Optimal Design of Computer Experiments for Metamodel Generation Using I-OPT™

Selden B. Crary[1], Peter Cousseau[2], David Armstrong[1], David M. Woodcock[3], Eva H. Mok[1]
Olivier Dubochet[4], Philippe Lerch[4], Philippe Renaud[2]

Abstract: We present a new and unique software capability for finding statistical optimal designs of deterministic experiments on continuous cuboidal regions. The objective function for the design optimization is the minimization of the expected integrated mean squared error of prediction of the metamodel that will be found, subsequent to the running of the computer simulations, using the best linear unbiased predictor (BLUP). The assumed response-model function includes an unknown, stochastic term, Z. We prove that this criterion, which we name I_Z-optimality, is equivalent to I-optimality for non-deterministic experiments, in the limit of zero correlations among the Z's for different inputs. An example is presented of metamodel generation for a micromachined-silicon flow sensor. The I_Z-optimal set of inputs is found, finite-element (FE) simulations run, and the metamodel generated using a BLUP fit. The method is compared to other approaches. I_Z-optimality, coupled with BLUP fitting, provides a highly efficient means of non-parametric metamodel generation. I_Z-optimal design searching and BLUP fitting are new options of the I-OPT™ program that is available on the World-Wide Web at URL http://www-personal.engin.umich.edu/∼crary/iopt.

keyword: design of computer experiments, I-optimality, microelectromechanical systems, MEMS, silicon flow sensor.

1 Introduction

The design and optimization of microsystems can require large numbers of computationally intensive simulations, such as discretized approximations of partial differential equations (finite-element or boundary-element analyses) or systems of coupled ordinary differential equations. Often it would be convenient if a simpler, but still reasonably accurate, functional approximation could be found that could be evaluated orders of magnitude more rapidly than the systems of equations it is replacing. Such surrogate functions, or metamodels, could represent components of a MEMS system and could be used effectively in design synthesis to allow for the rapid trial evaluation and then selection of components in a system. Alternatively, such metamodels could be used for rapid optimization, since the functional evaluations that normally dominate such optimization would be computed very rapidly.

In earlier work, we demonstrated two parametric methods for designing experiments for metamodel generation in the context of MEMS, namely I-optimal, single-domain, response-surface methodology, see Gianchandani and Crary (1998), and an algorithm for patch-wise functional approximation, see Crary and Phan (1998). These methods required the specification of either a model function or a set of basis functions prior to the search for a suitable designed experiment. By contrast, in this paper we draw upon a non- parametric method that uses an optimal linear predictor of the type introduced by Wold (1938) in the context of time series. Two conference reports included content found in the present study, see Crary, Cousseau, Armstrong, Woodcock, Dubochet, Lerch, and Renaud (1999) and Crary, Cousseau, Mok, Woodcock, and Renaud (1999).

A series of statistics papers has highlighted the method for deterministic computer experiments; see Sacks, Schiller and Welch (1989); Sacks, Welch, Mitchell, and Wynn (1989); Welch, Yu, Kang, and Sacks (1990); Currin, Mitchell, Morris, and Ylvisaker (1991); and Welch, Buck, Sacks, Wynn, Mitchell, and Morris (1992). There are both advantages and disadvantages to the method presented here, and these will be discussed below. This approach has been demonstrated previously with considerable success in geophysics; see Chapman, Welch, Bowman, Sacks, and Walsh (1994); in marine science, see Gough and Welch (1994); in aerospace engineering; see Simpson, Mauery, Korte, and Mistree (1998); in structural engineering; see Simpson, Allen, and Mistree (1998); for inkjet printhead design; see Salagame and Barton (1997); in semiconductor engineering; see Aslett, Buck, Duvall, Sacks, and Welch (1998); Bernardo, Buck, Liu, Nazaret, Sacks, and Welch (1992); and Currin, Mitchell, Morris, and Ylvisaker (1991); and for thermal energy-storage systems; see Currin, Mitchell, Morris, and Ylvisaker (1991). In an early chemical-kinetics example with two salient factors, Sacks, Schiller, and Welch (1989) compared the new approach with that of a traditional response-surface method (3×3 factorial design and least- squares fitting analysis, as reported in Miller and Frenklach (1983)) and demonstrated a remarkable 8- to 10-fold reduction in the variance of prediction of the fitting function.

Helpful reviews are available on statistics applied to computer

[1] Ctr. for Integrated MicroSystems, U. Mich., Ann Arbor, MI, USA
[2] Institute for Microsystems, EPFL, Ecublens, CH
[3] Lab. for Scientific Computation, U. Mich., Ann Arbor, MI, USA
[4] Leister Process Technologies, Sarnen, CH

experiments, see Koehler and Owen (1996), on the history of the correlation-function method, see Cressie (1990), on the history of BLUP fitting, see Kennedy (1991), and on metamodels, see Barton (1992) and Simpson (1998). The mathematics is covered well in the text by Searle, Casella, and McCulloch (1992). An interesting critique is given by Etman (1994). An alternative approach is to use moving least-squares methods, see Lancaster and Salkauskas (1981) and Etman (1994).

An interesting and related element-free approach to solving partial differential equations is taken by Lu, Belytschko, and Gu (1994) and Belytschko, Lu, and Gu (1994). Senturia (1998) and Senturia, Aluru, and White (1997) make reference to MEMS design using a basis- function approach.

Our work followed the Sacks, Schiller, and Welch (1989) closely, and we were able to duplicate much of their work on a demonstration MEMS example. However, there were also notable differences. The most important of these were that we were able to find superior designs to those published in their paper and that our design software is available to the public.

2 I-OPT version 4

Briefly, Version 4 of I-OPT includes a new capability for finding designs for deterministic experiments. I-OPT is a single program, compiled from both FORTRAN and C source code, that finds optimal designs minimizing the expected integrated mean-squared error of prediction (IMSE) of a metamodel, where the model function can contain an unknown part. For example, in two-factors the model function may be the following:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_2^2 + \beta_5 x_1 x_2 + Z(x_1, x_2), \quad (1)$$

where $Z(x_1,x_2)$, the departure from the second-degree model, is modeled as a stochastic process with covariance given by

$$\text{cov}\,[Z(s_1,s_2),Z(t_1,t_2)] = \sigma_z^2 \exp\left\{-\left[\theta_1(s_1-t_1)^2 + \theta_2(s_2-t_2)^2 + \theta_{12}(s_1-t_1)(s_2-t_2)\right]\right\}, \quad (2)$$

with θ_1, θ_2, θ_{12}, and σ_z^2 being parameters that must be set prior to the search for the optimal design. The setting of the θ's and σ_z^2 can be accomplished in any one of the following three ways: (1) through a set of preliminary simulations and a fitting using maximum likelihood, (2) through a so-called "robustness study", [see Sacks, Schiller, and Welch (1989)] or (3) in the course of sequential computer experimentation, again using maximum likelihood. Standard statistical methods are used to find the best linear unbiased predictor (BLUP) fit to the data.

One advantage of the method is that the error is not assumed to be random upon repetition of an experiment, as in more traditional approaches to design of experiments, including D-, G-, and I-optimality. Rather, the generally correct assumption is made that the differences between responses of replicated computer experiments are zero. This is consistent with the concept of deterministic computer experiments, although there are situations in which computer experiments give different results for the same inputs, see Gianchandani and Crary (1998). A second advantage of the method is that the form of the model function need not be specified prior to finding a design and initiating simulations. There is, however, the burden of establishing appropriate values for the θ's and σ_z^2, see Etman (1994) for a discussion of this issue.

3 Mathematical background

In this section we briefly review the IMSE-based optimality criteria for both non-deterministic and deterministic experiments.

