Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 58-15 (2024) 550–555

Experiment Design Taking Nonasymptotic Properties of the Model into Consideration ⋆

Masanori Oshima * **Sanghong Kim** ** **Yuri A. W. Shardt** *** **Ken-Ichiro Sotowa** *

* *Department of Chemical Engineering, Kyoto University, Kyoto 615-8510, Japan (e-mail: oshima.masanori.66a@st.kyoto-u.ac.jp, sotowa@cheme.kyoto-u.ac.jp).*
** *Department of Applied Physics and Chemical Engineering, Tokyo University of Agriculture and Technology, Tokyo 184-8588, Japan, (e-mail: sanghong@go.tuat.ac.jp)*
*** *Department of Automation Engineering, Technical University of Ilmenau, Ilmenau D-98684, Germany, (e-mail: yuri.shardt@tu-ilmenau.de)*

Abstract: Experiment design for identification is usually based on asymptotic theory, where an infinite number of samples is assumed. However, such an assumption does not hold in practical cases, and hence, the nonasymptotic properties of system identification should be considered. This paper proposes a new method for experiment design for identification based on the nonasymptotic confidence region of the system parameters calculated using the sign-perturbed-sums (SPS) method for multivariate autoregressive exogenous input (ARX) systems. The objective function based on the volume of the confidence region is introduced in the proposed method. Moreover, the proposed optimization problem is solved using Bayesian optimization because the proposed objective function is calculated from the data obtained only after the experiment. The validity of the proposed method was assessed in an experimental case study of a three-tank system, where the proposed method was compared with the existing D-optimal method. As a result, the model obtained using the proposed method reduced the mean squared control error of model predictive control by 22.9% from that of the existing method.

Keywords: System identification, experiment design, confidence region, finite-sample data, Bayesian optimization, sign-perturbed sums

1. INTRODUCTION

In chemical plants, model predictive control (MPC) (Morari and H. Lee, 1999) is sometimes used for efficiently operating the plant. However, it is often difficult to maintain the accuracy of the prediction model in the MPC for a long period because of the gradual change of the dynamics. The prediction model is usually a linear, time-invariant model in chemical plants. However, chemical plants generally exhibit time-varying dynamics, and the change of the dynamics over time is too slow to be considered during model construction. Hence, model update is necessary to maintain the accuracy of the prediction model.

Model building, including model updates, is often performed using an open-loop step test in chemical plants. The step test can be effectively used to build an initial model for the MPC, where there is little knowledge about the dynamics of the plant of interest. The step test improves the understanding of the plant by visualizing the plant dynamics.

However, the step test requires a huge amount of time and effort. In the open-loop step test, step signals are applied to each input of the targeted plant in turn. The large number of manipulated variables results in a significant increase in the number of step inputs. Moreover, after applying a step signal, it is necessary to wait until the system has reached the new steady state. Since chemical plants have slow dynamics whose time constant can be several days, the settling time tends to be significant. For the above reasons, the step test is unsuitable for updating models.

Experiment design for identification (Goodwin, 1977) can avoid such a problem. It determines the experimental condition, such as the amplitudes and frequencies of the input signals, solely based on maximization of the data-quality index. Thus, in the experiment design for identification, the multiple inputs can be varied simultaneously so that it is unnecessary to wait for the system to reach a new steady state. Moreover, experiment design for identification is an optimization-based procedure that can be easily automated to reduce the effort for model maintenance.

⋆ This work was supported by JST SPRING, Grant Number JPMJSP2110 and JSPS KAKENHI, Grant Number JP22K04816. In addition, M. O. was supported by a graduate exchange fellowship from JGP, Kyoto University.

10.1016/j.ifacol.2024.08.587

The data-quality index used in experiment design for identification is usually scalar measures of the Fisher information matrix (FIM), such as the determinant (D-optimal design) and the trace (A-optimal design) (Goodwin, 1977; Shardt, 2022). FIM corresponds to the inverse of the covariance matrix of the parameters derived from asymptotic theory, where the number of data samples is assumed to be infinite (Ljung, 1998). However, the assumption of the asymptotic theory does not generally hold, and the deviation of the actual situation from the situation assumed in the asymptotic theory increases as the available number of data samples decreases. This often occurs in systems where the experimental costs are high, such as chemical plants. Therefore, a new data-quality index that exactly quantifies the quality of the finite-sample data is desirable.

Recently, the sign-perturbed-sums (SPS) method was developed by Csáji et al. (2012). The SPS method provides the exact confidence region of the process parameters from the finite-sample data. The SPS method does not rely on the assumption that restricts the distribution of the noise innovation to a particular type, such as Gaussian distributions. It only assumes the noise innovation has a symmetric distribution. Moreover, the applicability of the SPS method was guaranteed to general linear systems (Csáji et al., 2012), closed-loop systems (Csáji and Weyer, 2015), and multivariate systems (Szentpéteri and Csáji, 2023). Hence, the data-quality index based on SPS has a wide range of applicability to practical problems.

