Optimal Input Signal Design for a Second Order Dynamic System Identification Subject to D-Efficiency Constraints

Wiktor Jakowluk

Optimal input signal design for a second order dynamic system identification subject to D-efficiency constraints

Wiktor Jakowluk

Bialystok University of Technology, Faculty of Computer Science
Wiejska 45a, 15-351 Bialystok, Poland
w.jakowluk@pb.edu.pl
http://wi.pb.edu.pl/pracownicy/

Abstract. System identification, in practice, is carried out by perturbing processes or plants under operation. That is why in many industrial applications an optimal input signal would be preferred for system identification. In this case, the objective function was formulated through maximisation of the Fisher information matrix determinant (D-optimality) expressed in conventional Bolza form. As setting such conditions of the identification experiment we can only say about the D-suboptimality, we quantify the plant trajectories using the D-efficiency measure. An additional constraint, imposed on D-efficiency of the solution, should allow to attain the most adequate contents of information from the plant which operating point is perturbed in the least invasive way. A simple numerical example, which clearly demonstrates the idea presented in the paper, is included and discussed.

Keywords: system identification, optimal input signal, D-optimality, D-efficiency

1 Introduction

The choice of an input signal used for actuation of the system is critical in the task of model building and parameter identification. System identification is the process of constructing an accurate and reliable dynamic mathematical model of the system from observed data and available knowledge. It is a common practice to perturb the system of interest and use the resulting data to build the model [1, 2]. The accuracy of parameter estimates is increased by the use of optimal excitation signals [3].

The pertinence of a model is the critical factor for proper tuning of a controller, usually performed as a model-based optimisation task. Inaccurate model can significantly influence the performance of the control loop, and finally deteriorate the quality of the plant product. The control performance assessment has a large impact on the economic aspect of the production process. It was found that about 66% - 80% of the advanced control systems are not able to achieve the desired performance [4].

The input design problem with respect to the intended model application, which is often a control task, has received considerable attention in the last two decades [5, 6].

It was reported that model development absorbs about 75% of the costs associated with advanced control projects [7]. System identification, in practice, is carried out by perturbing processes or plants under operation. In many industrial applications a plant friendly input signal would be preferred for system identification. Plant friendly identification experiments are those that satisfy plant or operator constraints on experiment duration, input and output amplitudes or input rate [8, 9]. Techniques for synthesising multi-harmonic signals with low crest factors, which are attractive from a plant friendly perspective, have been reported in [2]. It was demonstrated that plant friendliness demands are often in conflict with requirements for accurate and reliable identification [10]. Hence, plant friendly input design is inherently multi-objective in nature. There have been some reports on multi-objective optimisation based methods, applied to identification and control [11, 12].

However, the papers mentioned above present the optimisation methods of designing the parametric signals which meet the assumed friendliness criteria. In the experiments described in this paper we present different approach – a design of an optimal input signal via the optimisation procedure with respect to the cost function D-efficiency constraint has been attempted. In some our previous works, in the design of optimal and plant friendly inputs for system identification the sensitivity of the state variable to the unknown parameter has been maximised. The results of optimal and plant friendly input signal design utilising Mayer's canonical formulation of the performance index for the simple first-order inertial system case study were presented in [13, 14]. In this study we present the formulation of the performance index for optimal dynamic system identification and assess the qualitative measure of accuracy of parameter estimation in the second-order torsional spring system case study. In order to design an optimal actuation signal for the one degree of freedom torsional spring system parameter estimation, it was necessary to scale the model of the system. One of the problems to be solved was to find the value of the scaling factor between the angular position and the angular velocity of the real plant.

2 The D-efficiency constraint formulation

In the paper the design of optimal inputs for system identification with multiple unknown parameters is considered. In the design of optimal excitation signals for estimating more than one parameter, a suitable scalar function of the Fisher information matrix **M** must be selected as the performance criterion. The criterion, which is often used, is the trace of the matrix **M,** wherein the sum of diagonal elements of the Fisher information matrix is maximised. Other measures of identification performance are as follows [15]:

- A-optimality: $tr(\mathbf{M}^{-1})$, minimises the average variance of the parameters,
- E-optimality: $\lambda_{max}(\mathbf{M}^{-1})$, minimises the maximum eigenvalue of $\mathbf{M}^{-1}$,
- D-optimality: minimises the volume of the ellipsoidal confidence region of parameter estimates.