A few comments on notation are in order. There is no standard terminology for these optimality criteria, of the sort with which engineers are familiar through standards organizations, such as the IEEE or ASTM. Rather, consensus has built up through usage. In the case of non- deterministic experiments under IMSE-based optimality, the term "I-optimality" has been widely, but not universally, adopted. In the case of deterministic IMSE-based optimality, no conventional name has emerged. One complicating factor is that the optimality criteria, which are instantiated by an objective function, depend on the following: a model function exclusive of its error term; the portion of the model function that accounts for error, which may include random- noise terms and terms that account for systematic departures; knowledge of or an assumption about the nature of the random or systematic errors; and the goal of the optimization in terms of minimizing an average squared error, minimizing the worst-case of expected squared error, or some other objective. Nonetheless, the need for names arises.

For the purposes of this paper, we introduce the following three names for IMSE-based optimality, depending upon their error model:

- "I_ε-optimality" for non-deterministic experiments with error models containing only random error,
- "I_Z-optimality" for deterministic experiments with non-parametric error models representing departure from a given model, and
- "$I_{\varepsilon+Z}$-optimality" for non-deterministic experiments with error models containing both random error as well as non- parametric error representing departure from a given model.

In the Sec. 3.3 we prove that I_ε-optimality is a limiting case of I_Z-optimality, under simple assumptions. In light of this new result, which shows the connection between the various I_x criteria, it seems natural to use the term "I-optimality" to apply to the entire class of IMSE-based optimality.

3.1 I_ε-optimality

An I_ε-optimal design is a specified set of points in the design space at which measurements should be taken in order to minimizes the expected (possibly weighted) integrated mean squared error of prediction of a metamodel generated by (possibly generalized) least-squares fitting the responses to a linear statistical model. The fundamental assumptions are that the model function is known, is linear in the coefficients, and contains an error term $\boldsymbol{\varepsilon}(\mathbf{x})$, the distribution of which is assumed to have a mean of zero and a variance of $\sigma^2(\mathbf{x})$ and to be drawn independently and individually at random. Such a model can be written as the following:

$$Y(\mathbf{x}) = \sum_{i=1}^{k} \beta_i f_i(\mathbf{x}) + \varepsilon(\mathbf{x}), \tag{3}$$

where the β's are the linear coefficients and the f's are the functional terms in the model. The objective function can be expressed in differential form as the following:

$$\min_{\omega_N} \frac{1}{\Omega} \int_{\mathbf{x}\in\chi} E\left[\left(\hat{Y}(\mathbf{x}) - Y(\mathbf{x})\right)^2\right] w(\mathbf{x})\, dx_1 \ldots dx_d =$$
$$\min_{\omega_N} \frac{1}{\Omega} \int_{\mathbf{x}\in\chi} \left\{ \sum_{i=1}^{N} \left[\left(\frac{\partial \hat{Y}(\mathbf{x})}{\partial y_i} \right)^2 \sigma_i^2 \right] \right\} w(\mathbf{x})\, dx_1 \ldots dx_d,$$

where $w(\mathbf{x})$ is a weighting function and

$$\Omega = \int_{\mathbf{x}\in\chi} w(\mathbf{x})\, dx_1 \ldots dx_d,$$

which asks for the N-point design ω_N that minimizes the average over a domain χ of the total expected squared error, taken as a sum of independent variances of the contributions due to the variance σ_i^2 of each response y_i. This is a very natural definition, but is rarely encountered in the literature. More commonly, the definition is expressed in linear-algebraic form, such as the following, which we give for the restricted class of homoscedastic error models, i.e., error models where the error term $\varepsilon(\mathbf{x})$ is constant over the design region, $\varepsilon(\mathbf{x}) = \varepsilon$:

$$\min_{\omega_N} \frac{1}{\Omega} \int_{\mathbf{x}\in\chi} \mathbf{f}' \left(\mathbf{F}'\mathbf{F}\right)^{-1} \mathbf{f}\, w(\mathbf{x})\, dx_1 \ldots dx_d,$$

where

$$\mathbf{f}'_{\mathbf{x}} = [f_1(\mathbf{x}), \ldots, f_k(\mathbf{x})] \quad \text{and} \quad \mathbf{F} = [f_l(s_i)]_{1\le i\le N,\, 1\le l\le k},$$

or equivalently, via a matrix identity, as

$$\min_{\omega_N} \text{trace}\left[\left(\mathbf{F}'\mathbf{F}\right)^{-1} \cdot \mathbf{B}\right], \tag{4}$$

where

$$\mathbf{B} \equiv \frac{1}{\Omega} \int_{\mathbf{x}\in\chi} \mathbf{f}\mathbf{f}' w(\mathbf{x})\, dx_1 \ldots dx_d,$$

Readers who wish further detail are referred to Searle, Casella, and McCulloch (1992).

3.2 I_Z-optimality

When dealing with deterministic experiments there is no random error of the type modeled with $\varepsilon(\mathbf{x})$ in the section immediately above. Rather, a term $Z(x)$ is introduced that represents the unmodeled part of the response:

$$Y(\mathbf{x}) = \sum_{i=1}^{k} \beta_i f_i(\mathbf{x}) + Z(\mathbf{x}). \tag{5}$$

In order to define a objective function for the design problem, something must be known or assumed about the error term $Z(\mathbf{x})$. A plausible and convenient assumption that is often made assumes that the covariance between values of $Z(\mathbf{x})$ at two inputs has the following gaussian dependence on the separation between the inputs:

$$\text{cov}\,(Z(\mathbf{t}), Z(\mathbf{u})) = V(\mathbf{t}, \mathbf{u}) = \sigma_z^2 \exp\left\{ -\sum_{i=1}^{d} \left[\theta_i (t_i - u_i)^2 + \sum_{j=i+1}^{d} \theta_{i,j} (t_i - u_i)(t_j - u_j) \right] \right\}. \tag{6}$$

In contrast to the usual treatment, we explicitly introduce the cross terms, with pre-factors $\theta_{i,j}$, in Eq. 6, in order to properly account for the possibility that the elliptical contours of iso-covariance may have axes not aligned with the coordinate axes. After the θ's and σ_z^2 are specified and a set of input points selected as the design, $\mathbf{S} = \{\mathbf{s}_1, \ldots, \mathbf{s}_N\}$, the simulator is run and responses to $Y(\mathbf{x})$ recorded as a set $Y(\mathbf{s}_1), \ldots, Y(\mathbf{s}_N)$. Following Sacks, Schiller, and Welch (1989), we introduce the notation:

$$\mathbf{V} = [\text{cov}\,(Y(\mathbf{s}_i), Y(\mathbf{s}_j))]_{1\le i\le N,\, 1\le j\le N}$$
$$\mathbf{v}'_x = [V(\mathbf{s}_1, \mathbf{x}), \ldots, V(\mathbf{s}_N, \mathbf{x})]$$
$$\mathbf{y}' = [Y(\mathbf{s}_1), \ldots, Y(\mathbf{s}_N)].$$

The fit is performed using a linear predictor, $\mathbf{c}'\mathbf{y}$, that is a linear combination of the responses at the N design points and has a mean squared error of

$$E\left[\mathbf{c}'\mathbf{y} - Y(\mathbf{x})\right]^2 = \left(\mathbf{c}'\mathbf{F}\beta - \mathbf{f}'_{\mathbf{x}}\beta\right)^2 + [\mathbf{c}', -1] \begin{bmatrix} \mathbf{V} & \mathbf{v}_x \\ \mathbf{v}'_x & \sigma_z^2 \end{bmatrix} \begin{bmatrix} \mathbf{c} \\ -1 \end{bmatrix}. \tag{7}$$

