Research Article

Quality and Reliability Engineering International

(www.interscience.wiley.com) DOI: 10.1002/qre.1031

Published online 26 June 2009 in Wiley InterScience

Optimization-seeking Experimentations: Design of an *RL*-circuit Via the *Vs*-optimality Criterion

Hilla Ginsburg and Irad Ben-Gal*†

In this paper we explore the Vs-optimality criterion that was proposed in Ginsburg and Ben-Gal (*IIE Trans*. 2006; 38:445–461) as a new design-of-experiment (DOE) alphabetic optimality criterion. The Vs-optimality criterion seeks to minimize the variance of the optimal solution of an empirically fitted model. We show that the Vs-optimality citerion is well related to known alphabetic DOE criteria. However, it focuses on an 'optimization-seeking' experimental approach rather than an 'information-seeking' approach, which is often adopted by traditional optimality criteria. We illustrate the differences between these two approaches by a detailed example of a robust design of an *RL*-circuit. Copyright © 2009 John Wiley & Sons, Ltd.

Keywords: design-of-experiment; optimization; alphabetic optimality criteria; experimentations; sequential design

1. Introduction

What is the difference between designing experiments to gain *general knowledge about a system* vs designing experiments to gain knowledge about the *optimal solution of a system*? Or in other words, what is the difference between an experimentation strategy that is triggered by an 'information-seeking' approach (a 'scientist's viewpoint') to an experimentation strategy that is triggered by an 'optimization-seeking' approach (an 'engineer's viewpoint')? These types of questions are discussed in Reference[1] that proposed an experimentation strategy for robust design of empirically fitted models. A new design-of-experiment (DOE) alphabetic optimality criterion, termed as *V*s-optimal, was implemented there to minimize the variance of the optimal solution of a system. This optimal solution was defined as the solution that yields the most robust system's output.

The *V*s-optimality criterion is different from other DOE-optimality criteria, such as the *D*-optimality criterion that seeks to minimize the variance of the estimated system parameters. We claim that the Vs-optimality criterion is triggered by an 'optimization-seeking' approach rather than by an 'information-seeking' one. The difference between these approaches has a clear practical effect on the experimented design. An 'information-seeking' approach often leads to spread experimental points that are located towards the edges of the experimental region, as in *D*-optimal designs. An 'optimization-seeking' approach, however, often leads to concentrated experimental points in the vicinity of the optimal solution (e.g. in the vicinity of the best operational setting) of the system. Naturally, at the beginning of the design process, when there is almost no information about a system, the 'optimization-seeking' approach has little meaning, since the vicinity of the optimal setting for the system is unknown. In such a case, there is a need to gather some information first via experiments and only then, after some knowledge is obtained about the system, move on to the 'optimization-seeking' approach.

Another practical difference between the two approaches is found with respect to the estimation of the different parameters of the system. The *Vs*-optimality criterion prioritizes the model's coefficients and points into the coefficients that mostly affect the variance of the optimal solution and, thus, should be in the focus in each experimental stage. As seen below, the priority among different parameters might change along the experimental process, as new information about the investigated system is gathered.

As indicated in the example below, the proposed DOE-optimality criterion enables to combine, within the same experimental framework, 'information-seeking' as well as 'optimization seeking' principles. It is shown that when almost nothing is known about the parameters of the system, the proposed criterion leads to design matrices that are close to those obtained by *D*-optimal designs. However, as information is gathered, the same criterion will lead to experimental points that are closer to the area of the optimal solution of the system. This integration stands in contrast to the conventional canonical approach that implements the following two-step procedure. First, the experimenter estimates an empirical response model for the unknown system by using the conventional experimental matrices such as factorial designs. Second, the experimenter minimizes a loss function that is based on the estimated

Department of Industrial Engineering, Tel-Aviv University, Tel-Aviv 69978, Israel
**Correspondence to: Irad Ben-Gal, Department of Industrial Engineering, Tel-Aviv University, Tel-Aviv 69978, Israel.*
†E-mail: bengal@eng.tau.ac.il

model and obtains its optimal solution as if the system is deterministic. 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 are obtained for each set of experiments[1, 2].

In this paper we shortly describe the *Vs*-optimality criterion that was introduced in[1] and then we focus on illustrating the related concepts of 'optimization-seeking' designs. By relying on a simple example of a robust design of an *RL*-circuit, we illustrate how the optimal solution of the system depends on the unknown parameters. Using a sequential experimentation process, we show how the optimal solution is being updated as more information is gained about the system parameters. We indicate which of the system parameters has the highest effect on the optimal solution during different experimentation stages and how this impact changes over time as new information is gathered. The main contribution of the paper is, thus, in illustrating the 'optimization-seeking' approach by using a simple-enough example that can be described in an intuitive manner.

The paper is organized as follows. Section 2 presents known DOE-optimality criteria that are related to the *Vs*-optimality criterion. Section 3 outlines the *Vs*-optimality criterion for linear models and sketches a numerical procedure for non-linear models. Section 4 illustrates the optimization-seeking principles of the proposed approach by a detailed example of a robust design of an *RL*-circuit. Section 5 concludes the paper.

2. Related DOE-optimality criteria

This section presents some DOE alphabetic optimality criteria that are related to the proposed *Vs*-optimality criterion (see Reference[1]). We use the bold font to represent vectors or matrices.

