論 文

A Hybrid Selection Strategy Using Scalarization and Adaptive ϵ-Ranking for Many-objective Optimization

Hernán Aguirre — International Young Researcher Empowerment Center, Shinshu University
ahernan@shinshu-u.ac.jp

Kiyoshi Tanaka — Faculty of Engineering, Shinshu University
ktanaka@shinshu-u.ac.jp

keywords: hybrid strategy, scalarization, adaptive ϵ-ranking, many-objective optimization, MNK-landscapes

Summary

This work proposes a hybrid strategy in a two-stage search process for many-objective optimization. The first stage of the search is directed by a scalarization function and the second one by Pareto selection enhanced with Adaptive ϵ-Ranking. The scalarization strategy drives the population towards central regions of objective space, aiming to find solutions with good convergence properties to seed the second stage of the search. Adaptive ϵ-Ranking balances the search effort towards the different regions of objective space to find solutions with good convergence, spread, and distribution properties. We test the proposed hybrid strategy on MNK-Landscapes and DTLZ problems, showing that performance can improve significantly. Also, we compare the effectiveness of applying either Adaptive ϵ-Ranking or NSGA-II's non-domination sorting & crowding distance in the second stage, clarifying the necessity of Adaptive ϵ-Ranking. In addition, we include a comparison with two substitute assignment distance methods known to be very effective to improve convergence on many-objective problems, showing that the proposed hybrid approach can find solutions with similar or better convergence properties on highly complex problems, while achieving better spread and distribution.

1. Introduction

Recently, there is a growing interest on applying multi-objective evolutionary algorithms (MOEAs) to solve *many*-objective optimization problems with four or more objective functions. In general, conventional MOEAs [Deb 01, Coello 02] scale up poorly with the number of objectives and new evolutionary algorithms are being proposed [Ishibuchi 08]. Research has focused mainly on the effectiveness of selection, dimensionality reduction, incorporation of user preferences, and space partitioning.

The lack of discriminatory power by Pareto dominance and a poor scalability of density estimators affect seriously the performance of conventional Pareto MOEAs in many-objective spaces. Likewise, the number of scalarization functions required for many-objective optimization increases substantially the complexity of scalarization based MOEAs. These are serious issues that have motivated research on methods to improve selection in MOEAs for many-objective optimization (see below for related works).

Dimensionality reduction approaches try to reduce the number of objectives the evolutionary algorithm uses to solve the problem [Deb 06b, Brockhoff 07, López Jaimes 08]. Here the main challenge consists in how to determine a minimum set of objectives functions that preserves most of the characteristics of the original problem. When such reductions are possible, the (lower dimensional) problem becomes more amenable to the evolutionary algorithm and to the decision maker as well. However, even if dimensionality reduction is possible, there is no guarantee that the new dimension of the problem will be low enough to overcome the problems of weak selection pressure of conventional MOEAs.

Methods that focus on the incorporation of user preferences aim to provide the MOEA with a reference point so that it can concentrate its search on a small region of the Pareto front [Deb 06a]. Incorporation of user preferences is a very interesting and useful approach, however here the assumptions are that the user has a clear idea about the Pareto front and knows where to look for solutions. When such knowledge does not exist, we first need methods that can search effectively in many-objective spaces and generate an approximation of the Pareto front.

Space partitioning [Aguirre 09] instantaneously partitions the objective space into few lower dimensional subspaces and concurrently searches in each of them, aim-

ing to find good approximations of the true Pareto front in the original high-dimensional objective space. Space partitioning uses a partition strategy to define a schedule of subspace sampling, so that different subspaces can be emphasized at different generations. This approach has been used effectively to improve the effectiveness of selection in high dimensional spaces, and could be used as a framework for dynamic dimensionality reduction, and to realize parallel implementations of many-objective optimization algorithms.

In practice, the optimization of many-objective real-world problems is not a single step process. Rather, it often consists of a series of steps in which the above mentioned approaches could be used to complement and support each other. Thus, research to improve and combine these approaches is important to foster further developments on many-objective optimization.

This work focuses on the effectiveness of selection on many-objective optimization. Some methods have been proposed to improve Pareto selection for many-objective optimization by incorporating indicator functions or extensions of Pareto dominance [Zitzler 04, Emmerich 05, Corne 07, Kukkonen 07, Ishibuchi 07b, Koppen 07, Sato 07]. Most of these methods induce a different ranking based on information of how close solutions are to dominate other non-dominated solutions and have been proved effective to improve convergence at the expense of spread or vice-versa. To rank solutions, these methods compare each solution with all other solutions, bringing the computational order to $\mathcal{O}(M|P|^2)$, where M is the number of objectives and $|P|$ the population size. Other methods are based on scalarization functions that map the multi-objective problem to a single-objective one [Ishibuchi 07a, Hughes 05]. Since a scalarization function defines a search direction around a single point in objective space, to try to uniformly search the objective space and find good approximations of the Pareto front, very many scalarization functions must be specified. The computational order of ranking with one scalarization functions is $\mathcal{O}(M|P|)$. However, usually the number of scalarization functions for many-objective optimization is of the same order of the population size, making the overall computational order similar to $\mathcal{O}(M|P|^2)$.

In this work, we propose a hybrid strategy in a two-stage search process. The first stage of the search is directed by a scalarization function and the second one by Pareto selection enhanced with Adaptive ϵ-Ranking [Aguirre 09]. The scalarization function provides a computationally fast unifying search direction to drive the population towards central regions of objective space, aiming to find a subset of solutions with good convergence properties. On the other hand, Adaptive ϵ-Ranking uses the local information of the distribution of solutions to balance the search effort towards the different regions of objective space, increasing the discriminatory power of Pareto selection while introducing simultaneously a density estimator that scales-up well on high dimensional spaces, to find solutions with good convergence, spread, and distribution properties.

We study the effects of the scalarization and Adaptive ϵ-Ranking applied independently. Then, we study the effects of the proposed hybrid strategy, showing that it can significantly outperform its individual components. Also, since the hybrid strategy uses just one scalarization function during the first stage, it becomes considerably faster, which is an important issue for scalability on high-dimensional spaces. Fixing scalarization as the first stage strategy, we compare Adaptive ϵ-Ranking and NSGA-II's non-domination sorting $\&$ crowding distance as the strategy for the second stage, clarifying the necessity of Adaptive ϵ-Ranking. Also, we compare the proposed hybrid strategy with Subvector Dominance Assignment and Epsilon Dominance Assignment [Koppen 07], two substitute assignment distance methods known to be very effective to improve convergence on many-objective problems, showing that the hybrid approach can found solutions with similar or better convergence properties on highly complex problems, while achieving better spread and distribution. As benchmark instances we use MNK-Landscapes [Aguirre 07] with $4 \leq M \leq 10$ objectives, $N = 100$ bits, and $0 \leq K \leq 50$ epistatic interactions per bit. In addition, we test the proposed hybrid strategy on continuous problems with non-convex fronts using DTLZ functions with $4 \leq M \leq 10$ objectives and 100 variables. Parts of this work have been presented in [Aguirre 10b] and [Aguirre 10c].

2. Proposed Hybrid Strategy

2·1 Concept

Multi-objective optimizers seek to find trade-off solutions with good properties of convergence to the Pareto front, well spread and uniformly distributed along the front. These three properties are especially difficult to achieve in many-objective problems and most searching strategies compromise one in favor of the other. In addition, larger population sizes are likely to be required in order to create an appropriate approximation of the Pareto front in high dimensional spaces. Both, larger populations and high dimensionality, impose a serious challenge to the computational scalability of current algorithms.

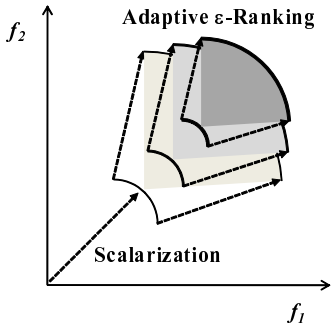


Fig. 1 Hybrid Strategy

Seeking to find approximations of the Pareto front fulfilling the three properties of convergence, spread, and distribution, rather than expecting a sole strategy to work efficiently for all tasks, it seems reasonable to distribute the search into different strategies that complement each other. The proposed hybrid strategy follows this approach, using one strategy from the domain of scalarization that focus on convergence-only and the other one from the domain of Pareto dominance and its extensions (Adaptive ϵ-Ranking) that in addition to convergence also pays attention to diversity. The strategies are deployed following a two-stage scenario, assigning one strategy to each stage, where the first strategy works to seed the second one, as illustrated in **Figure 1**. The expectation for the hybrid strategy is that better results could be achieved by diversifying the population after some degree of convergence has been achieved than by emphasizing convergence and diversity since the beginning of the search, where the population is random. Also, by simplifying the scalarization strategy to one scalarizing function, it is expected that the hybrid method could speed up the search.

2·2 Scalarization Strategy μF

Scalarization functions have played an essential role in solving multi-objective optimization problems. In scalarization, the multi-objective problem is transformed into one or a series of single objective optimization problems, involving possibly some parameters or constraints in order to capture some kind of preference information from the decision maker. Many different scalarization functions have been suggested in the literature based on different approaches and multi-objective optimization methods utilize them in various ways. The input requested from the decision maker may consist on trade-off information, marginal rates of substitution, desirable objective function values, or a reference point. Furthermore, the scalarization may be performed once or repeatedly as a part of an iterative process in which the decision maker is directly involved.