However, the choice of the experiment criterion is important, as it is possible that inputs obtained based on some criteria may not be persistently perturbing [1].

The input signal employed in the identification experiment should simultaneously yield two results: the acceptable accuracy of the system parameter estimates and the system should be perturbed in the least invasive (the most friendly) way. Such a compromise can be reached applying an approach, which relies on the notion of the D-efficiency [15]. Any optimality criterion can be associated with the efficiency function, defined as a measure of the relative performance of any given experiment **e** compared to that of the optimal experiment $\mathbf{e}^*$. The D-efficiency, which may be considered as a measure of the D-suboptimality of given input trajectories, is specified by

$$E_D(\mathbf{e}) = \left\{ \frac{\det(\mathbf{M}(\mathbf{e}))}{\det(\mathbf{M}(\mathbf{e}^*))} \right\}^{1/k}, \tag{1}$$

where k is the number of parameters to be identified, and $\mathbf{e}^*$ stands for the D-optimal trajectories which can be determined earlier. Following the reasoning and derivations presented in [16], we set a reasonable positive threshold $\eta < 1$ and impose the constraint on the D-efficiency value:

$$E_D(\mathbf{e}) \geq \eta. \quad . \tag{2}$$

Such an approach will yield a D-suboptimal, yet reasonable solution. The inequality (2) is equivalent to the constraint:

$$\Psi[\mathbf{M}(\mathbf{e})] \leq D, \tag{3}$$

where $\Psi[\mathbf{M}(\mathbf{e})] = \log(\det \mathbf{M}(\mathbf{e}))$ and $D = \Psi[\mathbf{M}(\mathbf{e}^*)] - k \log(\eta)$.

The objective of such an experiment is formulated through maximisation of the FIM determinant (D-optimality) with respect to D-efficiency inequality constraint (3).

The purpose of the current work is to formulate the optimisation problem for optimal input design with respect to the D-efficiency constraint. In that way (i.e. by setting such a constraint to control the level of the D-optimality loss) we can obtain the friendliest input signal, reducing the rapid changes of the mass or energy inflow to the system.

3 Optimal input design with respect to the cost function D-efficiency constraints

To illustrate the properties of the above approach to parameter identification, using the optimal input signal and with respect to the assumed level of D-optimality, we have selected the second-order linear dynamic system. The mass-spring damper system represents many similar physical plants. Thus the plant model order may be as high as six with either four, two, or no zeros.

The dynamic model for the one degree of freedom (1 DOF) plant is shown in below figure.

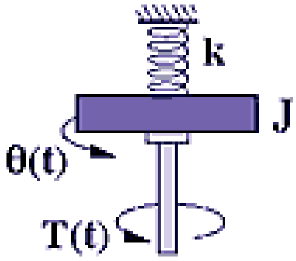

Fig. 1. The dynamic model for the one degree of freedom plant (free-clamped).

The equation of motion is as follows:

$$J_1\ddot{\theta} + c_1\dot{\theta}_1 + k_1\theta_1 = T(t), \tag{4}$$

where J_1, k_1, c_1, $T(t)$, θ_1 are, respectively: disk inertia, spring coefficient, damping ratio, force signal, position of the first disk of the plant. For notational convenience, let us introduce $x_1 = \theta_1$ and $x_2 = \dot{\theta}_1 = \dot{x}_1$. Then, the problem of synthesising an optimal input in the time domain for a torsional spring plant (1 DOF) can be described by the following single input, single output state space model:

$$\begin{aligned} &\dot{x}_1 = x_2; && x_1(0) = x_{10}; \\ &\dot{x}_2 = ax_1 + bx_2 + cu; && x_2(0) = 0; \\ &y(t) = x_1(t), && \end{aligned} \tag{5}$$

where $a = -k_1/J_1$, $b = -c_1/J_1$, $c = 1/J_1$. In order to design an optimal input signal, it was necessary to scale the model of the system as follows:

$$\begin{aligned} \xi_1 &= px_1 \rightarrow \dot{x}_1 = \frac{1}{p}\dot{\xi}_1; \\ \xi_2 &= gx_2 \rightarrow \dot{x}_2 = \frac{1}{g}\dot{\xi}_2. \end{aligned} \tag{6}$$

Utilising equation (6), the state space model (5) can be expressed as:

$$\begin{aligned} \dot{\xi}_1 &= \frac{p}{g}\xi_2; \\ \dot{\xi}_2 &= a\frac{g}{p}\xi_1 + b\xi_2 + cgu. \end{aligned} \quad . \tag{7}$$

