A Genetic Algorithm for Combinational Optimization Problems with Uncertainties

Kenta Hoshino and Hajime Igarashi, *Member, IEEE*

***Abstract*— This paper describes a genetic algorithm (GA) applied to combinational optimization problems in which the objective functions include uncertain constant parameters. In the present method, noises whose probabilistic distribution is assumed based on the problem environments are added to the parameters during the evaluations process. It is assumed that this allows us to make approximate evaluation of the expected values of the objective function under uncertainties. It is shown that the present method results in the robust solutions, which have higher expectation values than the ones obtained by the conventional GA, to the traveling salesman and knapsack problems with uncertainties.**

I. Introduction

IN the evolutional optimizations, which have been shown to be very effective for wide range of linear and non-linear optimization problems, uncertainties have to be often taken into consideration. In the design of industrial products, for example, it is important to evaluate degree of uncertainties in their performance, which result from time dependent changes in material properties, production errors, variations in material characteristics and so on. Moreover, there are various uncertainties, to which attentions are paid in optimization problems, for instance, relevant to noises introduced in measurement systems [1] and human-health risks [2].

According to [3], uncertainties are treated in the following ways in the evolutionary optimizations; (a) perturbations are added to the fitness values to consider noises in sensory measurements and system randomness, (b) perturbations are introduced in optimization variables to search for robust solutions which are insensitive to parameter changes, (c) uncertain terms are introduced in the objective function to consider unexpected errors due to approximation of objective functions for reduction of computational costs and (d) the objective function itself is changed to consider time-dependent, dynamical systems. This paper focuses on the robust optimization under uncertain environments, corresponding to the class (b).

To obtain robust optimal solutions which are insensitive to environmental changes, sensitivity of individuals has been evaluated using the Monte Carlo method [4-6], whose computational accuracy and efficiency have been improved using the Latin hypercube sampling [7]. It would be, however, ineffective to apply these explicit averaging approaches to optimization problems in which fitness evaluation carried out, for example, by computational mechanics and electromagnetism is expensive. Even if fitness evaluation is not very expensive, the total computational cost becomes inacceptable when the number of function evaluation is large, as experienced in high dimensional optimization problems. To overcome this difficulty, implicit averaging method for robust optimization has been introduced, in which perturbation noises are added to optimization variables to obtain their expected values assuming the population size is sufficiently large [8]. This method has been applied to optimization problems in which the fitness is evaluated with the aid of computational electromagnetism with improvement in elite selection [9].

In this paper, we will discuss the implicit sampling approach for combinational optimization problem containing probabilistic constants. An example of this problem can be found in investment activities; portfolio managements are performed to maximize the expected return and minimize the return risk. In this example, the prices of assets and amounts of purchase correspond to the probabilistic constants and optimization variables. Other examples can also be found optimization of transportation paths where transport costs depend on the traffic conditions, and scheduling problem in production systems in which probabilistic failures in producing machines are considered. In these problems, the expected values of the objective functions should be considered in contrast to usual optimizations. Those expected values would easily be evaluated in the evolutional optimization processes when the objective functions are linear and the probabilistic variables are independent. If it is not the case, however, it would be too expensive to evaluate them in optimization processes because the multiple integrations must be performed numerically.

In this paper, to treat the combinational optimization problems with uncertainty, an evolutional method based on the genetic algorithm (GA) will be introduced. In this method, noises are introduced to constants included in the objective function to express their stochastic nature. The present method differs from the robust optimizations [8, 9] in which

Manuscript received February 12, 2010.
K. Hoshino is with Graduate School of Information Science and Technology, Hokkaido University, Kita 14, Nishi 9, Kita-ku, Sapporo, 060-0814 Japan. phone: +81-11-706-6488; fax: +81-11-706-7670; e-mail: hoshino@em-si.eng.hokudai.ac.jp.
H. Igarashi is with Graduate School of Information Science and Technology, Hokkaido University, Japan, e-mail: igarashi@ssi.ist.hokudai.ac.jp.

noises are introduced not in the constants but in optimization variables, as will be mentioned in the next section. Although the probabilistic variables will be assumed to be independent in this paper, problems with dependent variables can also be treated with ease.

This paper will be organized as follows: the next section will describe the basic principle of GA for uncertain systems and formulation and algorithm of the present method, and the third section will provide numerical results for simple mathematical examples, traveling salesman and knapsack problems which contain uncertainties.