In multi-objective evolutionary algorithms the most frequently used scalarization functions are weighted sum fitness functions with non-negative weight vector, reference point-based fitness function, and ϵ-constrained fitness function. Weighted sum fitness functions, expressed by

$$fitness(\boldsymbol{x}) = w_1 f_1(\boldsymbol{x}) + w_2 f_2(\boldsymbol{x}) + \cdots w_M f_M(\boldsymbol{x}), \quad (1)$$

set with different weight vectors $\boldsymbol{w} = (w_1, w_2, \cdots, w_M)$ were successfully used to specify various search directions to search concurrently on them by multi-objective local search algorithms [Ishibuchi 98, Jaskiewics 02]. Very-many weighted sum fitness functions have also been used for many-objective optimization [Hughes 05], as mentioned in previous sections. The reference point-based fitness function is the distance from the original fitness values of a solution $\boldsymbol{f}(\boldsymbol{x})$ to a desired reference point in the objective space $\boldsymbol{f}^* = (f_1^*, f_2^*, \cdots, f_M^*)$, i.e.

$$fitness(\boldsymbol{x}) = distance(\boldsymbol{f}^*, \boldsymbol{f}(\boldsymbol{x})). \quad (2)$$

The incorporation of reference points into MOEAs was examined in [Deb 06a]. The ϵ-constrained fitness function is based on the wide spread ϵ-constrain approach, in which lower (or upper) bounds for some objectives are specified by inequality conditions. The incorporation of a ϵ-constrained fitness function into MOEAs was examined in [Ishibuchi 07a] using the following function

$$fitness(\boldsymbol{x}) = f_k(\boldsymbol{x}) - \alpha \sum_{i=1}^{k-1} \max\{0, f_i(\boldsymbol{x})\}, \quad (3)$$

where lower bounds were specified for the the first (k-1) objectives such that $f_i(\boldsymbol{x}) \geq \epsilon_i$, ϵ_i specifies the lower bound of the i-th objective, and α is a penalty with respect to the violation of the inequality condition. In the above fitness functions, the weights, the reference point, and the lower bounds for some objectives allow to specify one or more regions of interest for the evolutionary algorithm to focus on.

In our method, the role of the scalarization strategy is to provide a computationally-fast unifying search direction to drive the population towards central regions of objective space, so that solutions with good convergence could be found to seed the second stage of the search. That is, for computational reason we are interested on using only one scalarization function and our region of interest is the central part of objective space. A weighted sum function, set with equal weights for all objectives, can fulfill these requirements without requiring knowledge of the problem at hand [Ishibuchi 07b]. Thus, in this work we use the

following scalarizing function,

$$g = \frac{1}{M}\sum_{i=1}^{M} f_i, \tag{4}$$

where f_i denotes the i-th objective value and M the number of objectives.

2·3 *Adaptive ϵ-Ranking Strategy* $A\epsilon R^E$

Pareto ranking classifies the entire population in one or more sets of equally ranked solutions $\mathcal{F}_i$ ($i = 1, \cdots, N_F$), each set associated to rank i. On many-objective problems the number of Pareto non-dominated solutions increase substantially with the dimensionality of the objective space and $|\mathcal{F}_1|$ usually becomes larger than the size of the parent population $|P|$ from early generations [Aguirre 07].

ϵ-Ranking re-classifies the sets $\mathcal{F}_i$ ($i = 1, \cdots, N_F$) into sets $\mathcal{F}_j^\epsilon$ ($j = 1, \cdots, N_F^\epsilon$), $N_F^\epsilon \geq N_F$, using a randomized sampling procedure that favors a good distribution of solutions based on dominance regions wider than conventional Pareto dominance (ϵ-dominance). The sampling heuristic favors an effective search using the following criteria. (i) Extreme solutions are always part of the sample. (ii) Each (not extreme) sampled solution is the sole sampled representative of its area of influence, which is determined by ϵ-dominance. (iii) Sampling of (not extreme) solutions follows a random schedule. These criteria aim to balance the search effort towards the different regions of objective space, increasing the discriminatory power of Pareto selection while simultaneously introducing a density estimator that scales-up well on high dimensional spaces, to find solutions with good convergence and diversity (spread and distribution) properties.

The number of rank-1 solutions $|\mathcal{F}_1^\epsilon|$ after reclassification depends on the value set to ϵ (≥ 0). Larger values of ϵ imply that sampled solutions ϵ-dominate larger areas, increasing the likelihood of having more ϵ-dominated solutions excluded from the sample that form $\mathcal{F}_1^\epsilon$. Adaptive ϵ-Ranking adapts ϵ at each generation so that $|\mathcal{F}_1^\epsilon|$ is close to the size of the parent population $|\mathcal{P}|$. The adaptation rule takes advantage of the correlation between ϵ and the number of ϵ-nondominated solutions in the sample. Basically, if $|\mathcal{F}_1^\epsilon| > |\mathcal{P}|$ it increases the step of adaptation $\Delta \leftarrow \min(\Delta \times 2, \Delta_{max})$ and $\epsilon \leftarrow \epsilon + \Delta$. Otherwise, if $|\mathcal{F}_1^\epsilon| < |\mathcal{P}|$ it decreases $\Delta \leftarrow \max(\Delta \times 0.5, \Delta_{min})$ and $\epsilon \leftarrow \max(\epsilon - \Delta, 0.0)$. The appropriate value of ϵ that approaches $|\mathcal{F}_1^\epsilon|$ to $|\mathcal{P}|$ is expected to change as the evolution process proceeds, it is problem dependent, and affected by the stochastic nature of the search that alters the instantaneous distributions of solutions in objective space. Adaptation of ϵ and its step of adaptation Δ is important to properly follow the dynamics of the evolutionary process on a given problem.

3. Test Problems

3·1 *Multi-objective MNK-Landscapes*

A multi-objective MNK-Landscape [Aguirre 07] is a vector function that maps binary strings into real numbers $\boldsymbol{f}(\cdot) = (f_1(\cdot), f_2(\cdot), \cdots, f_M(\cdot)) : \mathcal{B}^N \rightarrow \Re^M$, where M is the number of objectives, $f_i(\cdot)$ is the i-th objective function, $\mathcal{B} = \{0, 1\}$, and N is the bit string length. $\boldsymbol{K} = \{K_1, \cdots, K_M\}$ is a set of integers where K_i ($i = 1, 2, \cdots, M$) is the number of bits in the string that epistatically interact with each bit in the i-th landscape. Each $f_i(\cdot)$ can be expressed as an average of N functions as follows

$$f_i(\boldsymbol{x}) = \frac{1}{N}\sum_{j=1}^{N} f_{i,j}(x_j, z_1^{(i,j)}, z_2^{(i,j)}, \cdots, z_{K_i}^{(i,j)}) \tag{5}$$

where $f_{i,j} : \mathcal{B}^{K_i+1} \rightarrow \Re$ gives the fitness contribution of bit x_j to $f_i(\cdot)$, and $z_1^{(i,j)}, z_2^{(i,j)}, \cdots, z_{K_i}^{(i,j)}$ are the K_i bits interacting with bit x_j in the string $\boldsymbol{x}$. The fitness contribution $f_{i,j}$ of bit x_j is a number between [0.0, 1.0] drawn from a uniform distribution. Thus, each $f_i(\cdot)$ is a non-linear function of $\boldsymbol{x}$ expressed by a Kauffman's NK-Landscape model of epistatic interactions [Kauffman 93]. In addition, it is also possible to arrange the epistatic pattern between bit x_j and the K_i other interacting bits. That is, the distribution $D_i = \{random, nearest\ neighbor\}$ of K_i bits among N. Thus, M, N, $\boldsymbol{K} = \{K_1, K_2, \cdots, K_M\}$, and $\boldsymbol{D} = \{D_1, D_2, \cdots, D_M\}$, completely specify a multi-objective MNK-Landscape.

3·2 *DTLZ Problems*

To study the performance of the proposed algorithm on continuous functions with non-convex Pareto fronts, the functions DTLZ2, DTLZ3 and DTLZ4 of the DTLZ test functions family [Deb 02] are used. These functions are scalable in the number of objectives and variables and thus allow for a many-objective study. DTLZ2 has a non-convex Pareto-optimal surface that lies inside the first quadrant of the unit hyper-sphere. DTLZ3 and DTLZ4 are variations of DTLZ2. DTLZ3 introduces a large number of local Pareto-optimal fronts in order to test the convergence ability of the algorithm. DTLZ4 introduces biases on the density of solutions to some of the objective-space planes in order to test the ability of the algorithms to maintain a good distribution of solutions. For a detailed description of these problems the reader is referred to [Deb 02].

4. Method of Analysis

In this work, we use the hypervolume $\mathcal{H}$ and the set coverage $\mathcal{C}$ [Zitzler 99] to evaluate the performance of the algorithms, complementing our analysis with the maximum $\max(f_i)$ and minimum $\min(f_i)$ fitness values found in each objective. The measure $\mathcal{C}$ provides information on convergence. $\mathcal{C}(\mathcal{A},\mathcal{B})$ gives the fraction of solutions in set $\mathcal{B}$ that are dominated at least by one solution in set $\mathcal{A}$.

