2014 IEEE Congress on Evolutionary Computation (CEC)
July 6-11, 2014, Beijing, China

Demonstrator Selection in a Social Learning Particle Swarm Optimizer

Ran Cheng
Department of Computing
University of Surrey
Guildford, Surrey, GU2 7XH
United Kingdom
Email: r.cheng@surrey.ac.uk

Yaochu Jin
Department of Computing
University of Surrey
Guildford, Surrey, GU2 7XH
United Kingdom
Email: yaochu.jin@surrey.ac.uk

***Abstract*—Social learning plays an important role in behavior learning among social animals. Different from individual (asocial) learning, social learning has the advantage of allowing individuals to learn behaviors from others without the extra costs of individual trial-and-error. Inspired by the natural social learning phenomenon, we have transplanted the social learning mechanism into particle swarm optimization (PSO) to develop a social learning PSO (SL-PSO). Unlike classical PSO variants, the SL-PSO is performed on a sorted swarm, and instead of merely learning from historical best positions, the particles are able to learn from anyone better (demonstrators) in the current swarm. A key mechanism in the SL-PSO is the learning strategy, where an imitator will learn from different demonstrators. However, in our previous work, little discussion has been focused on demonstrator selection, i.e., which demonstrators are to learn from by the imitator. In this paper, based on the analysis of the demonstrator selection in the SL-PSO, two demonstrator selection strategies are proposed. Experimental results show that, the proposed demonstrator selection strategies have significantly enhanced the performance of the SL-PSO in comparison to five representative PSO variants on a set of benchmark problems.**

I. INTRODUCTION

Social learning, different from individual (asocial) learning, has the advantage of allowing individuals to learn behaviors from others without incurring the costs of individual trial and error [1], which is able to accelerate learning rates [2], especially when the target (behavior) to learn is complex. More specifically, individual learning is a process of trial and error whilst social learning is a process taking advantage of mechanisms such as imitation, enhancement and conditioning [3].

Among various social learning mechanisms [4]–[6], imitation is considered to be the most distinctive one [7], because imitation, which operates across a whole community, could lead to population level similarities of behaviors [8], thus further triggering another significant mechanism, i.e., the culture transmission [9], [10]. In social learning theory, culture can be seen as the cumulative effect of countless processes of inter-individual transmission through imitation [11], which can be equally regarded as the global information shared by the whole population in swarm intelligence algorithms. Such global information, on the one hand, is a consequence of global search, on the other hand, can provide instructions for the whole population to move towards the global convergence.

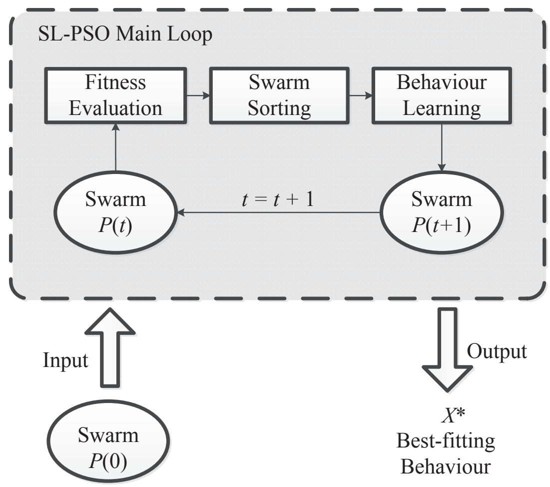


Fig. 1. The framework of the Social Learning Particle Swarm Optimizer (SL-PSO).

Due to the above mentioned attractive properties, it is quite natural to apply the social learning mechanisms to population-based stochastic optimization. Recently, the social learning mechanism has been transplanted into the particle swarm optimization [12] to develop a social learning particle swarm optimizer (SL-PSO) [13]. Unlike classical PSO variants, the proposed SL-PSO is performed on a sorted swarm, and instead of merely learning from historical best positions, the particles are able to learn from anyone better (demonstrators) in the current swarm. Due to the full contribution of the whole swarm, the learning procedure is expected to become more social. Experimental results have shown that the SL-PSO is able to perform robustly on a wide range of 47 benchmark functions, where the decision (search) space scales from 30-dimensional to 1000-dimensions [13].

It can be seen from Fig. 1 that, apart from fitness evaluations, the most important components in the framework of the SL-PSO are demonstrator selection and behavior learning, which operate to determine whom is to learn from and how to learn, respectively. In the original SL-PSO [13], a detailed description about the behavior learning is presented with a simple, random behavior selection strategy.

In this paper, the influence of the demonstrator selection on the search performance of SL-PSO has been examined. To address the weakness of the demonstrator selection strategy used in the original SL-PSO, two new demonstrator selection strategies based on different probability distributions are proposed to enhance the search ability of the SL-PSO by achieving a balance between convergence speed and population diversity.

This rest of this paper is organized as follows. In Section II, the original SL-PSO is introduced with the proposed demonstrator selection strategies. Experimental results on the proposed demonstrator selection strategies are presented with the comparisons to the original SL-PSO and five PSO variants in Section III. Section IV draws the conclusion.

II. DEMONSTRATOR SELECTION IN THE SL-PSO

As introduced in Section I, the most important components in the framework of the SL-PSO are demonstrator selection and behavior learning. In this section, details of the behavior learning are presented first, and then, based on a discussion of the original demonstrator selection strategy, two new demonstrator selection strategies are proposed.

