Designing Efficient Computer Experiments for Metamodel Generation

Selden B. Crary [1,2], David M. Woodcock[3], and Andreas Hieke[1]

[1]Crary Group, WebDOE™ Division, 320 Riverview Drive, Ann Arbor, MI 48104, USA
Voice: +1 734 995 1447, selden@crary.cc
[2]EECS Department, University of Michigan, Ann Arbor, MI, USA
[3]Laboratory for Advanced Scientific Computation, University of Michigan, Ann Arbor, MI, USA

ABSTRACT

We demonstrate the relative merits of I_Z-optimal designs for computer experiments relative to Latin hypercube designs, using two examples, both in two factors and eleven points, one a simple function from Welch et al. [1] and the other a micro-flowmeter response presented earlier by Crary et al. [2]. The relative reductions in variance of fitting are large, viz., 1.85 and 2.24, respectively. We show that, contrary to prevailing views, IMSE-optimal designs can have proximal design points, which we call "twin points," with the interesting property that the design is, in part, determined by the information resources available. We comment on this newly established link between optimality and information content. We outline a particle-interaction theory and demonstrate that ground-state twin points are observed in a simple particle-dynamics simulator, when the inter-particle potential is repulsive and growing with range, and there is a centrally directed single-particle force.

Keywords: IMSE-optimality, I-OPT, optimal design for computer experiments, kriging, metamodel

1 INTRODUCTION

We are concerned with the efficient generation of parsimonious surrogate metamodels from expensive deterministic simulations, using designed experimentation. Such metamodels may be used for a variety of purposes, such as assisting in the understanding of the function under study, optimizing a product or process under single or multiple objectives, or for rapid evaluation of the function within a computer-aided design environment.

In previous work, we demonstrated the use of I-OPT™ for finding designs for deterministic experiments under the objective of minimizing the integrated mean squared error (IMSE) of prediction of the metamodel over a domain of interest [2]. We named these designs I_Z-optimal designs, after demonstrating their relation to the class of response-surface designs with non-deterministic error models that are known in the statistics literature as I-optimal designs.

The statistical model considered both in the earlier work and in the present study is that the function of interest may contain both a known polynomial part and an unknown part. For example, in two-factors the model function may be the following:

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

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

$$\mathrm{cov}[Z(s_1, s_2), Z(t_1, t_2)] = \sigma_Z^2 \exp\{-[\theta_1 (s_1-t_1)^2 + \theta_2 (s_2-t_2)^2 + \theta_{12} (s_1-t_1)(s_2-t_2)]\},$$

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, et al. [3], 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. For an excellent review, see Koehler and Owen [4].

A difficulty with the earlier work was that we assumed, for convenience, that the values of the prefactors of the covariance were $\theta_1=\theta_2=1$ and $\theta_{12}=0$. The result was that all designs studied performed nearly equally well in generation of predictive metamodels, despite distinct differences in their theoretical expected integrated-mean-squared-error values. In particular, I_Z-optimal designs were generally no better than other classes of designs.

We now report our investigation of the relative performances of I_Z-optimal and Latin hypercube (LHC) designs after maximum-likelihood determination of the values of θ_1 and θ_2. We find that the theoretical performance is now matched by the significantly improved empirical performance, and this suggests the use of I_Z-optimal designs for metamodel generation when function evaluations are expensive.

2 DESIGNS

I_Z-optimal designs specify a set of inputs to a function-evaluation program so that upon fitting the results with a best-linear-unbiased-predictor (BLUP) fitting routine, the generated metamodel has the minimum expected integrated mean square error of prediction of the function at untried inputs. Such designs are computationally expensive, but recent advances in computational capabilities are encouraging that I_Z-optimal designs may soon become readily available [2].

LHC designs are based on a particular spatial spreading of points. In an N-point LHC design each factor range is divided into N levels, thus forming N^k (hyper-)rectangular cells in k factors. Points are placed in the cells in a way ensuring exactly one point at each level of each factor. An often-mentioned feature of LHC designs is that if the function of interest is independent of one or more factors, then after removal of the irrelevant factors, the projection of the LHC to the reduced design space maintains good spatial properties [4].

