

Designing experiments for robust-optimization problems: the V_s -optimality criterion

HILLA GINSBURG and IRAD BEN-GAL*

Department of Industrial Engineering, Tel-Aviv University, Tel-Aviv, 69978, Israel
E-mail: bengal@eng.tau.ac.il

Received August 2004 and accepted September 2005

We suggest an experimentation strategy for the robust design of empirically fitted models. The suggested approach is used to design experiments that minimize the variance of the optimal robust solution. The new design-of-experiment optimality criterion, termed V_s -optimal, prioritizes the estimation of a model's coefficients, such that the variance of the optimal solution is minimized by the performed experiments. It is discussed how the proposed criterion is related to known optimality criteria. We present an analytical formulation of the suggested approach for linear models and a numerical procedure for higher-order or nonpolynomial models. In comparison with conventional robust-design methods, our approach provides more information on the robust solution by numerically generating its multidimensional distribution. Moreover, in a case study, the proposed approach results in a better robust solution in comparison with these standard methods.

1. Introduction

In the last three decades, the Taguchi method of robust design has been widely applied to the design of various systems. In many cases, the exact underlying relationship between the design factors and the system response is unknown. Hence, there is a need to design and conduct experiments to gain information. The manner in which these experiments are performed clearly affects the obtained solution for the robust-design problem. Yet, there is no standard method by which to conduct these experiments. As a result, various experimentation strategies are being used that depend on the applied methodologies.

The Taguchi method proposes a set of experimental design matrices (“orthogonal arrays”) to estimate the effects of the design factors and select the combination that yields the highest *Signal-to-Noise* (S/N) ratio (Taguchi, 1978, 1986; Phadke, 1989). Later approaches use a standard canonical approach, in which the experimenter implements the following two-step procedure. First, the experimenter estimates an empirical response model for the unknown system by using conventional experimental matrices, such as factorial designs. These matrices are often based on known Design Of Experiment (DOE) optimality criteria, such as D -optimal designs (e.g., as in Myers and Montgomery (1995)). Second, he/she minimizes a loss function that is based on the estimated model and obtains its opti-

mal solution. The canonical approach is, thus, problematic as long as the estimated model deviates from the “real” unknown model. If the estimated model is noisy, different “optimal” solutions will be obtained for each set of experimental results. We aim to address this problem already at the experimental stage and combine the above two-step procedure in a unified DOE protocol. In particular, we suggest a DOE optimality criterion, termed V_s -optimal, that seeks to minimize the variance of the optimal solution rather than, for example, minimizing the variance of the regression coefficients as done by the D -optimal criterion. The proposed criterion minimizes the variance of the solution by prioritizing the estimation of various model coefficients. Thus, at each experimental stage, it indicates which coefficients should be estimated more accurately with respect to others to obtain a consistent solution.

The area of *robust optimization* (Kouvelis and Yu, 1997; Xu and Albin, 2003) addresses the above problem caused by canonical approach. Similar to the proposed approach, in robust optimization the coefficients of the response model are considered to be unknown and therefore are estimated and treated as random variables. However, the objective of robust optimization is to identify solutions that are insensitive to the estimation errors. A specific objective function is defined for loss minimization by the minimax criterion (Ben-Tal and Nemirovski, 1998). It requires the obtained solution to be small in the worst case, namely over a family of response models for which the coefficients are chosen from predefined confidence intervals.

*Corresponding author

The main difference between the proposed approach and robust optimization is that the former focuses on the *experimental stage*, whereas the latter focuses on the *optimization stage*. The robust-optimization approach uses conventional design matrices (e.g., factorial designs in Xu and Albin (2003)) for the estimation of the model coefficients. Then, based on these estimates, an optimization problem is solved and the output is the robust optimal solution. In comparison, a main output of the proposed approach is an experimental design matrix that minimizes the variance of the optimal solution and often diverges from conventional design matrices. Xu and Albin (2003) indicate that their robust-optimization approach can be contrasted with a sensitivity analysis in which the impact of the estimated objective function on the optimal solution is examined. In this work, we make one-step forward and try to measure the impact of the coefficients' estimates on the variance of the optimal solution. In particular, the proposed DOE optimality criterion can indicate which of the estimates has the highest impact on the optimal solution and can be used to observe how this impact changes over time, as new information is gathered through experiments. Although we do not apply the min-max criterion, we aim to reduce the dispersion of the optimal solution. Our numerical approach, which is applicable also for non-polynomial or high-order models, uses the first two moments of the estimated coefficients for a Monte Carlo simulation (or a parametric bootstrap) to generate a sample of optimal solutions, as exemplified in Section 5.

The rest of this paper is organized as follows. Section 2 presents a review of the related literature. Section 3 provides an example that motivates the use of the proposed approach. Section 4 presents an analytical implementation of the suggested approach for a linear-response model. Section 5 presents a numerical implementation of this approach for nonlinear models. It also compares the suggested approach to conventional robust-design methods. Section 6 suggests a practical framework for an iterative implementation of the suggested approach. Section 7 concludes the paper.

2. Literature review

Following is a literature review on conventional robust-design methods and related DOE optimality criteria.

2.1. Robust-design methods

Taguchi (1978) distinguishes between two types of factors: (i) *control factors* that can be freely selected by the designer; and (ii) *noise factors* that represent the uncontrollable factors, such as environmental conditions. His objective is to design systems that are insensitive to the noise factors (Taguchi, 1986; Phadke, 1989). Taguchi's "nominal-the-best" criterion, which we follow here, seeks for the system's output to be equal to a given target value. The corresponding loss function, $L(Y) = C(Y - T)^2$, implies a quadratic

loss whenever there is a deviation of the response, Y , from the given target value, T , where C is a predetermined cost constant. Taguchi's two-step procedure is based on the assumption that the control factors can be divided into different sets, depending on their effects on the response's mean and on the response's variance (see, e.g., Hunter (1985) and Anderson and Kraber (2002)). The first step is to maximize the S/N ratio. The second step is to adjust the mean value of the response to the given target value by using the so-called "location" factors.

Taguchi's work has been widely analyzed and extended during the last decades. Box and Meyer (1986) suggest a method to estimate the variance of the response and identify the factors that affect it with small nonreplicated designs. Leon *et al.* (1987) introduce the concept of *PerMIA*, a performance measure independent of adjustments. This measure is suggested as a generalization of Taguchi's S/N ratios during the analysis stage. Box (1988) claims that the statistical tools that are used by Taguchi are inefficient and unnecessarily complicated, and suggests working with two ratios based on the response mean and on the response variance independently. Pignatiello (1993) deals with multiple-characteristic functions and introduces priority-based approaches to be used when several quality characteristics are considered.

Other methods for robust design implement the canonical approach, which is often based on the *Response Surface Methodology* (RSM) (see, e.g., Box and Draper (1987) and Myers and Montgomery (1995)). A fitted empirical model enables an experimenter to find the optimal solution by means of traditional optimization techniques. In early RSM-based methods for robust design, the empirical models contained only the control factors (see, e.g., Nair and Pregibon (1988)). Later publications also model the noise factors explicitly, often by polynomial terms that represent random variables. For example, Steinberg and Bursztyn (1994, 1998) conduct a comparative study and demonstrate the importance of modeling the noise factors explicitly. We follow their approach and represent the noise factors by random variables in the model. McCaskey and Tsui (1997) analyze the Taguchi method under various response models and develop a robust-design procedure for dynamic systems with an additive response model. Tsui (1999) further investigates the response-model approach and related loss functions for the dynamic robust-design problem.

The popular *dual-response methodology* solves the robust-design problem by using two response models: one for the response's mean and one for the response's variance (see, e.g., Cho *et al.* (2000)). The analysis is performed in two stages. First, each model is optimized independently. The respective objective functions require a mean close to the target and a minimum variance. Second, the trade-off between the two independent optimal solutions is addressed to obtain the final solution. Myers and Montgomery (1995) suggest two strategies to deal with the introduced trade-off: (i) combining the two models into a single expected loss

function; or (ii) solving a constrained optimization problem that is aimed at minimizing the variance model, subject to constraining the mean to be close enough to the target value. Obviously, if location factors (that affect only the response's mean) exist, an experimenter can concentrate solely on minimizing the variance response model. In Section 5 we compare the proposed approach to the dual-response methodology. Kenett and Zacks (1998) approximate the mean and the variance of a nonlinear response model via Taylor series. Then, they find the robust solution analytically and compare it to solutions that are found by numerical Monte Carlo sampling. In this paper, we implement similar approximation and numerical procedures to address additional experimental design issues. Note that certain statistical aspects should be addressed when implementing response methods in a framework of computer experiments, as indicated in Sacks *et al.* (1989), Sanchez (2000) and Williams *et al.* (2000).

2.2. Related optimal DOE criteria

The Taguchi method implements experimental design by using tables of *orthogonal arrays*, such as the popular L_8 or the L_{16} (Taguchi, 1986). Over the years, these designs have been heavily criticized as being statistically inefficient with respect to known optimality criteria (see, e.g., Leon *et al.* (1987), Box (1988), and Cho *et al.* (2000)). Optimal DOE criteria are often considered in relation to the RSM-based approaches. These criteria aim at minimizing variance-related measures of the response model or of its coefficients (see, e.g., Feodorov (1972), Silvey (1980), and Chang (1994)). Known alphabetic optimal criteria are the A -, Q -, G -, V - and the most popular D -optimality criterion, which minimizes the variance of the estimated coefficients of the empirical model. Sebastiani and Settini (1998) present locally D -optimal experimental designs for a variety of nonlinear models. Atkinson and Donev (1992) discuss the *linear*-optimality, *c*-optimality and the D_s -optimality criteria that are related to our proposed criterion and are further justified by the suggested objective. All these optimality criteria are discussed in Section 5. To the best of our knowledge, the proposed DOE optimality criterion, which is to minimize the variance of the robust solution, has not been suggested explicitly in the literature.

3. Motivating example

We now demonstrate how a sequential estimation of the coefficients of a “real-world” engineering model affects the canonical optimal solution. The example is based on a problem presented by Kenett and Zacks (1998) that deals with a robust design of an RL electrical circuit. The objective is to select the control factors (in this case specifying the resistance (R) and inductance (L) values) such that the current in the circuit (I) is kept at a given target of $T = 10$ Amperes.