A. Details of the behavior learning

According to the social learning theory, imitators can learn behaviors from different demonstrators [?] as follows:

$$X_{i,j}(t+1) = \begin{cases} X_{i,j}(t) + \Delta X_{i,j}(t+1), \text{ if } p_i(t) \leq P_i^L \\ X_{i,j}(t), \text{ otherwise} \end{cases} \quad (1)$$

where $X_{i,j}(t)$ is j-th dimension of particle i's behavior vector in generation t, with $i \in \{1, 2, 3, ..., m\}$ and $j \in \{1, 2, 3, ..., n\}$, $\Delta X_{i,j}(t+1)$ being the behavior correction. Taking into account the fact that not all particles are willing to learn from others, a learning probability (P_i^L) for each particle i has been defined. As a result, particle i will learn (correct its behavior) only if a randomly generated number p_i satisfies $0 \leq p_i(t) \leq P_i^L \leq 1$.

Specifically, the j-th dimension of particle i's behavior vector is corrected as follows in generation $t+1$:

$$\Delta X_{i,j}(t+1) = r_1 \Delta X_{i,j}(t) + r_2 I_{i,j}(t) + r_3 \epsilon \times C_{i,j}(t), \quad (2)$$

where,

$$\begin{cases} I_{i,j}(t) = X_{k,j}(t) - X_{i,j}(t), \\ C_{i,j}(t) = \bar{X}_j(t) - X_{i,j}(t). \end{cases} \quad (3)$$

where r_1, r_2 and r_3 are three random numbers uniformly distributed in $[0, 1]$, and k is the index of the selected demonstrator. From the above equations, we can see that the behavior correction $\Delta X_{i,j}(t+1)$ consists of three parts, namely, learning inertia $\Delta X_{i,j}(t)$, which is the behavior correction in the previous generation, imitation $I_{i,j}(t)$ and social influence $\epsilon \times C_{i,j}(t)$, where ϵ is termed social influence factor.

Among the three behavior correction components, imitation $I_{i,j}(t)$ is the most important social learning element that reflects the difference between the imitator's behavior $X_{i,j}(t)$ and the demonstrator's behavior $X_{k,j}(t)$, and $C_{i,j}(t)$ is the difference between behavior $X_{i,j}(t)$ and the mean behavior $\bar{X}_j(t)$ of all particles ($\bar{X}_j(t) = \frac{\sum_{i=1}^{m} X_{i,j}}{m}$), which can be seen as the swarm-level social influence.

In the SL-PSO, there are three parameters that need to be defined, i.e., the swarm size m, the learning probability P_i^L, and the social influence factor ϵ. A good *scalability* should be one of the most important criteria in respect of one optimizer's robustness, because the search dimensionality of real-world problem, which are problem dependent, could vary a lot. Unfortunately, PSO has been shown to have poor scalability in comparison with some other widely used evolutionary algorithms [14]. Therefore, in the previous work, these parameters are adaptively related to the search dimensionality n as follows:

$$m = M + \lfloor \frac{n}{10} \rfloor, \quad (4)$$

$$P_i^L = (1 - \frac{i-1}{m})^{log(\sqrt{\frac{n}{M}})}. \quad (5)$$

$$\epsilon = \frac{n}{M} \times 0.01, \quad (6)$$

where M is the base population size for the SL-PSO to work properly. Typically, $M = 100$ is recommended.

B. Demonstrator selection

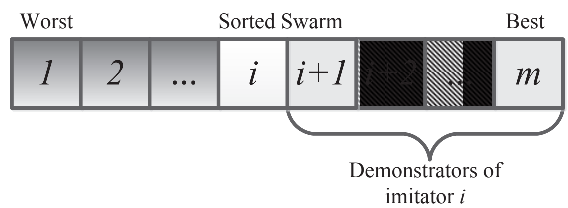


Fig. 2. The demonstrator selection strategy in the original SL-PSO.

In the original SL-PSO, before demonstrator selection, the swarm is first sorted (in the ascending order) according to the fitness values. Consequently, for the i-th particle (imitator), demonstrators are randomly chosen from the $(i+1)$-th particle to the m-th particle, where m is the swarm size, refer to Fig. 2. Accordingly, the probability that a particle is selected as a demonstrator will be:

$$Pr(k|i) = \frac{1}{m-i}, \quad (7)$$

where k is the demonstrator index, satisfying $i+1 \leq k \leq m$.

We can see that in this strategy, each demonstrator k for particle i has the same probability to be selected, independent of its fitness. This selection strategy has been shown to work effectively. However, a potential weakness in this selection strategy is that the better demonstrators have the same probability to be selected as the relatively weaker demonstrators, which may lead to a slow convergence. To address such a potential weakness, the most straightforward idea is to relate the selection probability to the demonstrators' fitness values, i.e., exerting different selection pressures on different demonstrators accordingly. Following this line, two selection strategies are proposed, where one is based on an uniform distribution, and the other one is based on a Gaussian distribution.

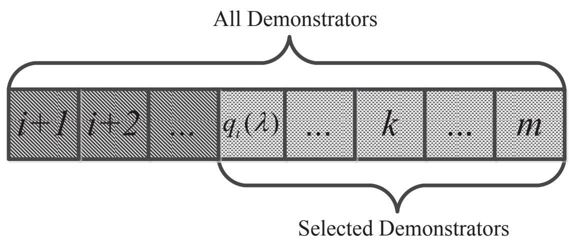


Fig. 3. The proposed Uniform Selection strategy. For the i-th particle, a number of $m - q_i(\lambda) + 1$ demonstrators are selected.

