A Statistical Study on Parameter Selection of Operators in Continuous State Transition Algorithm

Xiaojun Zhou, Chunhua Yang, *Member, IEEE*, and Weihua Gui, *Member, IEEE*

***Abstract*—The state transition algorithm (STA) has been emerging as a novel metaheuristic method for global optimization over the past few years. In our previous study, the parameter of transformation operator in continuous STA is kept constant or decreasing itself in a periodical way. In this paper, the optimal parameter selection of operators in continuous STA is taken into consideration. First, a statistical study with four benchmark 2-D functions is conducted to show how these parameters affect the search ability of the STA. Based on the experience gained from the statistical study, then, a new continuous STA with optimal parameter strategy is proposed to accelerate its search process. The proposed STA is successfully applied to 12 benchmarks with 20-D, 30-D, and 50-D space. A comparison with other metaheuristics has also demonstrated the effectiveness of the proposed method.**

***Index Terms*—Global optimization, metaheuristic, state transition algorithm (STA), statistical study.**

I. Introduction

THE STATE transition algorithm (STA) [1], [2] is a recently emerging metaheuristic method for global optimization and has found applications in nonlinear system identification and control [3], water distribution networks configuration [4]; sensor network localization [5]; PID controller design [6], [7]; overlapping peaks resolution [8]; image segmentation [9]; wind power prediction [10]; dynamic optimization [11], [12]; bi-level optimization [13]; and modeling and control of complex industrial processes [14]–[19], etc. In STA, a solution to an optimization problem is considered as a state, and an update of a solution can be regarded as a state transition. Unlike the population-based evolutionary algorithms [20]–[22], the standard STA is an individual-based optimization method. Based on an incumbent best solution, a neighborhood with special characteristics will be formed automatically when using a certain state transformation operator. A variety of state transformation operators, for example, rotation, translation, expansion, and axesion in continuous STA, or swap, shift, symmetry and substitute in discrete STA, are designed purposely for both global and local search. On the basis of the neighborhood, then, a sampling technique is used to generate a candidate set, and the next best solution is updated by using a selection technique based on the previous best solution and the candidate set. This process is repeated using state transformation operators alternatively until some terminal conditions are satisfied.

In this paper, the continuous STA is studied. As aforementioned, in continuous STA, there are four state transformation operators, and each transformation operator has certain geometric significance, i.e., the neighborhood formed by each transformation operator has certain geometric characteristic. To be more specific, the rotation transformation has the functionality to search in a hypersphere with the maximal radius α, called rotation factor; the translation transformation has the functionality to search along a line with the maximal length β, called translation factor; the expansion transformation has the functionality to search in a broader space controlled by the expansion factor γ; and the axesion transformation is designed to strengthen single-dimensional search regulated by the axesion factor δ. In our previous studies, the rotation factor is exponentially decreasing from a maximum value to a minimum value in a periodic way, and other transformation factors are kept constant at one [1]. To gain a better exploitation ability, all state transformation factors are exponentially decreasing from a maximum value to a minimum value in a periodic way in [5].

As is known to us, there exist several parameters in metaheuristic methods and parameter selection plays a significant role in their performance. For instance, crossover and mutation probability in genetic algorithms [23], inertia weight and acceleration factors in particle swarm optimization [24], [25], amplification factor and crossover rate in differential evolution [26]–[28], and neighborhood radius in artificial bee colony (ABC) [29]. In general, the parameter setting can be summarized to two types: parameter tuning and parameter control. The former is to find good parameter values before running these algorithms, and they remain fixed during the run. On the contrary, the later is to update parameter values in the process, and the types of update mechanisms can

Manuscript received December 31, 2017; revised April 11, 2018 and May 18, 2018; accepted June 21, 2018. Date of publication July 18, 2018; date of current version June 27, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61503416, Grant 61533021, Grant 61621062, and Grant 61725306, in part by the Innovation-Driven Plan in Central South University under Grant 2018CX12, in part by the 111 Project under Grant B17048, and in part by the Hunan Provincial Natural Science Foundation of China under Grant 2018JJ3683. This paper was recommended by Associate Editor H. Li. *(Corresponding author: Chunhua Yang.)*

The authors are with the School of Information Science and Engineering, Central South University, Changsha 410083, China (e-mail: michael.x.zhou@csu.edu.cn; ychh@csu.edu.cn; gwh@csu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCYB.2018.2850350

be deterministic, adaptive, or self-adaptive (for details, please refer to [30]–[32]).

To gain a better understanding of how the parameters of transformation operators in continuous STA affecting its performance, the parameter selection in continuous STA is focused in this paper. With four commonly used benchmark functions as cases, several properties of the operator parameters are observed from a statistical study. With the gained experience from the statistical results, a new continuous STA with optimal operator parameter selection strategy is proposed, and the proposed STA is successfully applied to other benchmarks with higher dimensions.

The remainder of this paper is organized as follows. In Section II, the standard continuous STA are described. Section III gives a statistical study to show how the operator parameters in continuous STA affecting its performance. The proposed STA with optimal operator parameter selection strategy is given in Section IV. In Section V, experimental results are given to testify the effectiveness of the proposed STA. Finally, the conclusion is drawn in Section VI.

II. Standard Continuous State Transition Algorithm

Consider the following continuous optimization problem with simple constraints:

$$\min_{\boldsymbol{x}\in\Omega} f(\boldsymbol{x}) \tag{1}$$

where $\Omega \subseteq \mathbb{R}^n$ is a closed and compact set, which is usually composed of lower and upper bounds of $\boldsymbol{x}$, i.e., $\Omega = \{\boldsymbol{x} \in \mathbb{R}^n | \underline{x_i} \le x_i \le \overline{x_i}, i = 1, \dots, n\}$.

In classical iterative methods for numerical optimization, a new candidate is generated based on a previous solution by using different optimization operators. In a state transition way, a solution can be regarded as a state, and an update of a solution can be considered as a state transition. On the basis of state space representation, the unified form of generation of solution in STA can be described as follows:

$$\begin{cases} \boldsymbol{s}_{k+1} = A_k \boldsymbol{s}_k + B_k \boldsymbol{u}_k \\ y_{k+1} = f(\boldsymbol{s}_{k+1}) \end{cases} \tag{2}$$

where $\boldsymbol{s}_k$ and $\boldsymbol{s}_{k+1}$ stand for a current state and the next state, respectively, corresponding to solutions of the optimization problem; $\boldsymbol{u}_k$ is a function of $\boldsymbol{s}_k$ and historical states; y_k is the fitness value at $\boldsymbol{s}_k$; A_k and B_k are state transition matrices, which can be considered as transformation operators; and f is the objective function or fitness function.

A. State Transformation Operators

Using state space representation and state transformation for reference, four special state transformation operators are designed to generate candidate solutions for an optimization problem [1], [33], and they are listed as follows.

1) Rotation Transformation:

$$\boldsymbol{s}_{k+1} = \boldsymbol{s}_k + \alpha \frac{1}{n\|\boldsymbol{s}_k\|_2} R_r \boldsymbol{s}_k \tag{3}$$

where α is a positive constant, called the rotation factor; $R_r \in \mathbb{R}^{n\times n}$, is a random matrix with its entries being uniformly distributed random variables defined on the interval $[-1, 1]$, and $\|\cdot\|_2$ is the L2-norm (or Euclidean norm) of a vector. This rotation transformation has the functionality to search in a hypersphere with the maximal radius α, which has been testified. The rotation transformation is designed for local search and can be used to guarantee local optimality and manipulate solution accuracy.

2) Translation Transformation:

$$\boldsymbol{s}_{k+1} = \boldsymbol{s}_k + \beta R_t \frac{\boldsymbol{s}_k - \boldsymbol{s}_{k-1}}{\|\boldsymbol{s}_k - \boldsymbol{s}_{k-1}\|_2} \tag{4}$$

where β is a positive constant, called the translation factor; $R_t \in \mathbb{R}$ is a uniformly distributed random variable defined on the interval [0,1]. It is not difficult to understand that the translation transformation has the functionality to search along a line from s_{k-1} to s_k at the starting point s_k with maximum length β. The translation operator is actually a line search, and it can be considered as a heuristic operator since there exists a possible better solution along the line if s_k is better than s_{k-1}.