Kolumbán and Csáji (2018) proposed the experiment design for identification based on SPS, where they minimized the expected volume of the confidence region. However, the targeted system was restricted to single-input, single-output finite-impulse-response (FIR) systems, whose regressors are composed of the past sequence of the input. Hence, it cannot be applied to multi-input, multi-output systems with regressors composed of the past sequence of the inputs and outputs.

This paper proposes a new experiment design for identification based on the SPS method for multivariate autoregressive exogenous input (ARX) systems. In the proposed method, the performance index of the data quality is based on the volume of the confidence region of SPS, and the constraints of the frequencies included in the input signals are imposed. Moreover, Bayesian optimization (BO) is used to solve the optimization problem for experiment design since the performance index is evaluated using the outcome of the identification experiment. The validity of the proposed method is examined through an experimental case study using the three-tank system (TTS).

2. BACKGROUND

2.1 Target system

The target system is an M-input, N-output ARX system. The outputs, inputs, and noise innovations at time index t are $\boldsymbol{y}_t = [y_{1,t}, \cdots, y_{N,t}]^\top$, $\boldsymbol{u}_t = [u_{1,t}, \cdots, u_{M,t}]^\top$, and $\boldsymbol{e}_t = [e_{1,t}, \cdots, e_{N,t}]^\top$. The process of interest is described as

$$\sum_{k=0}^{K} \boldsymbol{A}_k \boldsymbol{y}_{t-k} = \sum_{l=1}^{L} \boldsymbol{B}_l \boldsymbol{u}_{t-l} + \boldsymbol{e}_t, \tag{1}$$

where K and L are the maximum time delays of the outputs and inputs of the process, $\boldsymbol{A}_k \in \mathbb{R}^{N\times N}$ and $\boldsymbol{B}_l \in \mathbb{R}^{N\times M}$ are the process parameters, and $\boldsymbol{A}_0$ is the identity matrix.

Let $\underline{\tau}_{n,m} \in \{1, \cdots, L\}$ and $\overline{\tau}_{n,m} \in \{\underline{\tau}_{n,m}, \cdots, L\}$ be respectively the minimum and maximum time delays between $y_{n,t}$ and $u_{m,t}$. Then, the number d of the process parameters is

$$d = KN^2 + \sum_{n=1}^{N}\sum_{m=1}^{M} (\overline{\tau}_{n,m} - \underline{\tau}_{n,m} + 1), \tag{2}$$

and Eq. (1) is transformed into

$$\boldsymbol{y}_t = \boldsymbol{\Phi}_t^\top \boldsymbol{\theta}^* + \boldsymbol{e}_t, \tag{3}$$

$$\boldsymbol{\Phi}_t = \begin{bmatrix} \boldsymbol{\phi}_{1,t} & & 0 \\ & \ddots & \\ 0 & & \boldsymbol{\phi}_{N,t} \end{bmatrix}, \tag{4}$$

$$\begin{aligned} \boldsymbol{\phi}_{n,t} =& [-y_{1,t-1}, \cdots, -y_{1,t-K}, \cdots, \\ & - y_{N,t-1}, \cdots, -y_{N,t-K}, \\ & u_{1,t-\underline{\tau}_{n,1}}, \cdots, u_{1,t-\overline{\tau}_{n,1}}, \cdots, \\ & u_{M,t-\underline{\tau}_{n,M}}, \cdots, u_{M,t-\overline{\tau}_{n,M}}]^\top, \end{aligned} \tag{5}$$

where $\boldsymbol{\theta}^* \in \mathbb{R}^d$ and $\boldsymbol{\Phi}_t \in \mathbb{R}^{d\times N}$ are the true parameter vector and the regressor matrix, respectively.

2.2 Sign-perturbed sums

The SPS method gives the exact confidence region $\Theta_{\rm SPS}$ of $\boldsymbol{\theta}^*$ for a given confidence probability from a data set with T samples defined as $\mathbb{D}_0 = \{\boldsymbol{y}_t, \boldsymbol{u}_t \mid t = 1, \cdots, T\}$. In the SPS method, the following assumptions are made:

A.1 $\{e_{n,t}\}$ are independent, and the probability density function (PDF) of each $e_{n,t}$ is symmetric about 0.
A.2 The other external signals are independent of $\{e_{n,t}\}$.
A.3 The system has the ARX structure given by Eqs. (3) to (5).

