Supplementary File of "Fuzzy Classifier-Assisted Solution Transfer for Evolutionary Sequential Transfer Optimization"

Wu Lin, Qiuzhen Lin, *Member, IEEE*, Liang Feng, *Senior Member, IEEE*, Kay Chen Tan, *Fellow, IEEE*

Abstract

This file includes eleven tables and one figure. Table S-I and Table S-II show the detailed parameter settings of the test problems in two benchmark suites, respectively. Table S-III summarizes the comparison results of ESTOA-FCM with different values of parameters. Table S-IV presents the total running times of each algorithm on STOP1 to STOP12. Table S-V and Table S-VI provide the detailed numerical results of ESTOA-FCM with all compared algorithms on two benchmark suites, respectively. Table S-VII and Table S-VIII show the comparison results of ESTOA-FCM and its six variants. The detailed numerical results of ESTOA-FCM with different values of parameters are given in Table S-IX, Table S-X, and Table S-XI. Additionally, Fig. S-1 shows the convergence curves of all compared algorithms on three practical test problems.

S-I. Description of Test Problems

A. Benchmark Suite 1

In the commonly used multitasking benchmark suite [1], nine test problems (i.e., CIHS, CIMS, CILS, PIHS, PIMS, PILS, NIHS, NIMS, and NILS) are designed by considering the task similarity and the degree of global optima intersection, each of which has two different tasks. To construct the sequential transfer optimization problem (STOP), the target task is set to one existing task from one multitasking test problem while its associated knowledge base is formed by configuring another task with different global optima to generate the source tasks. As shown in S-Table I, a weight parameter $\tau \in [0, 1]$ is introduced as the perturbation factor to generate the global optima of the source tasks. Setting τ to 0 means that the global optimum of the generated source task is the same as that of the original task. When τ is set to 1, its global optimum is generated by randomly sampling in the normalized search space. Thus, using the Gaussian distribution to sample τ can diversify the relationship between the global optima of the source tasks and the original task. Here, the mean (μ), the standard deviation (σ), and the number of the source tasks (k) are 0.5, 0.1, and 100, respectively. The detailed parameter settings of F1-F18 are given in S-Table I.

B. Benchmark Suite 2

As introduced in [2], STOPs with diverse properties can be generated by using a problem generator with six necessary parameters, including task family ($\mathcal{TF}$), transfer scenario ($\mathcal{TS}$), optimum coverage of the image (ξ), similarity distribution ($\mathcal{SD}$), problem dimension (d), and the number of source tasks (k). By setting different parameters, the similarity distribution of the source tasks to the target task can be flexibly adjusted to mimic the diversity of similarity relationship in real-world problems. Thus, the problem generator generates a specific STOP by setting the parameters, i.e., $\mathcal{TF}$-$\mathcal{TS}$-ξ-$\mathcal{SD}$-d-k. Here, eight widely used single-objective optimization functions are employed as the candidate families for formulating the source and target tasks, i.e., $\mathcal{TF}$ = {Sphere, Ellipsoid, Schwefel 2.2, Quartic, Ackley, Rastrigin, Griewank, Levy}. In addition, there are two different transfer scenarios including the intra-family and the inter-family transfers, i.e., $\mathcal{TS}$ = {$\mathcal{T}_a$, $\mathcal{T}_e$}. The former shows that the source and target tasks are from the same family while the latter indicates that the families of the source tasks are different to that of the target task. The parameter $\xi \in [0, 1]$ determines the relative size of the image over the decision space. To create the STOPs with diverse similarity distributions, a weight parameter $\tau \in [0, 1]$ is used to control the relationship between the optimal solutions of the source and target tasks, which is given as follows:

$$\begin{cases} \mathbf{o}_t = \hat{\mathbf{x}}_{lb} + \mathbf{r} \times (\hat{\mathbf{x}}_{ub} - \hat{\mathbf{x}}_{lb}) \\ \mathbf{o}_{si}^b = \hat{\mathbf{x}}_{lb} + \mathbf{r} \times (\hat{\mathbf{x}}_{ub} - \hat{\mathbf{x}}_{lb}), \ \ i = 1,\ 2,\ ...,\ k \\ \mathbf{o}_{si} = \mathbf{o}_t \times \tau_i + \mathbf{o}_{si}^b \times (1 - \tau_i), \ \ i = 1,\ 2,\ ...,\ k \end{cases}, \tag{1}$$

where $\mathbf{o}_t$ and $\mathbf{o}_{si}$ are the optima of the target task and the ith source task, and τ_i is the weight parameter imposed on the optimum of the ith source task. Here, five different probability distributions of τ are built based on one or multiple Gaussian distributions with the pre-set mean and standard deviation, which are respectively presented, as follows:

$$p_1(\tau) = \begin{cases} \mathcal{N}(0.15,\ 0.1^2), \ i = 1,\ ...,\ \lfloor 2k/3 \rfloor \\ \mathcal{N}(0.45,\ 0.2^2), \ i = \lfloor 2k/3 \rfloor + 1,\ ...,\ k \end{cases} \tag{2}$$

$$p_2(\tau) = \begin{cases} \mathcal{N}(0.45,\ 0.2^2), \ i = 1,\ ...,\ \lfloor k/3 \rfloor \\ \mathcal{N}(0.70,\ 0.1^2), \ i = \lfloor k/3 \rfloor + 1,\ ...,\ k \end{cases} \tag{3}$$

$$p_3(\tau) = \mathcal{N}(0.45,\ 0.2^2), \ i = 1,\ ...,\ k \tag{4}$$

$$p_4(\tau) = \begin{cases} \mathcal{N}(0.15,\ 0.1^2), \ i = 1,\ ...,\ \lfloor k/3 \rfloor \\ \mathcal{N}(0.45,\ 0.1^2), \ i = \lfloor k/3 \rfloor + 1,\ ...,\ \lfloor 2k/3 \rfloor \\ \mathcal{N}(0.70,\ 0.1^2), \ i = \lfloor 2k/3 \rfloor + 1,\ ...,\ k \end{cases} \tag{5}$$

$$p_5(\tau) = \begin{cases} \mathcal{N}(0.15,\ 0.1^2), \ i = 1,\ ...,\ \lfloor k/2 \rfloor \\ \mathcal{N}(0.70,\ 0.1^2), \ i = \lfloor k/2 \rfloor + 1,\ ...,\ k \end{cases}, \tag{6}$$