Assuming that the parameter $p = 1$, from equation (6) we obtain $\xi_1 = x_1$. Then the above problem can be suitably modified by defining the state space model as:

$$\begin{aligned}
&\dot{x}_1 = \frac{1}{g} x_2;\\
&\dot{x}_2 = agx_1 + bx_2 + cgu;\\
&y(t) = x_1(t),
\end{aligned} \tag{8}$$

where $x_1 = x_1(t;\, a,\, b,\, c)$, $x_2 = x_2(t;\, a,\, b,\, c)$ and model parameters a, b, c are constant. The principle of the design of optimal input signals for system identification is to maximise the sensitivity of the state variable or the observation to the unknown parameter [3]. The justification for this approach is the Cramer-Rao lower bound, which provides a lower bound for the estimation error covariance. Providing this feature of the input, we obtain the parameter estimate or observation sensitivity which tends to be lowered for an optimal input

$$\operatorname{cov}([a,b,c]) \geq \mathbf{M}^{-1}. \tag{9}$$

The Fisher information matrix (FIM) for the torsional spring (1 DOF) model (8) can be expressed as:

$$\mathbf{M}(T) \cong \int_0^T \begin{bmatrix} x_{1a} \\ x_{1b} \\ x_{1c} \\ x_{2a} \\ x_{2b} \\ x_{2c} \end{bmatrix} \begin{bmatrix} x_{1a} & x_{1b} & x_{1c} & x_{2a} & x_{2b} & x_{2c} \end{bmatrix} dt, \tag{10}$$

where: $x_{ia} = \partial x_i/\partial a$, $x_{ib} = \partial x_i/\partial b$, $x_{ic} = \partial x_i/\partial c$, $i = 1, 2$.

Then the problem can be suitably modified by defining the augmented state equations as [3]:

$$\begin{array}{lll}
x_1 = x_1; & \dot{x}_1 = \dfrac{1}{g} x_2; & x_1(0) = x_{10};\\
x_2 = x_2; & \dot{x}_2 = agx_1 + bx_2 + cgu; & x_2(0) = 0;\\
x_3 = x_{1a}; & \dot{x}_3 = \dfrac{1}{g} x_6; & x_3(0) = 0;\\
x_4 = x_{1b}; & \dot{x}_4 = \dfrac{1}{g} x_7; & x_4(0) = 0;\\
x_5 = x_{1c}; & \dot{x}_5 = \dfrac{1}{g} x_8; & x_5(0) = 0;\\
x_6 = x_{2a}; & \dot{x}_6 = g(x_1 + ax_3) + bx_6; & x_6(0) = 0;\\
x_7 = x_{2b}; & \dot{x}_7 = g(ax_4) + x_2 + bx_7; & x_7(0) = 0;\\
x_8 = x_{2c}; & \dot{x}_8 = gax_5 + bx_8 + gu; & x_8(0) = 0.
\end{array} \tag{11}$$

An optimal input for exciting the torsional spring (1 DOF) system is formulated through maximisation of the Fisher information matrix determinant (D-optimality) in the form of a conventional integral-criterion optimal control problem. The problem of synthesising an optimal input signal for an inertial system, utilising Mayer's canonical formulation of the performance index, has been solved in literature [13, 14]

$$\mathbf{M}(t) = \int_0^t \begin{bmatrix} x_{1a} \\ x_{1b} \\ x_{1c} \\ x_{2a} \\ x_{2b} \\ x_{2c} \end{bmatrix} \begin{bmatrix} x_{1a} & x_{1b} & x_{1c} & x_{2a} & x_{2b} & x_{2c} \end{bmatrix} d\tau \ // \ \frac{d}{dt}, \tag{12}$$

the FIM can be modified as follows:

$$\dot{\mathbf{M}}(t) = \begin{bmatrix} x_{1a} \\ x_{1b} \\ x_{1c} \\ x_{2a} \\ x_{2b} \\ x_{2c} \end{bmatrix} \begin{bmatrix} x_{1a} & x_{1b} & x_{1c} & x_{2a} & x_{2b} & x_{2c} \end{bmatrix}, \ \mathbf{M}(0) = 0, \tag{13}$$

where

$$\mathbf{M}(t) = \begin{bmatrix} m_{11}(t) & \cdots & m_{16}(t) \\ \vdots & \ddots & \vdots \\ m_{61}(t) & \cdots & m_{66}(t) \end{bmatrix}, \tag{14}$$

and $m_{ij} = m_{ji}$.

