# A Fast Global Optimizer Based on Improved CS-RBF and Stochastic Optimal Algorithm

S. L. Ho[1], Shiyou Yang[2], Guangzheng Ni[2], and H. C. Wong[3]

[1]Department of Electrical Engineering, the Hong Kong Polytechnic University, Hong Kong
[2]College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China
[3]Industrial Center, the Hong Kong Polytechnic University, Hong Kong

**An improved compactly supported radial basis function is proposed as a response surface model in the study of computationally heavy design problems. A new interpolation formula is introduced to enhance the interpolation accuracy on boundary derivatives and the proposed response surface model is then combined with stochastic algorithms in the design of a fast global optimizer. Numerical results are reported to demonstrate the generality and the robustness of the proposed works.**

***Index Terms*—Compact support, global optimization, inverse problem, radial basis function, response surface model.**

## I. INTRODUCTION

DEVELOPMENT and utilization of stochastic or heuristic optimal algorithms in the design optimization of electromagnetic (EM) devices have advanced rapidly in the last three decades, mainly because of the need to address the multimodal nature of objective functions. Compared with their deterministic counterparts, the stochastic algorithms are computationally expensive since thousands of function evaluations are generally required before one can find the optimal solution. Moreover, design problems of EM devices usually involve sophisticated and tedious computer simulations. For example, it is very common to use finite-element analysis (FEA), repetitively, so as to evaluate the performance of an EM device. Consequently, one faces the dilemma of whether to give priority to computation burden or accuracy when selecting the most suitable optimal algorithm in solving computationally heavy design problems with multimodal objective functions.

To reduce the number of function evaluations involving computationally heavy procedures without sacrificing the computation accuracy, a lot of efforts are devoted to the use of response surface models or methodologies (RSM) and their applications in the development of rapid global optimizers [1], [2]. By reconstructing computational heavy design problems from the data on a set of sample points using RSM, one can solve the reconstructed problem efficiently. Obviously, the final solution obtained depends strongly on the RSM being used. To date, the most popular RSMs used in the study of EMs are those based on globally supported radial basis functions (RBFs) because of their interpolating power in dealing with both grid and scattered data. However, global RBFs have the inherent drawbacks of having the need to manipulate the full interpolation matrices. Due to difficulties in processing full matrices, the sample points of available RSMs could not exceed an upper limit [3]. Moreover, when the spacing between sample points is very small, the interpolation matrix becomes very ill-conditioned as well, giving rise to numerical singularity and low numerical accuracy. To circumvent the drawbacks of globally supported RBFs, the compactly supported RBF (CS-RBF) is improved and used to design an RSM for constructing a computationally heavy objective function. To enhance the interpolation accuracy for (partial) derivatives on boundaries of the solution domain, a new interpolating formula is introduced. Finally, the proposed CS-RBF-based RSM is combined with stochastic algorithms in the design of a fast global optimizer.

## II. INTERPOLATION USING CS-RBF

### A. Compactly Supported RBF

In essence, the pioneering work cocerning CS-RBFs is attributed to Wu [4] and Wendland [5] in the mid 1990s. The functions of CS-RBFs are strictly positive definite in $R^n$ for all values of $n$ which are less than or equal to some fixed value $d$ and can be constructed with any desired degree of smoothness $2k$. The family of CS-RBFs, constructed by Wendland, is used in the proposed study because, for a specific space dimension $d$, these functions posses the lowest positive degree among all of the piecewise polynomials of CS-RBFs which are positive definite on $R^d$ having a given order of smoothness. Generally, a CS-RBF is expressed as

$$\phi_{d,k}(r) \doteq (1-r)_+^n p(r) \quad (k \geq 0) \tag{1}$$

where $p(r)$ is a prescribed polynomial, $r = \|X\|$ is the Euclidean norm, and $\doteq$ denotes equality up to a constant factor. For example

$$(1-r)_+^n = \begin{cases} (1-r)^n, & (0 \leq r < 1) \\ 0, & (r \geq 1) \end{cases}. \tag{2}$$

It should be pointed out that the CS-RBF function, which is positive definite in $R^d$, is also positive definite in $R^k$ $(k < d)$. The explicit formulae for Wendland's CS-RBFs, $\phi_{l,k}(r)$, which possesses $2k$ smooth continuous derivatives ($l = [d/2 + k + 1]$, $d$, is the size of the space dimensions) for $k = 0, 1, 2, 3$, are given in Table I for the convenience of fellow researchers.

