

# A Hybrid One-Then-Two Stage Algorithm for Computationally Expensive Electromagnetic Design Optimization

G. I. Hawe (\*\*\*), J. K. Sykulski (\*\*)

(\*)Vector Fields Ltd., Oxford, U. K. OX5 1JE

(\*\*)School of Electronics and Computer Science,  
University of Southampton,  
Southampton, SO17 1BJ

**Abstract.** A novel kriging-assisted algorithm is proposed for computationally expensive single-objective optimization. The principle behind the algorithm is to use information about objective function space at the earliest possible opportunity. After constructing a very small experimental design, a one-stage optimization algorithm is used to select further points to evaluate in design variable space. These points are then used in lieu of a traditional space-filling experimental design to construct the initial kriging model for a normal two-stage optimization algorithm.

**Key words:** Optimization, Kriging, Design and Analysis of Computer Experiments (DACE).

## I. INTRODUCTION

In [1], surrogate-model assisted optimization algorithms are categorized into “two-stage” and “one-stage” varieties. Two-stage algorithms use an experimental design to construct an initial surrogate model (the first stage), and then use this model to determine where to sample next (the second stage), e.g. [2]. This approach has several drawbacks: experimental designs ignore information about objective function space, as they are constructed only to be space-filling in the design variable space, and this may be viewed as being wasteful; also, the surrogate model constructed from the points sampled may not be very accurate. One-stage algorithms do not fit a surrogate model to the observed data. Instead, they choose where to sample next by making a hypothesis about the location of the global minimum, and determining the credibility of surrogate models which would pass through it and the sampled points, e.g. [3]. The point chosen to be sampled is the hypothesised point which has the most credible surrogate model passing through it. This approach has the drawback that the computational cost of locating which point to evaluate next increases dramatically as the number of sampled points increases. In this paper, a hybrid one-then-two stage algorithm, which uses a one-stage algorithm to initialize a two-stage algorithm, is proposed for use with kriging surrogate models [4].

## II. ONE-THEN-TWO STAGE ALGORITHM

The proposed algorithm for locating the global minimum in  $n$ -dimensional design variable space  $D \subset \mathfrak{R}^n$  consists of three steps: initialization, one-stage experimental design, and two-stage optimization search.

### A. Initialization

The only purpose of the initialization step is to sample enough points to allow a non-trivial kriging model to be constructed (i.e. a model which is not a hyper-plane in  $\mathfrak{R}^n$ ). The space-filling Hammersley Sequence experimental design [5], of size  $2n$ , is used to select the points. The experimental design size of  $2n$  is much smaller than is normally used ( $10n$  is suggested in [2] for example), as the philosophy of this algorithm is to use information about objective function space to search for the minimum at the earliest possible opportunity.

### B. One-stage experimental design

Information about objective function space has now been obtained through sampling  $2n$  points, and the aim in this second step is to use this information to strategically choose where to sample next. Let the minimum objective function value of the  $2n$  sampled points be  $f_{\min}$ , and let the maximum be  $f_{\max}$ . A design vector  $\mathbf{x}^*$  is then hypothesized to exist in design variable space which has an objective function value  $f^* = f_{\min} - q(f_{\max} - f_{\min}) < f_{\min}$ , where  $0 < q \leq 1$ . Initially  $q$  is set to  $q = 1$ . Then, the likelihood of the  $2n$  sampled points conditional upon a kriging surface passing through  $(\mathbf{x}^*, f^*)$  is [1]:

$$\frac{1}{(2\pi)^{\frac{n}{2}} (\sigma^2)^{\frac{n}{2}} |\mathbf{C}|^{\frac{1}{2}}} \exp \left[ \frac{-(\mathbf{y} - \mathbf{m})\mathbf{C}^{-1}(\mathbf{y} - \mathbf{m})}{2\sigma^2} \right] \quad (1)$$

where:

$$\mathbf{m} = \mathbf{1}\mu + \mathbf{r}(f^* - \mu) \quad (2)$$

$$\mathbf{C} = \mathbf{R} - \mathbf{r}\mathbf{r}^T \quad (3)$$

are the conditional mean and correlation matrices. Here,  $\mathbf{R}$  is the correlation matrix,  $\mu$  is the mean and  $\sigma^2$  is the variance predicted by the kriging model. The next point to be evaluated is the  $x^*$  which maximizes (1). This value of  $x^*$  is the one which yields the most credible kriging surface which interpolates it (with objective function value  $f^*$ ) and the sampled points. This process is repeated for  $8n$  iterations until  $10n$  points in total have been evaluated. The value of  $q$  is cycled through using  $q = |\sin((i\pi/2n) + 0.01)|$  at the  $i^{\text{th}}$  iteration of this stage.