The noise factors are the input voltage (V) and frequency (f) that are modeled by Gaussian random variables, where $V \sim N(100, 3)$ and $f \sim N(55, 5/3)$. Kenett and Zacks (1998) consider the following known response model:

$$I(R, L, f, V) = \frac{V}{\sqrt{R^2 + 4\pi^2(fL)^2}}. \quad (1)$$

Let us assume, for illustration purposes, that the response model is known *a priori*, yet, some of its coefficients have to be estimated empirically. In particular, we consider the following response model with two unknown coefficients, a and b , and a noise term ε with a zero mean and a finite variance σ_ε^2 :

$$I(R, L, f, V) = \frac{V}{\sqrt{bR^2 + a(fL)^2 + \varepsilon}}, \quad (2)$$

Note that in following the canonical approach, the experimenter will take the following steps: (i) estimate a and b through experiments; (ii) plug the estimates into the response model of Equation (2) and then estimate the first two moments of I to formulate the loss function $E(\hat{I} - T)^2$; (iii) find the optimal values R^* and L^* that are functions of the estimates and therefore are random variables; and (iv) obtain the loss associated with the estimated optimal solution, $Loss(R^*, L^*)$. The estimated loss in step (ii) is approximated here by a second-order Taylor series expansion. The function itself is too long to be shown here (it does, however, appear in Ginsburg (2003)). It is a nonlinear function, whose arguments are the design factors and the first two moments of the noise factors and the coefficients, i.e.:

$$\begin{aligned} Loss(\hat{I}) &= E(\hat{I} - T)^2 = (\hat{\mu}_I - T)^2 + \hat{\sigma}_I^2 \\ &\approx g(\hat{\mu}_a, \hat{\sigma}_a, \hat{\mu}_b, \hat{\sigma}_b, \mu_V, \sigma_V, \mu_f, \sigma_f, R, L), \end{aligned}$$

where $\hat{\mu}_I$ and $\hat{\sigma}_I$ are the approximated mean and standard deviation of I , the first four arguments in the function $g(\cdot)$ denote the estimated means and standard deviations of the unknown coefficients, the next four arguments denote the given means and standard deviations of the noise factors and the last two arguments denote the control factors. Given this function, one can investigate how the uncertainties regarding the values of the coefficients a and b , as represented by $\hat{\sigma}_a$ and $\hat{\sigma}_b$, affect the estimation of the optimal solution (R^* , L^*) that is obtained by solving simultaneously the equations $R^* = \operatorname{argmin}_R\{E(\hat{I} - T)^2\}$ and $L^* = \operatorname{argmin}_L\{E(\hat{I} - T)^2\}$.

For illustration purposes, let us address a simplified problem by assuming that the experimenter considers a resistance of $R^* = 6 \Omega$. Given the deterministic model Equation (1), the first-order optimality condition yields an optimal value for the inductor in units of Henrys of $L_D^* = 0.023$ H, where the subscript “D” denotes “Deterministic”. Figure 1 presents the Taylor approximation of L^* ($\hat{\sigma}_a$, $\hat{\sigma}_b$), the optimal solution for the inductor, as a function of the standard deviations of the estimated coefficients, $\hat{\sigma}_a$ and $\hat{\sigma}_b$,

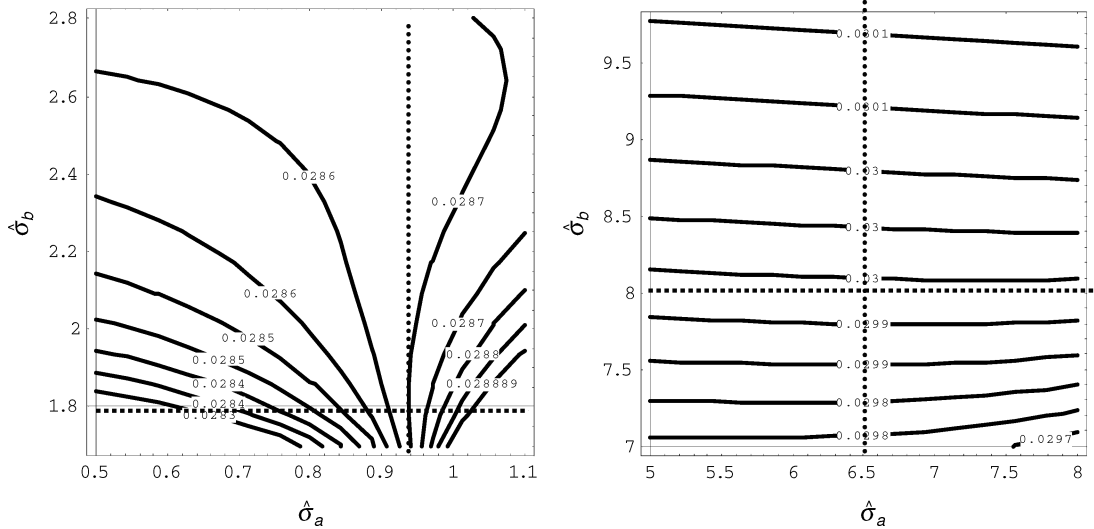


Fig. 1. The optimal solution for the RL circuit, $L^*(\hat{\sigma}_a, \hat{\sigma}_b)$ for a fixed $R = 6 \Omega$ as a function of $\hat{\sigma}_a$ and $\hat{\sigma}_b$ in two regions of the parameter uncertainty.

in the vicinity of $R^* = 6 \Omega$. The right-hand graph shows the value of L^* for one section of the experimental region, in which the values of the estimated standard deviations are relatively large. It shows a given solution, $L^* \approx 0.03H$, that is obtained for $\hat{\sigma}_a = 6.5$ and $\hat{\sigma}_b = 8$. Note that a reduction of one unit in $\hat{\sigma}_b$ while holding $\hat{\sigma}_a$ fixed results in a better optimal solution, L^* , which is closer to the deterministic solution. However, a reduction of one unit in $\hat{\sigma}_a$ while holding $\hat{\sigma}_b$ fixed does not change the value of the optimal solution. In general, for relatively large values of standard deviations, changes in $\hat{\sigma}_a$ do not result in a value change of L^* , whereas changes in $\hat{\sigma}_b$ do affect the optimal robust solution (Ginsburg, 2003). Thus, in order to reduce the variance of the estimated robust solution, $V(L^*)$ in this case, the experimenter should focus on estimating b as accurately as possible, rather than investing in a better estimation of a . The left-hand graph shows another section of the experimental region, in which the values of the estimated standard deviations are smaller. In this region the opposite situation occurs: to reduce the variance of L^* , the experimenter should focus on estimating a as accurately as possible rather than on estimating b . This phenomenon is also true for R^* , which is not presented here. Note that these conclusions cannot be anticipated *a priori* from the loss function itself.

This example shows that the optimal solution might depend on various coefficients' estimates in a manner that is not necessarily uniform or consistent. As new information is gathered through experiments, some of the model coefficients become more influential for a robust estimation of the optimal solution. On this ground, we formulate the suggested V_s -optimality criterion that takes into account the optimal solution for both linear and nonlinear models, as presented in the next sections.

4. A linear-response model

In this section we formulate the V_s -optimality criterion by implementing the suggested approach to a linear model given by:

$$Y(\mathbf{x}, \mathbf{Z}) = \beta_0 + \boldsymbol{\beta}'\mathbf{x} + \boldsymbol{\alpha}'\mathbf{Z} + \mathbf{x}'\boldsymbol{\Gamma}\mathbf{Z} + \varepsilon, \quad (3)$$

where Y is the unknown response, \mathbf{x} is a $(k \times 1)$ vector of the coded significant control factors, i.e., $\mathbf{x}' = (x_1, x_2, \dots, x_k)$, \mathbf{Z} is an $(m \times 1)$ vector of the significant noise factors, coded such that $E[Z_i] = 0, i = 1, \dots, m$, β_0 is a scalar, $\boldsymbol{\beta}'$ is a $(1 \times k)$ row vector of the control-factor's coefficients; $\boldsymbol{\alpha}'$ is a $(1 \times m)$ vector of the noise-factor's coefficients; $\boldsymbol{\Gamma}$ is a $(k \times m)$ matrix of the control-factor-by-noise-factor interactions that link the two types of factors and enable a reduction in the noise-factors' effects, and ε is a noise term with a mean of zero and a finite variance σ_ε^2 . Such a linear model can be obtained by traditional experimental designs. In most cases, we denote random variables by capital letters and use the bold font to represent vectors or matrices.

Given the response model, we now express the first-order optimality condition explicitly and then formulate the V_s -optimality criterion.

4.1. Optimization

The optimization stage of the robust-design problem is performed with respect to the introduced loss function, $L(Y) = E(Y - T)^2$, which depends on the mean and the variance of Y :

$$\begin{aligned} E(Y(\mathbf{x}, \mathbf{Z})) &= \beta_0 + \boldsymbol{\beta}'\mathbf{x}; \\ V(Y(\mathbf{x}, \mathbf{Z})) &= (\boldsymbol{\alpha}' + \mathbf{x}'\boldsymbol{\Gamma})\boldsymbol{\Sigma}(\boldsymbol{\alpha} + \boldsymbol{\Gamma}'\mathbf{x}) + \sigma_\varepsilon^2, \end{aligned} \quad (4)$$

where Σ is assumed to be the known ($m \times m$) variance-covariance matrix of \mathbf{Z} . Accordingly, the expected loss function is given by:

$$\begin{aligned} L(Y) &= E(Y - T)^2 = V(Y) + (E(Y) - T)^2 \\ &= V(Y) + E(Y)^2 - 2TE(Y) + T^2 \\ &= \boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha} + \boldsymbol{\alpha}'\Sigma\Gamma'\mathbf{x} + \mathbf{x}'\Gamma\Sigma\boldsymbol{\alpha} + \mathbf{x}'\Gamma\Sigma\Gamma'\mathbf{x} \\ &\quad + (\beta_0 + \boldsymbol{\beta}'\mathbf{x})(\beta_0 + \boldsymbol{\beta}'\mathbf{x}) - 2T(\beta_0 + \boldsymbol{\beta}'\mathbf{x}) + T^2. \end{aligned} \quad (5)$$

The first-order optimality condition for Equation (5) yields the optimal robust solution:

$$\Rightarrow \mathbf{x}^* = (\Gamma\Sigma\Gamma' + \boldsymbol{\beta}\boldsymbol{\beta}')^{-1} \cdot [\boldsymbol{\beta}(T - \beta_0) - \Gamma\Sigma\boldsymbol{\alpha}]. \quad (6)$$

Since the coefficients of the response model are empirically estimated, \mathbf{x}^* is a random function, $\mathbf{x}^* = \mathbf{g}(\hat{\boldsymbol{\theta}}_L)$, where $\mathbf{g}(\cdot)$ is a ($k \times 1$) vector function of the set of the linear-model's estimates, $\hat{\boldsymbol{\theta}}_L = \{\hat{\beta}_0, \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}, \hat{\Gamma}\}$ and, thus, \mathbf{x}^* is a random variable itself. Using a Taylor series expansion around the coefficients' estimates, one obtains an approximated robust solution for the system:

$$\begin{aligned} \mathbf{x}^* &= \mathbf{g}(\hat{\boldsymbol{\theta}}_L) \approx \mathbf{g}(\boldsymbol{\theta}_L) + \left[\frac{\partial \mathbf{g}(\boldsymbol{\theta}_L)}{\partial \boldsymbol{\theta}_L} \right] \cdot (\hat{\boldsymbol{\theta}}_L - \boldsymbol{\theta}_L) \\ &= \mathbf{g}(\boldsymbol{\theta}_L) + \mathbf{J} \cdot (\hat{\boldsymbol{\theta}}_L - \boldsymbol{\theta}_L), \end{aligned} \quad (7)$$

where \mathbf{J} is the ($k \times p$) Jacobian matrix of \mathbf{x}^* , which is estimated by $\partial \mathbf{g}(\hat{\boldsymbol{\theta}}_L) / \partial \boldsymbol{\theta}_L$, and $p = (k + 1)(m + 1)$ is the number of coefficients in the model.