One of the most popular DOE-optimality criteria is *D-optimality*. This criterion aims to minimize the variance of the estimated model parameters. Its related objective function is $\max\{\det(\mathbf{F}'\mathbf{F})\}$ or $\min\{\det((\mathbf{F}'\mathbf{F})^{-1})\}$, where $\mathbf{F}$ is the design matrix, det stands for the determinant and $\mathbf{F}'\mathbf{F}$ is the non-normalized information matrix of the design. Note that the *D*-optimality criterion is fully related to an 'information-seeking' approach and does not consider the optimal solution of the system nor assigns different priorities among the system parameters. The first modification of this criterion toward an 'optimization-seeking' approach is known as 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[3]. Considering the 'optimization-seeking' approach, the experimenter can use the D_s-*optimality* criterion with respect to the s most influential parameters on the variance of the optimal solutions. Atkinson and Donev[4] proposed a simple procedure to obtain D_s-optimal experiments.

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 linear-optimality, assigning equal weights to all the variance estimates. With respect to the 'optimization-seeking' approach, one can use the *linear-optimality* criterion and assign higher weights to system parameters that contribute more to the variance of the optimal solution.

The objective of the *c-optimality* criterion is in estimating a linear combination of the model's coefficients $\mathbf{c}'\boldsymbol{\theta}$ with a minimum variance (e.g. see Reference[4]), where $\boldsymbol{\theta}$ is the column vector of model coefficients and $\mathbf{c}'$ is a row vector of the linear combination coefficients. Thus, the *c-optimality* criterion minimizes $\mathrm{Var}[\mathbf{c}'\hat{\boldsymbol{\theta}}] \propto \mathrm{Var}[\mathbf{c}'\mathbf{M}^{-1}(\xi)\mathbf{c}]$, where $\mathbf{M}(\xi) = \mathbf{F}'\mathbf{F}/n$ is the normalized information matrix of the chosen design, ξ. If $\mathbf{c}$ is taken to obtain $f(\mathbf{x_0})$—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 $\mathbf{x_0}$. Following this line of thought, the Vs-optimality criterion focuses on $\mathbf{x_0} = \mathbf{x}^*$, which is the optimal solution of the model. If the location of $\mathbf{x}^*$ is known prior to the experiment, one can minimize the prediction variance by repeatedly performing all the succeeding experiments at $\mathbf{x}^*$. This procedure often results in a singular optimum design, which is often non-informative[4]. In the sequential framework of the proposed approach[1], the coefficients, and therefore $\mathbf{x}^*$, are approximated in each experimental stage. Thus, at each stage one can select $\mathbf{c}$ to obtain a locally c-optimal design at $f(\mathbf{x}^*)$. More discussions on locally c-optimum experiments can be found in Atkinson *et al.*[5] and Kitsos *et al.*[6]. The *Vs*-optimality criterion is, thus, closely related to the *c-optimality* criterion. On one hand the *Vs*-optimality criterion is less general and focuses on the optimal solution of the system; yet, on the other hand it extends the *c-optimality* criterion to cases of non-linear combinations of the model coefficients[1].

3. The *Vs*-optimality criterion

In this section we show how the V_s-optimality criterion applies to a linear model, as presented in Reference[1] in order for the paper to be self-contained. The end of the section shortly discusses the case of non-linear models that was elaborated in the above work. In most cases, we denote random variables by capital letters and use the bold font to represent vectors or matrices.

The investigated linear model is given by

$$Y(\mathbf{x}, \mathbf{Z}) = \beta_0 + \boldsymbol{\beta}'\mathbf{x} + \boldsymbol{\alpha}'\mathbf{Z} + \mathbf{x}'\boldsymbol{\Gamma}\mathbf{Z} + \varepsilon \tag{1}$$

where Y is the unknown response; $\mathbf{x}$ is a $(k \times 1)$ vector of the coded significant control factors, i.e. $\mathbf{x}' = (x_1, x_2, \ldots, x_k)$; $\mathbf{Z}$ is an $(m \times 1)$ vector of the significant noise factors, coded such that $E[Z_i] = 0$, $i = 1, \ldots, m$; 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 to reduce the noise-factors' effects; and ε is a noise term with mean zero and a finite variance σ_ε^2. Such a linear model can be obtained by traditional experimental designs.

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

Optimization: The optimization stage of the robust-design problem is performed with respect to the loss function, $L(Y)=E(Y-T)^2$, which depends on variance of Y and the deviation of its mean with respect to a given target value, T.

$$E(Y(\mathbf{x},\mathbf{Z}))=\beta_0+\boldsymbol{\beta}'\mathbf{x},\quad V(Y(\mathbf{x},\mathbf{Z}))=(\boldsymbol{\alpha}'+\mathbf{x}'\boldsymbol{\Gamma})\boldsymbol{\Sigma}(\boldsymbol{\alpha}+\boldsymbol{\Gamma}'\mathbf{x})+\sigma_\varepsilon^2 \tag{2}$$

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

$$\begin{aligned}L(Y)&=E(Y-T)^2=V(Y)+(E(Y)-T)^2=V(Y)+E(Y)^2-2TE(Y)+T^2\\&=\boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\alpha}+\boldsymbol{\alpha}'\boldsymbol{\Sigma}\boldsymbol{\Gamma}'\mathbf{x}+\mathbf{x}'\boldsymbol{\Gamma}\boldsymbol{\Sigma}\boldsymbol{\alpha}+\mathbf{x}'\boldsymbol{\Gamma}\boldsymbol{\Sigma}\boldsymbol{\Gamma}'\mathbf{x}+(\beta_0+\boldsymbol{\beta}'\mathbf{x})^2-2T(\beta_0+\boldsymbol{\beta}'\mathbf{x})+T^2\end{aligned} \tag{3}$$