TABLE S-I

PARAMETER SETTINGS OF F1-F18

Problem ID	Target Task			Source Tasks in Knowledge Base		
	Function	**Search Space**	**Normalized Global Optimum**	**Function**	**Search Space**	**Normalized Global Optimum**
F1	CIHS-T1	d = 50, $[-100, 100]^d$	$\mathbf{o}_t^1 = [0.5,...,0.5]^d$	CIHS-T2	d = 50, $[-50, 50]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^2 \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F2	CIHS-T2	d = 50, $[-50, 50]^d$	$\mathbf{o}_t^2 = [0.5,...,0.5]^d$	CIHS-T1	d = 50, $[-100, 100]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^1 \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F3	CIMS-T1	d = 50, $[-50, 50]^d$	$\mathbf{o}_t^3 = [0.5,...,0.5]^d$	CIMS-T2	d = 50, $[-50, 50]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^4 \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F4	CIMS-T2	d = 50, $[-50, 50]^d$	$\mathbf{o}_t^4 = [0.5,...,0.5]^d$	CIMS-T1	d = 50, $[-50, 50]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^3 \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F5	CILS-T1	d = 50, $[-50, 50]^d$	$\mathbf{o}_t^5 = [0.92,...,0.92]^d$	CILS-T2	d = 50, $[-500, 500]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^6 \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F6	CILS-T2	d = 50, $[-500, 500]^d$	$\mathbf{o}_t^6 = [0.92,...,0.92]^d$	CILS-T1	d = 50, $[-50, 50]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^5 \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F7	PIHS-T1	d = 50, $[-50, 50]^d$	$\mathbf{o}_t^7 = [0.5,...,0.5]^d$	PIHS-T2	d = 50, $[-100, 100]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^8 \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F8	PIHS-T2	d = 50, $[-100, 100]^d$	$\mathbf{o}_t^8 = \left[[0.5,...,0.5]^{d/2},\ [0.6,...,0.6]^{d/2}\right]$	PIHS-T1	d = 50, $[-50, 50]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^7 \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F9	PIMS-T1	d = 50, $[-50, 50]^d$	$\mathbf{o}_t^9 = \left[[0.5,...,0.5]^{d/2},\ [0.51,...,0.51]^{d/2}\right]$	PIMS-T2	d = 50, $[-50, 50]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^{10} \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F10	PIMS-T2	d = 50, $[-50, 50]^d$	$\mathbf{o}_t^{10} = [0.51,...,0.51]^d$	PIMS-T1	d = 50, $[-50, 50]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^9 \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F11	PILS-T1	d = 50, $[-50, 50]^d$	$\mathbf{o}_t^{11} = [0.5,...,0.5]^d$	PILS-T2	d = 25, $[-0.5, 0.5]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^{12} \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F12	PILS-T2	d = 25, $[-0.5, 0.5]^d$	$\mathbf{o}_t^{12} = [0.5,...,0.5]^d$	PILS-T1	d = 50, $[-50, 50]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^{11} \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F13	NIHS-T1	d = 50, $[-50, 50]^d$	$\mathbf{o}_t^{13} = [0.51,...,0.51]^d$	NIHS-T2	d = 50, $[-50, 50]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^{14} \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F14	NIHS-T2	d = 50, $[-50, 50]^d$	$\mathbf{o}_t^{14} = [0.5,...,0.5]^d$	NIHS-T1	d = 50, $[-50, 50]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^{13} \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F15	NIMS-T1	d = 50, $[-100, 100]^d$	$\mathbf{o}_t^{15} = [0.55,...,0.55]^d$	NIMS-T2	d = 50, $[-0.5, 0.5]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^{16} \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F16	NIMS-T2	d = 50, $[-0.5, 0.5]^d$	$\mathbf{o}_t^{16} = [0.5,...,0.5]^d$	NIMS-T1	d = 50, $[-100, 100]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^{15} \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F17	NILS-T1	d = 50, $[-50, 50]^d$	$\mathbf{o}_t^{17} = [0.5,...,0.5]^d$	NILS-T2	d = 50, $[-500, 500]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^{18} \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$
F18	NILS-T2	d = 50, $[-500, 500]^d$	$\mathbf{o}_t^{18} = [0.92,...,0.92]^d$	NILS-T1	d = 50, $[-50, 50]^d$	$\mathbf{o}_{si} = \mathbf{o}_t^{17} \times (1-\tau_i) + rand(1,d) \times \tau_i, i = 1,...,k$

TABLE S-II

PARAMETER SETTINGS OF STOP1-STOP12

Problem ID	Problem Specification ($\mathcal{TF}$-$\mathcal{TS}$-ξ-$\mathcal{SD}$-d-k)	The Proportions of Different Types of Source Tasks		
		Low ($0 \le \mathcal{S} \le 0.3$)	Medium ($0.3 < \mathcal{S} < 0.7$)	High ($0.7 \le \mathcal{S} \le 1$)
STOP1	Sphere-$\mathcal{T}_a$-1-$\mathcal{N}_1$-35-100	38.00%	57.00%	5.00%
STOP2	Ellipsoid-$\mathcal{T}_e$-1-$\mathcal{N}_1$-50-100	47.00%	49.00%	4.00%
STOP3	Schwefel-$\mathcal{T}_a$-1-$\mathcal{N}_1$-60-100	56.00%	40.00%	4.00%
STOP4	Quartic-$\mathcal{T}_e$-1-$\mathcal{N}_2$-35-100	3.00%	44.00%	53.00%
STOP5	Ackley-$\mathcal{T}_a$-1-$\mathcal{N}_2$-50-100	6.00%	48.00%	46.00%
STOP6	Rastrigin-$\mathcal{T}_e$-1-$\mathcal{N}_2$-60-100	2.00%	52.00%	46.00%
STOP7	Griewank-$\mathcal{T}_a$-1-$\mathcal{N}_3$-35-100	13.00%	75.00%	12.00%
STOP8	Levy-Te-$\mathcal{T}_e$-1-$\mathcal{N}_3$-35-200	9.00%	72.00%	19.00%
STOP9	Ellipsoid-$\mathcal{T}_a$-1-$\mathcal{N}_4$-50-100	25.00%	49.00%	26.00%
STOP10	Quartic-$\mathcal{T}_e$-1-$\mathcal{N}_4$-50-200	24.50%	51.50%	24.00%
STOP11	Ackley-$\mathcal{T}_a$-1-$\mathcal{N}_5$-60-100	36.00%	31.00%	33.00%
STOP12	Griewank-$\mathcal{T}_e$-1-$\mathcal{N}_5$-60-200	35.00%	32.50%	32.50%