### C. Two-stage optimization search

The  $10n$  points sampled so far are now used to construct a kriging model. (In this respect, the previous stage may be viewed as the experimental design stage). Utility functions, which balance the values predicted by a kriging model with the uncertainty in the model, are now used to select which points to evaluate next. The generalized expected improvement function [6]:

$$GE[I^g(x)] = s^g \sum_{k=0}^g (-1)^k \left( \frac{g!}{k!(g-k)!} \right) u^{g-k} T_k \quad (4)$$

where

$$T_k = -\phi(u) u^{k-1} + (k-1) T_{k-2} \quad (5)$$

with

$$T_0 = \Phi(u) \quad (6)$$

$$T_1 = -\phi(u) \quad (7)$$

where  $\Phi$  is the standard normal distribution function and  $\phi$  is the standard normal density function, is used first, as it places most emphasis on regions of high uncertainty, which is what is desired at the beginning of an optimization search. The level of emphasis placed on searching uncertain regions is determined by the value of the integer  $g$ , with higher values of  $g$  corresponding to more emphasis. A cooling strategy [7] is employed, using high values of  $g$  to begin with, and then progressively smaller values. When  $g = 1$ , the utility function is equivalent to the expected improvement utility function, as used in the EGO algorithm [2]. Upon reaching  $g = 1$ , we wish to place more emphasis on searching around the current minimum, and to do this the weighted expected improvement utility function [8] is used:

$$WE(I(x)) = w(f_{\min} - \hat{f}(x)) \Phi\left(\frac{f_{\min} - \hat{f}(x)}{s(x)}\right) + (1-w)s(x) \phi\left(\frac{f_{\min} - \hat{f}(x)}{s(x)}\right) \quad (8)$$

where  $\hat{f}(x)$  is the objective function value as predicted by the kriging surrogate model, and  $s(x)$  is the root mean squared error in this prediction. The first term in this function places emphasis on searching around the current minimum, whilst the second term places emphasis on searching regions of design variable space with high uncertainty in their values. The terms are balanced using the weighting parameter  $w$ . Using  $w=0.5$  in (8) is equivalent to using  $g = 1$  in (4), so we begin using (8) with a value of  $w = 0.6$  to select points. Then different values of  $w$  are cycled through, using  $w = 0.5 + 0.5 * |\cos(i\pi/2n)|$  at the  $i^{\text{th}}$  use of (8). The algorithm then proceeds either for a fixed number of iterations, or until an adequate solution has been found.

## III. RESULTS AND CONCLUSIONS

Results on test functions have been encouraging. A full analysis of the performance of the proposed algorithm and its application to electromagnetic device optimization will be discussed in the full paper.

## REFERENCES

- [1] D. R. Jones, "A Taxonomy of Global Optimization Methods Based on Response Surfaces", *Journal of Global Optimization*, vol. 21, 2001, pp. 345-383.
- [2] D. R. Jones, M. Schonlau, W. J. Welch "Efficient Global Optimization of Expensive Black-Box Functions", *Journal of Global Optimization*, vol. 21, 2001, pp. 455-492.
- [3] H. M. Gutmann, "A Radial Basis Function Method for Global Optimization", *Journal of Global Optimization*, vol. 19, no. 3, 2001, pp. 201-227.
- [4] L. Lebensztajn, C. A. R. Marretto, M. C. Costa, J-L. Coulomb, "Kriging: A Useful Tool for Electromagnetic Devices Optimization", *IEEE Transactions on Magnetics*, vol. 40, no. 2, 2004, pp 1196-1199.
- [5] J. R. Kalagnanam, U. M. Diwekar "An Efficient Sampling Technique for Off-line Quality Control", *Technometrics*, vol. 39, no. 3, 1997, pp. 308-319.
- [6] M. Schonlau, W. J. Welch, D. R. Jones, "Global versus Local Search in Constrained Optimization of Computer Models", *IMS Lecture Notes*, vol. 34, 1998 pp. 11-25.
- [7] M. J. Sasena, P. Papalambros, P. Goovaerts, "Exploration of Metamodelling Sampling Criteria for Constrained Global Optimization", *Engineering Optimization*, vol. 34, 2002, pp.263-278.
- [8] A. Sobester, S. J. Leary, A. J. Keane, "On the Design of Optimization Strategies Based on Global Response Surface Approximation Models", *Journal of Global Optimization*, vol. 33, 2005, pp. 31-59.