3) Expansion Transformation:

$$\boldsymbol{s}_{k+1} = \boldsymbol{s}_k + \gamma R_e \boldsymbol{s}_k \tag{5}$$

where γ is a positive constant, called the expansion factor; $R_e \in \mathbb{R}^{n\times n}$ is a random diagonal matrix with its entries obeying the Gaussian distribution (or normal distribution). In the standard STA, the mean equals zero and standard deviation equals one, i.e., the standard normal distribution is used. The expansion transformation has the functionality to search in the whole space in probability, and it is designed for global search.

4) Axesion Transformation:

$$\boldsymbol{s}_{k+1} = \boldsymbol{s}_k + \delta R_a \boldsymbol{s}_k \tag{6}$$

where δ is a positive constant, called the axesion factor; $R_a \in \mathbb{R}^{n\times n}$ is a random diagonal matrix with its entries obeying the Gaussian distribution and only one random position having nonzero value. The axesion transformation is designed to search along the axes, aiming to strengthen single-dimensional search [34].

B. Sampling Technique

The idea of sampling incorporated in continuous STA was first illustrated in [35]. It is found that for a given solution, a neighborhood will be automatically formed. To avoid enumerating all possible candidate solutions, representative samples can be used to reflect the characteristics of the neighborhood. Taking the rotation transformation for example, when independently executing the rotation operator for *SE* times, a total number of *SE* samples are generated in pseudocode as follows:

1: **for** $i \leftarrow 1$, *SE* **do**
2: State(:, i) $\leftarrow$ Best $+ \alpha \frac{1}{n\|\text{Best}\|_2} R_r$Best
3: **end for**

where Best is the incumbent best solution, and *SE* samples are stored in the matrix state.

C. Update Strategy

As mentioned above, based on the incumbent best solution, a total number of *SE* candidate solutions are generated, but it should be noted that these candidate solutions do not always belong to the domain Ω. To address this issue, these samples are projected into Ω through

$$x_i = \begin{cases} \overline{x_i}, & \text{if } x_i > \overline{x_i} \\ \underline{x_i}, & \text{if } x_i < \underline{x_i} \\ x_i, & \text{otherwise.} \end{cases} \tag{7}$$

As a result, the candidate solutions can be guaranteed to be always feasible. Next, a new best solution is selected from the candidate set by virtue of the fitness function, denoted as newBest. Finally, an update strategy based on greedy criterion is used to update the incumbent best as shown in the following equation:

$$\text{Best} = \begin{cases} \text{newBest}, & \text{if} \quad f(\text{newBest}) < f(\text{Best}) \\ \text{Best}, & \text{otherwise.} \end{cases} \tag{8}$$

D. Algorithm Procedure of the Standard Continuous STA

With the state transformation operators for both local and global search, sampling technique for time-saving and update strategy for convergence, the standard continuous STA can be described by the following pseudocodes:

State $\leftarrow$ initialization(SE,Ω)
Best $\leftarrow$ fitness(funfcn,State)
repeat
    if $\alpha < \alpha_{\min}$ **then**
        $\alpha \leftarrow \alpha_{\max}$
    end if
    Best $\leftarrow$ expansion(funfcn,Best,SE,β,γ)
    Best $\leftarrow$ rotation(funfcn,Best,SE,β,α)
    Best $\leftarrow$ axesion(funfcn,Best,SE,β,δ)
    $\alpha \leftarrow \frac{\alpha}{fc}$
until the specified termination criterion is met

As for detailed explanations, rotation($\cdot$) in above pseudocode is given for illustration purposes as follows:

oldBest $\leftarrow$ Best
fBest $\leftarrow$ feval(funfcn,oldBest)
State $\leftarrow$ op_rotate(Best,SE,α)
[newBest,fnewBest] $\leftarrow$ fitness(funfcn,State)
if fnewBest $<$ fBest **then**
    fBest $\leftarrow$ fnewBest
    Best $\leftarrow$ newBest
    State $\leftarrow$ op_translate(oldBest,newBest,SE,β)
    [newBest,fnewBest] $\leftarrow$ fitness(funfcn,State)
    if fnewBest $<$ fBest **then**
        fBest $\leftarrow$ fnewBest
        Best $\leftarrow$ newBest
    end if
end if

As shown in the above pseudocodes, initialization ($\cdot$) is used to make sure the initial solution is in the range Ω. The rotation factor α is decreasing periodically from a maximum value $\alpha_{\max}$ to a minimum value $\alpha_{\min}$ in an exponential way with base fc, which is called lessening coefficient. op_rotate($\cdot$) and op_translate($\cdot$) represent the implementations of proposed sampling technique for rotation and translation operators, respectively, and fitness($\cdot$) represents the implementation of selecting the new best solution from *SE* samples. It should be emphasized that the translation operator is only executed when a solution better than the incumbent best solution can be found in the *SE* samples from rotation, expansion or axesion transformation. In the standard continuous STA, the parameter settings are given as follows: $\alpha_{\max} = 1, \alpha_{\min} = $ 1e-4, $\beta = 1, \gamma = 1, \delta = 1, SE = 30,$ and $fc = 2$.

III. Statistical Study of the State Transformation Factors

As described in Section II, in the standard continuous STA, the state transformation factors like expansion factor γ, axesion factor δ are kept constant, and rotation factor α is decreasing periodically from a maximum value $\alpha_{\max}$ to a minimum value $\alpha_{\min}$ in an exponential way. In order to select the values of these parameters in a more effective manner, a statistical study of the state transformation factors is carried out to investigate the effect of parameter selection on the performance of state transition operators.

Four well-known benchmark functions are listed below.

1) Spherical function

$$f_1 = \sum_{i=1}^{n} x_i^2$$

where the global optimum $\boldsymbol{x}^* = (0, \ldots, 0)$ and $f(\boldsymbol{x}^*) = 0$, $-100 \le x_i \le 100, i = 1, \ldots, n$.

2) Rosenbrock function

$$f_2 = \sum_{i=1}^{n} \left(100\left(x_{i+1} - x_i^2\right)^2 + (x_i - 1)^2\right)$$

where the global optimum $\boldsymbol{x}^* = (1, \ldots, 1)$ and $f(\boldsymbol{x}^*) = 0$, $-30 \le x_i \le 30, i = 1, \ldots, n$.

3) Rastrigin function

$$f_3 = \sum_{i=1}^{n} \left(x_i^2 - 10\cos(2\pi x_i) + 10\right)$$

where the global optimum $\boldsymbol{x}^* = (0, \ldots, 0)$ and $f(\boldsymbol{x}^*) = 0$, $-5.12 \le x_i \le 5.12, i = 1, \ldots, n$.

4) Griewank function

$$f_4 = \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i}^{n} \cos|\frac{x_i}{\sqrt{i}}| + 1$$

where the global optimum $\boldsymbol{x}^* = (0, \ldots, 0)$ and $f(\boldsymbol{x}^*) = 0$, $-600 \le x_i \le 600, i = 1, \ldots, n$.

For a given solution Best_0, three state transition operators (rotation, expansion, and axesion) are performed, respectively, for *SE* times (yielding *SE* samples) independently on each benchmark function using different values of state transformation factors. To be more specific, there are five groups of given initial solutions, i.e., $\text{Best}_0 = (0.01, 0.01), (0.1, 0.1), (0.5, 0.5), (0.9, 0.9), (0.99, 0.99)$; the total number of samples is set at $SE = $ 1e6; and the value of state transformation operators is chosen from the set $\Omega = \{1,$ 1e-1, 1e-2, 1e-3, 1e-4, 1e-5, 1e-6, 1e-7, 1e-8$\}$.