II. Formulation and Algorithm

A. Genetic Algorithm for Uncertain Systems

Before introduction of the present method, we begin with the conventional GA for uncertain systems because they are based on a common assumption; the expected value $\widetilde{F}$ of an objective function f including probabilistic variables over the population at a generation can approximately be computed from

$$\widetilde{F} = \frac{1}{n}\sum_{i}^{n} f_i, \tag{1}$$

in the GA processes, where n denotes the number of total number of fitness evaluations at that generation. Namely, it is assumed that $\widetilde{F}$ can be obtained, like the Monte Carlo simulations, by stochastically changing the probabilistic variables, which obey the prescribed distribution function, in the fitness evaluation of each individual. In contrast to the standard GA whose fitness is determined from the value of f, this kind of GA searches for the optimal solution on the basis of $\widetilde{F}$.

As an example of the GA for uncertain systems, let us consider here the robust GA [8, 9] which searches for the optimal or quasi-optimal solution whose performance is insensitive to stochastic changes in the design parameters. In this method, the probabilistic variables correspond to the design parameter $\boldsymbol{x}$ which is expressed in the form of genes in the GA optimization. The expected value $\widetilde{F}$ of f over the population belonging to a generation is obtained from

$$\begin{aligned}\widetilde{F} &= \frac{1}{n}\sum_{i=1}^{n} f(\boldsymbol{x}_i + \boldsymbol{\delta}_i) \\ &= \int p(\boldsymbol{x})d\boldsymbol{x}\left[\int f(\boldsymbol{x}+\boldsymbol{\delta})q(\boldsymbol{\delta})d\boldsymbol{\delta}\right] \\ &= \int F(\boldsymbol{x})p(\boldsymbol{x})\,d\boldsymbol{x},\end{aligned} \tag{2}$$

where $\boldsymbol{\delta}$ is stochastic perturbation to $\boldsymbol{x}$, p and q are the probabilistic distribution functions of $\boldsymbol{x}$ and $\boldsymbol{\delta}$, defined as

$$\begin{aligned}p(\mathbf{x}) &= p_1(x_1)p_2(x_2)\ldots p_n(x_n), \\ q(\mathbf{x}) &= q_1(x_1)q_2(x_2)\ldots q_n(x_n),\end{aligned} \tag{3}$$

and $F(\boldsymbol{x})$ denotes the expected value of $f(\boldsymbol{x})$. In the GA process, the individuals evolves to maximize (minimize) $\widetilde{F}$. Figure 1 shows the profiles of a one-dimensional objective function f and corresponding F. It can be seen that f with a broader peak has the higher peak in F in comparison with f with a narrow peak. This suggests that the robust GA tends to converge to a broad peak whose height is relatively insensitive to perturbations in $\boldsymbol{x}$.

B. Present method

The present method searches for the optimal solution to combinational problems with the objective function which contains probabilistic constants. In this paper, we especially focus on an objective function of the form

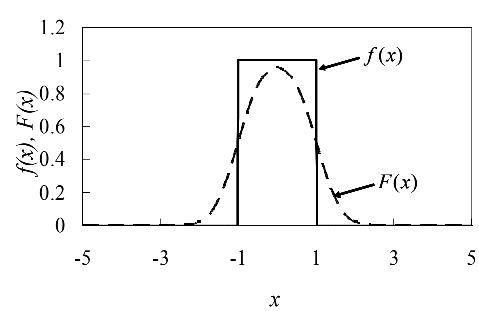
$$f(\boldsymbol{x},\boldsymbol{c}) = f\left(\sum_{j}^{m} c_j x_j\right), \tag{4}$$

where x_i and c_i, $i = 1,2,\ldots,m$, are the optimization variables, and coefficients which are assumed to be dependent probabilistic variables obeying given probabilistic distributions. The expected value of $f(\boldsymbol{x})$ is now given by

$$F(\boldsymbol{x}) = \int f\left(\sum_{j}^{m}(c_j + \delta_j)x_j\right)q(\boldsymbol{\delta})d\boldsymbol{\delta}. \tag{5}$$