The first-order optimality condition for (3) yields the optimal robust solution:

$$\Rightarrow\mathbf{x}^*=(\boldsymbol{\Gamma}\boldsymbol{\Sigma}\boldsymbol{\Gamma}'+\boldsymbol{\beta}\boldsymbol{\beta}')^{-1}\cdot[\boldsymbol{\beta}(T-\beta_0)-\boldsymbol{\Gamma}\boldsymbol{\Sigma}\boldsymbol{\alpha}] \tag{4}$$

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 $\mathbf{g}()$ is a $(k\times 1)$ vector function of the set of the Linear-model's estimates, $\hat{\theta}_L=\{\hat{\beta}_0,\hat{\boldsymbol{\alpha}},\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\Gamma}}\}$ 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:

$$\begin{aligned}\mathbf{x}^*=\mathbf{g}(\hat{\boldsymbol{\theta}}_L)&\approx\mathbf{g}(\boldsymbol{\theta}_L)+\left[\frac{\partial(\mathbf{g}(\boldsymbol{\theta}_L))}{\partial\boldsymbol{\theta}_L}\right]\cdot(\hat{\boldsymbol{\theta}}_L-\boldsymbol{\theta}_L)\\&=\mathbf{g}(\boldsymbol{\theta}_L)+\mathbf{J}\cdot(\hat{\boldsymbol{\theta}}_L-\boldsymbol{\theta}_L)\end{aligned} \tag{5}$$

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.

The Vs-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 that satisfies $\mathbf{F}^*=\arg\ \min\{V(\mathbf{x}^*)\}$ within the selected design region[1].

The required design matrix $\mathbf{F}^*$ that minimizes the variance of the optimal solution $V(\mathbf{x}^*)$ is of dimensions $(n\times p)$, where p is the number of parameters and $n\geqslant p$ is the number of experiments. Taking the variance of the expression in (5) yields the following $(k\times k)$ variance–covariance matrix that we aim at minimizing:

$$V(\mathbf{x}^*)\approx\left[\frac{\partial(\mathbf{g}(\hat{\boldsymbol{\theta}}_L))}{\partial\boldsymbol{\theta}_L}\right]\cdot V(\hat{\boldsymbol{\theta}}_L)\cdot\left[\frac{\partial(\mathbf{g}(\hat{\boldsymbol{\theta}}_L))}{\partial\boldsymbol{\theta}_L}\right]'=\mathbf{J}\cdot((\mathbf{F}'\mathbf{F})^{-1}\sigma_\varepsilon^2)\cdot\mathbf{J}' \tag{6}$$

This is a fundamental result of the new criterion that can be obtained analytically only for linear response models when (6) exists in a closed form.

3.1. A suggested procedure for high-order or non-linear models

In a case that the investigated system is not well described by a linear model, the experimenter might use higher-order models. For such model, Equation (6) does not exist in a closed form, and one should combine numerical methods as well. Following is a suggested numerical procedure:

First, estimate the model by replicated experiments, then calculate the means and the variances of the coefficients based on the observations. Use these means and variances to generate new sets of the model's coefficients via Monte-Carlo sampling. Next, compute the optimal solution(s) numerically for each generated set of model coefficients, plot the robust solutions, and if necessary, cluster them to obtain the final optimal solution at the center of each cluster. Finally, compute the variance of each robust solution. Further information is given in Reference[1].

4. An illustrative example: the *RL* circuit design

This section motivates the usage of the proposed criterion by considering a real engineering problem, when the form of the transfer function is not necessarily a standard polynomial one. We assume that the transfer function is known *a priori*; yet, some of its parameters have to be estimated empirically. We show that the design of the experiments depends heavily on our knowledge regarding the values of the parameters. In particular, we show that when almost nothing is known about the parameters the best strategy to obtain an optimal solution for the problem is first to estimate the parameters' values as accurately as possible, as required by the *D*-optimality criterion. However, when there exists a good knowledge about the parameters of the function, it is better to estimate the value of the response in the vicinity of the optimal robust solution. The above principle is analyzed from a sequential design prospective. The example is based on a problem presented by Kenett and Zacks[7], which deals with a design of an *RL* electrical circuit.

First, we introduce the RL system and its robust solution according to Kenett and Zacks[7]. Next, we motivate the use of the proposed criterion by considering a simple case of a single unknown parameter. Finally, we extend the discussion to a case study of two unknown parameters that was partially considered in Reference[1].

4.1. The investigated RL circuit

The investigated response, I, in this case study, is the current of an RL circuit whose transfer function is known *a priori* to be of the following form:

$$I(R,L,f,V)=\frac{V}{\sqrt{R^2+(2\pi fL)^2}} \tag{7}$$

where I is the output current of the circuit; R and L are the control factors (whose values should be determined to obtain a desired value of the current I) that represent the resistance and the inductance of the circuit elements, respectively; V and f are the noise factors that represent, respectively, the voltage amplitude and frequency input into the circuit. We follow Kenett and Zacks[7] who considered these noise factors as independent random variables having the following normal distributions: $V\sim N(100,3)$ and $f\sim N(55,5/3)$.

