Experiment designs to minimize input peak and crest factor in MIMO system identification

Häggblom, Kurt Erik

Published in:
IEEE CCTA 2019 : 3rd IEEE conference on control technology and applications : August 19-21, 2019, City University of Hong Kong, Hong Kong, China

DOI:
10.1109/CCTA.2019.8920499

Published: 01/01/2019

Document Version
Accepted author manuscript

Document License
Publisher rights policy

Experiment Designs to Minimize Input Peak and Crest Factor in MIMO System Identification

Kurt E. Häggblom, *Senior Member, IEEE*

***Abstract*— The quality of the data for system identification is of utmost of importance. Ideally, the output data should satisfy requirements regarding variance, and in the case of multiple-input multiple-output (MIMO) systems, correlation between the outputs. Usually, it is desired to achieve this by input perturbations with peak values as small as possible. A related measure is the crest factor, which is a measure of the (inverse) power of an input perturbation with given peak value. In this paper, the effect of minimizing the input peak and the crest factor, subject to desired output variances/covariances, is studied for various types of perturbation signals. The design procedure is completely data based; data are obtained by one or more preliminary experiments with the system to be identified. A model of an ill-conditioned distillation column is used for illustration.**

I. Introduction

A problem in the identification of multiple-input multiple-output (MIMO) systems is that the system outputs in an identification experiment may be strongly correlated if the inputs are perturbed simultaneously with (nearly) uncorrelated inputs. Standard practice is to use such inputs, as produced by the idinput command in [1]. See also [2]. However, the output correlation reduces identifiability. To maximize the information content of the outputs, their variances should be at some maximum level with sample correlations as small as possible (ideally zero). This concept is also used in partial least squares (PLS) regression, where the latent variables that extract maximum information have this property.

Some model-based experiment design methods have been proposed [3–5]. A drawback of the methods, besides numerical complexity, is that the resulting input perturbations are of some special type (not the typically used pseudo random binary signals, for example). Although the main design criteria are not related to the output distribution, the methods produce nearly uncorrelated outputs [6]. A model-based design method that directly addresses the output distribution was proposed by Häggblom [7, 8]. The optimizations are much simpler than in the previous methods.

In this paper, a data-based design method is used to obtain outputs with desired sample properties. Obviously, a data-based method for input design is preferable from a practical point of view. Data are obtained from one or more preliminary experiments with the system to be identified. The experiment can be a standard MIMO experiment with uncorrelated inputs, but better data for the design can be obtained by performing experiments with one input at a time.

K. E. Häggblom is with the Faculty of Science and Engineering, Åbo Akademi University, FI–20500 Turku, Finland (email: khaggblo@abo.fi).

The proposed method makes it possible to optimize some additional property besides output correlation. In this paper, minimization of input peak value and crest factor [3, 9, 10] subject to desired output variances are considered. Results for minimization of output peak, or constrained output peak when inputs are optimized, are also provided. Five different types of input perturbations are considered, namely, two kinds of random binary signals, a pseudo random binary signal, and two kinds of multi-sinusoidal signals. Results are illustrated by an ill-conditioned distillation column model.

II. Problem Formulation

Input design for identification of linear MIMO systems is considered. The system has n inputs $u(k)$ and n outputs $y(k)$, sampled at time instants $k = 1,\ldots,n_\mathrm{s}$. The variables are related by a dynamic relationship

$$y(k) = G(\mathrm{q})u(k)\,, \tag{1}$$

where $G(\mathrm{q})$ is a matrix of pulse transfer operators defined through the shift operator q. This relationship is assumed to be initially unknown and it is not implied that a model of this form is to be identified.

The input design uses an n-dimensional perturbation signal $\xi(k)$. In this paper, this signal is one of five different types of perturbation signals. The correlation between the individual signals $\xi_i(k)$, $i = 1,\ldots,n$, should preferably be small (ideally zero) because they serve as basis functions in the design. In practice, this is achieved by constructing each $\xi_i(k)$ from a base sequence $\xi_0(k)$, of length N, by shifting it (approximately) $(i-1)N/n$ positions in a circular way. As suggested by Ljung [9, p. 424], more than one period of the sequence may be used to give a total sequence length $n_\mathrm{s} = n_\mathrm{p}N$, where n_p is the number of periods.

The input $u(k)$ to be applied in the identification experiment is given by a linear transformation

$$u(k) = T\xi(k)\,, \tag{2}$$

where T is a constant matrix determined in the input design. The main objective of the input design is to make the output samples $y_i(k)$, $k = 1,\ldots,n_\mathrm{s}$, $i = 1,\ldots,n$, uncorrelated with $y_j(k)$, $j \neq i$, in the identification experiment. The transformation (2) makes the inputs $u_i(k)$, $i = 1,\ldots,n$, correlated with one another, but this is a lesser problem than correlated outputs because the inputs are assumed to be exactly known (being setpoints, for example), whereas the outputs generally are contaminated by noise and disturbances.