Constraining the fit to go through the responses at the designed points, i.e., $\mathbf{c}'\mathbf{F}\beta = \mathbf{f}'_x\beta$, gives a set of k constraint equations. Then minimizing Eq. 7 is reduced to minimizing the RHS of Eq. 7, subject to the constraints $\mathbf{F}'\mathbf{c} = \mathbf{f}_x$, a problem that is amenable to solution using a vector $\boldsymbol{\lambda}$ of k Lagrange multipliers. This gives the pair of $N \times 1$ vector equations $\mathbf{V}\mathbf{c} - \mathbf{v}_{\mathbf{x}} - \mathbf{F}\boldsymbol{\lambda} = 0$ and $\mathbf{F}'\mathbf{c} = \mathbf{f}_{\mathbf{x}}$, which can be written

Table 1 : I-OPT-generated, nine-point, putatively I_Z-optimal design for model in Eq. 1, assuming $\theta_1 = \theta_2 = 0$ and $\theta_{12} = 0$.

x_1	x_2
-0.719	0.874
0.013	0.642
0.782	0.782
-0.830	0.189
0.642	0.013
-0.236	-0.236
-0.753	-0.753
0.189	-0.830
0.874	-0.719

Table 2 : Nine-point design for model in Eq. 1, assuming $\theta_1 = \theta_2 = 1$ and $\theta_{12} = 0$ from Sacks, Schiller, and Welch (1989).

x_1	x_2
-0.74	0.90
0.00	0.66
0.80	0.80
-0.86	0.27
0.66	0.00
-0.34	-0.34
-0.78	-0.78
0.27	-0.86
0.90	-0.74

conveniently in the following matrix form:

$$\begin{bmatrix} \mathbf{0} & \mathbf{F}' \\ \mathbf{F} & \mathbf{V} \end{bmatrix} \begin{bmatrix} -\boldsymbol{\lambda} \\ \mathbf{c} \end{bmatrix} = \begin{bmatrix} \mathbf{f_x} \\ \mathbf{v_x} \end{bmatrix}. \tag{8}$$

We now seek to express the best (meaning minimum in a least-squares sense) linear unbiased predictor (BLUP) fit $\mathbf{c}'\mathbf{y}$ in a convenient form. The fit may be written, after rearrangement of Eq. 8, as the following:

$$\mathbf{c}'\mathbf{y} = \left[-\boldsymbol{\lambda}', \mathbf{c}'\right] \begin{bmatrix} \mathbf{0} \\ \mathbf{y} \end{bmatrix} = \left[\mathbf{f}'_\mathbf{x}, \mathbf{v}'_\mathbf{x}\right] \begin{bmatrix} \mathbf{0} & \mathbf{F}' \\ \mathbf{F} & \mathbf{V} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{0} \\ \mathbf{y} \end{bmatrix} \tag{9}$$

This can also be written as the sum of a generalized least squares term,

$$\mathbf{f}'_\mathbf{x}\hat{\boldsymbol{\beta}}, \quad \text{where} \quad \hat{\boldsymbol{\beta}} = \left(\mathbf{F}'\mathbf{V}^{-1}\mathbf{F}\right)^{-1}\mathbf{F}'\mathbf{V}^{-1}\mathbf{y},$$

and a term that enforces the unbiasedness constraint, i.e., that the metamodel go through the data at the designed points,

$$\mathbf{v}'_\mathbf{x}\mathbf{V}^{-1}\left(\mathbf{y} - \mathbf{F}\hat{\boldsymbol{\beta}}\right).$$

The IMSE of prediction under this model, given the true values for the θ's and normalizing for σ_z^2, is the following:

$$\text{IMSE} = \frac{1}{\sigma_z^2\Omega}\int_{\mathbf{x}\in\chi} E_\theta\left[\left(\hat{Y}(\mathbf{x}) - Y(\mathbf{x})\right)^2\right] w(\mathbf{x})dx_1 \ldots, dx_d.$$

After some algebra, starting with Eq. 7 and utilizing the two vector equations $\mathbf{Vc} - \mathbf{v_x} - \mathbf{F}\boldsymbol{\lambda} = \mathbf{0}$ and $\mathbf{F}'\mathbf{c} = \mathbf{f_x}$, as well as the matrix identity that led to Eq. 4, the IMSE, unnormalized for σ_z^2, is the following:

$$\text{IMSE} = \sigma_z^2 - trace\left\{\begin{bmatrix} \mathbf{0} & \mathbf{F}' \\ \mathbf{F} & \mathbf{V} \end{bmatrix}^{-1} \cdot \frac{1}{\Omega}\int_{\mathbf{x}\in\chi} \begin{bmatrix} \mathbf{f_x f'_x} & \mathbf{f_x v'_x} \\ \mathbf{v_x f'_x} & \mathbf{v_x v'_x} \end{bmatrix} w(\mathbf{x})dx_1 \ldots, dx_d.\right\} \tag{10}$$

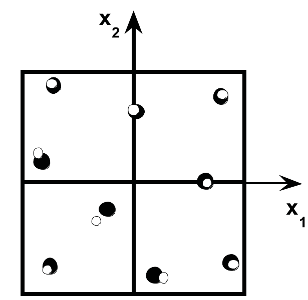

Figure 1 : I_Z-optimal design of Tab. 1 is shown plotted with black discs. The design picked off of Fig. 1a of the paper of Sacks, Schiller, and Welch (1989), and given in Tab. 2, is shown with white obscuring discs. Either design may be rotated by 90, 180, or 270 degrees about the origin to obtain other equally good designs to the corresponding unrotated design.

3.3 *Proof that as $\theta \to \infty$, I_Z-optimality $\to$ I_ε-optimality*

Referring to Eq. 10, we now prove that the conditions $\theta_i \to \infty$ and $\theta_{i,j}$ finite (∀ i,j) are sufficient for an I_Z-optimal design to be I_ε-optimal. *Proof*: In this limit, $\mathbf{V} \to \mathbf{I}$ and all the integrals are zero, except the one involving $\mathbf{f_x f'_x}$. This last integral, including the normalization, we define as $\mathbf{B}$, as in Eq. 4. Thus,

$$\lim_{\theta\to\infty} \text{IMSE} = \sigma_z^2 - trace\left\{\begin{bmatrix} \mathbf{0} & \mathbf{F}' \\ \mathbf{F} & \mathbf{I} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \mathbf{B} & \mathbf{0}' \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\right\}.$$

Only the upper-left partition of the inverse, $[0 - \mathbf{F}'\mathbf{IF}]^{-1}$, contributes to the trace, so the IMSE has the limit $\sigma_z^2 +$ trace $\left[(\mathbf{F}'\mathbf{F})^{-1} \cdot \mathbf{B}\right]$. Minimizing this is the same as minimizing the objective function for I-optimality.

Necessary conditions for the equivalence of I_Z- optimality and I_ε-optimality will be treated elsewhere. Generalization of the above proof to show that a design that is $I_{\varepsilon+Z}$-optimal is also I_ε-optimal is evident, under similar assumptions. A proof that D-optimal designs, for models of the type given in Eq. 5, maximize the minimum distance between any pair of points, i.e.,

Table 3 : I-OPT-generated, nine-point, putatively I_Z-optimal design for model in Eq. 1, assuming $\theta_1 = \theta_2 = 100$ and $\theta_{12} = 0$.

x_1	x_2
1.000	0.984
1.000	-0.121
0.749	-1.000
0.018	0.924
-0.952	-0.043
-0.869	-1.000
-1.000	1.000
*-0.008	0.103
*-0.017	-0.165

Table 4 : I-OPT-generated, nine-point, putatively I_ε-optimal design for the model in Eq. 10.

x_1	x_2
0.837	-1.000
-0.837	-1.000
-1.000	-0.044
1.000	-0.044
1.000	1.000
-1.000	1.000
0.000	1.000
*0.000	-0.045
*0.000	-0.045

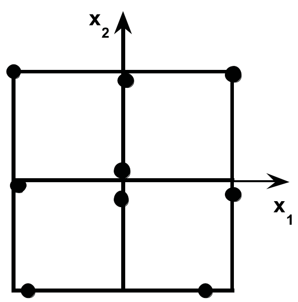

Figure 2 : I_Z-optimal design of Tab. 3. By invariance of the IMSE under interchange of the axes and reflections there are seven other equivalently good designs to the one shown.