Kenett and Zacks[7] derive a robust solution (R,L), based on a Taylor series approximation, in order to estimate the current's expected value, μ_I, its variance, V_I, and its associated loss function, $\text{Loss}=(\mu_I-T)^2+V_I$, with respect to a desired target value $T=10$[amp]. The control factors' settings are chosen based on a 3^2 factorial experiment. In each design combination, 500 values of V and f are simulated and the resulting current's mean, variance and loss are estimated. The best robust solution, according to Kenett and Zacks, is specified to be $R=9.5(\Omega)$ and $L=0.02$(H) with an associated loss of $\text{Loss}=0.1[\text{amp}]^2$.

Since the response function is known *a priori*, it is irrelevant to use conventional DOE tools for response surface methodology (RSM). Instead, we use DOE to estimate some function's parameters that are assumed unknown.

4.2. The case of a single unknown parameter

For illustration purpose, let us first consider the case where the coefficient of fL, which equals 2π in the original response (7), is unknown and has to be estimated by experimentation. Let us also assume that the current value is affected by a random-error term in the denominator as presented in (8).

$$I(R,L,f,V)=\frac{V}{\sqrt{R^2+a(fL)^2+\varepsilon}} \tag{8}$$

where a is the unknown parameter that has to be estimated and ε is the random-error term that has a zero mean and a finite variance σ_ε^2.

Estimating the unknown parameter: Recall that in the experimental stage, all the factors, including the noise factors, are assumed to be controllable. Therefore, in order to easily estimate a, one can apply the following transformation to obtain a simple linear function in a, $V^2/I^2-R^2=af^2L^2+\varepsilon$, or alternatively one can consider the following equation:

$$Y=af^2L^2+\varepsilon \quad \text{where } Y=\frac{V^2}{I^2}-R^2 \tag{9}$$

which can be written for a series of experiments as

$$\mathbf{Y}=a\cdot\mathbf{x}+\boldsymbol{\varepsilon} \tag{10}$$

where $\mathbf{Y}=(Y_1,Y_2,\ldots,Y_n)'$, $\mathbf{x}=(x_1,x_2,\ldots,x_n)'$, $x_i=f_i^2L_i^2$, $i=1,\ldots,n$. Thus, in each experiment, the designer selects the values of the factors V, f, R and L, observes I and estimates the value of a. The design matrix, $\mathbf{x}$, represents the chosen values for f and L during the n experiments and ε represents the vector of random errors. According to a simple linear regression the maximum likelihood estimator (MLE) for a is equal to

$$\hat{a}=(\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{Y} \quad \text{where } \boldsymbol{\sigma}_{\hat{a}}^2=\frac{\boldsymbol{\sigma}_\varepsilon^2}{\sum_{i=1}^n x_i^2} \tag{11}$$

As the number of experiments, n, increases $\mu_{\hat{a}}\to 4\pi^2$ and $\sigma_{\hat{a}}^2\to 0$.

Finding the robust solution: After estimating the unknown parameter, one can use this estimate in order to find the optimal robust solution, which minimizes the loss function with respect to the given target value $T=10$[amp] working with a transfer function that involves an unknown parameter and a random error component, as presented in (8). Note that the loss function is affected not only by the error but also by the uncertainty of the unknown parameter, as can be seen from a second-order Taylor approximation for the loss function,

$$\begin{aligned}\text{Loss}(R,L,T,\sigma_{\hat{a}},\mu_{\hat{a}},\mu_V,\sigma_V,\mu_f,\sigma_f)=(\mu_I-T)^2+V_I&=\frac{\mu_V^2\mu_f^2\mu_{\hat{a}}^2L^4(\mu_f^2\sigma_a^2+\mu_{\hat{a}}^2\sigma_f^2)}{M^3}\\&+\frac{\sigma_V^2}{M}+\left[\frac{1}{2}\cdot\left(\frac{3\mu_V\mu_f^4\mu_{\hat{a}}^2L^4}{M^{5/2}}-\frac{\mu_V\mu_f^2L^2}{M^{3/2}}\right)\cdot\sigma_{\hat{a}}^2+\frac{1}{2}\cdot\left(\frac{3\mu_V\mu_f^2\mu_{\hat{a}}^4L^4}{M^{5/2}}-\frac{\mu_V\mu_{\hat{a}}^2L^2}{M^{3/2}}\right)\cdot\sigma_f^2+\frac{\mu_V}{M^{1/2}}-T\right]^2\end{aligned} \tag{12}$$

where $M=\mu_{\hat{a}}^2\mu_f^2L^2+R^2$.