3 DIMENSIONAL SCALING

In order to make comparisons with literature sources, it is important to properly scale the quantities of interest. For example, the region considered by Sacks et al. [3] was $[0,1]^k$, by Welch et al. [1] was $[0,5]^k$, and by Crary et al. [2] was $[-1,1]^k$. A simple example will suffice. In the example presented on p. 17 of Welch et al. [1], a maximum-likelihood evaluation gave values of $[\theta_1,\theta_2]=[0.128,0.069]$. Since the θ's appear in the exponent multiplied by a square of the factors, the corresponding θ's when using $[-1,1]^k$ will be $[0.128,0.069]*(5/2)^2=[0.800,0.43125]$. The quantity σ_Z^2 scales with the output variable, so if the same output units are used in a comparison, σ_Z^2 will be invariant, that is, unchanged after a change in the scale of the factors.

4 EXISTENCE OF NON-SPATIAL IMSE-OPTIMAL DESIGNS – "TWIN POINTS"

We digress to report a recent discovery.

It is easily understood that repeated running of a deterministic simulator at a given set of inputs is non-informative, and it has been generally assumed, as a sort of informal corollary, that design-optimality criteria for deterministic experiments have a tendency to keep points from being close to one another. However, we have recently discovered and reported [5] a perfectly ordinary case for which the optimal design specification is that two sets of inputs should be taken as close together as possible, without actually being identical, and we say such a design contains "twin points."

The design was found using a quadruple-precision version of I-OPT, for two factors, eleven points, on the design region $[-1,1]^2$, and using the deterministic model $Y = \beta_0 + Z(x)$, with $[\theta_1,\theta_2]=[0.128,0.069]$. The design is shown in Fig. 1. Other designs with twin points have been observed for different values of the θ's, including a set with different symmetry properties from the one reported here. In fact, a complex phase diagram can be constructed for this system.

How can this unusual type of design be understood? First, it is clear that when two points become identical, there may be a step singularity in the objective function, since the design reduces from a specification of N points to a specification in N-1 points. Thus, when we perform a search for an optimal design, we should allow for such a discontinuity and search the (N*k)-dimensional search space considering that it contains (hyper-)planes of possibly step-wise singularities wherever two design points coincide.

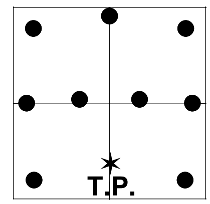

Figure 1: Twin points (labeled "T.P.") discovered in the IMSE-optimal design for the criterion and parameters described in the text immediately above. There is one equivalent design related to the one shown by a rotation of π radians about the origin.

4.1 Supporting Evidence for Twin Points

We now provide three pieces of evidence, beyond simply increasing the precision of floating-point operations, that support our discovery and interpretation of twin points.

First, the IMSE of the single, N=1 design point in one factor was determined symbolically and then evaluated numerically over the domain [-1,1]. As shown in Fig. 2, the IMSE function has a clear minimum at the origin, where IMSE=0.5063.... We could conceive of a restoring force acting to return an errant point to the origin.

Second, for the same model but with N=2 points, one of which is constrained at the origin, another symbolic computation, along with a final-step numerical evaluation, showed that the IMSE had a well-defined limit, as the location of the second point approaches the first, i.e., as $x^{(2)}\rightarrow x^{(1)}=0$. See Figure 3. Indeed the slope is zero in the relevant limit:

$$\left(\frac{\partial IMSE}{\partial x^{(2)}}\right)_{\left(x^{(1)}\equiv 0;\ x^{(2)}\rightarrow 0^+ \text{ or } 0^-\right)} = 0$$

$$\left(IMSE\right)_{\left(x^{(1)}\equiv 0,\, x^{(2)}\rightarrow 0^+ \text{ or } 0^-\right)} = 0.2749...$$

Also, when both points are identically at the origin, the problem reduces to the one-point problem of Fig. 2, with the IMSE jumping to the value 0.5063... . Allowing the second point to move, the symbolic computation gives the same two-point design as was found numerically using I-OPT, that is, $[x^{(1)},x^{(2)}]=[0,0.625...]$. We could conceive of a weak repulsive force pushing the two proximal points apart.

Third, we can introduce the concept of an inter-design-point "potential energy" and a "zero-point energy" and consider the resulting Newtonian dynamics of the design points. We are guided by the semi-classical physics of systems of particles with dominant two-particle interactions and zero-point energies arising from confinement of individual particles' quantum-mechanical wave functions to a fixed (hyper-)cubical region. We will present our predictive theory in greater detail elsewhere and provide only an outline here.