$\mathcal{H}$ is a measure of convergence and diversity, calculated as the volume of the M-dimensional region in objective space enclosed by the set of non-dominated solutions and a dominated reference point. If the reference point is very close to the Pareto front, non-dominated solutions around the center region of the Pareto front are relatively emphasized in the hypervolume calculation. On the other hand, if the reference point is far from the Pareto front, non-dominated solutions along the extreme regions of the Pareto front are emphasized in the hypervolume calculation. The hypervolume has become widely used to analyze the performance of multi-objective optimizers. However, results on the hypervolume are usually reported using a single reference point, which provides only a partial vision of the results obtained. In many-objective problems, particularly, it is difficult to grasp the convergence and diversity properties of the solutions obtained and reporting results using one reference point could often lead to overstated and sometimes erroneous conclusions about the overall performance of the algorithms. Analysis of hypervolume varying the reference point provides more insights on the distribution of the obtained solutions and helps clarify the relative contribution to the hypervolume of solutions that converge to the central regions of the space and those that contribute to diversity (spread). To enrich our analysis, we compute the hypervolume using different reference points. The reference point $\boldsymbol{r}_{d_R} = (r_1, r_2, \cdots, r_M)$ is calculated by

$$r_i = (1.0 - d_R) \times \min(f_i), i = 1, 2, \cdots, M, \quad (6)$$

where $\min(f_i)$ is the minimum value of the i-th objective function observed in the joined sets of Pareto optimal solutions found by the algorithms we compare, and d_R is a parameter that determines the distance of the reference point to the minimum values found for each objective function $(\min(f_1), \min(f_2), \cdots, \min(f_M))$. In this work, we use $d_R = \{0.01, 0.1, 0.3, 0.5, 0.7, 1.0\}$ to set the reference point $\boldsymbol{r}_{d_R} = (r_1, r_2, \cdots, r_M)$. Note that we maximize all objective functions and the allowed range for all f_i is $[0.0, 1.0]$. Hence, $d_R = 0.01$ means that the reference point is $\boldsymbol{r}_{d_R} = 0.99 \times (\min(f_1), \min(f_2), \cdots, \min(f_M))$ and thus very close to the Pareto front, whereas $d_R = 1.0$ means that $\boldsymbol{r}_{d_R} = (0.0, 0.0, \cdots, 0.0)$ and far from the Pareto front. To calculate $\mathcal{H}$, we use Fonseca et al. [Fonseca 06] algorithm.

5. Performance of the Proposed Hybrid Strategy on MNK-Landscapes

5·1 Preparation

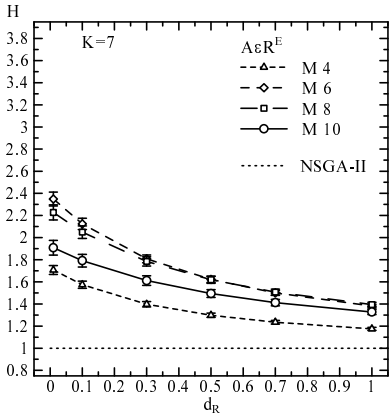
The performance of the algorithms is verified on MNK-Landscapes with $4 \leq M \leq 10$ objectives, $N = 100$ bits, number of epistatic interactions $K = \{0, 1, 3, 5, 7, 10, 15, 25, 35, 50\}$ ($K_1, \cdots, K_M = K$), and *random* epistatic patterns among bits in all objectives ($D_1, \cdots, D_M = random$). Results presented below show the average performance of the algorithms on 50 different problems randomly generated for each combination of M, N and K. In the plots, error bars show 95% confidence intervals on the mean.

In this work, we implement the proposed hybrid strategy using NSGA-II [Deb 00] as a host algorithm, modifying it accordingly to include the scalarization and $A\epsilon R^E$ strategies. During the first stage, selection is based solely on the scalarization function, whereas in the second stage $A\epsilon R^E$ is applied after Pareto dominance. All algorithms used in our study are set with parent and offspring populations of size $|\mathcal{P}| = |\mathcal{Q}| = 100$, two point crossover for recombination with rate $p_c = 0.6$, and bit flipping mutation with rate $p_m = 1/N$ per bit. The number of evaluations is set to 3×10^5 ($T = 3000$ generations). In $A\epsilon R^E$, initially $\epsilon = 0.0$, the initial value of the step of adaptation is $\Delta_0 = 0.005$ (0.5%) and its maximum and minimum values are set to $\Delta_{max} = 0.05$ (5%) and $\Delta_{min} = 0.0001$ (0.01%).

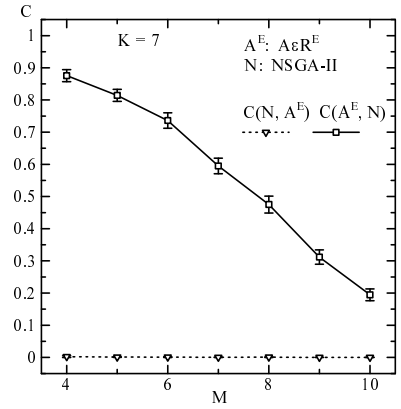
5·2 Effects of Individual Components

§ 1 Adaptive ϵ-Ranking Strategy

In this section we discuss the performance of Adaptive ϵ-Ranking Strategy ($A\epsilon R^E$) using NSGA-II as a reference for comparison. **Figure 2(a)** shows the normalized hypervolume $\mathcal{H}$ between $A\epsilon R^E$ and NSGA-II varying the reference point for $K = 7$ and $M = \{4, 6, 8, 10\}$ landscapes. From this figure it can be seen that $A\epsilon R^E$ attains better $\mathcal{H}$ for all values of M regardless of the reference point. Also, note the increasing slopes of the $\mathcal{H}$ curves as the reference point gets closer to the Pareto front, i.e. varying d_R from 1.0 to 0.01. These results suggest that solutions by $A\epsilon R^E$ are better than solutions by NSGA-II particularly in the central regions of the objective space. Notice that the slope of the $\mathcal{H}$ curve becomes steeper by increasing the number of objectives from $M = 4$ to $M = 6$, but

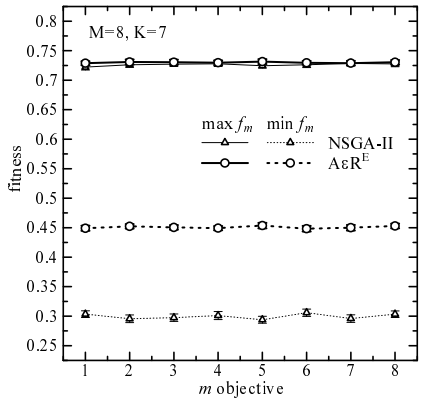


(a) Normalized $\mathcal{H}$

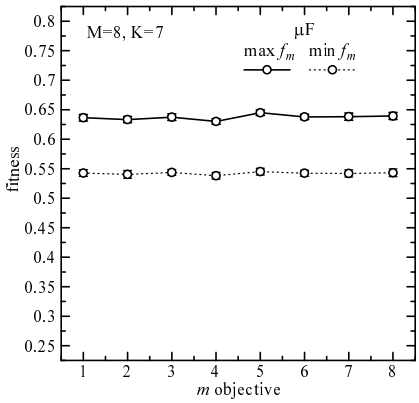


(b) Set coverage $\mathcal{C}$

Fig. 2 Adaptive ϵ-Ranking, $K = 7$, $4 \leq M \leq 10$.

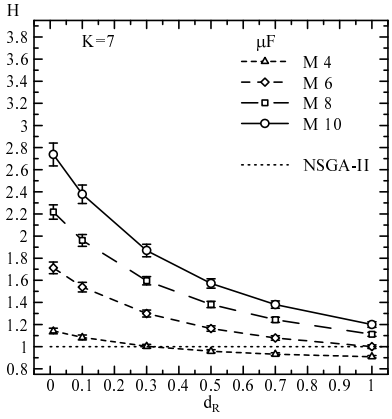


(a) Adaptive ϵ-Ranking

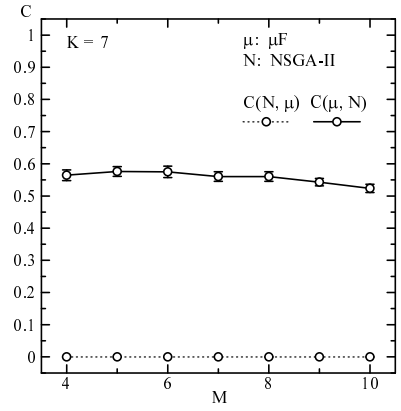


(b) Scalarization strategy

Fig. 3 $\max(f_m)$ and $\min(f_m)$, $K = 7$.



(a) Normalized $\mathcal{H}$



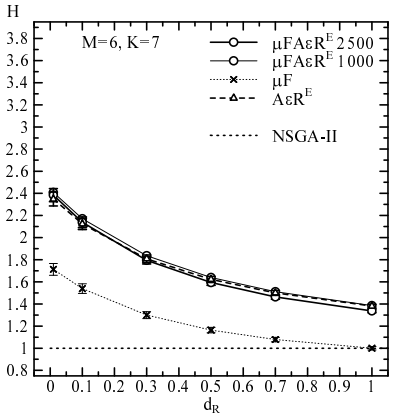
(b) Set coverage $\mathcal{C}$

Fig. 4 Scalarization strategy, $K = 7$ and $4 \leq M \leq 10$.

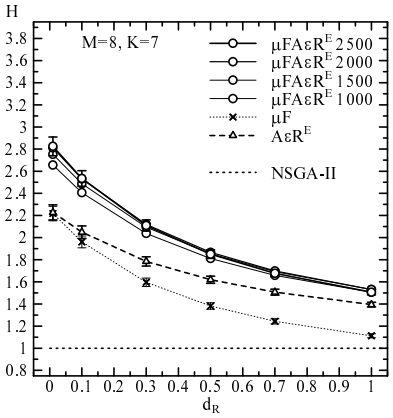
it gradually recedes for $M = 8$ and $M = 10$ compared to $M = 6$. This is an indication that the convergence abilities of $A\epsilon R^E$ reduce for $M > 6$, especially to the central regions of objective space.