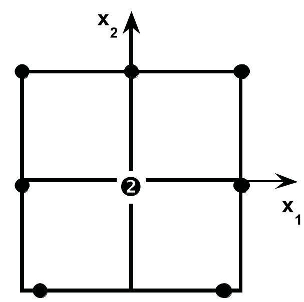

Figure 3 : I_ε-optimal design of Tab. 4. The design may be rotated by 90, 180, or 270 degrees about the origin to obtain other equally good designs to the one shown.

are "maximin", in this limit was presented by Mitchell, Sacks, and Ylvisaker (1994).

4 Validation of the I_Z-optimality capability of I-OPT

We tested I-OPT on the first problem given in Sacks, Schiller, and Welch (1989), namely, finding the nine-point I_Z-optimal design over the square $[-1,1]^2$, with θ_1, θ_2, and σ_z^2 taken as unity and $\theta_{12} = 0$. The design found using I-OPT is given in Tab. 1 and shown plotted using large black disks in Fig. 1. This design differs somewhat from the design given in the earlier paper, which is given in Tab. 2 (based on picking off points from their Fig. 1a.) and is shown plotted as small white obscuring discs in Fig. 1. I-OPT gave a normalized integrated variance NIV=0.04650 for the I_Z- optimal design and a somewhat higher value, NIV=0.04878, for the design from Sacks, Schiller, and Welch (1989).

Because of the discrepancy between our putatively optimal design and that given in Sacks, Schiller, and Welch (1989), we sought an independent check of the numerical correctness of I-OPT. We wrote a complete IMSE- evaluation program in a commercial, symbolic-manipulation software system (Maple V, Release 5) that performed all the needed matrix operations and evaluated all the required moment integrals of gaussian functions, thus obviating the need for tables of integrals. The numerical evaluations of the IMSE's as computed by I-OPT were confirmed.

As a further check on the correctness of I-OPT, we made comparison with the design given by Sacks, Schiller, and Welch (1989) in their Fig. 1b, which was for the identical problem specification as was the design in their Fig. 1a, but with $\theta_1 = \theta_2 = 100$ instead of unity. That design contains nine points, all of which were nearly on a 3×3 grid. We found a very different design, using I-OPT, as given in Tab. 3 and Fig. 2.

While this latter design has 8 points spread out approximately on a 3×3 grid, one point in the middle of one side is missing, and there is, instead, a second centrally located point (centrally located points are denoted by asterisks). While this design may seem unusual, this type of design has been seen in earlier work on I_ε-optimal designs generated using both I-OPT (see Crary, Hoo, and Tennenhouse (1992), and Crary, Clark, and Kuether (1999)) and Gosset [see Hardin and Sloane (1993)]. The second centrally located point is understood as providing additional support in the central region, where the prediction would be weak without it. Specifically, the nine-point putatively I_ε-optimal design on the two-unit square for the model

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 2 + \beta_4 x_2^2 2 + \beta_5 x_1 x_2 + \varepsilon \quad (11)$$

is as given in Tab. 4 and Fig. 3, and this design has exactly one replicated point (denoted by asterisks in the table), which is in the central region.

We made the plausible conjecture that as θ grows significantly larger than unity the I_Z-optimal designs increasingly resemble I_ε-optimal designs. Runs of I-OPT for larger and larger values of θ provided anecdotal confirmation of this conjecture. Subsequently we proved that in the limit $\theta \to \infty$ the optimal I_Z-optimal design is I_ε-optimal. The proof is given in Sec. 3.3.

5 World-Wide-Web presence

The I_Z-optimality search capability of I-OPT was confirmed to be correct for several additional anecdotal cases in one and two factors using symbolic-manipulation software. This evidence, along with the demonstration of the correct asymptotics as $\theta \to \infty$, provide the basis for the University of Michigan authors to place and maintain a demonstration version of Version 4 of I-OPT, which includes capabilities for OLS and BLUP fitting, on the World-Wide Web at URL: http://www-personal.engin.umich.edu/~crary/iopt

6 Example

To evaluate the quality of metamodels created using an I_Z- optimal design and BLUP fitting, a test case involving a micromachined flow sensor was chosen.

6.1 Description of the flow sensor

Flow sensors find applications in many fields, including industrial process control, automotive applications, security, and biomedical instrumentation, see Lerch, Dubochet, and Renaud (1997). There are many flow measurement principles, but most silicon flow sensors are based on thermal effects. Fig. 4 is a photomicrograph of a thermal anemometer developed by Leister Process Technologies in Switzerland. A silicon nitride membrane (0.3 *μm* thick) is fabricated by anisotropic backside etching of silicon (0.5 *mm* thick). Three nickel-film thermoresistors are structured on the dielectric membrane and covered with a silicon nitride passivation layer (0.2 *μm*). The package includes a channel (3 *mm* deep and 0.8 *mm* high) so that gas flows perpendicularly to the thermoresistors.

The principle of a hot wire anemometer is based on correlating the heat transfer from a heated wire to a fluid with the rate of flow. The energy losses due to a moving fluid increase with fluid velocity. In the device described above, the center wire is the heating element and the outside wires are resistive temperature sensors. An electronic circuit establishes the average temperature of the upstream and downstream thermoresistors and maintains the heater at a constant temperature above this average. Such a sensor can detect flows from less than 0.10 ml/min up to 20 ml/min, the time constant is less than 1 *ms*, and the power consumption of the sensor is 200 *mW*.

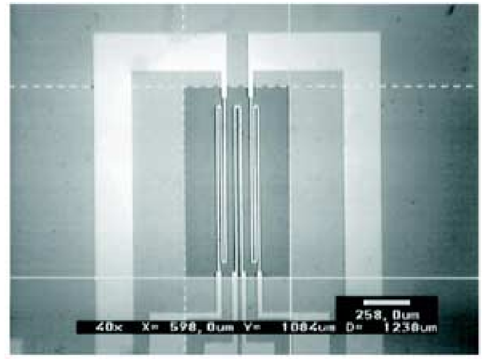

Figure 4 : Photomicrograph of the gas flow sensor. The three serpentine thermoresistors are shown on the (dark) membrane region.

6.2 Modeling

A two-dimensional, finite-element model of the cross section perpendicular to the thermoresistors in the direction of flow was developed, and the ANSYS commercial finite-element package was used to simulate the thermal behavior of the sensor. The model contained the silicon-nitride membrane with its passivation layer, supported at both ends by the anisotropically etched silicon. The flow to be measured passes above the membrane, enters at ambient temperature, and is considered fully developed in a parabolic profile. The air that is trapped below the membrane and between the silicon walls is modeled as stationary. The outside edges of the bulk silicon are kept at ambient temperature. The nickel film thermoresistors are not included in the model, so the appropriate heating boundary conditions (either heat flux or constant temperature) are applied directly to the silicon nitride membrane. An example of a simulated temperature distribution (logarithmic scale), for a specific flow rate, membrane thickness, and a heater temperature of 50 ^{o}C above ambient is shown in Fig. 5.

