

OPTIMAL INPUT DESIGN FOR IDENTIFICATION OF CONTINUOUS-TIME SYSTEMS

Masami Iwase* Makoto Shigi*
Shoshiro Hatakeyama*

* Dept. of Computers and Systems Engineering
Tokyo Denki University, Hatoyama, Hiki-gun, Saitama,
350-0394, JAPAN, iwase@k.dendai.ac.jp

Abstract: In this paper, we present an optimal input design method for the identification of single input single output continuous-time transfer functions. In the proposed algorithm, we show that this optimal input design problem can be rewritten as that for discrete-time systems proposed by Antoulas et al. (Antoulas and Anderson, 1999; Antoulas, 1997; Antoulas and Astolfi, 1998), if the input is approximated by the finite Fourier series expansion. The derivative signals of input/output are required to design the optimal input, and filters are used to obtain the derivative signals. We also numerically consider the effect of the filters to the optimality. Through numerical examples, its effectiveness is verified.

Copyright ©2005 IFAC

Keywords: Continuous time systems, Identification, Input Design

1. INTRODUCTION

The persistently exciting signals are usually chosen as input signals for system identification (Katayama, 1994; Suda and Nakamizo, 1988). However, the persistently exciting condition should be different if target systems and/or chosen identification methods are different. For example, considering the identification of discrete-time systems, the maximum length sequence is used frequently, whereas we know that if the sequence is used for the identification of continuous-time systems, the persistently exciting condition might not be satisfied (Yang, 1998; Iikubo *et al.*, 2002; Iwase *et al.*, 2002). Even in the discrete-time case, the maximum period and minimum pulse width of the maximum length sequence have to be chosen to adapt it to the target system.

This kind of problems can be formulated as an optimal input design problem what input identifies a system with highest accuracy. We can find one answer in Antoulas's paper (Antoulas and Anderson, 1999). They introduced the misfit

function, which can be regarded as the distance between signal space and noise space, into the identification problem of discrete-time systems. Then they gave a criterion using singular values of an input-output data matrix representing the system in the kernel form, and proposed a design method of the optimal input with the criterion. However, continuous-time system case has not been discussed. Continuous-time systems are often preferred rather than discrete-time systems from the viewpoint of control system design and analysis, since the relation between pole-zeros and time responses of system is able to be understood intuitively (Haverkamp *et al.*, 1996; Haverkamp *et al.*, 1997).

Therefore, in this paper, we discuss the optimal input for continuous-time transfer function, and propose a design method given by extending the Antoulas's method to the continuous-time case. First, we consider a representation of the optimal input. In our method, input and output signals are differentiated by some filters, and are measured at each sampling intervals. The sampled data can

be approximated by using the Fourier series expansion. Using the approximated input, we show that the identification problem for continuous-time systems can be reformulated as that for discrete-time systems proposed by Antoulas et al. Then the optimization algorithm discussed in their paper (Antoulas and Anderson, 1999) can be easily extended to our case. Through numerical simulations, we verify the effectiveness and robustness of the proposed method, and compare the method with a traditional case using the maximum length sequence.

This paper is organized as follows. In section 2, we describe how to differentiate and measure the input-output signals with filters. Using the input-output data, a data matrix representing the system and a criterion evaluating the optimality are defined. In section 3, we discuss how to design the optimal input. Especially, we show that our problem is equal to the Antoulas's problem for discrete-time system by considering the Fourier series expanded input. In next section, we present simulation results and verify the robustness and effectiveness. Finally, some concluding remarks are given.

2. PRELIMINARIES AND PROBLEMS

2.1 Input-output Data

The proposed algorithm requires the derivatives of input-output signals. We obtain the derivatives not by such numerical methods as difference approximation but by filters $s^i H(s)$, and assume that $H(s)$ is strictly proper and its relative degree is high enough to realize $s^i H(s)$ in the state space representation. The filters $s^i H(s)$ are set as shown in Fig.1 and let $\hat{v}^{(i)}(t)$, $\hat{y}^{(i)}(t)$ be the output of the filters. $\hat{v}(t)$, $\hat{y}(t)$ are generated from $v(t)$, $y(t)$ as follows:

$$\hat{v}(t) = \int_0^t h(\tau)v(t-\tau) d\tau \quad (1)$$

$$\hat{y}(t) = \int_0^t h(\tau)y(t-\tau) d\tau. \quad (2)$$

i -th derivatives of (1) and (2) are

$$\hat{v}^{(i)}(t) = \int_0^t h^{(i)}(\tau)v(t-\tau) d\tau, \quad (3)$$

$$\hat{y}^{(i)}(t) = \int_0^t h^{(i)}(\tau)y(t-\tau) d\tau. \quad (4)$$

The filters need to be discretized in order to be implemented practically. However, in this paper, the discretized filters can be considered as continuous-time systems because the filters are implemented on fast hardware devices such as DSP. On the

