Optimization of Noisy Fitness Functions by means of Genetic Algorithms using History of Search with Test of Estimation

Yasuhito Sano[1] and Hajime Kita[2]
[1]Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, Nagatsuta, Yokohama 226-8502, JAPAN
sano@fe.dis.titech.ac.jp
[2]National Institution for Academic Degrees, Otsuka 3-29-1, Bunkyo, Tokyo 112-0012, JAPAN
kita@niad.ac.jp

Abstract - This paper discusses optimization of functions with uncertainty by means of Genetic Algorithms (GAs). In practical application of such GAs, possible number of fitness evaluation is quite limited. The authors have proposed a GA utilizing history of search (Memory-based Fitness Evaluation GA: MFEGA) so as to reduce the number of fitness evaluation for such applications of GAs. However, it is also found that the MFEGA faces difficulty when the optimum resides outside of the region where population covers because the MFEGA uses the history of search for estimation of fitness values. In this paper, the authors propose the tested-MFEGA, an extension of the MFEGA that tests validity of the estimated fitness value so as to overcome aforesaid problem. Numerical experiments show that the proposed method outperforms a conventional GA of sampling fitness values several times even when the original MFEGA fails.

I. Introduction

On-line optimization attracts attention as a means of the system optimization in the real world. However, there exist some difficulties in on-line optimization. That is, (1) uncertainty of evaluation value, (2) limitation of number of evaluation, (3) avoidance of extreme and dangerous trials. Moreover, optimization through computer simulation using random numbers also attract attention as another technique for optimization of a complicated real problem. This method shares similar difficulties (1) and (2) with on-line optimization.

The genetic algorithms (GAs) are applied to various optimization problems because of their broad applicability, including optimization of the function accompanied by uncertainty [1 - 4, 6, 8 - 10].

However if the usual GAs are applied to optimization of fitness function accompanied by uncertainty, they face problems of slow and unreliable convergence. Hence, formulation of a problem that takes the system optimization in the real world into consideration, and construction of optimization techniques that have a good performance to this problem are required.

Considering practical requirement of limitation on number of fitness evaluations, we have proposed the Memory-based Fitness Estimation GA: (MFEGA) a genetic algorithm that reduces uncertainty by estimating the fitness value with reference to the search history, and performs more efficient and exact optimization [9][10]. However, since the MFEGA uses the information acquired in the past, MFEGA faces a difficulty of premature convergence when the optimal solution locates the outside of the region that population covers.

The cause of such premature convergence is that an individual locating outside of region which population covers is not selected properly because estimation of its fitness value is inaccurate. In order to choose such individual properly, to use the sample fitness value of the individual is rather effective than to use the estimated fitness because the sample value may be more accurate than the estimated value. In this paper, taking the above point into consideration, we proposed a method of search using the sampled value instead of the estimated value when the between these two values is larger than a certain threshold.

This paper is organized as follows. Section 2 gives mathematical definitions and discusses some conventional methods for optimization of noisy fitness functions. Section 3 states a drawback of MFEGA and describes a new algorithm we propose. Section 4 shows the results of computer simulations. Section 5 concludes this study and suggests some perspectives.

II. Optimization of Noisy Fitness Functions

A. Formulation of Problem

In online optimization of real systems, the following three points should be taken into consideration.

(1) Some uncertainty exists in the system.
(2) The number of fitness evaluation is restricted.
(3) Trial of extreme solution candidates should be avoided.

0-7803-7282-4/02/$10.00

Also, in applications of GAs to optimization of system through large-scale computer simulation using random numbers, we have to take points **(1)** and **(2)** into consideration. As for uncertainty of the system, it is categorized into the following three cases, 1) an uncertain change invades into the input to a system, 2) an uncertain fluctuation is involved in the observable output, and 3) system itself or its environment changes. In this paper, we consider the case 2) where an uncertain change follows on the output. GAs for search of robust solution based on the formulation of case 1) in also discussed [5], which is closely related the issue treated in this paper.