4.2. The V_s -optimality criterion

At this stage, the proposed DOE criterion can be formulated explicitly. This DOE stage involves two sub-problems: (i) defining the appropriate design region; and (ii) selecting the optimal design matrix \mathbf{F}^* that satisfies $\mathbf{F}^* = \arg \min_{\{\mathbf{F}\}} \{V(\mathbf{x}^*)\}$ within the selected design region.

The new design region in which the control factors are selected can be defined as a confidence region around the estimated optimal solution \mathbf{x}^* . Such a definition is consistent with \mathbf{x}^* being a random vector, whose realizations lie within a confidence region. This approach is in agreement with the common robust-optimization approach (see, e.g., Ben-Tal and Nemirovski (1998)). The design region can be either spherical, such as $\mathbf{x} \in \mathbf{x}^* \pm 3 \cdot \sqrt{V(\mathbf{x}^*)}$, or a hypercube, such as $-1 \leq \mathbf{I}_{k+m} \mathbf{u} \leq 1$, where \mathbf{I}_{k+m} is an identity matrix with dimensions $(k + m) \times (k + m)$ and \mathbf{u} is a vector representing the k control and the m noise factors.

Our next step is to search in the design region for the optimal design matrix, \mathbf{F}^* , that minimizes the variance of the optimal solution, $V(\mathbf{x}^*)$. The required design matrix has dimensions of $(n \times p)$, where p is the number of parameters and $n \geq p$ is the number of experiments. Taking the variance of the expression in Equation (7) yields the following $(k \times k)$ variance-covariance matrix that we aim to

minimize:

$$\begin{aligned} V(\mathbf{x}^*) &\approx \left[\frac{\partial \mathbf{g}(\hat{\boldsymbol{\theta}}_L)}{\partial \boldsymbol{\theta}_L} \right] V(\hat{\boldsymbol{\theta}}_L) \left[\frac{\partial \mathbf{g}(\hat{\boldsymbol{\theta}}_L)}{\partial \boldsymbol{\theta}_L} \right]' \\ &= \mathbf{J} ((\mathbf{F}'\mathbf{F})^{-1} \sigma_\varepsilon^2) \mathbf{J}'. \end{aligned} \quad (8)$$

This is a fundamental result of our proposed criterion that can be obtained analytically only for linear response models when Equation (7) exists in a closed form.

4.3. An illustrative example

Let us look at a simplified linear model that contains a single control factor ($k = 1$), a single noise factor ($m = 1$) and a single interaction term:

$$Y(x, Z) = \beta_0 + b_1 x + a_1 Z + x\gamma Z + \varepsilon. \quad (9)$$

Following Equation (6), the optimal solution is given by:

$$x^* = \frac{b_1(T - \beta_0) - \gamma a_1}{\gamma^2 + b_1^2}. \quad (10)$$

The Jacobian matrix of x^* in this case, with respect to $\hat{\boldsymbol{\theta}}_L = \{\beta_0, b_1, a_1, \gamma\}$, is the following ($1 \times p$) row vector:

$$\begin{aligned} \mathbf{J} &= \left[-\frac{b_1}{\gamma^2 + b_1^2}, \quad \frac{T - \beta_0}{\gamma^2 + b_1^2} - \frac{2b_1(b_1(T - \beta_0) - a_1\gamma)}{(\gamma^2 + b_1^2)^2}, \right. \\ &\quad \left. -\frac{\gamma}{\gamma^2 + b_1^2}, \quad -\frac{a_1}{\gamma^2 + b_1^2} - \frac{2\gamma(b_1(T - \beta_0) - a_1\gamma)}{(\gamma^2 + b_1^2)^2} \right]. \end{aligned} \quad (11)$$

Let us further simplify this vector by considering an on-target product, $\beta_0 = T$, an equal effect of the control and the noise factors, $b_1 = a_1$, and a respectively 50% effect of the interaction, e.g., $\gamma = a_1/2 = b_1/2$. Now, the Jacobian matrix of x^* is further simplified to:

$$\begin{aligned} \mathbf{J} &= \left[-\frac{4}{5a_1}, \quad \frac{16}{25a_1}, \quad -\frac{2}{5a_1}, \quad -\frac{12}{25a_1} \right] \\ &= \left[2t, \quad -\frac{8t}{5}, \quad t, \quad \frac{6t}{5} \right], \end{aligned} \quad (12)$$

where $t \equiv -2/5a_1$. Note that the partial derivatives (the components in the Jacobian, matrix) are substantially different, e.g., the following square ratios of the partial derivatives in Equation (12) are [4, 2.56, 1, 1.44]. Such a vector implies that when designing the next experiment, one should invest four times more effort in estimating β_0 than in estimating a_1 . This makes sense due to our assumption that the system is on-target. The next-in-importance parameters are (in descending order) the control factor's coefficient, b_1 , the interaction coefficient, γ , and the noise factor's coefficient, a_1 .

4.4. Designing the next experiment

The optimal design matrix can now be found by standard nonlinear programming methods: via Equation (8) aiming

at:

$$\mathbf{F}^* = \min_{\mathbf{F}} \{V(\mathbf{x}^*)\} = \min_{\mathbf{F}} \left\{ \det[\mathbf{J}((\mathbf{F}'\mathbf{F})^{-1}\sigma_\varepsilon^2)\mathbf{J}'] \right\}.$$

The variables are the $n(k+m)$ independent elements of the design matrix that represent the values of the control and the noise factors; the optimization is constrained within the design region, which is proportional to $V(\mathbf{x}^*)$. Formulating the first-order optimality condition for Equation (8), yields a set of high-degree equations that often cannot be solved analytically. For example, even the simple model in Equation (9) yields a set of fourth-degree equations that needs to be solved simultaneously to obtain the optimal solution. We implement a Matlab-based numerical optimization procedure, which is based on a standard library of quasi-Newton routines (with polynomial complexity). Since this optimization problem is not necessarily convex, the obtained solutions are often local minima that depend on the starting point of the search. Therefore, we use several starting points for the search to try to avoid local minima. In this example, we start with 100 designs (a larger number of starting points did not result in a different solution) that are randomly chosen (by a uniform distribution) within a hypercube design region, plus another starting point set at a D -optimal design. The best solution is then chosen among all the obtained results. The search algorithm has a polynomial complexity, which implies that it can be used to solve larger problems in general. If the number of variables is large, one can also use a simpler objective function, such as minimizing the trace of $\mathbf{F}'\mathbf{F}$ instead of its determinant, thus, following the A -optimality criterion.

To illustrate the difference between our optimal design, \mathbf{F}^* , and the conventional D -optimal design, denoted here by \mathbf{D}^* , let us consider the model in Equation (9) with $\beta_0 = 8$, $b_1 = 0.18$, $a_1 = -0.1$, $\gamma = 0.5$ and a target $T = 3$. Executing the numerical optimization routine with the smallest $n > p$ results in the optimal design matrix:

$$\mathbf{F}^* = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -0.75 & -0.75 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & -0.94 & 0.94 \\ 1 & 1 & -0.17 & -0.17 \end{bmatrix}. \quad (13)$$

The values are coded between (-1) and $(+1)$ and rounded to two digits after the decimal point. A comparative D -optimal design with a similar number of experiments ($n = 5$) is composed of four factorial points and one replication, e.g.:

$$\mathbf{D}^* = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}. \quad (14)$$

As expected by the definitions of the criteria, the variance of the robust solution which is based on the D -optimal design is larger (almost twice in this case) than the variance of the robust solution which is based on the V_s -optimality criterion, i.e., $V(\mathbf{x}^*|\mathbf{D}^*) = 78.42\sigma_\varepsilon^2$ whereas $V(\mathbf{x}^*|\mathbf{F}^*) = 39.42\sigma_\varepsilon^2$.

Further numerical studies (Ginsburg, 2003) indicate that the V_s -optimality criterion often yields a design matrix which differs from traditionally used D -optimal designs.

5. A second-order (or a higher-order) model

In this section we consider nonlinear response models. In RSM, such models are often fitted iteratively, at a stage when the linear model cannot capture the curvature information in the vicinity of the optimal solution (Myers and Montgomery, 1995).

Consider a second-order response model of the following form:

$$Y(\mathbf{x}, \mathbf{Z}) = \beta_0 + \beta' \mathbf{x} + \mathbf{x}' \mathbf{B} \mathbf{x} + \alpha' \mathbf{Z} + \mathbf{x}' \Gamma \mathbf{Z} + \varepsilon, \quad (15)$$

where \mathbf{B} is a $(k \times k)$ matrix of the control factors' quadratic terms and interactions. Apply the first-order condition for optimality to the approximated loss function:

$$\begin{aligned} \frac{\partial E(Y - T)^2}{\partial \mathbf{x}} &= 0, \\ \Rightarrow [\Gamma \Sigma \alpha + \beta_0 \beta - T \beta] + [\Gamma \Sigma \Gamma' + \beta \beta' + 2\beta_0 \mathbf{B} - 2T \mathbf{B}] \mathbf{x} \\ &+ 3\mathbf{x} \beta' \mathbf{B} \mathbf{x} + 2\mathbf{B} \mathbf{x} \mathbf{x}' \mathbf{B} \mathbf{x} = 0, \end{aligned} \quad (16)$$

which represents a system of k equations of a third-order degree that, in general, cannot be solved analytically. Since there is no closed form for \mathbf{x}^* , as for Equation (7) in the linear case, we apply a numerical approach to obtain the V_s -optimal design. The new experimental region is a confidence region around the optimal solution $\mathbf{x}^* = \mathbf{g}(\hat{\theta}_Q)$, where $\hat{\theta}_Q$ is the set of the estimates of the Quadratic-model coefficients, $\hat{\theta}_Q = \{\hat{\beta}_0, \hat{\alpha}, \hat{\beta}, \hat{\mathbf{B}}, \hat{\Gamma}\}$. However, since the vector function $\mathbf{x}^* = \mathbf{g}(\hat{\theta}_Q)$ for quadratic or higher-order models is unknown analytically, one should use numerical methods to estimate the optimal solution and its variance. Here, we implement a Monte Carlo sampling or *parametric bootstrap* (Efron, 1979) to generate several sets of the model's coefficients, assuming that they are normally distributed. Then, we numerically calculate this optimal solution, \mathbf{x}^* , for each generated set of the estimated coefficients. The procedure repeats itself to provide many solutions and yield an empirical distribution of the optimal solution. For illustration purposes, let us now present the suggested procedure, which can be implemented in a similar manner to that for higher-order models.