Figure 2(b) shows results using the $\mathcal{C}$ coverage measure. Note that $\mathcal{C}(A^E, N)$, the fraction of NSGA-II's solutions dominated by $A\epsilon R^E$'s solutions, is almost 0.9 for $M = 4$ and reduces progressively with M until it approaches 0.2 for $M = 10$. On the contrary, $\mathcal{C}(N, A^E)$ is zero for all M, which means that no solution by $A\epsilon R^E$ is dominated by NSGA-II's solutions. These results confirm the superiority of $A\epsilon R^E$ over NSGA-II and corroborate the decreasing convergence power of $A\epsilon R^E$ for large values of M.

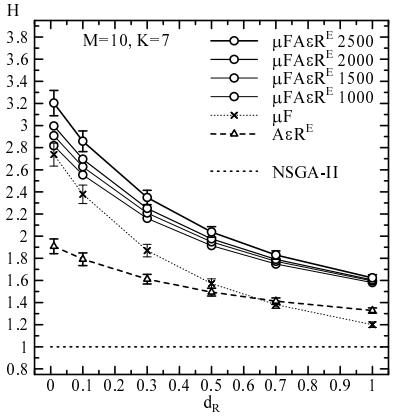
Figure 3(a) shows the maximum and minimum fitness, $\max(f_m)$ and $\min(f_m)$, of solutions in the Pareto front found by NSGA-II and $A\epsilon R^E$ for $K = 7$ and $M = 8$ landscapes. From this figure, it can be seen that NSGA-II and $A\epsilon R^E$ achieve similar $\max(f_m)$. However, $\min(f_m)$ is lower by NSGA-II. Similar values of $\max(f_m)$ sug-



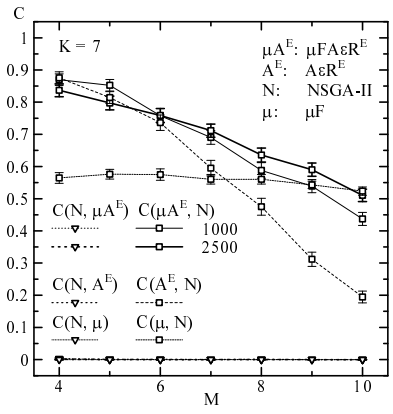
(a) Normalized $\mathcal{H}$, $M = 6$



(b) Normalized $\mathcal{H}$, $M = 8$



(c) Normalized $\mathcal{H}$, $M = 10$



(d) Set coverage $\mathcal{C}$

Fig. 5 Proposed Hybrid Strategy, $K = 7$ and $4 \leq M \leq 10$.

gest that spread by the algorithms is comparable, but the lower values of $\min(f_m)$ suggest that solutions by NSGA-II seem to be trapped in lower local optima. Another interesting property of $A\epsilon R^E$ is that solutions in the Pareto front are ϵ-nondominated, which gives a good distribution of solutions.

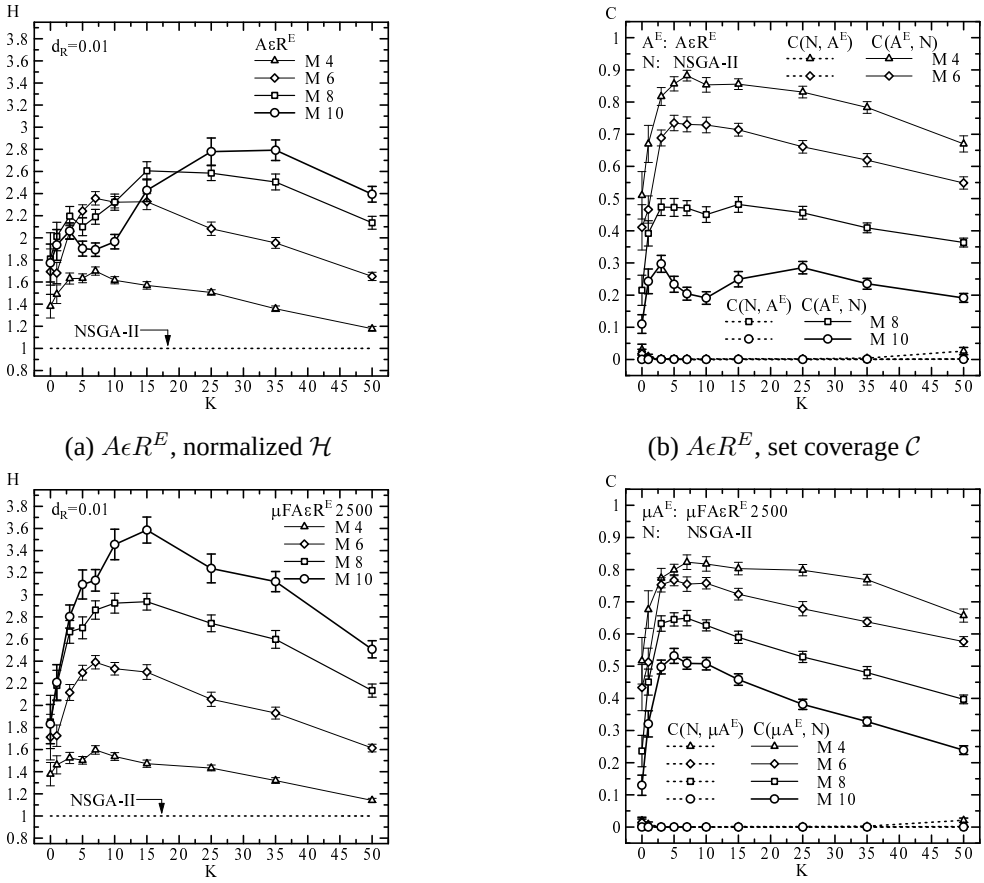
§ 2 Scalarization Strategy

In this section we analyze the scalarization strategy (μF). **Figure 4(a)** shows the normalized $\mathcal{H}$ between μF and NSGA-II varying the reference point on $K = 7$ and $M = \{4, 6, 8, 10\}$ landscapes. Comparing $\mathcal{H}$ by looking at **Figure 4(a)** and **Figure 2(a)**, it can be seen that on $M = \{4, 6\}$ landscapes μF is significantly worse than $A\epsilon R^E$ for any value of d_R. On $M = 8$ landscapes, μF is still worse than $A\epsilon R^E$ for $d_R \geq 0.1$, but similar to $A\epsilon R^E$ for $d_R = 0.01$. However, on $M = 10$ landscapes, μF is better than $A\epsilon R^E$ for $d_R < 0.5$. These results suggest that μF gets better compared to $A\epsilon R^E$ in terms of convergence to central regions when the number of objectives is above eight.

Figure 4(b) shows $\mathcal{C}$ between μF and NSGA-II. Note that for any number of objectives, the values of $\mathcal{C}(\mu F, N)$ are similar and above 0.5, meaning that more than half of the solutions by NSGA-II are dominated by solutions of μF. Comparing with **Figure 2(b)**, note that $\mathcal{C}(\mu, N) < \mathcal{C}(A^E, N)$ for $M \leq 7$, but $\mathcal{C}(\mu, N) > \mathcal{C}(A^E, N)$ for $M \geq 8$. These results are in accordance with the observations made for $\mathcal{H}$, and confirm the better convergence properties of μF on landscapes with more than eight objectives. However, note that μF converges to a narrow area, as shown in **Figure 3(b)** that plots the $\max(f_m)$ and $\min(f_m)$ of the non-dominated set found by μF. Overall, these results shows that the scalarization strategy μF converges well, albeit to a narrow region. The similar values of $\mathcal{C}(\mu F, N)$ for all M is an interesting property of μF. It shows that this strategy in terms of convergence can scale up to a large number of objectives, suggesting that it could be useful as part of the hybrid strategy.

5·3 Effects of the Hybrid Strategy

In this section we analyze the hybrid strategy that combines in a two-stage process scalarization and Adaptive ϵ-Ranking ($\mu F A\epsilon R^E$). $\mu F A\epsilon R^E$ first starts with μF and then at generation t_S it switches to $A\epsilon R^E$. **Figure 5 (a)-(c)** show $\mathcal{H}$ by $\mu F A\epsilon R^E$ on $M = \{6, 8, 10\}$ landscapes, respectively, varying $t_S = \{1000, 1500, 2000, 2500\}$ and keeping the total number of the generations fixed to $T =$



(a) $A\epsilon R^E$, normalized $\mathcal{H}$ (b) $A\epsilon R^E$, set coverage $\mathcal{C}$

(c) $\mu FA\epsilon R^E$, normalized $\mathcal{H}$ (d) $\mu FA\epsilon R^E$, set coverage $\mathcal{C}$

Fig. 6 Adaptive ϵ-Ranking $A\epsilon R^E$ and Proposed Hybrid Strategy $\mu FA\epsilon R^E$, $0 \leq K \leq 50$ and $4 \leq M \leq 10$.

3000. The same figures also include results by μF and $A\epsilon R^E$ for comparison. Note that on $M = 6$ the inclusion of μF does not improve $\mathcal{H}$ (actually, on $M = 4$ for which results are not shown, $\mathcal{H}$ reduces by including μF, with larger reductions observed for late switching times t_S). However, switching from μF to $A\epsilon R^E$ during the run can improve $\mathcal{H}$ substantially ($d_R = 0.01$) on $M = 8$ and $M = 10$, with a late switching time ($t_S = 2500$) working better than an early one.