The problem is formulated as follows:

$$\min_{\boldsymbol{x} \in \boldsymbol{R}} \langle F(\boldsymbol{x}) \rangle, \tag{1}$$

$$F(\boldsymbol{x}) = f(\boldsymbol{x}) + \delta \tag{2}$$

where $\boldsymbol{R}$ represent the search space, f and F are fitness function and its observation. In the following, we consider the search space $\boldsymbol{R}$ is continuous. δ is additive random fluctuation whose mean is zero, and $\langle\rangle$ represents expectation over δ. We call $f(\boldsymbol{x})$ 'the true fitness value', and $F(\boldsymbol{x})$ 'the sampled fitness value', respectively.

B. Genetic Algorithms for Noisy Fitness Functions

In optimization of noisy fitness function with usual GAs, a problematic situation occurs. That is, a good solution may not necessarily be survivor of selection since it is evaluated by a sampled fitness value. Hence, we need some devices of reducing influence of fluctuation for optimization of noisy fitness function. Further, from the viewpoint of practical use, limitation of number of fitness evaluation should be also considered.

The following techniques have been proposed as approaches to this problem.

B.1 Using Mean of Several Samples for Each Individual

Fitzpatrick et al. have proposed the method that samples fitness value two or more times from the same individual, and uses the average value as an estimated fitness value [2]. It reduces the variance of fitness values without particular additional assumption, and more accurate search is achieved. However, in practical usage of GAs, such the method is not a desirable because it requires larger samples.

Branke has proposed a method that only the best individual is evaluated by the average of several samples of fitness [3]. With this method, evaluations of unimportant solution are reduced. Stagge has used a technique in which the number of samples is decided using t-test so as to confirm significance of good solution [4]. However, even with these techniques, basically large number of samples are required.

B.2 Referring to the Sample Value of Other Individuals

Tanooka et al. have used a method of referring to parent's sampled fitness values for search of the robust optimal solution [5]. This technique can raise the accuracy of a solution without increasing the number of evaluation. However, since an individual with the apparently good fitness survives by selection, there is a problem that bias due to selection is included in the estimated value. Moreover, Branke has also proposed the technique of referring to the sampled fitness values of the individuals which are near the point of interest in the current and previous generations [3]. However, since referred samples are small in number, there is a problem that sufficient accuracy is not attained.

B.3 Using a Threshold Value

Markon et al. have proposed a technique that individual is not replaced unless the sampled fitness value of a child is less than parent's sampled values by the width beyond a threshold [8]. With this technique, failure of preservation of the best solution in the search can be reduced. If sufficiently many numbers of sampling is allowed, it is proved that this technique always finds the optimal solution for the sphere fitness function. However, this method dose not pay attention to restriction of possible fitness evaluation in practical applications.

B.4 Memory Based Fitness Estimation GA

Considering difficulties of other techniques in practical applications having severe limitation of available number of fitness evaluation, we have proposed the Memory-based Fitness Evaluation GA (MFEGA) [9]. MFEGA estimates the true fitness value using a search history. The composition of MFEGA is as follows.

1. A stochastic model about the uncertainty of the fitness function is introduced. (See Appendix I).
2. The sampled value of fitness value are acquired during search, and stored as the search history.
3. Fitness of an individual is estimated by a statistical technique referring the search history. (See Appendix II and III)

III. Problem of Using a Search History, and Proposal of tested-MFEGA

A. Drawback of MFEGA

Estimation technique used in the MFEGA assumes that points of interest locates inside the region covered by the

0-7803-7282-4/02/$10.00 ©2002 IEEE

individual in the history. If points of interest, i.e., candidates for further search, locate outside the region, estimation used in MFEGA may become harmful, and cause some premature convergence due to poor estimation of fitness value of good solution outside the region.

In this paper, we assume that the variance of a noise contained in the observable fitness value is constant in the search space. Therefore, in the place where estimation error becomes large rather than a noise, we should use rather the sample values itself than the estimation. In this paper, we propose the following method. If there is sufficient difference between the sample values, we remove bad sampled individuals from the candidate of selection.