The output correlation depends on the output covariance matrix, which for an $n \times n$ system is defined by $n(n+1)/2$ independent parameters. This implies that the same number of adjustable elements of T is sufficient to produce uncorrelated outputs. Hence, it is possible to optimize some quantity besides output correlation if a full T matrix with unconstrained elements is used. The extra degrees of freedom are here used to minimize the input peak value $\max_{i,k} |u_i(k)|$ and the crest factor [3, 9, 10] $\max_i u_{i,\mathrm{CF}}$, where

$$u_{i,\mathrm{CF}} = \frac{\max_k |u_i(k)|}{u_{i,\mathrm{RMS}}} = \frac{\max_k |u_i(k)|}{\sqrt{\frac{1}{n_\mathrm{s}} \sum_1^{n_\mathrm{s}} u_i(k)^2}} . \tag{3}$$

For reference, results of minimizing the output peak value $\max_{i,k} |y_i(k)|$ are also given.

III. Design Procedure

The input design objective is to determine the perturbation signal $\xi(k)$ and the constant transformation matrix T to give the outputs specified properties. When $u(k)$, given by (2), is applied to the system (1), the obtained outputs $y(k)$, $k = 1, \ldots, n_\mathrm{s}$, can be collected in a matrix Y of size $n_\mathrm{s} \times n$. In particular, the sampled covariance matrix $P = \operatorname{cov} Y$ is desired to be diagonal, which means no correlation between the outputs. It is assumed that the desired output variances are $\operatorname{var} y_i = 1$, $i = 1, \ldots, n$. Thus, the objective is to obtain $P = I$.

A. Decision Variables

In the design optimization, T is a decision matrix with elements t_{ij}. To facilitate the design, the vector

$$x = \operatorname{vec} T \tag{4}$$

is introduced. Thus, $x_\ell = t_{ij}$, $\ell = i + n(j-1)$. The matrix X is defined

$$X = x^\mathrm{T} \otimes I_n , \tag{5}$$

where x^T is the transpose of x, $\otimes$ is the Kronecker product, and I_n is the n-dimensional identity matrix.

B. Output Covariance

Assume the perturbation $\xi_j(k)$, $k = 1, \ldots, n_\mathrm{s}$, is applied to the input $u_i(k)$ with $u_m(k) = 0$, $m \neq i$. This produces an $n_\mathrm{s} \times n$ matrix of outputs Y_ℓ, $\ell = i + n(j-1)$, where $y(k)^\mathrm{T}$ occupies the kth row. In total, all combinations of $\xi_j(k)$ and $u_i(k)$ yield the output matrices Y_ℓ, $\ell = 1, \ldots, n^2$. If the system is linear, the output obtained by applying all inputs simultaneously according to (2) is given by

$$Y = \sum_{\ell=1}^{n^2} x_\ell Y_\ell = Y_0 X^\mathrm{T} , \tag{6}$$

where $Y_0 = [Y_1, \ldots, Y_{n^2}]$. The output covariance matrix is

$$P = X P_0 X^\mathrm{T} , \tag{7}$$

where $P_0 = \operatorname{cov} Y_0$. If P_0 is known, (7) can be solved to yield $P = I$. This solution gives T through (4) and (5).

C. Optimization

Equation (7) contains n^2 unknowns in X whereas the number of independent equations is $n(n+1)/2$ due to the symmetry of P. Thus, (7) is underdetermined and the solution is not unique. This can be utilized for optimization. It is, for example, possible to minimize the output peak $\max_{i,k} |y_i(k)|$, the input peak $\max_{i,k} |u_i(k)|$, or the crest factor $\max_i u_{i,\mathrm{CF}}$, subject to (7) and $P = I$.

The output peak can be minimized by minimizing $\max \operatorname{abs}(Y)$, where Y is given by (6). Similarly, the input peak can be minimized by minimizing $\max \operatorname{abs}(U)$, where

$$U = \Xi T^\mathrm{T} , \quad \Xi = \left[\xi(1),\ \xi(2),\ \ldots,\ \xi(n_\mathrm{s})\right]^\mathrm{T} . \tag{8}$$

The crest factor for u_i, i.e. $u_{i,\mathrm{CF}}$, can be expressed in terms of U_i, which is the ith column of U.