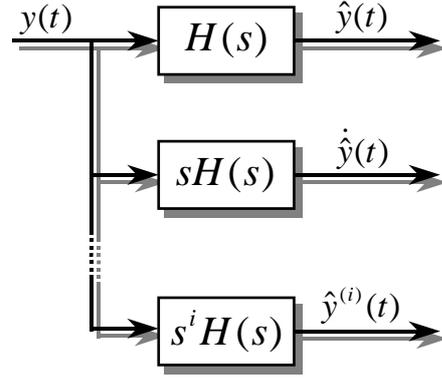


Fig. 1. Block diagram of differential Filter

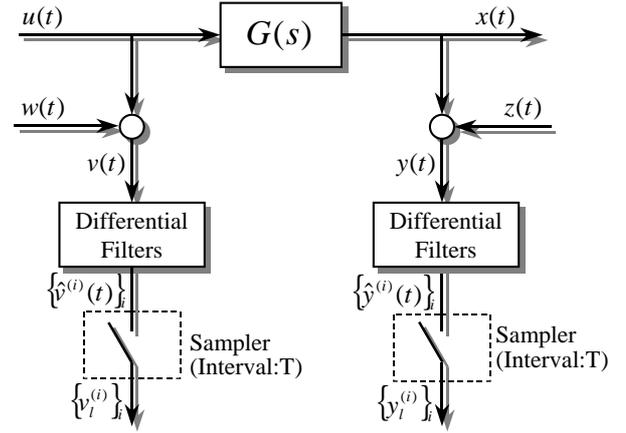


Fig. 2. Block diagram of system configuration

other hand, sampling interval in measurement of $\hat{v}^{(i)}(t)$, $\hat{y}^{(i)}(t)$ can be longer than filters' sampling. If a lowpass filter is chosen as $H(s)$, it can be expected to suppress high-frequency noises contaminating input-output signals. How to design the filters is an important topic for this kind of research. But, due to lack of space, we'd like to focus on only the optimal input design here. On the filtered derivatives, please see the literature (Gawthrop, 1987; Gawthrop, 1990) for example.

2.2 Problem

Let us consider a single input single output linear continuous-time transfer function:

$$G(s) = \frac{p(s)}{q(s)} = \frac{p_0 + p_1 s + \dots + p_m s^m}{q_0 + q_1 s + \dots + q_n s^n} \quad (5)$$

where n is the degree of the denominator, and m is that of the numerator, where $n \geq m$. The noise-free input and output signals of the system (5) are defined as $u(t)$ and $x(t)$, respectively. Using the polynomials p and q , (5) can be rewritten as follows:

$$\begin{bmatrix} p(s) & -q(s) \end{bmatrix} \begin{bmatrix} u(t) \\ x(t) \end{bmatrix} = 0 \quad (6)$$

The parameters vector θ of $G(s)$ is defined as

$$\theta = [p_0 \cdots p_m -q_0 \cdots q_n]^T$$

and an input-output data matrix is also defined as

$$\mathcal{M} = \frac{1}{\sqrt{N+1}} \begin{pmatrix} \mathcal{M}_1 \\ \mathcal{M}_2 \end{pmatrix} \quad (7)$$

$$\mathcal{M}_1 = \begin{bmatrix} \hat{u}_0 & \hat{u}_1 & \cdots & \hat{u}_N \\ \dot{\hat{u}}_0 & \dot{\hat{u}}_1 & \cdots & \dot{\hat{u}}_N \\ \vdots & \vdots & \ddots & \vdots \\ \hat{u}_0^{(m)} & \hat{u}_1^{(m)} & \cdots & \hat{u}_N^{(m)} \end{bmatrix} \quad (8)$$