1) The Uniform Selection (US): To relate the selection probability to the demonstrators' fitness values, one simple measure is to directly remove the first several less fitter demonstrators, thus increasing the selection probabilities of the following fitter ones. Based on this idea, an Uniform Selection strategy is proposed, where the indexes of the selected demonstrator follow a uniform distribution defined as:

$$Pr(k|\lambda, i) = \frac{1}{m - q_i(\lambda)}, \tag{8}$$

with

$$q_i(\lambda) = m - \lceil \lambda \times (m - i - 1) \rceil, \tag{9}$$

where i is the imitator index, k is the corresponding demonstrator index, and λ is a parameter that controls the number of demonstrators that can be selected, refer to Fig. 3.

One extreme case in this selection strategy is when $\lambda = 0$, where all imitators learn from the best demonstrator (m-th particle). By contrast, when $\lambda = 1$, this selection strategy will become the same as the one in the original SL-PSO.

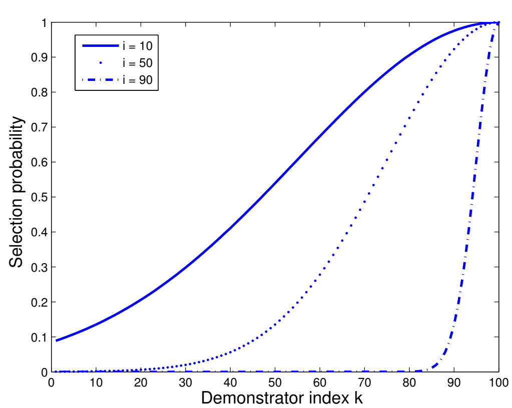


Fig. 4. The corresponding demonstrator selection probabilities (Gaussian distribution) with different imitator $i = 10$, $i = 50$ and $i = 100$. Population size $m = 100$ and control parameter $\theta = 0.5$ is used.

2) The Gaussian Selection (GS): Instead of entirely removing some worse fitting demonstrators, an alternative method is to reduce their selection probabilities. Naturally, a Gaussian distribution can well meet such a requirement. In the SL-PSO, since the population is already sorted according to the fitness values, refer to Fig. 2, the Gaussian distribution can be directly related to the demonstrator indexes as follows:

$$Pr(k|\theta, i) = \frac{1}{\sigma_i(\theta)\sqrt{2\pi}} e^{-\frac{(k-m)^2}{2\sigma_i(\theta)^2}}, \tag{10}$$

with

$$\sigma_i(\theta) = \theta \times (m - i), \tag{11}$$

where i is the imitator index and k is the corresponding demonstrator index, satisfying $i + 1 \leq k \leq m$; θ is a predefined parameter that controls the size of σ_i. A more intuitive observation can be referred to Fig. 4.

It should be noted that, since the Gaussian distribution is defined in $\mathbb{R}$, theoretically, the sampled values are continuously distributed. However, in practice, to ensure the feasibility of the randomly sampled demonstrator indexes, only rounded discrete values in $\{i + 1, i + 2, ..., m\}$ are used.

III. Experimental Results

To examine the performance of the proposed demonstrator selection strategies, a group of numerical experiments are conducted on 12 widely used benchmark functions [15], refer to Tables I, where the first 5 are uni-modal functions and the rest are multi-modal functions. All test functions are 30-dimensional and the results are obtained from 30 independent runs. For each single run, the maximum number of fitness evaluations (FEs) is set to 200,000 as the terminal condition.

Firstly, in order to investigate the settings of the control parameters λ and θ in the Uniform Selection (US) strategy and Gaussian Selection (GS) strategy , respectively, a group of λ and θ values have been tested. Thereafter, to examine the effectiveness of the US and GS, further comparisons have been made with the original SL-PSO and 5 representative PSO variants.

All experiments have been conducted on a PC with an Intel Core i3-2328 2.2GHz CPU and Microsoft Windows 7 Enterprise SP1 64-bit operating system, and the programmes are implemented in Matlab 2010a. It should also be noted that, apart from the parameters introduced in the demonstrator selection strategies, no other specific parameter setting is needed for the SL-PSO to operate.

A. Parameter setting of λ in the US strategy

To adopt the Uniform Selection strategy, it is essential to find a proper setting for λ. To investigate the parameter setting of λ, five different values ($\lambda = 0.1, 0.3, 0.5, 0.7, 0.9$) are tested on the 12 benchmark functions using SL-PSO with the US strategy (denoted as SL-PSO-US).

As shown in Table II, the overall best results are obtained with the setting $\lambda = 0.7$. In comparison, the best results on f_1 and f_2 are obtained with $\lambda = 0.1$, which is the smallest λ value. It seems that the performance of SL-PSO on f_1 and f_2