Let $\boldsymbol{y}_i, i \in \{1, 2, \cdots, F\}$ be family individuals those are candidates of selection, let $F(\boldsymbol{y}_i)$ be sample values of $\boldsymbol{y}_i$ and let $\boldsymbol{y}_{\text{Best}} = \arg\min_{\boldsymbol{y}_i} F(\boldsymbol{y}_i)$ be the best individual in the family. We divide the family into $\boldsymbol{Y}_{Accept}$ and $\boldsymbol{Y}_{Reject}$ by the following rule before selection by the estimated value.

$$\begin{cases} \textbf{If } F(\boldsymbol{y}_i) - F(\boldsymbol{y}_{\text{Best}}) < Z \\ \qquad\qquad \textbf{then } \boldsymbol{Y}_{Accept} = \boldsymbol{Y}_{Accept} \cup \boldsymbol{y}_i \\ \textbf{Else} \qquad\quad \boldsymbol{Y}_{Reject} = \boldsymbol{Y}_{Reject} \cup \boldsymbol{y}_i \end{cases} \tag{3}$$

where Z is a threshold value using the view of Z-test, $\boldsymbol{Y}_{accept}$ are candidates for selection and $\boldsymbol{Y}_{reject}$ are not.

As for the value of Z, we use the value which makes probability of error of the first kind be 0.3 or less by Z-test. We determined this level by preliminary experiment. In practical application of on-line optimization of the system, generally some preliminary experiment to determine this value will be possible. This technique is called the tested-MFEGA.

B. Algorithm of the tested-MFEGA

The fitness estimation method used in MFEGA can be combined with any type of crossovers and generation alternation models since it is independent of crossovers and generation alternation models. We use the UNDX for crossover proposed by Ono [6] and the generation alternation model replacing an individual partially.

The algorithm of the tested-MFEGA is as follows.

1. Initialize the population of M individuals $\boldsymbol{x}_1, ..., \boldsymbol{x}_M$ randomly.
2. Let evaluation counter $e = 0$. Set the maximal number of evaluations to E.
3. Let history $H = \emptyset$.
4. Choose two individuals $\boldsymbol{x}_{p1}$ and $\boldsymbol{x}_{p2}$ from the population and let them be family individuals $\boldsymbol{y}_1, \boldsymbol{y}_2$.
5. Produce C children $\boldsymbol{y}_{2+i}, i = 1, 2, \cdots C$ by applying the crossover to the parents $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$.
6. Sample fitness values $F(\boldsymbol{y}_i)$ for $i = 1, ..., C + 2$.
7. Let $e = e + C + 2$.
8. Store the sampled values into the history H, i.e., $H = H \cup \{(\boldsymbol{y}_i, F(\boldsymbol{y}_i)) | i = 1, ..., C + 2\}$.
9. Select the individual $\boldsymbol{h}_{\min}$ having the smallest sampled fitness value from H.
10. Estimate k' by maximization of Eq. (13) in Appendix III using $\boldsymbol{h}_{\min}$.
11. Obtain estimation of the fitness values $\tilde{f}(\boldsymbol{y}_i)$ by Eq. (12) shown in Appendix II.
12. Let $\boldsymbol{Y}_{Reject} = \emptyset$ and $\boldsymbol{Y}_{Accept} = \emptyset$ and divide $\boldsymbol{y}_i$ into $\boldsymbol{Y}_{Accept}$ and $\boldsymbol{Y}_{Reject}$ by rule (3).
13. Substitute the parents individuals $\boldsymbol{x}_{p1}$ and $\boldsymbol{x}_{p2}$ with individual having two smallest $\tilde{f}(\boldsymbol{y}_i)$ in $\boldsymbol{Y}_{Accept}$.
14. If $e \leq E$, go to Step 4, otherwise terminate the algorithm.

IV. Experiment

A. Basic Behaviors of MFEGA and tested-MFEGA

First, we compare the basic performances of the MFEGA and the tested-MFEGA with other conventional GAs. We used the following fitness function:

$$F_{\text{Sphere}}(\boldsymbol{x}) = \sum_{i=1}^{D} x_i^2 + \delta \tag{4}$$

$$\delta \sim N(0, \sigma_N^2) \tag{5}$$

where D is the dimension of search space, and δ is an additive noise following the normal distribution with zero mean and variance of σ_N^2.

The following four algorithms are compared.