### B. Interpolation Using CS-RBF

The interpolation of a function using CS-RBFs can be implemented simply due to their positively definite property. Mathe-

Digital Object Identifier 10.1109/TMAG.2006.871429

TABLE I
EXPLICIT FORMULAE OF SOME WENDLAND'S CS-RBFs.

| | |
|---|---|
| $\phi_{l,0}(r) \doteq (1-r)_+^l$ | $C^0 \bigcap PD_d$ |
| $\phi_{l,1}(r) \doteq (1-r)_+^{l+1}[(l+1)r+1]$ | $C^2 \bigcap PD_d$ |
| $\phi_{l,2}(r) \doteq (1-r)_+^{l+2}[(l^2+4l+3)r^2+(3l+6)r+3]$ | $C^4 \bigcap PD_d$ |
| $\phi_{l,3}(r) \doteq (1-r)_+^{l+3}[(l^3+9l^2+23l+15)r^3$ $+(6l^2+36l+45)r^2+(1l+45)r+15]$ | $C^6 \bigcap PD_d$ |

matically, the reconstruction of a function $f(X) : R^d \rightarrow R$ on the basis of its values $f_i$ at a set of sample points $X_i \in R^d (i = 1, 2, \ldots, N)$ in terms of some CS-RBF $\phi$ is

$$f(X) = \sum_{j=1}^{N} c_j \phi(\|X - X_j\|). \tag{3}$$

As observed in [6], although the interpolation of a function using (3) performs very well in the inner region of the parameter space, it will give rise to significant errors in the derivatives on the boundary. To solve this problem, the information of the derivatives on the boundary is incorporated, and the interpolation becomes

$$f(X) = \sum_{j=1}^{N} c_j \phi(\|X - X_j\|) + \sum_{k=1}^{N_B} d_k \phi'(\|X - X_j\|) \tag{4}$$

where $N$ is the number of the total sample points, $N_B$ is the number of the boundary sample points used only for derivative fitting, $\phi'(\cdot)$ is the first-order derivatives of $\phi$ with respect to the radial variable $r$.

The coefficients $c_j$ and $d_k$ are determined from the following matrix equation:

$$\begin{bmatrix} c_j \\ d_k \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}^{-1} \begin{bmatrix} f_j \\ f'_k \end{bmatrix} \tag{5}$$

where $f'_k$ is the first-order derivative of function $f$ with respect to the radial variable $r$ at boundary point $k$, and

$$\begin{aligned} (\phi_{11})_{ij} &= \phi(\|X_i - X_j\|) \\ &\quad (i, j = 1, 2, \ldots, N) \\ (\phi_{12})_{ij} &= \phi'(\|X_i - X_j\|) \\ &\quad (i = 1, 2, \ldots, N; j = 1, 2, \ldots, N_B) \\ (\phi_{21})_{ij} &= \phi'(\|X_i - X_j\|) \\ &\quad (j = 1, 2, \ldots, N; i = 1, 2, \ldots, N_B) \\ (\phi_{22})_{ij} &= \phi''(\|X_i - X_j\|) \\ &\quad (i, j = 1, 2, \ldots, N_B) \end{aligned} \tag{6}$$

where $\phi''(\cdot)$ is the second-order derivatives of $\phi(\cdot)$ with respect to the radial variable $r$.

It should be noted that the available information of (partial) derivatives is related to the coordinate variables rather than to the radial one. Therefore, one should derive the radial derivatives as required by (5) with respect to their coordinate variables. For example, if the partial derivative of function $f$ with the $i$th coordinate variable $x_i$ of Cartesian system $f'_{x_i}$ is given, then the derivative of function $f$ with respect to the radial variable $r$, $f'$, can be determined from

$$f' = \frac{r}{x_i} f'_{x_i}. \tag{7}$$

For high dimensional problems, the partial derivatives may include numerical errors if they are determined using a numerical approach. Therefore, the computed derivatives of a function with respect to the radial variable on the same sample point may not be identical if partial derivative with a different coordinate direction is used in (7). In such a case, the averaged value of the derivatives obtained from the derivative information of different coordinate directions is used. Also, because of the compactly supported nature of CS-RBFs, the coefficient matrix $\phi$ in (5) is sparse. Hence, the addition or removal of some sample points causes only a local change in the interpolation results.

In many high dimensional problems, the partial derivatives of the objective function for different coordinates may vary significantly. In such cases, the interpolation performances of the CS-RBFs as formulated in (4) are often degraded. To eliminate this problem, an improved CS-RBF with the introduction of a scale parameter is proposed. The high dimensional CS-RBFs being proposed are defined as

$$\phi_j(r) = \phi\left(\sqrt{\sum_{i=1}^{d}((x_i - (x_i)_j)/D_i)^2}\right) \tag{8}$$

where $D_i$ is a scale parameter which is inversely proportional to the $i$th partial derivative of the function at the point $X_j$.