Then the equivalent optimal control problem utilising Mayer's canonical formulation, which maximises the performance index with respect to the D-efficiency equality constraint, is

$$J = \det\left[\mathbf{M}(T_f)\right], \tag{15}$$

subject to:

$$\begin{gathered} \det\left[\mathbf{M}(T_f)\right] = D; \\ -5 \le u(t) \le 5, t \in [0, T], \end{gathered} \tag{16}$$

where D is D-efficiency constant.

4 Experimental results for torsional spring case study

For numerical solution of the above optimal control problems one of existing packages for solving dynamic optimisation tasks, such as Riots_95, Dircol or Miser, can be employed. We were employing the Riots_95 package [17], which is implemented in Matlab, has efficient tools for solving the constrained problems of dynamic optimisation and can be easily merged with other Matlab facilities (e.g. simulation of models developed in Simulink environment, with graphical user interface). The Matlab toolbox Riots solves a very large class of finite-time optimal controls problems that includes: trajectory and end-point constraints, variable initial conditions, free final time tasks and problems with cost functions endpoint. System dynamics can be integrated with fixed step-size Runge-Kutta method, a discrete-time solver or a variable step-size method. The software automatically computes gradients for all functions with respect to the controls and any free initial conditions.

The main program in Riots is based on thick sequential quadratic programming (SQP). Hence, the program is not well-suited for high discretisation levels. One of the major limitations of Riots is that it is not well-appointed to deal with problems whose dynamics are unstable.

All computations were performed using low-cost PC (Atom, 1.66 GHz, 1 GB RAM) running Windows 7 and Matlab 7.12 (R2011a). Optimal and sub-optimal signals are computed for nominal values of parameters a = -88.95, b = -0.42, c = 52.02 assumed termination time T_f = 10 seconds and the scaling factor $g = 10^5$ utilising SQP algorithm. The system is assumed to be at an initial state $x_1(0) = 0.393$, the initial value of the input signal is $u(0) = 1$ and $-5 \leq u(t) \leq 5$. The system dynamics was integrated using the fixed step-size fourth-order Runge-Kutta method with grid intervals of 0.2 seconds. The D-optimal input signals obtained for different desired values of the D-efficiency constant D (according to (3)) are shown in below figures. As the minimisation algorithms implemented in RIOTS_95 guarantee convergence only to a local minimum, we made use of the typical way to reduce the risk of "trapping", i.e. repeating the computations several times, starting from different initial conditions from the range [-0.4, 0.4] around the 'nominal' initial state $x_1(0) = 0.393$. The D-optimal excitation signal obtained when there was no constraint on the D-efficiency value (i.e. the coefficient $\eta = 1$ in the inequality (2)), is shown in black colour in Figure 1. It corresponds to the optimal experiment $\mathbf{e}^*$, where the maximal possible value of the FIM determinant is obtained and in such a case the value $D_{eff} = 100\% \cdot D_{eff,\,opt}$. The control signals obtained for decreasing values of D-efficiency from the interval [70%, 100%] of its maximum value are shown in Figures 2-5. As we can see, when the desired value of D-optimality decreases, the shape of the optimal input signal substantially changes. While for the optimal experiment (in the sense of (2)) there are the abrupt changes of the input, the control signal obtained for $D_{eff} < 70\% \cdot D_{eff,\,opt}$ is almost constant, so the FIM determinant component was dominated by D-efficiency constraint value of the maximised performance index.

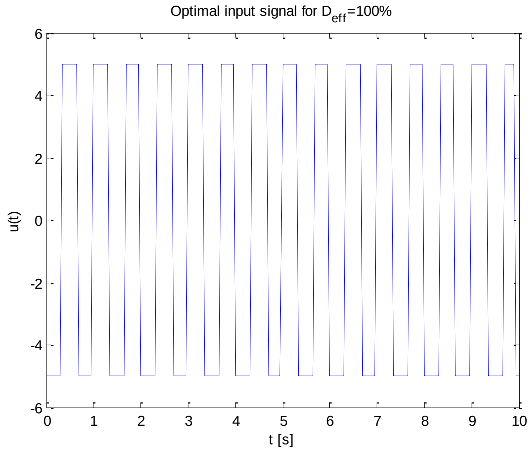

Fig. 2. Optimal input signal to the torsional spring system.