To evaluate the influence of the parameter selection on the performance of state transformation operators, the following

TABLE I
STATISTICAL RESULTS OF SUCCESS RATE AND DESCENT RATE FOR THE ROTATION TRANSFORMATION (SPHERICAL PROBLEM)

Best_0	Index	α = 1	α = 0.1	α = 0.01	α = 1e-3	α = 1e-4	α = 1e-5	α = 1e-6	α = 1e-7	α = 1e-8
$(0.01, 0.01)$	ρ_s	0.0012	0.0919	0.4550	0.4953	0.5002	0.5000	0.4991	0.5001	0.4999
	ρ_d	5.0448e-1	5.0347e-1	2.7785e-1	3.2465e-2	3.2970e-3	3.2960e-4	3.2990e-5	3.2923e-6	3.3056e-7
$(0.1, 0.1)$	ρ_s	0.0921	0.4555	0.4956	0.4990	0.4996	0.4993	0.4996	0.5001	0.5012
	ρ_d	5.0313e-1	2.7801e-1	3.2470e-2	3.2949e-3	3.3048e-4	3.3026e-5	3.2996e-6	3.2944e-7	3.2962e-8
$(0.5, 0.5)$	ρ_s	0.4144	0.4899	0.4992	0.5000	0.4999	0.5006	0.4989	0.5008	0.5003
	ρ_d	4.4908e-1	6.3938e-2	6.5821e-3	6.5914e-4	6.5918e-5	6.6006e-6	6.5986e-7	6.5950e-8	6.6128e-9
$(0.9, 0.9)$	ρ_s	0.4498	0.4947	0.4994	0.5002	0.5000	0.4994	0.4995	0.4996	0.4989
	ρ_d	3.0238e-1	3.6061e-2	3.6623e-3	3.6555e-4	3.6682e-5	3.6639e-6	3.6656e-7	3.6734e-8	3.6637e-9
$(0.99, 0.99)$	ρ_s	0.4543	0.4955	0.4986	0.5000	0.4995	0.4999	0.5001	0.5008	0.4999
	ρ_d	2.7968e-1	3.2731e-2	3.3249e-3	3.3378e-4	3.3307e-5	3.3289e-6	3.3337e-7	3.3349e-8	3.3323e-9

TABLE II
STATISTICAL RESULTS OF SUCCESS RATE AND DESCENT RATE FOR THE ROTATION TRANSFORMATION (ROSENBROCK PROBLEM)

Best_0	Index	α = 1	α = 0.1	α = 0.01	α = 1e-3	α = 1e-4	α = 1e-5	α = 1e-6	α = 1e-7	α = 1e-8
$(0.01, 0.01)$	ρ_s	0.0700	0.2665	0.4766	0.4985	0.5001	0.4997	0.5005	0.4991	0.4999
	ρ_d	2.9603e-1	4.3729e-2	6.1371e-3	6.6152e-4	6.6718e-5	6.6619e-6	6.6708e-7	6.6658e-8	6.6683e-9
$(0.1, 0.1)$	ρ_s	0.1891	0.5082	0.5019	0.4996	0.5003	0.5000	0.5005	0.4993	0.4999
	ρ_d	3.9041e-1	2.2187e-1	2.6801e-2	2.7281e-3	2.7272e-4	2.7264e-5	2.7332e-6	2.7328e-7	2.7319e-8
$(0.5, 0.5)$	ρ_s	0.3884	0.5049	0.5005	0.5003	0.5003	0.5004	0.4996	0.5010	0.4997
	ρ_d	6.4979e-1	2.3369e-1	2.5464e-2	2.5633e-3	2.5620e-4	2.5615e-5	2.5619e-6	2.5595e-7	2.5623e-8
$(0.9, 0.9)$	ρ_s	0.1121	0.5006	0.4999	0.5004	0.5001	0.4999	0.5005	0.4994	0.4996
	ρ_d	6.4261e-1	6.3423e-1	1.0225e-1	1.0622e-2	1.0642e-3	1.0650e-4	1.0645e-5	1.0649e-6	1.0633e-7
$(0.99, 0.99)$	ρ_s	0.0052	0.1146	0.5004	0.5003	0.4996	0.4992	0.5003	0.4994	0.5003
	ρ_d	5.1273e-1	6.5350e-1	6.2945e-1	1.0034e-1	1.0419e-2	1.0446e-3	1.0448e-4	1.0457e-5	1.0441e-6

TABLE III
STATISTICAL RESULTS OF SUCCESS RATE AND DESCENT RATE FOR THE ROTATION TRANSFORMATION (RASTRIGIN PROBLEM)

Best_0	Index	α = 1	α = 0.1	α = 0.01	α = 1e-3	α = 1e-4	α = 1e-5	α = 1e-6	α = 1e-7	α = 1e-8
$(0.01, 0.01)$	ρ_s	0.0012	0.0918	0.4552	0.4948	0.4993	0.5000	0.5003	0.5000	0.4995
	ρ_d	5.0017e-1	5.0278e-1	2.7796e-1	3.2470e-2	3.2959e-3	3.3012e-4	3.2908e-5	3.3004e-6	3.3061e-7
$(0.1, 0.1)$	ρ_s	0.0932	0.4606	0.4960	0.4992	0.5000	0.4995	0.4997	0.4996	0.4998
	ρ_d	4.9612e-1	2.7255e-1	3.1446e-2	3.1888e-3	3.1848e-4	3.1876e-5	3.1924e-6	3.1915e-7	3.1902e-8
$(0.5, 0.5)$	ρ_s	0.9999	0.9926	0.6778	0.5173	0.5014	0.5004	0.4994	0.5000	0.5008
	ρ_d	4.2649e-1	8.0885e-3	1.3463e-4	8.6621e-6	8.1893e-7	8.1569e-8	8.1507e-9	8.1457e-10	8.1412e-11
$(0.9, 0.9)$	ρ_s	0.0849	0.4581	0.4950	0.5002	0.5003	0.5004	0.5005	0.5001	0.4998
	ρ_d	3.1378e-1	1.8020e-1	2.1018e-2	2.1261e-3	2.1318e-4	2.1297e-5	2.1310e-6	2.1274e-7	2.1323e-8
$(0.99, 0.99)$	ρ_s	0.0003	0.0266	0.4140	0.4903	0.4986	0.4995	0.5005	0.4998	0.4999
	ρ_d	2.5399e-3	2.4397e-3	2.1969e-3	3.1457e-4	3.2336e-5	3.2384e-6	3.2463e-7	3.2484e-8	3.2424e-9

two indexes are introduced:

$$\rho_s = \frac{N_s}{SE} \tag{9}$$

$$\rho_d = \frac{|\text{ave} - \text{fBest}_0|}{|\text{fBest}_0|} \tag{10}$$

where ρ_s and ρ_d are called success rate and descent rate, respectively. N_s is the number of samples whose objective function values are smaller than that of the Best_0. ave is the average function value of the N_s samples, and fBest_0 represents the function value for Best_0.

The statistical results for different values of state transformation factors can be found from Tables I–IV.

As indicated in these tables, the following properties can be observed.

1) As the decrease of a state transformation factor below a certain threshold, the descent rate ρ_d shows a declining trend.
2) The success rate remains almost steadily high if a state transformation factor is below a certain threshold.
3) The success rate of the rotation transformation is not high until the rotation factor is below a threshold when current solution Best_0 is approaching the global optimal solution.

To be more specific, let us take the rotation transformation for example, the changes of success rate ρ_s and descent rate ρ_d with the rotation factor α are illustrated in Fig. 1 when current solution Best_0 is approaching the global optimum. Here, Best_0 equals to (0.01, 0.01), (0.99, 0.99), (0.01, 0.01), and (0.01, 0.01) for f_1, f_2, f_3, and f_4, respectively. By taking a closer look at these two figures, it is not difficult to find that there exists a tradeoff between the success rate and the descent rate. For instance, when $\alpha = 1$, the success rate ρ_s is quite low, while the descent rate ρ_d is quite high. On the contrary, when $\alpha \in$ {1e-5, 1e-6, 1e-7, 1e-8}, the success rate ρ_s is quite high, while the descent rate ρ_d is quite low.