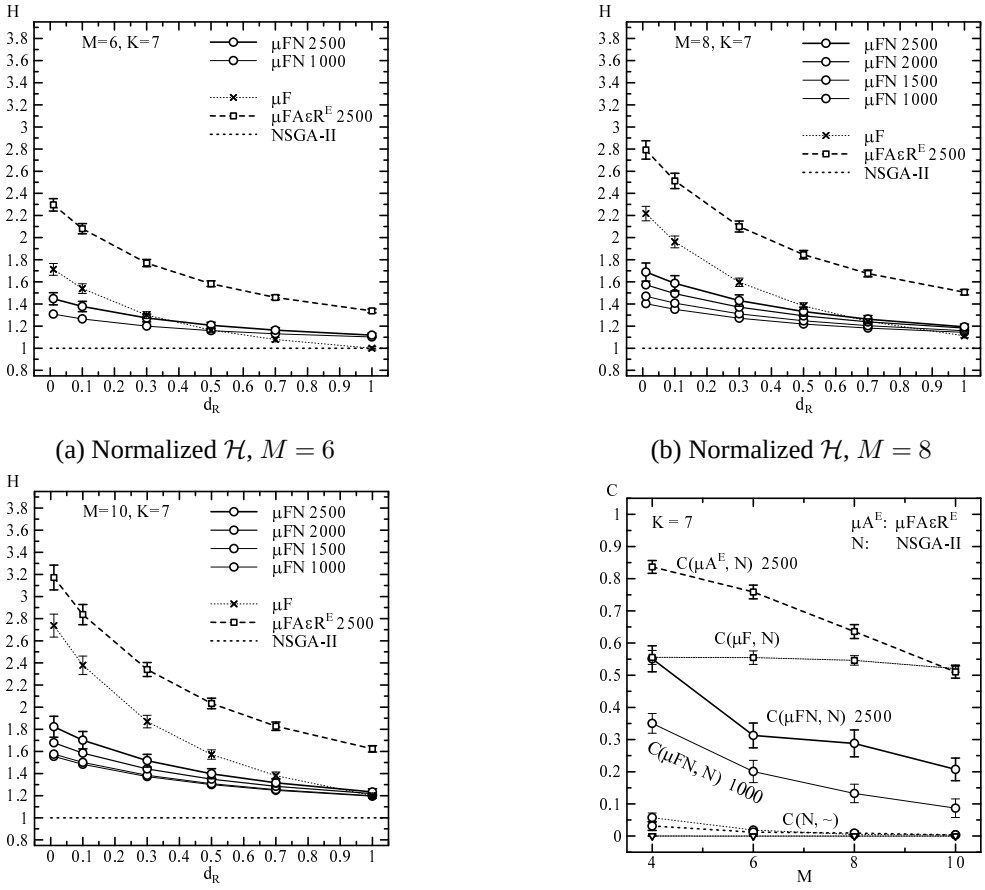
Figure 5 (d) shows $\mathcal{C}$ values between $\mu FA\epsilon R^E$ and NSGA-II for $t_S = \{1000, 2500\}$. We also include results by μF and $A\epsilon R^E$ for comparison. Results on $\mathcal{C}$ confirm our observations on $\mathcal{H}$ and give a clearer picture of the effects of including μF. Note that $\mu FA\epsilon R^E$ shows relatively better convergence than $A\epsilon R^E$ and the importance of late switching times t_S for $M \geq 7$. Also, it can be seen that convergence is better than μF for $M \leq 9$ if $t_S = 2500$. For $M = 10$ similar convergence to μF is observed. However, $\mu FA\epsilon R^E$ $t_S = 2500$ shows significantly better $\mathcal{H}$ than μF on $M = 10$ as shown in **Figure 5 (c)**.

Figure 6 show results by $A\epsilon R^E$ and $\mu FA\epsilon R^E$ ($t_S = 2500$), varying K from 0 to 50 to observe the scalability of the algorithms on problems of increased epistasis. From these figures, note that similar to $K = 7$, on $M = 4$ performance deteriorates slightly by including μF, especially in terms of convergence, whereas on $M = 6$ $\mathcal{H}$ and $\mathcal{C}$ are similar by both algorithms. On the other hand, on $M = 8$ and $M = 10$ the inclusion of μF leads to better $\mathcal{H}$ and $\mathcal{C}$ on a broad range of K ($K \geq 1$). Since μF is just one scalarazing function, its computation is faster than Pareto ranking based approaches. Thus, the hybrid strategy $\mu FA\epsilon R^E$ ($t_S = 2500$) is also substantially faster than $A\epsilon R^E$, which becomes relevant for scalability on high dimensional spaces.

6. Scalarization and NSGA-II

In this section we analyze a two-stage method using the scalarization strategy μF in the first stage and NSGA-II's non-domination sorting & crowding distance strategy in the second stage. We call this two-stage method μFN for short. By comparing the performance of μFN and the proposed $\mu FA\epsilon R^E$, we aim to clarify the necessity of Adaptive ϵ-Ranking $A\epsilon R^E$.

Figure 7 shows results by μFN varying the switching time between stages $t_S = \{1000, 1500, 2000, 2500\}$ on land-



(a) Normalized $\mathcal{H}$, $M = 6$

(b) Normalized $\mathcal{H}$, $M = 8$

(c) Normalized $\mathcal{H}$, $M = 10$

(d) Set coverage $\mathcal{C}$

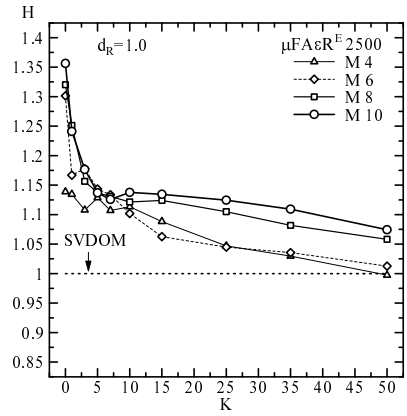
Fig. 7 Two-stage Strategy using Scalarization and NSGA-II's Non-domination Sorting & Crowding Distance, $K = 7$ and $4 \leq M \leq 10$.

scapes with $M = \{4, 6, 8, 10\}$ objectives and $K = 7$ epistatic interactions. We also include results by NSGA-II, scalarization μF alone, and the proposed hybrid $\mu F A\epsilon R^E$ ($t_S = 2500$) for comparison.

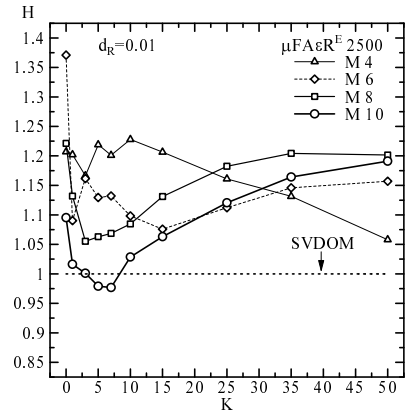
From **Figure 7 (a)-(c)** note that the two-stage method μFN achieves better hypervolume $\mathcal{H}$ than NSGA-II and that later switching times work better than early ones. However, hypervolume by μFN is considerably lower than by μF alone, especially for reference points close to the Pareto front. These results on hypervolume contrast sharply with those achieve by the hybrid $\mu F A\epsilon R^E$, which are considerably better than μF as shown in the same figures. From **Figure 7 (d)**, a similar conclusion can be reached on the set coverage $\mathcal{C}$. Note that $\mathcal{C}(\mu FN, N)$ is in the range [0.1,0.35] for $t_S = 1000$ and [0.2,0.55] for $t_S = 2500$, which means that μFN dominates NSGA-II's solutions. However, $\mathcal{C}(\mu F, N)$ and $\mathcal{C}(\mu FA^E, N)$ are superior. (NSGA-II does not dominate solutions by other algorithms, labeled as $\mathcal{C}(N, \sim)$ in the figure). These results suggest that the second stage based on NSGA-II's non-domination sorting & crowding distance cannot take advantage of the well-converged solutions found by μF in the first stage of the search, whereas Adaptive ϵ-Ranking $A\epsilon R^E$ adds to the good performance of the scalarization strategy μF further enhancing convergence, spread, and distribution of solutions.

7. Comparison with Substitute Assignment Distance Methods

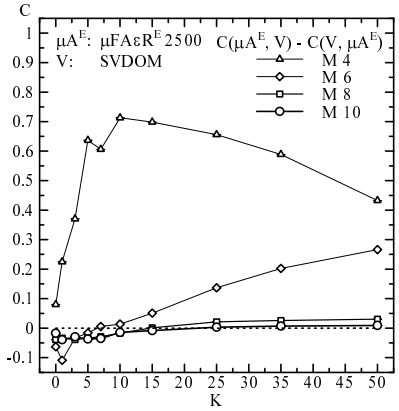
In this section we compare performance between the hybrid strategy $\mu F A\epsilon R^E$ and two substitute distance assignment methods, namely Subvector Dominance Assignment (SVDOM) and Epsilon Dominance Assignment (EPSDOM) [Koppen 07]. Similar to the hybrid strategy, SVDOM and EPSDOM were initially proposed using the NSGA-II framework. SVDOM and EPSDOM keep Pareto dominance as the primary ranking of solutions, but replace the diversity estimator with a substitute assignment distance to assign the secondary ranking of solutions favoring convergence exclusively. Our motivation is to understand the effectiveness of the hybrid strategy that seeks to balance convergence, spread, and distribution against two highly effective strategies such SVDOM and EPSDOM that focus on



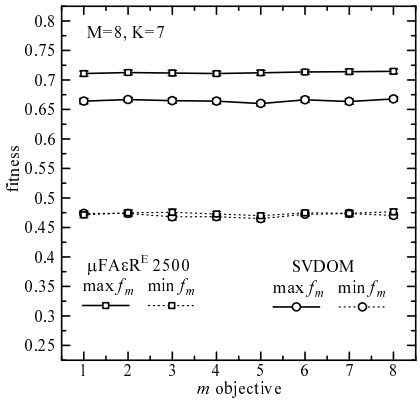
(a) Normalized $\mathcal{H}$, $d_R = 1.0$



(b) Normalized $\mathcal{H}$, $d_R = 0.01$



(c) Set coverage $\mathcal{C}$ (difference)



(d) $\max(f_m)$ and $\min(f_m)$, $M = 8$ $K = 7$

Fig. 8 Proposed Hybrid Strategy $\mu FA\epsilon R^E$ 2500 and SVDOM, $0 \leq K \leq 50$ and $4 \leq M \leq 10$.