Note that assumption A.1 holds for various PDFs such as the Gaussian, the Laplacian, and the uniform distributions. The detailed procedure of the SPS method is

M.1 Determine $R \in \{2, 3, \cdots\}$ and $\tilde{R} \in \{1, 2, \cdots, R-1\}$, then the confidence probability is $p = 1 - \tilde{R}/R$.
M.2 Generate $(R-1)$ time-series of random-sign matrices $\{\boldsymbol{\Xi}_{1,t}\}_{t=1}^T, \cdots, \{\boldsymbol{\Xi}_{R-1,t}\}_{t=1}^T$. The random-sign matrix $\boldsymbol{\Xi}_{r,t} \in \mathbb{R}^{N\times N}$ is a diagonal matrix, whose diagonal components take the values $+1$ and -1 with a probability of 0.5 each.
M.3 Determine $\Theta_{\rm SPS}$ according to the following equation:

$$\Theta_{\rm SPS} = \left\{\boldsymbol{\theta} \in \mathbb{R}^d \mid \mathrm{Rank}(\boldsymbol{\theta}) \le R - \tilde{R}\right\}, \tag{6}$$

where $\mathrm{Rank}(\boldsymbol{\theta})$ is defined as follows:

M.3.1 Calculate $\{\boldsymbol{\epsilon}_{0,t}(\boldsymbol{\theta})\}_{t=1}^T$ defined as

$$\boldsymbol{\epsilon}_{0,t}(\boldsymbol{\theta}) = \boldsymbol{y}_t - \boldsymbol{\Phi}_t^\top \boldsymbol{\theta} \in \mathbb{R}^N. \tag{7}$$

M.3.2 Calculate $\{\boldsymbol{\epsilon}_{1,t}(\boldsymbol{\theta})\}_{t=1}^T, \cdots, \{\boldsymbol{\epsilon}_{R-1,t}(\boldsymbol{\theta})\}_{t=1}^T$ as

$$\boldsymbol{\epsilon}_{r,t}(\boldsymbol{\theta}) = \boldsymbol{\Xi}_{r,t} \boldsymbol{\epsilon}_{0,t}(\boldsymbol{\theta}). \tag{8}$$

M.3.3 Aquire the input-output data $\mathbb{D}_r(\boldsymbol{\theta})$ from the simulation of the system with $\boldsymbol{\theta}$ using $\boldsymbol{\epsilon}_{r,t}(\boldsymbol{\theta})$ as a noise innovation vector, for $\forall r \in \{1, 2, \cdots, R-1\}$. Note that the initial condition of the system for the simulation is the same as when the original data $\mathbb{D}_0$ is obtained.

M.3.4 Prepare R regressor matrices $\{\boldsymbol{\Phi}_{\rho,t}\}_{\rho=0}^{R-1}$, where $\boldsymbol{\Phi}_{0,t} = \boldsymbol{\Phi}_t$, and $\boldsymbol{\Phi}_{r,t}\,(r > 0)$ is composed of the data $\mathbb{D}_r(\boldsymbol{\theta})$ in the same way as $\boldsymbol{\Phi}_t$.

M.3.5 Calculate $\{\boldsymbol{S}_\rho(\boldsymbol{\theta})\}_{\rho=0}^{R-1}$ defined as follows:

$$\boldsymbol{S}_\rho(\boldsymbol{\theta}) = \left(\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{\Phi}_{\rho,t}\boldsymbol{\Phi}_{\rho,t}^\top\right)^{-\frac{1}{2}} \cdot \left(\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{\Phi}_{\rho,t}\boldsymbol{\epsilon}_\rho(t,\boldsymbol{\theta})\right). \quad (9)$$

M.3.6 Arrange $\{\|\boldsymbol{S}_\rho(\boldsymbol{\theta})\|\}_{\rho=0}^{R-1}$ from smallest to largest, and let $||\boldsymbol{S}_{\rho_i}(\boldsymbol{\theta})||$ be the i-th smallest.

M.3.7 Define $\mathrm{Rank}(\boldsymbol{\theta})$ as

$$\mathrm{Rank}(\boldsymbol{\theta}) = i, \text{ if } \|\boldsymbol{S}_0(\boldsymbol{\theta})\| = \|\boldsymbol{S}_{\rho_i}(\boldsymbol{\theta})\|. \quad (10)$$

2.3 How to calculate the volume of the confidence region using the SPS method

Let the indicator function $I_{\mathrm{SPS}} : \mathbb{R}^d \to \{0, 1\}$ for SPS be

$$I_{\mathrm{SPS}}(\boldsymbol{\theta}) = \begin{cases} 1, & \mathrm{Rank}(\boldsymbol{\theta}) \le R - \tilde{R}, \\ 0, & \mathrm{Rank}(\boldsymbol{\theta}) > R - \tilde{R}. \end{cases} \quad (11)$$

Then, from Eq. (6), the volume V_{SPS} of Θ_{SPS} is

$$V_{\mathrm{SPS}} = \int_{\boldsymbol{\theta}\in\mathbb{R}^d} I_{\mathrm{SPS}}(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}. \quad (12)$$

If the system of interest has a regressor uncorrelated with the noise innovation as in a FIR system, the integral of Eq. (12) is efficiently calculated using the fact that Θ_{SPS} is a star-convex set whose star center is the least-squares estimate $\hat{\boldsymbol{\theta}}_{\mathrm{LS}}$ (Csáji et al., 2015). Otherwise, it is difficult to calculate the integral of Eq. (12) since it is not even guaranteed that Θ_{SPS} is a connected set.