### C. Merit of the Proposed CS-RBF Interpolation

To show the merit and the generality of the proposed CS-RBF in function interpolations, a two-dimensional (2-D) mathematical function as defined below is deliberately designed with the partial derivative of the $x$-coordinate variable being much smaller than that of the $y$-coordinate variable. Moreover

$$f(x, y) = \sin(x)\sin(20y)(x \in [0, \pi], y \in [0, \pi/20]). \tag{9}$$

For comparative purpose, this mathematical function is reconstructed using the proposed CS-RBF with and without the scale parameter $D_i$ as defined in (8). The CS-RBF used in this case study is a simple one with $C^0$ smoothness and with $d$ being equal to or smaller than 3, and is defined as

$$\phi(r) = (1 - r)_+^2. \tag{10}$$

For elucidative simplicity, the derivative term as defined on the right side of (4) is deliberately excluded. In the numerical implementations, the scale parameters for $D_x$ and $D_y$ are set, respectively, to 1 and 0.05. Twelve equidistant points in each coordinate direction are used as sample points. To evaluate the performances of different interpolation schemes, the following metrics are used:

$$\text{Metric-A}(MA) = \sum_{i=1}^{N} (f_i^{\text{ex}} - f_i^{\text{com}})^2 \tag{11}$$

$$\text{Metric-B}(MB) = \sum_{i=1}^{N} |f_i^{\text{ex}} - f_i^{\text{com}}| \tag{12}$$

where $f_i^{\text{ex}}$ and $f_i^{\text{com}}$ are the exact and computed values, respectively, of the test function at sample point $i$ using an interpolation scheme; $N$ is the number of observation points.

In the evaluation, 40 equidistance sample points along each coordinate direction are used as observing points to calculate

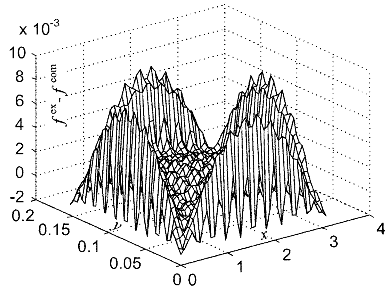


Fig. 1. Error distribution of the interpolated mathematical function using the proposed CS-RBF with scale parameters.

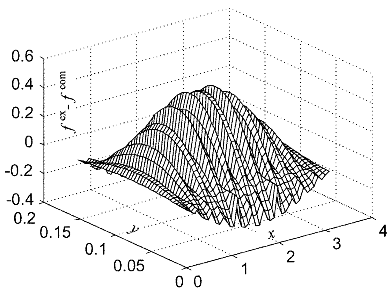


Fig. 2. Error distribution of the interpolated mathematical function using the proposed CS-RBF without scale parameters.

these matrices. The error distribution of the interpolated functions using the proposed CS-RBF with and without the scale parameter are shown, respectively, in Figs. 1 and 2, and the corresponding performance comparison results are given in Table II.

From this case study, it is obvious that, even if a very simple CS-RBF is used, the interpolation errors of the proposed CS-RBF is very small, whereas those of the same one without a scale parameter are too large for the interpolated function to be used as a response surface when the function has significant differences in the derivatives of different coordinate directions. In other words, the example is a good demonstration of the robustness and generality of the improved CS-RBF.

## III. Efficient Optimizer Based on Improved CS-RBF

### A. RSM Based on Multistep CS-RBF Interpolations

Generally, the sample points of an objective function should be distributed irregularly in the parameter space such that the point densities are higher in regions where the local optima are likely to exist. Thus, every CS-RBF should have the ability to adjust its support according to the point density around it. Correspondingly, the CS-RBFs are of the form $\phi(\cdot/\alpha)$ for $\alpha > 0$ ($\alpha$ will be called the resolution parameter hereafter). For this purpose, the multistep method as proposed in [7] is used in this paper. Thus, one shall first decompose the set of sample points $X$ into a nested sequence as follows:

$$X^1 \subset X^2 \subset \cdots X^{M-1} \subset X^M = X \tag{13}$$

with the subset $X^k$ of $X$ being given as

$$X^k = \left\{ X_1^{(k)}, X_2^{(k)}, \ldots, X_{N_k}^{(k)} \right\} \quad (1 \le k \le M) \tag{14}$$

and the interpolation problem is then decomposed into $M$ steps as described below.