where k is the number of source tasks and $\lfloor \cdot \rfloor$ is the operator of rounding down. The similarity between the global optimum of the ith source task and the target task is measured by

$$\mathcal{S}_i = 1 - \max_j \left(|\, \mathbf{o}_t^j - \mathbf{o}_{si}^j \,| \right), \qquad (7)$$

where $\mathbf{o}_t^j$ and $\mathbf{o}_{si}^j$ denote the jth variables of $\mathbf{o}_t$ and $\mathbf{o}_{si}$, and $|\cdot|$ is the absolute value. The computed similarity degrees in the ranges, i.e., [0, 0.3], (0.3, 0.7), and [0.7, 1], are considered to be high, medium, and low, respectively. Therefore, $\mathcal{SD}$ = $\{p_1(\tau),\ p_2(\tau),\ p_3(\tau),\ p_4(\tau),\ p_5(\tau)\}$ is employed to mimic the diversity of similarity relationships of the source task to the target task in various STOPs. Moreover, d and k are the

TABLE S-III
SUMMARIZED RESULTS OF ESTOA-FCM WITH DIFFERENT VALUES OF α, K, AND c

Different Parameter Settings	STOP1-STOP12 (+/-/~)
$\alpha = 0.0$ vs $\alpha = 0.2$	0/2/10
$\alpha = 0.4$ vs $\alpha = 0.2$	0/3/9
$\alpha = 0.6$ vs $\alpha = 0.2$	0/1/11
$\alpha = 0.8$ vs $\alpha = 0.2$	0/3/9
$\alpha = 1.0$ vs $\alpha = 0.2$	1/10/1
$K = 1$ vs $K = 5$	1/8/3
$K = 3$ vs $K = 5$	0/4/8
$K = 7$ vs $K = 5$	0/0/12
$K = 9$ vs $K = 5$	0/2/10
$c = 5$ vs $c = 1$	0/11/1
$c = 10$ vs $c = 1$	0/11/1
$c = 20$ vs $c = 1$	0/12/0

"+" (or "-") indicates the number of test problems on which ESTOA-FCM with the corresponding parameter is better (or worse) than ESTOA-FCM with the suggested parameter, and "~" indicates the number of test problems on which they obtain similar performance.

problem's dimensionality and the number of source tasks in the knowledge base, respectively. Note that different STOPs can be configured with different dimensions, but the source and target tasks have the same dimension in the same STOP. In the problem generator [2], a conventional evolutionary algorithm is employed as the basic solver to optimize all the source tasks and their evaluated solutions are collected to form the knowledge base. According to the above parameter settings, a benchmark suite consisting of STOP1-STOP12 is designed to provide a set of representative test problems, where their knowledge bases are configured with different proportions of the source tasks with low, medium, and high similarity for their target tasks. The detailed parameter settings are given in Table S-II.

S-II. PARAMETER SENSITIVITY ANALYSIS

A. The Effect of α

To study the impact of α in the progressional representation, the comparisons of ESTOA-FCM using the progressional representation with different values of α from {0, 0.2, 0.4, 0.6, 0.8, 1} are done on STOP1-STOP12. The summarized results are collected in Table S-III, while the mean objective values and standard deviations are listed in Tables S-IX of the supplementary file. As summarized in Table S-III, compared to ESTOA-FCM with other values of α that are smaller than 1, i.e., α = 0, 0.4, 0.6, and 0.8, ESTOA-FCM with $\alpha = 0.2$ shows similar performance on most test problems (i.e., 10, 9, 11, and 9 cases out of 12 test problems). However, it can be observed that the significant performance deterioration of ESTOA-FCM with $\alpha = 1$. In particular, ESTOA-FCM with $\alpha = 0.2$ achieves significantly better performance on 10 test problems when compared to that with $\alpha = 1$. In addition, there is only one case on which ESTOA-FCM with $\alpha = 1$ is better than ESTOA-FCM with $\alpha = 0.2$. The above comparison results show that the current generation's population plays a critical role in using the progressional representation to estimate the population distribution. In summary, setting α to a value less than 1 is suggested. Thus, α is set to 0.2 in this study.

TABLE S-IV
THE TOTAL RUNNING TIME OF ALL COMPARED ALGORITHMS

Algorithm	Running Time (s)
ESTOA-ED	75
ESTOA-WD	85
ESTOA-KLD	341
ESTOA-MMD	200
ESTOA-AD	297
ESTOA-INB	305
ESTOA-FCM	109

B. The Effect of K

To study the impact of the number of nearest neighbours in FKNN, the comparisons of ESTOA-FCM using FKNN with different values of K from {1, 3, 5, 7, 9} are done on STOP1-STOP12. The summarized results are collected in Table S-III, while the mean objective values and standard deviations are listed in Table S-X of the supplementary file. As shown in Table S-III, ESTOA-FCM with $K = 5$ performs better than that with $K = 1$ on most test problems (i.e., 8 out of 12 test problems), and achieves similar results on 3 cases. In addition, compared to ESTOA-FCM with K = 3, 7, and 9, ESTOA-FCM with $K = 5$ achieves similar results on most test cases, i.e., 8, 12, and 10 cases out of 12 test cases, respectively. Meanwhile, there is no performance degradation on other test problems. The above comparison results show that ESTOA-FCM with a smaller value of K, i.e., $K = 1$, will cause the performance degradation of ESTOA-FCM on some test problems. In addition, the performance of ESTOA-FCM on most test problems is not very sensitive to FKNN with a larger value of K. Thus, setting K to 5 is suggested in this study.