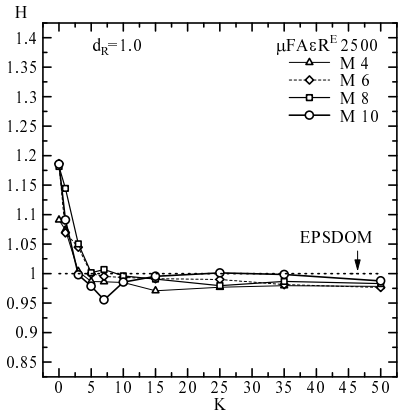
convergence exclusively [Aguirre 10a].

The substitute distance that determines the secondary ranking of solutions in SVDOM and EPSDOM are based on measurement procedures that calculate the highest degree to which a solution is nearly Pareto dominated by any other non-dominated solution [Koppen 07] (closeness to dominance). SVDOM re-ranks a non-dominated solution based on the number of objectives that are better in other non-dominated solutions. On the other hand, EPSDOM re-ranks a non-dominated solution based on the magnitude that other non-dominated solutions need to improve in order to dominate it. The computational order of calculating the secondary ranking by these methods is $\mathcal{O}(M|P|^2)$, in addition to the $\mathcal{O}(M|P|^2)$ order needed to calculate Pareto dominance. The reader is referred to [Koppen 07] for details on SVDOM and EPSDOM.

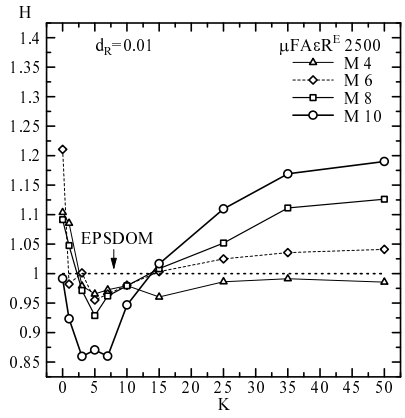
Figure 8 (a)-(d) show the normalized hypervolume between $\mu FA\epsilon R^E$ 2500 and SVDOM for $d_R = 1.0$ (reference point far away from the Pareto front) and $d_R = 0.01$ (reference point close to the Pareto front), the differential cover set $\mathcal{C}$ value, and $\max(f_i)$ and $\min(f_i)$ in all objectives, respectively. Similarly, **Figure 9 (a)-(d)** show results for $\mu FA\epsilon R^E$ and EPSDOM.

Firstly, we analyze results by $\mu FA\epsilon R^E$ and SVDOM. From **Figure 8 (a)** and **(b)** it can be seen that overall on all M the hypervolume by $\mu FA\epsilon R^E$ is better than by SVDOM. It should be highlighted that **Figure 8 (a)** shows the hypervolume calculated setting a reference point far away from the Pareto front, which emphasizes the contribution of solutions located along the extreme regions of the Pareto front. On the other hand, in **Figure 8 (b)** a reference point close to the Pareto front is used to emphasize the contribution of solutions located in central regions of the Pareto front. From **Figure 8 (c)** note that $\mathcal{C}(\mu A^E, V)$ is larger than $\mathcal{C}(V, \mu A^E)$ on $M = 4$ and $M = 6$ for most K, whereas similar values are observed on $M = 8$ and $M = 10$. Looking at **Figure 8 (d)** note that $\max(f_i)$ by SVDOM is smaller in all objectives than by the hybrid strategy $\mu FA\epsilon R^E$. These results suggest that overall the hybrid strategy can find better solutions than SVDOM in a broader region of objective space.

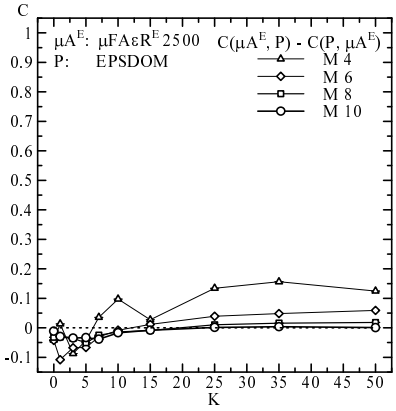
Next, we analyze results by $\mu FA\epsilon R^E$ and EPSDOM. From **Figure 9 (a)** it can be seen that for very small K on all M the hypervolume by $\mu FA\epsilon R^E$ is better than by EPSDOM, whereas for larger K similar hypervolume is achieved by both algorithms. From **Figure 9 (b)**, where



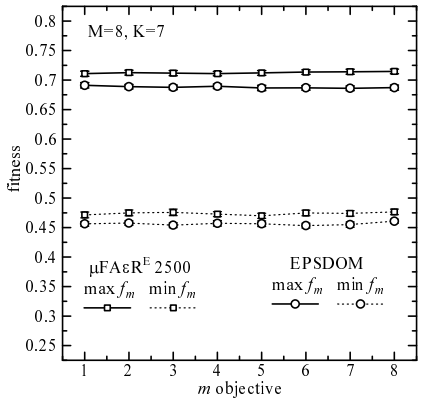
(a) Normalized $\mathcal{H}$, $d_R = 1.0$



(b) Normalized $\mathcal{H}$, $d_R = 0.01$



(c) Set coverage $\mathcal{C}$ (difference)



(d) $\max(f_m)$ and $\min(f_m)$, $M = 8$ $K = 7$

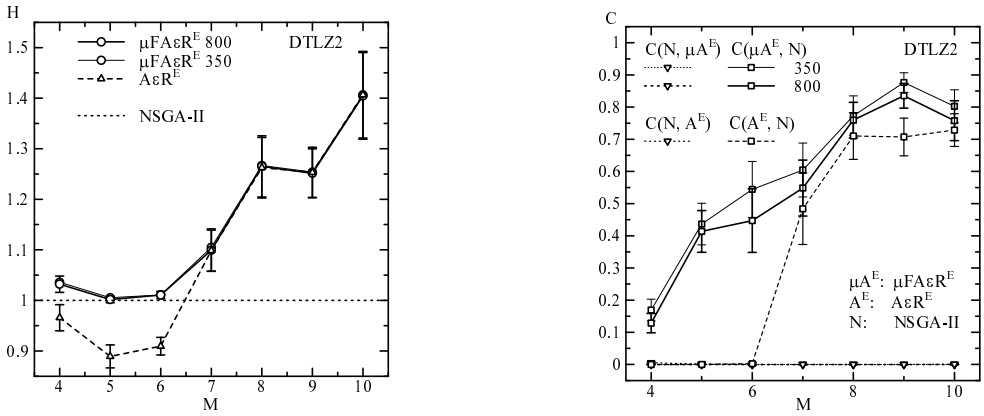
Fig. 9 Proposed Hybrid Strategy $\mu FA\epsilon R^E$ 2500 and EPSDOM, $0 \leq K \leq 50$ and $4 \leq M \leq 10$.

solutions in the central regions are relatively emphasized, it can be seen that EPSDOM achieves better hypervolume than $\mu FA\epsilon R^E$ for small values of K ($1 \leq K \leq 10$), especially for $M = 10$. On the other hand, $\mu FA\epsilon R^E$ achieves better hypervolume than EPSDOM for medium and large K and $M \geq 6$. Note in **Figure 8 (b)** that SVDOM's best performance is also for small K. These results suggest that information of closeness to dominance is more effective to improve convergence on problems of moderate complexity than on highly complex ones. From **Figure 9 (b)** it should also be noted that for highly complex problems (large K) the hybrid approach performs very well, which is interesting because non-convex regions in the Pareto front increase with K [Aguirre 07] and are supposed to hinder performance of approaches that include scalarization by weighting sum strategies. From **Figure 9 (c)** note that $\mathcal{C}(\mu A^E, V)$ is larger than $\mathcal{C}(V, \mu A^E)$ on $M = 4$ and $M = 6$ especially for large K, whereas similar values are observed on $M \geq 6$. Note also that $\mathcal{C}(V, \mu A^E)$ is slightly better on $K \leq 7$. Looking at **Figure 9 (d)** note that $\max(f_i)$ by EPSDOM is smaller in all objectives than by the hybrid strategy $\mu FA\epsilon R^E$, albeit better than SVDOM. These results suggest that for medium and large K the hybrid strategy can find similar or better solutions than EPSDOM in a broader region of objective space. For small K there is a slight advantage by EPSDOM in terms of convergence to central regions, especially for $M = 10$ objectives.