In this work, optimizations are done with the MATLAB software [11] and the YALMIP toolbox [12]. In YALMIP, minimization of $p = \max_{i,k} |y_i(k)|$, subject to the appropriate constraints, is easily formulated as

$$\min_x p \ \text{ s.t. } -p \le Y \le p ,\ P = I ,\ (5), (6), (7). \tag{9}$$

Note that the inequalities in (9) apply to every element of Y. Similarly, minimization of $q = \max_{i,k} |u_i(k)|$ is formulated

$$\min_x q \ \text{ s.t. } -q \le U \le q ,\ P = I ,\ (5), (7), (8). \tag{10}$$

The crest factor is minimized by

$$\min_x r \ \text{ s.t. } \begin{cases} r \operatorname{diag} \operatorname{cov} U \ge (U \circ U)^\mathrm{T} \\ P = I, (5), (7), (8) \end{cases} , \tag{11}$$

where $\circ$ is the Hadamard product and $\max_i u_{i,\mathrm{CF}} = r^{1/2}$. In all optimizations, further constraints can be included (e.g., $-q \le U \le q$ with q fixed in (9)).

The main problem in the optimizations is that they are non-convex and nonlinear. In practice, the nonlinearity can be handled by linearization, as done in [8] in a model-based approach to experiment design. An alternative is to solve the nonlinear problem by the fmincon function of MATLAB. In both methods, an in initial guess $x = x_0$ is required. As the problem is non-convex, different initial guesses often give different final (sub)optimal solutions. Because the optimizations are fast, a large set of initial guesses can be tried. In this way, both methods tend to give the same best solution, which therefore is likely to be the global optimum.

A third alternative is to use the branch-and-bound algorithm bmibnb provided by YALMIP. In addition to fmincon, the bmibnb solver uses Gurobi [13]. The method has the advantage that an initial guess is not required. Instead, some reasonable bounds on the decision variables have to be provided. In all comparisons of the three methods to solve this problem, the branch-and-bound solution is the same as, or only slightly worse than (less than 1 % worse), the best solution found by the other methods. The results reported in this paper are those obtained by the branch-and-bound method.

D. Obtaining Data

Data for the design is generated by one or more experiments with the system to be identified. The most reliable method is to make an experiment with every combination $u_i(k) = \xi_j(k)$, $i = 1,\ldots,n$, $j = 1,\ldots,n$, one at a time. This results in $n \times n$ experiments, each one yielding a matrix of sampled outputs Y_ℓ, $\ell = i + (j-1)n$. The overall covariance matrix P_0 can then be calculated and used in (7) for the input design.

For $n > 2$, this is a lot of experiments, and even for $n = 2$, four experiments might be undesirable. An alternative is to make n experiments with $u_i(k) = \xi_j(k)$, $i = 1,\ldots,n$ and j arbitrary (e.g., $j = i$ or $j = 1$). For each experiment, a finite impulse-response (FIR) model can be determined in a simple way. This makes it possible to simulate all combinations $u_i(k) = \xi_j(k)$ to obtain the required output data matrices.

It is possible to take this one step further and make only one experiment with $u(k) = \xi(k)$. Then, all inputs are perturbed simultaneously. If the components $\xi_i(k)$, $i = 1,\ldots,n$, are essentially uncorrelated with one another, it is possible to determine the required FIR models from this single experiment and proceed as above. This kind of experiment is the standard identification experiment for MIMO systems [1, 2], but here the data is used to design a better experiment.

E. Effect of Noise

It is the presence of noise and disturbances in the data that makes system identification challenging [9, 10, 14]. In a system with strongly correlated outputs, the outputs contain almost the same information. Noise and disturbances may then prevent small differences in the system dynamics to be observed. The effect of this may be disastrous, especially if the identified model is to be used for control design.

The objective of the proposed input design is to produce outputs that are not strongly correlated, ideally not correlated at all. In the presence of output noise, it is then much easier to catch the differences between the outputs. The transformation (2) makes the inputs correlated, but this is a lesser problem if the inputs are not contaminated with noise. This is the case if the inputs are setpoints of flow rates, for example.

If noise is present in the data matrices Y_ℓ, $\ell = 1,\ldots,n^2$, it will generally affect the transformation matrix T in the input design. The result of the experiment based on T may then not produce uncorrelated outputs. However, even if they are correlated, it is an advantage as long as the output correlation is reduced from what it would be in a standard identification experiment, for example.

In a related study, the effect of output noise was explicitly studied for some systems with different degrees of ill-conditioning and $n = 2$ or 3. It was found that the effect of a signal-to-noise (SNR) ratio = 50 was negligible, regardless of the system. Even for very ill-conditioned systems, as the one in the case study to follow, SNR = 20 could be handled well, even when only one initial experiment was used to obtain the data via FIR models. With two initial experiments, even SNR = 10 gave good results.