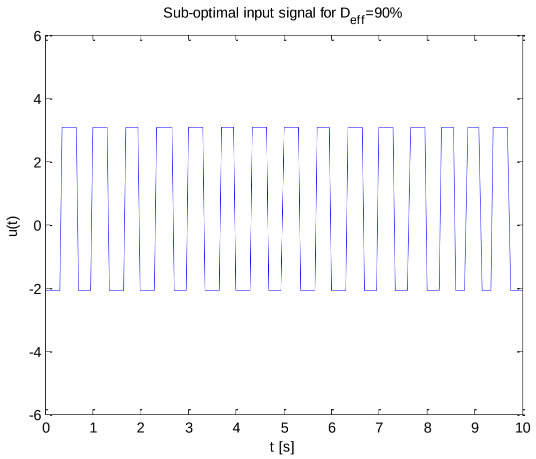

Fig. 3. Sub-optimal input signal to the torsional spring system for D_{eff} = 90%·$D_{eff, opt.}$

The optimal excitation signal obtained when there was no constraint on the D-efficiency component (i.e., for D_{eff} = 100% and $J = 4.14 \times 10^{32}$) is shown in Figure 2. The D-efficiency constraint value increased (Fig. 5) to obtain the critical value of the maximised performance index $J = 2.89 \times 10^{32}$ at the level of D_{eff} = 70%. For comparison, Figure 3 shows the sub-optimal input signal, which corresponds to the objective function value $J = 3.72 \times 10^{32}$ at the level of D_{eff} = 90%. Figure 4 contains the graphical display of the non-optimal signal obtained for $J = 3.10 \times 10^{32}$, where the FIM determinant component was maximised to the level of D_{eff} = 80%.

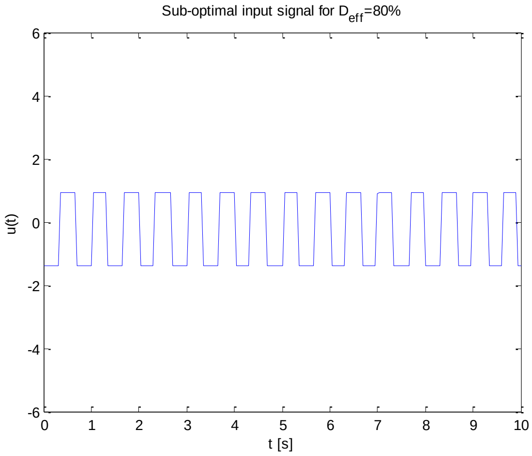

Fig. 4. Sub-optimal input signal to the torsional spring system for D_{eff} = 80%·$D_{eff,\ opt.}$

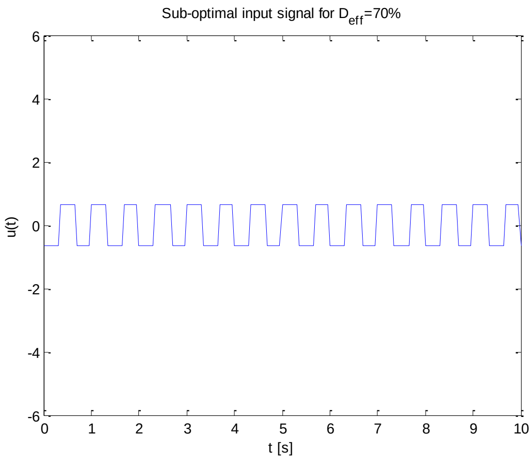

Fig. 5. Sub-optimal input signal to the torsional spring system for D_{eff} = 70%·$D_{eff,\ opt.}$

According to expectation, the constraint (16) was active in the optimal solution as shown in the above figures.

The signals $u(t)$ computed for different values of D-efficiency as solutions of the optimisation task (15) with the set of constraints (16), were then used as inputs in the parameter identification procedure. Figure 6 summarises the flow of information in the system identification process: we act on the physical system through the input $u(t)$ and collect information through the observations of its output $y(t)$. The presence of the white noise with different variance from the interval $0.0 \leq \sigma^2 \leq 0.7$ makes the observations random variables. The model corresponds to the theoretical representation of the system (8), which depends on a vector of unknown parameters $\boldsymbol{\theta} = [a, b, c]^T$. The objective of the system identification task is to find the best values of model parameters θ in terms of the performance criterion. The two hundred runs have been made for minimisation of the integral (within the time period from $t_0 = 0$ to the termination time $T_f = 10$ sec) of the squared difference between the output of the system and the output of the model. The initial state of the torsional spring (1 DOF) model was chosen from the interval $-0.4 \leq x_1(0) \leq 0.4$ [rad] and angular velocity from the interval $0 \leq x_2(0) \leq 4$ [rad/s]. The optimisation was performed using the Nelder-Mead method.