TABLE IV
STATISTICAL RESULTS OF SUCCESS RATE AND DESCENT RATE FOR THE ROTATION TRANSFORMATION (GRIEWANK PROBLEM)

Best_0	Index	α = 1	α = 0.1	α = 0.01	α = 1e-3	α = 1e-4	α = 1e-5	α = 1e-6	α = 1e-7	α = 1e-8
$(0.01, 0.01)$	ρ_s	0.0014	0.0967	0.4677	0.4966	0.4993	0.4996	0.5000	0.5007	0.5003
	ρ_d	5.0031e-1	5.0462e-1	2.8800e-1	3.4337e-2	3.4921e-3	3.4973e-4	3.4962e-5	3.4957e-6	3.5030e-7
$(0.1, 0.1)$	ρ_s	0.0962	0.4676	0.4975	0.4997	0.5001	0.4992	0.5003	0.4995	0.4999
	ρ_d	5.0587e-1	2.8807e-1	3.4234e-2	3.4888e-3	3.4825e-4	3.4831e-5	3.4904e-6	3.4869e-7	3.4899e-8
$(0.5, 0.5)$	ρ_s	0.4250	0.4937	0.4993	0.4998	0.4996	0.5003	0.5000	0.5007	0.5006
	ρ_d	4.4869e-1	6.4092e-2	6.5932e-3	6.6154e-4	6.6073e-5	6.6104e-6	6.6212e-7	6.6257e-8	6.6127e-9
$(0.9, 0.9)$	ρ_s	0.4532	0.4958	0.4994	0.5005	0.4998	0.4995	0.4996	0.5002	0.4994
	ρ_d	2.7822e-1	3.1695e-2	3.2066e-3	3.2065e-4	3.2064e-5	3.2016e-6	3.2054e-7	3.2012e-8	3.2048e-9
$(0.99, 0.99)$	ρ_s	0.4550	0.4958	0.4999	0.5003	0.5002	0.4996	0.4999	0.5003	0.5002
	ρ_d	2.4931e-1	2.7546e-2	2.7818e-3	2.7817e-4	2.7754e-5	2.7808e-6	2.7839e-7	2.7824e-8	2.7801e-9

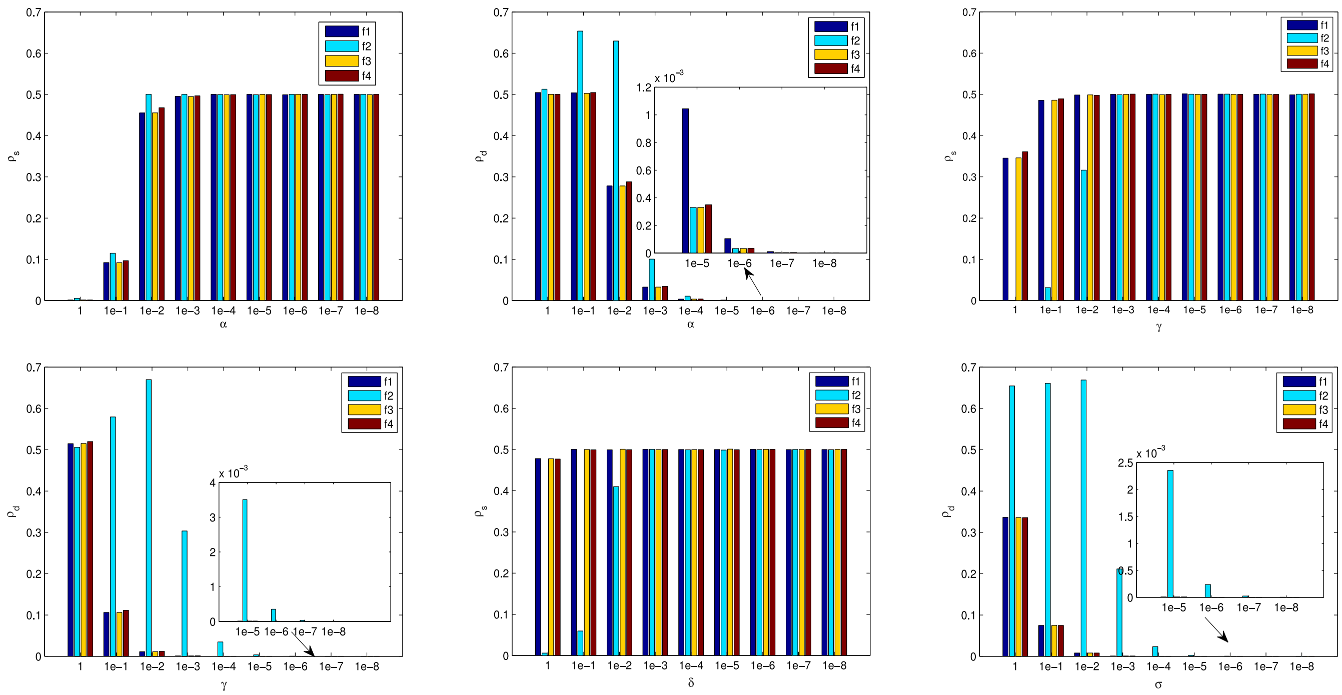

Fig. 1. Changes of success rate ρ_s and descent rate ρ_d with α, γ, and δ, respectively, when approaching the global minima.

Remark 1: Property 3 can provide additional support to the way in changing the rotation factor in the standard continuous STA, i.e., α is not kept constant but decreasing periodically from a maximum value $\alpha_{\max}$ to a minimum value $\alpha_{\min}$. Anyway, it is obvious that the way in changing the state transformation factors is not in an optimal manner.

IV. STATE TRANSITION ALGORITHM WITH OPTIMAL PARAMETER SELECTION

As inspired by the statistical study of the state transformation factors, in this section, an optimal parameter selection strategy is proposed to accelerate the search of the standard continuous STA.

A. Optimal Parameter Selection for the State Transformation Factors

In classical iterative methods for numerical optimization, the following iterative formula is usually adopted:

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + a_k \boldsymbol{d}_k \tag{11}$$

where $\boldsymbol{d}_k$ is the search direction, and a_k is the step size. For gradient-based algorithms, the search direction is relevant to the gradient of current iterative point, for instance, the steepest descent method, $\boldsymbol{d}_k = -\nabla f(\boldsymbol{x}_k)$, and the step size a_k is often restricted to the range [0,1]. It can be found that the pattern of iterative formula in continuous STA is similar to that of (11), as shown in the following equation:

$$\left.\begin{array}{l}\frac{1}{n\|\boldsymbol{x}_k\|_2}R_r \boldsymbol{s}_k\\ R_t\frac{\boldsymbol{s}_k-\boldsymbol{s}_{k-1}}{\|\boldsymbol{s}_k-\boldsymbol{s}_{k-1}\|_2}\\ R_e \boldsymbol{s}_k\\ R_a \boldsymbol{s}_k\end{array}\right\}\Rightarrow \tilde{\boldsymbol{d}}_k,\qquad \left.\begin{array}{c}\alpha\\ \beta\\ \gamma\\ \delta\end{array}\right\}\Rightarrow \tilde{a}_k$$

and a big difference is that the search direction is not determined. Compared with gradient-based algorithms, the STA can be used for global optimization lies in at least two aspects: 1) the search is in all directions and 2) the search can go to any length. While compared with the traditional trust region method, the similarity is that some parts of the STA (except the translation transformation) can be considered as a special kind of trust region method, but the differences are: 1) the STA utilizes the original function not its quadratic approximation information and 2) the search direction in STA is stochastic.

For rotation and translation transformation, the search zone is restricted in a hypersphere or along a line, which are controlled by the corresponding transformation factors. For expansion and axesion transformation, although the search zone can be expanded to the whole space in probability due to the Gaussian distribution, the search zone is restricted and manipulated by the expansion and axesion factors as well. That is to say, in practical numerical computation, the neighborhood formed by the state transformation operators is controlled by the transformation factors to a large extent, which are also testified by the statistical study. To simplify the parameter selection and accelerate the search process, the values of these parameters are all taken from the set Ω = {1, 1e-1, 1e-2, 1e-3, 1e-4, 1e-5, 1e-6, 1e-7, 1e-8}, and the parameter value with the corresponding smallest objective function value is chosen.