C. The Effect of c

To study the impact of the number of transferred solutions at each transferable generation, the comparisons of ESTOA-FCM with different values of c from {1, 5, 10, 20} are done on STOP1-STOP12. The summarized results are collected in Table S-III, while the mean objective values and standard deviations are listed in Table S-XI of the supplementary file. It is observed from Table S-III that ESTOA-FCM with $c = 1$ achieves significantly better performance on most test problems (i.e., 11, 11, and 12 out of 12 test problems) when compared to that with other values of c. The comparison results show that selecting more transferred solutions from one source task will result in performance degradation. Thus, setting c to 1 is suggested in this study.

S-III. COMPUTATION COMPLEXITY ANALYSIS

For distance metrics, the computational complexities of ED, WD, KLD, and MMD are $O(Nd)$, $O(Nd)$, $O(N^3d^3)$, and $O(N^2d)$, respectively. Here, N and d are the population size and the problem's dimensionality, respectively. Considering machine learning models, their computational complexities are significantly influenced by four essential components, such as training data construction, model training, solution usefulness measurement, and training data update. Unlike AD and INB, which require the construction of a multivariate Gaussian distribution and the computation of probability distributions, FKNN does not introduce additional time complexity during the model training phase, as it operates directly on the training data without the need for an explicit parameter estimation process. Table S-IV presents the total running time of each algorithm for solving STOP1 to STOP12. All compared algorithms are implemented in MATLAB R2020b and run on a computer with AMD Ryzen 9 5900X 12-Core Processor. It is observed that the running speed of ESTOA-FCM is slightly slower than that of ESTOA-ED and ESTOA-WD, while it can significantly outperform the running speed of other algorithms. The comparative analysis thus indicates that introducing FKNN into ESTOA-FCM does not lead to a substantial increase in computational overhead.

References

[1] B. Da *et al.*, "Evolutionary Multitasking for Single-objective Continuous Optimization: Benchmark Problems, Performance Metric, and Baseline Results," Dept. Comput. Sci., Nanyang Technol. Univ., Singapore, Rep., 2016.

[2] X. Xue, C. Yang, L. Feng, K. Zhang, L. Song, and K. C. Tan, "A Scalable Problem Generator for Sequential Transfer Optimization," arXiv preprint. Oct. 2023. https://doi.org/10.48550/arXiv.2304.08503.

TABLE S-V

MEAN OBJECTIVE VALUES AND STANDARD DEVIATIONS OBTAINED BY ESTOA-FCM, EA, AND SIX ESTO ALGORITHMS

Problem		EA	ESTOA-ED	ESTOA-WD	ESTOA-KLD	ESTOA-MMD	ESTOA-AD	ESTOA-INB	ESTOA-FCM
F1	mean	1.05e+00(-)	1.02e+00(~)	1.02e+00(~)	1.02e+00(~)	1.03e+00(-)	1.03e+00(-)	1.03e+00(-)	**1.01e+00**
	std	1.79e-02	2.92e-02	3.08e-02	2.98e-02	3.80e-02	1.67e-02	1.90e-02	**2.86e-02**
F2	mean	5.52e+02(~)	5.57e+02(-)	5.52e+02(~)	5.39e+02(~)	5.62e+02(-)	**5.36e+02(~)**	5.39e+02(~)	5.36e+02
	std	4.18e+01	3.98e+01	4.55e+01	3.62e+01	4.43e+01	**3.26e+01**	3.32e+01	3.78e+01
F3	mean	9.44e+00(-)	5.50e+00(-)	5.57e+00(-)	5.36e+00(-)	6.02e+00(-)	5.41e+00(-)	5.26e+00(~)	**5.09e+00**
	std	1.83e+00	7.20e-01	7.67e-01	4.35e-01	8.42e-01	4.31e-01	4.87e-01	**5.01e-01**
F4	mean	5.53e+02(-)	5.41e+02(~)	5.37e+02(~)	5.31e+02(~)	5.50e+02(~)	**5.28e+02(~)**	5.58e+02(-)	5.39e+02
	std	4.07e+01	3.19e+01	3.23e+01	2.82e+01	4.59e+01	**3.81e+01**	3.94e+01	3.20e+01
F5	mean	2.13e+01(~)	2.13e+01(~)	2.13e+01(~)	2.13e+01(~)	**2.13e+01(~)**	2.13e+01(-)	2.13e+01(~)	2.13e+01
	std	4.12e-02	8.53e-02	9.47e-02	3.92e-02	**3.36e-02**	4.36e-02	3.94e-02	4.55e-02
F6	mean	**1.99e+03(~)**	2.11e+03(~)	2.18e+03(~)	2.19e+03(~)	2.14e+03(~)	2.15e+03(~)	2.10e+03(~)	2.09e+03
	std	**5.21e+02**	5.17e+02	4.93e+02	4.64e+02	5.48e+02	4.44e+02	4.61e+02	4.42e+02
F7	mean	5.58e+02(-)	5.58e+02(-)	5.63e+02(-)	5.39e+02(~)	5.61e+02(-)	5.42e+02(~)	5.59e+02(-)	**5.34e+02**
	std	4.63e+01	5.05e+01	4.22e+01	3.75e+01	3.18e+01	2.81e+01	3.68e+01	**3.52e+01**
F8	mean	2.33e+02(-)	2.01e+02(~)	1.90e+02(~)	**1.65e+02(~)**	2.38e+02(-)	1.68e+02(~)	2.50e+02(-)	1.82e+02
	std	6.24e+01	7.43e+01	7.80e+01	**4.31e+01**	1.35e+02	3.77e+01	6.51e+01	4.94e+01
F9	mean	1.02e+01(-)	5.31e+00(-)	5.12e+00(-)	5.32e+00(-)	6.15e+00(-)	5.36e+00(-)	9.83e+00(-)	**4.79e+00(~)**
	std	2.28e+00	5.40e-01	5.80e-01	3.24e-01	1.36e+00	4.32e-01	2.26e+00	**3.19e-01**
F10	mean	3.03e+05(-)	1.23e+05(-)	1.08e+05(-)	6.99e+04(-)	7.80e+04(-)	7.11e+04(-)	3.01e+05(-)	**3.89e+04(~)**
	std	1.65e+05	7.56e+04	6.50e+04	3.26e+04	4.16e+04	3.88e+04	1.38e+05	**2.16e+04**
F11	mean	9.67e+00(-)	7.37e+00(~)	7.71e+00(~)	7.97e+00(-)	**7.33e+00(~)**	7.80e+00(~)	8.75e+00(-)	7.50e+00(~)
	std	1.68e+00	9.47e-01	1.09e+00	1.09e+00	**8.96e-01**	1.08e+00	1.76e+00	1.01e+00
F12	mean	9.64e+00(-)	7.05e+00(~)	6.87e+00(~)	**6.44e+00(~)**	7.16e+00(~)	6.84e+00(~)	7.64e+00(-)	6.72e+00(~)
	std	1.91e+00	1.64e+00	1.47e+00	**1.37e+00**	2.20e+00	1.30e+00	2.11e+00	1.34e+00
F13	mean	3.43e+05(-)	9.03e+04(-)	9.80e+04(-)	6.95e+04(-)	1.18e+05(-)	7.37e+04(-)	7.46e+04(-)	**2.37e+04(~)**
	std	1.67e+05	6.68e+04	6.24e+04	2.78e+04	7.87e+04	2.85e+04	3.94e+04	**1.16e+04**
F14	mean	5.51e+02(-)	5.53e+02(-)	5.51e+02(~)	**5.31e+02(~)**	5.49e+02(~)	5.35e+02(~)	5.51e+02(~)	5.36e+02(~)
	std	3.61e+01	3.09e+01	3.69e+01	**3.34e+01**	4.02e+01	3.38e+01	4.33e+01	3.53e+01
F15	mean	1.05e+00(-)	1.04e+00(~)	1.05e+00(~)	1.04e+00(~)	1.06e+00(-)	1.05e+00(~)	**1.03e+00(+)**	1.05e+00(~)
	std	1.97e-02	1.90e-02	1.57e-02	1.98e-02	2.21e-02	1.63e-02	**2.83e-02**	1.71e-02
F16	mean	3.41e+01(-)	2.14e+01(~)	2.04e+01(~)	2.03e+01(~)	2.32e+01(-)	**2.00e+01(~)**	3.23e+01(-)	2.11e+01(~)
	std	4.86e+00	2.98e+00	3.53e+00	2.12e+00	5.26e+00	**2.93e+00**	4.05e+00	3.45e+00
F17	mean	5.62e+02(~)	5.70e+02(~)	5.72e+02(~)	5.69e+02(~)	5.62e+02(~)	**5.59e+02(~)**	5.67e+02(~)	5.62e+02(~)
	std	3.63e+01	4.03e+01	4.21e+01	3.67e+01	3.55e+01	**4.15e+01**	4.49e+01	3.34e+01
F18	mean	2.08e+03(~)	2.14e+03(~)	2.05e+03(~)	2.25e+03(~)	2.12e+03(~)	2.14e+03(~)	**1.97e+03(~)**	2.04e+03(~)
	std	5.58e+02	4.51e+02	4.28e+02	4.83e+02	5.21e+02	4.23e+02	**4.96e+02**	4.83e+02
Best/All		1/18	0/18	0/18	3/18	2/18	4/18	2/18	6/18
+/-/~		0/13/5	0/7/11	0/5/13	0/5/13	0/9/9	0/6/12	1/10/7	\