6.3 Objectives of the metamodel generation

In order to make contact with the previously published paper of Sacks, Schiller, and Welch (1989), we chose to build metamodels using various nine-point designs in two factors and both OLS and BLUP fitting. In a conference report [Crary, Cousseau, Armstrong, Woodcock, Dubouchet, Lerch, and Renaud (1999)] we had chosen the two factors to be the membrane thickness and thermoresistor-sensor separation. In the present paper, we chose to replace the membrane thickness with a factor that was more challenging to model due to nonlinear effects, namely the flow speed. As in the earlier report,

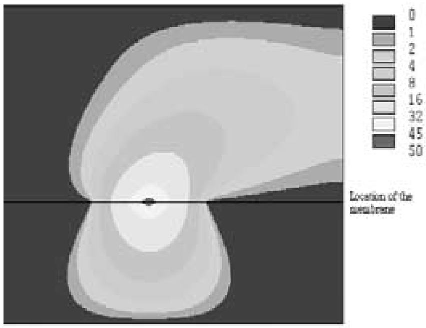

Figure 5 : An example of the above-ambient temperature distribution, as determined by FEA. The heater is located at the middle of the membrane, and flow is present above the membrane. The temperature key is in degrees Celsius.

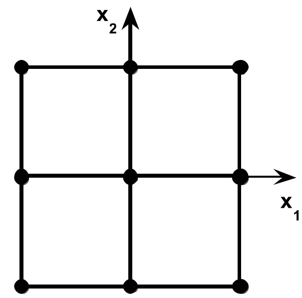

Figure 6 : 3^2 factorial design. IMSE for model (1) is 0.1227.

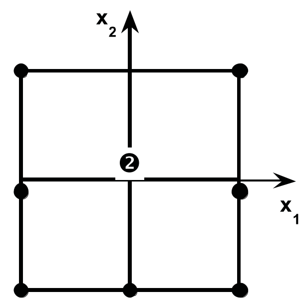

Figure 7 : A-optimal design. The design may be rotated by 90, 180, or 270 degrees about the origin to obtain other equally good designs to the one shown.

the single response chosen was the temperature difference between the upstream and downstream wires.

We already had a nominal design and were seeking a metamodel that could be used for variations from the nominal design, as might be particularly useful in design synthesis. We chose ranges of the factors that would be challenging for simple second-degree bivariate polynomial functional approximation, but for which an effective nine- point design might be expected to perform reasonably well. The flow rate varied from 0 to 20 ml/min, and the distance of the thermoresistors from the heater varied from 50 μm to 290 μm.

6.4 Experimental designs

In addition to the I_Z-optimal design found with I-OPT (Tab. 1 and Fig. 1), the design picked off of Fig. 1a of Sacks, Schiller and Welch (1989) (Tab. 2 and Fig. 1)), and the 3^2 factorial design shown in Fig. 6, we used a variety of other designs, including I-OPT-generated I_Z-optimal designs based on simpler model functions and some rotations of these designs, which are also optimal due to reflection and permutation symmetries of the problem statement. In addition, we explored the designs given in this sub-section. In all cases, as in the earlier study, the region of prediction was the same as the design region. The reported values of IMSE assume model (1).

6.4.1 Optimal designs of non-deterministic experiments for parameter estimation

These designs do not make reference to the region of prediction.

- *A-optimality* (Fig. 7): minimize the trace of $(\mathbf{F}'\mathbf{F})^{-1}$
  Design = $[(\pm1,-0.1017),(\pm1,\pm1),(0,-1),2\times(0,0.1671)]$
  Found with I-OPT. IMSE=0.1762

- *D-optimality*: minimize the determinant of $(\mathbf{F}'\mathbf{F})^{-1}$
  [same as 3^2 factorial]
  Found with I-OPT. IMSE=0.4500

6.4.2 Optimal designs of non-deterministic experiments for prediction

These designs make explicit reference to the region of prediction.

- *G-optimality* (Fig. 8): minimize the maximum expected variance, $\mathbf{f}'(\mathbf{F}'\mathbf{F})^{-1}\mathbf{f}$, over the region of prediction
  $[(\pm1,\pm1),\ (0,0),\ (-1,0.4190),\ (1,-0.4190),\ (-0.4190,-1),(0.4190,1)]$
  Given in Haines (1987). IMSE=0.1554

- *I_ε-optimality* (Fig. 3): minimize the average (taken as an integral) of the expected variance, $\mathbf{f}'(\mathbf{F}'\mathbf{F})^{-1}\mathbf{f}$, over the region of prediction
  $[(\pm1,-0.0442),\ 2\times(0,-0.0447),\ (0,1),\ (\pm1,1),\ (\pm0.8367,-1)]$
  Found with I-OPT. IMSE=0.1579

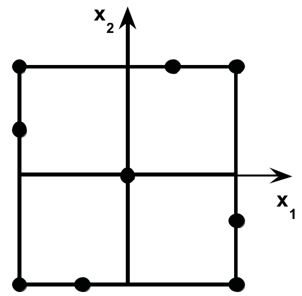

Figure 8 : G-optimal design. An equivalently good design may be obtained by a reflection, resulting in a design with the opposite chirality.

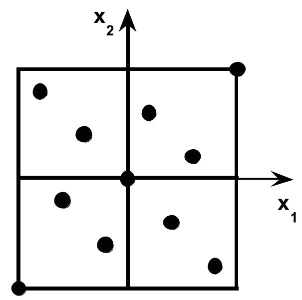

Figure 9 : Latin hypercube design. An equivalently good design may be obtained by a reflection about either axis.

6.4.3 *Designs that are based on spreading the points apart (spatial designs):*

- *S-optimality*: maximize the geometric mean of the distances of points from their nearest neighbors [same as 3^2 factorial]
  Found with I-OPT.

- *Maximin criterion*: maximize the minimum distance of any point from its nearest neighbors [same as 3^2 factorial]
  Found with I-OPT.

- *Maximum entropy sampling*: Shewry and Wynn (1987) define this criterion based on information theory and show that for deterministic experiments it is equivalent to D-optimality. [same as 3^2 factorial]

6.4.4 *Designs that have high degrees of symmetry by definition (classical designs):*

- *Latin hypercube design* (Fig. 9): after dividing the design region into a regular N × N mesh, choose a design that has exactly one point in each row and exactly one point in each column
  $[N = 11 : (-1,-1), (-0.8,0.8), (-0.6,-0.2), (-0.4,0.4), (-0.2,-0.6), (0,0), (0.2,0.6), (0.4,-0.4), (0.6,0.2), (0.8,-0.8), (1,1)]$
  Welch, Buck, Sacks, Wynn, Mitchell, and Morris (1992) used this low-IMSE design. IMSE=0.0423

- *Central-composite design* (Fig. 10): points are taken at the 2^k vertices of the k-dimensional cuboidal region of prediction, at $2k$ locations a distance a from the center of the region along the axes in the positive and negative directions, and at n_c points at the origin. For deterministic experiments, $n_c = 1$, as duplicate points are non-informative. For our example, no design points were allowed outside of the two-unit square, so a CCD was used with the $2k$ points on the boundary and the 2^k vertex points chosen for convenience as points included in other designs, as follows:
  $[(0,0), (\pm1,0), (0,\pm1), (\pm0.7530,\pm0.7530)]$
  IMSE=0.0822

Table 5 : Coefficients of fitting equation (11)

Index	β	γ
0	0.6078739017	
1	0.3223577673	0.2464299622
2	0.0579050968	-0.1010820946
3	-0.1131439765	-0.1517626950
4	-0.3357223872	-0.3714921774
5	-0.0059478897	0.3781090349
6		0.0408797695
7		0.2062421237
8		-0.1856936929
9		-0.0616302248

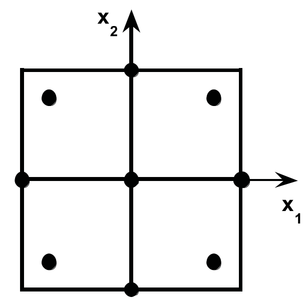

Figure 10 : Central composite design.