B. Proposed STA

Let us denote the optimal parameter as $\tilde{a}^*$, and then we have

$$\tilde{a}^* = \arg\min_{\tilde{a}_k \in \Omega} f\left(\boldsymbol{x}_k + \tilde{a}_k \tilde{\boldsymbol{d}}_k\right). \tag{12}$$

In theory, the neighborhood formed by the state transformation operators has infinite candidate solutions; however, only SE samples are used for evaluation in practice. That is to say, for a given parameter value, only SE samples are taken into consideration. In order to further utilize the parameter more completely, the selected parameter value is kept for a period of time, denoted as T_p. To be more specific, the detailed procedures of the proposed STA can be outlined as follows:

```
repeat
    Best ← expansion_w(funfcn,Best,SE,Ω)
    Best ← rotation_w(funfcn,Best,SE,Ω)
    Best ← axesion_w(funfcn,Best,SE,Ω)
until the specified termination criterion is met
```

In the meanwhile, rotation_w($\cdot$) in above pseudocode is given for further explanations

```
[Best,α] ← update_alpha(funfcn,Best,SE,Ω)
for i ← 1, T_p do
    Best ← rotation(funfcn,Best,SE,α)
end for
```

where the function update_alpha represents the implementation of selection the optimal parameter value of rotation factor. The proposed STA differs from the standard STA in threefold.

1) The periodical way of diminishing the transformation factors is no longer used.
2) The optimal parameter is selected for every state transformation except the translation operator.
3) The optimal parameter is kept to utilize for a period of time.

V. Experimental Results

In order to testify the effectiveness of the proposed STA, the following additional benchmark functions are used for test.

5) Ackley function

$$f_5(\boldsymbol{x}) = 20 + e - 20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\right) - \exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)\right)$$

where the global optimum $\boldsymbol{x}^* = (0, \ldots, 0)$ and $f(\boldsymbol{x}^*) = 0$, $-32 \le x_i \le 32$, $i = 1, \ldots, n$.

6) High conditioned elliptic function

$$f_6(\boldsymbol{x}) = \sum_{i=1}^{n}\left(10^6\right)^{\frac{i-1}{n-1}} x_i^2$$

where the global optimum $\boldsymbol{x}^* = (0, \ldots, 0)$ and $f(\boldsymbol{x}^*) = 0$, $-100 \le x_i \le 100$, $i = 1, \ldots, n$.

7) Michalewicz function

$$f_7(\boldsymbol{x}) = -\sum_{i=1}^{n}\sin(x_i)\sin\left(\frac{ix_i^2}{\pi}\right)^{20}$$

where the global optimum is unknown, $0 \le x_i \le \pi$, $i = 1, \ldots, n$.

8) Trid function

$$f_8(\boldsymbol{x}) = \sum_{i=1}^{n}(x_i - 1)^2 - \sum_{i=2}^{n} x_i x_{i-1}$$

where the global optimum $x_i^* = i(n+1-i)$ and $f(\boldsymbol{x}^*) = -\frac{n(n+4)(n-1)}{6}$, $-n^2 \le x_i \le n^2$, $i = 1, \ldots, n$.

9) Schwefel function

$$f_9(\boldsymbol{x}) = \sum_{i=1}^{n}\left[-x_i\sin\left(\sqrt{|x_i|}\right)\right]$$

where the global optimum $\boldsymbol{x}^* = (420.9687, \ldots, 420.9687)$ and $f(\boldsymbol{x}^*) = -418.9829n$, $-500 \le x_i \le 500$, $i = 1, \ldots, n$.

10) Schwefel 1.2 function

$$f_{10}(\boldsymbol{x}) = \sum_{i=1}^{n}\left(\sum_{j=1}^{i} x_j\right)^2$$

where the global optimum $\boldsymbol{x}^* = (0, \ldots, 0)$ and $f(\boldsymbol{x}^*) = 0$, $-100 \le x_i \le 100$, $i = 1, \ldots, n$.

11) Schwefel 2.4 function

$$f_{11}(\boldsymbol{x}) = \sum_{i=1}^{n}\left[(x_i - 1)^2 + \left(x_1 - x_i^2\right)^2\right]$$

where the global optimum $\boldsymbol{x}^* = (1, \ldots, 1)$ and $f(\boldsymbol{x}^*) = 0$, $0 \le x_i \le 10$, $i = 1, \ldots, n$.