"+", "-", and "~" indicate that the results of the corresponding algorithm are better than, worse than, and similar to that of ESTOA-FCM, respectively. The best result on each test problem is highlighted in bold.

TABLE S-VI

MEAN OBJECTIVE VALUES AND STANDARD DEVIATIONS OBTAINED BY ESTOA-FCM, EA, AND SIX ESTO ALGORITHMS

Problem		EA	ESTOA-ED	ESTOA-WD	ESTOA-KLD	ESTOA-MMD	ESTOA-AD	ESTOA-INB	ESTOA-FCM
STOP1	mean	1.65e+02(~)	2.76e+02(-)	2.59e+02(-)	1.69e+02(~)	2.74e+02(-)	**1.57e+02**(~)	2.16e+02(~)	1.67e+02
	std	1.09e+02	1.46e+02	1.13e+02	1.20e+02	1.25e+02	**1.07e+02**	1.55e+02	9.52e+01
STOP2	mean	4.72e+03(-)	3.96e+03(-)	3.40e+03(-)	3.63e+03(-)	3.55e+03(-)	3.20e+03(~)	4.35e+03(-)	**2.71e+03**
	std	2.09e+03	1.58e+03	1.05e+03	1.56e+03	1.23e+03	1.40e+03	2.23e+03	**1.01e+03**
STOP3	mean	2.23e+01(~)	3.23e+01(-)	2.91e+01(-)	2.44e+01(~)	2.75e+01(~)	**2.05e+01**(+)	2.24e+01(~)	2.57e+01
	std	8.16e+00	1.39e+01	9.83e+00	9.67e+00	9.10e+00	**7.17e+00**	8.26e+00	8.81e+00
STOP4	mean	4.05e+00(-)	5.20e-01(-)	6.11e-01(-)	5.34e-01(-)	5.58e-01(-)	3.96e-01(~)	6.36e-01(-)	**3.79e-01**
	std	5.08e+00	2.18e-01	2.52e-01	2.38e-01	1.94e-01	2.09e-01	3.78e-01	**1.51e-01**
STOP5	mean	7.43e+00(-)	4.24e+00(-)	4.33e+00(-)	**3.68e+00**(~)	4.37e+00(-)	4.09e+00(-)	7.57e+00(-)	3.80e+00
	std	1.49e+00	6.32e-01	6.52e-01	**5.03e-01**	6.33e-01	5.59e-01	1.16e+00	6.08e-01
STOP6	mean	1.46e+02(-)	6.93e+01(~)	6.99e+01(-)	9.02e+01(-)	7.27e+01(~)	8.19e+01(-)	9.37e+01(-)	**6.73e+01**
	std	1.79e+01	2.58e+01	3.91e+01	1.60e+01	2.46e+01	1.98e+01	2.53e+01	**1.28e+01**
STOP7	mean	1.17e+00(-)	1.06e+00(~)	1.06e+00(~)	1.07e+00(-)	1.07e+00(-)	1.11e+00(-)	1.17e+00(-)	**1.05e+00**
	std	8.32e-02	3.85e-02	2.97e-02	3.88e-02	3.41e-02	9.11e-02	1.06e-01	**3.56e-02**
STOP8	mean	6.64e+00(-)	1.33e+00(~)	1.68e+00(-)	3.15e+00(-)	1.73e+00(~)	3.14e+00(-)	7.16e+00(-)	**1.27e+00**
	std	4.21e+00	7.09e-01	8.54e-01	2.35e+00	1.45e+00	2.39e+00	4.28e+00	**6.42e-01**
STOP9	mean	4.17e+03(-)	1.32e+03(~)	1.85e+03(-)	**1.30e+03(~)**	1.96e+03(-)	1.83e+03(-)	4.67e+03(-)	1.37e+03
	std	2.21e+03	4.95e+02	7.72e+02	**6.30e+02**	6.89e+02	1.00e+03	2.72e+03	4.68e+02
STOP10	mean	1.05e+02(-)	6.85e+00(~)	7.38e+00(~)	1.29e+01(-)	9.26e+00(~)	1.17e+01(-)	7.12e+01(-)	**6.04e+00**
	std	8.15e+01	7.21e+00	4.90e+00	1.04e+01	8.08e+00	8.50e+00	8.46e+01	**2.89e+00**
STOP11	mean	9.67e+00(-)	5.49e+00(-)	5.40e+00(-)	5.71e+00(-)	5.68e+00(-)	5.17e+00(~)	9.40e+00(-)	**5.04e+00**
	std	1.29e+00	1.01e+00	9.34e-01	9.40e-01	1.18e+00	7.78e-01	1.53e+00	**8.41e-01**
STOP12	mean	2.88e+00(-)	1.47e+00(-)	1.45e+00(~)	1.98e+00(-)	1.47e+00(-)	1.70e+00(-)	2.82e+00(-)	**1.38e+00**
	std	8.30e-01	2.52e-01	2.00e-01	6.05e-01	3.36e-01	2.90e-01	7.60e-01	**1.30e-01**
Best/All		0/12	0/12	0/12	2/12	0/12	2/12	0/12	8/12
+/-/~		0/10/2	0/7/5	0/9/3	0/8/4	0/8/4	1/7/4	0/10/2	\