TABLE II
Performance Comparison of Different Interpolation Schemes

| Deviances | Proposed CS-RBF | Proposed CS-RBF without scale |
|---|---|---|
| MA | $1.33\times10^{-2}$ | 54.43 |
| MB | 3.65 | 211.43 |

Starting with $k = 1$, one will match the error function at the $k$th step as

$$f - (s^1 + s^2 + \cdots s^{k-1}) \tag{15}$$

on $X^k$ by computing the coefficients of the $k$th interpolant

$$s^k(Y) = \sum_{j=1}^{N_k} c_j^k H\left( \left\| Y - X_j^{(k)} \right\| \Big/ \alpha_k \right) + \sum_{j=1}^{N_k^B} d_j^k H'\left( \left\| Y - X_j^{(k)} \right\| \Big/ \alpha_k \right) \tag{16}$$

after the value of $\alpha_k$ of the basis function has been chosen.

It follows naturally that:

$$f|_X = (s^1 + s^2 + \cdots + s^M)|_X. \tag{17}$$

The details of the numerical implementations of this multistep method are referred to [7]. This approach allows one to choose a relatively large scale at the lowest level to capture the overall behavior of the function, and by decreasing it during the process of the procedure, finer and finer details of the function are obtained step by step, thereby providing a hierarchical construction procedure with reasonable computing time. The resolution parameter of the proposed CS-RBF at a substep is determined in such a way that the influence of a CS-RBF covers at least 20 sample points. Also, a cluster algorithm is used to decompose the sample points into a nested sequence of subsets [7].

### B. Efficient Optimizer Based on the Improved CS-RBF

For the efficient solution of computationally heavy design problems, an iterative procedure using the improved CS-RBF based RSM and stochastic algorithms is proposed by the following steps.

- Step 1) Generate a set of sample points; compute the objective function/derivative values using computationally heavy algorithms at these sample/boundary points. Decompose the sample points into a nested sequence of subsets.
- Step 2) Reconstruct the optimal problem using the proposed CS-RBF, and solve it using a stochastic method, report the searched local/global optimal solutions.
- Step 3) Solve directly the original optimal problem using a deterministic method starting from the searched local optimal solutions to find the final ones.

## IV. Numerical Example and Conclusions

To validate the proposed algorithm in solving computationally heavy design problems, the proposed CS-RBF is used to solve the Team Workshop problem 22 of a superconducting

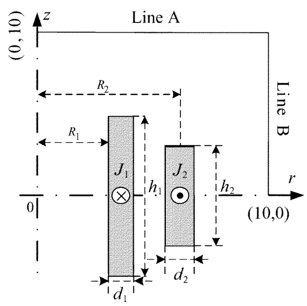


Fig. 3. Schematic diagram of the SMES.

TABLE III
FINAL OPTIMAL RESULTS SEARCHED BY USING THE PROPOSED METHOD

| Results | $R_1$(m) | $R_2$(m) | $h_1/2$(m) | $h_2/2$(m) | $d_1$(m) | $d_2$(m) |
|---|---|---|---|---|---|---|
| Proposed | 1.5706 | 2.1021 | 0.7852 | 1.4208 | 0.6002 | 0.2582 |
| By IGTE | 1.5703 | 2.0999 | 0.7846 | 1.4184 | 0.5943 | 0.2562 |

TABLE IV
FINAL OPTIMAL RESULTS SEARCHED BY USING THE PROPOSED METHOD

| Results | $J_1$(MA/m$^2$) | $J_2$(MA/m$^2$) | $B^2_{stray}$ (T$^2$) | *Energy*(MJ) | $f_{obj}$ |
|---|---|---|---|---|---|
| Proposed | 17.2469 | -12.9653 | $2.2005\times10^{-10}$ | 180.2020 | $6.6270\times10^{-3}$ |
| By IGTE | 17.3367 | -12.5738 | $2.1913\times10^{-10}$ | 179.9924 | $5.5203\times10^{-3}$ |

magnetic energy storage (SMES) configuration with eight free parameters, as shown in Fig. 3. The problem is formulated as

$$\begin{aligned}\text{minimize} \quad & f = \frac{B^2_{\text{stray}}}{B^2_{\text{norm}}} + \frac{|\text{Energy} - E_{\text{ref}}|}{E_{\text{ref}}} \\ \text{subject to} \quad & J_i \le (-6.4|(B_{\max})_i| + 56)(\text{A/mm}_2) \\ & \qquad (i = 1, 2) \qquad (18)\end{aligned}$$

where Energy is the stored energy in the SMES device, $E_{\text{ref}} = 180\ \text{MJ}$, $B_{\text{norm}} = 2 \times 10^{-4}\ \text{T}$, $B^2_{\text{stray}}$ is a measure of the stray fields which is evaluated along 22 equidistant points of line $A$ and line $B$ of Fig. 3 as follows:

$$B^2_{\text{stray}} = \sum_{i=1}^{22} (B_{\text{stray}})^2_i \Big/ 22. \qquad (19)$$