In the case study shown below, the integral is approximated using the following approach based on the Monte Carlo integration (MCI) (Gunther and Friedrich, 2014). Using MCI, Eq. (12) is approximated as

$$V_{\mathrm{SPS}} \simeq \frac{V_{\mathrm{I}}}{I}\sum_{i=1}^{I} I_{\mathrm{SPS}}(\boldsymbol{\theta}_i), \quad (13)$$

where V_{I} is the volume of the integral domain $\Theta_{\mathrm{I}} \subset \mathbb{R}^d$, and $\{\boldsymbol{\theta}_i\}_{i=1}^{I}$ are points randomly sampled in Θ_{I}. Θ_{I} is determined so that it includes $\hat{\boldsymbol{\theta}}_{\mathrm{LS}}$, which is always included in Θ_{SPS} (Csáji et al., 2012). In the calculation, enlargement of Θ_{I} and random sampling of $\{\boldsymbol{\theta}_i\}_{i=1}^{I}$ are repeated until the increase in the number of points where I_{SPS} returns 1 gets smaller than a sufficiently small tolerance.

3. EXPERIMENT DESIGN FOR IDENTIFICATION

3.1 Existing objective functions

In existing methods for experiment design for identification, scalar measures, such as the determinant and trace, of FIM are maximized (Goodwin, 1977; Shardt, 2022). The definition of FIM for ARX systems is

$$\boldsymbol{F}_T = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{\Phi}_t\boldsymbol{\Phi}_t^\top. \quad (14)$$

When the determinant of FIM is used as the maximized objective function, the experiment design is called a D-optimal design. A typical optimization problem solved in the D-optimal design is

$$\mathrm{P1}: \min_{\boldsymbol{\chi}} J_{\mathrm{Dopt}} = -\log(\det \boldsymbol{F}_T), \quad (15)$$

where $\boldsymbol{\chi}$ is the vector of the design variables, and *log* is the natural logarithm. The optimization problem of Eq. (15) can be solved before a data-acquisition experiment when $\boldsymbol{\Phi}_t$ is independent of $\boldsymbol{e}_t$. Otherwise, the evaluation of J_{Dopt} requires the outcome of the data-acquisition experiment.

The determinant of FIM is related to the volume V_{As} of the confidence region obtained by asymptotic theory since it is given by

$$V_{\mathrm{As}} \propto [\det(\boldsymbol{F}_T)]^{-\frac{1}{2}}\, T^{-\frac{d}{2}}. \quad (16)$$

Therefore, the D-optimal design of Eq. (15) finds an optimal experimental condition $\boldsymbol{\chi}$ based on asymptotic theory.

3.2 Proposed objective function

In this paper, a new method based on the SPS method is proposed. V_{SPS} is used for calculating the objective function J_{SPS}. The proposed objective function is defined based on the relationship between J_{Dopt} and V_{As}. From Eq. (16), $\det(\boldsymbol{F}_T)$ is proportional to $(V_{\mathrm{As}}T^{d/2})^{-2}$. Hence, the D-optimal objective function is expressed using V_{As} as

$$J_{\mathrm{Dopt}} = 2\log(V_{\mathrm{As}}T^{d/2}) + C_{\mathrm{Dopt}}, \quad (17)$$

where $C_{\mathrm{Dopt}} \in \mathbb{R}$ is a constant. Based on Eq. (17), the proposed optimization problem is

$$\mathrm{P2}: \min_{\boldsymbol{\chi}} J_{\mathrm{SPS}} = 2\log(V_{\mathrm{SPS}}T^{d/2}). \quad (18)$$

V_{SPS} always depends on a realization of the noise innovation from Eq. (7) in step M.3.1 regardless of the dependency of $\boldsymbol{\Phi}_t$ on $\boldsymbol{e}_t$. Hence, the evaluation of the objective function related to V_{SPS} always requires an outcome of a data-acquisition experiment.

3.3 Designing identification experiment using BO

The outcome of a data-acquisition experiment is often required for evaluating the objective function of the design problems as mentioned above. In fact, it is mandatory in P2. To overcome this, the use of BO is proposed in this paper. BO is suitable for experiment design problems whose objective function is expensive to evaluate (Shahriari et al., 2016; Greenhill et al., 2020).

The BO algorithm does not require the equational relationship between the objective function and optimized variables before solving the optimization problem. Hence, the objective function whose value is determined only after the experiment as in our situation can be used in the BO algorithm.

Let the objective function of the experiment design for identification be J, which is either J_{Dopt} or J_{SPS}, and $\boldsymbol{\chi}_i$

Table 1. Definition of the input signals for checking the dependency of $V_{\mathrm{SPS}}T^{d/2}$ on T. The filter $F_{\mathrm{BW}}(q|n,\underline{\omega},\overline{\omega})$ is the n-th order Butterworth filter with upper and lower cut-off frequencies of $\overline{\omega}$ and $\underline{\omega}$. In addition, $\{v_t\}$ is white noise, and $v_t \sim \mathcal{N}(0,10)$.