7 Simulations

The thickness of the membrane was 0.5 μm, and the heater temperature was kept at a constant 50 ^oC above ambient. The temperature difference between the centers of the upstream and downstream wires was computed as the response of interest for the metamodel generation. As stated previously, the

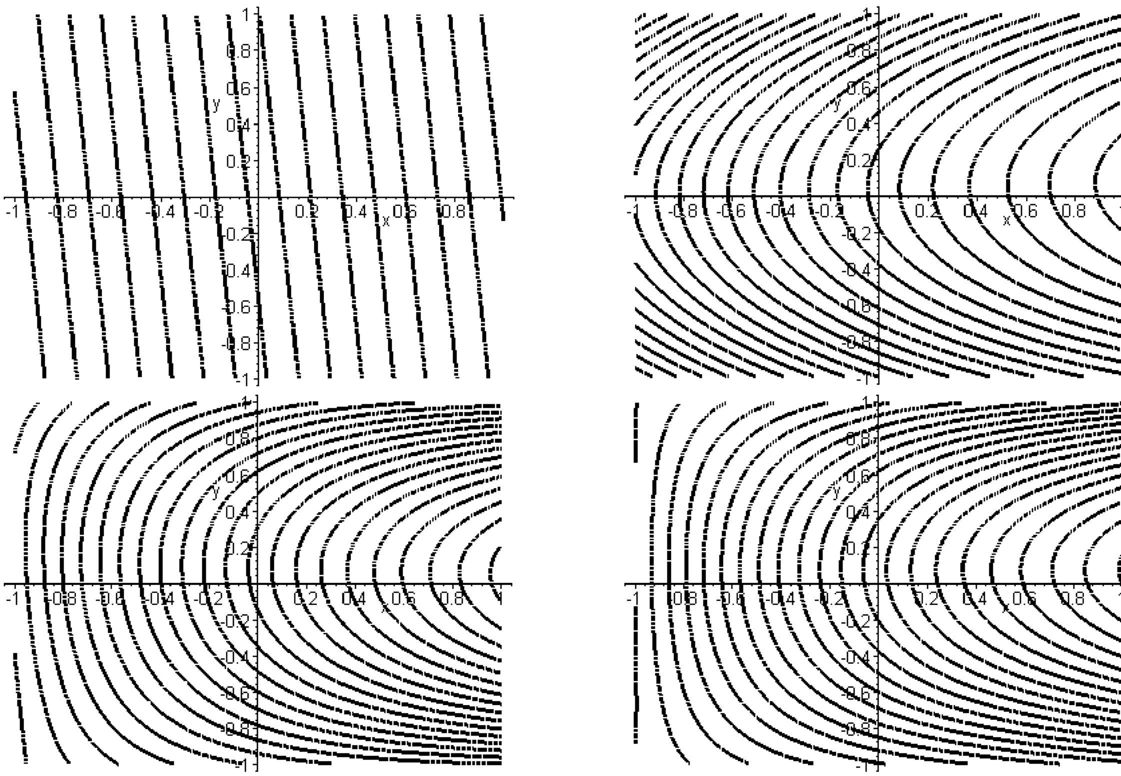

Figure 11 : Contours of the functional approximation based on a series of OLS fits to 121 responses on a regular 11x11 mesh. The contour values increase from left to right for all fits. Upper left: first-degree fit, contours from 0.25 to 1.00, e.r.m.s. error 160mK. Upper right: second-degree fit, contours from -0.20 to 0.90, e.r.m.s. error 78mK. Lower left: third-degree fit, contours from 0.00 to 1.05, e.r.m.s. error 9.6mK. Lower right: fourth-degree, contours from 0.00 to 1.05, e.r.m.s. error 1.1mK.

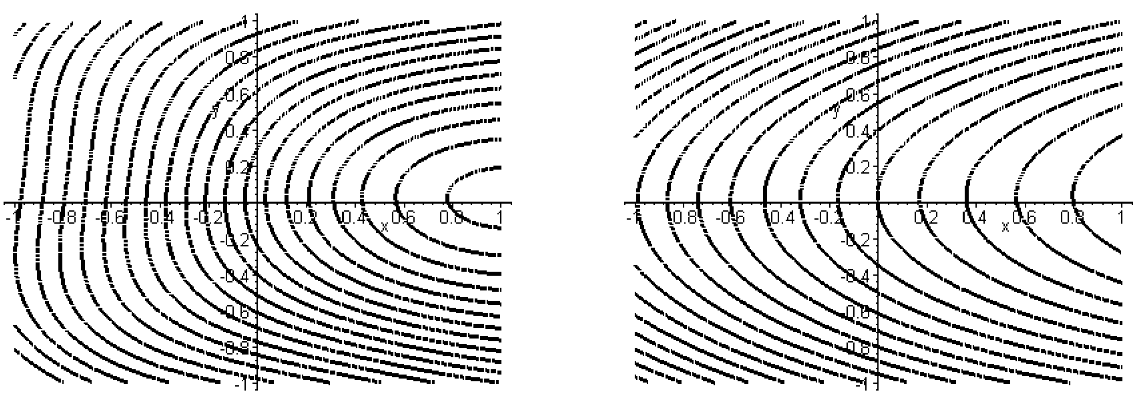

Figure 12 : Left: contours of the functional approximation based on the BLUP fit (full-second-degree model + $Z(x)$) to the 9- point I_Z-optimal design of Tab. 1; contours from -0.05 to 0.95; e.r.m.s. error over 121-point mesh of validation points was 38.5mK. Right: similar plot based on the full-second-degree OLS fit to the 3^2 factorial design; contours from -0.10 to 0.80; e.r.m.s. error over 121-point mesh of validation points was 97.2mK. The superiority of the combination of I_Z-optimal design and BLUP fit, compared to the 3^2 factorial design and OLS fit combination, is evident.

flow rate varied from 0 to 20 ml/min, and the distance of the thermoresistors from the heater varied from 50 μm to 290 μm.

Simulations were run at all the design points mentioned above, as well as on a 11×11 regular grid covering the square, for validation.