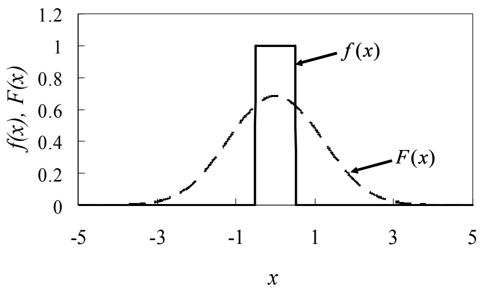
If f is a linear function, then (5) reduces to

$$F(\boldsymbol{x}) = \sum_{j}^{m} E(c_j)f(x_j) \tag{6}$$



(a) broad peak

Fig. 1 Example of profiles of f and F



(b) narrow peak
Fig. 1 continued

which can easily evaluated for each $\boldsymbol{x}$, otherwise its evaluation based on the numerical integration in m-dimensional space would be computationally expensive for GA and other optimization methods. In the present method, (5) is evaluated based on (1) where noises are introduced to c_j in the GA processes. Hence, the individuals are expected to converge to one of the peaks of $F(\boldsymbol{x})$.

In the present method, the elite selection method presented in [9] is employed to avoid accidental convergence to a solution with non-best value of F.

(i) At the first generation, t=0, the best individual is found with respect to the value $f(\boldsymbol{x},\boldsymbol{c}+\boldsymbol{\delta})$ and reserve it as an elite $\boldsymbol{x}_e$.

(ii) At generation t+1, the candidate $\boldsymbol{x}_{t+1}$ for new elite is found with respect to $f(\boldsymbol{x},\boldsymbol{c}+\boldsymbol{\delta})$.

(iii) Compare $f(\boldsymbol{x}_e,\boldsymbol{c}+\boldsymbol{\delta})$ with $f(\boldsymbol{x}_{t+1},\boldsymbol{c})$. If $f(\boldsymbol{x}_{t+1},\boldsymbol{c})$ is better than $f(\boldsymbol{x}_e,\boldsymbol{c}+\boldsymbol{\delta})$, then $\boldsymbol{x}_e$ is replaced by $\boldsymbol{x}_{t+1}$.

(iv) Go to the next generation and return to (ii).

III. Optimization Results

A. A Simple Example

To test the accuracy of the present method, the one-dimensional mathematical function

$$f(x,c)=cf_1(x)+(1-c)f_2(x) \tag{7}$$

is considered, where c obeys the normal distribution $N(\mu,\sigma^2)=(0.6,0.04)$. The functions f_1, f_2 are the Gauss functions given by

$$f_i(x)=\exp\left(-\frac{(x-x_i)^2}{\sigma_i{}^2}\right), \tag{8}$$

where $x_1=2.0$, $x_2=0.8$, $\sigma_1=\sigma_2=0.3$. The expected value F of f can be evaluated to be

$$F(x)=0.6f_1(x)+0.4f_2(x). \tag{9}$$

The profile of $F(x)$ is shown in Fig.2. For comparison, the present method and simple GA (SGA) are applied to (7) and (9), respectively. The GA parameters are summarized as follows: number of individuals in population N_{pop}=100, number of generations N_g=200, crossover probability P_c=0.15, mutation probability P_m=0.05, generation gap g=0.8. The trial numbers for each of the present method and SGA are 100.

The results are summarized in TABLE I, from which it can be seen that the mean values of the objective functions are almost same. Hence we can conclude that the present method gives sufficiently accurate solutions to this simple problem. Figure 3 shows the time evolution in the fitness of the elite solutions during the present method and SGA. In the former case, there are fluctuations in the elite fitness, which are due to the introduced noises.

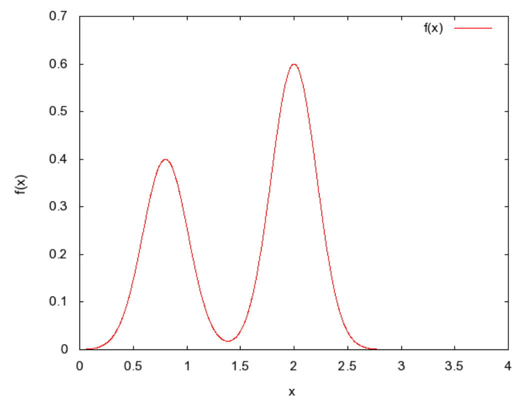


Fig.2 Profile of F defined in (9).

TABLE I Optimal solutions to one-dimensional problem

methods	F_{opt}	$\sigma_F{}^2$
SGA	2.000	1.156×10^{-11}
present	1.999	5.011×10^{-5}

F_{opt} denotes the optimized value of (9) for SGA and posterior expected value of (7) obtained by Monte Carlo method for the present method, respectively, averaged over 100 trials, and σ_F is the standard deviation of F.