"+", "-", and "~" indicate that the results of the corresponding algorithm are better than, worse than, and similar to that of ESTOA-FCM, respectively. The best result on each test problem is highlighted in bold.

TABLE S-VII

MEAN OBJECTIVE VALUES AND STANDARD DEVIATIONS OBTAINED BY ESTOA-FCM AND ITS THREE VARIANTS

Problem	Variant-I	Variant-II	Variant-III	ESTOA-FCM
STOP1	1.64e+02(1.01e+02)~	1.56e+02(1.19e+02)~	1.07e+02(6.32e+01)+	1.67e+02(9.52e+01)
STOP2	3.20e+03(1.27e+03)-	4.43e+03(2.39e+03)-	2.79e+03(1.58e+03)~	2.71e+03(1.01e+03)
STOP3	2.03e+01(7.15e+00)+	2.02e+01(7.35e+00)+	1.95e+01(8.02e+00)+	2.57e+01(8.81e+00)
STOP4	2.05e+00(2.10e+00)-	2.61e+00(2.28e+00)-	4.83e-01(1.67e-01)-	3.79e-01(1.51e-01)
STOP5	4.84e+00(9.19e-01)-	7.88e+00(1.22e+00)-	4.92e+00(7.91e-01)-	3.80e+00(6.08e-01)
STOP6	1.82e+02(3.80e+01)-	1.47e+02(2.08e+01)-	1.14e+02(2.64e+01)-	6.73e+01(1.28e+01)
STOP7	1.27e+00(2.14e-01)-	1.17e+00(1.35e-01)-	1.07e+00(2.78e-02)-	1.05e+00(3.56e-02)
STOP8	2.42e+00(2.19e+00)-	6.74e+00(3.75e+00)-	2.90e+00(2.11e+00)-	1.27e+00(6.42e-01)
STOP9	1.67e+03(7.13e+02)-	4.11e+03(2.01e+03)-	1.72e+03(7.05e+02)-	1.37e+03(4.68e+02)
STOP10	1.15e+01(7.44e+00)-	1.08e+02(9.12e+01)-	1.43e+01(1.22e+01)-	6.04e+00(2.89e+00)
STOP11	5.66e+00(1.04e+00)-	9.91e+00(1.64e+00)-	5.51e+00(9.49e-01)-	5.04e+00(8.41e-01)
STOP12	1.90e+00(4.25e-01)-	2.73e+00(7.82e-01)-	1.50e+00(2.11e-01)-	1.38e+00(1.30e-01)
+/-/~	1/10/1	1/10/1	2/9/1	\

"+" (or "-") indicates that the results of the corresponding variant are better (or worse) than that of ESTOA-FCM, and "~" indicates that they obtain similar performance.