$$\mathcal{M}_2 = \begin{bmatrix} \hat{x}_0 & \hat{x}_1 & \cdots & \hat{x}_N \\ \dot{\hat{x}}_0 & \dot{\hat{x}}_1 & \cdots & \dot{\hat{x}}_N \\ \vdots & \vdots & \ddots & \vdots \\ \hat{x}_0^{(n)} & \hat{x}_1^{(n)} & \cdots & \hat{x}_N^{(n)} \end{bmatrix} \quad (9)$$

using the differentiated signals $\hat{u}^{(i)}$ and $\hat{x}^{(i)}$. Then (6) is equal to

$$\theta^T \mathcal{M} = 0 \quad (10)$$

which is called the kernel form of the system. With noise perturbing the measurements $\hat{v}^{(i)}$, $\hat{y}^{(i)}$ following Fig.2, the data matrix can be written as

$$\hat{\mathcal{M}} = \mathcal{M} + \bar{\mathcal{M}} \quad (11)$$

where $\bar{\mathcal{M}}$ is the noise data matrix consisting of $\hat{v}^{(i)}$ and $\hat{y}^{(i)}$, which has the same structure of \mathcal{M} . We assume that $\|\bar{\mathcal{M}}\|_2 \leq \epsilon$ with an enough small real number ϵ .

Let us consider the system representation of kernel form in frequency domain. The input signal $u(t)$ is approximated with the finite Fourier series expansion:

$$u(t) = \sum_{k=-N/2}^{N/2} c_k \mu_k(t) \quad (12)$$

where c_k ($k = -N/2, \dots, N/2$) are the expansion coefficients, and N is expansion length. If we could set amplitude of input arbitrarily, we could improve the signal-noise ratio arbitrarily using the enough large input. But this assumption is nonsense. Therefore, for normalization purposes, we will assume that $u(t)$ has unit power, i.e.

$$\sum_k |c_k|^2 = 1$$

$\mu_k(t)$ are the bases functions of the expansion:

$$\mu_k(t) = e^{jk\omega t} \quad (13)$$

where $j = \sqrt{-1}$ and $\omega = 2\pi/(N+1)$. The output $y(t)$ corresponding to the input (12) is

$$y(t) = \sum_{k=-N/2}^{N/2} \psi_k c_k \mu_k(t) \quad (14)$$

where $\psi_k = G(j\omega k)$ are the frequency transfer function of $G(s)$. To represent the data matrix (7) briefly in the frequency domain, some matrices are introduced. A Vandermonde matrix $V_i \in C^{(i+1) \times (N+1)}$ consisting of μ_k is given by

$$V_i = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \mu_{-N/2}(1) & \mu_{-N/2+1}(1) & \cdots & \mu_{N/2}(1) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{-N/2}(i) & \mu_{-N/2+1}(i) & \cdots & \mu_{N/2}(i) \end{bmatrix}$$

, and W_i is also given by

$$W_i = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ j(-\frac{N}{2})\omega & j(-\frac{N}{2}+1)\omega & \cdots & j(\frac{N}{2})\omega \\ \vdots & \vdots & \ddots & \vdots \\ (j(-\frac{N}{2})\omega)^i & (j(-\frac{N}{2}+1)\omega)^i & \cdots & (j(\frac{N}{2})\omega)^i \end{bmatrix}$$

Then (8) and (9) can be rewritten as

$$\mathcal{M}_1 = W_m C \bar{V}_N^*, \quad \mathcal{M}_2 = W_n \Psi C \bar{V}_N^*$$

where $C = \text{diag}(\{c_k\}_{k=-N/2, \dots, N/2})$ and $\Psi = \text{diag}(\{\psi_k\}_{k=-N/2, \dots, N/2})$, and (7) can be

$$\mathcal{M} = \frac{1}{\sqrt{N+1}} \begin{bmatrix} W_m \\ W_n \Psi \end{bmatrix} C \bar{V}_N^* \quad (15)$$

We consider the modified data matrix \mathcal{R} :

$$\mathcal{R} = \frac{1}{\sqrt{N+1}} \mathcal{V} C, \quad \mathcal{V} = \begin{bmatrix} W_m \\ W_n \psi \end{bmatrix}. \quad (16)$$

Obviously, \mathcal{R} satisfies

$$\mathcal{M} = \mathcal{R} \bar{V}_N^* \quad (17)$$

and the singular values of \mathcal{M} and \mathcal{R} are same because V_N is unitary. In addition, \mathcal{R} has the same left kernel of \mathcal{M} , θ :

$$\theta^T \mathcal{R} = \theta^T \mathcal{V} C = 0 \quad (18)$$

Therefore, we will use \mathcal{R} instead of \mathcal{M} in the following discussion.