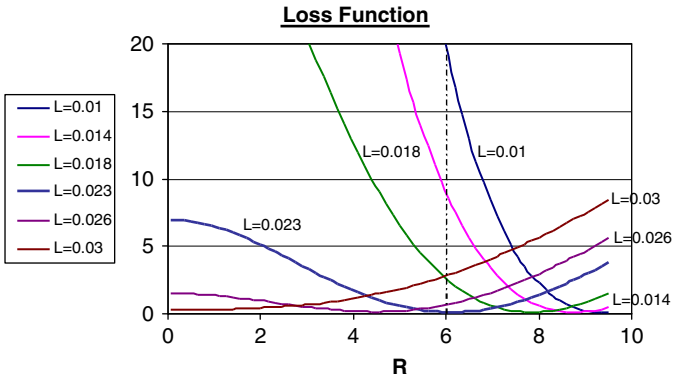

Figure 1. Loss function's values as a function of the resistance for several coil magnitudes. The optimal solution for the deterministic case, $L_D^*=0.023$, is represented by the thick line. This figure is available in colour online at www.interscience.wiley.com/journal/qre

Evaluating the loss at the point of the expected values of the noise factors, $\mu_V=100$, $\mu_f=55$, plugging $\mu_{\hat{a}}=4\pi^2$ and the variances of the noise factors, $\sigma_V=3$, $\sigma_f=\frac{5}{3}$, to Equation (12) results in a loss function that depends on the uncertainty of the unknown parameter $\sigma_{\hat{a}}$ and on the control factors R and L.

Let us suppose, for example, that from practical reasons, the designer considers to work with resistor of 6(Ω), i.e. $R=6$. Had the value $a=4\pi^2$ been known *a priori* and the transfer function presented in (7) had been deterministic, then a straight forward numerical optimization could yield an optimal value of the inductance $L_D^*=0.023$(H) (where D denotes 'deterministic'), as can be seen in Figure 1. The figure shows the loss function's values as a function of the resistance for several coil magnitudes. It can be seen that in the vicinity of $R=6$ the loss function is minimized at $L=0.023$, which is represented by the thick line. Note that if such a coil does not exist, the designer might either choose a close value, e.g. 0.02(*H*) (although not necessarily optimal) or consider working with another combination such as $(R,L)=(8,0.018)$.

4.3. *Analyzing the effect of the parameter's uncertainty*

Recall that for the case where the parameter a is unknown, the loss function is affected by its related uncertainty, denoted by $\sigma_{\hat{a}}$. The considered second-order Taylor series approximation in (12) for the loss function in the vicinity of $R=6$ and the deterministic optimal inductance solution $L_D^*=0.023$ is given by

$$\begin{aligned}&\text{Loss}(R=6,L=0.023,T=10,\mu_{\hat{a}}=4\pi^2,\mu_V=100,\sigma_V=3,\sigma_f=\tfrac{5}{3})\\&=(77791.6-4946.66\cdot\sigma_{\hat{a}}^2-9.85\cdot\sigma_{\hat{a}}^4)\cdot L^2-(3607.73-195\cdot\sigma_{\hat{a}}^2-1.22\cdot\sigma_{\hat{a}}^4)\cdot L+41.96-0.825\cdot\sigma_{\hat{a}}^2-0.017\cdot\sigma_{\hat{a}}^4\end{aligned}\tag{13}$$

Since a is estimated by means of a simple linear regression analysis, its standard deviation following the $(n+1)$th experiment can be written as

$$\sigma_{\hat{a}}=\frac{\sigma_\varepsilon}{\sqrt{k+x^2}}\tag{14}$$

where $k=\sum_{i=1}^{n}x_i^2$ represents the summation of the squares of past values $x_i=f_i^2L_i^2$, $i=1,\ldots,n$, that were already selected in the previous n experiments, as represented in Equation (10), and x is the next experimental point that will be used to re-evaluate the unknown coefficient.

Let us now analyze the loss function in the vicinity of $R=6$ and $L=0.023$ as a function of: (i) the initial uncertainty in the parameter value (σ_ε), (ii) the past experiment values (represented by k) and (iii) the value of the next experimental point x. Moreover, since the noise factors are assumed to be controllable during the experimental stage, we can keep the circuit's frequency, f, at its mean value, $\mu_f=55$, and focus on finding the optimal value of the coil, L. Figure 2 shows the loss function $\text{Loss}(R,L,\sigma_{\hat{a}},T=10,\mu_{\hat{a}}=4\pi^2,\mu_V=100,\sigma_V=3,\sigma_f=\frac{5}{3})$ against various values of L for $k=25$. The two graphs represent different initial uncertainties in the value of parameter a. The lower graph is depicted for an initial low uncertainty, $\sigma_\varepsilon=1$, and the upper graph is depicted for an initial high uncertainty, $\sigma_\varepsilon=30$, implying that the loss function grows with σ_ε, as expected.

It can be seen that when σ_ε is low with respect to k, the value of L that minimizes the loss is in the vicinity of the optimal solution for the deterministic case, $L_D^*=0.023$, thus, according to the 'optimization-seeking' approach. However, if σ_ε is large, the parameter a should be estimated previous to the optimization, and the optimal value of L is toward the edges of the design region, as often happened in *D*-optimal designs, thus, according to the 'information-seeking' approach.