IV. Case Study Setup

A. Model of Ill-Conditioned Distillation Column

The experiment designs with various types of perturbations are illustrated by the use of an ill-conditioned distillation column model [15]

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned}, \tag{12a}$$

$$A = \begin{bmatrix} -\frac{1}{194} & 0 \\ 0 & -\frac{1}{15} \end{bmatrix},\ B = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix},\ C = \begin{bmatrix} \frac{87.8}{194} & \frac{1.4}{15} \\ \frac{108.2}{194} & -\frac{1.4}{15} \end{bmatrix}. \tag{12b}$$

The condition number of the static gain matrix is 142.

B. Perturbation Data

In this study, data for the experiment designs are obtained by four experiments with every combination of u_i and ξ_j for every perturbation type (see below) as described in Section III*D*. The sampling time is $T_s = 1$ time unit and one period of the input sequences consists of $N = 1020$ samples giving the period length $P = NT_s = 1020$. In all cases, the signal amplitudes are normalized to 1. In the optimizations, input sequences consisting of two sequential periods, with opposite signs, are used. This reduces harmful startup effects.

Next, the types of perturbations are defined. In this study, they are selected because they have quite different properties regarding input peaks and crest factors. Small values indicate that the perturbation is efficient [3, 9, 10]. Signal design guidelines, with further references, are given in [6, 16].

1) Random Binary Signal

A random binary signal (RBS) switches between two levels (a and $-a$) in a random way. In this paper, the idinput command of MATLAB's System Identification Toolbox [1] is used to generate a RBS with the amplitude $a = 1$.

A perfect RBS of infinite length has no autocorrelation, but a finite-length RBS will contain some autocorrelation. In practice, this autocorrelation might be quite significant [7]. In this study, 99 sequences were generated and the sequence with the median autocorrelation was selected.

2) Correlated Binary Signal

A correlated binary signal (CBS) contains autocorrelation by design. In this study, a RBS of length 255 was first generated as described above. The CBS was constructed by holding every sample of the RBS constant for four samples. This signal is an intermediate between a RBS and a PRBS.

3) Pseudo Random Binary Signal

A pseudo random binary (PRB) signal is a deterministic, finite-length, binary signal that mimics the properties of white noise. This property is guaranteed, unlike that of a RBS. The PRB signal switches between two levels (a and $-a$) with a minimum switching time T_{sw}, which is some integer multiple of the sampling time T_s. In this study, $T_{sw} = 4$ is used in accordance with [9, 10]. The PRB signal is a maximum-length sequence of length $2^8 - 1 = 255$ [9]. The PRB sequence is generated by the idinput command [1].

4) Multi-Sinusoidal Signal with Schroeder Phase Shift

A multi-sinusoidal signal is determined as a sum of n_f sinusoidals with different (equally-spaced) frequencies. To prevent excessive cumulation of the amplitudes, which increases the crest factor, the phase shifts are selected in some special way. One way is to use Schroeder phase-shift spacing [9, 10] (SMS). In this study, $n_\mathrm{f} = 102$ with frequencies ranging from $2\pi / P$ to $2\pi n_\mathrm{f} / P$ are used. Essentially, the SMS covers the same frequency range as the PRB signal.

5) Multi-Sinusoidal Signal with Random Phase Shifts

Although the phase-shift spacing in the SMS is selected to restrict the amplitude cumulation, it is still a problem. An alternative is to use random phase shifts (RMS). With the same design parameters as for the SMS, the idinput command was used to generate 99 RMS signals, and the one with the smallest amplitude build-up was selected.

V. Results

The results of the case study for the noise-free case are presented in this section. After presenting the result of a standard identification experiment, the results of optimizations according to various criteria with different kinds of input perturbations are given. The results include the obtained transformation matrix T, the output range y_R defined as $y_\mathrm{R} = [\max_{i,k} y_i(k) \quad \min_{i,k} y_i(k)]^\mathrm{T}$, the input range u_R defined similarly, and the crest factor $u_\mathrm{CF} = \max_i u_{i,\mathrm{CF}}$. In all cases, the desired sample output variance $P = I$ was obtained.

A. Standard Identification Experiment

A standard identification experiment using (almost) uncorrelated PRB inputs is illustrated in Fig. 1. As can be seen, the outputs are very strongly correlated. This is also shown by the scatter plot of y_1 vs. y_2 in Fig. 2.

B. Minimization of Output Peak

Table I summarizes the results for output peak minimization. As can be seen, the maximum/minimum output and input values are symmetrically distributed. The input peak values and the crest factor are smallest for the PRB signal. Scatter plots are shown in Fig. 3.

C. Minimization of Input Peak

Table II summarizes the results for input peak minimization. The results concerning u_R and u_CF are clearly improved for RBS and CBS perturbations. However, the output range y_R is increased. For SMS and RMS, there is no significant difference from output peak minimization. Scatter plots are shown in Fig. 4.