5.1. The unknown (“real”) model

Let us consider the following response model which represents a “real” and unknown (to the experimenter) system. It contains two control factors (x_1 and x_2) and two noise factors (z_1 and z_2), all coded between (-1) and $(+1)$:

$$\begin{aligned}
 Y &= \beta_0 + [\beta_1 \ \beta_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [\alpha_1 \ \alpha_2] \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \\
 &+ [x_1 \ x_2] \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [x_1 \ x_2] \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \\
 &+ \varepsilon \\
 &= 5 + [-2 \ 4] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [1 \ -5] \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \\
 &+ [x_1 \ x_2] \begin{bmatrix} 1 & -7 \\ -7 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [x_1 \ x_2] \begin{bmatrix} -10 & 18 \\ -15 & 14 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \\
 &+ \varepsilon.
 \end{aligned} \tag{17}$$

The noise factors are assumed to be controllable during the experiments and independent, having an identity covari-

ance matrix. The underlying model enables us to measure how close is the “real” optimal solution, which is derived from Equations (16) and (17), to the solution of the proposed approach, as well as to solutions of other traditional methods. Otherwise, the model in Equation (17) is only used to generate the simulated experiments.

Figure 2(a–c) plots the response model of Equation (17) as a function of the two control factors, x_1 and x_2 . The noise factors are set to their mean values that are equal to zero. Figure 2(a) shows the mean value of the response function. The dashed lines in the upper-right and lower-left corners show the desired target which is $T = -10$. Note that within the region of interest (indicated by the inner square) there is no point at which the expected value of the response is equal to the target value. In order for the response to approach this target value, the control factors should be set at the edges of the region of interest, i.e., around $(-1, -1)$ or $(+1, +1)$; Fig. 2(b) shows the variance of the response function, which is minimized along the upper-left to the lower-right diagonal (dashed); Fig. 2(c) shows the associated loss function that balances the response’s mean and

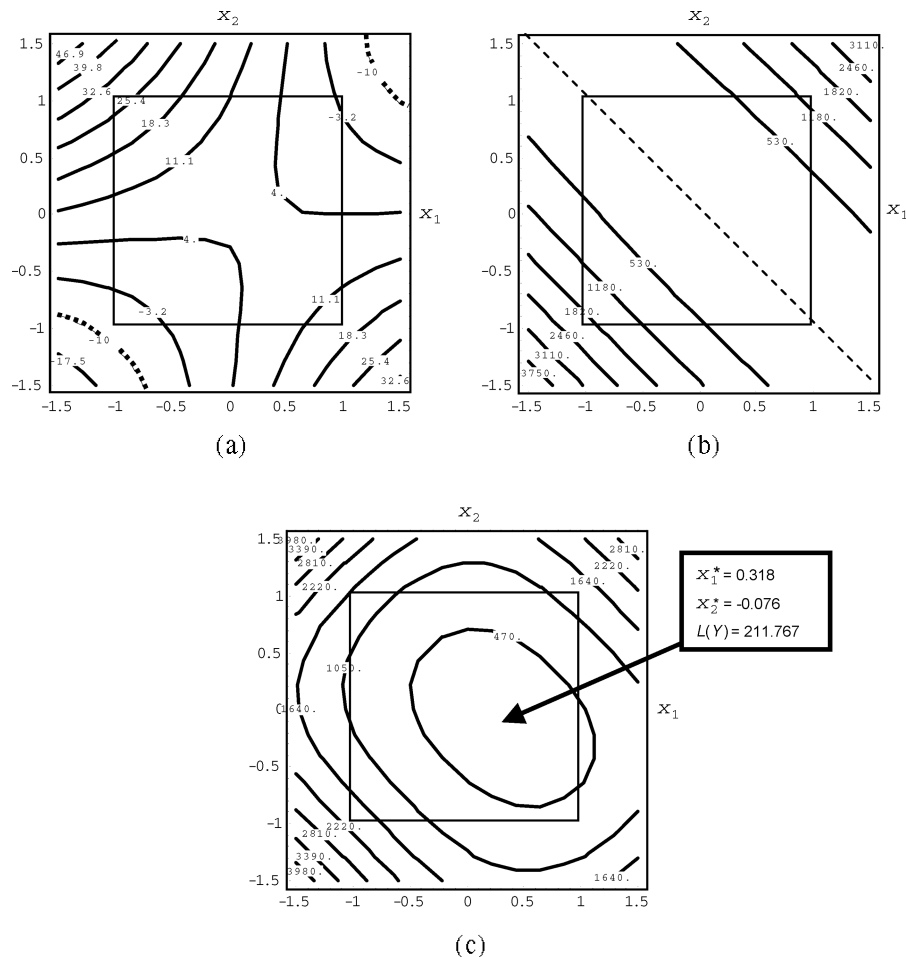


Fig. 2. (a) The response’s mean; (b) the response’s variance; and (c) the associated loss function of the “real” system with respect to the control factors x_1 and x_2 and the target value of $T = -10$. Noise factors are set to their mean values that are equal to zero.

variance. The loss reaches its minimum close to the center of the interest region at $\mathbf{x}_a^* = (0.318, -0.076)$ (where the subscript “a” stands for “accurate”) with an associated loss value of 211.77.

5.2. Numerical procedure steps

Given the underlying function, we outline the proposed numerical procedure. The procedure is further exemplified in the next section.

- Step 1.* Estimate the unknown model by replicated experiments (in this example the underlying model of Equation (17) is used to generate the experimental observations); estimate the means and the variances of the model’s coefficients based on the replicated observations.
- Step 2.* Generate new sets of the model’s coefficients via Monte Carlo sampling or parametric bootstrap.
- Step 3.* Compute the optimal solution(s) numerically for each generated set of the model’s coefficients.
- Step 4.* Plot the robust solution(s) as a function(s) of the control factors. For nonlinear models, several robust solutions might exist, thus, cluster them. The center of each cluster represents one optimal solution.
- Step 5.* Compute the variance of each solution from its cluster numerically.
- Step 6.* Design the next experiment to minimize the variance of the optimal solution(s). This step is further addressed below.

Next, the procedure steps are illustrated by a running example.

5.3. Estimating the model and the coefficients’ moments by replicated experiments

In this step, we create an initial data reservoir based on the coefficients’ estimates that are obtained from w replica-

tions of an r -treatment experiment. Practically, we choose $w = 10$ and $r = 50$, thus, we replicate a 50-treatment experiment for ten times (the optimal selection of w and r will be a subject of future research). The design is presented in the Appendix. The purpose of the replicated experiment is to estimate the means and the standard deviations of the model’s coefficients. These estimates are used in later steps to generate new sets of coefficients. Table 1 presents the coefficients’ estimates for each of the ten replicated experiments, as well as their sample mean-values and sample standard deviations.

5.4. Generating new sets of coefficients via Monte Carlo sampling

Once the initial reservoir set is obtained, we re-generate new sets of the model’s coefficients by Monte Carlo sampling (or a parametric bootstrap). In particular, we use the inverse transform method for a Gaussian random variable with the estimated means and standard deviations as its arguments to generate an additional, say 100 sets of coefficients (one simple way to generate these values is by the “NORMINV” function in Excel). Table 2 presents ten of the 100 generated sets of coefficients (each set in a row). The generation is immediate and one can easily generate more sets of parameters.

5.5. Computing the optimal solutions

For each of the generated sets of coefficients, we compute the optimal solution, \mathbf{x}^* . The solutions are derived numerically by solving the set of equations given by Equation (16). The solutions for 15 (out of the 100) sets of coefficients are presented in Table 3. We calculate the loss values for each solution, $L(\mathbf{x}^*)$, as shown in the last column. Once again, using Matlab procedures requires a polynomial time to numerically solve these nonlinear equations.

Table 1. The coefficients’ estimates, sample means and sample standard deviations estimated from the 50-treatment experiment

| Coefficient | Estimated values | | | | | | | | | | | |
|-------------|------------------|-----------|-----------|------------|------------|----------|----------|----------|---------------|---------------|---------------|---------------|
| | β_0 | β_1 | β_2 | α_1 | α_2 | B_{11} | B_{22} | B_{12} | Γ_{11} | Γ_{12} | Γ_{21} | Γ_{22} |
| 1 | 5.918 | -1.55 | 4.314 | 0.865 | -5.248 | 1.835 | 0 | -14.133 | -9.804 | 17.967 | -15.097 | 14.337 |
| 2 | 6.263 | -1.519 | 4.197 | 0.536 | -5.494 | 1.663 | 0 | -14.027 | -10.609 | 18.043 | -15.019 | 13.625 |
| 3 | 5.087 | -1.592 | 3.872 | 1.307 | -4.687 | 1.314 | 1.771 | -13.987 | -10.037 | 18.184 | -15.177 | 14.312 |
| 4 | 6.18 | -2.532 | 4.321 | 1.386 | -5.544 | 0 | 1.795 | -14.069 | -9.916 | 18.098 | -14.958 | 13.725 |
| 5 | 5.163 | -2.095 | 3.7 | 0.972 | -5.222 | 0 | 2.563 | -13.902 | -10.406 | 17.641 | -14.976 | 13.986 |
| 6 | 4.997 | -1.839 | 3.894 | 0.93 | -4.784 | 0 | 2.64 | -13.951 | -10.767 | 18.28 | -15.8 | 14.56 |
| 7 | 6.702 | -1.318 | 3.56 | 0.995 | -5.515 | 0 | 0 | -14.035 | -10.713 | 18.456 | -14.631 | 13.953 |
| 8 | 5.713 | -2.198 | 4.2 | 0.638 | -4.704 | 0 | 2.233 | -14.494 | -10.391 | 18.412 | -14.625 | 13.751 |
| 9 | 4.223 | -2.467 | 4.587 | 0.864 | -4.744 | 1.256 | 2.943 | -13.875 | -9.625 | 17.048 | -14.735 | 12.962 |
| 10 | 4.89 | -1.987 | 3.969 | 1.049 | -5.267 | 0 | 2.881 | -13.88 | -9.684 | 17.889 | -14.687 | 13.8 |
| Average | 5.514 | -1.91 | 4.061 | 0.954 | -5.121 | 1.517 | 2.404 | -14.035 | -10.195 | 18.002 | -14.971 | 13.901 |
| Stdv | 0.764 | 0.416 | 0.316 | 0.261 | 0.356 | 0.278 | 0.483 | 0.182 | 0.434 | 0.415 | 0.353 | 0.45 |