8 Analyses

The data were fit using OLS, first using just a constant, then with a first-degree bivariate polynomial, and then again with a full-second-degree bivariate polynomial. Then the data sets

Table 6 : Empirical root-mean-square values for the residuals (in mK) of various fits based on various designs. The maximum value of the response (temperature difference) over the range was 1.5K

DESIGNS	ANALYSES (FITS)					
(all N=9, except Latin hypercube with N=11)	OLS deg = 0	OLS deg = 1	OLS deg = 2	BLUP deg = 0+Z	BLUP deg = 1+Z	BLUP deg = 2+Z
3^2 factorial, D-optimal, S-optimal, Maximin, and Max entropy	314.7	205.8	97.2	60.4	47.3	28.5
A-optimal	301.8	184.3	98.4	57.3	43.9	31.3
G-optimal	321.8	224.0	113.1	75.6	64.8	50.1
I_ε-optimal	300.0	182.4	93.1	59.6	48.1	31.2
Latin hypercube, N=11	292.3	172.0	101.1	70.8	61.4	49.1
CCD	294.9	165.7	81.3	58.5	54.2	35.3
I_Z-optimal (6 terms + $Z(x)$) $\beta_o+\beta_1x_1+\beta_2x_2+\beta_3x_1^2+\beta_4x_2^2+\beta_5x_1x_2+Z(x)$	292.3	159.9	81.3	70.3	55.1	38.5
rotated ccw 90 ^{o}	292.2	159.5	81.1	56.3	45.8	38.4
rotated ccw 180 ^{o}	292.2	159.8	80.7	57.7	48.3	40.1
rotated ccw 270 ^{o}	292.3	159.8	80.7	68.3	53.2	35.2
I_Z-optimal (3 terms + $Z(x)$) $\beta_o+\beta_1x_1+\beta_2x_2+Z(x)$	292.5	160.8	81.9	60.0	53.0	43.4
rotated ccw 90 ^{o}	292.4	160.0	82.2	74.1	59.3	38.9
I_Z-optimal (1 term + $Z(x)$) $\beta_o+Z(x)$	292.6	160.4	79.6	72.6	55.4	36.9
rotated ccw 90 ^{o}	292.7	160.4	82.9	74.2	60.2	37.5
Sacks, Schiller, Welch	293.5	161.9	81.1	65.7	52.2	37.5

were fit using the BLUP, first using a model with constant plus unmodeled part $Z(x)$, then with a first-order bivariate polynomial plus $Z(x)$, and finally with a full-second-degree bivariate function plus $Z(x)$.

For the case of the full-second-degree bivariate function plus $Z(x)$ BLUP fit, the fit function has $6+N$ terms, as follows for $N=9$:

$$Y=\beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_1^2+\beta_4x_2^2+\beta_5x_1x_2+\sum_{i=1}^{9}\gamma_i\exp\left(-\left(x_1-s_1^{(i)}\right)^2-\left(x_2-s_2^{(i)}\right)^2\right), \tag{12}$$

where the $s_1^{(i)}$ and $s_2^{(i)}$ are the x_1 and x_2 coordinates of the i'th design point, respectively. The coefficients for the I_Z-optimal design of Tab. 1 (e.g., $s_1^{(1)}=-0.719$ and $s_2^{(1)}=0.874$) are given in Tab. 5. Eq. 12 can be rapidly evaluated in its present form or can be rearranged to be evaluated with fewer numerical operations. As described in the Sec. 3, the BLUP passes through all of the data and provides an interpolation elsewhere.

The empirical root mean square (e.r.m.s) errors of the ninety fits over the 121 validation points are reported in Tab. 6.

9 Interpretation

As a means of exploring the response function, we made a series of contour plots of the function based on fits of increasing degree, up to degree four, and these are shown in Fig. 11. The final fit in the lower-right-hand side of this figure is representative of the actual function, and shows, in particular, the zero response along the left border, where the value of x_1 (the flow speed) is at its minimum value (zero). The difficulty that bivariate polynomials up to third degree have in incorporating this characteristic of the response is evident. The BLUP fits do considerably better, in large measure due to their extra degrees of freedom.

In Fig. 12 are shown the contours based upon the second-degree plus $Z(x)$ BLUP fit to the responses at the 9- point I-OPT-generated design points for the full-second- degree model function plus $Z(x)$. It can be seen that the BLUP fit is approximately as good as the third-degree OLS fit in approximating the response function. For comparison, the contours of the relatively poorer second-degree OLS fit to the 3^2 factorial are also shown.

The earlier finding that the I_Z-optimal designs fit with BLUP were superior to designs for non-deterministic error- model functions (see Sacks, Schiller, and Welch (1989)) was specifically upheld in the analyses. For example, the OLS second-degree fit for the factorial design gave an e.r.m.s error of 97.2 *mK*, whereas the average e.r.m.s error of the four (similar but rotated) BLUP fits assuming the second-degree model plus $Z(x)$ based on the design assuming a second- degree model plus $Z(x)$ was only 38.1 *mK*, for a reduction in average variance of $(97.2/38.1)^2 = 6.5$.

Also, the higher the assumed order of the I_Z-optimal models for which designs were found and subsequently fit, the better the fit. At zero, first, and second degree, the e.r.m.s errors, averaged over rotations, as given in Tab. 6 were 73.4, 56.2, and 38.1 *mK*, respectively. This anecdotal finding is in contrast to statements made and repeated elsewhere, for example in Welch, Buck, Sacks, Wynn, Mitchell, and Morris (1992); Bernardo, Buck, Liu, Nazaret, Sacks, and Welch (1992); Etman (1994); and Simpson, Mauery, Korte, and Mistree (1998) and references cited therein; where use of the simplest model function, $Y = \beta_0 + Z(x)$ is promoted.

The question of whether this improvement is the result of the type of design used, or just a matter of a better fitting method, requires further exploration. Generally speaking, the power of the BLUP fits is evident even with the simplest fit, i.e., zero-degree polynomial (constant) plus $Z(x)$, since for every design in the table the BLUP outperformed the best of the OLS second-degree polynomial fits. In addition, with increasing order, the BLUP fits improved further, although the I_Z-optimal designs were not notably superior to the other designs, and, in fact, the I_Z-optimal design based on the zero- degree (constant) model plus $Z(x)$ outperformed, perhaps surprisingly, both of the other classes of models used for generating I_Z-optimal designs, when all the analyses were performed with second-degree fits.

We find that additional research will be needed to elucidate the best design of experiments approach for this anecdotal case, and in this regard, we agree with Salagame and Barton (1997) who reached the same conclusion for a different problem.

We note that we did not see any appreciable difference in the e.r.m.s. error of the fits based on designs that were rotated versions of their base design.

There is also the issue of the computational resources required to find the best values of the θ's via maximum likelihood and to search for the optimal designs using an optimal-design search engine, such as I-OPT. This issue will be discussed at length in a future paper, but it is clear that the required resources grow rapidly with problem size, although this can be alleviated by the creation and use of stored public libraries of optimal designs, which will be an inevitable development in information technology. Alternative design approaches that avoid the computational burdens, in addition to those mentioned in this paper, have been proposed, see Salagame and Barton (1997) for factorial hypercube designs and Kalagnanam and Diwekar (1997), and references therein, for low-discrepancy designs, e.g., Hammersley points.

10 Conclusions

I-OPT will provide a heretofore-missing public tool for researchers investigating the use of optimal-design-of- experiments approaches to deterministic experimentation.

We propose that the term "I-optimality" be used for IMSE-based optimality in both non-deterministic and deterministic settings and that "I_ε-optimality" and "I_Z- optimality" be adopted for the more specific settings. This proposal is based on our proof that the I_ε-optimality is a limiting case of I_Z-optimality, as well as practical considerations.

In our example of the use of various designs of experiments and fitting methods for metamodel generation for a specific device, we found that the inclusion of explicit bivariate-polynomial terms in Eq. 1 was very helpful in improving the fit.

The design that performed best for our example problem was not the I_Z-optimal design. Further investigation is needed to fully elucidate the design issues for this and similar problems. The value of capturing salient engineering knowledge before applying a black-box approach should not be underestimated.

Acknowledgement: This work was supported, in part, by a contract from Tanner Research, Pasadena, CA and by the MINAST (Micro and Nano Systems) MicroSim 4.01 project, a priority program funded by the Swiss government. The first author thanks the following for pointing out relevant literature: M. Sasena of the Univ. of Michigan, Prof. T. Allen of Ohio State Univ., Prof. T. Simpson of Penn. State Univ., and Prof. W. Welch of the Univ. of Waterloo.

References

Aslett, R.; Buck, R. J.; Duvall, S. G.; Sacks, J.; Welch, W. J. (1998): Circuit optimization via sequential computer experiments: design of an output buffer. *Appl. Statist.*, vol. 47, pp. 31–48.