TABLE V
COMPARISONS AMONG VARIOUS ALGORITHMS ON TEST FUNCTIONS

Fcn	Dim	GL-25	CLPSO	SaDE	ABC	Standard STA	Proposed STA
f_1	20	2.5523e-10 ± 1.5883e-10	6.2546e-43 ± 1.4407e-42	6.3533e-188 ± 0	2.7287e-16 ± 6.1809e-17	0 ± 0	0 ± 0
	30	1.7872e-8 ± 1.0381e-8	1.9944e-40 ± 1.8175e-40	5.5498e-184 ± 0	5.6618e-16 ± 7.3169e-17	0 ± 0	0 ± 0
	50	2.3336e-6 ± 1.5613e-6	8.3697e-63 ± 6.6314e-63	4.8082e-190 ± 0	1.3115e-15 ± 1.4686e-16	0 ± 0	0 ± 0
f_2	20	15.9120 ± 0.2273	1.3524 ± 1.5792	0.7973 ± 1.6361	0.0871 ± 0.1254	0.0327 ± 0.0019	3.2981e-07 ± 1.0312e-06
	30	25.9785 ± 0.1774	3.3395 ± 4.4690	1.3895 ± 2.1499	0.0523 ± 0.0672	0.0711 ± 0.0128	1.0027e-07 ± 1.0502e-07
	50	46.3067 ± 0.4004	38.4515 ± 31.7815	16.2265 ± 21.2962	0.0634 ± 0.1142	2.5228 ± 1.2541	1.0660e-07 ± 8.0190e-08
f_3	20	88.3377 ± 10.1747	0 ± 0	0.2985 ± 0.4678	4.2633e-15 ± 1.0412e-14	0 ± 0	0 ± 0
	30	177.1109 ± 12.2431	5.6843e-15 ± 1.7496e-14	1.0945 ± 0.8479	8.5265e-14 ± 4.3251e-14	0 ± 0	0 ± 0
	50	365.4491 ± 12.8696	0 ± 0	5.5220 ± 2.6516	1.0601e-12 ± 1.6704e-12	0 ± 0	8.5265e-15 ± 2.7817e-14
f_4	20	0.2620 ± 0.1020	5.2736e-16 ± 1.4066e-15	0.0034 ± 0.0051	1.3711e-15 ± 2.2887e-15	0 ± 0	0 ± 0
	30	0.0178 ± 0.0797	0 ± 0	0.0041 ± 0.0098	7.9936e-16 ± 6.3277e-16	0 ± 0	0 ± 0
	50	2.0621e-6 ± 1.2609e-6	0 ± 0	0.0229 ± 0.0338	1.5432e-15 ± 6.5721e-16	0 ± 0	0 ± 0
f_5	20	2.9519e-6 ± 8.5896e-7	6.0396e-15 ± 7.9441e-16	2.6645e-15 ± 0	2.4336e-14 ± 3.6267e-15	7.1054e-16 ± 1.8134e-15	1.2434e-15 ± 1.7857e-15
	30	1.7312e-5 ± 4.9711e-6	7.2831e-15 ± 1.6704e-15	0.4004 ± 0.5677	4.7073e-14 ± 5.2189e-15	2.4869e-15 ± 7.9441e-16	2.6645e-15 ± 0
	50	1.5638e-4 ± 5.1830e-5	1.2790e-14 ± 1.3015e-15	1.8811 ± 1.8811	1.0214e-13 ± 8.3914e-15	2.6645e-15 ± 0	2.6645e-15 ± 0
f_6	20	1.0920e-7 ± 8.5896e-7	9.6865e-40 ± 1.0089e-39	2.9962e-182 ± 0	2.8100e-16 ± 2.2693e-17	0 ± 0	0 ± 0
	30	4.3319e-6 ± 4.1372e-6	9.6865e-40 ± 1.0089e-39	7.5626e-180 ± 0	5.0197e-16 ± 5.0710e-17	0 ± 0	0 ± 0
	50	1.4938e-4 ± 9.1548e-5	1.0767e-59 ± 9.2087e-60	3.6584e-188 ± 0	1.2513e-15 ± 1.3320e-16	0 ± 0	0 ± 0
f_7	20	-10.7121 ± 0.4311	-19.6363 ± 0.0013	-19.6204 ± 0.0210	-19.6359 ± 0.0013	-19.2512 ± 0.7144	-19.6370 ± 4.5865e-15
	30	-13.5080 ± 4.1372e-6	-29.5405 ± 0.0422	-29.5668 ± 0.0439	-29.6083 ± 0.0121	-29.2917 ± 0.5761	-29.3322 ± 0.4810
	50	-18.2114 ± 0.8100	-49.2281 ± 0.1068	-49.3694 ± 0.1155	-49.5258 ± 0.0239	-48.9364 ± 0.7706	-49.2284 ± 0.5118
f_8	20	-1.2099e3 ± 63.9563	-1.2126e3 ± 198.0180	-1.5200e3 ± 0.0030	-1.4934e3 ± 17.7408	-1.5200e3 ± 7.6568e-10	-1.5200e3 ± 1.0526e-09
	30	-2.1886e3 ± 419.3236	-2.3303e3 ± 786.8715	-4.8684e3 ± 30.0984	-3.7616e3 ± 379.9890	-4.9300e3 ± 1.0317e-8	-4.9300e3 ± 1.9630e-8
	50	-3.8943e3 ± 1.0972e3	-4.6039e3 ± 2.1276e3	-1.6988e4 ± 1.9831e3	-6.2731e3 ± 2.9887e3	-2.2050e4 ± 2.4728e-5	-2.2050e4 ± 1.2703e-06
f_9	20	-3.4543e3 ± 262.3109	-8.3797e3 ± 3.7325e-12	-8.3678e3 ± 52.9672	-8.3797e3 ± 1.7705e-12	-8.3797e3 ± 2.0013e-12	-8.3797e3 ± 2.4688e-12
	30	-4.2340e3 ± 206.4148	-1.2569e4 ± 1.8662e-12	-1.2534e4 ± 55.6852	-1.2569e4 ± 5.0878e-10	-1.2569e4 ± 3.9808e-12	-1.2569e4 ± 3.9808e-12
	50	-5.5094e3 ± 308.3849	-2.0949e4 ± 7.4650e-12	-2.0848e4 ± 123.1747	-2.0949e4 ± 2.2694e-4	-2.0949e4 ± 6.9328e-12	-2.0949e4 ± 6.8824e-12
f_{10}	20	680.7399 ± 158.9219	6.0736 ± 3.4399	3.4609e-29 ± 1.5188e-28	122.8894 ± 57.5833	0 ± 0	5.4347e-323 ± 0
	30	6.0084e3 ± 891.0211	192.7349 ± 40.2446	1.7348e-18 ± 3.7702e-18	1.4779e3 ± 417.2093	0 ± 0	2.9857e-207 ± 0
	50	2.6141e4 ± 3.4905e3	1.7640e3 ± 334.6602	1.1549e-11 ± 2.5786e-11	1.3111e4 ± 2.0672e3	4.9187e-16 ± 2.1481e-15	1.3290e-143 ± 2.4407e-143
f_{11}	20	2.6106e-9 ± 2.0321e-9	5.9122e-6 ± 2.7178e-6	1.8933e-30 ± 1.2854e-30	4.6418e-15 ± 1.2107e-15	3.4527e-13 ± 1.9644e-13	2.8663e-21 ± 1.5596e-21
	30	1.8994e-9 ± 1.5212e-9	7.2339e-5 ± 1.6458e-5	6.9050e-30 ± 1.9571e-30	7.9455e-15 ± 1.9605e-15	4.3734e-13 ± 2.3603e-13	4.7017e-21 ± 1.8546e-21
	50	6.2123e-12 ± 1.0570e-11	4.6982e-7 ± 1.0856e-7	4.1750e-29 ± 2.8959e-29	2.3607e-14 ± 7.8298e-15	8.2000e-13 ± 2.3292e-13	8.3954e-21 ± 2.8656e-21
f_{12}	20	0.0047 ± 0.0012	0 ± 0	0 ± 0	0 ± 0	0 ± 0	0 ± 0
	30	0.0302 ± 0.0077	0 ± 0	0.2097 ± 0.3033	0 ± 0	0 ± 0	0 ± 0
	50	0.2307 ± 0.0805	0 ± 0	1.7162 ± 0.8589	2.7711e-14 ± 9.7534e-15	0 ± 0	0 ± 0

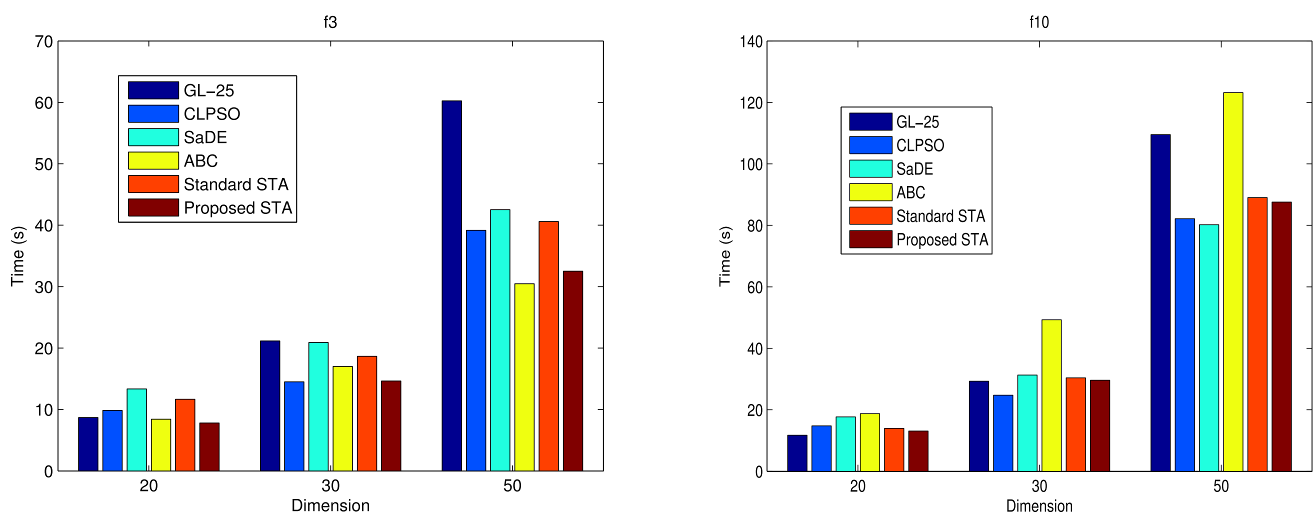

Fig. 2. Average elapsed time for different metaheuristic methods with respect to f_3 and f_{10}, respectively.

12) Weierstrass function

$$f_{12}(\boldsymbol{x}) = \sum_{i=1}^{n}\sum_{k=0}^{k_{\max}}\left[a^k \cos\left(2\pi b^k (x_i + 0.5)\right)\right] - n\sum_{k=0}^{k_{\max}} a^k \cos\left(\pi b^k x_i\right)$$

where $a = 0.5, b = 3$, and $k_{\max} = 20$, the global optimum $\boldsymbol{x}^* = (0, \ldots, 0)$, and $f(\boldsymbol{x}^*) = 0$, $-0.5 \leq x_i \leq 0.5, i = 1, \ldots, n$.

Other metaheuristics are used for comparison, including the GL-25 [36], CLPSO [37], SaDE [38], and ABC [39], with the same parameter settings as in these literatures. The parameters in the proposed STA are given by experience as follows: $SE = 30$ and $T_p = 10$ (additional experiments have testified the validity of these parameter values). The number of decision variables n of the benchmark functions is set to 20, 30, and 50, and the corresponding maximum function evaluations is set at 5e4*n*log(n). A total of 20 independent runs are conducted in the MATLAB (version R2010b) software platform on Intel Core i3-2310M CPU @2.10 GHz under Window 7 environment. The statistic results are given in Table V and some typical instances with respect to elapsed time and iterative curves are illustrated in Figs. 2–4.