Table 2. Ten out of the 100 generated sets of the model's coefficients based on the sample estimates

| Run | Generated values | | | | | | | | | | | |
|-----|------------------|-----------|-----------|------------|------------|----------|----------|----------|---------------|---------------|---------------|---------------|
| | β_0 | β_1 | β_2 | α_1 | α_2 | B_{11} | B_{22} | B_{12} | Γ_{11} | Γ_{12} | Γ_{21} | Γ_{22} |
| 1 | 5.133 | -1.961 | 3.972 | 1.658 | -5.184 | 1.573 | 2.267 | -13.85 | -10.086 | 17.498 | -14.701 | 14.86 |
| 2 | 5.76 | -1.822 | 3.863 | 0.964 | -5.284 | 2.18 | 2.471 | -13.678 | -10.974 | 17.438 | -14.72 | 14.02 |
| 3 | 6.786 | -1.848 | 4.132 | 0.941 | -5.586 | 1.819 | 2.48 | -13.991 | -10.357 | 18.417 | -14.792 | 14.143 |
| 4 | 5.35 | -1.852 | 3.429 | 0.896 | -5.511 | 1.267 | 2.816 | -13.964 | -9.842 | 18.454 | -15.021 | 14.623 |
| 5 | 5.197 | -2.189 | 3.997 | 0.52 | -5.011 | 1.694 | 2.835 | -13.76 | -10.592 | 18.543 | -15.944 | 14.345 |
| 6 | 5.529 | -1.523 | 3.862 | 1.016 | -5.538 | 1.703 | 2.353 | -14.093 | -9.901 | 17.934 | -14.858 | 12.649 |
| 7 | 4.721 | -1.82 | 4.141 | 1.006 | -4.382 | 1.37 | 2.083 | -14.222 | -9.693 | 17.548 | -15.325 | 14.124 |
| 8 | 6.136 | -1.974 | 3.958 | 1.128 | -5.784 | 1.332 | 1.636 | -13.911 | -10.846 | 17.81 | -15.377 | 13.905 |
| 9 | 6.726 | -1.533 | 4.304 | 1.009 | -5.502 | 1.676 | 1.889 | -14.052 | -10.458 | 17.747 | -15.14 | 14.272 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 100 | 6.062 | -1.638 | 3.959 | 1.221 | -5.023 | 1.466 | 2.979 | -14.02 | -10.067 | 17.537 | -14.907 | 14.167 |

5.6. Plotting and clustering the optimal solutions

When it is unknown how many optimal solutions might exist for Equation (16), one can draw a scatter plot of the solutions to estimate how many optimal \mathbf{x}^* solutions potentially exist by using the number of blobs (clusters). We plot only 90% of the solutions in order to neutralize outliers' effects that can result from the randomized sampling. In this example, we choose those solutions with the lowest loss values although alternative criteria, such as the distance of a solution from the center of each blob, can be also applied. If the partition for blobs is not trivial, one can use known clustering methods (e.g., see Duda *et al.* (2001)) to assign solutions to the blobs. An optimal solution is defined as the center of gravity of each blob. Since all the solutions are equally weighted in this case, the center of gravity is obtained by simply averaging all the solutions' coordinates.

Table 3. Fifteen out of the 100 simulation-based optimal solutions

| Run | X_1^* | X_2^* | Loss value |
|-----|---------|---------|------------|
| 1 | 0.320 | -0.045 | 216.735 |
| 2 | 0.288 | -0.056 | 241.011 |
| 3 | 0.305 | -0.062 | 271.667 |
| 4 | 0.309 | -0.042 | 225.698 |
| 5 | 0.299 | -0.089 | 218.954 |
| 6 | 0.297 | -0.046 | 234.646 |
| 7 | 0.289 | -0.076 | 205.584 |
| 8 | 0.332 | -0.061 | 248.537 |
| 9 | 0.302 | -0.064 | 272.114 |
| 10 | 0.297 | -0.050 | 248.531 |
| 11 | 0.266 | -0.050 | 244.912 |
| 12 | 0.270 | -0.070 | 220.806 |
| 13 | 0.265 | -0.065 | 229.662 |
| 14 | 0.308 | -0.070 | 284.773 |
| 15 | 0.309 | -0.089 | 206.068 |

5.7. Computing the variance of each optimal solution

Finally, we numerically compute the variance, $V(\mathbf{x}^*)$, for each of the blobs. The diagonal elements of $V(\mathbf{x}^*)$ are calculated by the variance of each column's numbers, and the off-diagonal elements are calculated by the covariance terms between pairs of columns in Table 3.

Figure 3 presents a scatter plot on the control-factors' plane. It shows the 90% of the optimal solutions with the lowest loss values. The scatter plot presents a single blob, implying that there is a single \mathbf{x}^* solution for the system, as expected from the quadratic unknown function in Equation (17). The center of gravity of this blob, which is equal to (0.305, -0.071), is our recommended solution at this stage, with an associated loss value of 211.83. The accurate optimal solution $\mathbf{x}_a^* = (0.318, -0.076)$ with a loss of 211.77, which is based on the "real" (unknown to the experimenter) model of Equation (17), is denoted by the \times sign. Note that the accurate solution is located within the blob and is close to its center, implying that the proposed procedure provides a good solution in this case.

5.8. A comparison with traditional methods

The above robust-design problem was also solved by the following traditional methods (see the literature review in Section 2):

- (i) Taguchi's (original) method: picking the control factors' combination that yields the largest S/N ratio.
- (ii) a variation of Taguchi's two-step procedure which is based on both the S/N ratios and the expected values. Practically, since the "real" function does not include location factors, we chose a solution that yields the highest S/N ratio, unless there exists an alternative solution, where for a unit loss of the S/N ratio, the expected value is closer to the target in two units or more.

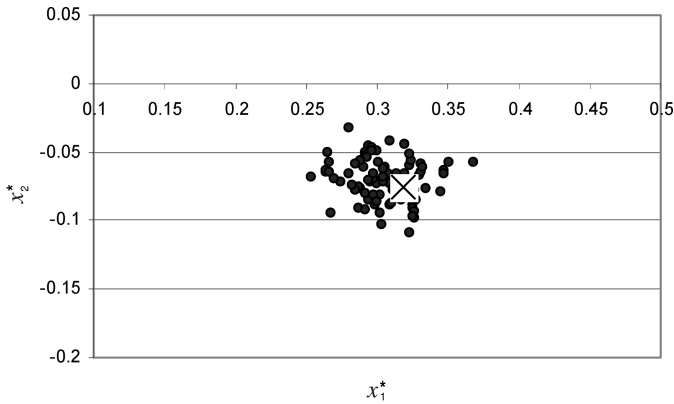


Fig. 3. The 90% of the \mathbf{x}^* solutions with the lowest loss values, as obtained by the Monte Carlo sampling procedure. The “x” sign shows the location of the “real” (unknown) optimal solution.

- (iii) a mean response approach: identifying within the design region those solutions whose expected value is as close as possible to the given target value (i.e., either solutions whose expected value equals the target value or solutions that are at the edges of the design region) and among them choosing the solution that yields the smallest variance.
- (iv) a variance response approach: finding the solution that minimizes the variance response model.
- (v) an extended mean response approach: a similar method to (iii), with the distinction that the expected value must be equal to the target value, even if this implies picking solutions that lie beyond the interest region.
- (vi) a dual-response method: constructing a response model for the loss function, which is based on both the expected value of the response and its variance and optimizing this model explicitly (Myers and Montgomery, 1995).

Table 4 demonstrates the proposed solutions. It compares the “real” optimal solution (had the model in Equation (17) been known) to the solutions of the suggested approach and of the traditional methods. The “real” optimal solu-

tion is presented in column 2; the proposed approach is presented in column 3; and the above-mentioned methods are presented in columns 4–9 (labeled by (i)–(vi) respectively). In order to maintain a fair comparison, in each of the investigated methods we use the same coefficients’ estimates, as obtained from the ten replicated experiments. For each method we present the best obtained solution (x_1^* and x_2^* in the second and third rows) along with their associated loss values (fourth row). In the last row, the “real” loss value of 211.77 (second column) is considered as a reference loss and therefore obtains a score of 100%. The loss values of the other methods are scored as a percentage of the “real” loss and are higher than 100%.

Note from Table 4, that the suggested approach yields a solution which is very close to the “real” optimal solution. Moreover, it obtains the smallest loss among all other traditional methods (with a score of 100.03%). The dual response method of Myers and Montgomery (in the last column), that minimizes the loss model explicitly, yields a similar solution to the proposed one (with a score of 100.04% which is an insignificant difference). Note, however, that the optimal solution in this case is a point-estimate for \mathbf{x}^* and it does not include additional information on the spread of \mathbf{x}^* , as obtained by the cluster of the 90% of the calculated solutions with the lowest loss values obtained by the suggested approach (in Fig. 3). The variance response method results in a slightly larger loss value (109.29%), in contrast to the mean-response model that results in a significantly higher loss (633.24%). This observation depends, of course, on the specific-model’s characteristics; in systems with an on-target value within the region of interest the mean response model might obtain better results than the variance response method. With regard to Taguchi’s methods, the variation of the two-step approach yields a relatively small loss value (118.53%), which results in a much larger loss (727.69%). This observation is in agreement with Steinberg and Burszty (1994). The extended mean model yields the largest loss (1199.31%) in this case. Note from Fig. 2 that the expected value is equal to the target only outside the region of interest (upper-right and lower-left dashed contour lines). However, the feasible solutions on these contour lines have very large

Table 4. Comparative study of robust solutions and their associated loss values, as obtained by the different robust-design methods

| Value | Method | | | | | | | | |
|-----------------------------------|-------------------------|-------------------|-----------------------|----------------|------------------|---------------------|-------------------------|-----------------|--|
| | | | Taguchi’s methods | | Response methods | | | | |
| | “Real” unknown solution | Proposed approach | (i) Original one-step | (ii) Two steps | (iii) Mean model | (iv) Variance model | (v) Extended mean model | (vi) Loss model | |
| x_1^* | 0.318 | 0.305 | −1 | 0 | 1 | 0.496 | −1.009 | 0.304 | |
| x_2^* | −0.076 | −0.071 | 1 | 0 | 1 | −0.274 | −1.247 | −0.069 | |
| Loss value | 211.77 | 211.83 | 1541 | 251 | 1341 | 231.436 | 2539.74 | 211.844 | |
| Percentage of “real” solution (%) | 100 | 100.03 | 727.69 | 118.53 | 633.24 | 109.29 | 1199.31 | 100.04 | |

variances that result in a large loss, as can be seen in Fig. 2(b) and Fig. 2(c).