TABLE I

BENCHMARK TEST FUNCTIONS USED IN THE EXPERIMENTS

Name	Function	Search range
Sphere	$f_1(X) = \sum_{i=1}^{n} x_i^2$	$[-100, 100]^n$
Schwefel 2.22	$f_2(X) = \sum_{i=1}^{n} \lvert x_i \rvert + \prod_{i=1}^{n} x_i$	$[-10, 10]^n$
Schwefel 1.2	$f_3(X) = \sum_{i=1}^{n} (\sum_{j=1}^{i} x_j)^2$	$[-100, 100]^n$
Schwefel 2.21	$f_4(X) = \max \lvert x_i, i \leq i \leq n \rvert$	$[-100, 100]^n$
Rosenbrock	$f_5(X) = \sum_{i=1}^{n-1} (100(x_i^2 - x_{i+1})^2 + (x_i - 1)^2)$	$[-30, 30]^n$
Step	$f_6(X) = \sum_{i=1}^{n} \lfloor x_i + 0.5 \rfloor^2$	$[-100, 100]^n$
Schwefel	$f_7(X) = 418.9829 \cdot n + \sum_{i=1}^{n} -x_i \cdot \sin(\sqrt{\lvert x_i \rvert})$	$[-500, 500]^n$
Rastrigin	$f_8(X) = \sum_{i=1}^{n} (x_i^2 - 10\cos(2\pi x_i) + 10)$	$[-5.12, 5.12]^n$
Ackley	$f_9(X) = -20\exp(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}) - \exp(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)) + 20 + e$	$[-32, 32]^n$
Griewank	$f_{10}(X) = \sum_{i=1}^{n} \frac{x_i^2}{4000} - \prod_{i=1}^{n} \cos(\frac{x_i}{\sqrt{i}}) + 1$	$[-600, 600]^n$
Penalized 1	$f_{11}(X) = \frac{\pi}{n}\left\{10\sin^2(\pi y_1) + \sum_{i=1}^{n-1} (y_i - 1)^2[1 + 10\sin^2(\pi y_{i+1})] + (y_n - 1)^2\right\} + \sum_{i=1}^{n} u(x_i, 10, 100, 4),$ $y = 1 + \frac{1}{4}(x_i + 1)$	$[-50, 50]^n$
Penalized 2	$f_{12}(X) = 0.1\left\{\sin^2(3\pi x_1) + \sum_{i=1}^{n-1} (x_i - 1)^2[1 + \sin^2(3\pi x_{i+1})] + (x_n - 1)^2[1 + \sin^2(2\pi x_n)]\right\} + \sum_{i=1}^{n} u(x_i, 5, 100, 4)$	$[-50, 50]^n$

In f_{11} and f_{12}, $u(x_j, a, k, m) = \begin{cases} k(x_j - a)^m, & x_j > a \\ 0, & -a \leq x_j \leq a \\ k(-x_j - a)^m, & x_j < -a \end{cases}$

All functions are scalable to the search dimension denoted by n.
The global optimum is 0 for all functions.

deteriorates with the increase of λ. This is because the smaller a λ is, the less diverse a population will be, and consequently a faster convergence can be achieved, so that the performance on simple uni-modal functions like f_1 and f_2 can be considerably improved given limited number of fitness evaluations.

However, the drawback of a too small λ can also be reflected from the results on multi-modal functions (f_6 to f_{12}) or even the difficult uni-modal functions (f_3 to f_5): on the one hand, SL-PSO tends to be trapped into local optima (e.g., f_7 and f_8); on the other hand, the performance tends to be unstable (e.g., f_{10} and f_{12}). Therefore, we suggest that the settings of λ not be too small, and a range $\lambda \in (0.5, 1)$ is recommended.

B. Parameter setting of θ in the GS strategy

Similar to the parameter λ in the Uniform Selection strategy, different θ values will cause different selection probability distributions, thus leading to different search behaviors of the SL-PSO. To investigate the parameter setting of θ, five different values ($\theta = 0.2, 0.3, 0.4, 0.5, 0.6$) are tested on the 12 benchmark functions using SL-PSO with the GS strategy (denoted as SL-PSO-GS).

As shown in Table II, the overall best results are obtained with the setting $\theta = 0.4$. In comparison, the best results on f_1 and f_2 are obtained with $\theta = 0.2$, which is the smallest θ value. Similar to the patterns shown in the results obtained with the Uniform Selection strategy, with the increase of θ, SL-PSO's performance on simple uni-modal functions (f_1 and f_2) shows deterioration, but the performance on more difficult functions (f_3 to f_{12}) becomes better.

From the average ranks of the overall results, it can be seen that the differences between the results obtained by different θ settings are not as significant as the results obtained by different λ settings. Normally, the performance on multi-modal functions are closely related to the population diversity. A Gaussian distribution can guarantee a more widely spreading selection range covering more demonstrators, such that more population diversity can be maintained; by contrast, a uniform distribution is only able to cover some of the demonstrators,

TABLE II
THE STATISTICAL RESULTS AND THE RANKS OF OPTIMIZATION ON THE TEST FUNCTIONS USING SL-PSO WITH DIFFERENT λ AND θ VALUES OF THE UNIFORM SELECTION STRATEGY (SL-PSO-US) AND THE GAUSSIAN SELECTION STRATEGY (SL-PSO-GS) RESPECTIVELY. THE BEST MEAN RESULTS ARE HIGHLIGHTED.