- MFEGA
- tested-MFEGA
- Standard GA: A GA using a single fitness sample value for evaluation of each individual.
- Sample 10-GA: A GA using the mean of 10 fitness value samples for evaluation of each individual.

So as to provide reference performance where a noise is not contained, the result using Standard GA for fitness function without noise is also shown as Noiseless GA.

In this paper, we used the UNDX for crossover in all GAs and the generation alternation model replacing an individual partially. Since the search ability of the UNDX is sufficiently high, we didn't employ mutation. Conditions of experiments are shown in Table 1. The number of fitness evaluation E is decided considering a practical application of on-line adaptation [10].

For each method, 20 trials are performed, and we compared the studied methods by the mean of the following

0-7803-7282-4/02/$10.00 ©2002 IEEE

TABLE I
CONDITIONS OF EXPERIMENT

Generation Size	$G = 100$ $(E = 700)$
Population Size	$P = 30$
Children Size	$C = 5$
Dimension	$D = 10$
Range of Initial Population	$[\,-r/2 : r/2]$ $r = 1.0$
Variance of Noise	$\sigma_N^2 = 1.0$

value.

$$f_{\text{Best}} = \begin{cases} f(\arg\min_{\boldsymbol{y}} F(\boldsymbol{y})); \\ \qquad \text{for Standard GA and Noiseless GA} \\ f(\arg\min_{\boldsymbol{y}} \tilde{f}(\boldsymbol{y})); \\ \qquad \text{for Sample 10-GA,} \\ \qquad \text{tested-MFEGA and MFEGA} \end{cases}$$

Figure 1 shows evolution of the means of 20 trials of the performance index f_{Best}. For Standard GA, f_{Best} is improved until around 1000 evaluations, but it stagnated around 0.2 after. The convergence of Sample 10-GA is very slow while steady improvement continued. Contrary to this, the MFEGA converges quickly and finds more accurate solutions than the other two methods. As for the tested-MFEGA, although its convergence speed from 400 evaluations to 1000 evaluations is inferior to the MFEGA, it is quite high as compared with the Standard GA and the Sample 10-GA. Moreover, the tested-MFEGA shows steady improvement even after 1000 evaluation while the search of the MFEGA stagnates. Finally, more accurate solution is found by the tested-MFEGA than the MFEGA.

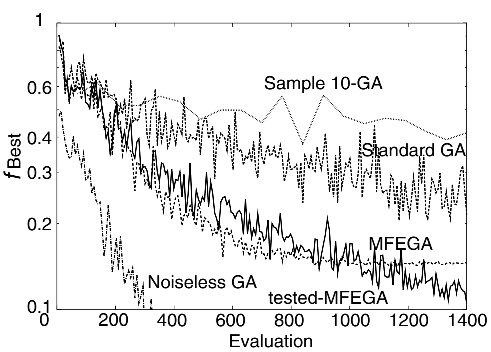

Fig. 1. Search precision for F_{Sphere}

B. Influence of Offset of the Optimum

The MFEGA estimates a fitness value as a weighted sum of the sampled fitness values in the search history. Therefore estimation will be poor for the point outside the region that covered by the history, and the performance of the MFEGA will be degraded. In this section, we show the degradation of the performance for optimization of fitness function whose optimal value has some offset:

$$F_{\text{Offset}}(\boldsymbol{x}) \;=\; \sum_{i}^{D}(x_i - o)^2 + \delta \qquad (6)$$

$$\delta \;\sim\; N(0, \sigma_N^2) \qquad (7)$$

where $o \in \{0.3, 0.5, 0.7, 1.0, 1.3\}$ is a parameter of offset. When $o = 0.3$, the optimum resides inside of region covered by the initial populationwhen $o = 0.7 \sim 1.3$ the optimum resides outside, and $o = 0.5$, it locates at the boundary.