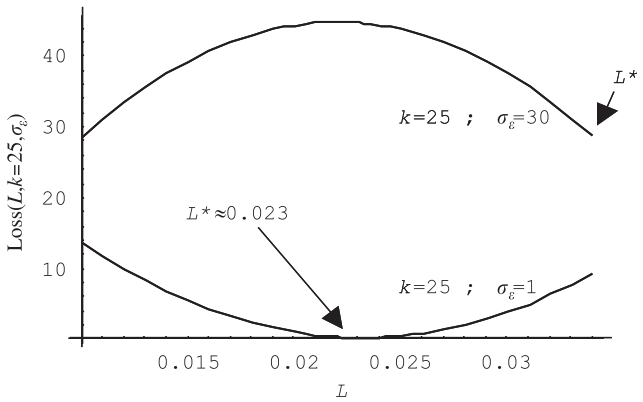

Figure 2. Loss function with respect to the values of L and σ_ε for $k=25$

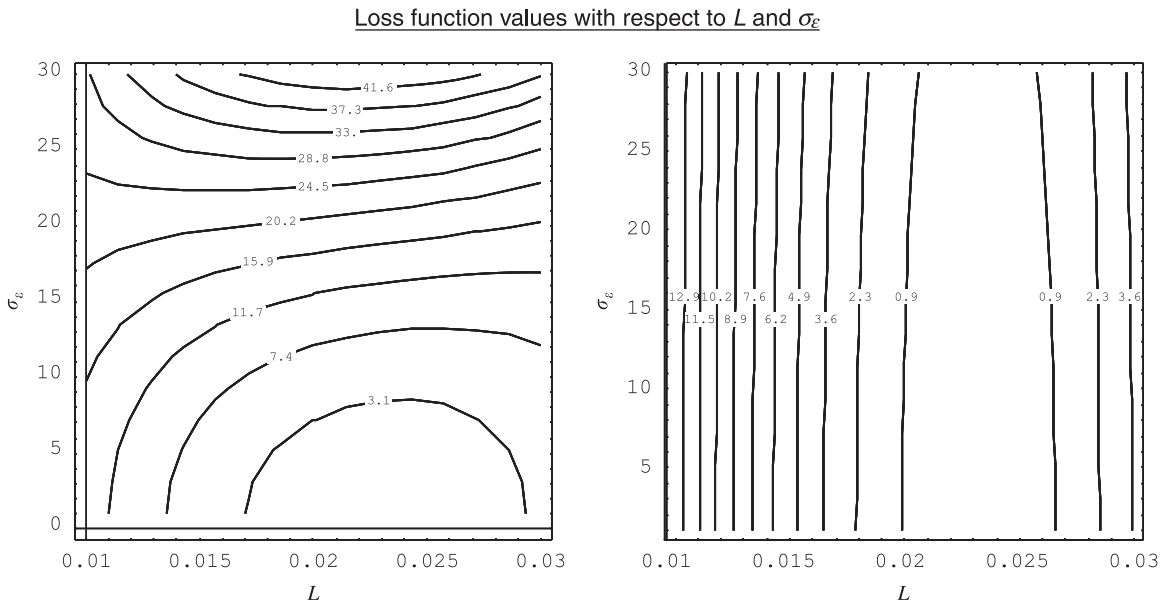

Figure 3. Loss function with respect to the values of L,k and σ_ε: (i) $k=25$ and (ii) $k=2500$

This result can also be seen from the analytical minimization of the loss function in (13) with respect to L:

$$L^* = \frac{0.5\cdot(3607.73 - 195\sigma_{\hat{a}}^2 - 1.22\sigma_{\hat{a}}^4)}{77791.6 - 4946.66\sigma_{\hat{a}}^2 - 9.85\sigma_{\hat{a}}^4} \tag{15}$$

Indeed, substituting $\sigma_{\hat{a}}=0$ in (15) yields $L^*=0.0232$ while substituting $\sigma_{\hat{a}}=30$ yields $L^*=0.0468$, which is approximately twice as large.

Figure 3 illustrates the effect of k (the sum of past experimental square values for L). The left graph shows a contour plot of the loss function with respect to the values of L and σ_ε for $k=25$, and the right graph shows a contour plot of the loss function with respect to the values of L and σ_ε for $k=2500$. The left graph leads to the same conclusions as obtained from Figure 2, i.e. the robust solution is located in the vicinity of 0.023 for low σ_ε and toward the edges of the region as σ_ε grows. The right graph shows that when k is large enough with respect to σ_ε, the best value of L stays fixed in the vicinity of $L_D^*=0.023$ with a very small effect of σ_ε. Hence, when enough experimental points were used in order to estimate the unknown coefficient, its standard deviation goes lower and has a smaller effect on the optimal robust solution.

4.4. The case of two unknown parameters

The response of the current, I, is now considered to have two unknown parameters, a and b, where a is the coefficient of $(fL)^2$ (which equals to $4\pi^2$ in the known response in (7)) and b is the coefficient of R^2 (which is equal to 1 in the known response in (7)). As in the