8. Performance of the Proposed Hybrid Strategy on DTLZ Problems

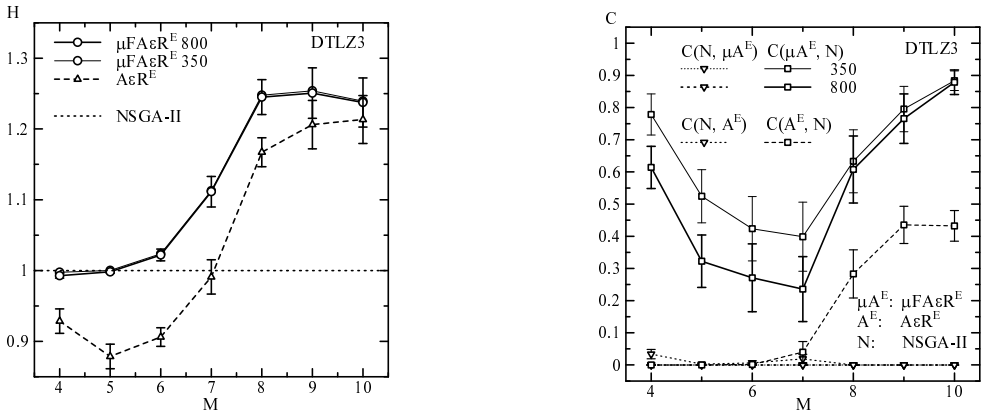
In previous sections, we have studied the hybrid strategy using MNK-Landscapes, showing that it can significantly improve performance on many-objective non-linear combinatorial optimization problems. These problems are known to have convex fronts for low values of non-linearity K, but non-convex regions in the fronts and convergence-difficulty increase with the non-linearity of the problem.

In this section, we study the hybrid strategy on many-objective continuous problems with large number of variables and non-convex Pareto fronts, using instances with different characteristics of convergence-difficulty and bias on density of solutions. Our aim is to verify whether a simple scalarization strategy could be helpful as part of the two-stage hybrid approach on non-convex problems, where simple scalarization functions alone are known not



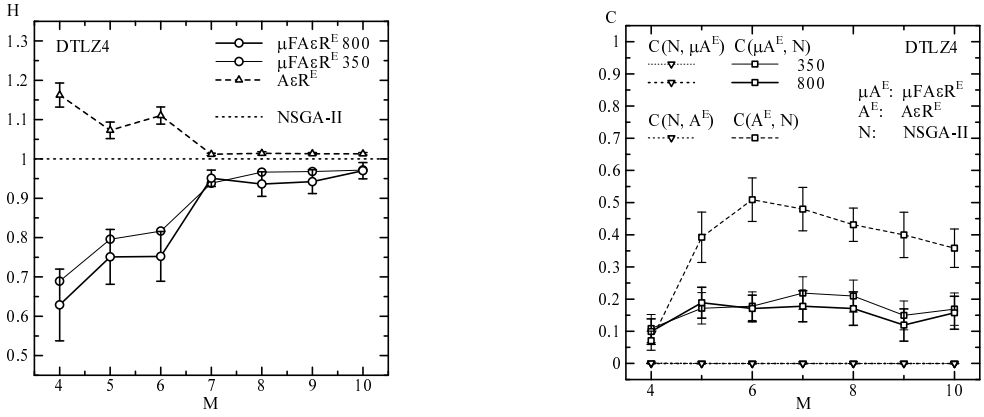
(a) Normalized $\mathcal{H}$ (b) Set coverage $\mathcal{C}$

Fig. 10 NSGA-II, Adaptive ϵ-Ranking $A\epsilon R^E$, and hybrid method $\mu FA\epsilon R^E$ on problem DTLZ2, $4 \leq M \leq 10$.



(a) Normalized $\mathcal{H}$ (b) Set coverage $\mathcal{C}$

Fig. 11 NSGA-II, Adaptive ϵ-Ranking $A\epsilon R^E$, and hybrid method $\mu FA\epsilon R^E$ on problem DTLZ3, $4 \leq M \leq 10$.



(a) Normalized $\mathcal{H}$ (b) Set coverage $\mathcal{C}$

Fig. 12 NSGA-II, Adaptive ϵ-Ranking $A\epsilon R^E$, and hybrid method $\mu FA\epsilon R^E$ on problem DTLZ4, $4 \leq M \leq 10$.

to perform well. Also, we want to verify the extent to which difficulty on convergence and bias on distribution of solutions affect the scalarization strategy on continuous domains.

The performance of the algorithms is verified on DTLZ2, DTLZ3, and DTLZ4 continuous functions [Deb 02], setting the total number of variables to 100 and varying the number of objectives M from 4 to 10. Results below show the average performance of the algorithms on 50 different runs. In the plots, error bars show 95% confidence intervals on the mean.

Below we discuss the performance of Adaptive ϵ-Ranking $A\epsilon R^E$ and the hybrid strategy $\mu FA\epsilon R^E$ using NSGA-II as a reference for comparison. We study the hybrid strategy varying the strategies' switching time, $t_S = \{350, 500, 650, 800\}$, and keeping the total number of the generations

fixed to $T = 1000$. In the following, for the sake of clarity we only present results for $t_S = \{350, 800\}$. All algorithms are set with parent and offspring population of size 100, Simulated Binary Crossover (SBX) for recombination with rate $p_c = 1$ per individual, and polynomial mutation (PM) with rate $1/n$ per decision variable. Initial settings for $A\epsilon R^E$ are the same used for MNK-Landscapes, as indicated in Section 5·1. The reference point to calculate the hypervolume is set to $(1.01 \times \max(f_1), \cdots, 1.01 \times \max(f_M))$, computing $\max(f_m)$, $m = 1, \cdots, M$, from the solutions generated in all runs by the algorithms. Note that DTLZ functions are minimization problems.

Firstly, we analyze results for function DTLZ2, which has a non-convex Pareto-optimal surface that lies inside the first quadrant of the unit hyper-sphere. **Figure 10(a)** and **(b)** show the normalized hypervolume $\mathcal{H}$ and set coverage measure $\mathcal{C}$, respectively. Looking at $A\epsilon R^E$ and NSGA-II, from these figures note that for $M = 6$ objectives or less $A\epsilon R^E$ is worse than NSGA-II on $\mathcal{H}$, but similar on $\mathcal{C}$, $\mathcal{C}(N, A^E) = 0$ and $\mathcal{C}(A^E, N) = 0$. On the other hand, $A\epsilon R^E$ is significantly better than NSGA-II for more than 6 objectives, both on $\mathcal{H}$ and $\mathcal{C}$, $\mathcal{C}(N, A^E) = 0$ and $\mathcal{C}(A^E, N) > 0.45$. These results suggest that NSGA-II achieves better spread that improves hypervolume for $M \leq 6$, but it cannot find solutions that dominate $A\epsilon R^E$'s solutions. Increasing M above 6, $A\epsilon R^E$ achieves better spread and convergence than NSGA-II. Next, we look at results by the hybrid strategy $\mu F A\epsilon R^E$. From the same **Figure 10**, note that $\mu F A\epsilon R^E$ improves $\mathcal{H}$ compared to $A\epsilon R^E$ for $M < 6$, and also improves $\mathcal{C}$ for all values of M. These results show that the inclusion of the simple scalarization strategy μF helps convergence of the algorithm. It is also worth mentioning that the switching time t_S used in $\mu F A\epsilon R^E$ leads to slightly different $\mathcal{C}$ values for $M > 6$, but those differences are not captured by $\mathcal{H}$. This is because the reference point used to calculate $\mathcal{H}$ is located relatively far away from the Pareto front of $\mu F A\epsilon R^E$, in which case differences in convergence are difficult to detect with the $\mathcal{H}$ measure.

Secondly, we analyze results for DTLZ3, a variation of DTLZ2 that introduces a large number of local Pareto-optimal fronts in order to test the convergence ability of the algorithm. Results for DTLZ3 are shown in **Figure 11(a)** and **(b)**. From these figures note that the inclusion of the scalarization function μF improves $\mathcal{H}$ and $\mathcal{C}$ for all values of M. These results suggest that on non-convex continuous problems of increased convergence-di-fficulty, the introduction of the scalarization strategy μF in the hybrid approach works effectively to improve the performance of the algorithm. Similar to DTLZ2, the effects of the switching time t_S can be observed on $\mathcal{C}$ but not on $\mathcal{H}$. Notice that better values of $\mathcal{C}$ are observed for the early switching time. This is because μF focuses in an increasingly narrower area of objective space for larger t_S.

Thirdly, we analyze results for DTLZ4, a problem with a biased density of solutions to some of the objective space planes, in order to test the ability of the algorithms to maintain a good distribution of solutions. Results are shown in **Figure 12(a)** and **(b)**. Note that $A\epsilon R^E$ performs better than NSGA-II. However, the introduction of the scalarization function μF that focuses on convergence-only deteriorates performance of the hybrid approach $\mu F A\epsilon R^E$, compared to $A\epsilon R^E$. These results suggest that on problems with highly biased distribution of solutions simple strategies that favor convergence without paying attention to diversity could mislead the algorithm.