In the numerical implementation, 2000 sample points are first generated using a simulated algorithm on the original problem in which the objective function is obtained using finite-element simulations. These sampling points, together with 80 additional ones which are uniformly distributed on the boundaries for derivative fitting, and their function values are then used to reconstruct the optimal problem using a $C^2 \cap PD_9$ CS-RBF $\phi(r) = (1-r)^7_+(7r+1)$. The problem is then solved efficiently using a tabu search method to find the close solutions of the "optimal ones." Finally, the simplex method is run directly on the original problem to find the final solution. It is found that 35 iterations are required before the simplex algorithm converges, and the final solutions are reported in Tables III and IV. The distributions of the computed magnetic flux density along line $A$ and line $B$ is depicted in Fig. 4. It should be pointed out that the dimensional sizes of different parameter spaces are all scaled to $[0, 1]$ in this case study in our numerical implementation. From these numerical results, it is clear that the proposed fast optimal algorithm can virtually reach the same optimal solutions as those obtained by the Institut für Grundlagen und Theorie der Elektrotechnik, Graz University of Technology (IGTE), which validates the feasibility of the proposed method in solving computationally expensive engineering design problems.

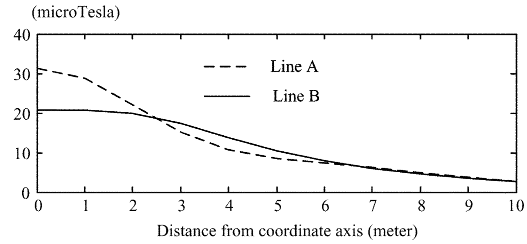


Fig. 4. Distribution of the magnetic flux density along line A and line B.

In summary, the proposed fast optimal algorithm is: 1) more computationally efficient and robust when compared with globally supported RBF-based ones because of its compact support nature and its ability to eliminate the ill-conditioned problems of the interpolation matrices; 2) can obtain the "best" tradeoff between accuracy and efficiency of an interpolation when compared to available CS-RBF-based ones because of its multilevel interpolation capability. Hence, the proposed method is not only very promising for rapid and robust global optimizations for EM design problems; it is also ideal for general engineering problems in which the objective/constraint functions must be determined by using computationally expensive algorithms such as three-dimensional finite-element analysis.

## ACKNOWLEDGMENT

The authors acknowledge the funding support of the Research Grant Council of the Hong Kong Special Administrative Region, China, under Project PolyU 5242/04E.

## REFERENCES

[1] D. N. Dyck and D. A. Lower, "Response surface modeling of magnetic device performance using function value and gradient," *Int. J. Appl. Electromagn. Mech.*, vol. 9, pp. 241–248, 1998.
[2] K. Rashid, M. Farina, J. A. Ramirez, J. K. Sykulski, and E. M. Freeman, "A comparison of two generalized response surface methods for optimization in electromagnetics," *COMPEL*, vol. 20, pp. 740–752, 2001.
[3] P. Alotto, A. Bertoni, and M. Nervi, "An efficient hybrid algorithm for the optimization of problems with several local minima," *Int. J. Numer. Meth. Eng.*, vol. 50, pp. 847–868, 2001.
[4] Z. Wu, "Multivariate compactly supported positive definite radial functions," *Adv. Comput. Math.*, vol. 4, pp. 283–292, 1995.
[5] H. Wendland, "Piecewise polynominal, positive definite and compactly supported radial functions of minimal degree," *Adv. Comput. Math.*, vol. 4, pp. 389–396, 1995.
[6] X. Zhang, K. Z. Song, and X. Liu, "Meshless methods based on collocation with radial basis functions," *Comput. Mech.*, vol. 26, pp. 333–343, 2000.
[7] M. S. Floater and A. Iske, "Multistep scattered data interpolation using compactly supported radial functions," *J. Comput. Appl. Math.*, vol. 73, pp. 65–78, 1996.

Manuscript received June 20, 2005; revised November 20, 2005 (e-mail: eeslho@polyu.edu.hk).