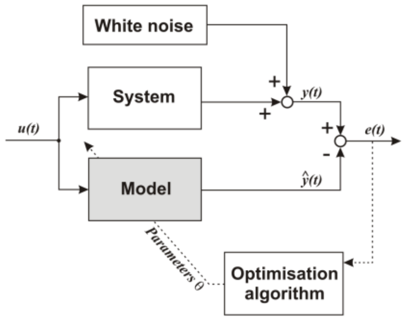

Fig. 6. Flow of information in the parameter identification system.

Figures 7(a)-7(d) show the result of the simulation experiments, i.e. the optimal values of parameters a and b computed as results of optimisation for each run, when the system starts from different initial state condition and the measurement noise influencing the system output has different variance. Figure 7(a) shows the results with the input signal obtained for the maximal value of FIM determinant (i.e. $D_{eff} = 100\% \cdot D_{eff,\,opt}$), the successive figures show the results (for the same combinations of initial states and noise variance) with the input signals were computed when a certain loss of D-optimality was assumed as ($D_{eff} / D_{eff,\,opt} = 90\%$, 80% and 70%, respectively). Analysis of the confidence regions of the torsional system parameter estimates confirms the following regularities. The optimal input signal, obtained for $D_{eff} / D_{eff,\,opt} = 100\%$, yields the minimal volume of the ellipsoidal confidence region of parameter estimates. When the desired ratio of $D_{eff} / D_{eff,\,opt}$ decreases (i.e. we accept bigger loss of D-optimality), the cluster occupied by the optimal values of identified model parameters increases its size – for the same initial conditions and noise characteristics as in the above experiment. Decreasing the desired ratio of $D_{eff} / D_{eff,\,opt}$ yields the input signal, which is more “friendly” for the plant, i.e. in that way we avoid abrupt changes of the control valve settings in the real-life identification experiments. The results of the simulation experiments for other combinations of the torsional system parameters *a*, *b* and *c* are very similar to those shown in Figure 7. The purpose of this case study was to show that the requirements of high friendliness of the input signal and the accuracy of parameter estimation are, in some sense, opposite.

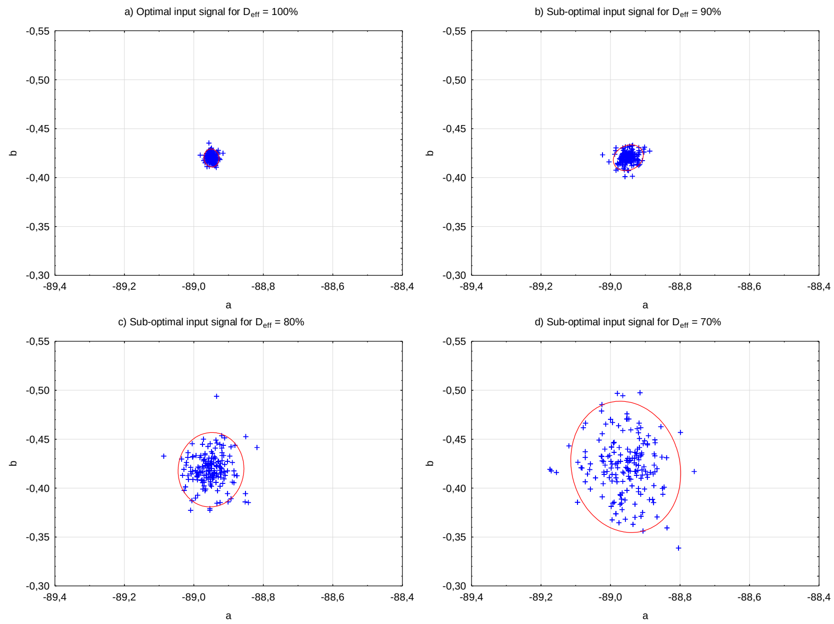

Fig. 7. Confidence regions of the torsional spring model parameter estimates; the model was excited utilising: (a) Optimal input signal ($D_{eff} = D_{eff,\ opt}$). (b) Sub-optimal input signal ($D_{eff} = 90\% \cdot D_{eff,\ opt}$). (c) Sub-optimal input signal ($D_{eff} = 80\% \cdot D_{eff,\ opt}$). (d) Sub-optimal input signal ($D_{eff} = 70\% \cdot D_{eff,\ opt}$).