D. Minimization of Crest Factor

Table III summarizes the results for crest factor minimization. The results are very similar to input peak minimization. A reason for this is that the outputs are constrained to satisfy $P = I$. Since input peak minimization is an easier optimization problem than crest factor minimization, the latter can be dispensed with. Scatter plots are shown in Fig. 5.

TABLE I. MINIMIZATION OF OUTPUT PEAK

	T	y_R	u_R	u_CF
RBS	$\begin{bmatrix} -1.8853 & 1.8809 \\ -1.9125 & 1.7852 \end{bmatrix}$	$\begin{bmatrix} 2.8329 \\ -2.8377 \end{bmatrix}$	$\begin{bmatrix} 3.7662 \\ -3.7662 \end{bmatrix}$	1.3896
CBS	$\begin{bmatrix} -1.6500 & 2.2785 \\ -1.8436 & 2.1012 \end{bmatrix}$	$\begin{bmatrix} 2.6945 \\ -2.6853 \end{bmatrix}$	$\begin{bmatrix} 3.9447 \\ -3.9447 \end{bmatrix}$	1.4305
PRB	$\begin{bmatrix} 1.4526 & 0.3649 \\ 1.4097 & 0.4367 \end{bmatrix}$	$\begin{bmatrix} 2.4382 \\ -2.4414 \end{bmatrix}$	$\begin{bmatrix} 1.8464 \\ -1.8464 \end{bmatrix}$	1.2522
SMS	$\begin{bmatrix} 0.7238 & 2.1008 \\ 0.5783 & 2.0890 \end{bmatrix}$	$\begin{bmatrix} 3.0325 \\ -2.9383 \end{bmatrix}$	$\begin{bmatrix} 2.6862 \\ -2.6862 \end{bmatrix}$	1.9955
RMS	$\begin{bmatrix} -1.9163 & 2.5224 \\ -1.7470 & 2.5909 \end{bmatrix}$	$\begin{bmatrix} 2.7225 \\ -2.7474 \end{bmatrix}$	$\begin{bmatrix} 3.3947 \\ -3.3947 \end{bmatrix}$	2.6313

TABLE II. MINIMIZATION OF INPUT PEAK

	T	y_R	u_R	u_CF
RBS	$\begin{bmatrix} 0.0753 & 2.4941 \\ -0.0748 & 2.4946 \end{bmatrix}$	$\begin{bmatrix} 3.2562 \\ -3.3065 \end{bmatrix}$	$\begin{bmatrix} 2.5694 \\ -2.5694 \end{bmatrix}$	1.0306
CBS	$\begin{bmatrix} 0.0868 & 3.0119 \\ -0.1561 & 2.9426 \end{bmatrix}$	$\begin{bmatrix} 3.4829 \\ -3.4799 \end{bmatrix}$	$\begin{bmatrix} 3.0986 \\ -3.0986 \end{bmatrix}$	1.0528
PRB	$\begin{bmatrix} 1.4709 & 0.0401 \\ 1.4613 & -0.0496 \end{bmatrix}$	$\begin{bmatrix} 2.7370 \\ -2.7167 \end{bmatrix}$	$\begin{bmatrix} 1.5110 \\ -1.5110 \end{bmatrix}$	1.0330
SMS	$\begin{bmatrix} -0.0385 & 2.1635 \\ 0.1022 & 2.1515 \end{bmatrix}$	$\begin{bmatrix} 3.0035 \\ -3.1781 \end{bmatrix}$	$\begin{bmatrix} 2.1913 \\ -2.1913 \end{bmatrix}$	1.6793
RMS	$\begin{bmatrix} -0.1024 & 3.1833 \\ -0.2769 & 3.1366 \end{bmatrix}$	$\begin{bmatrix} 3.0278 \\ -2.9972 \end{bmatrix}$	$\begin{bmatrix} 3.1230 \\ -3.1230 \end{bmatrix}$	2.4204