9. Conclusions

We have shown that a two-stage hybrid strategy that uses scalarization and Pareto dominance enhanced with Adaptive ϵ-Ranking can significantly improve performance on many-objective MNK-Landscapes, especially for a large number of objectives. Also, we showed that it is feasible to simplify the scalarization strategy, which could reduce overall computational cost. In addition, we compared the hybrid strategy with Subvector Dominance Assignment and Epsilon Dominance Assignment, two highly effective methods to improve convergence on many-objective problems, showing that the hybrid approach can find solutions with similar or better convergence properties on highly complex MNK-Landscapes, while achieving better spread and distribution. We also showed that the proposed hybrid strategy can significantly improve performance on many-objective continuous functions with non-convex Pareto fronts, especially on problems having a large number of local fronts that increase the difficulty to converge. However, the simple scalarization function can be misleading on problems with large bias on distribution of solutions, in which case Adaptive ϵ-Ranking alone is superior. In the future, we would like to look into adaptive strategies to switch between stages avoiding the negative effects of large biases on distribution of solutions, and try other scenarios to deploy the individual strategies.

Acknowledgments

This study was performed through Special Coordination Funds for Promoting Science and Technology of the Ministry of Education, Culture, Sports, Science and Technology, of the Japanese Government.

♢ References ♢

[Aguirre 07] Aguirre, H., and Tanaka, K.: “Working Principles, Behavior, and Performance of MOEAs on MNK-Landscapes”, *European Journal of Operational Research*, Elsevier, vol. 181(3), pp. 1670-1690 (2007)

[Aguirre 09] Aguirre, H., and Tanaka, K.: “Adaptive ϵ-Ranking on MNK-Landscapes”, *Proc.2009 IEEE Symposium on Computational Intelligence in Multicriteria Decision Making*, IEEE Service Center, pp.104-111 (2009)

[Aguirre 10a] Aguirre, H., and Tanaka, K.: “Space Partitioning Evolutionary Many-Objective Optimization: Performance Analysis on MNK-Landscapes”, *Transactions of the Japanese Society for Artificial Intelligence*, vol. 25, no. 2, pp.366-373 (2010)

[Aguirre 10b] Aguirre, H., and Tanaka, K.: “A Hybrid Scalarization and Adaptive ϵ-Ranking Strategy for Many-objective Optimization”, *Parallel Problem Solving from Nature XI*, Part II, Springer, LNCS, vol. 6239, pp. 11-20 (2010)

[Aguirre 10c] Aguirre, H., and Tanaka, K.: “Performance Study of a Two-stage Hybrid Method for Many-objective Optimization on Non-convex Continuous Problems”, *Proc. International Symposium on Intelligent Systems iFAN 2010*, 4-pages(2010)

[Brockhoff 07] Brockhoff, D., and Zitzler, E.: “Are All Objectives Necessary? On Dimensionality Reduction in Evolutionary Multiobjective Optimization”, *Parallel Problem Solving from Nature IX*, Springer, LNCS, vol. 4193, pp. 533-542 (2006)

[Coello 02] Coello, C., Van Veldhuizen, D., and Lamont, G.: *Evolutionary Algorithms for Solving Multi-Objective Problems*. Kluwer Academic Publishers, Boston (2002)

[Corne 07] Corne, D., and Knowles, J.: “Techniques for Highly Multiobjective Optimization: Some Non-dominated Points are Better than Others”, *Proc. 2007 Genetic and Evolutionary Computation Conference”*, pp.773-780 (2007)

[Deb 00] Deb, K., Agrawal, S., Pratap, A., and Meyarivan, T.: “A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II”, *KanGAL report 200001* (2000)

[Deb 01] Deb, K.: *Multi-Objective Optimization using Evolutionary Algorithms*, John Wiley & Sons, Chichester, West Sussex, England (2001)

[Deb 02] Deb, K., Thiele, L., Laumanns, M., and Zitzler, E.: “Scalable Multi-Objective Optimization Test Problems”, *Proc. 2002 Congress on Evolutionary Computation*, IEEE Service Center, pp.825–830 (2002)

[Deb 06a] Deb, K., and Sundar,J.: “Preference Point Based Multiobjective Optimization Using Evolutionary Algorithms”, *Proc. 2006 Genetic and Evolutionary Computation Conference*, pp. 635-642 (2006)

[Deb 06b] Deb, K., and Saxena, K.: “Searching for Pareto-optimal Solutions Through Dimensionality Reduction for Certain Large-dimensional Multi-objective Optimization Problems”, *Proc. 2006 IEEE Congress on Evolutionary Computation*, pp. 3353-3360 (2006)

[Emmerich 05] Emmerich, M., Beume, N., and Naujoks, B.: “An EMO Algorithm Using the Hypervolume Measure as Selection Criterion”, *Proc. 3rd Intl. Conf. on Evolutionary Multi-Criterion Optimization*, Springer, LNCS, vol. 3410, pp. 62-76 (2005)

[Fonseca 06] Fonseca, C., Paquete, L., and López-Ibáñez, M.: “An Improved Dimension-sweep Algorithm for the Hypervolume Indicator”, *Proc. 2006 IEEE Congress on Evolutionary Computation*, IEEE Service Center, pp.1157-1163 (2006)

[Hughes 05] Hughes, E. J.: “Evolutionary Many-Objective Optimization: Many Once or One Many?”, *Proc. 2005 IEEE Congress on Evolutionary Computation*, IEEE Service Center, Vol.1, pp.222-227 (2005)

[Ishibuchi 98] Ishibuchi, H., Murata,T.: “A Multi-Objective Genetic Local Search Algorithm and Its Applications to Flowshop Scheduling”, *IEEE Trans. on Systems, Man, and Cybernetics - Part C: Applications and Reviews*, vol. 28, pp. 392-403 (1998)

[Ishibuchi 07a] Ishibuchi, H., Tsukamoto, N., and Nojima, Y.: “Iterative Approach to Indicator-based Multiobjective Optimization”, *Proc. 2007 IEEE Congress on Evolutionary Computation*, pp. 3697-3704 (2007)

[Ishibuchi 07b] Ishibuchi, H., and Nojima, Y.: “Optimization of Scalarizing Functions Through Evolutionary Multiobjective Optimization”, *Proc. Fourth Intl. Conf. on Evolutionary Multi-Criterion Optimization*, Springer, LNCS, vol. 4403, pp. 51-65 (2007)

[Ishibuchi 08] Ishibuchi, H., Tsukamoto, N., and Nojima, Y.: “Evolutionary Many-Objective Optimization: A Short Review”, *Proc. IEEE Congress on Evolutionary Computation*, IEEE Service Center, pp.2424-2431 (2008)

[Jaskiewics 02] Jaskiewics, A.: “Genetic Local Search for Multi-Objective Combinatorial Optimization”, *European Journal of Operational Research*, vol. 137, pp.50-71 (2002)

[Kauffman 93] Kauffman, S. A.: *The Origins of Order: Self-Organization and Selection in Evolution*, Oxford University Press (1993)

[Koppen 07] Koppen, M., and Yoshida, K.: “Substitute Distance Assignments in NSGA-II for Handling Many-Objective Optimization Problems”, *Proc. 4th Intl. Conf. on Evolutionary Multi-Criterion Optimization*, Springer, LNCS, vol.4403, pp,727-741 (2007)

[Kukkonen 07] Kukkonen, S., and Lampinen, J.: “Ranking-dominance and Many Objective Optimization”, *Proc. 2007 IEEE Congress on Evolutionary Computation*, pp. 3983-3990 (2007)

[López Jaimes 08] López Jaimes, A., Coello, C., and Chakraborty, D.: “Objective Reduction Using a Feature Selection Technique”, *Proc. 2008 Genetic and Evolutionary Computation Conference*, ACM Press, pp.674-680 (2008)

[Sato 07] Sato, H., Aguirre H., and Tanaka K.: “Controlling Dominance Area of Solutions in Multiobjective Evolutionary Algorithms and Performance Analysis on Multiobjective 0/1 Knapsack Problems”, *IPSJ Trans. Mathematical Modeling and its Applications*, vol. 48, no. SIG15 (TOM 18), pp.137-152 (2007)

[Zitzler 99] Zitzler, E.: *Evolutionary Algorithms for Multiobjective Optimization: Methods and Applications*, PhD thesis, Swiss Federal Institute of Technology, Zurich (1999)

[Zitzler 04] Zitzler, E., and Kunzli, S. : “Indicator-based Selection in Multiobjective Search”, *Proc. 8th Intl. Conf. on Parallel Problem Solving from Nature*, Springer, LNCS, vol.3242, pp. 832-842 (2004)

〔担当委員：渡邉 真也〕

Received August 19, 2010.

Author’s Profile

Aguirre, Hernán (Member)

received his Engineer degree in computer systems from Escuela Politécnica Nacional, Quito, Ecuador in 1992. From 1997 to 2003 he was a research scholar sponsored by the Japanese Ministry of Education, Culture, Sports, Science and Technology. He received the M.S. and Ph.D. degrees from Shinshu University, Japan, in 2000 and 2003, respectively. Currently, he is an assistant professor at Shinshu University. His research interests include evolutionary computation, multidisciplinary design optimization, computational intelligence, information security, parallel computing, and their applications. He is a member of IEEE, ACM-SIGEVO, and IPSJ.

Tanaka, Kiyoshi (Member)

received B.S and M.S. degrees in Electrical Engineering and Operations Research from National Defense Academy, Yokosuka, Japan, in 1984 and 1989, respectively. In 1992, he received Dr. Eng. degree from Keio University, Tokyo, Japan. In 1995, he joined the Department of Electrical and Electronic Engineering, Faculty of Engineering, Shinshu University, Nagano, Japan, where he is currently a professor. His research interests include image and video processing, information hiding, evolutionary computation, chaos & fractals, and their applications. He is a member of IEEE, IEICE, IPSJ and IIEEJ. He is the vice-chairman of journal editing committee of IIEEJ.