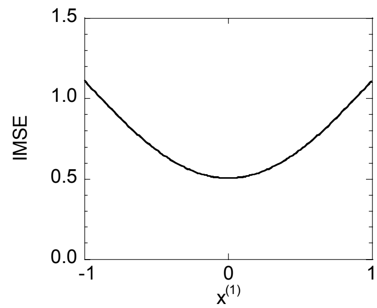

Figure 2: IMSE of a single design point in one factor for the model $Y = \beta_0 + Z(x)$ with $[\theta_1,\theta_2]=[1.0,1.0]$

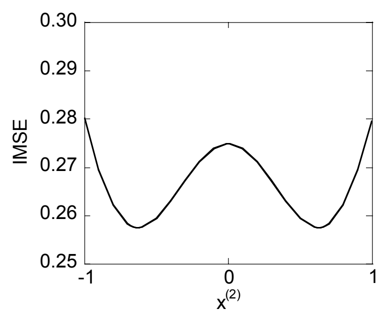

Figure 3: IMSE of a pair of design points, one of which is constrained to lie at the origin.

We find it useful to associate the negative inverse of the IMSE function as the nearly additive "energy" of the system. For the two-particle case of Fig. 3 that has the first point held at the origin (x_1=0), we introduce a two-particle interaction, $-(I_{1,2})^{-1}$, as the total energy of the two-particle system (written as $-I_2^{-1}$) minus the two single-particle zero-point energies:

$$-I_{1,2}^{-1}(x_2) = -I_2^{-1}(x_2) + I_1^{-1}(x_1{=}0) + I_1^{-1}(x_2) \quad ,$$

where we have introduced I as a shorthand for IMSE. This suggests the following additive rule for $-\mathrm{IMSE}^{-1}$:

$$-I^{-1}(x_1,x_2) = -I_1^{-1}(x_1) - I_1^{-1}(x_2) - I_{1,2}^{-1}(|x_2 - x_1|) \quad ,$$

that is, that the total "energy" is the sum of two single-particle zero-point energy terms and a two-particle interaction term that depends only on the distance between points.

We have had success with using this approximate model for prediction of simple designs, such as the optimal design for the same model function but with three design points. Further work is needed to fully elucidate the benefits and limitations of this approach, although it could be used to speed the search for I_Z-optimal, or possibly other, designs.

We have also used the two-particle and zero-point potentials as inputs to a two-dimensional Newtonian particle simulator developed previously by one of us (A.H.), using up to a few tens of points. We found that repulsive potentials that give forces that grow with inter-particle separation from zero at close range, combined with a single-particle, origin-directed attractive force, gave frequent twin-point equilibrium final configurations upon annealing. Occasionally the system gave a twin-point ground-state configuration.

It thus appears that twin-point configurations may be common in systems with competing inter-particle repulsive and single-particle centering effects.

When designs with twin points are used in practice, the design points of the twin should be spaced by some ε sufficiently large that the numerical simulations and subsequent fitting procedures, which involve matrix inversions, are not numerically compromised.

A final remark is in order. It is, of course, always necessary to have sufficient computational precision to perform the required steps of a design search and the subsequent metamodel fit. Optimal designs with twin points, however, introduce a new notion. Such designs depends *fundamentally* on the precision to which the computations are performed, and thus on the information capacity of the computation. In such situations, the meaning of "optimal" has a new meaning because of a fundamental coupling between information and optimality.

5 EXAMPLE FROM STATISTICS LITERATURE

We now return to the main theme of the paper, viz., examples and comparisons of metamodel generation.

Our first example is taken from p. 17 of Welch et al. [1], where the following function of the two variables $x^{(1)}$ and $x^{(2)}$ was introduced:

$$Y(\tilde{x}^{(1)},\tilde{x}^{(2)}) = [30 + \tilde{x}^{(1)}\sin(\tilde{x}^{(1)})]\,(4 + e^{-\tilde{x}^{(2)}})$$
$$0 \le \tilde{x}^{(1)},\tilde{x}^{(2)} \le 5. \quad (1)$$

The function is plotted in Figure 4, after a simple transformation of variables to the region $[x^{(1)}, x^{(2)}]=[-1,1]^2$, which we use throughout this paper.

First we chose to use a simple neutral set of points for estimation of the θ's, rather than the points of the LHC design, since the latter would bias the subsequent analysis. We used the 3x3 mesh of points $\{-2/3,0,2/3\}^2$ and determined the maximum likelihood θ's to be $[\theta_1, \theta_2]=[0.506,0.138]$, following the procedure outlined in Welch et al. [1]. The I_Z-optimal design found subsequently with I-OPT is shown in Fig. 5b. The ratio of

IMSE's of the LHC to the I_Z-optimal design was 2.35; σ_Z^2 was 717.