TABLE III. MINIMIZATION OF CREST FACTOR

	T	y_R	u_R	u_CF
RBS	$\begin{bmatrix} 2.5015 & 0.0628 \\ 2.5509 & -0.0689 \end{bmatrix}$	$\begin{bmatrix} 3.4267 \\ -3.3006 \end{bmatrix}$	$\begin{bmatrix} 2.6198 \\ -2.6198 \end{bmatrix}$	1.0254
CBS	$\begin{bmatrix} 3.1364 & 0.1132 \\ 3.0160 & -0.1026 \end{bmatrix}$	$\begin{bmatrix} 3.6451 \\ -3.6226 \end{bmatrix}$	$\begin{bmatrix} 3.2496 \\ -3.2496 \end{bmatrix}$	1.0341
PRB	$\begin{bmatrix} 1.4709 & 0.0488 \\ 1.4616 & -0.0449 \end{bmatrix}$	$\begin{bmatrix} 2.7364 \\ -2.7161 \end{bmatrix}$	$\begin{bmatrix} 1.5158 \\ -1.5158 \end{bmatrix}$	1.0299
SMS	$\begin{bmatrix} 2.2353 & -0.0383 \\ 2.2040 & 0.0910 \end{bmatrix}$	$\begin{bmatrix} 3.2148 \\ -3.1875 \end{bmatrix}$	$\begin{bmatrix} 2.2630 \\ -2.2630 \end{bmatrix}$	1.6708
RMS	$\begin{bmatrix} -0.2360 & 3.1957 \\ -0.0654 & 3.1547 \end{bmatrix}$	$\begin{bmatrix} 3.0541 \\ -3.0767 \end{bmatrix}$	$\begin{bmatrix} 3.1646 \\ -3.1646 \end{bmatrix}$	2.4101

TABLE IV. OUTPUT MINIMIZATION WITH INPUT CONSTRAINT

	T	y_R	u_R	u_CF
RBS	$\begin{bmatrix} 0.4165 & 2.4590 \\ 0.6182 & 2.3726 \end{bmatrix}$	$\begin{bmatrix} 2.9856 \\ -2.9314 \end{bmatrix}$	$\begin{bmatrix} 2.9908 \\ -2.9908 \end{bmatrix}$	1.2302
CBS	$\begin{bmatrix} -0.4832 & 2.8874 \\ -0.7199 & 2.7801 \end{bmatrix}$	$\begin{bmatrix} 3.3608 \\ -3.3557 \end{bmatrix}$	$\begin{bmatrix} 3.5000 \\ -3.5000 \end{bmatrix}$	1.2266
PRB	$\begin{bmatrix} 1.4903 & 0.0578 \\ 1.4609 & 0.1391 \end{bmatrix}$	$\begin{bmatrix} 2.4668 \\ -2.4877 \end{bmatrix}$	$\begin{bmatrix} 1.6000 \\ -1.6000 \end{bmatrix}$	1.0904
SMS	$\begin{bmatrix} -0.2023 & 2.1512 \\ -0.0590 & 2.1495 \end{bmatrix}$	$\begin{bmatrix} 2.9827 \\ -3.1319 \end{bmatrix}$	$\begin{bmatrix} 2.3000 \\ -2.3000 \end{bmatrix}$	1.7571
RMS	$\begin{bmatrix} -0.3621 & 3.1583 \\ -0.5332 & 3.0961 \end{bmatrix}$	$\begin{bmatrix} 2.9645 \\ -2.9340 \end{bmatrix}$	$\begin{bmatrix} 3.2000 \\ -3.2000 \end{bmatrix}$	2.4587

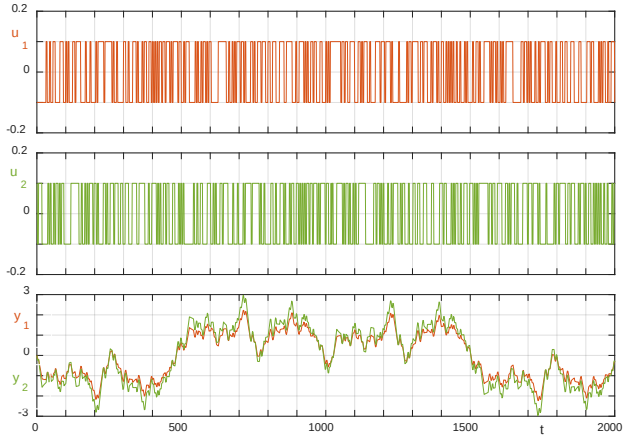

Figure 1. Uncorrelated PRB inputs (upper two panels) and correlated outputs (lower panel) of standard identification experiment.

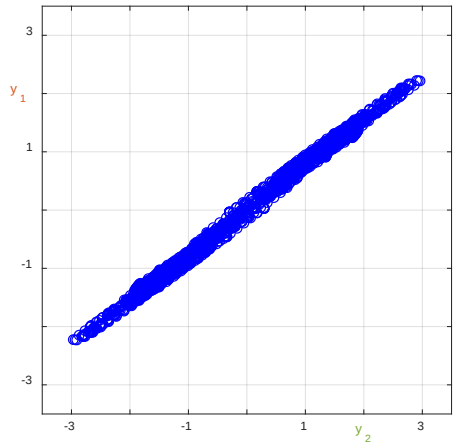

Figure 2. Scatter plot of outputs in standard identification experiment.