TABLE S-VIII

MEAN OBJECTIVE VALUES AND STANDARD DEVIATIONS OBTAINED BY ESTOA-FCM AND ITS THREE VARIANTS

Problem	Variant-IV	Variant-V	Variant-VI	ESTOA-FCM
STOP1	2.04e+02(1.19e+02)~	1.87e+02(1.09e+02)~	2.17e+02(1.49e+02)~	1.67e+02(9.52e+01)
STOP2	4.22e+03(2.07e+03)-	3.54e+03(1.83e+03)-	3.19e+03(1.45e+03)~	2.71e+03(1.01e+03)
STOP3	2.20e+01(9.43e+00)+	2.17e+01(7.20e+00)+	2.03e+01(6.74e+00)+	2.57e+01(8.81e+00)
STOP4	1.49e+00(1.05e+00)-	4.39e-01(1.73e-01)~	5.78e-01(2.79e-01)-	3.79e-01(1.51e-01)
STOP5	6.97e+00(1.20e+00)-	3.98e+00(6.28e-01)~	5.27e+00(7.48e-01)-	3.80e+00(6.08e-01)
STOP6	1.48e+02(2.19e+01)-	8.06e+01(1.66e+01)-	6.86e+01(1.49e+01)~	6.73e+01(1.28e+01)
STOP7	1.20e+00(1.25e-01)-	1.13e+00(9.53e-02)-	1.19e+00(1.06e-01)-	1.05e+00(3.56e-02)
STOP8	6.26e+00(4.04e+00)-	3.11e+00(2.55e+00)-	2.91e+00(1.83e+00)-	1.27e+00(6.42e-01)
STOP9	4.44e+03(1.67e+03)-	1.68e+03(7.48e+02)-	2.68e+03(1.26e+03)-	1.37e+03(4.68e+02)
STOP10	1.13e+02(8.75e+01)-	9.72e+00(8.94e+00)~	4.27e+01(3.72e+01)-	6.04e+00(2.89e+00)
STOP11	7.79e+00(1.32e+00)-	5.06e+00(7.26e-01)~	5.66e+00(9.76e-01)-	5.04e+00(8.41e-01)
STOP12	2.83e+00(7.22e-01)-	1.61e+00(3.29e-01)-	1.43e+00(2.02e-01)~	1.38e+00(1.30e-01)
+/-/~	1/10/1	1/6/5	1/7/4	\

"+" (or "-") indicates that the results of the corresponding variant are better (or worse) than that of ESTOA-FCM, and "~" indicates that they obtain similar performance.

TABLE S-IX

MEAN OBJECTIVE VALUES AND STANDARD DEVIATIONS OBTAINED BY ESTOA-FCM WITH DIFFERENT VALUES OF α

Problem	α = 0.0	α = 0.4	α = 0.6	α = 0.8	α = 1.0	α = 0.2
STOP1	1.62e+02(1.11e+02)~	1.96e+02(1.30e+02)~	1.80e+02(9.64e+01)~	1.63e+02(1.17e+02)~	1.74e+02(1.17e+02)~	1.67e+02(9.52e+01)
STOP2	2.81e+03(1.13e+03)~	3.19e+03(1.16e+03)-	3.00e+03(1.09e+03)~	2.87e+03(9.64e+02)~	3.94e+03(2.27e+03)-	2.71e+03(1.01e+03)
STOP3	2.36e+01(9.15e+00)~	2.45e+01(9.45e+00)~	2.39e+01(8.56e+00)~	2.36e+01(7.91e+00)~	2.08e+01(8.34e+00)+	2.57e+01(8.81e+00)
STOP4	3.49e-01(1.55e-01)~	4.58e-01(1.97e-01)-	6.15e-01(2.94e-01)-	4.95e-01(2.49e-01)-	2.93e+00(2.72e+00)-	3.79e-01(1.51e-01)
STOP5	3.66e+00(5.20e-01)~	3.85e+00(6.21e-01)~	3.88e+00(5.11e-01)~	3.67e+00(5.43e-01)~	8.40e+00(9.62e-01)-	3.80e+00(6.08e-01)
STOP6	8.21e+01(1.89e+01)-	8.22e+01(2.12e+01)-	7.14e+01(1.97e+01)~	8.39e+01(2.59e+01)-	1.78e+02(1.98e+01)-	6.73e+01(1.28e+01)
STOP7	1.08e+00(3.96e-02)-	1.06e+00(6.55e-02)~	1.06e+00(3.19e-02)~	1.07e+00(2.26e-02)-	1.57e+00(2.24e-01)-	1.05e+00(3.56e-02)
STOP8	1.15e+00(5.28e-01)~	1.39e+00(7.10e-01)~	1.46e+00(7.38e-01)~	1.51e+00(8.02e-01)~	1.34e+01(8.54e+00)-	1.27e+00(6.42e-01)
STOP9	1.27e+03(4.88e+02)~	1.43e+03(6.54e+02)~	1.40e+03(5.06e+02)~	1.26e+03(5.90e+02)~	4.32e+03(2.19e+03)-	1.37e+03(4.68e+02)
STOP10	7.01e+00(5.30e+00)~	5.64e+00(2.84e+00)~	5.25e+00(2.64e+00)~	5.64e+00(3.12e+00)~	8.98e+01(6.79e+01)-	6.04e+00(2.89e+00)
STOP11	5.00e+00(8.22e-01)~	5.20e+00(9.10e-01)~	5.26e+00(8.41e-01)~	4.94e+00(9.80e-01)~	9.38e+00(1.33e+00)-	5.04e+00(8.41e-01)
STOP12	1.35e+00(1.28e-01)~	1.41e+00(1.58e-01)~	1.41e+00(1.29e-01)~	1.40e+00(1.91e-01)~	2.85e+00(7.06e-01)-	1.38e+00(1.30e-01)
+/-/~	0/2/10	0/3/9	0/1/11	0/3/9	1/10/1	\

"+" (or "-") indicates that the results of ESTOA-FCM with the corresponding parameter are better (or worse) than that of ESTOA-FCM with the suggested parameter, and "~" indicates that they obtain similar performance.