Type of input	Symbol	Mathematical definition
2nd-order-PE input	S1	$u(t) = 10\sin(2\pi t/15)$
4th-order-PE input	S2	$u(t) = \frac{10(u^*(t)-\sum_{t=1}^{T} u^*(t)/T)}{(\max\{u^*(t)\}-\min\{u^*(t)\})/2}$ $u^*(t) = \sin(2\pi t/20) + \sin(2\pi t/10 + \pi/2)$
High-order-PE input with high-pass filter	S3	$u(t) = F_{\mathrm{BW}}(q|2, 0.05, 1)v(t)$
White-noise input	S4	$u(t) = v(t)$
High-order-PE input with bandpass filter	S5	$u(t) = F_{\mathrm{BW}}(q|2, 0.05, 0.1)v(t)$
High-order-PE input with low-pass filter	S6	$u(t) = \frac{1}{1-0.95q^{-1}}v(t)$

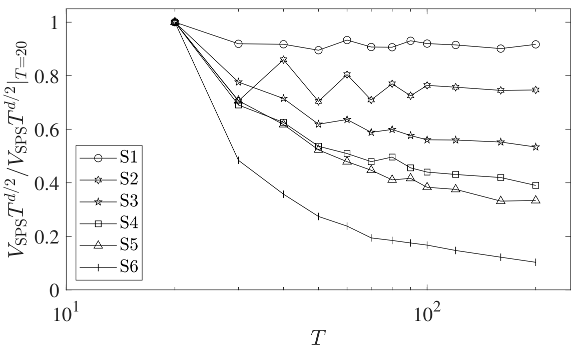

Fig. 1. The plots of $V_{\mathrm{SPS}}T^{d/2}$ as a function of the number of samples T.

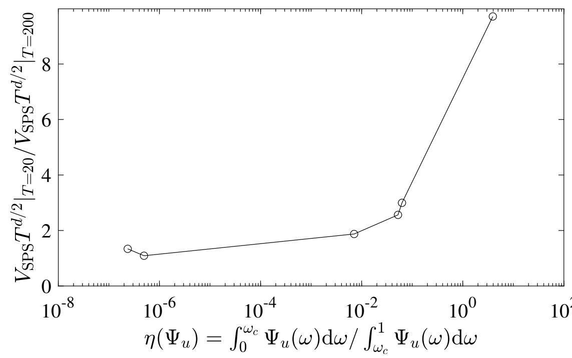

Fig. 2. The plot of the decrease in $V_{\mathrm{SPS}}T^{d/2}$ against the increase of T as a function of the power ratio $\eta(\Psi_u)$.

and J_i be the values of $\boldsymbol{\chi}$ and J in the i-th iteration of BO. Then, the following procedure is repeated N_{BO} times after the evaluation of N_{rand} sets of J for the randomly-sampled $\boldsymbol{\chi}$:

- B.1 Modelling the relationship between $\boldsymbol{\chi}$ and J is performed using Gaussian process (GP) regression based on the data of $\boldsymbol{\chi}$ and J stored up to the current i-th iteration.
- B.2 The acquisition function, whose maximum indicates the most promising design variables based on the data obtained up to the current iteration, is derived from the GP model.
- B.3 Maximizing the acquisition function, the next point $\boldsymbol{\chi}_i$ of $\boldsymbol{\chi}$ is suggested.
- B.4 Using the excitation signal defined by $\boldsymbol{\chi}_i$, a data-acquisition experiment is performed, and the value J_i of J is evaluated. Note that the random-sign matrices $\{\boldsymbol{\Xi}_{r,t}\}$ are regenerated at every iteration in the case of P2.
- B.5 $\boldsymbol{\chi}_i$ and J_i are stored.

Let T_{BO} be the number of samples obtained in step B.4 in each iteration. Then, $(N_{\mathrm{rand}} + N_{\mathrm{BO}})T_{\mathrm{BO}}$ samples of data are obtained up to the final iteration of BO and used for system identification. Note that the design variables determined in B.3 of the i-th iteration of BO depend on the noise innovations observed in B.4 of the $(i-i')$-th iterations of BO for $\forall i' \in \{1, 2, \cdots, i-1\}$. Hence, the excitation signals defined using $\boldsymbol{\chi}_i$ depend on the noise innovations observed in the $(i-i')$-th iterations, as well. This dependency may make system identification using the final data difficult. However, the excitation signals and noise innovations are still independent in each iteration of BO, that is, the assumption A.2 of the SPS method holds in step B.4 for case P2.

3.4 Constraint to reduce the dependency of J_{SPS} on the number of samples

The objective function J of the experiment design for identification using BO is preferred to be independent of T_{BO}. Using J independent of T_{BO} allows us to decrease T_{BO} and increase $(N_{\mathrm{rand}}+N_{\mathrm{BO}})$ to improve the optimality of the solution while keeping $(N_{\mathrm{rand}}+N_{\mathrm{BO}})T_{\mathrm{BO}}$ constant. Hence, the dependency of J on T_{BO} should be mitigated.