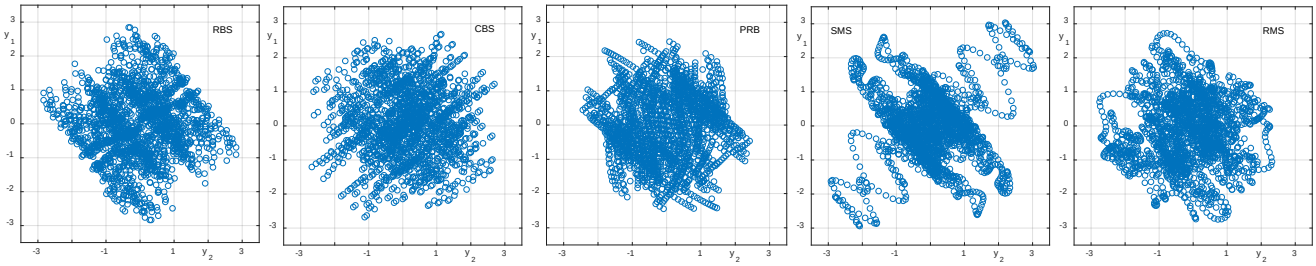

Figure 3. Scatter plots for minimized output peak.

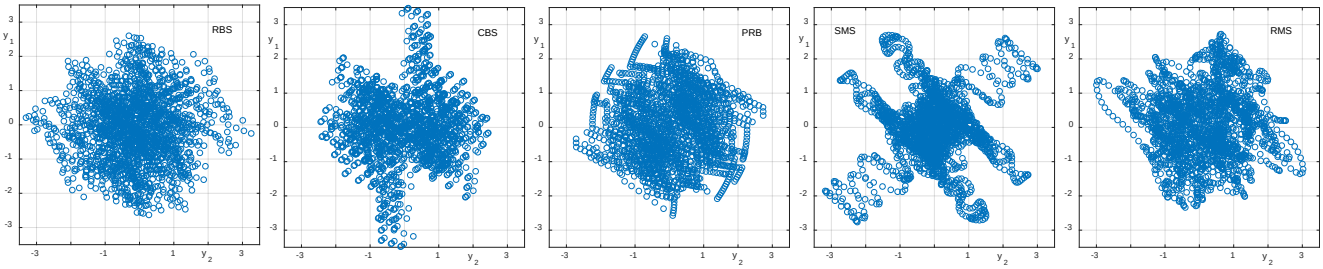

Figure 4. Scatter plots for minimized input peak.

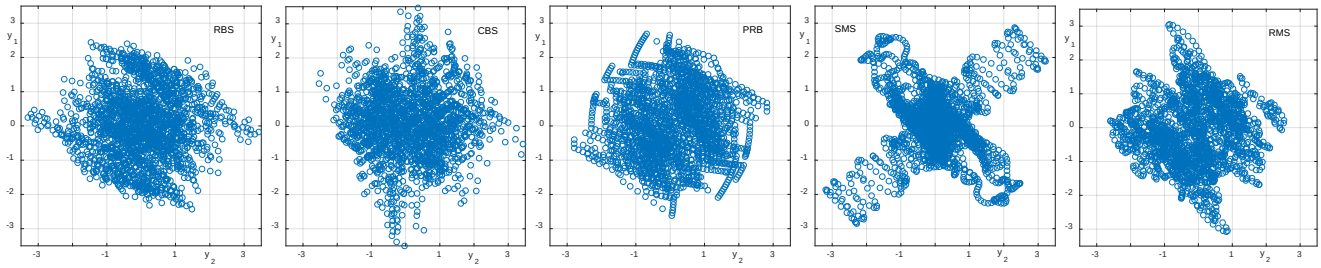

Figure 5. Scatter plots for minimized crest factor.

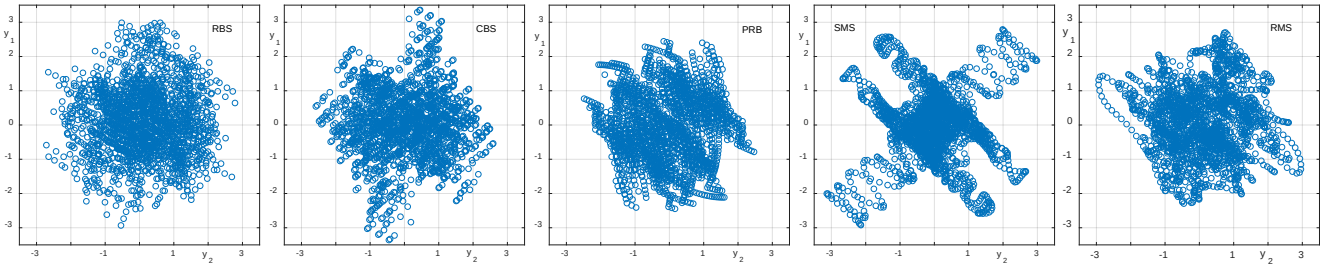

Figure 6. Scatter plots for minimized output peak with constraint on input peak.