Now we will describe the main problem how to design an optimal input for identification of $G(s)$. If input-output signals are not contaminated with noise, $\theta^T \mathcal{R} = 0$ is satisfied, and the smallest singular value of \mathcal{R} is equal to 0. Otherwise, $\theta^T \mathcal{R} \neq 0$ and the smallest singular value is not 0 because the noise data matrix $\bar{\mathcal{M}} \neq 0$ and

$$\epsilon \geq \|\bar{\mathcal{M}}\|_2 > 0 \quad (19)$$

Since the noise cannot be removed perfectly in general, inputs less correlative to the noise are

desirable. When the second smallest singular value is bigger enough than the smallest singular value, we can consider that the noise-effect is relatively decreased. Then, we formulate the optimal input design problem as maximizing the second smallest singular value.

Problem:

Given a single input single output linear continuous-time transfer function $G(s)$. The problem is to design an optimal input $u_*(t)$, defined for $t \geq 0$, which will maximize the second smallest singular value σ_* of the data matrix \mathcal{R} defined by (16).

In the next section, we will discuss the answer to this problem.

3. MAIN RESULT

To solve the problem mentioned in the previous section, we introduce the data covariance matrix \mathcal{D} :

$$\mathcal{D} = \mathcal{R}\mathcal{R}^* = \frac{1}{N+1} \mathcal{V}\mathcal{B}\mathcal{V}^* \quad (20)$$

$$\mathcal{B} = \mathcal{C}\mathcal{C}^* = \text{diag}(\beta_0, \dots, \beta_k) \quad (21)$$

where $\beta_k = c_k c_k^* = |c_k|^2$. From the structure of (21), \mathcal{D} is a square matrix with real entries. Furthermore the kernel of \mathcal{D} is the same as that of \mathcal{R} , i.e. $\theta^T \mathcal{D} = 0$.

Because $\beta_k = |c_k|^2$, to maximize the second-smallest singular value σ_* of \mathcal{R} is equal to maximize the second-smallest eigenvalue λ_* of \mathcal{D} . Then, the problem can be converted into an eigenvalue maximizing problem by considering the second-smallest eigenvalue λ_* of \mathcal{D} :

$$\max_{\mathcal{B}} \lambda_*(\mathcal{D}) \text{ where } \mathcal{B} = \text{diag}(\beta_i) \geq 0, \text{Trace} \mathcal{B} = 2\mathcal{B}$$

At this point, we obtain the same formulation as the Antoulas' formulation of the optimal input design for discrete-time systems. Therefore, we also obtain the optimization method of the above eigenvalue maximizing problem from their paper. The Antoulas' optimization method can be applied to our problem easily. Here, we describe the method briefly.

Optimization Algorithm (Antoulas and Anderson, 1999):

- (1) Preliminaries: pick a system.
- (2) Give an initial choice of $\beta_i, i = 1, \dots, N$.
- (3) Compute eigenvalue value decomposition of \mathcal{D} . Let the eigenvalue decomposition of the corresponding data covariance matrix $\mathcal{D}(\beta_i)$ to the given $\{\beta_i\}$ be given by

$$\mathcal{D}(\beta_i) = \mathcal{W}(\beta_i)\Lambda(\beta_i)\mathcal{W}^*(\beta_i) \quad (23)$$

$$\mathcal{W}(\beta_i) = [w_1 \ w_2 \ \dots] \quad (24)$$

$$\Lambda(\beta_i) = \text{diag}(\lambda_1, \lambda_2, \dots) \quad (25)$$

- (4) Compute $\partial\lambda_2/\partial\beta_i$ according to

$$\frac{\partial\lambda_j}{\partial\beta_i} = |w_j^* \mathcal{V}e_{i+1}|^2 - |w_j^* \mathcal{V}e_{N+1}|^2 \quad (26)$$

where e_i denotes the i -th unit vector in R^{N+1} .