5.9. Designing the next experiment

Following the same arguments as in the linear case, the next design region can be defined by a confidence region around the optimal solution. The new DOE criterion aims at minimizing the variance of the optimal solution, $V(\mathbf{x}^*)$, which depends on the estimators of the Jacobian matrix's terms. Since a numerical procedure is used to obtain \mathbf{x}^* for nonlinear models, the relations between the Jacobian terms and the variance of the optimal solution cannot be expressed by a closed-form equation, as for the linear case in Equation (8). Therefore, the search for the coefficients' estimates with the largest effects on $V(\mathbf{x}^*)$ is also derived numerically.

We suggest the following procedure to locate those coefficients with the largest effect on the optimal solution. Namely, at each stage of this procedure, the experimenter picks one (or a combination) of the coefficients' estimators and reduces its (their) estimated standard deviation(s) by one-half or by any other selected ratio. The selected ratio can depend on the current variance of the estimator, as well as on the required number of experiments to reduce the solution's variance. Then, the experimenter repeats the Monte Carlo sampling procedure (that is presented in Table 2) to obtain additional \mathbf{x}^* solutions. Consider the following illustrative example; the standard deviation of β_0 in Table 1 is reduced by one-half from 0.764 to 0.382. Since the random generation of the coefficients' values is based on their sample standard deviations, such a reduction implies that the

generated coefficients will be closer to their mean values. This procedure leads to a reduced variance of the optimal solutions, denoted by $V_r(\mathbf{x}^*)$, in comparison with the original variance of the optimal solution, $V(\mathbf{x}^*)$, which was obtained prior to the reduction. At the end of this procedure, the experimenter obtains a set of new $V_r(\mathbf{x}^*)$ values, each of which is associated with one (or with a combination) of the coefficients' estimates. A comparison among these values with respect to the original $V(\mathbf{x}^*)$ can indicate which of the coefficients' estimates has the largest influence on the variance of the optimal solution.

Table 5 illustrates the above procedure. It presents the obtained optimal solutions, x_1^* and x_2^* , and their estimated variance-covariance components: $V(x_1^*)$, $V(x_2^*)$ and $Cov(x_1^*, x_2^*)$ (all multiplied by a factor of 10^3). The first row presents the original values prior to any reduction in the standard deviations of the coefficients. The next 12 rows present the optimal solutions and their variances following a 50% reduction in the standard deviation of each corresponding coefficient. For example, the second row presents the obtained values for the case where the standard deviation of β_0 is reduced by one-half. The last three rows in Table 5 present the optimal solutions and their reduced variance components following a simultaneous reduction in the standard deviations of some combinations of coefficients. Row 14 presents a simultaneous reduction of 50% in the variances of the two most influential estimators, β_1 and β_2 . Row 15 repeats the same simulated experiment when reducing 25% rather than 50% of the variances of β_1 and β_2 . Finally, row 16 repeats the analysis following a simultaneous reduction of 25% in the variances of β_2 and α_1 . The left part of the table presents the absolute values of the

Table 5. The reductions in the variance components of $V(\mathbf{x}^*)$ resulting from the reductions in the standard deviations of the coefficients' estimators

| Reduction | Absolute values | | | | | As percentage of the original values | | |
|---|-----------------|---------|---------------------------------|---------------------------------|--|--------------------------------------|------------|---------------------|
| | x_1^* | x_2^* | $V(x_1^*)$ ($\times 10^3$) | $V(x_2^*)$ ($\times 10^3$) | $Cov(x_1^*, x_2^*)$ ($\times 10^3$) | $V(x_1^*)$ | $V(x_2^*)$ | $Cov(x_1^*, x_2^*)$ |
| No reduction | 0.305 | -0.071 | 0.476 | 0.235 | -0.035 | | | |
| β_0 (50%) | 0.305 | -0.072 | 0.473 | 0.232 | -0.030 | 0.993 | 0.987 | 0.870 |
| β_1 (50%) | 0.304 | -0.071 | 0.297 | 0.226 | 0.012 | 0.625 | 0.960 | -0.354 |
| β_2 (50%) | 0.305 | -0.070 | 0.473 | 0.150 | -0.013 | 0.993 | 0.640 | 0.361 |
| α_1 (50%) | 0.305 | -0.071 | 0.467 | 0.201 | -0.053 | 0.981 | 0.856 | 1.526 |
| α_2 (50%) | 0.305 | -0.071 | 0.372 | 0.219 | -0.077 | 0.781 | 0.933 | 2.208 |
| B_{11} (50%) | 0.305 | -0.071 | 0.419 | 0.229 | -0.017 | 0.881 | 0.976 | 0.475 |
| B_{22} (50%) | 0.305 | -0.071 | 0.479 | 0.229 | -0.037 | 1.006 | 0.973 | 1.066 |
| B_{12} (50%) | 0.305 | -0.071 | 0.472 | 0.237 | -0.032 | 0.992 | 1.010 | 0.927 |
| Γ_{11} (50%) | 0.305 | -0.071 | 0.472 | 0.224 | -0.043 | 0.990 | 0.951 | 1.234 |
| Γ_{12} (50%) | 0.305 | -0.071 | 0.463 | 0.232 | -0.041 | 0.973 | 0.986 | 1.168 |
| Γ_{21} (50%) | 0.305 | -0.071 | 0.474 | 0.234 | -0.034 | 0.996 | 0.995 | 0.981 |
| Γ_{22} (50%) | 0.305 | -0.071 | 0.477 | 0.238 | -0.030 | 1.002 | 1.012 | 0.858 |
| β_1, β_2 ($2 \times 50\%$) | 0.304 | -0.070 | 0.288 | 0.135 | 0.043 | 0.606 | 0.576 | -1.218 |
| β_1, β_2 ($2 \times 25\%$) | 0.305 | -0.071 | 0.368 | 0.176 | 0.011 | 0.773 | 0.751 | -0.304 |
| β_2, α_1 ($2 \times 25\%$) | 0.305 | -0.071 | 0.468 | 0.166 | -0.031 | 0.983 | 0.705 | 0.881 |

variance components of the optimal solutions, whereas the right part of the table presents these values as percentages of the original $V(\mathbf{x}^*)$ value. For example, the most influential (single) estimator when aiming to reduce the variance of x_1^* is β_1 , since a reduction of 50% in the standard deviation of β_1 results in a reduction of 37.5% in $V(x_1^*)$ (the reduced variance of x_1^* is equal to 0.297×10^{-3} which equals 62.5% of the original value). Similarly, a reduction of 50% in the standard deviation of β_2 results in a reduction of 36% in $V(x_2^*)$. These results are expected, yet, note that the next-importance coefficients are α_2 for x_1^* and α_1 for x_2^* and not the interaction coefficients, as could have been theoretically anticipated. A simultaneous reduction of 50% in the variances of β_1 and β_2 (row 14) results in the most significant reduction in the solution's variances (around 40%) with a negative correlation between x_1^* and x_2^* (-1.218). A smaller reduction of 25% in the variances of these estimators (row 15), which is probably cheaper in practice, yields a smaller reduction in the solution's variances (around 25%), yet with a lower correlation between x_1^* and x_2^* . A simultaneous reduction of 25% in the variances of β_2 and α_1 (last row) is presented to illustrate the flexibility of the simulated procedure in the selection of experiments. The proposed numerical procedure is computationally tractable similar to other Monte Carlo simulations. It can be further extended by applying simple factorial experiments to analyze the effects of single or combined reductions in the estimators' variances. Note that some entries in the right part are larger than 100% as a result of the randomization effects in the simulations.

Once the relations between the variance of the coefficients' estimators and the variance of the optimal solution have been evaluated numerically, the next design matrix can be constructed. Note again, that in the linear case these relations are expressed by the closed-form expression of Equation (8), which enables us to design the next experiment by direct optimization methods, as the design in Equation (13). Instead, in the nonlinear case, we suggest the use of the information gathered by the Monte Carlo simulations and associate it with known DOE-optimality criteria. Three such implementations are suggested next and include the linear-optimality, the c -optimality and the D_s -optimality criteria. These criteria are related to the V_s -optimality criterion and can fit well the suggested sequential framework that is proposed in Section 6. Moreover, known procedures for obtaining (locally) linear-, c -, and D_s -optimal designs for nonlinear models exist in the literature and can be used, as discussed next.

The linear-optimality criterion seeks to minimize a weighted average of the variances of the coefficients' estimates. Thus, the known A -optimality criterion is a special case of the linear-optimality criterion that assigns equal weights to all the variance estimates (Atkinson and Donev, 1992). Numerical experiments, such as the one in Table 5, provide ways to assign unequal weights to different variance's estimates. One way to obtain these weights is to nor-

malize the variance of the optimal solutions (using the last 3 columns in Table 5) by the maximum simulated value, and then define the weights as being inversely proportional to these normalized values. Thus, coefficients' estimates that contribute more to the variance of the optimal solution will obtain higher weights implying larger efforts to estimate them in the next experiment.

The objective according to the c -optimality criterion is to estimate a linear combination of the model's coefficients, $\mathbf{c}'\boldsymbol{\theta}$, with a minimum variance. Thus, the c -optimality criterion minimizes $V[\mathbf{c}'\hat{\boldsymbol{\theta}}] \propto V[\mathbf{c}'\mathbf{M}^{-1}(\boldsymbol{\xi})\mathbf{c}]$, where \mathbf{c} is a $(p \times 1)$ vector and $\mathbf{M}(\boldsymbol{\xi}) = \mathbf{F}'\mathbf{F}/n$ is the information matrix of the chosen design, $\boldsymbol{\xi}$. If \mathbf{c} is taken to be $f(\mathbf{x}_0)$ (the response at a specific design point \mathbf{x}_0) this criterion is reduced to minimizing the variance of the prediction of the response at \mathbf{x}_0 . In this work, we are interested in $\mathbf{x}_0 = \mathbf{x}^*$, which is the optimal solution of the model. If the location of \mathbf{x}^* is known prior to the experiment, one can minimize this prediction by repeatedly performing all the following experiments at \mathbf{x}^* . This procedure often results in a singular optimum design, which is noninformative about all the other aspects of the model and the design region (Atkinson and Donev, 1992). In the sequential framework of the proposed approach (presented in Section 6), the coefficients, and therefore \mathbf{x}^* , are approximated in each experimental stage. Thus, at each stage one can use $\mathbf{c} = f(\mathbf{x}^*)$ to obtain a locally c -optimal design. Locally c -optimum experiments for both linear and nonlinear models are discussed in the literature (Atkinson and Donev, 1992; Atkinson *et al.*, 1992; Kitsos *et al.*, 1988). Extended procedures are proposed in the case of nonlinear models and in the case where the objective is to minimize the variance of a nonlinear combination of the model's coefficients. The nonlinear function is often expanded to a Taylor series in a manner which is similar to our proposed procedures. Since c -optimal designs might be singular, we propose to regularize the information matrix through additions of small multiples of the identity matrix as in Atkinson and Donev (1992).