E. Minimization of Output Peak with Input Peak Constraint

Minimization of the output peak results in large input peaks, and vice versa. As a compromise, one can minimize one signal peak with a constraint on the other signal peak. Table IV gives results of output peak minimization subject to the input peak constraints indicated by u_{R}. Scatter plots are shown in Fig. 6.

VI. Conclusion

A data-based method for design of experiments for identification of MIMO systems was described. Previous methods have (mostly) been model-based. Obviously, a model-free method for experiment design is preferable from a practical viewpoint.

The required data can be obtained from one or more preliminary experiments with the system. The input design yields uncorrelated outputs, which is good for identifiability. In addition, input and output peak values as well as the crest factor can be minimized subject to desired output variances with no output correlation. Minimizing these quantities is an advantage in process operation.

Five types of perturbation signals were considered, namely, random binary signals (RBS), correlated binary signals (CBS), pseudo random binary (PRB) signals, multi-sinusoidal signals with Schroeder phase shift (SMS), and multi-sinusoidal signals with random phase shifts (RMS). In this study, no advantage concerning peak values and the crest factor was found with the multi-sinusoidal signals. When the crest factor was minimized, the RBS and CBS was close to the PRB signal, but with larger peak values. In all respects, the PRB signal was the superior type of perturbation in this study.

The optimization problems are straightforward to formulate and can be solved easily using standard MATLAB software. In this work, the YALMIP toolbox was also used.

References

[1] L. Ljung, *System Identification Toolbox User's Guide*. Natick, MA: The MathWorks, Inc., 2019.

[2] R. Isermann and M. Münchhof, *Identification of Dynamic Systems*. Berlin and Heidelberg: Springer, 2011.

[3] D. E. Rivera, H. Lee, H. D. Mittelmann, and M. W. Braun, "High-purity distillation: using plant-friendly multisine signals to identify a strongly interactive process," *IEEE Control Syst. Mag*., vol. 27, no. 5, pp. 72–89, Oct. 2007.

[4] M. L. Darby and M. Nikolaou, "Identification test design for multivariable model-based control: An industrial perspective," *Control Eng. Pract*., vol. 22, pp. 165–180, Jan. 2014.

[5] A. Kumar and S. Narasimhan, "Optimal input signal design for identification of interactive and ill-conditioned systems," *Ind. Eng. Chem. Res.*, vol. 55, no. 14, pp. 4000–4010, March 2016.

[6] K. E. Häggblom, "Input designs to obtain uncorrelated outputs in MIMO system identification," in *Proc 13th Int. Symp. on Process Systems Engineering – PSE 2018*, San Diego, CA, USA, 2018, pp. 637–642.

[7] K. E. Häggblom, "A new optimization-based approach to experiment design for dynamic MIMO identification," *IFAC-PapersOnLine*, vol. 50, no 1, pp. 7321–7326, July 2017.

[8] K. E. Häggblom, "Easy ways to design inputs to obtain uncorrelated outputs in MIMO system identification," *IFAC-PapersOnLine*, vol. 51, no 15, pp. 227–232, July 2018.

[9] L. Ljung, *System Identification: Theory for the User*. Upper Saddle River, NJ: Prentice Hall, 1999.

[10] R. Pintelon and J. Schoukens, *System Identification: A Frequency Domain Approach*. Hoboken, NJ: Wiley, 2012.

[11] The MathWorks, Inc, *Optimization Toolbox™ User's Guide*. Natick, MA: The MathWorks, Inc., 2018.

[12] J. Löfberg, "YALMIP: A toolbox for modeling and optimization in MATLAB", in *Proc. IEEE Int. Symp. Computer Aided Control Systems Design (CACSD)*, Taipei, Taiwan, 2004, pp. 284–289.

[13] Gurobi Optimization, LLC, *Gurobi Optimizer Reference Manual*. http://www.gurobi.com, 2019.

[14] T. Söderström, *Errors-in-Variables Methods in System Identification*. Cham: Springer, 2018.

[15] S. Skogestad and M. Morari, "Understanding the dynamic behavior of distillation columns." *Ind. Eng. Chem. Res.*, vol. 27, no. 10, pp. 1848–1862, Oct. 1988.

[16] K. E. Häggblom, "Evaluation of experiment designs for MIMO identification by cross-validation," *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 308–313, June 2016.