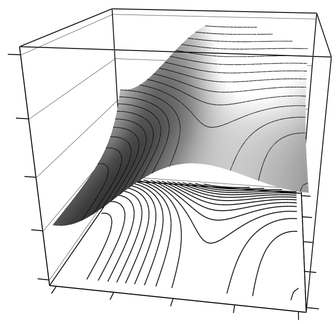

Figure 4: 3D plot, with contours, of the function in Eq. 1 transformed to the domain $[x^{(1)}, x^{(2)}]=[-1,1]^2$

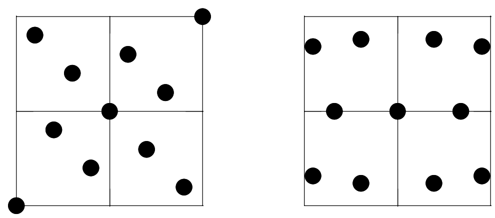

Figure 5: (a) left, LHC and (b) right, I_Z-optimal design used in the example drawn from Welch et al. The $x^{(1)}$ axes are horizontal.

The function was evaluated at each set of 11 points, BLUP fits performed using I-OPT [2], and the fits tested at 441 points on a 21x21 grid to calculate the empirical mean squared errors (EMSE's) of the fits, as in Welch et al. [1], with the slight modification that a weighted average was used in which the 72 points on the side boundaries were weighted by ½ and the four corner points were weighted by ¼, in order to obtain a slightly better estimate of the average error over the region. The eleven LHC design points were all in the set of samples for the EMSE, whereas the I_Z-optimal design had only a single point in the 21x21 mesh, viz., the origin, so the relative performance of the LHC design was biased high.

Nonetheless, the average variance of the BLUP fit based on the I_Z-optimal design was 1.85 times lower than the fit based on the LHC design. This is in good agreement with the expected ratio of 2.35 based on the IMSE's found using I-OPT.

6 MICROSYSTEMS-DESIGN EXAMPLE

We performed the same type of analysis on the excellent fourth-degree bivariate-polynomial response function, Fig. 6, found for a micro-flowmeter presented in our earlier work [2]. The analysis yielded the following results: $[\theta_1, \theta_2]=[0.108,0.478]$; the I_Z-optimal design was found, see Fig. 7; the ERMSE of the new fits were 9.7 mK and 14.5 mK for the I_Z-optimal and LHC designs, respectively (over a range of zero to 150 mK), representing significant improvements over the ERMSE's reported previously (56 to 76 mK for all designs [2]) for BLUP fits to the same model function but with different θ's; the ratio of expected IMSE's was 2.41, and the empirical ratio of IMSE's was 2.24.

We conclude that I_Z-optimal designs and BLUP fitting are especially attractive for metamodel generation.

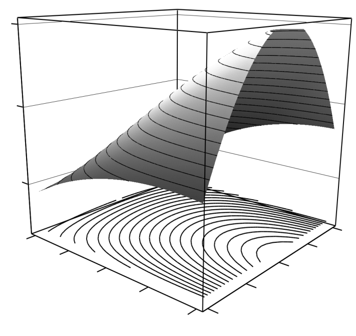

Figure 6: The response function for the micro-flowmeter example

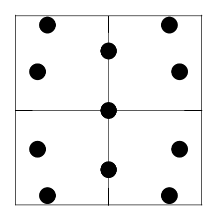

Figure 7: The I_Z-optimal design for the micro-flowmeter example. The $x^{(1)}$ axis is horizontal.

7 REFERENCES

[1] W.J. Welch, R.J. Buck, J. Sacks, H.P. Wynn, and T.J. Mitchell, M.D. Morris, "Screening, Predicting, and Computer Experiments," Technometrics, 34, 15-25, 1992.

[2] S.B. Crary, P. Cousseau, D. Armstrong, D.M. Woodcock, E.H. Mok, O. Dubochet, P. Lerch, and P. Renaud, "Optimal Design of Computer Experiments for Metamodel Generation Using I-OPT™," Computer Modeling in Engineering and Sciences, 1, 1-14, 2000.

[3] J. Sacks, S.B. Schiller, W.J. Welch, "Design of Computer Experiments," Technometrics, 31, 41-47, 1989.

[4] J.R. Koehler and A.B. Owen, "Computer Experiments," in S. Ghosh and C.R. Rao, editors, Handbook of Statistics, 13, 261-308, Elsevier Science, New York, 1996.

[5] D.M. Woodcock and S.B. Crary, "Navigating Smooth IMSE Landscapes Using High Precision Floating Point," (oral only) Joint Statistical Meetings, Indianapolis, IN, 13-17 Aug 2000.