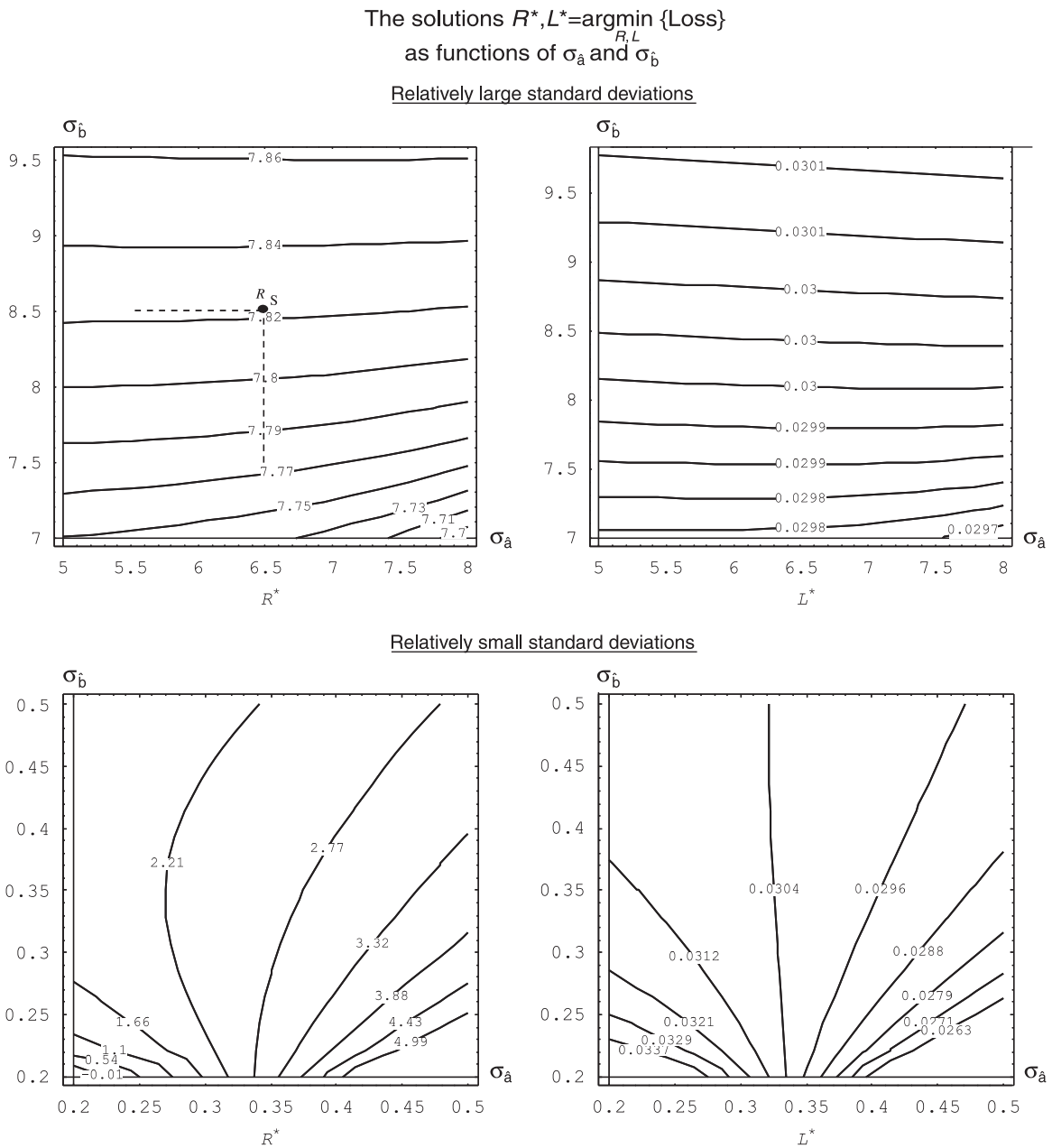

Figure 4. R^* and L^* as functions of $\sigma_{\hat{a}}$ and $\sigma_{\hat{b}}$ two design regions

previous section, ε is assumed to have a zero mean and a finite variance σ_ε^2. The response can be now written as

$$I(R,L,f,V)=\frac{V}{\sqrt{bR^2+a(fL)^2+\varepsilon}} \tag{16}$$

Following (8), one can obtain a simple linear equation:

$$\frac{V^2}{I^2}=(fL)^2+bR^2+\varepsilon$$

$$\Rightarrow Y=a(fL)^2+bR^2+\varepsilon \quad \text{where } Y=\frac{V^2}{I^2} \tag{17}$$

A second-order Taylor approximation for the loss function with respect to the target value, $T=10$, can be derived as follows:

$$L_I(R,L,T,\sigma_{\hat{a}},\mu_{\hat{a}},\sigma_{\hat{b}},\mu_{\hat{b}},\mu_V,\sigma_V,\mu_f,\sigma_f)=(\mu_I-T)^2+V_I$$
$$=\frac{\mu_V^2\mu_f^4\mu_{\hat{a}}^2\sigma_{\hat{a}}^2L^4}{M^3}+\frac{\mu_V^2\sigma_{\hat{b}}^2R^4}{4M^3}+\frac{\mu_V^2\mu_f^2\mu_{\hat{a}}^4\sigma_f^2L^4}{M^3}+\frac{\sigma_V^2}{M}+\left[\frac{1}{2}\cdot\left(\frac{3\mu_V\mu_f^4\mu_{\hat{a}}^2L^4}{M^{5/2}}-\frac{\mu_V\mu_f^2 3^2}{M^{3/2}}\right)\cdot\sigma_{\hat{a}}^2\right.$$
$$\left.+\frac{1}{2}\cdot\frac{3\mu_V\sigma_{\hat{b}}^2R^4}{4M^{5/2}}\cdot\sigma_{\hat{b}}^2+\frac{1}{2}\cdot\left(\frac{3\mu_V\mu_f^2\mu_{\hat{a}}^4L^4}{M^{5/2}}-\frac{\mu_V\mu_{\hat{a}}^2L^2}{M^{3/2}}\right)\cdot\sigma_f^2+\frac{\mu_V}{M^{1/2}}-T\right]^2 \quad (18)$$

where $M=\mu_{\hat{a}}^2\mu_f^2L^2+\mu_{\hat{b}}R^2$.