		The Uniform Selection Strategy					The Gaussian Selection Strategy				
		$\lambda = 0.1$	$\lambda = 0.3$	$\lambda = 0.5$	$\lambda = 0.7$	$\lambda = 0.9$	$\theta = 0.2$	$\theta = 0.3$	$\theta = 0.4$	$\theta = 0.5$	$\lambda = 0.6$
f_1	Mean	**3.47E-150**	5.65E-133	4.52E-117	3.17E-103	1.27E-90	**1.14E-135**	1.01E-121	6.58E-109	3.14E-95	5.53E-83
	Best	2.46E-151	3.04E-133	1.51E-117	1.21E-104	3.65E-91	6.45E-137	2.32E-122	1.84E-110	2.64E-96	5.69E-84
	Dev	3.99E-150	4.16E-133	3.94E-117	2.96E-103	1.28E-90	1.52E-135	9.25E-122	5.88E-109	2.75E-95	6.02E-83
	Rank	1	2	3	4	5	1	2	3	4	5
f_2	Mean	**9.13E-78**	2.55E-68	2.55E-60	3.11E-53	7.87E-47	**3.46E-70**	5.46E-63	2.78E-56	8.21E-50	1.31E-43
	Best	1.60E-78	7.41E-69	9.68E-61	1.38E-53	3.56E-47	1.05E-70	2.42E-63	9.34E-57	2.83E-50	1.74E-44
	Dev	9.10E-78	1.90E-68	1.30E-60	2.97E-53	4.83E-47	2.21E-70	4.49E-63	3.72E-56	5.36E-50	1.37E-43
	Rank	1	2	3	4	5	1	2	3	4	5
f_3	Mean	6.37E+02	6.19E-01	1.52E-06	**3.09E-07**	2.63E-06	1.11E-04	1.04E-06	**2.23E-08**	2.45E-07	4.37E-05
	Best	2.28E+02	2.24E-03	6.54E-07	1.55E-08	5.06E-07	1.71E-06	4.26E-08	2.94E-09	7.36E-09	1.24E-05
	Dev	2.31E+02	1.23E+00	1.34E-06	1.90E-07	1.39E-06	1.74E-04	1.76E-06	2.89E-08	4.51E-07	4.09E-05
	Rank	5	4	2	1	3	5	3	1	2	4
f_4	Mean	6.89E+00	9.91E-02	3.45E-04	**3.34E-30**	2.19E-25	3.57E-01	4.31E-02	**1.86E-31**	2.06E-27	1.76E-22
	Best	5.30E+00	3.50E-03	3.91E-35	7.28E-31	1.17E-25	8.13E-03	1.76E-15	4.31E-32	3.44E-28	1.01E-22
	Dev	1.36E+00	1.02E-01	7.70E-04	2.20E-30	1.62E-25	5.77E-01	9.28E-02	1.65E-31	1.36E-27	6.84E-23
	Rank	5	4	3	1	2	5	4	1	2	3
f_5	Mean	2.75E+01	4.52E+01	**1.86E+01**	1.94E+01	3.04E+01	3.30E+01	**1.71E+01**	1.80E+01	6.93E+01	4.22E+01
	Best	1.61E+01	1.93E+01	1.82E+01	1.92E+01	1.83E+01	1.70E+01	1.68E+01	1.77E+01	1.84E+01	1.77E+01
	Dev	9.15E+00	3.43E+01	3.47E-01	1.30E-01	2.40E+01	3.35E+01	3.44E-01	2.49E-01	5.01E+01	3.24E+01
	Rank	4	5	1	2	3	3	1	2	5	4
f_6	Mean	6.00E-01	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**
	Best	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
	Dev	5.48E-01	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
	Rank	5	1	1	1	1	1	1	1	1	1
f_7	Mean	2.45E+03	1.41E+03	1.69E+03	1.23E+03	**9.95E+02**	2.07E+03	1.69E+03	1.54E+03	**1.47E+03**	1.51E+03
	Best	1.78E+03	9.28E+02	1.42E+03	8.29E+02	7.11E+02	1.24E+03	1.19E+03	9.28E+02	9.48E+02	1.30E+03
	Dev	4.85E+02	4.96E+02	3.93E+02	2.44E+02	2.15E+02	6.87E+02	3.09E+02	4.58E+02	4.26E+02	1.29E+02
	Rank	5	4	3	2	1	5	4	3	1	2
f_8	Mean	1.97E+01	1.03E+01	6.96E+00	**6.37E+00**	6.77E+00	1.45E+01	1.51E+01	1.17E+01	1.21E+01	**9.55E+00**
	Best	1.59E+01	7.96E+00	4.97E+00	3.98E+00	3.98E+00	1.19E+01	1.39E+01	1.09E+01	8.95E+00	7.96E+00
	Dev	3.61E+00	2.29E+00	1.86E+00	2.29E+00	1.91E+00	1.81E+00	1.09E+00	8.32E-01	2.58E+00	1.51E+00
	Rank	5	4	3	1	2	4	5	2	3	1
f_9	Mean	1.19E-14	7.64E-15	6.22E-15	6.22E-15	**5.51E-15**	6.22E-15	6.22E-15	**5.51E-15**	**5.51E-15**	**5.51E-15**
	Best	6.22E-15	6.22E-15	6.22E-15	6.22E-15	2.66E-15	6.22E-15	6.22E-15	2.66E-15	2.66E-15	2.66E-15
	Dev	3.18E-15	3.18E-15	0.00E+00	0.00E+00	1.59E-15	0.00E+00	0.00E+00	1.59E-15	1.59E-15	1.59E-15
	Rank	5	4	2	2	1	4	4	1	1	1
f_{10}	Mean	7.38E-03	0.00E+00	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**
	Best	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
	Dev	1.28E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
	Rank	5	1	1	1	1	1	1	1	1	1
f_{11}	Mean	2.07E-02	**1.57E-32**	**1.57E-32**	**1.57E-32**	**1.57E-32**	2.07E-02	**1.57E-32**	**1.57E-32**	**1.57E-32**	**1.57E-32**
	Best	1.57E-32	1.57E-32	1.57E-32	1.57E-32	1.57E-32	1.57E-32	1.57E-32	1.57E-32	1.57E-32	1.57E-32
	Dev	4.64E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00	4.64E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00
	Rank	5	1	1	1	1	5	1	1	1	1
f_{12}	Mean	4.39E-03	3.30E-03	2.20E-03	**1.35E-32**	**1.35E-32**	**1.35E-32**	**1.35E-32**	**1.35E-32**	**1.35E-32**	**1.35E-32**
	Best	1.35E-32	1.35E-32	1.35E-32	1.35E-32	1.35E-32	1.35E-32	1.35E-32	1.35E-32	1.35E-32	1.35E-32
	Dev	6.02E-03	5.31E-03	4.91E-03	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
	Rank	5	4	3	1	1	1	1	1	1	1
Ave. Rank		4.25	3	2.2	**1.6**	2	3	2.4	**1.7**	2.2	2.4