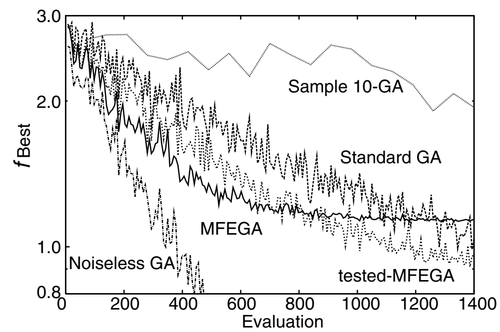

Fig. 2. Search Precision for F_{Offset}, $o = 0.5$

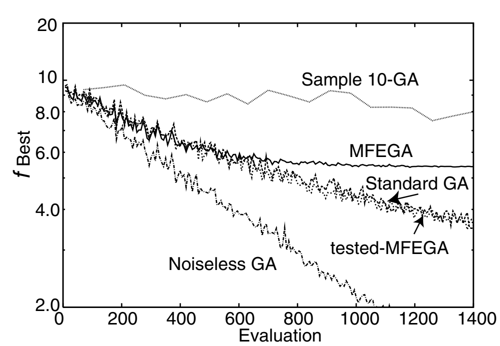

Fig. 3. Search Precision for F_{Offset}, $o = 1.0$

Fig. 2 is the search performance with $o = 0.5$. The Sample 10-GA shows convergence like previous experiment. The MFEGA achieves quick convergence up to 1000 evaluation. However stagnation is shown afterward. On the other hand the Standard GA achieves continuous improvement, and finally it outperforms the MFEGA. As for the tested-MFEGA, although convergence speed in early stages of search is slower than the MFEGA, the

0-7803-7282-4/02/$10.00 ©2002 IEEE

tested-MFEGA shows always better performance than the Standard GA. After 900 evaluations, it outperforms the MFEGA.

Fig. 3 shows the result with $o = 1.0$. The performance of the Sample 10-GA is quite poor also in this case. The MFEGA converges in early stage. The performance of the tested-MFEGA is almost same with that of the Standard GA. This is because almost all the samples are rejected by the Z-test.

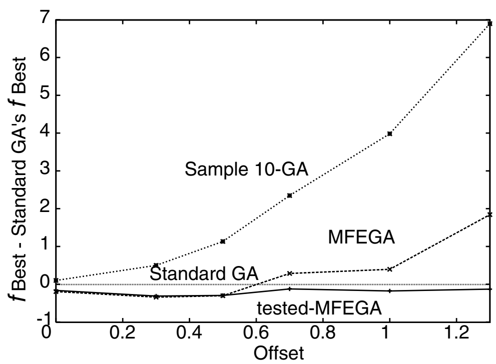

Fig. 4. Comparison of a search performance

Fig. 4 shows the comparison of performance f_{Best} of all the methods against to the parameter o. The difference of the performance between each method and that of Standard GA is plotted. It is shown that the Sample 10-GA is always the worst. The MFEGA is the best method for $o = 0.3$ and 0.5, and it is outperformed by the Standard GA for $o = 0.7, 1.0, 1.3$. Performance of tested-MFEGA shows good performance for all the parameter values.

From these results, we can confirm that tested-MFEGA shows robustness in search ability.

V. Conclusion

The Memory-based Fitness Evaluation GA (MFEGA) proposed by the authors is a genetic algorithm for optimization of noisy fitness function considering limitation of number of fitness evaluation in practical applications. While the MFEGA shows good performance in general, it is also observed that the MFEGA shows some premature convergence due to poor estimation of fitness value. This paper discusses a typical situation of such phenomena, and propose an improved method called the tested-MFEGA. Numerical simulation shows that the tested-MFEGA works well even in the situation where the original MFEGA faces difficulties.

Our subjects of future study are development for the technique of further improvement of convergence of the MFEGA, and to show effectiveness of the tested-MFEGA in practical problems. Development of a genetic algorithm for optimization of fitness function using the search history for the problem of optimum changes randomly is another line of future study [11].

Finally, the authors would like to acknowledge helpful discussions by Prof. Shigenobu Kobayasi of Tokyo Institute of Technology and Mr. Jürgen Branke of University of Karlsruhe.

References

[1] I. Kamihara, M. Yamaguchi and H. Kita: On-line Adaptation of Vehicles by Means of an Evolutionary Control System, *Proc. IEEE International Conference on Systems, Man, and Cybernetics (SMC '99)*, pp. 553-558 (1999).