B. Traveling Salesman Problem with Uncertainty

Let us consider the traveling salesman problem in which

$$f=\frac{1}{T},\quad T=\sum_{j}^{m}c_j x_j \tag{10}$$

is maximized, where c_j denotes the cost attributed to path j which connects two different cities, and x_j is the optimization variable which takes one of the two states (1, 0) corresponding to connection and disconnection, respectively. As usual, all the cities must be visited only once. The cost c_j is assumed to be a probabilistic variable to which noise

$\xi | \Delta c_j |$ is added, where the value of $\Delta c_j = N(0, c_j{}^2)$ is determined at the beginning of the present optimization, and ξ is a random variable taking 1 or 0 which changes during the optimization. In real situations, this probability in c_j would be attributed to traffic conditions changing in a probabilistic way. In the traveling salesman problem solved by SGA, one searches for the optimal solution $\boldsymbol{x}$ which maximizes f in (10) without considering the uncertainty in c_j. When c_j changes probabilistically, there possibly exists another individual $\boldsymbol{x}'$ whose expected value of $f(\boldsymbol{x}')$ is possibly larger than that of $f(\boldsymbol{x})$. The aim of the present method applied to this problem is to find such a solution.

Figure 4 shows the distribution of ten cities between which45 possible connections present. The GA parameters are as follows: N_{pop}=200, N_g=100, P_c=0.15, P_m=0.05, g=0.7.

The number of trials for each of the present method and SGA is 100. TABLE II summarizes the resultant values of f, defined by (10) and its expected value of F computed by the Monte Carlo method after the optimization. It can be seen from TABLE II that both values of f and F obtained by the present method are smaller than those obtained by SGA. It is concluded from this result that the present method has no superiority over the conventional GA as long as (10) is considered.

In practical traffic problems, extremely high transport costs which could be caused by traffic accidents and natural disasters should be avoided. To take this risk into account, (10) is modified to

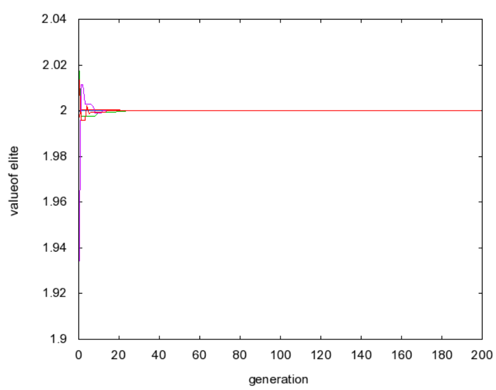
$$f = \begin{cases} 1/T & \text{if} \quad T \le T_S, \\ \varepsilon & \text{othersiwe}, \end{cases} \tag{11}$$

where ε and T_s represent penalty and threshold, which are set to 0.01, 0.12, respectively. The results for this modified problem is summarized in TABLE III, where the present method now yields better result in comparison with SGA with respect to F. Figure 5 shows the best solution of each method. Figure 6 shows the histograms of these solutions which are again obtained by the posterior Mote Carlo simulation, where the abscissa and ordinate represent the value of f and frequency. Figure 6 suggests that the solution obtained by SGA, shown in Fig. 5 (a), has no negative effects from the penalty term in (11) in contrast to that obtained by SGA.

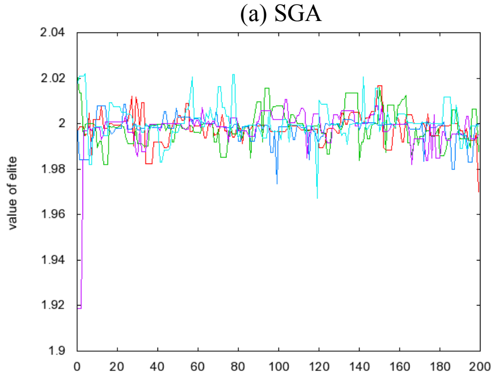
TABLE II Results for raveling salesman problem with (10)

methods	$f_{\rm opt} \times 10^{-1}$	$F_{\rm opt} \times 10^{-1}$
SGA	1.528	1.316
present	1.506	1.308

The symbols $f_{\rm opt}$ and $F_{\rm opt}$ represent the optimized value of f defined by (10) and its posterior expected value.