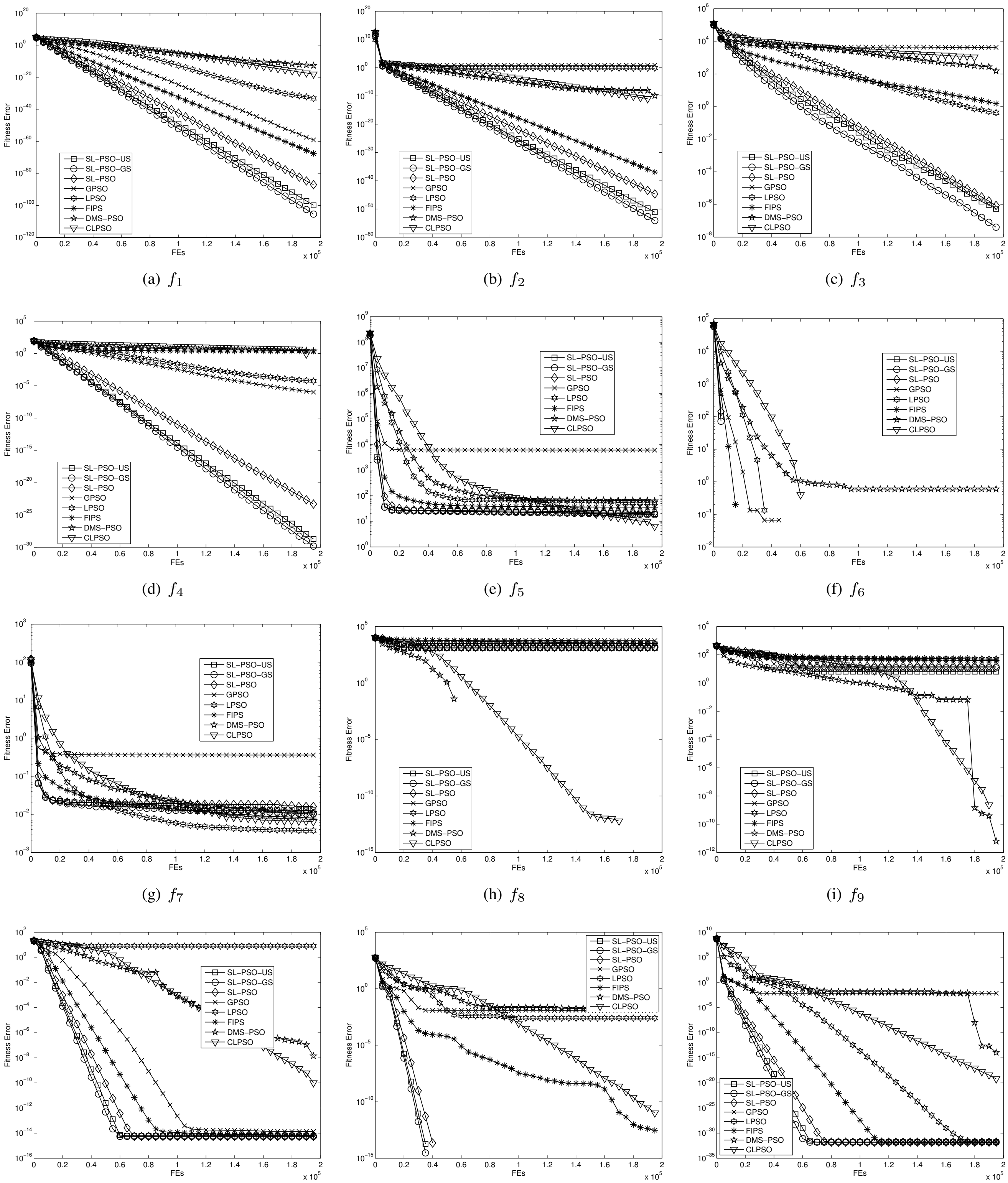
thus causing some loss of population diversity. Therefore, the Gaussian Selection strategy shows more stable performance, regardless of the small differences of θ values.

C. Comparison with PSO variants

In this section, the SL-PSO using the proposed demonstrator selection strategies is compared with the original SL-PSO and five representative PSO variants, including the global version PSO (GPSO) [16], the local version PSO (LPSO) [17], the fully informed PSO (FIPS) [18], the dynamic multi-swarm PSO (DMS-PSO) [19] and the comprehensive learning PSO (CLPSO) [20]. The parameter settings for these PSO variants are summarized in Table III. Empirical settings of $\lambda = 0.7$ and $\theta = 0.4$ are used in the two proposed demonstrator selection strategies respectively.

TABLE III
PARAMETER SETTINGS FOR PSO VARIANTS

Algorithm	Parameter Settings
GPSO	$\omega = 0.9\ 0.4, c_1 = c_2 = 2.0$
LPSO	$\omega = 0.9\ 0.4, c_1 = c_2 = 2.0$
FIPS	$\chi = 0.729, \sum c_i = 4.1$
DMS-PSO	$\omega = 0.729, c_1 = c_2 = 1.49445, m = 3, R = 15$
CLPSO	$\omega = 0.9\ 0.7, c_1 = c_2 = 1.49445$



(a) f_1 (b) f_2 (c) f_3

(d) f_4 (e) f_5 (f) f_6

(g) f_7 (h) f_8 (i) f_9

(j) f_{10} (k) f_{11} (l) f_{12}

Fig. 5. The convergence profiles of the compared algorithms

TABLE IV
THE STATISTICAL RESULTS AND THE RANKS OF OPTIMIZATION ON THE TEST FUNCTIONS USING THE UNIFORM SELECTION STRATEGY (SL-PSO-US) AND THE GAUSSIAN SELECTION STRATEGY (SL-PSO-GS), IN COMPARISON WITH THE ORIGINAL SL-PSO AND PSO VARIANTS. THE BEST MEAN RESULTS ARE HIGHLIGHTED.