[2] J. M. Fitzpatrick and J. J. Greffenstette: Genetic algorithms in noisy environments, *Machine Learning*, Vol. 3 pp.101-120 (1988).

[3] J. Branke: Creating Robust Solutions by Means of Evolutionary Algorithms *Proc. of 5th PPSN*, pp. 119-128 (1998).

[4] P. Stagge: Averaging Efficiently in the Presence of Noise, *Proc. of 5th PPSN*, pp. 188-197 (1998).

[5] K. Tanooka, H. Tamaki, S. Abe and S. Kitamura: A Continuous Age Model of Genetic Algorithms Applicable to Optimization Problems with Uncertainties, *Proc. IEEE International. Conference on Systems, Man, and Cybernetics (SMC '99)*, pp. 637-642 (1999).

[6] I. Ono and S. Kobayashi: A Real-Coded Genetic Algorithm for Function Optimization Using Unimodal Normal Distribution Crossover, *Proc. of 7th ICGA*, pp. 246-253 (1997).

[7] N. Mori, H. Kita and Y. Nishikawa: Adaptation to a Changing Environment by Means of the Feedback Thermodynamical Genetic Algorithm *Proc. of 5th PPSN*, pp. 149-159 (1998).

[8] S. Markon, V. Arnold, T. Bäck, T. Beielstein and H.-G. Beyer: Thresholding-a Selection Operator for Noisy ES, *Proc. of IEEE Congress of Evolutionary Computation* pp. 465-472 (2001).

[9] Y. Sano and H. Kita: Optimization of Noisy Fitness Functions by means of Genetic Algorithms using History of Search. *Proc. of 6th PPSN*, pp. 571-580 (2000).

[10] Y. Sano and H. Kita, I. Kamihira and M. Yamaguchi: Optimization of an Engine Controller by means of a Genetic Algorithm using History of Search. *Proc. of SEAL 2000*, pp. 2929-2934 (2000).

[11] N. Mori, S. Imanishi, H. Kita and Y. Nishikawa: Adaptation to a Changing Environment by Means of the memory based thermodynamical genetic algorithm; *Proc. of 7th ICGA*, pp. 299-306 (1997).

Appendix

I. Stochastic Model of Fitness Functions

The MFEGA has adopted a simple stochastic model that fitness values of individuals distribute randomly around the fitness value of the individual of interest, and have assumed that the variance of the fitness value depends only on the distance from the individual of interest. Let $\boldsymbol{x}$ be an individual we want to estimate its fitness value, and let $\boldsymbol{h}$ be an individual in the history of search, whose distance from $\boldsymbol{x}$ is d and sampled fitness value is $F(\boldsymbol{h})$. We assume the following model:

$$
\begin{aligned}
f(\boldsymbol{h}) &\sim N(f(\boldsymbol{x}), kd) && (8)\\
F(\boldsymbol{h}) = f(\boldsymbol{h}) + \delta &\sim N(f(\boldsymbol{x}), kd + \sigma^2) && (9)
\end{aligned}
$$

where $f(\boldsymbol{h})$ is the true fitness of individual $\boldsymbol{x}$, and k is a positive parameter. The additive noise δ is assumed to follow $N(0, \sigma^2)$. As shown in Fig.5, Eqs. (8) and (9) mean that the true fitness $f(\boldsymbol{h})$ distributed randomly around the $f(\boldsymbol{x})$ following normal distribution of a variance proportional to the distance d, and hence observation of the fitness $F(\boldsymbol{h})$ follows the normal distribution of a variance $kd + \sigma^2$ considering the additive observation noise.

