterative feedback tuning: Theory and application. *IEEE Control Systems Magazine*, volume 42, pages 843–847. 1999.

M. Jun and M.G. Safonov. Automatic PID tuning. *Proc. of the 38th Conference on Decision & Control*, pages 3250–3255,1999.

M. Saeki. Unfalsified control approach to parameter space design of PID controllers. *Trans. of the Society of Instrument and Control Engineers*, volume 40, pages 398–404. 2004.

M. Saeki, A. Takahashi, O. Hamada, and N. Wada. Unfalsified parameter space design of PID controllers for nonlinear plants. *Proc. of the 2004 IEEE International Conference on Control Applications*, pages 1521–1526, 2004.

M. Saeki, O. Hamada, ,N. Wada, and I. Masubuchi. PID gain tuning based on falsification using bandpass filters. *Proc. of SICE-ICCAS, Busan, Korea*, pages 4032–4037, 2006.

M.G. Safonov and T.C. Taso. The unfalsified control concept and learning. *IEEE Trans. on Automatic Control*, volume Ac-42, pages 843–847, 1997.

R. Wang, A. Paul, M. Stefanovic, and M.G. Safonov. Cost-detectability and stability of adaptive control systems. *Proc. of the 44th Conference on Decision & Control*, 2005.

M. Saeki. MIMO PID controller design for model-free loop-shaping problem. *Proc. of CCS2008, Kyoto, Japan*, March, 2008 (to be presented).