To achieve this, the dependency of J_{SPS} on T was examined through a numerical case study explained below. The system of interest is a SISO ARX system, that is,

$$y_t = \frac{0.5q^{-1}}{1-0.9q^{-1}}u_t + \frac{1}{1-0.9q^{-1}}e_t,\ e_t \sim \mathcal{N}(0, 0.1). \quad (19)$$

The system is operated in open-loop to obtain the input-output data used for calculating $V_{\mathrm{SPS}}T^{d/2}$, which is the main component of J_{SPS}. The types of input signals used as u_t are summarised in Table 1. The number of samples T is changed to a value in $\mathbb{T} = \{20,\ 30,\ \cdots,\ 100,\ 120,\ 160,\ 200\}$. The parameters R and $\tilde{R}$ of SPS were set to 800 and 20. The mean values of $V_{\mathrm{SPS}}T^{d/2}$ for each input are evaluated by the 200-times calculation of the objective function.

The dependency of mean of $V_{\mathrm{SPS}}T^{d/2}$ on T is shown in Fig. 1. As T increases, the mean values of $V_{\mathrm{SPS}}T^{d/2}$ tend to decrease, and the degree of decrease depends on the type of input. To quantify this, the ratio η of the input power in the frequency band located below ω_{c} to that in the frequency band located above ω_{c} is introduced, that is,

$$\eta(\Psi_u) = \frac{\int_0^{\omega_{\mathrm{c}}} \Psi_u(\omega)\mathrm{d}\omega}{\int_{\omega_{\mathrm{c}}}^1 \Psi_u(\omega)\mathrm{d}\omega}, \tag{20}$$

where Ψ_u is the power spectrum of u, and $\omega_c = 1/20$. Note that the denominator of ω_c is the minimum value of $\mathbb{T}$. Then, the slope of $V_{\mathrm{SPS}}T^{d/2}$ between $T = 20$ and $T = 200$ is plotted as a function of η as shown in Fig. 2. The slope of $V_{\mathrm{SPS}}T^{d/2}$ increases as the power ratio tends to increase. This means that the degree of dependency of $V_{\mathrm{SPS}}T^{d/2}$ on T is increased by the power in the frequency band located below ω_{c} and decreased by the power in the frequency band located above ω_{c}.

From the discussion above, the upper limit on η effectively reduces the dependency of $V_{\mathrm{SPS}}T^{d/2}$ on T. Hence, the constraint,

$$\eta(\Psi_{u_m}) < \epsilon, \text{ for } m = 1, \cdots, M, \tag{21}$$

where ω_{c} for calculating η is $1/T_{\mathrm{BO}}$, are introduced into the problems P1 and P2 when using BO.

4. EXPERIMENTAL CASE STUDY

4.1 Target system

The targeted system is the TTS as shown in Fig. 3. The TTS is comprised of three tanks from T1 to T3 and two pumps, P1 and P2. The output variables are the liquid levels $y_{1,t}$, $y_{2,t}$, and $y_{3,t}$ of T1 to T3, and the input variables are the pump voltages $u_{1,t}$ and $u_{2,t}$ of P1 and P2. The system is operated in open-loop during the identification experiment.

4.2 Methods

In this experimental case study, the proposed method for experiment design for identification was compared with the existing D-optimal design. Both methods assumed the first-order ARX model structure, that is,

$$\begin{bmatrix} 1+a_1q^{-1} & & 0 \\ & 1+a_2q^{-1} & \\ 0 & & 1+a_3q^{-1} \end{bmatrix} (\boldsymbol{y}_t - \overline{\boldsymbol{y}}) = \begin{bmatrix} b_{1,1}q^{-1} & b_{1,2}q^{-2} \\ b_{2,1}q^{-1} & b_{2,2}q^{-1} \\ b_{3,1}q^{-1} & b_{3,2}q^{-1} \end{bmatrix} (\boldsymbol{u}_t - \overline{\boldsymbol{u}}) + \boldsymbol{e}_t, \tag{22}$$

where $(\boldsymbol{u}_t - \bar{\boldsymbol{u}})$ and $(\boldsymbol{y}_t - \bar{\boldsymbol{y}})$ are the centered inputs and outputs, and sampling interval is 20.4 s. Hence, the objective function of the D-optimal design also depends on the realization of the noise innovation. Thus, the optimization problems of both methods were solved using BO. Note that the type of acquisition function used in BO was the expected-improvement function (Mockus, 1998).

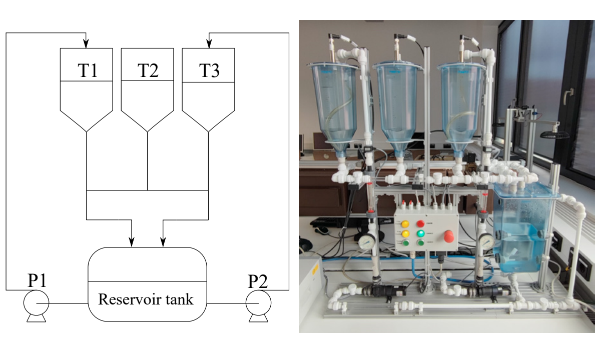

Fig. 3. (left) Process flow diagram and (right) picture of the three-tank system.