(a) SGA



(b) present method

Fig. 3 Time evolution of elite fitness

TABLE III Results for raveling salesman problem with (11)

methods	$f_{\rm opt} \times 10^{-1}$	$F_{\rm opt} \times 10^{-1}$
SGA	1.533	1.061
present	1.528	1.116

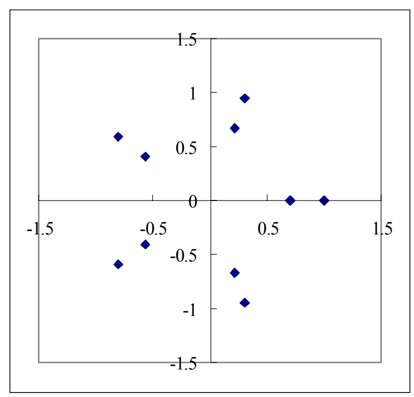


Fig. 4 Cities for traveling salesman problem

C. *Knapsack Problem with Uncertainty*

We finally consider the knapsack problem, whose parameters are summarized in TABLE IV. In this problem, the total price in the knapsack

$$P = \sum_i p_i n_i \tag{12}$$

is made as close to P_{center} as possible under the constraint that the total weight is less than $w_{\max}$, where p_i and n_i denote price and number of article i. Moreover, penalties are given if P is too far from P_{center}. Consequently, the objective function f is defined as

$$f(P) = \begin{cases} 0 & \text{if } \displaystyle\sum_{i \in \{A,B,C,D\}} w_i n_i > w_{\max}, \\ 10 & \text{else if } P < P_{\text{small}} \quad P > P_{\text{large}}, \\ g(P) & \text{otherwise,} \end{cases} \tag{13}$$

where

$$g(P) = \begin{cases} P & P < P_{\text{center}}, \\ -P + b & P \geq P_{\text{center}}. \end{cases} \tag{14}$$

The parameters in (13) and (14) are as follows: $w_{\max}$=150, P_{center}=1500, P_{small}=1250, P_{large}=1750, b=2 P_{center}. The profile of f given in (13) is shown in Fig. 7. In addition, the maximum number of articles in the knapsack is set to 15. The prices p_i, assumed to be probabilistic variables here, are added by noises obeying the normal distribution $N(0, \sigma^2)$, $\sigma = 0.15 p_i$. The GA parameters are as follows: N_{pop}=200, N_g=200, P_c=0.20, P_m=0.02, g=0.8.

The performance of the optimized solutions is summarized in TABLE V, which shows the mean values over 100 trials. While the optimal values of f are the same for both methods, the expected value F_{opt} for the present method is better than that for SGA. Examples of the solution (A, B, C, D) are: (7, 7, 3, 9), (1, 7, 14, 3) for SGA, and (14, 2, 11, 6), (14,6,6,7) for the present method.

TABLE IV Setting of knapsack problem

	A	B	C	D
price p_i	10	50	60	100
weight w_i	1	5	6	10

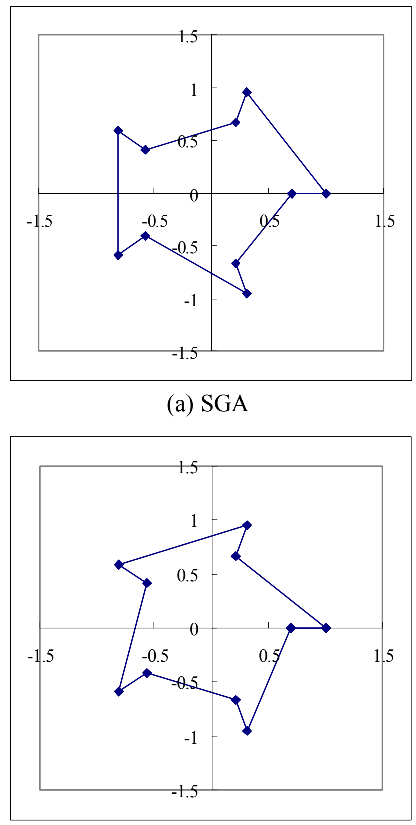
TABLE V Results for knapsack problem with (13)

methods	f_{opt}	F_{opt}
SGA	1500	1297
present	1500	1336

IV. Conclusions

This paper has described an evolutional optimization based on GA applied to the problems with respect to the objective functions including probabilistic constants. In this method, noises have been introduced to the constants during the GA processes.