		SL-PSO-US	SL-PSO-GS	SL-PSO	GPSO	LPSO	FIPS	DMS-PSO	CLPSO
f_1	Mean	3.2E-103	**6.6E-109**	4.24E-90	1.25E-61	8.48E-35	6.20E-70	3.30E-14	4.76E-19
	Best	1.2E-104	1.8E-110	3.74E-91	1.88E-64	3.40E-40	5.44E-71	3.32E-20	2.01E-19
	Dev	3E-103	5.9E-109	5.26E-90	2.82E-61	2.85E-34	1.44E-69	1.27E-13	1.92E-19
	Rank	2	1	3	5	6	4	8	7
f_2	Mean	3.11E-53	**2.78E-56**	1.5E-46	7.33E+00	6.67E-01	1.13E-38	8.48E-11	7.54E-12
	Best	1.38E-53	9.34E-57	1.02E-46	4.46E-36	3.16E-25	3.94E-39	1.66E-13	2.44E-12
	Dev	2.97E-53	3.72E-56	5.34E-47	1.39E+01	2.58E+00	5.70E-39	1.84E-10	2.50E-12
	Rank	2	1	3	8	7	4	6	5
f_3	Mean	3.09E-07	**2.23E-08**	4.66E-07	4.22E+03	3.65E-01	1.21E+00	9.79E+01	1.13E+03
	Best	1.55E-08	2.94E-09	1.7E-07	3.24E-08	3.97E-02	3.89E-01	2.97E+01	6.68E+02
	Dev	1.9E-07	2.89E-08	2.48E-07	5.08E+03	3.83E-01	6.59E-01	7.31E+01	2.89E+02
	Rank	2	1	3	8	4	5	6	7
f_4	Mean	3.34E-30	**1.86E-31**	1.17E-24	8.49E-07	4.42E-05	2.37E+00	1.90E+00	4.31E+00
	Best	7.28E-31	4.31E-32	4.2E-25	3.20E-08	1.37E-05	6.87E-01	8.17E-01	3.15E+00
	Dev	2.2E-30	1.65E-31	8.37E-25	1.01E-06	2.32E-05	1.17E+00	7.85E-01	6.84E-01
	Rank	2	1	3	4	5	7	6	8
f_5	Mean	1.94E+01	1.80E+01	2.15E+01	6.05E+03	5.18E+01	3.53E+01	5.60E+01	**9.28E+00**
	Best	1.92E+01	1.77E+01	1.86E+01	2.17E-01	4.11E+00	1.11E+01	2.88E-02	9.35E-01
	Dev	1.30E-01	2.49E-01	3.41E+00	2.32E+04	3.68E+01	2.71E+01	3.28E+01	1.03E+01
	Rank	3	2	4	8	6	5	7	1
f_6	Mean	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**	**0.00E+00**	5.33E-01	**0.00E+00**
	Best	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
	Dev	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	9.15E-01	0.00E+00
	Rank	1	1	1	1	1	1	8	1
f_7	Mean	1.23E+03	1.54E+03	1.50E+03	5.62E+03	3.07E+03	2.98E+03	5.74E-08	**6.06E-13**
	Best	8.29E+02	9.28E+02	1.40E+03	3.08E+03	2.02E+03	1.66E+03	5.71E-08	0.00E+00
	Dev	2.44E+02	4.58E+02	9.10E+01	2.19E+03	7.80E+02	7.87E+02	6.02E-10	8.88E-13
	Rank	3	5	4	8	7	6	2	1
f_8	Mean	6.37E+00	1.17E+01	1.55E+01	4.65E+01	5.02E+01	3.86E+01	**2.70E-13**	5.83E-09
	Best	3.98E+00	1.09E+01	1.19E+01	1.49E+01	1.01E+01	1.90E+01	0.00E+00	1.54E-09
	Dev	2.29E+00	8.32E-01	3.19E+00	2.55E+01	2.25E+01	1.04E+01	8.41E-13	5.02E-09
	Rank	3	4	5	7	8	6	1	2
f_9	Mean	6.22E-15	**5.51E-15**	**5.51E-15**	1.36E-14	7.67E+00	6.69E-15	6.11E-09	2.99E-10
	Best	6.22E-15	2.66E-15	2.66E-15	6.22E-15	6.22E-15	6.22E-15	3.13E-11	1.16E-10
	Dev	0.00E+00	1.59E-15	1.59E-15	4.34E-15	9.79E+00	1.83E-15	1.89E-08	9.47E-11
	Rank	3	1	1	5	8	4	7	6
f_{10}	Mean	**0.00E+00**	**0.00E+00**	**0.00E+00**	1.21E-02	2.46E-03	2.07E-13	1.76E-02	8.40E-12
	Best	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.00E-15
	Dev	0.00E+00	0.00E+00	0.00E+00	1.58E-02	6.64E-03	5.03E-13	2.56E-02	1.45E-11
	Rank	1	1	1	7	6	4	8	5
f_{11}	Mean	**1.57E-32**	**1.57E-32**	**1.57E-32**	6.91E-03	**1.57E-32**	**1.57E-32**	9.32E-15	3.61E-20
	Best	1.57E-32	1.57E-32	1.57E-32	1.57E-32	1.57E-32	1.57E-32	1.71E-22	1.57E-20
	Dev	0.00E+00	0.00E+00	0.00E+00	2.68E-02	2.83E-48	2.83E-48	3.61E-14	1.87E-20
	Rank	1	1	1	8	1	1	7	6
f_{12}	Mean	**1.35E-32**	**1.35E-32**	**1.35E-32**	7.32E-04	7.32E-04	**1.35E-32**	1.46E-03	3.31E-19
	Best	1.35E-32	1.35E-32	1.35E-32	1.35E-32	1.35E-32	1.35E-32	2.32E-19	1.84E-19
	Dev	0.00E+00	0.00E+00	0.00E+00	2.84E-03	2.84E-03	2.83E-48	3.87E-03	8.67E-20
	Rank	1	1	1	6	6	1	8	5
Ave. Rank		2	**1.6**	2.5	6.3	5.4	4	6.2	4.5