The input signals $\{u_{1,t}\}$ and $\{u_{2,t}\}$ for both methods were given by

$$u_{1,t} = \frac{0.3}{\max\{\overline{\omega} - \underline{\omega}, 0.1\}} F_{\mathrm{BW}}(q|5, \underline{\omega}, \overline{\omega}) v_{1,t} + u_{\mathrm{ss},1}, \tag{23}$$

$$u_{2,t} = \frac{0.3}{\max\{\overline{\lambda} - \underline{\lambda}, 0.1\}} F_{\mathrm{BW}}(q|5, \underline{\omega}, \overline{\omega}) v_{2,t} + u_{\mathrm{ss},2}, \tag{24}$$

$$v_{m,t} \sim \mathcal{N}(0,1),\ 0\,\mathrm{V} \le u_{m,t} \le 10\,\mathrm{V},\ m = 1, 2 \tag{25}$$

where $F_{\mathrm{BW}}(q|5, \underline{\omega}, \overline{\omega})$ is the fifth-order Butterworth filter whose lower and upper cutoff frequencies are respectively $\underline{\omega}$ and $\overline{\omega}$, and $u_{\mathrm{ss},m}$ is the steady-state value of $u_{m,t}$. $\boldsymbol{\chi} = [\overline{\omega}, \underline{\omega}, \overline{\lambda}, \underline{\lambda}]^\top$ was determined by solving the proposed and existing experiment design problems. Moreover, the frequency constraint of Eq. (21) was applied to both experiment design methods, where ϵ was set to 0.05.

Only the objective function was changed to $J_{\mathrm{SPS}}|_{T=T_{\mathrm{BO}}}$ and $J_{\mathrm{Dopt}}|_{T=T_{\mathrm{BO}}}$ to switch the experiment design method. Regarding J_{SPS}, the parameters R and $\tilde{R}$ of SPS were 200 and 10, respectively.

N_{BO}, N_{rand}, and T_{BO} were 8, 2, and 40, respectively. Thus, the input-output data with 400 samples as shown in Fig. 4 were obtained. Based on the data, the prediction model with the first-order BJ structure was identified using the prediction-error method (Ljung, 1998).

The identified models from the two different experiment design methods were compared using the control performance of MPC. The optimization problem of MPC was

$$\min_{\boldsymbol{u}_{t_0}, \boldsymbol{u}_{t_0+1}} \sum_{t=t_0+1}^{t_0+20} \left(\|\hat{\boldsymbol{y}}_t - \boldsymbol{r}_t\|^2_{\boldsymbol{W}_y} + \|\boldsymbol{u}_{t-1} - \boldsymbol{u}_{t-2}\|^2 \right) \tag{26}$$

subject to

$$5\,\mathrm{V} < u_{m,t} < 10\,\mathrm{V},\ m = 1, 2, \tag{27}$$

$$\boldsymbol{r}_t = (1 - 0.8^{t-t_0})\boldsymbol{s}_t + 0.8^{t-t_0}\boldsymbol{y}_{t_0}, \tag{28}$$

$$\boldsymbol{u}_t = \boldsymbol{u}_{t_0+1},\ t \ge t_0 + 2, \tag{29}$$

where $\hat{\boldsymbol{y}}_t$, $\boldsymbol{r}_t$, and $\boldsymbol{s}_t$ are respectively the predicted outputs, the reference trajectories, and the setpoints, and $\boldsymbol{W}_y = \mathrm{diag}(1, 0, 1)$.

4.3 Results and Discussion

The time-series plot of the MPC operation for comparing the models obtained from both experiment design methods is shown in Fig. 5, where the setpoints of $y_{1,t}$ and $y_{3,t}$ were changed stepwise from 12 to 14 cm. The MPC response using the model of the existing method has a larger overshoot and longer settling time than that of the proposed method.

The mean squared control error (MSE), rise time, settling time, and overshoot in Fig. 5 are summarised in Table 2. The mean $J_{\mathrm{SPS}}|_{T=T_{\mathrm{BO}}}$ and $J_{\mathrm{Dopt}}|_{T=T_{\mathrm{BO}}}$ over the 10 iterations of BO are shown in Table 2 as well. From Table 2, it is shown that the proposed method gives the lower values of J_{SPS} and MSE than the existing method. This suggests that the proposed method provides a better prediction

Table 2. The performance indices of the existing and proposed experiment design methods.