- (5) Determine I, K according to

$$I := \max_i \left\{ |w_2^* \mathcal{V}e_i|^2 : \beta_i < 1 \right\} \quad (27)$$

$$K := \min_i \left\{ |w_2^* \mathcal{V}e_i|^2 : \beta_i > 0 \right\} \quad (28)$$

- (6) Update β_i according to the following rule with small $\epsilon > 0$ in order to maximize λ_2 .

$$\begin{cases} \beta_I \rightarrow \beta_I + \epsilon \\ \beta_K \rightarrow \beta_K - \epsilon \\ \beta_i \rightarrow \beta_i, \quad i \neq I, K \end{cases} \quad (29)$$

- (7) Compute change and percentage change of λ_2 using

$$\delta\lambda_2 = \epsilon \cdot \left(|w_2^* \mathcal{V}e_I|^2 - |w_2^* \mathcal{V}e_K|^2 \right) \quad (30)$$

- (8) To repeat procedure with new values β_i , back to 3.
- (9) Stop when β_i do not change significantly.

Please see the detail in the paper (Antoulas and Anderson, 1999). Note that the above algorithm may converge to a local extremum of λ_2 depending on the initial condition β_i .

4. SIMULATIONS

Let us present some numerical examples in order to verify the effectiveness. The system configuration in Fig.2 was used in the simulation. As systems to be identified, a transfer function

$$G(s) = \frac{1}{s^2 + 2s + 3} \quad (31)$$

was considered. In the simulation, we also investigated the effect of choice of the filter $H(s)$ for obtaining the derivatives of input-output signals. Then we considered the following filter with two parameters α and n :

$$H_{\alpha,n}(s) = \frac{\alpha^n}{(s + \alpha)^n}. \quad (32)$$

The candidates of the pole parameter α are from 0.1 to 9.0, and the one of the order parameter n are from 2 to 7. The simulation conditions were set as follows.

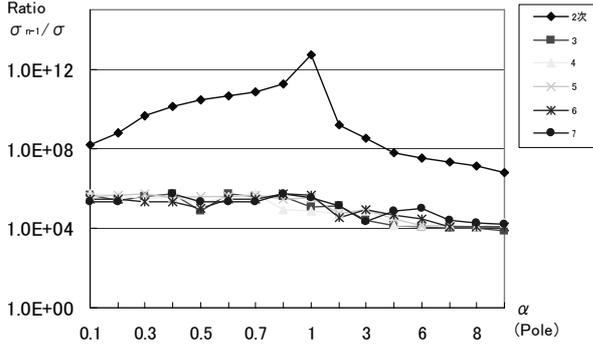


Fig. 3. Ratio of the minimum singular value and the second minimum singular value under each pair of the parameters (α, n) . (This results is of noise-free case.)

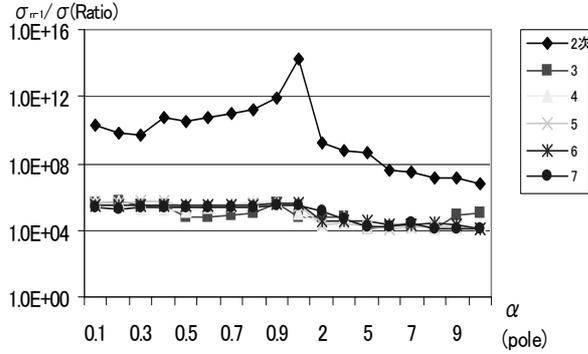


Fig. 4. Ratio of the minimum singular value and the second minimum singular value under each pair of the parameters (α, n) . (This results is of noisy case.)