Another related DOE optimality criterion is the D_s -optimality, which is often used for model selection when the interest is in estimating a subset s of the coefficients as precisely as possible (Feodorov, 1972). In our case, the experimenter can use the D_s -optimality criterion with respect to the s most influential coefficients on the variance of the optimal solutions, thus, avoiding the heuristics that are required by the linear-optimality criterion to obtain the weights of the coefficients' estimates. For example, let us consider the control factors in model Equation (17) and the experimental results in Table 5. These experiments reveal that the most influential estimators, with respect to the variance of the optimal solution, are β_1 and β_2 . Thus, these are the two coefficients of interest in the next D_s -optimal experiment. Accordingly, we follow the procedure in Atkinson and Donev (1992) to obtain the required D_s -optimal experiment. First, the control factors in Equation (17) are

divided into two groups:

$$E(Y) = f'(x)\beta = \underbrace{\beta_1 x_1 + \beta_2 x_2}_{f'_1(x)\beta_1} + \underbrace{\beta_0 + B_{11}x_1^2 + B_{12}x_1x_2 + B_{22}x_2^2}_{f'_2(x)\beta_2}, \quad (18)$$

with corresponding rows in the design matrix of $f'_1(x) = (x_1, x_2)$, $f'_2(x) = (1, x_1^2, x_1x_2, x_2^2)$. The information matrix of a design ξ with n experiments, $\mathbf{M} = \mathbf{F}'\mathbf{F}/n$, is divided respectively:

$$\mathbf{M}(\xi) = \begin{pmatrix} \mathbf{M}_{11}(\xi) & \mathbf{M}_{12}(\xi) \\ \mathbf{M}'_{12}(\xi) & \mathbf{M}_{22}(\xi) \end{pmatrix}. \quad (19)$$

The covariance matrix for the least-squares estimates of β_1 (or β_2) is \mathbf{M}_{11}^{-1} (or, respectively, \mathbf{M}_{22}^{-1}) which is the $(s \times s)$ left upper (or respectively, $(p - s) \times (p - s)$ right-lower) submatrix of \mathbf{M}^{-1} . Finally, the scaled standardized variance of a D_s -optimal design ξ^* should satisfy the equation:

$$f'(x)\mathbf{M}^{-1}(\xi^*) - f'_2(x)\mathbf{M}_{22}^{-1}(\xi^*)f_2(x) \leq s, \quad (20)$$

with equality at the points of support of the design. Applying this criterion to the control factors in Equation (17) with β_1 and β_2 being the two coefficients of interest, and, for example, $n = 8$, results in the following optimal design:

$$\mathbf{D}_s^* = \begin{bmatrix} +1 & +1 & +1 & +1 & +1 & +1 \\ -1 & +1 & +1 & +1 & -1 & +1 \\ -1 & -1 & +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & +1 & -1 & +1 \\ +1 & +1 & +1 & +1 & +1 & +1 \\ -1 & +1 & +1 & +1 & -1 & +1 \\ +1 & -1 & +1 & +1 & -1 & +1 \\ -1 & -1 & +1 & +1 & +1 & +1 \end{bmatrix}.$$

Further mathematical procedures can be used to improve the D_s -optimal design, as discussed in Silvey (1980) and Pazman (1986).

6. A suggested experimental framework

In this section we propose a sequential procedure, which is not necessarily optimal, for the implementation of the suggested approach. The procedure fits other DOE schemes that are suggested when the optimal design depends on the unknown coefficients of the model. An example for such a general scheme is given in Atkinson and Donev (1992): (i) start with a preliminary estimate based on past knowledge or experiments: either the prior point estimate, θ_0 , or a prior distribution for θ ; (ii) linearize the model by a Taylor series expansion; (iii) find the optimum design for the linearized

model; and (iv) execute several trials of the optimum design for the linearized model. If the new estimate of θ is sufficiently accurate, stop the process. Otherwise, repeat step (ii) for the new estimate.

Note that the optimality of a sequential experimentations scheme depends on a precise selection of the number of experiments, n , in each stage. This problem, however, is left open despite the vast literature on experimental design. Several heuristic approaches have been suggested to select the number of the experiments in a sequential design scheme. For example, Box *et al.* (1978) give a general recommendation, known as the “25% rule of the thumb”, according to which, not more than one-quarter of the experimental budget should be used in the first design stages. Atkinson and Donev (1992) follow the same approach, while using \sqrt{n} experiments in the first stage, when n is the total number of predetermined experiments. Other approaches determine n by defining required confidence intervals for the estimates of the model at each stage. The optimal solution to this problem has to address the trade-off between the costs of future experiments versus the expected value of the information to be obtained. An example for such an optimal scheme is suggested in Ben-Gal and Caramanis (2002). There, the gain in information obtained through experiments, is measured by the reduction in the entropy of the estimators of the coefficients. The optimal experimentation strategy is obtained by a stochastic dynamic-programming scheme. In this paper, we do not attempt to solve this problem and rely on known approaches to determine the number of experiments in the various stages.

The proposed sequential scheme contains six stages that enable us to divide a given experimental budget and to refine the computed solution. These stages are presented in Fig. 4 and described next.

Stage 0: Definition of the system. The experimenter obtains the inputs for the procedure: (i) a general description of the system, its response and target; (ii) candidate control and noise factors (types, ranges of variation, levels); (iii) the objective function (e.g., minimizing a loss function, as considered in this paper); and (iv) other relevant information. The experimenter determines the stopping condition, which can be based, for example, on the marginal changes in the loss function, on the number of conducted iterations, or on budget-related criteria.

Stage 1: Screening experiment. A screening experiment is conducted over the region of interest in order to identify the significant control and noise factors and estimate their effects (the noise factors are assumed to be controllable during this experimental phase). Once the significant factors are identified and if budget permits, a larger initial design can be carried out, like the one presented in the Appendix. A linear-response model can be fitted at this stage and used for further analytical optimization.

Stage 2: Finding an analytical predictive model. A predictive model for the response is obtained by using RSM. The model is linear following the first step or of a higher-order

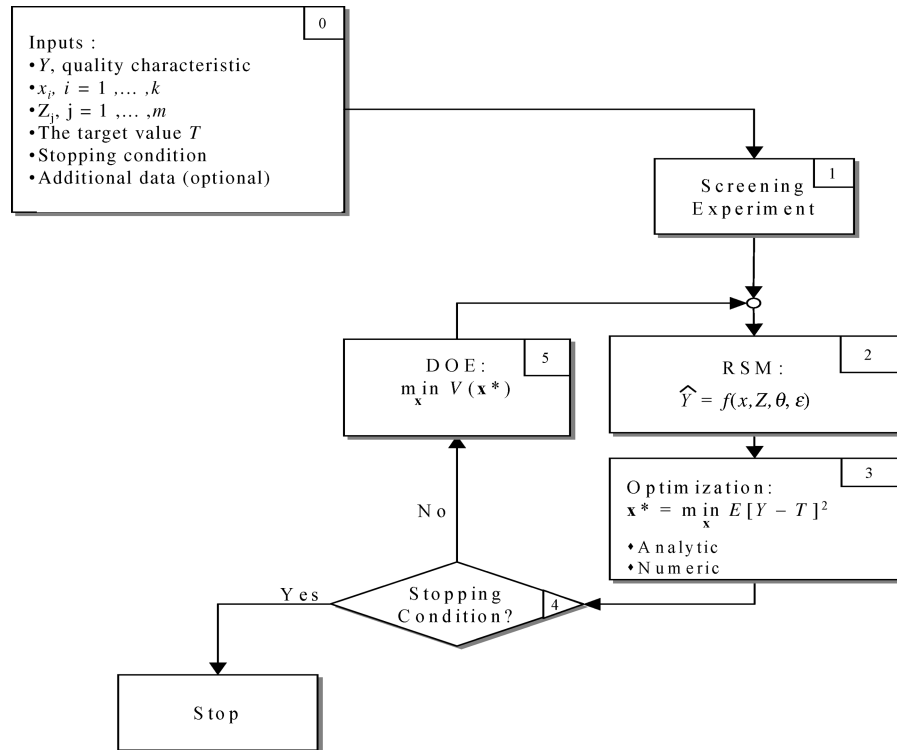


Fig. 4. A proposed sequential scheme for the suggested approach.

following stage 5 that is described below. The noise factors are modeled as random variables.

Stage 3: Optimization. The optimal robust solution (minimizing the objective function) is obtained analytically (Section 4) or numerically (Section 5).

Stage 4: Checking the stopping condition. The stopping condition is checked. If it is satisfied, the procedure ends. Otherwise, a new iteration begins at stage 5.

Stage 5: Designing the next experiment. The V_s -optimality criterion is used to define a new experiment within the selected design region. The experiment's design matrix is obtained numerically either from a linear model (Section 4) or from a higher-order model (Section 5) by associating it with other criteria (linear-optimality, c -optimality or D_s -optimality, as discussed in Section 5). Following the new experiment, the experimenter updates the model to include new significant terms and better estimated coefficients and returns to stage 2.

The number of experiments that are allocated to each stage is defined by using one of the known methods discussed above. The computation time to solve the procedure is polynomial as mentioned in Sections 4 and 5.

7. Conclusions

In this paper we consider the problem of robust design in empirically fitted models. We suggest the V_s -optimality

DOE criterion, which aims to minimize the variance of the optimal solution. The suggested criterion prioritizes the estimations of various model coefficients in order to obtain a consistent optimal solution. It is associated with known optimal design criteria and provides further justifications for their implementations.

We propose an analytical implementation of the suggested approach for a linear-response model, as well as a numerical procedure for nonlinear models. Various examples illustrate some potential advantages of the suggested approach. One advantage is that it provides a multidimensional distribution of the optimal solution, \mathbf{x}^* . This multidimensional distribution carries important operational information with respect to a point estimate for the optimal solution that is obtained from traditional methods. Moreover, most of the obtained information can be gathered through computer re-samplings that are relatively cheap and fast. Another advantage is that the suggested approach enables us to design, in advance, the next experiment and to estimate more accurately the most influential coefficients in various models, including high-order ones. Also, the proposed approach can be used to solve various problems that are considered in the field of robust optimization, with the proviso that the proposed approach is focused on the experimental stage rather than on the optimization stage. A potential research direction is to implement the suggested V_s -optimality criterion within a robust-optimization framework to ensure that the obtained solutions are not only

consistent but also immune against inaccuracies in the model's estimates. Another challenge is to integrate an optimal selection of the number of experiments in each stage of the sequential framework that is proposed in Section 6.