Experiment design method	Data-quality indices		Control performance of MPC								
	Mean	Mean	MSE [cm^2]			Rise time [s]		Settling time [s]		Overshoot [%]	
	$J_{\mathrm{SPS}}\|_{T=\tilde{T}}$	$J_{\mathrm{Dopt}}\|_{T=\tilde{T}}$	y_1	y_3	Total	y_1	y_3	y_1	y_3	y_1	y_3
Existing method	7.58	7.28	0.152	0.140	0.292	183.6	183.6	938.4	938.4	35.7	28.8
Proposed method	1.10	11.96	0.115	0.111	0.225	204.0	183.6	387.6	448.8	12.1	19.9

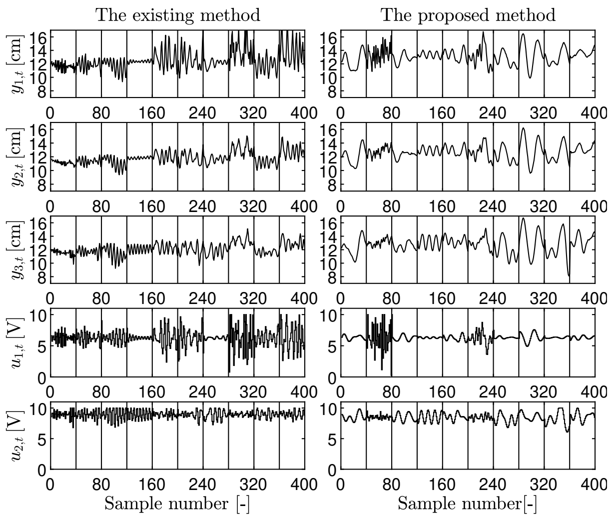

Fig. 4. Time-series plots of the data obtained in the identification experiments using both the proposed and the existing design methods.

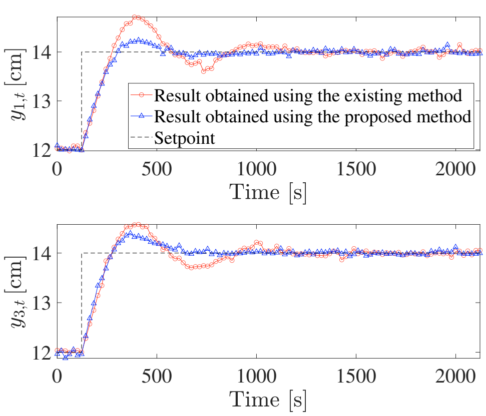

Fig. 5. Comparison of the experiment design methods in the MPC operation.

model. On the other hand, the mean J_{Dopt} is smaller in the existing method than in the proposed method. This implies that the objective function of the D-optimal design is misleading when the number of samples is small.

5. CONCLUSION

In this paper, a new method for experiment design for identification based on a nonasymptotic confidence region of the system identification model was proposed. Based on the volume of the confidence region calculated using sign-perturbed sums, the objective function for the experiment design for identification was derived. Moreover, the frequency constraints for input signals were introduced to reduce the dependency of the objective function on the number of samples. To solve the proposed design problem, Bayesian optimization was used since the objective function requires the outcome of a data-acquisition experiment for evaluation.

The proposed method was compared with the conventional D-optimal design method in the experiment using the three-tank system. It was shown that the prediction model obtained from the proposed method provides better control performance for the liquid level when using model predictive control than the conventional method.

REFERENCES

Csáji, B.C., Campi, M.C., and Weyer, E. (2012). Sign-perturbed sums (SPS): A method for constructing exact finite-sample confidence regions for general linear systems. *Proceedings of the IEEE Conference on Decision and Control*, 7321–7326.

Csáji, B.C., Campi, M.C., and Weyer, E. (2015). Sign-perturbed sums: A new system identification approach for constructing exact non-asymptotic confidence regions in linear regression models. *IEEE Transactions on Signal Processing*, 63(1), 169–181.

Csáji, B.C. and Weyer, E. (2015). Closed-loop applicability of the sign-perturbed sums method. *54th IEEE Conference on Decision and Control (CDC)*, 1441–1446.

Goodwin, G.C. (1977). *Dynamic system identification: experiment design and data analysis*, volume 136. Academic Press.

Greenhill, S., Rana, S., Gupta, S., Vellanki, P., and Venkatesh, S. (2020). Bayesian optimization for adaptive experimental design: A review. *IEEE Access*, 8, 13937–13948.

Gunther, L. and Friedrich, P. (2014). *Introduction to Quasi-Monte Carlo Integration and Application.* Birkhäuser. doi:10.1007/978-3-319-03425-6.

Kolumbán, S. and Csáji, B.C. (2018). Towards d-optimal input design for finite-sample system identification. *18th IFAC Symposium on System Identification SYSID 2018*, 51(15), 215–220.

Ljung, L. (1998). *System Identification: Theory for the User.* Pearson, 2nd edition.

Mockus, J. (1998). The application of bayesian methods for seeking the extremum. *Towards Global Optimization*, 2, 117.

Morari, M. and H. Lee, J. (1999). Model predictive control: past, present and future. *Computers & Chemical Engineering*, 23(4), 667–682.

Shahriari, B., Swersky, K., Wang, Z., Adams, R.P., and de Freitas, N. (2016). Taking the human out of the loop: A review of bayesian optimization. *Proceedings of the IEEE*, 104(1), 148–175.

Shardt, Y.A.W. (2022). *Statistics for Chemical and Process Engineers-A Modern Approach.* Springer Cham.

Szentpéteri, S. and Csáji, B.C. (2023). Non-asymptotic state-space identification of closed-loop stochastic linear systems using instrumental variables. *Systems & Control Letters*, 178, 105565.