Barton, R. R. (1992): Metamodels for simulation input-output relations. In Swain, J. J.; Goldsman, D.; Crain, R. C.; ; Wilson, J. R.(Eds): *Proc. 1992 Winter Simulation Conf.*, pp. 289–299.

Belytschko, T.; Lu, Y. Y.; Gu, L. (1994): Element-free Galerkin methods. *Int'l. J. Numer. Methods Engrg.*, vol. 37, pp. 229–256.

Bernardo, M. C.; Buck, R.; Liu, L.; Nazaret, W. A.; Sacks, J.; Welch, W. J. (1992): Integrated circuit design optimization using a sequential strategy. *IEEE Transactions on Computer-Aided Design*, vol. 11, pp. 361–372.

Chapman, W. L.; Welch, W. J.; Bowman, K. P.; Sacks, J.; Walsh, J. E. (1994): Arctic ice variability: model sensitivities and a multidecadal simulation. *J. Geophysical Res.*, vol. 99, pp. 919–935.

Crary, S.; Cousseau, P.; Armstrong, D.; Woodcock, D.; Dubochet, O.; Lerch, P.; Renaud, P. (1999): Optimal design of computer experiments for the generation of microsystem macromodels using IMSET^{TM} and non-parametric fitting. In *Proc. Second International Conference on Modeling and Simulation of Microsystems*, pp. 184–189, San Juan, PR.

Crary, S.; Cousseau, P.; Mok, E.; Woodcock, D.; Renaud, P. (1999): Critical comparison of design- and analysis-of-experiments methods for MEMS metamodel generation. In *SPIE Conference on Micromachined Devices and Components V*, volume 3876, Santa Clara, CA.

Crary, S. B.; Clark, J. R.; Kuether, K. (1999): *I-OPT User's Manual*. on-line version available at URL http://www-personal.engin.umich.edu/∼crary/iopt, see Ex. 2.

Crary, S. B.; Hoo, L.; Tennenhouse, M. (1992): I-optimality algorithm and implementation. *Computational Statistics*, vol. 2, pp. 209–214.

Crary, S. B.; Phan, K. N. (1998): The COBBLE^{TM} algorithm: macromodels for MEMS via stepwise, sequential, design of experiments for patched functional approximation. In *Late News Poster Session Supplementary Digest, Solid-State Sensor and Actuator Workshop*, pp. 1–2, Hilton Head Island, SC.

Cressie, N. (1990): The origins of kriging. *Math. Geol.*, vol. 22, pp. 239–252.

Currin, C.; Mitchell, T.; Morris, M.; Ylvisaker, D. (1991): Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments. *J. Amer. Stat. Assoc.*, vol. 86, pp. 953–963.

Etman, L. F. P. (1994): Design and analysis of computer experiments: the method of Sacks et al. In *Engineering Mechanics Report WFW 94.098*, Eindhoven University of Technology.

Gianchandani, Y. B.; Crary, S. B. (1998): Parametric modeling of a microaccelerometer using I-optimal design of experiments for finite-element analysis. *IEEE J. of MEMS*, vol. 7, pp. 274–282.

Gough, W. A.; Welch, W. J. (1994): Parameter space exploration of an ocean general circulation model using an isopycnal mixing parameterization. *J. Marine Res.*, vol. 52, pp. 773–796.

Haines, L. M. (1987): The application of the annealing algorithm to the construction of exact optimal designs for linear-regression models. *Technometrics*, vol. 29, pp. 439–447.

Hardin, R. H.; Sloane, N. J. A. (1993): A new approach to the construction of optimal designs. *J. Statistical Planning and Inference*, vol. 37, pp. 339–369.

Kalagnanam, J. R.; Diwekar, U. M. (1997): An efficient sampling technique for off-line quality control. *Technometrics*, vol. 39, pp. 308–319.

Kennedy, B. W. (1991): C. R. Henderson: the unfinished legacy. *J. Dairy Sci.*, vol. 74, pp. 4067–4081.

Koehler, J. R.; Owen, A. B. (1996): Computer experiments. In Ghosh, S.; Rao, C. R.(Eds): *Design and Analysis of Experiments*, pp. 261–308, New York. Elsevier.

Lancaster, P.; Salkauskas, K. (1981): Surfaces generated by moving least squares methods. *Mathematics of Computation*, vol. 37, pp. 141–158.

Lerch, P.; Dubochet, O.; Renaud, P. (1997): From simple devices to promising microsystem applications. In *IEEE Proceedings of the 21st International Conference on Microelectronics (MIEL '97)*, pp. 59–62, Nug, YUG.

Lu, Y. Y.; Belytschko, T.; Gu, L. (1994): A new implementation of the element free Galerkin method. *Comput. Methods Appl. Mech. Engrg*, vol. 113, pp. 397–414.

Miller, D.; Frenklach, M. (1983): Sensitivity analysis and parameter estimation in dynamic modeling of chemical kinetics. *Int. J. of Chem. Kinetics*, vol. 15, pp. 677–696.

Mitchell, T.; Sacks, J.; ; Ylvisaker, D. (1994): Asymptotic bayes criteria for nonparametric response surface generation. *Annals of Statistics*, vol. 22, pp. 634–651.

Sacks, J.; Schiller, S. B.; Welch, W. J. (1989): Design of computer experiments. *Technometrics*, vol. 31, pp. 41–47.

Sacks, J.; Welch, W. J.; Mitchell, T. J.; Wynn, H. P. (1989): Design and analysis of computer experiments. *Statistical Science*, vol. 4, pp. 409–435.

Salagame, R. R.; Barton, R. R. (1997): Factorial hypercube designs for spatial correlation regression. *J. Appl. Stat.*, vol. 24, pp. 453–473.

Searle, S. R.; Casella, G.; McCulloch, C. E. (1992): *Variance components*. Wiley, New York.

Senturia, S. D. (1998): CAD challenges for microsensors, microactuators, and microsystems. *Proc. IEEE*, vol. 86, pp. 1611–1626.

Senturia, S. D.; N., A.; White, J. (1997): Simulating the behavior of MEMS devices: computational methods and needs. *IEEE Computat. Sci. and Engrg.*, vol. 4, pp. 30–43.

Shewry, M. C.; Wynn, H. P. (1987): Maximum entropy sampling. *J. Appl. Stat.*, vol. 14, pp. 165–170.

Simpson, T. W. (1998): *A concept exploration method for product family design*. dissertation, Georgia Institute of Tech., Atlanta, GA.

Simpson, T. W.; Allen, J. K.; Mistree, F. (1998): Spatial correlation metamodels for global approximation in structural design optimization. In *Proc. 1998 ASME Design Engineering Technical Conference*, pp. 1–13, Atlanta, GA.

Simpson, T. W.; Mauery, T. M.; Korte, J. J.; Mistree, F. (1998): Comparison of response surface and kriging models for multidisciplinary design optimization. In *AIAA-98-4755, 7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, volume 1, pp. 381–391, St. Louis, MO.

Welch, W. J.; Buck, R. J.; Sacks, J.; Wynn, H. P.; Mitchell, T. J.; Morris, M. D. (1992): Screening, predicting, and computer experiments. *Technometrics*, vol. 34, pp. 15–25.

Welch, W. J.; Yu, T.-K.; Kang, S. M.; Sacks, J. (1990): Computer experiments for quality control by parameter design. *J. Quality Technol.*, vol. 22, pp. 15–22.

Wold, H. (1938): *A study in the analysis of stationary time series*. Almqvist and Wiksells, Uppsala, Chapter II, Sec. 18.