The statistical results are summarized in Table IV. It can be seen that the SL-PSO with the proposed demonstrator selection strategies shows best overall performance, in comparison with the original SL-PSO and 5 PSO variants. Although the DMS-PSO and CLPSO performs best on f_8 and f_7, respectively, the SL-PSO-GS performs best on 9 out of 12 functions. With the proposed demonstrator selection strategies, it can be seen that, the SL-PSO shows fast convergence on uni-modal functions; on the other hand, the SL-PSO maintains stable performance on multi-modal functions, which can be further confirmed from the convergence profiles shown in Fig. 5.

IV. CONCLUSIONS

This paper has introduced two demonstrator selection strategies in the social learning particle swarm optimizer (SL-PSO). Experiments have been conducted on 12 widely used benchmark functions in comparison with the original SL-PSO and 5 representative PSO variants. The results have shown that, with the proposed demonstrator selection strategies, the performance of SL-PSO has been enhanced.

In future work, more studies on different demonstrator selection strategies might be of interest, and it should also be meaningful to investigate the influence of different demonstrator topological structures on the performance of SL-PSO.

REFERENCES

[1] M. A. Montes de Oca and T. Stützle, "Towards incremental social learning in optimization and multiagent systems," in *Proceedings of the Conference Companion on Genetic and Evolutionary Computation*. ACM, 2008, pp. 1939–1944.

[2] L. Lefebvre, "The opening of milk bottles by birds: evidence for accelerating learning rates, but against the wave-of-advance model of cultural transmission," *Behavioural Processes*, vol. 34, no. 1, pp. 43–53, 1995.

[3] C. Heyes, E. Ray, C. Mitchell, and T. Nokes, "Stimulus enhancement: controls for social facilitation and local enhancement," *Learning and Motivation*, vol. 31, no. 2, pp. 83–98, 2000.

[4] S. Mineka and M. Cook, "Social learning and the acquisition of snake fear in monkeys," *Social learning: Physhological and Biological Perspectives*, pp. 51–73, 1998.

[5] H. P. Young, "Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning," *The American Economic Review*, vol. 99, no. 5, pp. 1899–1924, 2009.

[6] C. F. Bond and L. J. Titus, "Social facilitation: a meta-analysis of 241 studies." *Psychological Bulletin*, vol. 94, no. 2, p. 265, 1983.

[7] T. R. Zentall, "Imitation in animals: evidence, function, and mechanisms," *Cybernetics & Systems*, vol. 32, no. 1-2, pp. 53–96, 2001.

[8] A. Whiten, "Primate culture and social learning," *Cognitive Science*, vol. 24, no. 3, pp. 477–508, 2000.

[9] C. M. Heyes and B. G. Galef Jr, *Social learning in Animals: the Roots of Culture*. Access Online via Elsevier, 1996.

[10] M. V. Flinn, "Culture and the evolution of social learning," *Evolution and Human Behavior*, vol. 18, no. 1, pp. 23–67, 1997.

[11] D. Sperber and H. Caton, *Explaining culture: A naturalistic approach*. Cambridge Univ Press, 1996.

[12] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Joint Conference on Neural Networks*, vol. 4. IEEE, 1995, pp. 1942–1948.

[13] R. Cheng and Y. Jin, "A social learning particle swarm optimization algorithm for scalable optimization," submitted to *Information Sciences*, 2013.

[14] J. Vesterstrom and R. Thomsen, "A comparative study of differential evolution, particle swarm optimization, and evolutionary algorithms on numerical benchmark problems," in *Proceedings of IEEE Congress on Evolutionary Computation*, vol. 2. IEEE, 2004, pp. 1980–1987.

[15] X. Yao, Y. Liu, and G. Lin, "Evolutionary programming made faster," *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 2, pp. 82–102, 1999.

[16] Y. Shi and R. C. Eberhart, "Empirical study of particle swarm optimization," in *Proceedings of IEEE Congress on Evolutionary Computation*. IEEE, 1999, pp. 1945–1950.

[17] J. Kennedy and R. Mendes, "Population structure and particle swarm performance," in *Proceedings of IEEE Congress on Evolutionary Computation*, vol. 2. IEEE, 2002, pp. 1671–1676.

[18] R. Mendes, J. Kennedy, and J. Neves, "The fully informed particle swarm: simpler, maybe better," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 204–210, 2004.

[19] J. Liang and P. Suganthan, "Dynamic multi-swarm particle swarm optimizer," in *Proceedings of IEEE Swarm Intelligence Symposium*. IEEE, 2005, pp. 124–129.

[20] J. J. Liang, A. Qin, P. N. Suganthan, and S. Baskar, "Comprehensive learning particle swarm optimizer for global optimization of multimodal functions," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 3, pp. 281–295, 2006.