TABLE S-X

MEAN OBJECTIVE VALUES AND STANDARD DEVIATIONS OBTAINED BY ESTOA-FCM WITH DIFFERENT VALUES OF K

Problem	K = 1	K = 3	K = 7	K = 9	K = 5
STOP1	1.38e+02(7.46e+01)~	1.58e+02(9.55e+01)~	1.75e+02(1.30e+02)~	1.40e+02(8.76e+01)~	1.67e+02(9.52e+01)
STOP2	3.36e+03(1.35e+03)-	3.10e+03(1.31e+03)~	2.94e+03(1.19e+03)~	2.95e+03(1.65e+03)~	2.71e+03(1.01e+03)
STOP3	2.03e+01(7.35e+00)+	2.32e+01(8.05e+00)~	2.62e+01(9.60e+00)~	2.70e+01(1.03e+01)~	2.57e+01(8.81e+00)
STOP4	5.20e-01(2.93e-01)-	4.55e-01(2.10e-01)-	4.33e-01(1.87e-01)~	4.58e-01(2.03e-01)-	3.79e-01(1.51e-01)
STOP5	4.19e+00(6.49e-01)-	4.15e+00(6.47e-01)-	3.81e+00(5.85e-01)~	3.85e+00(6.29e-01)~	3.80e+00(6.08e-01)
STOP6	1.08e+02(2.10e+01)-	8.03e+01(1.73e+01)-	7.33e+01(1.67e+01)~	6.50e+01(1.52e+01)~	6.73e+01(1.28e+01)
STOP7	1.09e+00(3.07e-02)-	1.07e+00(3.42e-02)-	1.06e+00(3.19e-02)~	1.07e+00(2.76e-02)-	1.05e+00(3.56e-02)
STOP8	1.67e+00(1.05e+00)-	1.23e+00(6.03e-01)~	1.27e+00(7.26e-01)~	1.40e+00(8.07e-01)~	1.27e+00(6.42e-01)
STOP9	1.36e+03(4.69e+02)~	1.48e+03(6.68e+02)~	1.31e+03(5.01e+02)~	1.34e+03(5.45e+02)~	1.37e+03(4.68e+02)
STOP10	8.62e+00(5.24e+00)-	6.82e+00(4.78e+00)~	7.23e+00(4.76e+00)~	6.31e+00(4.39e+00)~	6.04e+00(2.89e+00)
STOP11	6.13e+00(8.93e-01)-	5.06e+00(7.87e-01)~	5.11e+00(7.76e-01)~	4.99e+00(8.63e-01)~	5.04e+00(8.41e-01)
STOP12	1.41e+00(1.33e-01)~	1.42e+00(1.70e-01)~	1.40e+00(1.65e-01)~	1.38e+00(1.46e-01)~	1.38e+00(1.30e-01)
+/-/~	1/8/3	0/4/8	0/0/12	0/2/10	\

"+" (or "-") indicates that the results of ESTOA-FCM with the corresponding parameter are better (or worse) than that of ESTOA-FCM with the suggested parameter, and "~" indicates that they obtain similar performance.

TABLE S-XI

MEAN OBJECTIVE VALUES AND STANDARD DEVIATIONS OBTAINED BY ESTOA-FCM WITH DIFFERENT VALUES OF *c*

Problem	***c* = 5**	***c* = 10**	***c* = 20**	***c* = 1**
STOP1	5.28e+02(3.50e+02)-	1.09e+03(9.72e+02)-	2.98e+03(2.48e+03)-	1.67e+02(9.52e+01)
STOP2	7.90e+03(3.91e+03)-	1.36e+04(1.05e+04)-	3.19e+04(3.60e+04)-	2.71e+03(1.01e+03)
STOP3	4.76e+01(2.44e+01)-	8.93e+01(3.97e+01)-	1.59e+02(8.42e+01)-	2.57e+01(8.81e+00)
STOP4	4.63e-01(2.03e-01)-	5.85e-01(2.75e-01)-	6.87e-01(3.19e-01)-	3.79e-01(1.51e-01)
STOP5	4.20e+00(7.59e-01)-	4.42e+00(7.39e-01)-	4.67e+00(7.13e-01)-	3.80e+00(6.08e-01)
STOP6	7.22e+01(2.43e+01)~	7.93e+01(3.16e+01)~	9.05e+01(3.52e+01)-	6.73e+01(1.28e+01)
STOP7	1.26e+00(1.59e-01)-	1.35e+00(3.30e-01)-	1.58e+00(4.62e-01)-	1.05e+00(3.56e-02)
STOP8	5.34e+00(3.36e+00)-	8.84e+00(7.03e+00)-	1.47e+01(1.46e+01)-	1.27e+00(6.42e-01)
STOP9	8.93e+03(5.14e+03)-	1.36e+04(8.65e+03)-	3.50e+04(2.12e+04)-	1.37e+03(4.68e+02)
STOP10	9.19e+01(1.13e+02)-	1.04e+02(1.54e+02)-	3.37e+02(4.14e+02)-	6.04e+00(2.89e+00)
STOP11	7.26e+00(1.89e+00)-	7.83e+00(1.74e+00)-	9.60e+00(2.59e+00)-	5.04e+00(8.41e-01)
STOP12	3.79e+00(2.31e+00)-	3.81e+00(1.56e+00)-	8.21e+00(6.69e+00)-	1.38e+00(1.30e-01)
+/-/~	0/11/1	0/11/1	0/12/0	\

"+" (or "-") indicates that the results of ESTOA-FCM with the corresponding parameter are better (or worse) than that of ESTOA-FCM with the suggested parameter, and "~" indicates that they obtain similar performance.

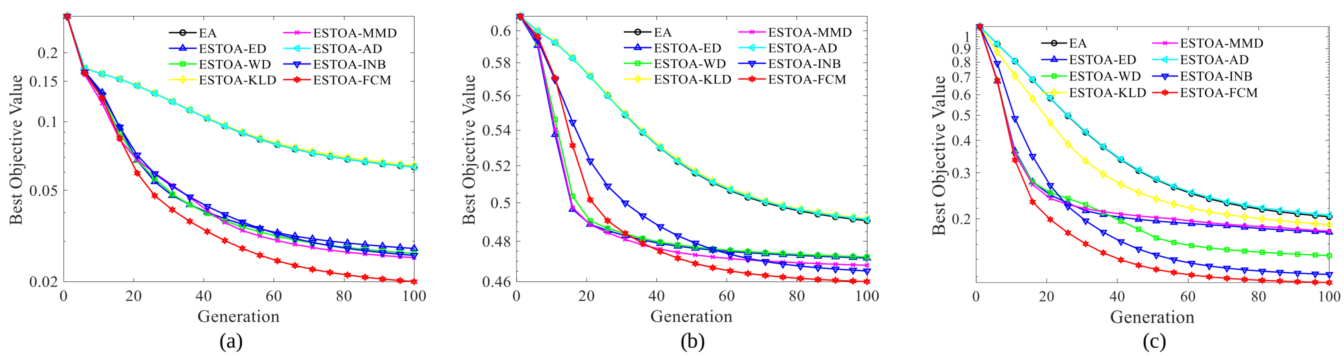

Fig. S-1. Convergence curves of all compared algorithms on three representative test problems with (a) d = 100, (b) d = 150, and (c) $d \in [50,100]$.