The suggested approach can be associated with a broader scientific dilemma regardless of the considered objective function (a robust design in this case): "should one investigate experimental efforts to learn as much as possible about the system; including, for example, design regions that are distant from the optimal solution? or should one concentrate on the optimal solution itself and direct the investigation such that this solution is obtained more consistently?" We believe that the latter approach might be useful for practitioners who are less interested in the theoretical investigation of the response model and prefer to focus on the solution itself.

References

- Anderson, M.J. and Kraber, S.L. (2003) Using design of experiments to make processes more robust to environmental and input variations. *Paint and Coating Industry*, **19**(2), 52–60.
- Atkinson, A.C., Chaloner, K., Herzberg, A.M. and Juritz, J. (1992) Optimum experimental designs for properties of a compartmental model. *Biometrics*, **49**(1), 325–337.
- Atkinson, A.C. and Donev, A.N. (1992) *Optimum Experimental Designs*, Oxford University Press, New York.
- Ben-Gal, I. and Caramanis, M. (2002) A stochastic dynamic programming framework for an efficient adaptive design of experiments. *IIE Transactions on Quality and Reliability*, **34**(12), 1087–1100.
- Ben-Tal, A. and Nemirovski, A. (1998) Robust convex optimization. *Mathematics of Operation Research*, **23**, 769–805.
- Box, G. (1988) Signal-to-noise ratios, performance criteria and transformation. *Technometrics*, **30**(1), 1–17.
- Box, G. and Draper, N. (1987). *Empirical Model Building and Response Surfaces*, Wiley, New York.
- Box, G., Hunter, W.G. and Hunter, J.S. (1978) *Statistics for Experimenters*, Wiley, New York.
- Box, G. and Meyer, R.D. (1986) Dispersion effects from fractional designs. *Technometrics*, **28**(1), 19–27.
- Chang, S.I. (1994) Some properties of multiresponse D-optimal designs. *Journal of Mathematical Analysis and Applications*, **184**(2), 256–262.
- Cho, B.R., Kim, Y. J., Kimbler, D.L. and Phillips, M.D. (2000) An integrated joint optimization procedure for robust and tolerance. *International Journal of Production Research*, **38**(10), 2309–2325.
- Duda, R.O., Hart, P.E. and Stork, D.G. (2001). *Pattern Classification*. Wiley, New York, NY.
- Efron, B. (1979) Bootstrap methods: another look at the jackknife. *Annals of Statistics*, **7**(1), 1–26.
- Feodorov, V.V. (1972) *Theory of Optimal Experiments*, Academic Press, New York, NY.
- Ginsburg, H. (2003) An approach for designing experiments in robust design problems, Unpublished M.S.C. thesis, Industrial Engineering, Tel Aviv University. Available at <http://www.eng.tau.ac.il/~bengal/>.
- Hunter, J.S. (1985) Statistical design applied to product design. *Journal of Quality Technology*, **17**(4), 210–221.
- Kenett, R.S. and Zacks, S. (1998) *Modern Industrial Statistics: Design and Control of Quality and Reliability*, Brooks/Cole, Pacific Grove, CA.
- Kitsos, C.P., Titterton, D.M. and Torsney, B. (1988) An optimal design problem in rhythmometry. *Biometrics*, **44**, 657–671.
- Kouvelis, P. and Yu, G. (1997) *Robust discrete optimization and its applications*, Kluwer, Dordrecht, The Netherlands.
- Leon, R.V., Shoemaker, A.C. and Kacker, R.N. (1987) Performance measures independent of adjustment. *Technometrics*, **29**(1), 253–285.
- McCaskey, S.D. and Tsui, K.L. (1997) Analysis of dynamic robust design experiments. *International Journal of Production Research*, **35**(6), 1561–1574.
- Myers, R.H. and Montgomery, D.C. (1995) *Response Surface Methodology Process and Product Optimization Using Designed Experiments*. Wiley, New York, NY.
- Nair, V.N. and Pregibon, D. (1988) Analyzing dispersion effects from replicated factorial experiments. *Technometrics*, **30**, 247–257.
- Pazman, A. (1986) *Foundation of Optimum Experimental Design*, Reidel, Dordrecht. The Netherlands.
- Phadke, M.S. (1989) *Quality Engineering Using Robust Design*, Prentice Hall, Englewood Cliffs, NJ.
- Pignatiello, J. (1993) Strategies for robust multiresponse quality engineering. *IIE Transactions*, **25**(3), 5–15.
- Sacks, J., Welch, W.J., Mitchell, T.J. and Wynn, H.P. (1989) Design and analysis of computer experiments. *Statistical Science*, **4**, 409–435.
- Sanchez, S.M. (2000) Robust design: seeking the best of all possible worlds, in *Proceeding of the 2000 Winter Simulation Conference*, Piscataway, NJ, pp. 69–76.
- Sebastiani, P. and Settimi, R. (1998) First-order optimal designs for non-linear models. *Journal of Statistical Planning and Inference*, **74**, 177–192.
- Silvey, S.D. (1980) *Optimal Design: An Introduction to the Theory for Parameter Estimation*, Chapman and Hall, London, UK.
- Steinberg, D.M. and Bursztyn, D. (1994) Dispersion effects in robust-design experiments with noise factors. *Journal of Quality Technology*, **26**(1), 12–20.
- Steinberg, D.M. and Bursztyn, D. (1998) Noise factors, dispersion effects and robust design. *Statistica Sinica*, **8**, 67–85.
- Taguchi, G. (1978) Off-line and on-line quality control systems, in *Proceeding of the International Conference on Quality*, Tokyo, Japan.
- Taguchi, G. (1986) *Introduction to Quality Engineering*, Asian Productivity Organization, Tokyo, Japan.
- Tsui, K.L. (1999) Modeling and analysis of dynamic robust design experiments. *IIE Transactions*, **31**, 1113–1122.
- Williams, B.J., Santner, T.J. and Notz, W.I. (2000) Sequential design of computer experiments to minimize integrated response functions. *Statistica Sinica*, **10**, 1133–1152.
- Xu, D. and Albin, S.L. (2003) Robust optimization of experimentally derived objective functions. *IIE Transactions*, **35**, 793–802.

Appendix 1

The implemented 50-treatments experiment shown in Table A1 contains: (i) 36 factorial combinations (three levels of each of the two control factors and two levels of each of the two noise factors); (ii) a two-level factorial design of the control factors at each (equal) level of the noise factors (12 combinations); and (iii) two center points. The design depends on the experimental budget available at various stages of the sequential procedure that is presented in Section 6.

Table A1. The 50-treatment experimental design that is used to create the input data reservoir (and replicated for ten times)

| <i>Treatment</i> | <i>Factor values</i> | | | |
|------------------|----------------------|-------|-------|-------|
| | x_1 | x_2 | z_1 | z_2 |
| 1 | -1 | -1 | -1 | -1 |
| 2 | -1 | -1 | -1 | 1 |
| 3 | -1 | -1 | 1 | -1 |
| 4 | -1 | -1 | 1 | 1 |
| 5 | -1 | 0 | -1 | -1 |
| 6 | -1 | 0 | -1 | 1 |
| 7 | -1 | 0 | 1 | -1 |
| 8 | -1 | 0 | 1 | 1 |
| 9 | -1 | 1 | -1 | -1 |
| 10 | -1 | 1 | -1 | 1 |
| 11 | -1 | 1 | 1 | -1 |
| 12 | -1 | 1 | 1 | 1 |
| 13 | 0 | -1 | -1 | -1 |
| 14 | 0 | -1 | -1 | 1 |
| 15 | 0 | -1 | 1 | -1 |
| 16 | 0 | -1 | 1 | 1 |
| 17 | 0 | 0 | -1 | -1 |
| 18 | 0 | 0 | -1 | 1 |
| 19 | 0 | 0 | 1 | -1 |
| 20 | 0 | 0 | 1 | 1 |
| 21 | 0 | 1 | -1 | -1 |
| 22 | 0 | 1 | -1 | 1 |
| 23 | 0 | 1 | 1 | -1 |
| 24 | 0 | 1 | 1 | 1 |
| 25 | 1 | -1 | -1 | -1 |
| 26 | 1 | -1 | -1 | 1 |
| 27 | 1 | -1 | 1 | -1 |
| 28 | 1 | -1 | 1 | 1 |
| 29 | 1 | 0 | -1 | -1 |
| 30 | 1 | 0 | -1 | 1 |
| 31 | 1 | 0 | 1 | -1 |
| 32 | 1 | 0 | 1 | 1 |
| 33 | 1 | 1 | -1 | -1 |
| 34 | 1 | 1 | -1 | 1 |
| 35 | 1 | 1 | 1 | -1 |
| 36 | 1 | 1 | 1 | 1 |
| 37 | -1 | -1 | -1 | -1 |
| 38 | -1 | 1 | -1 | -1 |
| 39 | 1 | -1 | -1 | -1 |
| 40 | 1 | 1 | -1 | -1 |
| 41 | -1 | -1 | 0 | 0 |
| 42 | -1 | 1 | 0 | 0 |
| 43 | 1 | -1 | 0 | 0 |
| 44 | 1 | 1 | 0 | 0 |
| 45 | -1 | -1 | 1 | 1 |
| 46 | -1 | 1 | 1 | 1 |
| 47 | 1 | -1 | 1 | 1 |
| 48 | 1 | 1 | 1 | 1 |
| 49 | 0 | 0 | 0 | 0 |
| 50 | 0 | 0 | 0 | 0 |

Biographies

Hilla Ginsburg is an industrial engineer. She holds B.Sc. (2001) and M.Sc. (2003) degrees in industrial engineering and an M.B.A (2005) degree, all from Tel-Aviv University. Hilla has been lecturing in Tel Aviv University and in the Open University, where she serves as a coordinator. Her research interests include design-of-experiments and design-of-manufacturing-processes.

Irad Ben-Gal is a Senior Lecturer at the Department of Industrial Engineering in Tel-Aviv University. He is the head of the Computer In-

tegrated Manufacturing (CIM) lab at Tel-Aviv University. He holds a B.Sc. (1992) degree from Tel-Aviv University, M.Sc. (1996) and Ph.D. (1998) degrees from Boston University. Irad is a member of the Institute for Operations Research and Management Sciences (INFORMS) and the Institute of Industrial Engineers (IIE). Irad has worked for several years in various industrial organizations. His research interests include quality control, design-of-experiments, testing procedures, and application of information theory to industrial and bioinformatics problems.

Contributed by the Process Optimization Department