Evaluating the loss at the point of the expected values of the noise factors, $\mu_V=100$, $\mu_f=55$ and plugging to Equation (18) both the MLE values, $\mu_{\hat{a}}=4\pi^2$, $\mu_{\hat{b}}=1$, and the variances of the noise factors, $\sigma_V=3$, $\sigma_f=\frac{5}{3}$, result in a similar expression to (12) with additional terms related to $\sigma_{\hat{b}}^2$.

Figure 4 shows a contour plot of the optimal robust solutions $\{R^*,L^*\}=\text{argmin}_{R,L}\{\text{Loss}\}$ as a function of $\sigma_{\hat{a}}$ and $\sigma_{\hat{b}}$. The upper graphs show the values of R^* and L^* for one part of the design region, in which the values of these sample standard deviations are relatively high. The lower graphs show another part of the design region, in which the values of these sample standard deviations are relatively low. The left upper graph shows a given solution, $R_S^*\approx 7.822(\Omega)$, that is obtained for $\sigma_{\hat{a}}=6.5$ and $\sigma_{\hat{b}}=8.5$. Note that a reduction of one unit of $\sigma_{\hat{b}}$ while holding $\sigma_{\hat{a}}$ fixed resulted in a new optimal solution $R_S^*\approx 7.77(\Omega)$. However, a reduction of one unit in $\sigma_{\hat{a}}$ while holding $\sigma_{\hat{b}}$ fixed does not result in any significant change in the optimal robust solution. This phenomenon is also true for L^*, as can be seen from the upper right figure. Thus, in general, for relatively large values of the standard deviations, changes in $\sigma_{\hat{a}}$ do not result in significant changes of the values R^* and L^*, while changes in $\sigma_{\hat{b}}$ affect significantly the optimal robust solutions. In other words, in order to reduce the variance of the obtained solutions, $V(\mathbf{x}^*)$, one should focus primarily on estimating b as accurately as possible rather than investing in a better estimation of a. The lower graphs show the opposite situation for relatively low standard deviations. In order to reduce the variances of R^* and L^* one should focus on estimating a as accurately as possible rather than investing in the estimation of b. In case of intermediate standard deviations' values, there are no general recommendations, and the decision how to allocate the next experiments is left to the experimenter. Note that these conclusions cannot be anticipated from the loss function itself and are the result of the suggested methodology.

5. Conclusions

This paper illustrates the *Vs*-optimality criterion, which was first introduced in Reference[1]. The paper illustrates how the *Vs*-optimality criterion is related to an 'optimization seeking' experimental approach rather than to an 'information-seeking' approach. The proposed criterion is formulated for a case that the investigated system is linear. The principles of the proposed approach are illustrated by a running example of a robust design of an *RL*-circuit. One potential research direction is to seek ways to implement the V_s-optimality criterion within a fully 'automated' sequential optimization framework and not necessarily for robust optimization.

References

1. Ginsburg H, Ben-Gal I. Designing experiments for robust optimization problems: The Vs-optimality criterion. *IIE Transactions* 2006; **38**:445–461.
2. Xu D, Albin SL. Robust optimization of experimentally derived objective functions. *IIE Transactions* 2003; **35**:793–802.
3. Feodorov V. *Theory of Optimal Experiments*. Academic Press: New York, 1972.
4. Atkinson AC, Donev AN. *Optimum Experimental Designs*. Oxford University Press: New York, 1992.
5. Atkinson AC, Chaloner K, Herzberg AM, Juritz J. Optimum experimental designs for properties of a compartmental model. *Biometrics* 1992; **49**(1):325–337.
6. Kitsos CP, Titterington DM, Torsney B. An optimal design problem in rythmometry. *Biometrics* 1988; **44**:657–671.
7. Kenett RS, Zacks S. *Modern Industrial Statistics*: *Design and Control of Quality and Reliability*. Cengage Learning: New York, 1998.

Authors' biographies

Hilla Ginsburg is an industrial engineer. She holds BSc (2001) and MSc (2003) degrees in Industrial Engineering and an MBA (2005) degree, all from Tel-Aviv University. Hilla has been lecturing in Tel Aviv University and in the Open University. At present she is working as a project manager In Medatech Information Technologies which deals with the implementation of ERP information systems.

Her research interests include design-of-experiments and design of manufacturing processes.

Professor Irad Ben-Gal is the head of the CIM Lab in the Department of Industrial Engineering at Tel Aviv University. He holds a BSc (1992) degree from Tel-Aviv University, and MSc (1996) and PhD (1998) degrees from Boston University. He is a member of the Institute for Operations Research and Management Sciences (INFORMS), the Institute of Industrial Engineers (IIE), The European Network for Business and Industrial Statistics (ENBIS) and an elected member in the International Statistical Institute (ISI). He is a Department

Editor in the *IIE Transactions on Quality and Reliability* and serves on the Editorial Boards of several other professional journals. He has published more than 80 scientific papers and received several best papers awards. His papers have been published in *IIE Transactions, International Journal of Production Research, Technometrics, IEEE Transaction, Quality and Reliability Engineering International, Journal of Statistical Planning and Inference,* as well as *Bioinformatics* and *BMC Bioinformatics*. Prof. Ben-Gal received several research grants, among them from General Motors, IEEE, the Israeli Ministry of Science, and the European Community. He has worked for several years in industrial organizations. His research interests include statistical methods for control and analysis of stochastic processes; Applications of Information Theory to industrial problems; Machine Learning and Automation and Computer Integrated Manufacturing systems.