- Sampling intervals: 10.0 [msec]
- Total time of measurement: 80.0 [sec]
- Number of samples: 16000 [sample]

The measurement noise was assumed to zero-mean white, and the noise-to-signal ratio was set as follows:

$$\frac{\|\text{noise}\|_2}{\|\text{signal}\|_2} = 10\% \quad (33)$$

To verify the effectiveness, the proposed method was compared with the case where the maximum length sequence was used as input. The amplitude and the maximum period of the used maximum length sequence were set as

- Amplitude: ± 1.0
- Maximum period $2^9 - 1$

The designed optimal input for $G(s)$ under some pairs of the parameters (α, n) are shown in are shown in Figs.5–7. Fig.3 shows the ratio of the

Table 1. Analysis results (Noise free case)

Input (α, n)	M-seq (3, 5)	Opt. Input (3, 5)	Opt. Input (3, 2)	Opt. Input (1, 2)
σ_{n-1}/σ_n	-	8.14×10^3	3.56×10^7	5.36×10^{11}
num	0.9875	0.9892	1.0200	1.0000
den	1.0000	1.0000	1.0000	1.0000
Err	2.9753	1.9832	2.0009	2.0100
Err	5.40×10^{-4}	3.58×10^{-4}	7.23×10^{-5}	1.00×10^{-5}

Table 2. Analysis results (Noisy case)

Input (α, n)	M-seq (3, 5)	Opt. Input (3, 5)	Opt. Input (3, 2)	Opt. Input (1, 2)
σ_{n-1}/σ_n	-	8.14×10^3	4.41×10^7	1.80×10^{13}
num	1.0013	0.9881	1.0340	1.0000
den	1.0000	1.0000	1.0000	1.0000
Err	2.0011	1.9810	2.0010	2.0099
Err	2.9431	2.9646	3.0062	2.9999
Err	3.52×10^{-4}	4.40×10^{-4}	7.24×10^{-5}	1.00×10^{-5}

minimum singular value and the second minimum singular value under each pair of the parameters in the noise-free case. This ratio is the measure of the optimality of the input for our identification method. Fig.4 is also the graph of the ratio in the noisy case.

From the figures, we can find that the ratio of the singular values in the proposed case becomes bigger than that of the maximum length sequence case, and moreover that the ratio clearly depends on the choice of the filter's parameters (α, n) . To compare the results, some quantities are listed into the tables 1 and 2. In the tables, to compare the accuracy of identification, the relative square error is defined as

$$\text{Err} = \sqrt{\sum_i \left(\frac{\hat{\theta}_i - \theta_i}{\theta_i} \right)^2} \quad (34)$$

where $\hat{\theta}_i$ are the identified parameters and θ_i are the ideal parameters. From the result, we can see the identification accuracy is improved in our method. Especially, in the noise-contaminated case, we failed in identification by using the maximum length sequence, whereas succeeded by using the optimal input. Note that the optimality of the input clearly depends on the choice of the filters, and that if one chose better parameters, the accuracy also could be improved. The optimal input can be considered robust against the noise. Thus, it is shown that the proposed method is effective for the input signal design for continuous-time systems.

5. CONCLUSION

In this paper, we presented a design method of the optimal input for the identification of single input single output continuous-time transfer functions. As a criterion of optimality, we used the singular value of the input-output data matrix representing the system in the kernel form. Our method is to design the input signal in order to