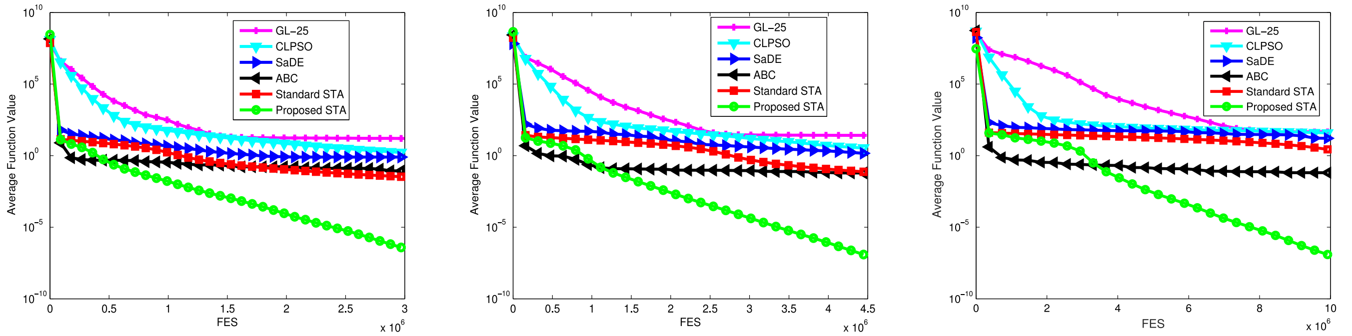

Fig. 3. Average iterative curves for different metaheuristic methods with respect to the Rosenbrock function.

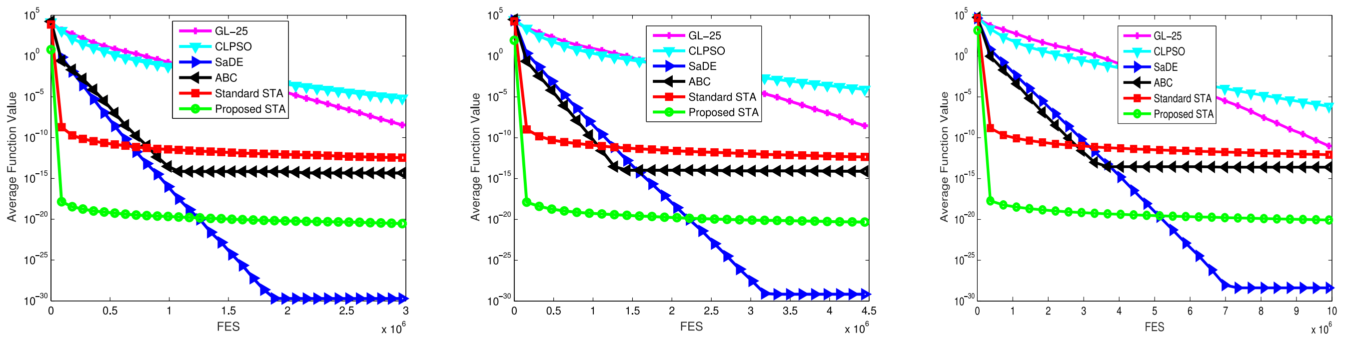

Fig. 4. Average iterative curves for different metaheuristic methods with respect to the Schwefel 2.4 function.

From the experimental results, it can be found that the proposed STA is superior to the basic STA among most of these test problems. The global search ability (see the Michalewicz function) and the solution accuracy (see the Rosenbrock function) has greatly improved. It can also be comparable to other metaheuristics except for the Michalewicz function. However, it should be noted that only mean and standard deviation are given for comparison. Actually, for the Michalewicz function, the results obtained from the proposed STA hit the known global solution for more than 50% of the total runs.

VI. Conclusion

In this paper, the optimal parameter selection of operators in continuous STA was considered to improve its search performance. First, a statistical study with four benchmark cases was conducted to investigate how these parameters affect the performance of continuous STA. And several properties are observed from the statistical study. With the experience gained from the statistical results, then, a new continuous STA with optimal parameter strategy was proposed to accelerate its search process. The proposed STA was successfully applied to other benchmarks. Comparison with other metaheuristics was conducted to demonstrate the effectiveness of the proposed method as well.

It should be noted that the parameter T_p is given by experience that needs further study, and the parameter selection of operators in continuous STA is still a challenging problem, since the proposed optimal parameter selection strategy can only be considered as a local vision. From an overall perspective, the parameter set should be taken into consideration as well, and it is not necessarily restricted to one below. Furthermore, it can be found that the STA does not work steadily for the Michalewicz function and global search ability should be strengthened further. In our future work, the upper bound of the parameter set will be considered as well, and an adaptive parameter selection strategy and appropriate utilization of transformation operators can also be alternative choices.

The MATLAB source codes of the standard STA and the proposed STA are available upon request from the first author, or can be downloaded from MATLAB central file exchange, or from X. Zhou's homepage as follows: https://www.mathworks.com/matlabcentral/fileexchange/ and http://faculty.csu.edu.cn/michael_x_zhou/zh_CN/index.htm.

References

[1] X. Zhou, C. Yang, and W. Gui, "State transition algorithm," *J. Ind. Manag. Optim.*, vol. 8, no. 4, pp. 1039–1056, 2012.
[2] X. Zhou, D. Y. Gao, C. Yang, and W. Gui, "Discrete state transition algorithm for unconstrained integer optimization problems," *Neurocomputing*, vol. 173, pp. 864–874, Jan. 2016.
[3] X. Zhou, C. Yang, and W. Gui, "Nonlinear system identification and control using state transition algorithm," *Appl. Math. Comput.*, vol. 226, pp. 169–179, Jan. 2014.
[4] X. Zhou, D. Y. Gao, and A. R. Simpson, "Optimal design of water distribution networks by a discrete state transition algorithm," *Eng. Optim.*, vol. 48, no. 4, pp. 603–628, 2016.
[5] X. Zhou, P. Shi, C.-C. Lim, C. Yang, and W. Gui, "A dynamic state transition algorithm with application to sensor network localization," *Neurocomputing*, vol. 273, pp. 237–250, Jan. 2018.
[6] F. Zhang, C. Yang, X. Zhou, and W. Gui, "Fractional-order PID controller tuning using continuous state transition algorithm," *Neural Comput. Appl.*, vol. 29, no. 10, pp. 795–804, 2018.
[7] G. Saravanakumar *et al.*, "Tuning of multivariable decentralized PID controller using state transition algorithm," *Stud. Informat. Control*, vol. 24, no. 4, pp. 367–378, 2015.