5 Conclusions

An optimal input signal design problem for system identification was formulated in the paper and the method of the problem solution was outlined. In the presented approach the input signal is a solution of a dynamic optimisation problem, where the FIM determinant is maximised, at the same time providing a guaranteed level of D-efficiency. The second order dynamic system case study of optimal input signal design with guaranteed D-efficiency was presented in the paper. The experiments confirm that we can provide a compromise between the friendliness of the plant excitation in the identification process and the accuracy of estimates (observed as a reasonably small volume of the confidence ellipsoid) for a wide range of measurement noise at the system output. One of the most important steps in the approach presented in the paper was the transcription of the proposed problem formulation into an equivalent optimal control task expressed in the Lagrange form with the appropriate set of constraints. The optimal input trajectories were then computed using one of existing packages for solving dynamic optimisation problems. An optimal input signal design for solving free final time parameter estimation problem in the time domain will be presented.

Acknowledgement

The work was supported by the Dean's Project No. S/WI/1/13, Faculty of Computer Science, Department of Digital Media and Computer Graphics, Bialystok University of Technology.

References

1. Ljung, L.: System identification: Theory for the user. Prentice Hall, USA (1999)
2. Pintelon, R., Schoukens, J.: System identification: A frequency domain approach. John Wiley & Sons, USA (2001)
3. Kalaba, R., Spingarn, K.: Control, identification, and input optimization. Plenum Press, New York, USA (1982)
4. Hugo, A.J.: Process controller performance monitoring and assessment. Control. Arts Inc., http://www.controlarts.com
5. Hildebrand, R., Gevers, M.: Identification for control: Optimal input design with respect to worst-case v-gap cost function. SIAM Journal on Control Optimization, 41, 1586-1608 (2003)
6. Pronzato, L.: Optimal experimental design and some related control problems. Automatica, 44, 303-325 (2008)
7. Hussain, M.: Review of the applications of neural networks in chemical process control-simulation and on-line implementation. Artificial Intelligence in Engineering, 13, 55-68 (1999)
8. Narasimhan, S., Rengaswamy, R.: Multi-objective input signal design for plant friendly identification of process systems. Proceeding of the American Control Conference 2006, pp. 4891-4896. Boston, Massachusetts (2004)
9. Narasimhan, S., Rengaswamy, R.: Multi-objective optimal input design for plant friendly identification. Proceeding of the American Control Conference 2008, pp. 1304-1309. Seattle, Washington (2008)
10. Rivera, D., Lee, H., Braun, M., Mittelmann, H.: Plant friendly system identification: A challenge for the process industries. In SYSID 2003; Rotterdam, Netherlands (2003)
11. Steenis, R., Rivera, D.: Plant-Friendly Signal Generation for System Identification Using a Modified Simultaneous Perturbation Stochastic Approximation (SPSA) Methodology. IEEE Transactions on Control Systems Technology, 19, 1604-1612 (2011)
12. El-Kady, M., Salim, M., El-Sagheer, A.: Numerical treatment of multiobjective optimal control problems. Automatica, 39, 47-55 (2003)
13. Jakowluk, W.: Design of an optimal input signal for plant-friendly identification of inertial systems. Przegląd Elektrotechniczny, 85(6), 125-129 (2009)
14. Jakowluk, W.: Plant friendly input design for parameter estimation in an inertial system with respect to D-efficiency constraints. Entropy, 16(11), 5822-5837 (2014)
15. Atkinson, A., Donev, A., Tobias, R.: Optimum experimental design with SAS. Oxford University Press, Oxford (2007)
16. Uciński, D., Chen, Y.Q.: Sensor Motion Planning in Distributed Parameter Systems Using Turing's Measure of Conditioning. Proceedings of the 45th IEEE Conference on Decision and Control 2006, San Diego, CA. WeB03.6, pp. 759-764. CD-ROM (2006)
17. Schwartz, A.. Polak, E., Chen, Y.: A Matlab toolbox for solving optimal control problems. Version 1.0 for Windows, May (1997), http://www.schwartz-home.com/~adam/RIOTS/