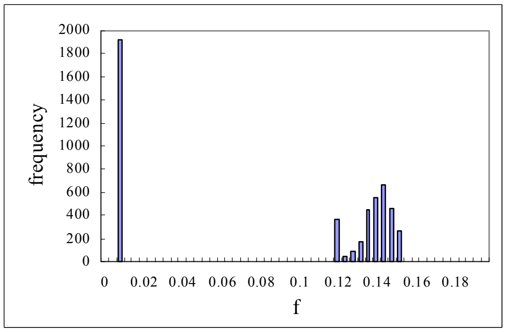
The present method is applied to the one-dimensional mathematical problem, and its result is in good agreement with that obtained by SGA in which the exact expected value of the probabilistic constant is set in the objective function. The present method is also applied to the traveling salesman and knapsack problems containing the probabilistic constants in their objective functions. It has been shown that the expected values of the objective functions, computed by the Monte Carlo simulation after the optimization, obtained by the present method are better than those obtained by SGA when the objective functions include the penalty terms.



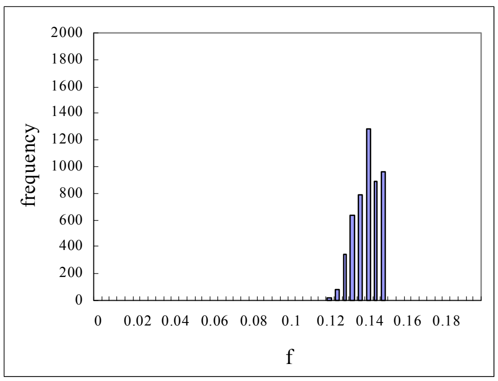
(a) SGA

(b) present method

Fig.5 Results of traveling salesman problem



(a) SGA



(b) present method

Fig.6 Histograms for traveling salesman problem

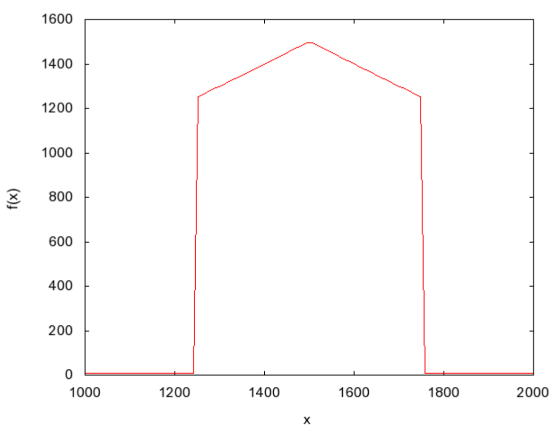


Fig. 7 Profile of *f* for knapsack problem defined by (13)

REFERENCES

[1] T. Kobayashi, D. L. Simon, “A Hybrid Neural Network-Genetic Algorithm Technique for Aircraft Engine Performance Diagnostics,” *Proc. 37th Joint Propulsion Conference and Exhibition*, ARL-TR-1266, AIAA-2001-3763, 2001.

[2] G. Gopalakrishnan, “Optimal Sampling in a Noisy Genetic Algorithm for Risk-Based Remediation Design,” .Doctor thesis in University of Illinois at Urbana Champaign.

[3] Y. Jin and J. Branke, “Evolutionary optimization in uncertain environments - a survey”, *IEEE Transaction on Evolutionary Computation*, vol.9, no.3, pp. 303-317, 2005.

[4] A. N. Aizawa, B. W. Wah, “Scheduling of genetic algorithms in a noisy environment,” *Evolutionary Computation*, vol.2, no.2, pp.97-122, 1994.

[5] A. Thompson, “On the automatic design of robust electronics through artificial evolution,” *Proc. Int. Conf. Evolvable Syst.*, 1998, pp. 13-24.

[6] G. Steiner, A. Weber, C. Magele, “Managing uncertainties in electromagnetic design problems with robust optimization,” *IEEE Transactions on Magnetics*, vol.40, no.2, pp. 1094-1097, 2004.

[7] D. H. Loughlin and S. R. Ranjithan, “Chance-constrained genetic algorithms,” *Proc. Genetic Evol. Comput. Conf.*, 1999, pp. 369–376.

[8] S. Tsutsui and A. Ghosh, “Genetic Algorithms with Robust Solution Searching Scheme”, *IEEE Transaction on Evolutionary Computation*, vol.1, no.3, pp. 201-208, 1997.

[9] T. Maruyama and H. Igarashi, “An Effective Robust Optimization Based on Genetic Algorithm,” *IEEE Transaction on Magnetics*, vol.44, no.6, pp. 990-993, 2008.