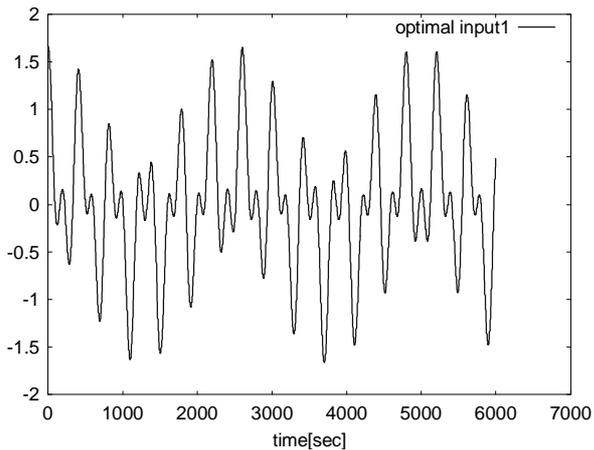


Fig. 5. Designed optimal input for $G(s)$ under $(\alpha, n) = (3, 5)$

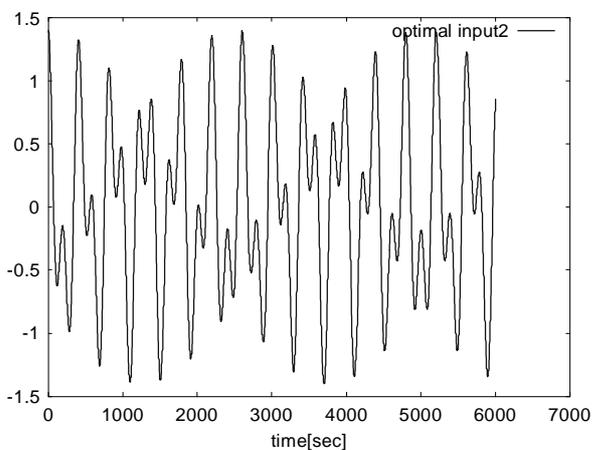


Fig. 6. Designed optimal input for $G(s)$ under $(\alpha, n) = (3, 2)$

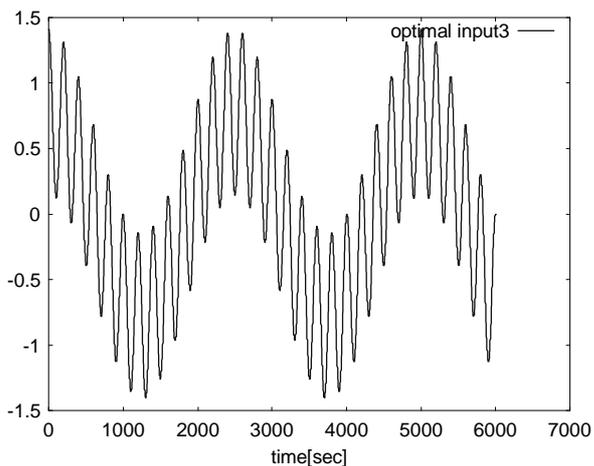


Fig. 7. Designed optimal input for $G(s)$ under $(\alpha, n) = (1, 3)$

maximize the second smallest singular value. Considering the finite Fourier series expanded input signal, our problem could be reformulated as the problem proposed by Antoulas et al. for discrete-time system. Through the numerical simulations, the effectiveness of the proposed method was ver-

ified. Especially, the optimal input generated by our method was robust to the noise, and made identification be success even though failed if the maximum length input was used.

REFERENCES

- Antoulas, A. C. (1997). On modeling for robust control. *Proc. of the IFAC conference on SYSID '97, Kyusyu* pp. 1707–1709.
- Antoulas, A. C. and A. Astolfi (1998). Optimizing the ev of the data matrix: A case study in non-smooth analysis. *Proc. of the IEEE CDC '98, Tampa, Fl.* pp. 3378–3383.
- Antoulas, A. C. and B. D. O. Anderson (1999). On the choice of inputs in identification for robust control. *Automatica* **35**, 1009–1031.
- Gawthrop, P. J. (1987). Continuous-time self-tuning control design – mechanical engineering research studies - engineering control series 1 –. *Reseach Studies Pr.*
- Gawthrop, P. J. (1990). P. j. gawthrop, continuous-time self-tuning control design – mechanical engineering research studies - engineering control series 2 –. *Reseach Studies Pr.*
- Haverkamp, B. R., C. Chou, M. Verhaegen and R. Johansson (1996). Identification of continuous-time mimo state space models from sampled data in the presence of process and measurement noise. *Proc. of the 35th IEEE CDC* pp. 1539–1544.
- Haverkamp, B. R., C. Chou, M. Verhaegen and R. Johansson (1997). Continuous-time subspace model identification using laguerre filtering. *Proc. of the 11th IFAC/IFORS Symposium on Identification and System Parameter Estimation* pp. 1143–1148.
- Iikubo, H., M. Iwase, Y. Shiino and S. Hatakeyama (2002). On an identification method for continuous-time transfer functions using sampled data. *Proc. of SICE Annual Conference in Osaka* pp. in CD-ROM.
- Iwase, M., H. Iikubo, S. Hatakeyama and K. Furuta (2002). An identification method for continuous-time transfer functions based on nonlinear optimization. *Proc. of IECON'02, Nov. 5-8 in Spain* pp. in CD-ROM.
- Katayama, T. (1994). Introduction to system identification. *Asakura Bookstore Co. Ltd.* pp. 39–48.
- Suda, N. and T. Nakamizo (1988). Computrol. *Corona Co. Ltd.* pp. 79–84.
- Yang, Zi-Jiang (1998). Subspace model identification of continuous-time systems with the aid of the ω -operator. *Trans. of the SICE* pp. 546–554.