[8] G. Wang *et al.*, "State-transition-algorithm-based resolution for overlapping linear sweep voltammetric peaks with high signal ratio," *Chemometrics Intell. Lab. Syst.*, vol. 151, pp. 61–70, Feb. 2016.
[9] J. Han, C. Yang, X. Zhou, and W. Gui, "A new multi-threshold image segmentation approach using state transition algorithm," *Appl. Math. Model.*, vol. 44, pp. 588–601, Apr. 2017.
[10] C. Wang, H. Zhang, W. Fan, and X. Fan, "A new wind power prediction method based on chaotic theory and Bernstein neural network," *Energy*, vol. 117, pp. 259–271, Dec. 2016.
[11] J. Han, C. Yang, X. Zhou, and W. Gui, "Dynamic multi-objective optimization arising in iron precipitation of zinc hydrometallurgy," *Hydrometallurgy*, vol. 173, pp. 134–148, Nov. 2017.
[12] M. Huang, X. Zhou, T. Huang, C. Yang, and W. Gui, "Dynamic optimization based on state transition algorithm for copper removal process," *Neural Comput. Appl.*, 2017, doi: 10.1007/s00521-017-3232-0.
[13] Z. Huang, C. Yang, X. Zhou, and W. Gui, "A novel cognitively inspired state transition algorithm for solving the linear bi-level programming problem," *Cogn. Comput.*, 2018, doi: 10.1007/s12559-018-9561-1.
[14] Y. Wang, H. He, X. Zhou, C. Yang, and Y. Xie, "Optimization of both operating costs and energy efficiency in the alumina evaporation process by a multi-objective state transition algorithm," *Can. J. Chem. Eng.*, vol. 94, no. 1, pp. 53–65, 2016.
[15] Y. Xie, S. Wei, X. Wang, S. Xie, and C. Yang, "A new prediction model based on the leaching rate kinetics in the alumina digestion process," *Hydrometallurgy*, vol. 164, pp. 7–14, Sep. 2016.
[16] S. Xie, C. Yang, X. Wang, and Y. Xie, "Data reconciliation strategy with time registration for the evaporation process in alumina production," *Can. J. Chem. Eng.*, vol. 96, no. 1, pp. 189–204, 2018.
[17] C. Yang, S. Deng, Y. Li, H. Zhu, and F. Li, "Optimal control for zinc electrowinning process with current switching," *IEEE Access*, vol. 5, pp. 24688–24697, 2017.
[18] F. Zhang, C. Yang, X. Zhou, and H. Zhu, "Fractional order fuzzy PID optimal control in copper removal process of zinc hydrometallurgy," *Hydrometallurgy*, vol. 178, pp. 60–76, Jun. 2018.
[19] X. Zhou, J. Zhou, C. Yang, and W. Gui, "Set-point tracking and multi-objective optimization-based PID control for the goethite process," *IEEE Access*, to be published, doi: 10.1109/ACCESS.2018.2847641.
[20] J. Wang, G. Liang, and J. Zhang, "Cooperative differential evolution framework for constrained multiobjective optimization," *IEEE Trans. Cybern.*, to be published, doi: 10.1109/TCYB.2018.2819208.
[21] Y. Wang, D.-Q. Yin, S. Yang, and G. Sun, "Global and local surrogate-assisted differential evolution for expensive constrained optimization problems with inequality constraints," *IEEE Trans. Cybern.*, to be published, doi: 10.1109/TCYB.2018.2809430.
[22] J. Zhang, X. Zhu, Y. Wang, and M. Zhou, "Dual-environmental particle swarm optimizer in noisy and noise-free environments," *IEEE Trans. Cybern.*, to be published, doi: 10.1109/TCYB.2018.2817020.
[23] A. E. Eiben and S. K. Smit, "Parameter tuning for configuring and analyzing evolutionary algorithms," *Swarm Evol. Comput.*, vol. 1, no. 1, pp. 19–31, 2011.
[24] J. Sun, W. Fang, X. Wu, V. Palade, and W. Xu, "Quantum-behaved particle swarm optimization: Analysis of individual particle behavior and parameter selection," *Evol. Comput.*, vol. 20, no. 3, pp. 349–393, 2012.
[25] W. Zhang, D. Ma, J.-J. Wei, and H.-F. Liang, "A parameter selection strategy for particle swarm optimization based on particle positions," *Expert Syst. Appl.*, vol. 41, no. 7, pp. 3576–3584, 2014.
[26] R. Mallipeddi, P. N. Suganthan, Q.-K. Pan, and M. F. Tasgetiren, "Differential evolution algorithm with ensemble of parameters and mutation strategies," *Appl. Soft Comput.*, vol. 11, no. 2, pp. 1679–1696, 2011.
[27] R. A. Sarker, S. M. Elsayed, and T. Ray, "Differential evolution with dynamic parameters selection for optimization problems," *IEEE Trans. Evol. Comput.*, vol. 18, no. 5, pp. 689–707, Oct. 2014.
[28] Y. Wang, Z. Cai, and Q. Zhang, "Differential evolution with composite trial vector generation strategies and control parameters," *IEEE Trans. Evol. Comput.*, vol. 15, no. 1, pp. 55–66, Feb. 2011.
[29] D. Karaboga and B. Gorkemli, "A quick artificial bee colony (qABC) algorithm and its performance on optimization problems," *Appl. Soft Comput.*, vol. 23, pp. 227–238, Oct. 2014.
[30] Á. E. Eiben, R. Hinterding, and Z. Michalewicz, "Parameter control in evolutionary algorithms," *IEEE Trans. Evol. Comput.*, vol. 3, no. 2, pp. 124–141, Jul. 1999.
[31] K. De Jong, "Parameter setting in EAs: A 30 year perspective," in *Parameter Setting in Evolutionary Algorithms*. Heidelberg, Germany: Springer, 2007, pp. 1–18.
[32] G. Karafotias, M. Hoogendoorn, and Á. E. Eiben, "Parameter control in evolutionary algorithms: Trends and challenges," *IEEE Trans. Evol. Comput.*, vol. 19, no. 2, pp. 167–187, Apr. 2015.
[33] X. Zhou, C. Yang, and W. Gui, "Initial version of state transition algorithm," in *Proc. IEEE 2nd Int. Conf. Digit. Manuf. Autom. (ICDMA)*, 2011, pp. 644–647.
[34] X. Zhou, C. Yang, and W. Gui, "A new transformation into state transition algorithm for finding the global minimum," in *Proc. IEEE 2nd Int. Conf. Intell. Control Inf. Process. (ICICIP)*, vol. 2. Harbin, China, 2011, pp. 674–678.
[35] X. Zhou, C. Yang, and W. Gui, "A comparative study of STA on large scale global optimization," in *Proc. IEEE 12th World Congr. Intell. Control Autom. (WCICA)*, Guilin, China, 2016, pp. 2115–2119.
[36] C. García-Martínez, M. Lozano, F. Herrera, D. Molina, and A. M. Sánchez, "Global and local real-coded genetic algorithms based on parent-centric crossover operators," *Eur. J. Oper. Res.*, vol. 185, no. 3, pp. 1088–1113, 2008.
[37] J. J. Liang, A. K. Qin, P. N. Suganthan, and S. Baskar, "Comprehensive learning particle swarm optimizer for global optimization of multimodal functions," *IEEE Trans. Evol. Comput.*, vol. 10, no. 3, pp. 281–295, Jun. 2006.
[38] A. K. Qin, V. L. Huang, and P. N. Suganthan, "Differential evolution algorithm with strategy adaptation for global numerical optimization," *IEEE Trans. Evol. Comput.*, vol. 13, no. 2, pp. 398–417, Apr. 2009.
[39] D. Karaboga and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: Artificial bee colony (ABC) algorithm," *J. Glob. Optim.*, vol. 39, no. 3, pp. 459–471, 2007.

Xiaojun Zhou received the bachelor's degree in automation from Central South University, Changsha, China, in 2009, and the Ph.D. degree in applied mathematics from Federation University Australia, Ballarat, VIC, Australia, in 2014.

He is currently an Associate Professor with Central South University. His current research interests include modeling, optimization and control of complex industrial process, optimization theory and algorithms, state transition algorithm, duality theory, and their applications.

Chunhua Yang (M'09) received the M.S. degree in automatic control engineering and the Ph.D. degree in control science and engineering from Central South University, Changsha, China, in 1988 and 2002, respectively.

From 1999 to 2001, she was a Visiting Professor with the University of Leuven, Leuven, Belgium. Since 1999, she has been a Full Professor with the School of Information Science and Engineering, Central South University. From 2009 to 2010, she was a Senior Visiting Scholar with the University of Western Ontario, London, ON, Canada. Her current research interests include modeling and optimal control of complex industrial process, fault diagnosis, and intelligent control system.

Weihua Gui (M'09) received the B.Eng. degree in automatic control engineering and the M.Eng. degree in control science and engineering from Central South University, Changsha, China, in 1976 and 1981, respectively.

From 1986 to 1988, he was a Visiting Scholar with Universitat-GH-Duisburg, Duisburg, Germany. He has been a Full Professor with the School of Information Science and Engineering, Central South University, since 1991. His current research interests include modeling and optimal control of complex industrial process, distributed robust control, and fault diagnoses.

Mr. Gui is a member of the Chinese Academy of Engineering.