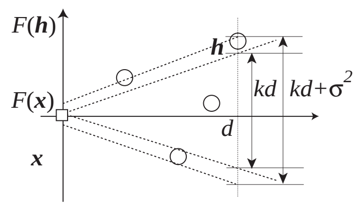

Fig. 5. A Stochustic Model of Fitness Functions

II. Estimation of Fitness by the Maximum Likelihood Method

If the parameters k and σ^2 in Eqs. (8) and (9) are known in advance, the fitness value $f(\boldsymbol{x})$ can be estimated by the maximum likelihood method using the search history based on the above model. Let $\boldsymbol{h}_l, l = 1, ..., H$ be the H individuals in the search history and $F(\boldsymbol{h}_l)$ and $d_l, l = 1, 2, ..., H$ be their sampled fitness values and distances from the individual of interest $\boldsymbol{x}$, respectively. The probability of obtaining $F(\boldsymbol{h}_1), ..., F(\boldsymbol{h}_H)$ is represented by

$$\prod_{l=1}^{H} p(F(\boldsymbol{h}_l), d_l) \tag{10}$$

where $p(F(\boldsymbol{h}_l), d_l)$ is the probability density function of $F(\boldsymbol{h}_l)$ given by the normal distribution:

$$
\begin{aligned}
&p(F(\boldsymbol{h}_l), d_l)\\
&= \frac{1}{\sqrt{2\pi(kd_l + \sigma^2)}} \exp\left(-\frac{1}{2}\frac{(F(\boldsymbol{h}_l) - f(\boldsymbol{x}))^2}{kd_l + \sigma^2}\right)
\end{aligned} \tag{11}
$$

Treating Eq. (10) as the likelihood w.r.t. $f(\boldsymbol{x})$, estimation of $f(\boldsymbol{x})$, say $\tilde{f}(\boldsymbol{x})$, can be obtained by maximizing Eq. (10) for $f(\boldsymbol{x})$ as a weighted average of the sampled fitness values:

$$\tilde{f}(\boldsymbol{x}) = \frac{\sum_{l=1}^{H} \frac{F(\boldsymbol{h}_l)}{k'd_l + 1}}{\sum_{l=1}^{H} \frac{1}{k'd_l + 1}} = \frac{F(\boldsymbol{x}) + \sum_{l=2}^{H} \frac{1}{k'd_l + 1} F(\boldsymbol{h}_l)}{1 + \sum_{l=2}^{H} \frac{1}{k'd_l + 1}} \tag{12}$$

where $k' = k/\sigma^2$.

III. Estimation of the Model Parameters

In actual situation, the parameters k (or k') and σ^2 in the Eq. (12) of MFEGA are unknown in advance, and it is needed to estimate these parameters. We employ the maximum likelihood technique for estimation of k' . Taking Eq. (10) as the likelihood of the parameters k', and a logarithm likelihood of the parameters is calculated from Eq. (10) as follows:

$$
\begin{aligned}
\log L \;=\; &-\frac{1}{2}\Bigg(H \log 2\pi + \sum_{l=1}^{H} \log \sigma^2 (k'd_l + 1)\\
&+ \sum_{l}^{H} \frac{(F(\boldsymbol{h}_l) - f(\boldsymbol{x}))^2}{\sigma^2(k'd_l + 1)}\Bigg)
\end{aligned} \tag{13}
$$

Differentiating $\log L$ by σ^2, we obtain the following equation.

$$\frac{\partial}{\partial \sigma^2} \log L = -\frac{1}{2}\left(\frac{H}{\sigma^2} - \sum_{l=1}^{H} \frac{(F(\boldsymbol{h}_l) - f(\boldsymbol{x}))^2}{\sigma^4(k'd_l + 1)}\right) \tag{14}$$

Let the RHS of Eq. (14) be 0 and solving it for σ^2, the optimal σ^2 is represented as follows:

$$\sigma^2 = \sum_{l=1}^{H} \frac{(F(\boldsymbol{h}_l) - f(\boldsymbol{x}))^2}{H(k'd_l + 1)} \tag{15}$$

Substituting σ^2 in Eq. (13) with Eq. (15), the logarithm likelihood of k' is obtained. Since this equation includes unknown variables $f(\boldsymbol{x})$ and d_l, we set $\boldsymbol{x}$ at the individual having the minimum sampled fitness considering that the usage of the model is optimization of $f(\boldsymbol{x})$. The fitness value of $f(\boldsymbol{x})$ is simply estimated by the average of the sampled fitness values of five individuals near by. Distance d_l is easily calculated if $\boldsymbol{x}$ is decided. For maximization of the likelihood, a numerical hill climb method w.r.t. the logarithm of k' is used, considering positiveness of k'